You R Cute: The Influence Of Societal Perception On The Search For Online Romantic Partners

by
Samantha DiPiero

A Thesis Submitted to the Faculty of

The Harriet L. Wilkes Honors College

in Partial Fulfillment of the Requirements for the Degree of

Bachelor of Arts in Liberal Arts and Sciences

with a Concentration in Psychology

Wilkes Honors College of

Florida Atlantic University

Jupiter, Florida

May 2018

You R Cute: The Influence Of Societal Perception On The Search For Online Romantic Partners

by

Samantha DiPiero

This thesis was prepared under the direction of the candidate's thesis advisor, Dr. Kevin Lanning, and has been approved by members of her supervisory committee. It was submitted to the faculty of The Honors College and was accepted in partial fulfillment of the requirements for the degree of Bachelor of Arts in Liberal Arts and Sciences.

SUPERVISORY COMMITTEE:

____

Dr. Kevin Lanning

____

Dr. Wairimũ Njambi

____

Dean Ellen Goldey, Wilkes Honors College

____

Date

Acknowledgements

I want to first thank Dr. Lanning for constantly believing in me and pushing me to achieve things I did not think were possible, like learning a completely new programming language in one semester. His support throughout this whole project was amazing, I do not think I would have achieved as much as I did without his guidance. I would also like to thank Dr. Njambi for being one of the most influential professors I have ever had at the Honors College. Dr. Njambi, Dr. Lanning, and Dr. Earles have helped me develop both as a student and as a person, my experience would not have been the same without all of their support and guidance, I will always be grateful to them.

I also could not have completed this project without the help of my thesis partners, Karly Hauser and Sonia Figueroa. Learning R with them, working through code errors, and meeting with them every week was an extremely rewarding experience. It was nice to know that I was not alone in facing my issues with the dataset or with the programming.

I would also like to thank my family for dealing with me in my extremely stressed out state throughout this writing process. My parents were extremely supportive throughout this whole process, and my brother and sister provided a bunch of very necessary comedic relief when things got too stressful. Mom and Dad, thank you for listening to me freak out about thesis and for supporting me throughout my whole college experience. Dominic, thank you for bringing me various cats in times of stress, and doing work with me. And Gigi for staying up way too late with me, watching makeup tutorials, binging Hulu and Netflix, and keeping me on track. I can’t wait to see what you and Dominic achieve! I am so blessed to have such an amazingly supportive family that believes in me no matter what! And to my second family, Brian, Melissa, and Cameron, thank you for all of your kind words of support and supplies of love, laughs, and board games.

Finally I would like to thank all of my amazing friends that have supported me throughout my time at the Honors College, including Loren, Piper, and Ian. Loren for all of your kind words and calm demeanor when I was on the verge of exploding. Piper for always listening to me vent about my struggles with the code, and for taking Disney breaks with me to help destress. And Ian for being one of the most amazing human beings I have ever met, and providing support, love, guidance, fancy waters, ice cream, and mac and cheese. Also for being my behind the scenes editor making my thesis more understandable for readers who don't have code experience. I could not have completed my thesis without all of your love and support. And to everyone else who has provided me with memes, cute animal videos, and support, you are all amazing!

ABSTRACT

Author: Samantha DiPiero

Title: You R Cute: A Statistical Analysis Of The Influence Of Societal Perception On The Searching For Online Romantic Partners

Institution: Wilkes Honors College of Florida Atlantic University

Thesis Advisor: Dr. Kevin Lanning

Degree: Bachelor of Arts in Liberal Arts and Sciences

Concentration: Psychology

Year: 2018

While online dating is very convenient, it causes an evolutionary mismatch, our ancestors did not have websites to search for partners. However, there are some similarities that can be seen between current and ancestral times in relation to societal perception, especially in relation to parental investment theory. Using an OkCupid dataset of 2620 questions, 166 were selected in order to evaluate how much of an influence societal perception (the influence of family and friends, substance use, education, stigma, religion, appearance, and morals) has on searching for a romantic partner. The data was evaluated in four separate stages: Initially just 100 cases were examined to reduce computational burden, then an additional 2000 were used to look at the structure of key variables. Of the remaining cases, 60% were used to develop a prediction model and 40% to test this model. It was found that women tended to be more selective on six of the seven scales that were measured, this selectiveness can be attributed to parental investment theory.

Table of Contents

List of Tables and Charts

You R Cute: A Statistical Analysis Of The Influence Of Societal Perception On The Searching For Online Romantic Partners

Introduction

It appears that in our current society technology is taking over every facet of human interaction, starting with friends on Facebook, all the way through finding a romantic partner with Match.com or OkCupid. The internet essentially takes away the issue of proximity in forming relationships (Ansari, 2015). Now individuals can be in different countries and communicate as if they were right next to each other. It also allows for individuals to be more selective of those they wish to interact with, this extends through finding a romantic partner. In ancestral times when groups were nomadic proximity had a lot to do with romantic relationships. Furthermore, the approval of the immediate kin group was also necessary. Online dating has provided humans with the power to choose in ways they have not experienced before, but there are some leftover preferences that have been carried over from ancestral times.

Evolutionary Psychology

Ancestral Times

In ancestral times, finding a mate was a very important process. The most noteworthy difference between current times and back then is that an individual would need to be in a close proximity to his or her possible mate. Currently it is possible to have long distance online relationships, individuals are able to meet via online dating sites that use algorithms that link individuals up based on compatibility. With this new development, individuals are facing things they never had to before. However, in these times there are things that are likely to be equally as important to individuals. One such thing is the social perception of one’s partner. In ancestral times, individuals traveled in groups of around 150, because of these groups, the sense of community was imperative for group success (Geher, 2014). That being said, groups were mostly comprised of kin, these kinship ties were the central force of the groups, holding them together and making sure no one would betray each other or do anything that was deemed unacceptable by the group.

When a member of the group would find a mate, it was important that the mate be accepted by the group in order to maintain balance and cohesiveness. While humans no longer travel in groups of 150, group morals and values are still held in high regard by individuals. This loosely translates to the impact the social perception of one's mate has on the relationship, and the importance of social perception to the individual looking for a mate. A person who comes from a religious family will likely not look for a non-religious mate, or a mate that does not share the same religion. A person who grew up in a home where family was held above all else will likely not look for a mate who does not consider family to be important. Furthermore, the opinions of the individual’s kin and society will likely have an impact on compatibility and longevity of a relationship. The importance of acceptance by one's kin is something that has not changed much from ancestral times (Kerr & Levine, 2008). Individuals live further away, but a person will likely not maintain a relationship with someone who their friends and family do not like or approve of.

Parental Investment Theory

Parental investment theory was initially posited by Robert Trivers in 1972. Trivers defines parental investment as, “any investment by the parent in an individual offspring that increases the offspring's chance of surviving (and hence reproductive success) at the cost of the parent's ability to invest in other offspring. So defined, parental investment includes the metabolic investment in the primary sex cells but refers to any investment (such as feeding or guarding the young) that benefits the young" (Trivers 1972, p. 55). From the basic definition provided by Trivers, it can be seen that parental investment weighs more heavily on females in humans, and in many other species. Females from the beginning have a larger risk going into any sort of relationship, so it makes sense for them to be wary and have an idea of what they will get out of the relationship (Clark, Dover, Geher, & Presson, 2005). If the woman gets pregnant, they have that up-front investment and they have to carry the child to term, or terminate the pregnancy, but that is still a form of investment that the male does not have to experience, Trivers elaborates,

> "In the human species, for example, a copulation costing the male virtually nothing may trigger a nine-month investment by the female that is not trivial, followed, if she wishes, by a fifteen-year investment in the offspring that is considerable. Although the male may often contribute parental care during this period, he need not necessarily do so. After a nine-month pregnancy, a female is more or less free to terminate her investment at any moment but doing so wastes her investment up until then. Given the initial imbalance in investment the male: may maximize his chances of leaving surviving offspring by copulating and abandoning many females, some of whom, alone or with the aid of others, will raise his offspring" (Trivers 1972, p.62).

Females are more involved in almost every step of a child's development. This would lead males to be less needed and allow a window for them to leave. Strictly speaking, according to evolutionary psychology, the main goal for males in any species is to spread as much of their genetic material as possible. This explains why the parental investment of males is so low, because they are meant to continuously spread their genes. Males compete for the attention of females in order to get the possibility of spreading their genetic material, and if the tables were turned females would do the same thing (Trivers 1972). However, since in humans the females are the more invested sex in parental investment terms, they have the ability to be more selective when it comes to mates, and evolution favors those females who are more selective. Evolution also favors those that take care of their kin.

Kinship Ties

Evolutionary psychology dictates that there are no ties more important than kinship ties. There have been numerous studies done on animals that prove that animals are more likely to be aggressive when a predator is attacking one of its kin, they are also more likely to take risks in order to protect their kin, this is considered altruism. One explanation for altruism is kin-selection, documented by William Hamilton, "perhaps relatively high levels of genetic interrelatedness among individuals within a species relates in an important way to the amount of altruism we expect to see within that species" (Geher 2014, p. 107). The idea of kin-selection is connected to another one of Hamilton's theories, inclusive fitness (Gorelik, Shackelford, & Salmon, 2010).

Inclusive fitness posits, "that there are two distinct methods by which an organism can increase its fitness or levels of reproductive success" (Geher 2014, p. 107). One of the methods is direct fitness, "by which an organism facilitates its own survival and/or reproduction" (Geher 2014, p. 107). The other is indirect fitness, "which includes processes by which an organism facilitates the survival and/or reproductive ability of its kin (i.e., genetically related individuals)" (Geher 2014, p. 107). The concept of indirect fitness allows one explanation for altruism, the organism is not going to let its kin die off or struggle, the relatedness dictates how likely others of the species are to help out that individual. Evolutionarily, it makes sense that an individual of a species would want to ensure the survival of its kin, and help them even if it does have some sort of personal cost.

Family ties are imperative for development of an individual because the family is the first exposure an individual has to any sort of social norms or rules (Conger, Cui, Bryant, & Elder, 2000). Maintaining family relationships over time was imperative in ancestral times due to the small groups that individuals traveled in. The relationships also provided a support network for individuals to depend on if things were difficult or they needed help (Fuller-Iglesias, Webster, & Antonucci, 2015; Sander, Schupp, & Richter, 2017).

Societal Perception

Stemming from ancestral times and kinship ties, humans are likely to stay within social norms when searching for a partner (Fischer, 2009; Sprecher, 2010). Having the support of family and friends allows for an individual to be comfortable in a relationship or any other decision they will make (Sprecher & Felmlee, 1992; Preston, Gottfried, Oliver, Gottfried, Delany, & Ibrahim, 2016). Furthermore, the comfort levels are increased when being part of a relationship that is socially acceptable. If an individual were to deviate from the norm, for example by breaking the law, they would be looked down on and excluded from society due to the stigma of breaking the law. Furthermore an individual that has been stigmatized will have a harder time fitting into groups (Kerr & Levine, 2008).

Stemming from ancestral times, an aversion to exclusion is woven into human beings. In ancestral times, being excluded from a group would likely end up with death because the individual loses the social network they have depended on for so long (Kerr & Levine, 2008). The need to fit in socially makes people sensitive to social perception. The severity of exclusion has decreased over time, but individuals still avoid it. This could be why online dating had such a slow pick up, because for the longest time when computers were first made publicly attainable, dating websites were seen as sleazy or weird so individuals avoided them (Ansari, 2015). However, once computers were more trusted, online dating became an important method for people to use to look for a partner.

Online Dating

Background

The internet has become an integral part of society, from social media, to professional networking, and even dating sites. People are interacting mostly through online sources rather than having face to face conversations. It makes sense that people would also turn to the internet to look for a romantic partner because it is so deeply rooted in almost every other aspect of human interaction (Finkel, Eastwick, Karney, Reis, & Sprecher, 2012). The development of online dating is outlined by Aziz Ansari in his book “Modern Romance”, this form of dating originated in the 1960s. Much like a personality test, individuals filled out extensive questionnaires and would enter the responses into a computer, which would then provide a match based off of whatever algorithm was being used. According to Ansari, this craze died out in the 1980s and was replaced by classified ads in newspapers. Individuals would put up a blurb about themselves in the newspapers at the time describing what they were like, and the partners they were looking for. The listers of the ads included a phone number for individuals who were interested could call them and hopefully it would turn into some sort of meeting or date.

Another type of dating that was present at this time was video dating. Individuals would go into studios and record themselves talking about their interests and partners they were interested in. This form of dating was more personal than the newspaper ad, with the bonus of being able to see the individual in the video. However, this form of dating did not catch on and quickly died out. Once the internet was becoming accessible to everyone, one of the first online dating websites, Match.com, was created. There were some notable issues in the beginning of this era of online dating, “Initially Match.com was hampered by the same stigma that had kept people away from previous computer dating services. During the Internet boom of the late 1990s, though, people’s relationship to computers and online culture changed dramatically, and more and more people were getting comfortable using computers for basic tasks” (Ansari 2015, p.78-79). After people got used to computers, online dating picked up a lot of users.

Popularity

Ansari explains that between 2005 and 2012 more than one third of couples who got married in the United States had met online. Online dating had become, “the biggest way people met their spouses. Bigger than work, friends, and school *combined*” (Ansari 2015, p. 79). Online dating is widely used because it is easy, and because everyone that is using an online dating site is looking to meet new people, which takes away the fear of rejection. It is easily used by individuals from every generation, from 18 year olds all the way through individuals in their 70s (Alterovitz & Mendelsohn, 2011). Furthermore, it is easier to filter out the criteria for possible matches beyond just what the algorithm asks for. After answering however many questions the site requires, many sites allow the individual to fine tune the possible matches. For example, one could choose to control for height of a partner, age of a partner, or proximity to a partner . These are all things that were not available without online dating, which presents a new concept, the shift from mostly face to face interactions to mostly online interactions.

Face to Face vs. Online Interactions

There is a paradigm shift from face to face interactions to online interactions with the introduction of the internet. Before computers were commonplace, individuals had a lot more variables influencing whom they met and interacted with. One of the most important variables was proximity (Festinger, Schachter, & Black, 1968). Individuals who were in the same general area they were more likely to form relationships with each other. Ansari dove deeper into the connection between proximity and relationships referring to a study conducted by James Bossard in 1932, "*One-third* of the couples who got married [in Philadelphia] had lived within a five-block radius of each other before they got married. One out of six had lived within the same block. Most amazingly, one of every eight married couples had lived in the *same building* before they got married" (Ansari 2015, p. 14). Other studies were conducted outside of large cities, and similar trends occurred, people were likely to live within a close proximity to their partners, and were only likely to expand their search as far as they had to in order to find someone.

With the increasing popularity of online dating, the proximity variable has been somewhat removed from the equation. People can interact with others from different cities, states, or even countries. Individuals who engage in computer mediated relationships are likely to have a large distance between them (Merkle & Richardson, 2000; Huxhold, Fiori, & Windsor, 2012). Another variable that facilitates relationships is physical attractiveness (Oesch & Miklousic, 2012; (Menkin, Fobles, Wiley & Gonzaga, 2015). If individuals were close together and were physically attracted to one another, then a relationship was likely to form. However, in online relationships an individual is allowed to conceal their physical appearance so an emotional connection is made first (Merkle & Richardson, 2000). The internet has taken the face-to-face relationship that is rooted in proximity and physical appearance, and created computer mediated relationships in which individuals get to know each other quickly, and have the proximity and attractiveness variables become somewhat obsolete when a relationship is beginning (Merkle & Richardson, 2000; Gibbs, Ellison & Lai, 2011). However, in ancestral times, the main indicator of a relationship was the proximity one individual had to another.

OkCupid

Background

OkCupid is an online dating website, it matches individuals based on their answers to multiple-choice questions. The answers to these questions were posted on each individuals online profile and displayed for possible matches to see. The creators of OkCupid (Chris Coyne, Christian Rudder, Sam Yagan, and Max Krohn) were all college students when they made the site. Rudder explains, "I started it [OkCupid] with friends. We were all mathematically minded, and the site succeeded in large part because we applied that mind-set to dating; we brought some analysis and rigor to what had historically been the domain of love 'experts' and warlocks like Dr. Phil" (Rudder 2015, p. 15). The website's algorithm matches individuals based on their answers to hundreds, sometimes thousands, of multiple-choice questions. It allows the individual to select the answer that they choose, and those they deem unacceptable from a possible match. Using this, the algorithm matches individuals the best it can in order to create a match. Being that this is an online system, privacy is not assured, in fact, OkCupid has all of the users data set to be public as a default. This allowed for a group of researchers led by Emil O. W. Kirkegaard to publish the 'OkCupid dataset'. The dataset contains a little over 2,600 variables and describes a little under 70,000 users of OkCupid (Kirkegaard & Bjerrekær, 2016).

Structure

OkCupid's structure is somewhat simplistic for the user. When an individual sets up an OkCupid profile, he/she answers several basic questions including those about gender, sexuality, and what kind of relationship he/she is looking for. After the initial questions, the individual is then asked more specific questions related to sexual preferences, age preferences, trait preferences, and other things along those lines (McKinlay, 2014). OkCupid's site has thousands of questions that individuals answer over time. The user is also able to rate how important a question is to him/her. For example, if there was a question asked about how important family is, the user would also be able to say what the acceptable answer would be from his/her prospective partner. If family is really important to a user, then they would likely rate the question highly and want his/her potential match to also hold family in high regard. The user is also able to go back and change answers as they see fit.

After the user answers the questions, the algorithm of OkCupid takes the information and matches the user to another user based on compatibility and shared interests. These various matches are able to talk to each other through OkCupid and the hope is that the interactions on the website will lead to dates in real life, and maybe even a proper relationship. That being said, OkCupid is a public platform designed to help people meet each other. The information filled out by each user is visible to the public, which is practical because the goal of the dating website is to meet new people. It is not like making one's Netflix feed public information, the goal is to connect compatible individuals who have not yet met which makes the public default the most practical.

OkCupid Dataset

Publication

The OkCupid dataset was published in 2016, and contains a little over 2,600 variables and describes a little under 70,000 users of OkCupid. The set contains the question numbers, the questions, the possible responses, number of responses, and some other information about each user.

Privacy

There was a lot of controversy about the dataset that was published. The main issue was that although names and images of users were not published, based on answers to questions some individuals might be able to be identified. However, the information on OkCupid is all public, and when filling out the questions on OkCupid, under each of the first questions it says that the information will be public. According to Christian Rudder, the default setting of the website is a public setting. The function of the website is to form new connections, and the public nature of the information is not hidden in a long contract, it is a stand-alone statement under each question which means the dataset is within the parameters of OkCupid's guidelines.

Hypothesis

An evolutionary approach to the OkCupid data could provide insight into the importance of social acceptance and perception on possible mates. Using this approach it will be easier to look at the questions through a psychological lens while addressing the possible reasons for the answers provided by the individuals. Furthermore, coming from evolutionary psychology, parental investment theory is ideal to evaluate using the questions in the OkCupid dataset. Based off of parental investment theory, it is hypothesized that women will be more influenced by societal perception and more likely to be more selective when looking for a partner because they have more of an up-front risk when entering a new relationship, this will be measured through individuals' various responses to the questions from the OkCupid dataset and the demographic information that was provided.

Methods

Participants and Design

The OkCupid data set that was used included the responses of 68,371 individuals (ages 18 through 100**)** to 2,620 multiple-choice questions on OkCupid.com. From these questions, 166 were selected to measure the influence of societal perception, along with three demographic variables. After selecting the questions, the data set was split into three groups of different sizes in order to create and test scales to measure societal perception. The final scales were then run through the statistical program R to find group means for each scale and demographic variable.

Data Set

In order to evaluate the influence of societal perception and kin acceptance on the acquisition of a mate, a list of pertinent questions was selected from the overall OkCupid dataset in order to provide a more focused analysis. These questions range from, “Is it possible to agree to disagree on religious practices in a relationship?” to “Would you consider dating someone who had spent considerable time in a mental health facility? (i.e. mental hospital/institution.)” These questions evaluate the opinion of the individual about stigma, religion, and other topics that relate to societal influence.

Selection. After reading each of the questions from the dataset, 166 were selected in order to evaluate the influence of societal perception on an individual. Several characteristics of questions were important when selecting them for the analysis, including questions that ask about the individuals opinions on moral topics, questions that look at how the individual values others’ opinions, questions that measure how important religion and family are to the individual, and other areas that evaluate how much the individual cares about others thoughts and opinions.

Three demographic variables were also selected to evaluate, these included age, gender, and ethnicity.

Separation. After selecting the questions, the dataset was separated into three groups, one called 'tiny' which included 100 participants, one called 'secondSmall' which included 2,000 participants, and one called 'NOTsecondSmall' which included 66,271 participants. The groups were comprised of randomly selected individuals, and there was no replacement which allowed for a new group of respondents each time, providing a more accurate scale. Using the group 'tiny' the questions were separated into scales and narrowed down the questions that were chosen based on their relevance and how many individuals answered them. Using 'secondSmall' the scales were closely evaluated to ensure that they were accurately measuring what they were created to measure. This was achieved through calculating an alpha value for each scale, and looking at the correlations between each question. Finally, using 'NOTsecondSmall' the scales were run once again to test the alphas and correlations, and this set was used to produce the results.

Recodes. In order to allow R to calculate values, all of the questions had to be recoded. Each question was analyzed, and the possible responses were all assigned a numerical value. The values ranged from 0 to 1.5 depending on the question. The questions were all selected and evaluated, and the responses were ranked on a scale from 0 to 1.5. The responses on the higher end of the scale represent how much of an answer that demonstrates a high impact of societal perception. For example, if someone did not care what others thought about his/her decision, his/her response to the question would be 0, while someone who is highly influenced by what others think would score higher, around a 1 or 1.5. If a question had only two possible answers, they were recoded as either a 0 or a 1, if a question had three possible answers, the recodes were

0, 0.5, or 1, and if there was a question with four possible answers, they were recoded as 0, 0.5, 1, or 1,5. For example, question 6988 “If someone sends you an e-mail (or has a profile) full of spelling and grammatical errors, are you less likely to talk to them?” has only two possible answers, “Yes” or “No”. In this case, “Yes” would be recoded as 1 and “No” would be recoded as 0, the code looks like this:

```
samData$q6988<- car::recode(q6988, "'Yes'= 1; 'No'= 0").
```

For a question with three possible answers, like question 19,566 “Would you consider dating a High School dropout?” the responses are “Yes”, “Yes, if they are doing well”, or “No”, they would be recoded like this:

```
samData$q19566<- car::recode(q19566, "'Yes'= 0; 'Yes, if they are doing well'= 0.5;
'No'= 1")
```

For other questions and recodes see Appendix A.

In all of these cases, a higher recode means that the answer is indicative of a response that fits with social norms.

For the demographic variables a different approach was used. The demographic variables were recoded as characters rather than numbers. This allowed for the data to be more understandable when making visualizations. For example, the ages were divided into groups and recoded as different generations, the generation ranges were adapted from the PEW Research Center (Dimock, 2018). The “d_age” recodes are as follows:

```
demographicData$d_age <- car::recode( d_age," 18:34 ='1_Millennials'; 35:50 =
'2_Generation X'; 51:69 = '3_Baby Boomers'; NA='6_Other'; else= '6_Other'")
```

There are a number of things that are different in this recode than the others. First, the numerical values were recoded as words, but when this was done R would produce tables in alphabetical order rather than sorting them by age. This was solved by assigning both a numerical and a character value to the recode, so when the data was run, it would be arranged from youngest to oldest.

Scales

The questions were separated into seven different groups: family and friends, appearance, education, moral, religion, stigma, and substance use. Each of these groups contains a facet of societal influence that might have an impact on the individuals answering the questions. Initially the questions were separated into eight groups: family and friend, appearance, education, moral, religion, stigma, substance use and race. However, after looking at the groups, the race scale only comprised of three questions, which was not enough to be an accurate measure, so the race group was eliminated and replaced with the ethnicity demographic data.

The family and friend scale measured how much influence the opinions of the family and friends had on the participant. An example of a question from the family and friend scale would be “Are you very close to your family?” Family and friend influences have been shown to be extremely important to individuals. Individuals learn all they know from an early age from their family and friends. They learn moral values, social norms, and how to interact/communicate with others (Conger, Cui, Bryant, & Elder, 2000). This scale was crafted to measure the importance of family and friends to the individual searching for a partner. The measure shows how dependent the user is on their family and friends for advice and approval.

The appearance scale measured the importance of physical appearance to the participant, an example question would be “If you meet someone and they are everything you are looking for, except their body type, do you give them a chance?”. It has been found in other studies that both men and women are more concerned with appearance in short-term relationships, but women are more selective about appearance in most cases (Regan, Levin, Sprecher, Christopher, & Cate, 2000). Appearance is an indicator mostly for short-term relationships, but for online dating is an ideal way to narrow down the choices (Oesch & Miklousic, 2012; Menkin, Fobles, Wiley & Gonzaga, 2015).

Next, the education scale measured the importance of education level to the participant, a question would be “If your significant other has not reached the same educational level as you, is this a turn off?”. It has been found that individuals prefer their partners to have achieved a similar education level to themselves (Boberg, 2008). Education is an important aspect in forming relationships. Having similar education levels is a good predictor of relationship formation. It has been found that people with college degrees place a higher value on education level than those without college degrees (Boberg, 2008). This guided the formation of the education scale, so it does not measure an individuals’ intelligence, but rather how important a comparable education level is to them in a partner.

The moral scale measured how strong the morals of the participant were in relation to social norms. An example question would be “Is it wrong to cultivate or manufacture illegal drugs?”. Individuals with similar morals tend to have better relationships, and women tend to look more closely for moral similarities in their partner (Koleva, 2012). Moral baselines are important to the formation of a relationship. Individuals with similar morals are likely to get along much better than those who do not (Koleva, 2012). Both men and women care about similar morals, but women tend to be more selective when it comes to this aspect of a romantic partner (Koleva, 2012).

The religion scale measured the importance of religion to the participant, for example “Does a relaxed attitude about spirituality/religion bother you?”. Religion has been found to have an impact on both the formation and longevity of relationships (Braithwaite, Coulson, Spjut, Dickerson, Beck, Dougal, Debenham, & Jones, 2015). Due to the relationship between religious beliefs and romantic relationships, the religion scale measures for similar attitudes and level of religious belief.

The stigma scale measured the aversion of the participant to any possible stigmatizing characteristic, an example question would be "Would you consider dating someone who had a very shady, questionable past?". Stigma aversion has been present since ancestral times because individuals do not want to be excluded from their groups, although now there are less dire consequences (Kerr & Levine, 2008). The stigma scale measures an aversion to stigma when looking for a partner. The avoidance of stigmatized individuals is related to the human beings need for acceptance and their reliance on the social network they have built (Kerr & Levine, 2008). If an individual is looking for a long term relationship they will be more averse to stigma (Regan, Levin, Sprecher, Christopher, & Cate, 2000).

Finally, the substance use scale measured the participants aversion to substance use, including both drugs and alcohol, a question from this scale would be "Could you date someone who does drugs?" It has been found that individuals that engage in substance use are less reliable when it comes to their personal lives (Geher, 2014). For an individual trying to form a romantic relationship, it is important that they find a dependable and trustworthy partner (Regan, Levin, Sprecher, Christopher, & Cate, 2000).

Testing. After selecting the questions for the scales and defining them all, the accuracy of the scales had to be measured. This was achieved by calculating an alpha for each of the scales in R, see Appendix B. The language used for testing the scales is selecting the questions associated with the stigma scale, isolating them from the rest of 'samData' and adjusting for the missing responses. After this is achieved, the stigma scale will then undergo a test for Cronbach's alpha, which measures the internal consistency of the scale. A good alpha is above .6, anything below that is questionable and indicates that the scale should be reevaluated (Kline, 1999). The alpha for the stigma scale came out to be .89, which is a good alpha, indicating that the stigma scale has solid internal consistency and is measuring for stigma aversion variables. The alphas of the other scales include: family and friend is .63, appearance is .78, education is .71, moral is .64, religion is .8, and substance use is .63. After looking at the scale alphas, it was time to view the means for the individuals.

Visualizations

Using the dplyr library in R, the means of each group were able to be calculated, the code was written to look at the scales given demographic information (Wickham, Francois, Henry, & Muller, 2017). For example, if one was to look for the means of the family and friend scale in relation to gender, the code would look like this:

```
all %>%
  group_by(d_gender) %>%
  summarise_at(vars(familyandfriend), funs(mean(., na.rm=TRUE), n=n()))
```

When looking at other scales or demographic information, all that has to be changed is the contents of the parentheses in the 'group_by' command (changing d_gender to a different demographic variable), and the 'vars' command (changing the family and friend scale to a different scale name). If one wanted to look at more than one demographic variable, then all that would be changed is an addition of a comma after d_gender, and naming the other variable, for example, group_by(d_gender, d_age) would show the compilation of the gender and age demographics. The output of this function is just a normal table (see Appendix C), not the prettiest thing, so the next/final step was creating better data visualizations that demonstrate the relationships between variables.

After looking at a basic outline of the data, producing better visualizations is important both aesthetically and for the representation of relationships between variables. In order to create data visualizations that demonstrated these relationships For the more complex visualizations the ggplot2 library was used (Wickham & Chang, 2016). It is an expansion from the initial code that produced a simple table. In order to make more complex visualizations, it was necessary to have a more complex code. After loading the 'group_by' syntax, the code was written to create a new variable that looked at specific columns in the table, including the demographic information, the mean, the standard deviation, the number of individuals, and minimum and maximum values of each group. Furthermore, code was added to create error bars showing the spread around the mean for each group. After this, the axis labels were made, and the finished can be found in Appendices D-G. The output of the code shows the relationship between the gender demographic information and the family and friend scale. The genders are color coded so it is easy to identify when looking at the table, and it is easier to see the differences between the groups.

Results

The main focus of this research was to look at the differences between males and females in their sensitivities to social perception. After looking at just gender information, other analyses were run to see if there were any other patterns in the data.

Gender

When initially looking at the results for gender, it was easily seen that there was indeed a difference between men and women and their averages for each scale.

Family and Friend. Table 1 shows the differences between men and women in relation to the family and friend scale. The effect size of this scale is .08.

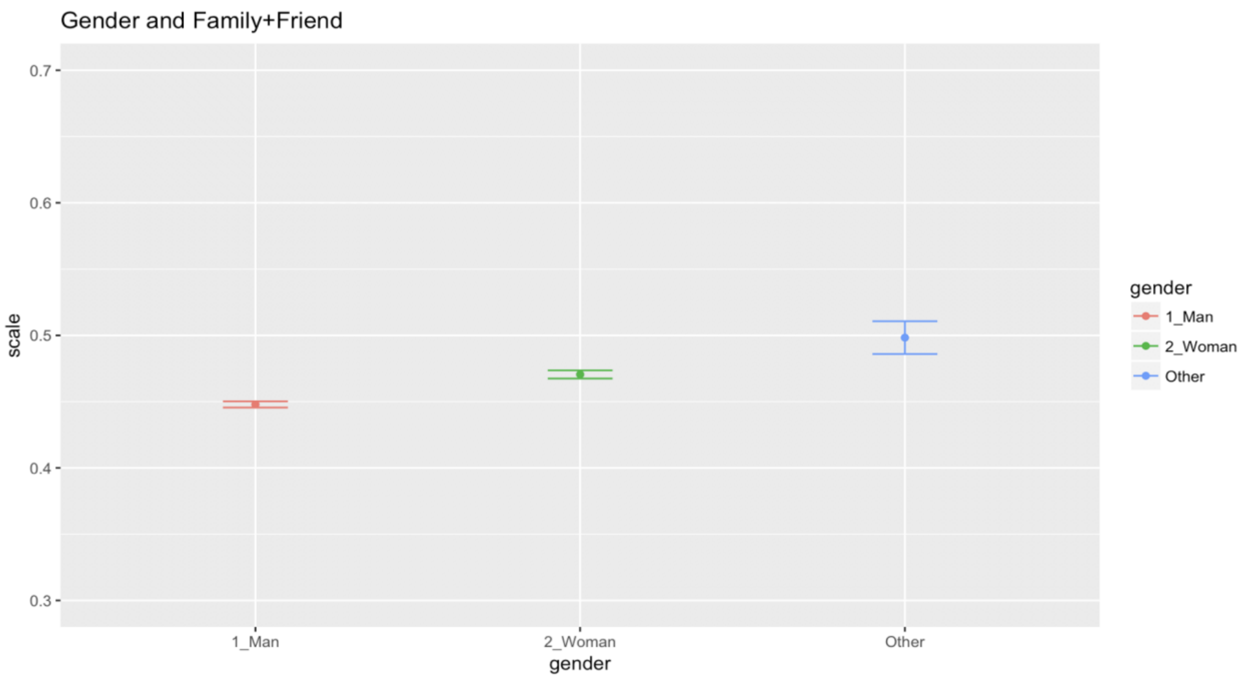


Table 1. Gender Means on Family and Friend Scale

In the table it is seen that the women's mean is higher than the men's. The family and friends scale measures the influence that family and friends have on an individual in his or her life. The women's mean is at .47, and the men's is at .45. This is only a .02 difference, but because of the nature of the dataset virtually any finding is significant. In this table, 38,995 men, 25,170 women, 2,106 others (it is important to note that the group "Other" in the gender demographic, contains individuals that responded 'other' to the gender question, as well as asexual, agender, or genderfluid, and those who did not respond which were listed as NA). The Other/NA group is composed of individuals that did not answer the demographic information questions, but did answer the other questions present in the dataset. An important thing to note is that the “Other” variable has much larger error bars, this is attributed to the size of the group. If there is a smaller group any sort of variability will cause a large effect. This is why in the larger groups, the error bars are very close to the mean, while with the other group, there is more variation. From this table it can be conclusively seen that women are more influenced by their family and friends about their decisions.

Appearance. Table 2 shows the relationship between gender and appearance. The appearance scale measures how important physical appearance is to the individual searching for a partner online. The effect size of this scale is .10.

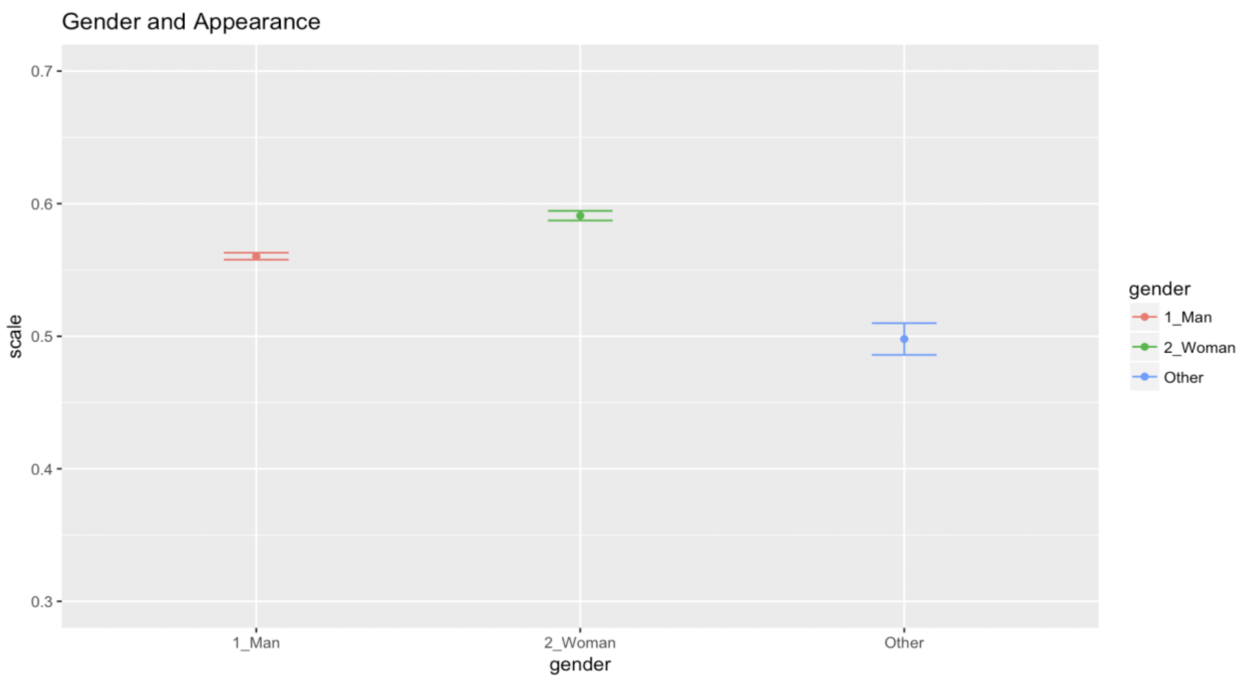


Table 2. Gender Means on Appearance Scale

From this table it is once again illustrated that women have a higher average than men. The women’s average is .59, the men’s average is .56, once again, a small difference still represents a lot. Also, it is seen again that there is more variance in the smaller group of “Other”, this is a recurring trend that is present throughout the results due to the size of the group. This table is interesting because it goes against a previous study that showed that men are more picky about appearance than women (Buss, 1989).

Education. The education scale measured the importance of education to the individual, the means are shown in Table 3. The effect size is .47.

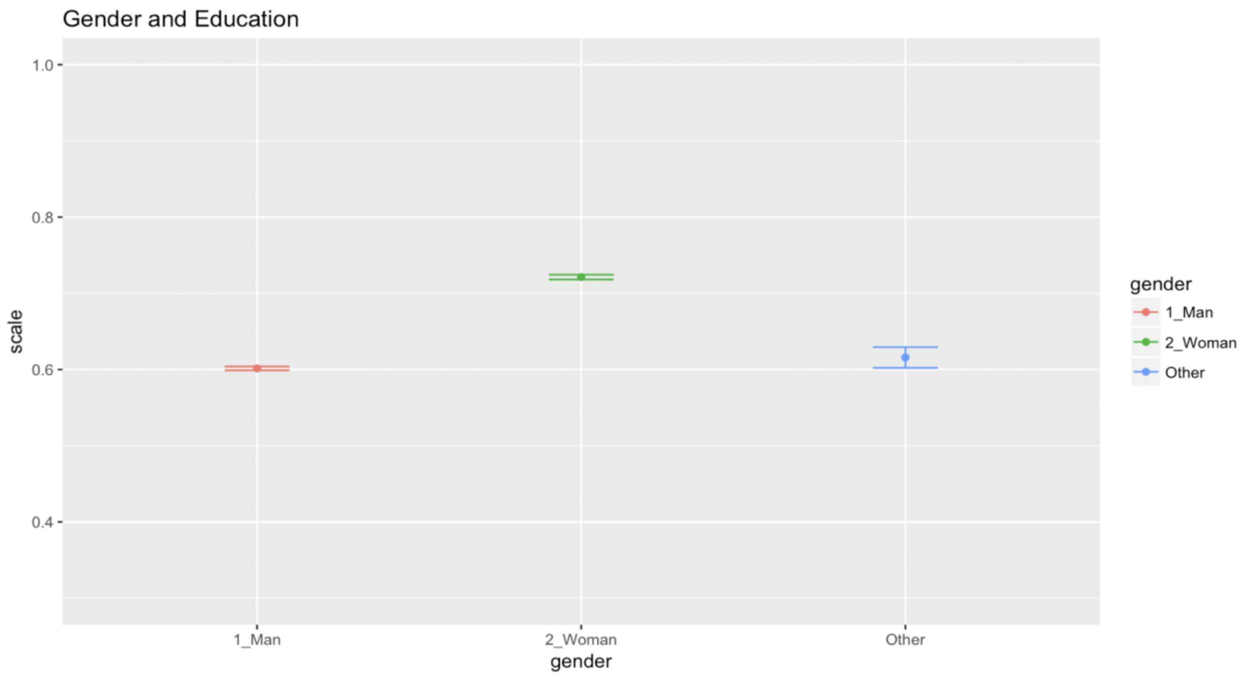


Table 3. Gender Means on Education Scale (Men= .60, Women= .72)

Moral. The moral scale measured for the importance of morals to the individual, and how important it was that the individuals morals matched with his or her potential partner. The means of the moral scale are shown in Table 4. The effect size is .45.

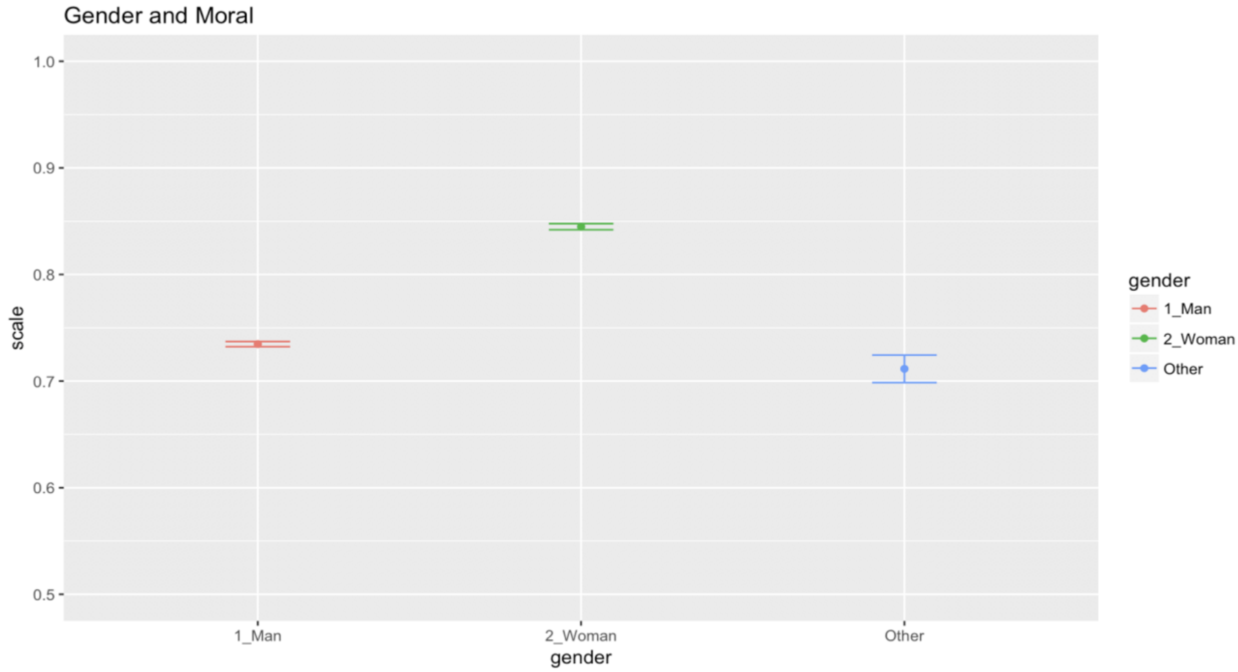


Table 4. Gender Means on Moral Scale (Men= .73, Women= .84)

Religion. The religion scale measured the religiosity of the individual, as well as how important religion was to him or her when looking for a partner. The means are shown in Table 5. The effect size is .03.

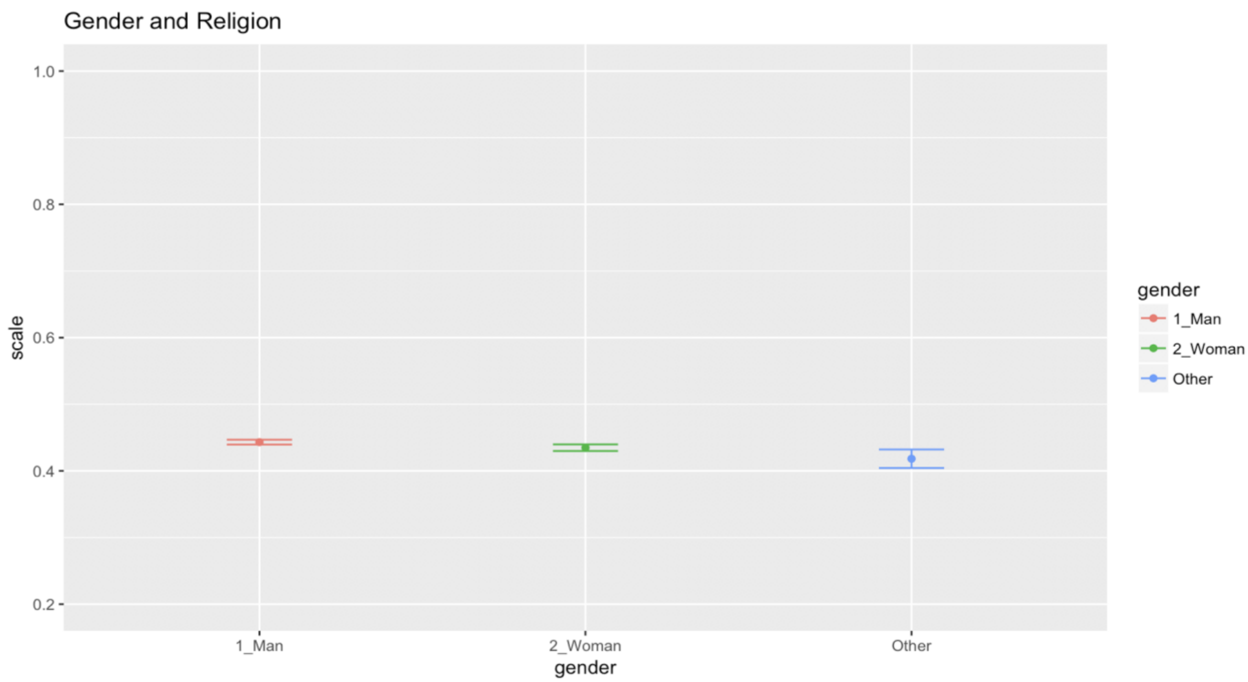


Table 5. Gender Means on Religion Scale

There is an interesting finding in this table, it shows that the men have a higher mean score on this scale than the women. The men have an average of .44 while the women have an average of .43, this is the only instance in which men have a higher mean score than women. The only possible error is that the religion questions had a very low response rate that could have skewed the data, but from what the dataset contained, it shows that men are more influenced by religion than women.

Stigma. The stigma scale measured for stigma aversion, meaning that it measures how likely someone is to avoid stigma. The means are shown in Table 6. The effect size is .09.

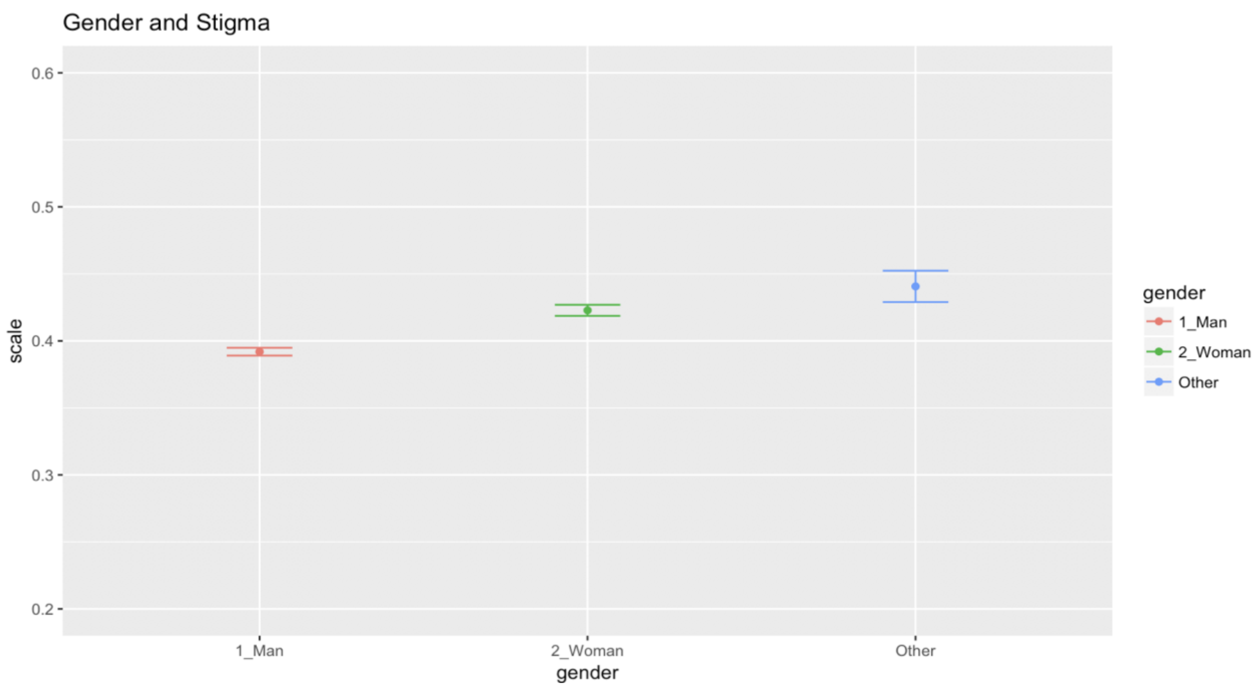


Table 6. Gender Means on Stigma Scale (Men= .39, Women= .42)

Substance Use. The substance use scale measured the individuals aversion to substance use, both legal and illegal. The means are shown in Table 7. The effect size is .19.

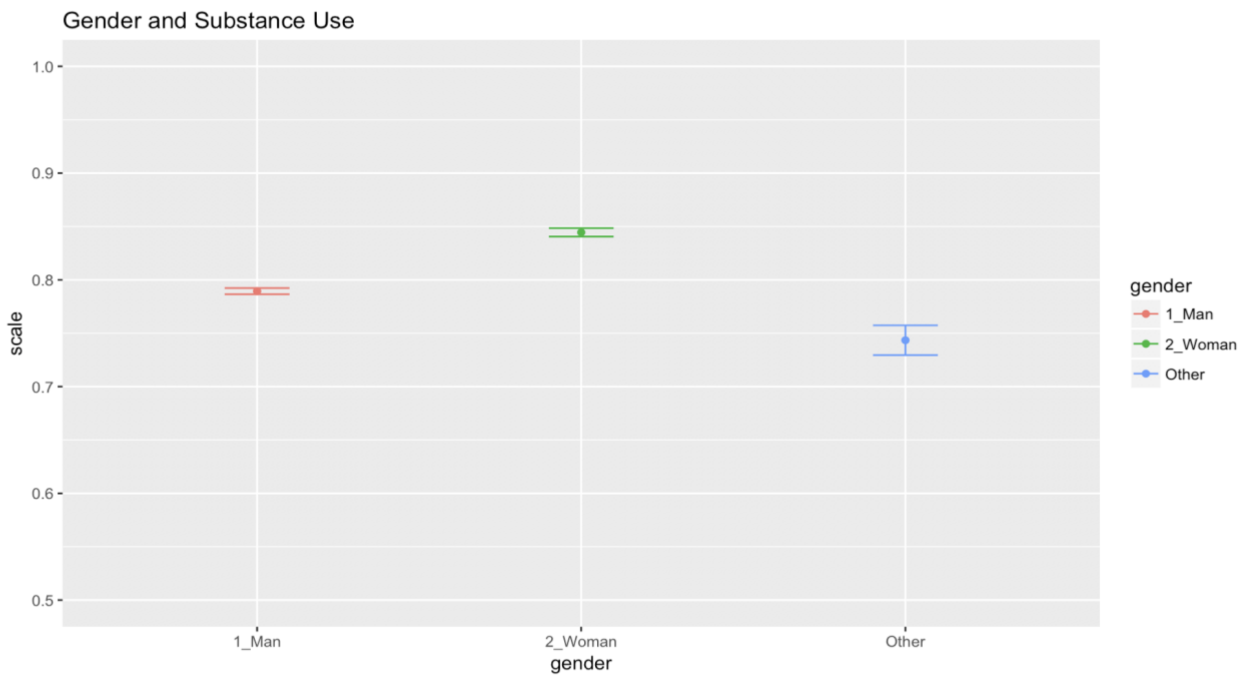


Table 7. Gender Means on Substance Use Scale (Men= .79, Women= .84)

Gender and Scales

It can be seen that in six of the seven scales, women had higher average scores. The only instance in which men had a higher mean than women was on the religion scale. After looking at the individual gender means, it seemed necessary to look at more complex variables, for example more than one demographic variable.

Combined Demographic Means

After looking at just gender, the age demographic was added in to evaluate any generational effects that might be present in the data.

Gender, Age & Family + Friend. Table 8 shows the combined averages of gender, age and the family and friend scale.

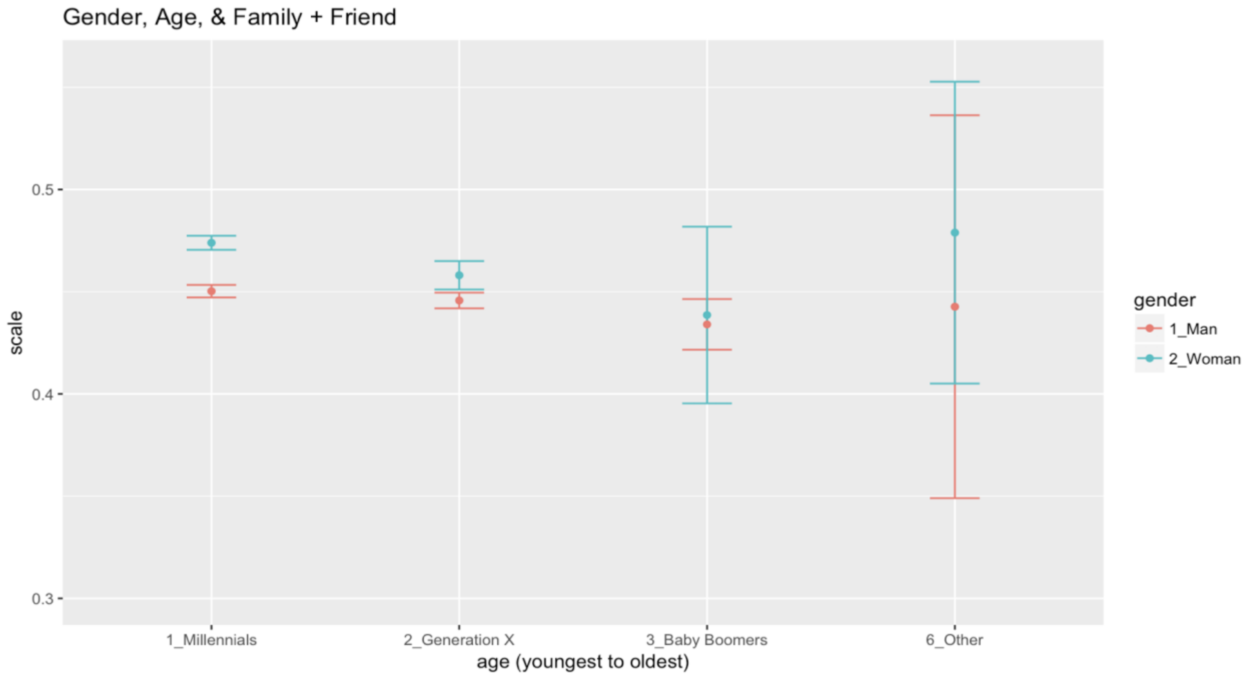


Table 8. Gender and Age Means on Family + Friend Scale

It can be seen that throughout the generations, women still have higher means than men on the family and friend scale. The means with the larger error bars are representing smaller groups which is why there is so much more variation, while the ones that represent larger groups have smaller error bars, showing less variation and a more representative sample.

Gender, Age & Appearance. Table 9 shows the relationship between gender, age, and the averages on the appearance scale

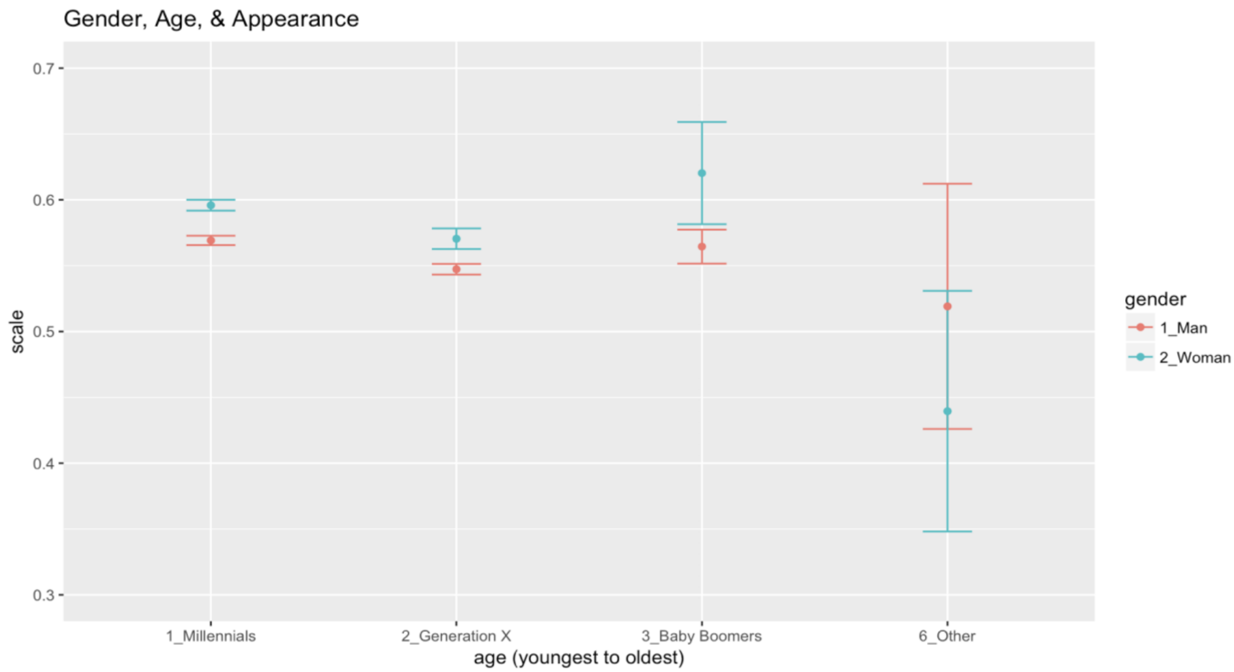


Table 9. Gender and Age Means on Appearance Scale

It can be seen that in all age groups except '6_Other' that women have significantly higher average scores on the appearance scale. The difference in the '6_Other' group could be due to the size of the group, which is not a large enough sample to show a significant effect. Even in the generational breakdown it shows that women are more concerned with appearance than men. The only group where this is not the case is with the '6_Other' group, this might be due to the age differences, because this group has people over the age of 90. This finding goes against what Buss found in his 1989 study in which he found that across cultures males are more selective about appearance than women. This is one of the most interesting results in this paper, as it directly goes against a previous study. This could be due to the nature of online dating because individuals might see it as a more short-term rather than long-term relationship which would support findings that women are more selective for appearance in short-term relationships (Regan et al. 2000).

Gender, Age & Education. Table 10 shows the relationship between gender, age, and education.

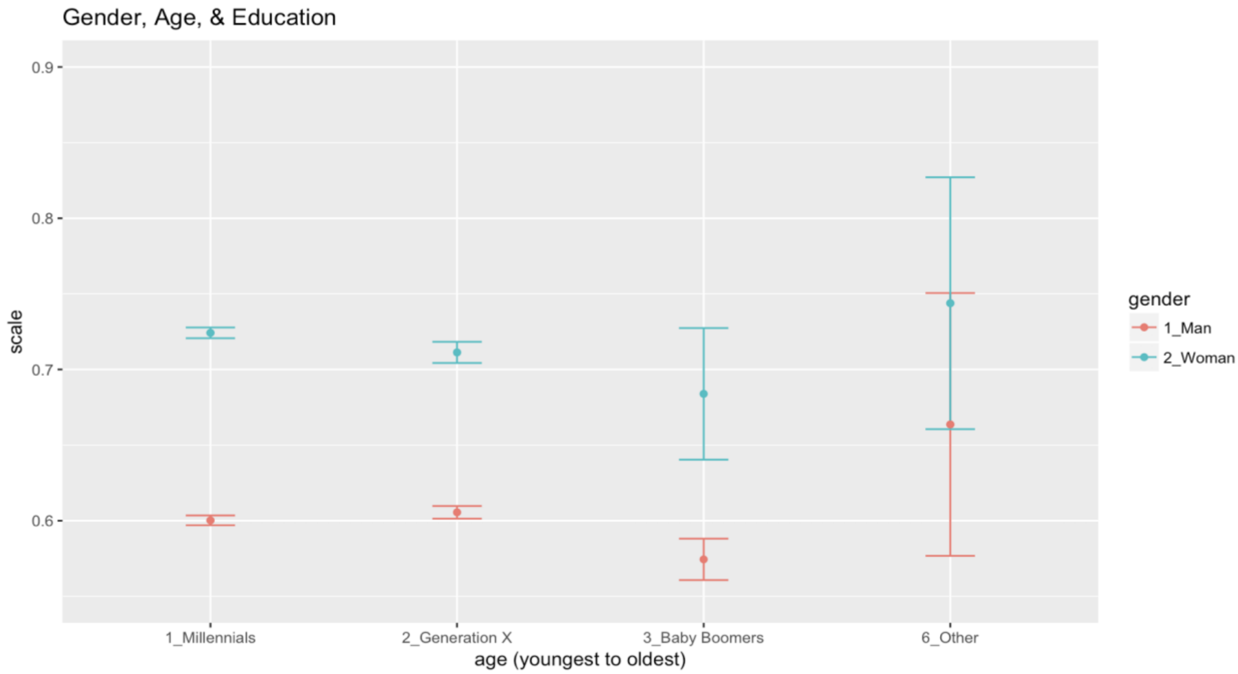


Table 10. Gender and Age Means on Education Scale

The table demonstrates that women are higher on the education scale, the other group is once again surpassing both genders, but this is attributed to the size of the group rather than the actual effect size.

Gender, Age & Moral. Table 11 shows the relationship between gender, age, and the averages on the moral scale.

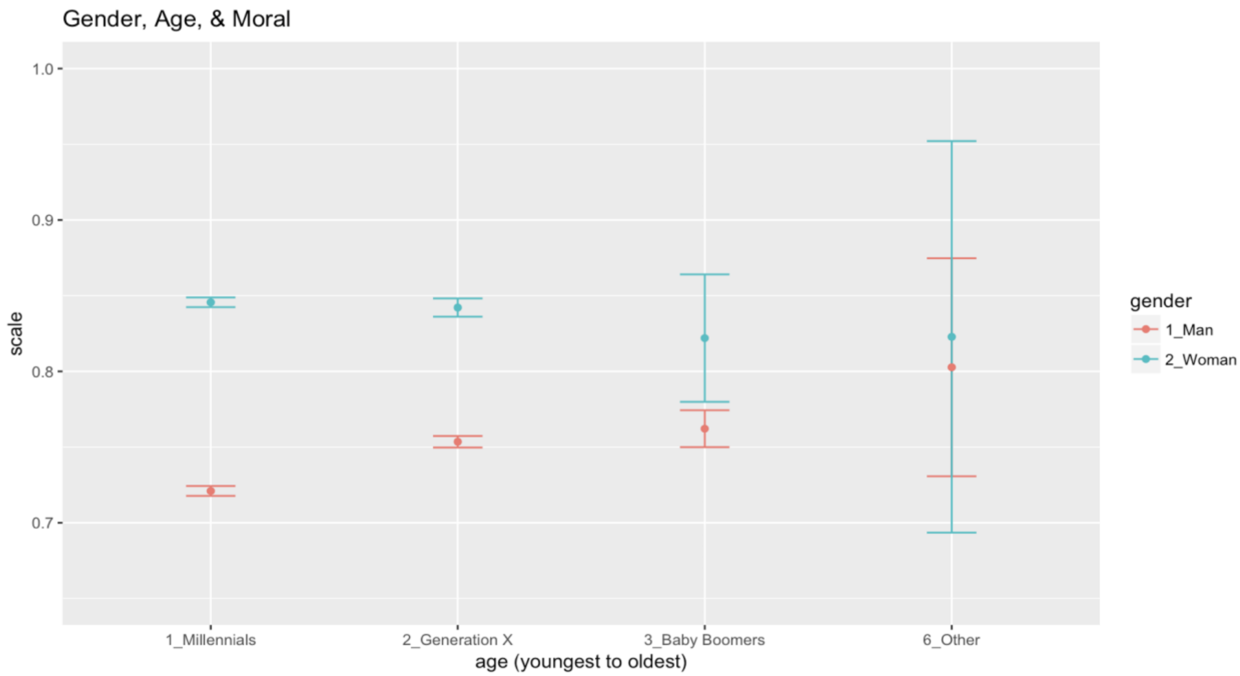


Table 11. Gender and Age Means on Moral Scale

Gender, Age & Religion. Table 12 shows the age and gender averages on the religion scale.

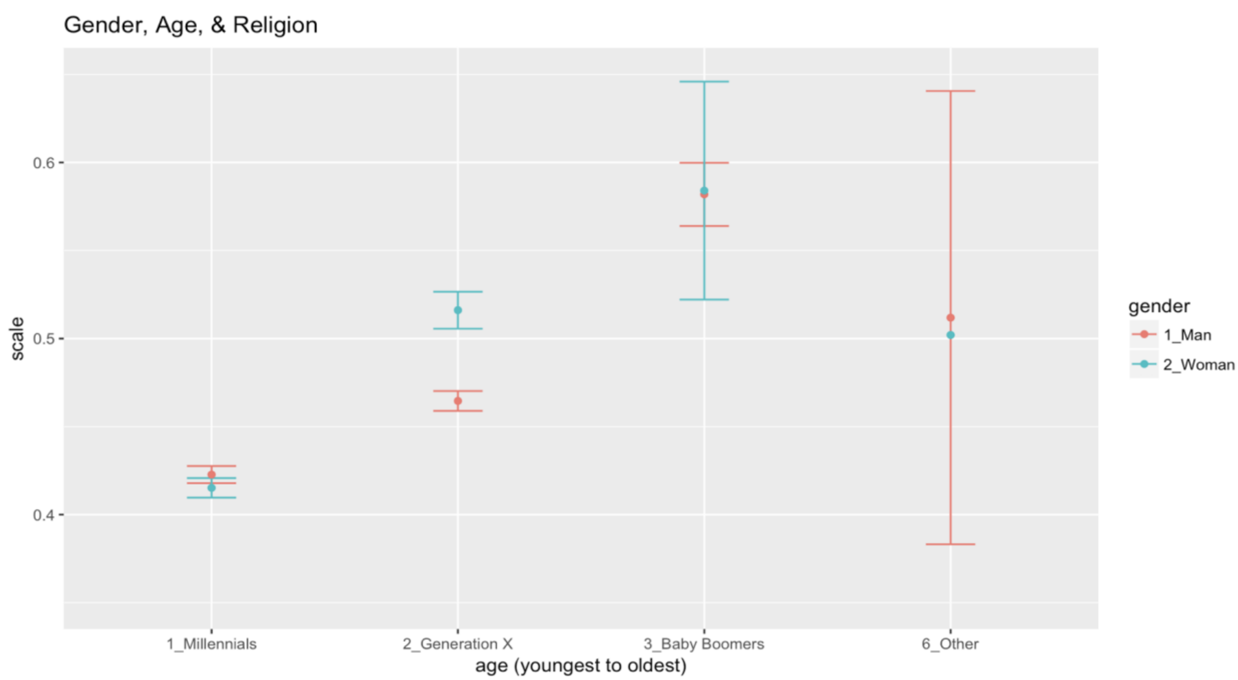


Table 12. Gender and Age Means on Religion Scale

This is of interest because in the previous section, the religion scale was the only one that men hat a higher mean for than women. It is seen here that the Millennial men have a higher average on the scale than women, and in the Baby Boomers group, the average for religion is essentially the same. This explains the trend found in the gender and religion data.

Gender, Age & Stigma. Table 13 shows the relationship between gender and age averages on the stigma scale.

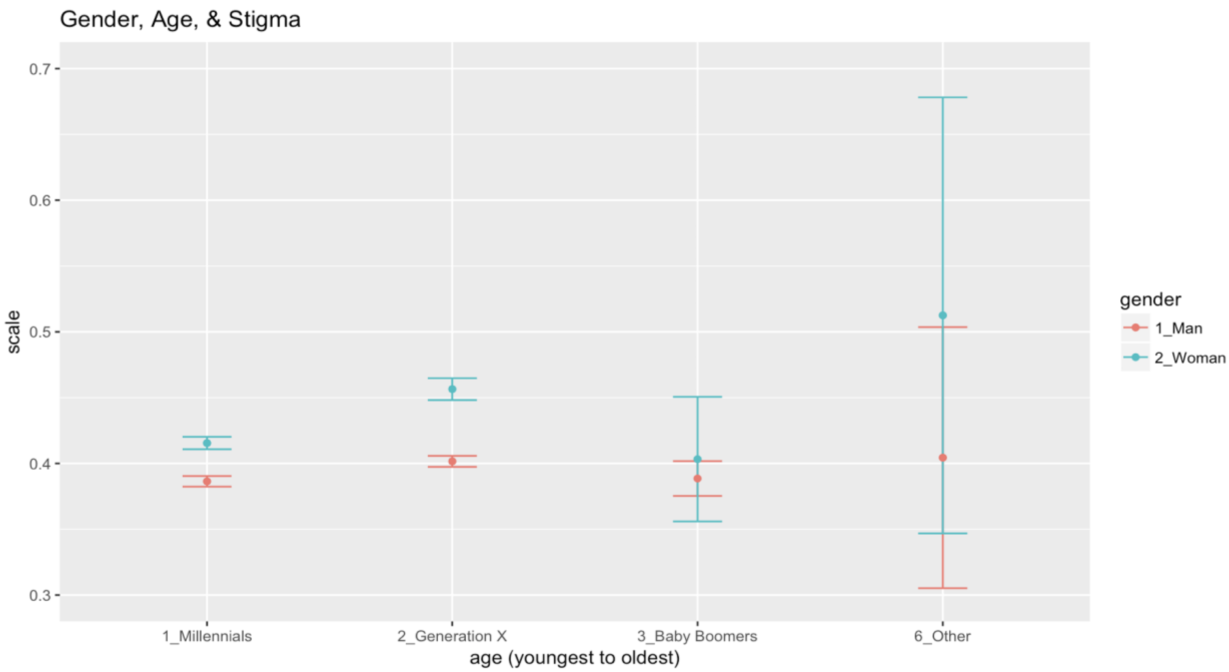


Table 13. Gender and Age Means on Stigma Scale

Gender, Age & Substance Use. Table 14 shows the relationship between gender and age averages on the substance use scale.

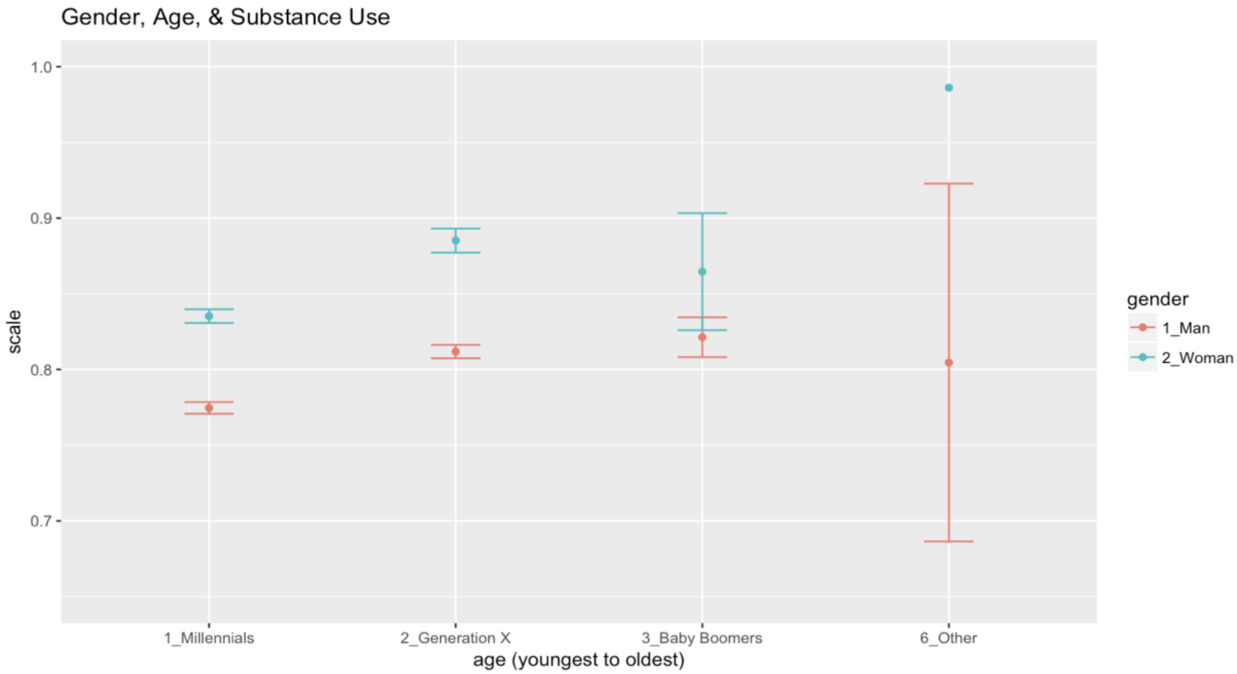


Table 14. Gender and Age Means on Substance Use Scale

Gender, Ethnicity & Education. Table 15 shows the relationship between gender and ethnicity on education scale averages.

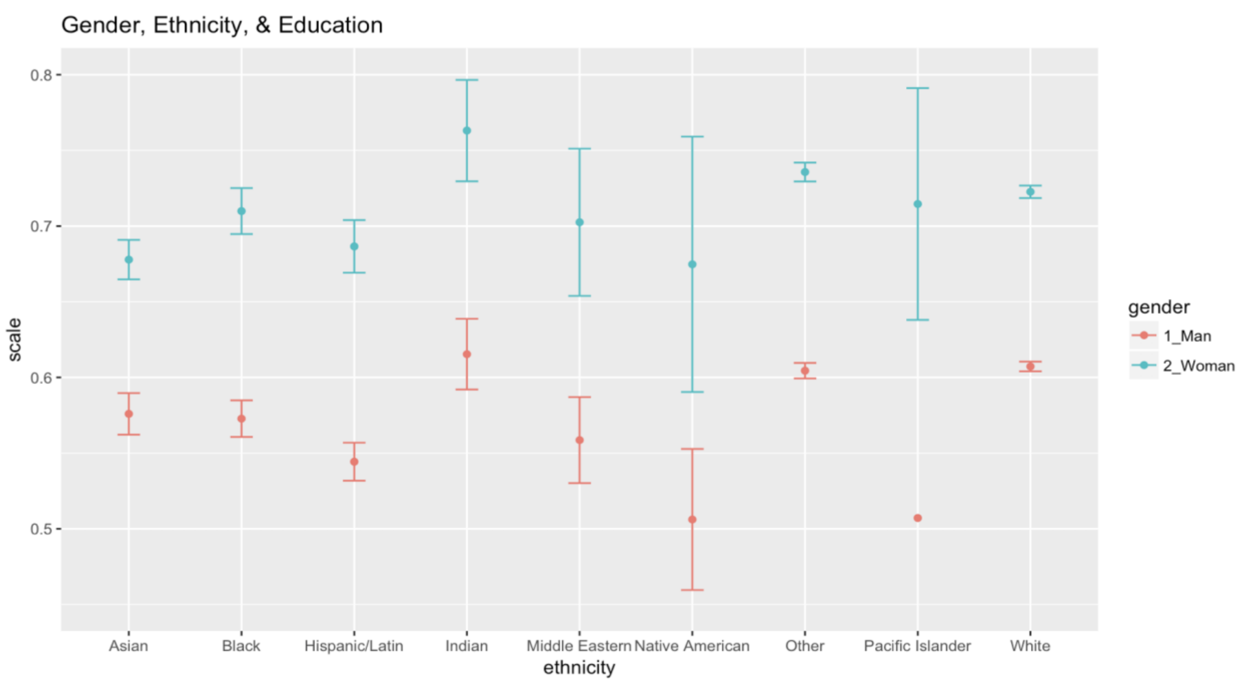


Table 15. Gender and Ethnicity Means on Education Scale

Gender, Ethnicity, & Moral. Table 16 shows the relationship between gender and ethnicity on the moral scale means.

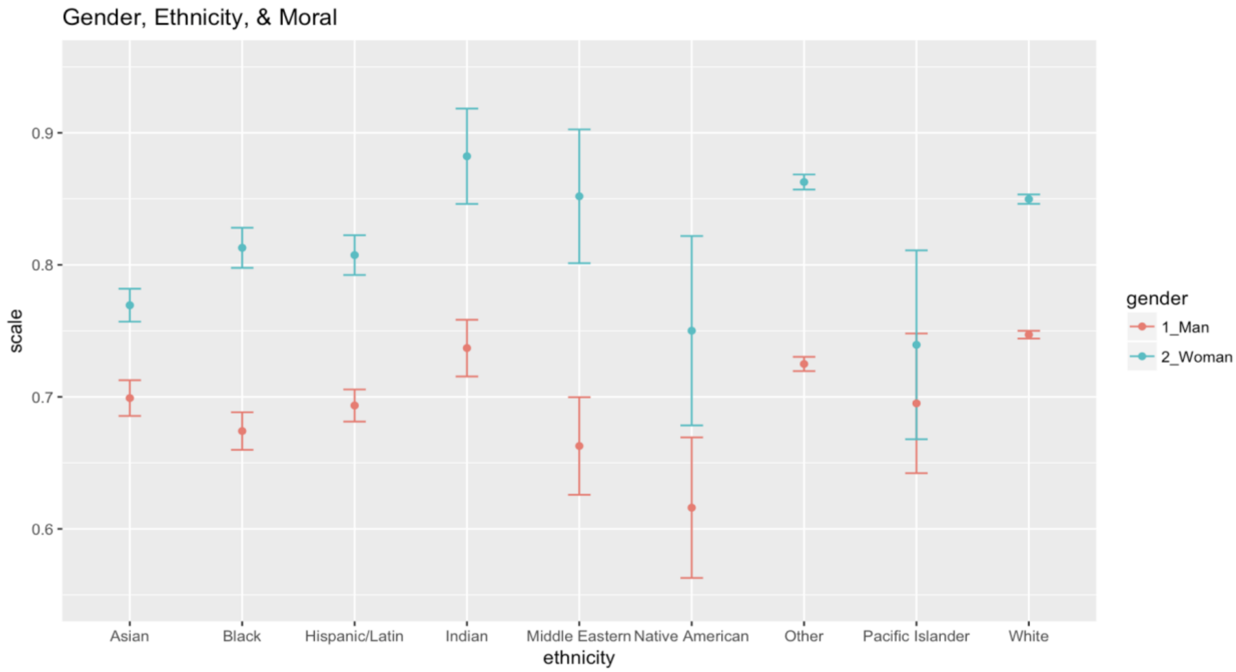


Table 16. Gender and Ethnicity Means on Moral Scale

While this is a busy table, it is seen that consistently women are higher on the scale than men in each ethnic group. Furthermore, it is important to note that the groups with larger error bars have less individuals in them which accounts for the variance in the data.

Gender, Ethnicity, & Religion. Table 17 shows the relationship between gender and ethnicity on the means of the religion scale.

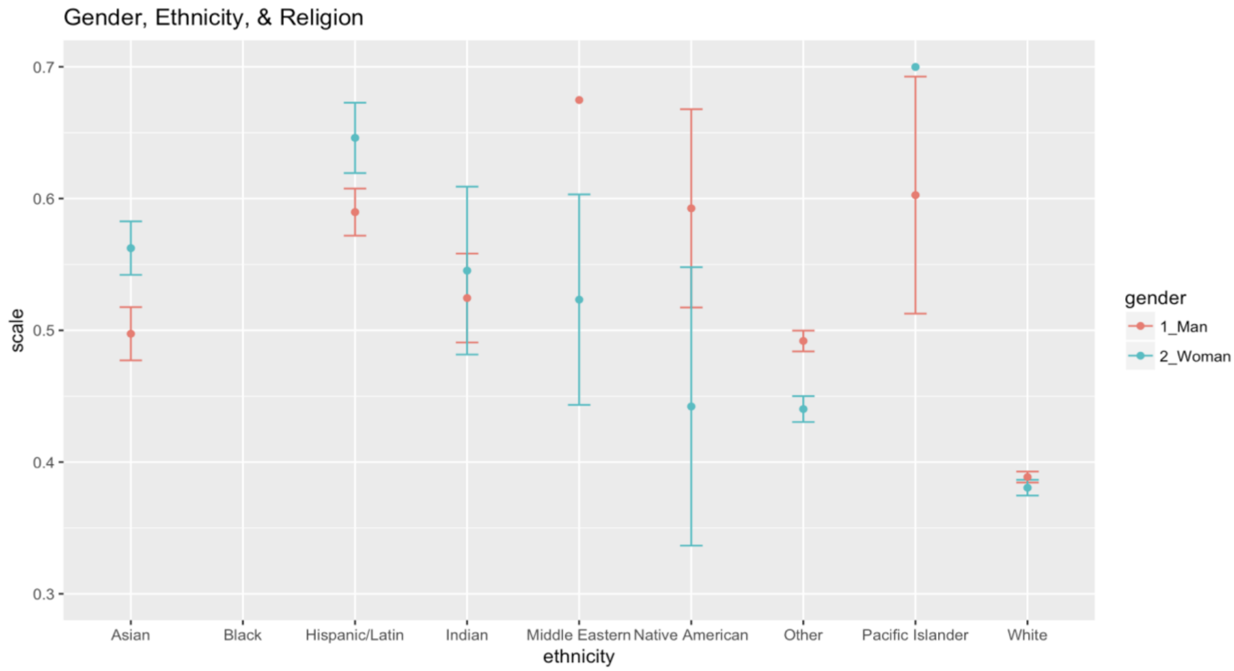


Table 17. Gender and Ethnicity Means on Religion Scale

Once again, religion is the only scale in which more men scored highly than women, it is seen here that the men from the Middle Eastern, Native American, and White ethnicities rated religion higher and had higher averages.

Gender, Ethnicity, & Stigma. Table 18 shows the relationship between gender and ethnicity means on the stigma scale.

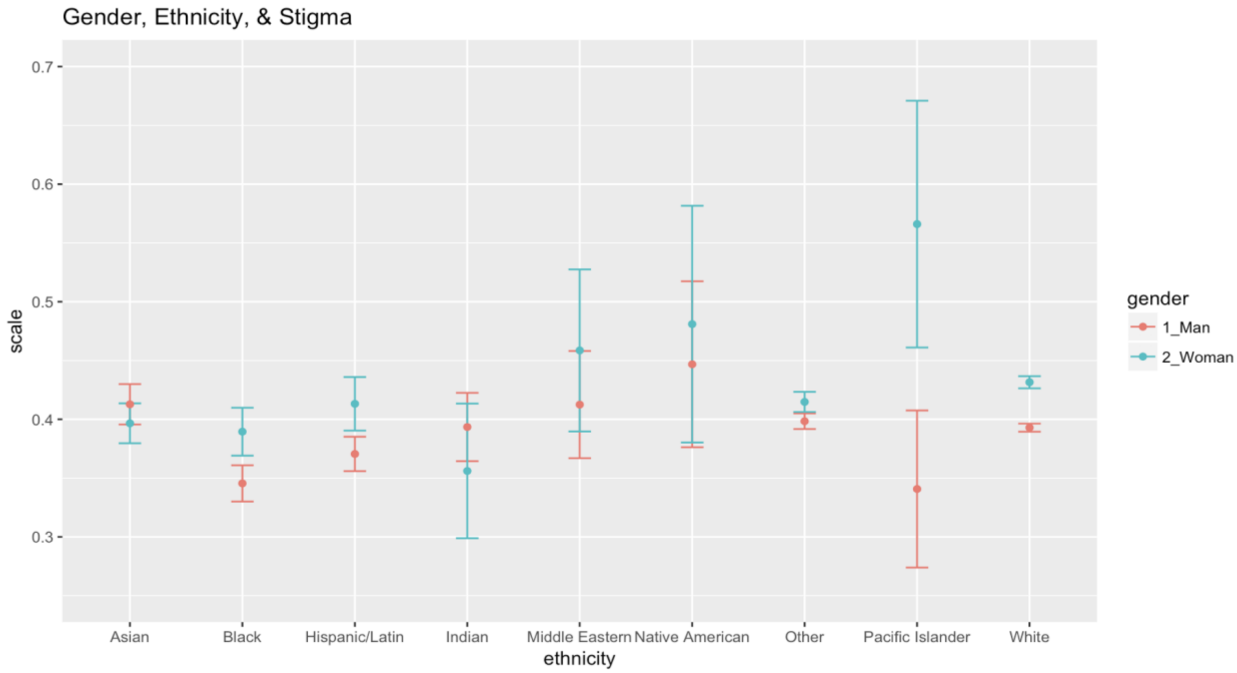


Table 18. Gender and Ethnicity Means on Stigma Scale

There is an interesting trend in this table, it is seen that both Asian men and Indian men scored higher than the females of that ethnicity on the stigma scale. More research should be done to see why this occurred. But in other ethnicities, women still maintained higher averages than men.

Discussion

Gender, Generation, & Ethnicity

From the results, it can be seen that women are more selective about various aspects of societal perception than men. The only instance in which men scored higher was in the Religion scale. The women were more concerned with appearance, morals, education, and family and friend approval. They were also more averse to stigmatized individuals and ones that used substances. The results show that in the larger groups the means are significantly representative of the population. This is seen through the size of the error bars. In the gender demographic, the error bars were small for both men and women, but were larger for the ‘NA’ and ‘Other’ groups. This is because both the ‘NA’ and ‘Other’ groups had smaller sizes, which allows for even a slight variation to have an immense impact on the means. With larger groups a more representative sample is produced, which allows for predictions to be made from the data.

When the age demographic was included, it was easier to see the specific effects of the scales. Across all the generations on all the scales, women still scored higher than men. The only exception was, once again, in the religion scale. This is because in the "Millennial" group, men scored higher on the scale, this was also true of the "Baby Boomers". A possible explanation for this trend is that the religion questions had the highest rates of unanswered questions from each individual. In order to have solid evidence for the difference in religion averages, further studies should be conducted to evaluate the topic. Another dimension that was added when factoring in the age demographic was the obvious breakdown of individuals that used the dating site. It is seen that it is more common among the younger generations, and tapers off for the older generations. After looking at the age demographic, it seemed interesting to look at the ethnicity demographic.

The ethnicity demographic data was a little messy. The groups could not be combined because they are independent ethnicities and combining them would cause them to lose representation, even though the representation was limited due to the small size of some groups. From the seven scales, a few of the scales stood out in relation to the ethnicity and gender data these included the education, moral, religion, and stigma scales. For the education scale it was shown that no matter the ethnicity, women scored higher on the education scale than men, this trend was also echoed in the moral scale. There is a significant gap between the scores of the two genders. For the religion scale it was seen that men from the Middle Eastern, Native American, and White ethnicities rated religion higher and had higher averages. This echoes the data from the other demographic analyses. Finally, the stigma scale was interesting because it showed two ethnicities in which males were more concerned with stigma than women the ethnicities are Asian men and Indian men. This shows that there are many ways the data can be evaluated, but the overarching theme is that women are more selective when it comes to searching for a partner, which connects to parental investment theory.

Parental Investment Theory

The findings in this study provide support for Trivers' theory of parental investment. Women appear to be more likely to be influenced by societal perception when choosing a romantic partner. Trivers' posited that women are rewarded for their selectivity when it comes to finding a mate. This is because women have a much higher risk when entering a relationship. If a woman gets pregnant, it has much more of an effect on her than it does on the male. This investment is very long term, including the 9 months of carrying the child through raising it into young adulthood. Men have the ability to help along the way, but women have to deal with the most up front responsibility. Because of this, women have evolved to be more selective when it comes to mates. If they pick an individual that conforms with social norms, that has a stable life, is attractive, and is able to support a family, then the woman does not have to worry as much about the state of the relationship if she gets pregnant, however, if she is not selective, then it could have a different outcome.

Trivers explains that evolutionarily the goal of males in each species is to spread as much of their genetic material as possible. A main tenant of evolution is making sure that an individual's genes are passed on to other generations. In most species, the females are the ones that carry the offspring to term and then take care of them after they are born. Males do not experience the same pressure. If a woman is not selective, there are several consequences that she might face. First, if a woman gets into a relationship with someone that her kin group does not like or approve of, in ancestral times, she would likely have been exiled which would have made her life much more difficult, and would basically be a death sentence. Another consequence of not being selective is that a woman could get into a relationship and get pregnant and the man could leave, causing her to be the only one responsible for the child, in which case she could terminate the pregnancy or choose to carry the child to term, either choice involves considerable investment on the woman's part. It is unfortunate that women have to deal with the consequences of not being selective. The consequences are due to the woman's investment in her offspring. The women have a disproportionately large investment up front, and sometimes men make up for it in the end, but the woman has to take the risk in the first place. The results support the idea that women are more selective when it comes to finding a partner as a function of parental investment theory. This is a trend that crosses through generations.

Challenges and Limitations

Throughout this project there were many challenges and limitations that became obvious. Initially, the whole language of R had to be learned in order to achieve any of the data analysis. Learning the program took considerable time and effort, but once achieved the project ran more smoothly. Another challenge was the sheer size of the dataset. The dataset contained almost 70,000 responses to about 2,600 questions, going through and looking at each question took time, and when recoding them there were many grammatical errors in questions that had to be accounted for or the program would not run the analysis. Furthermore, the dataset had a lot of holes in it. Due to the size of the dataset, a lot of individuals did not answer that many questions. The initial researchers that published the data tried to filter out only users that answered a considerable chunk of questions, but even still there were many blank answers that had to be calculated. A challenge that came along with this was that a lot of questions had to be cut from the possible pool because of the lack of responses. It does not make sense to use a question that only a few people answered. After filtering through those questions there were some minor programming challenges. The programming challenges mostly had to do with syntax, and learning about R throughout the process of the project. A main challenge was the source of the data.

The data came from OkCupid which is an online dating site and had significant limitations. It is a useful wealth of information, however it is not ideal for trying to run a psychological study. A lot of the questions, while being related to psychology, were not measuring for any sort of psychological idea. The limitation was that the questions in the dataset sometimes would have gaps, for example when a race scale was trying to be constructed, it was found that in the whole set of about 2,600 questions, only five of them had to do with race, and they had very few respondents. Furthermore, another limitation was that not every individual answered every question for both the regular questions and the demographic information. If all of the questions were answered then it would have provided a more accurate measure of the effects of societal perception. If the questions were phrased differently or intentionally created to measure for the influence of societal perception that would have been an improvement.

References

Alterovitz, S. S., & Mendelsohn, G. A. (2009). Partner preferences across the life span: Online dating by older adults. *Psychology of Popular Media Culture,1*, 89-95.

Ansari, A., & Klinenberg, E. (2015). *Modern Romance*. London: Penguin Books.

Boberg, J. L. (2008). Picking a partner in young adulthood: The impact of education and gender on mate preferences. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, *68*(8-B), 5558.

Braithwaite, S. R., Coulson, G. L., Spjut, K., Dickerson, W., Beck, A. R., Dougal, K., . . . Jones, D. (2015). The influence of religion on the partner selection strategies of emerging adults. *Journal of Family Issues, 36*(2), 212-231.

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences,12*, 1-49.

Clark, S. C., Dover, A. M., Geher, G., & Presson, P. K. (2005). Perceptions of self and of ideal mates: Similarities and differences across the sexes. *Current Psychology,24*(3), 180-202.

Conger, R. D., Cui, M., Bryant, C. M., & Elder, G. H., Jr. (2000). Competence in Early Adult Romantic Relationships: A Developmental Perspective on Family Influences. *Journal of Personality and Social Psychology,79*(2), 224-237.

Dimock, M. (2018, March 1). Defining generations: Where Millennials end and post-Millennials begin. Retrieved from http://www.pewresearch.org/fact-tank/2018/03/01/defining-generations-where-millennials-end-and-post-millennials-begin/

Festinger, L., Scachter, S., & Black, K. (1968). The Operation of Group Standards. *Group Dynamics,*152-164.

Finkel, E. J., Eastwick, P. W., Karney, B. R., Reis, H. T., & Sprecher, S. (2012). Online Dating: A Critical Analysis From the Perspective of Psychological Science. *Psychological Science in the Public Interest,13*(1), 3-66.

Fischer, I. (2009). Friend or Foe: Subjective Expected Relative Similarity as a Determinant of Cooperation. *Journal of Experimental Psychology: General,138*(3), 341-350.

Fuller-Iglesias, H. R., Webster, N. J., & Antonucci, T. C. (2015). The Complex Nature of Family Support Across the Life Span: Implications for Psychological Well-Being. *Developmental Psychology,51*(3), 277-288.

Gibbs, J. L., Ellison, N. B., & Lai, C. (2011). First Comes Love, Then Comes Google: An Investigation of Uncertainty Reduction Strategies and Self-Disclosure in Online Dating. *Communication Research,38*(1), 70-100.

Gorelik, G., Shackelford, T. K., & Salmon, C. A. (2010). New Horizons in the Evolutionary Science of the Human Family. *Review of General Psychology,14*(4), 330-339.

Huxhold, O., Fiori, K. L., & Windsor, T. D. (2013). The Dynamic Interplay of Social Network Characteristics, Subjective Well-Being, and Health: The Costs and Benefits of Socio-Emotional Selectivity. *Psychology and Aging,28*(1), 3-16.

Kerr, N. L., & Levine, J. M. (2008). The Detection of Social Exclusion: Evolution and Beyond. *Group Dynamics: Theory, Research, and Practice,12*(1), 39-52.

Kirkegaard, E. O., & Bjerrekær, J. D. (2016). The OKCupid dataset: A very large public dataset of dating site users. *Open Differential Psychology*.

Kline, P. (1999). The handbook of psychological testing (2nd ed.). London: Routledge.

Koleva, S. P. (2012). Birds of a moral feather: The role of morality in romantic attraction and relationship satisfaction. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 72*(12-B), 7743.

McKinlay, C. (2014). *Optimal Cupid: Mastering the Hidden Logic of OkCupid*. Christopher McKinlay.

Menkin, J. A., Robles, T. F., Wiley, J. F., & Gonzaga, G. C. (2015). Online Dating Across the Life Span: Users' Relationship Goals. *Psychology and Aging,30*(4), 987-993.

Merkle, E. R., & Richardson, R. A. (2000). Digital Dating and Virtual Relating: Conceptualizing Computer Mediated Romantic Relationships. *Family Relations,49*, 187-192.

Oesch, N., & Miklousic, I. (2012). The Dating Mind: Evolutionary Psychology and the Emerging Science of Human Courtship. *Evolutionary Psychology,10*(5), 899-909.

Preston, K., Gottfried, A. W., Oliver, P. H., Gottfried, A., Delany, D. E., & Ibrahim, S. M. (2016). Positive Family Relationships: Longitudinal Network of Relations. *Journal of Family Psychology,30*(7), 875-885.

Regan, P. C., Levin, L., Sprecher, S., Christopher, F. S., & Cate, R. (2000). Partner preferences: What characteristics do men and women desire in their short-term sexual and long-term romantic partners? *Journal of Psychology & Human Sexuality*, *12*(3), 1-21.

Rudder, C. (2015). *Dataclysm: Who We Are - When We Think No One's Looking*. Broadway Books.

Sander, J., Schupp, J., & Richter, D. (2017). Getting Together: Social Contact Frequency Across the Life Span. *Developmental Psychology,53*(8), 1571-1588.

Sprecher, S., & Felmlee, D. (1992). The Influence of Parents and Friends on the Quality and Stability of Romantic Relationships: A Three-Wave Longitudinal Investigation. *Journal of Marriage and the Family,54*(4), 888-900.

Sprecher, S. (2011). The Influence of Social Networks on Romantic Relationships: Through the Lens of the Social Network. *Personal Relationships,18*, 630-644.

The R Project for Statistical Computing. (n.d.). Retrieved from https://www.r-project.org/

Trivers, R. (1972). *Parental investment and sexual selection*. Cambridge, MA: Biological Laboratories, Harvard University.

Wickham, H., & Chang, W. (2016, December 30). Package 'ggplot2'. Retrieved from https://cran.r-project.org/web/packages/ggplot2/ggplot2.pdf

Wickham, H., Francois, R., Henry, L., & Muller, K. (2017, September 28). Package 'dplyr'. Retrieved from https://cran.r-project.org/web/packages/dplyr/dplyr.pdf

Appendix

Appendix A. Recodes

Read in and Recodes

Sam DiPiero

December 13, 2017

Add descriptive text here

Break questions into blocks corresponding to different anticipated scales or content areas

Initial car::recodes for OKCupid dataset

```
# change your code chunk options (e.g., echo =)
# so that the output is the way you like it
# consider naming your chunks too
knitr::opts_chunk$set(echo = TRUE)
library(car)

## Warning: package 'car' was built under R version 3.4.4

## Loading required package: carData

## Warning: package 'carData' was built under R version 3.4.4

library(psych)

## Warning: package 'psych' was built under R version 3.4.4

##
## Attaching package: 'psych'

## The following object is masked from 'package:car':
##
##     logit

library(knitr)
library(papeR)

## Loading required package: xtable

##
## Attaching package: 'papeR'
```

```
## The following object is masked from 'package:utils':
##
##     toLatex
```

```
library(tidyverse)
```

```
## ── Attaching packages ──────────────────────────── tidyverse 1.2.1 ──

## ✔ ggplot2 2.2.1     ✔ purrr   0.2.4
## ✔ tibble  1.4.2     ✔ dplyr   0.7.4
## ✔ tidyr   0.8.0     ✔ stringr 1.3.0
## ✔ readr   1.1.1     ✔ forcats 0.3.0

## ── Conflicts ─────────────────────────────── tidyverse_conflicts() ──
## ✖ ggplot2::%+%()     masks psych::%+%()
## ✖ ggplot2::alpha()   masks psych::alpha()
## ✖ dplyr::filter()    masks stats::filter()
## ✖ dplyr::lag()       masks stats::lag()
## ✖ dplyr::recode()    masks car::recode()
## ✖ purrr::some()      masks car::some()
## ✖ dplyr::summarise() masks papeR::summarise()
## ✖ dplyr::summarize() masks papeR::summarize()
```

```
#library(lavaan)
```

This is the first block from start.rmd. Copied here to retain file paths.

```
# options for whole doc
knitr::opts_chunk$set(echo = TRUE)
options(digits = 2)
#setwd("/Users/samanthadipiero/OneDrive/z_OKCupid")
getwd()
```

```
## [1] "/Users/samanthadipiero/OneDrive/z_OKCupid"
```

```
#maindir <-("/Users/samanthadipiero/OneDrive/z_OKCupid")
#secondSmall <- read.csv("C:/Users/samanthadipiero/OneDrive/z_OKCupid/data/se
condSmall.csv")

# next line is for my surface pro
 maindir <-getwd()#"C:/Users/the_l/OneDrive/z_OKCupid"
scriptsdir <-(file.path(maindir,"scripts"))
resultsdir <-(file.path(maindir,"results"))
datadir <- (file.path(maindir,"data"))
NOTsecondSmall <- read.csv(file.path(datadir, 'NOTsecondSmall.csv'))
```

All Sam items

```
mydemographics <- c("d_ethnicity","d_age","d_gender")

myvars<- c( "q13", "q14", "q16", "q41", "q42", "q43", "q65", "q70", "q71", "q
```

```
77", "q78", "q79", "q80", "q122", "q124", "q125", "q126", "q127", "q128", "q1
29", "q134", "q155", "q156", "q193", "q210", "q252", "q298", "q332", "q350",
"q382", "q383", "q395", "q397", "q401", "q427", "q499", "q501", "q546", "q723
", "q868", "q1117", "q1119","q1138", "q1287", "q1495", "q1618", "q2017", "q54
17", "q6258", "q6971", "q6988", "q7079", "q7920", "q8197", "q8215", "q9415",
"q9418","q9589", "q9688", "q12560", "q12949", "q13006", "q13032", "q13033", "
q13077", "q13669", "q14649", "q14663", "q14884", "q15063", "q15261", "q15280"
, "q15372", "q15742", "q15743", "q15778", "q15872", "q15889", "q16288", "q163
17", "q16714", "q16807", "q17168", "q18699", "q18875", "q18935", "q18955", "q
19125", "q19148", "q19198", "q19365", "q19566", "q19690", "q19815", "q19820",
"q19915", "q20121", "q20424", "q20530", "q20661", "q20755", "q21729", "q22991
", "q23401", "q23779", "q25107", "q26600", "q26742", "q27915", "q28537", "q31
488", "q33107", "q33625", "q35281", "q35555", "q35856", "q35997", "q36235", "
q36355", "q36671", "q37440", "q37936", "q38072", "q38186", "q39189", "q40194"
, "q41393", "q42190", "q42406", "q42665", "q43237", "q43261", "q46301", "q466
47", "q48260", "q48372", "q48553", "q49016", "q49038", "q49660", "q52682", "q
57694", "q57854", "q57886", "q59456", "q59457", "q61705", "q61786", "q64923",
"q68746", "q70218", "q80404", "q81259", "q81847", "q82618", "q82755", "q83547
", "q85440", "q85583", "q85773", "q85932", "q86761", "q212814", "d_ethnicity"
, "d_age", "d_gender")

#Family and Friends
#"q350", "q401", "q427", "q1495", "q9589", "q14649", "q15261", "q15742", "q15
778", "q16288", "q16714", "q19125", "q20424", "q20755", "q22991", "q23779", "
q35281", "q35555", "q35856", "q37440", "q42406", "q43261", "q48553", "q49660"
, "q61705", "q85440", "q85773",

#Appearance
#"q122", "q124", "q125", "q126", "q127", "q134", "q298", "q382", "q383", "q70
79", "q13669", "q14663", "q15872", "q18875", "q18935", "q19365", "q19820", "q
20121", "q26742", "q28537", "q35997", "q38186", "q49038", "q52682", "q57854",
"q64923", "q70218", "q128", "q129","q332", "q36355",

#Education
#"q1119", "q6988", "q18955", "q19566", "q20530", "q40194", "q41393", "q49016"
, "q82755", "q83547", "q80404",

#Moral
#"q16", "q65", "q70", "q71", "q868", "q9418", "q18699", "q23401", "q27915", "
q36671", "q38072", "q57694", "q86761", "q13077", "q20661", "q68746", "q15372"
, "q212814",

#Religion
#"q41", "q42", "q43", "q193", "q210", "q13032", "q13033", "q15889", "q19198",
"q19915", "q21729", "q25107", "q43237","q48372", "q61786", "q72000", "q82618"
, "q85583", "q85932", "q57886",

#Stigma
#"q13", "q14", "q155", "q156", "q395", "q397", "q499", "q501", "q546", "q723"
, "q1117", "q1618", "q2017", "q5417", "q6258", "q7920", "q8197", "q8215", "q1
```

```
2560", "q12949", "q15063", "q15743", "q16317", "q16807", "q19148", "q19690",
"q31488", "q33107", "q33625", "q36235", "q42190", "q81847", "q6971", "q9415",
"q48260", "q1287", "q14884", "q15280", "q17168", "q19815", "q252", "q37936",
"q46301", "q59456", "q81259", "q1138",

#Substance Use
#"q77", "q78", "q79", "q80", "q42665", "q59457", "q9688", "q13006", "q26600",
"q39189", "q46647",

#samData <- tiny[myvars]
samData <- NOTsecondSmall[myvars]
library("car")
attach(samData)

demographicData <- NOTsecondSmall[mydemographics]


library(car)
attach(demographicData)

## The following objects are masked from samData:
##
##     d_age, d_ethnicity, d_gender
```

Examine Components of Demographic Variables

```
table(d_age)

table(d_gender)

table(d_ethnicity)
```

Recode Demdata Demographic Variables

fix this section as you like it

```
demographicData$d_gender<-car::recode(d_gender,"'Man'= '1_Man'; 'Cis Man'= '1
_Man'; 'Woman'= '2_Woman'; 'Cis Woman'= '2_Woman'; 'Trans Woman'= '2_Woman';
'Transfeminine'= '2_Woman'; 'Trans Man'= '1_Man';  NA= 'Other'; else= 'Other'
")

table(demographicData$d_gender)

##
##   1_Man 2_Woman   Other
##   38995   25170    2106

gender<-demographicData['d_gender']
```

```
table(gender)

## gender
##   1_Man 2_Woman   Other
##   38995   25170    2106

demographicData$d_age <- car::recode( d_age," 18:34 ='1_Millennials'; 35:50 =
'2_Generation X'; 51:69 = '3_Baby Boomers'; NA='6_Other'; else= '6_Other'")

table(demographicData$d_age)

##
##  1_Millennials 2_Generation X 3_Baby Boomers        6_Other
##          44633          18254           1392           1992

age<-demographicData['d_age']



demographicData$d_ethnicity<-car::recode(d_ethnicity,"'White'= 'White'; 'Blac
k'= 'Black'; 'Mixed'= 'Mixed'; 'Pacific Islander'= 'Pacific Islander'; 'Hispa
nic / Latin'= 'Hispanic/Latin'; 'Indian'= 'Indian'; 'Asian'= 'Asian'; 'Middle
Eastern'= 'Middle Eastern'; 'Native American'= 'Native American'; NA= 'Other'
;  else= 'Other'" )

table(demographicData$d_ethnicity)

##
##            Asian            Black   Hispanic/Latin           Indian
##             3091             2610             2715              697
##   Middle Eastern  Native American            Other Pacific Islander
##              376              135            19099              124
##            White
##            37424

ethnicity<-demographicData['d_ethnicity']

write.csv(demographicData,'demographicData1.csv',row.names=FALSE)
```

Sam’s car::recodes

```
#Family and Friend
samData$q350<- car::recode(q350, "'Yes'= 1; 'No'= 0")

samData$q401<- car::recode(q401, "'Yes'= 1; 'No'= 0")

samData$q427<- car::recode(q427, "'Yes'= 1; 'No'= 0")

samData$q1495<- car::recode(q1495, "'Family'= 1; 'Friends'= 0.5")
```

```
samData$q9589<- car::recode(q9589, "'A Friend'= 0.5; 'A Parent'= 1; 'A Siblin
g'= 1")

samData$q14649<- car::recode(q14649, "'Yes'= 1; 'No'= 0")

samData$q15261<- car::recode(q15261, "'Yes'= 1; 'No'= 0")

samData$q15742<- car::recode(q15742, "'A lot'= 1.5; 'Some'= 1; 'Very little'=
0.5; 'Not at all'= 0")

samData$q15778<- car::recode(q15778, "'Yes'= 1; 'No'= 0; \"I'm Not Sure\"= 0.
5")

samData$q16288<- car::recode(q16288, "\"Yes, I'd be dating them not their fam
ily.\"= 0; 'No, What the heck are they hiding?'= 1; 'Dunno, depends.'= 0.5")

samData$q16714<- car::recode(q16714, "'Lose your friend'= 0; 'Lose your lover
'= 1; \"I'm Not Sure\"= 0.5")

samData$q19125<- car::recode(q19125, "'Very Much So'= 1.5; 'Often, yes'= 1; '
A little bit'= 0.5; 'Not at all / I hate the mainstream'= 0")

samData$q20424<- car::recode(q20424, "'Yes'= 0; 'No'= 1")

samData$q20755<- car::recode(q20755, "'Yes'= 0; 'No'= 1")

samData$q22991<- car::recode(q22991, "'Yes'= 1; 'No'= 0; \"I'm Not Sure\"= 0.
5")

samData$q23779<- car::recode(q23779, "'Often'= 1; 'Rarely'= 0; 'Sometimes'= 0
.5")

samData$q35281<- car::recode(q35281, "'God'= 1; 'My family or friends'= 1.5;
'Myself'= 0.5; \"I don't worry about such things\"= 0")

samData$q35555<- car::recode(q35555, "'Yes.'= 1; 'No.'= 0; \"I'm Not Sure.\"=
0.5")

samData$q35856<- car::recode(q35856, "'Yes.'= 0; 'Only if there were good rea
sons for no contact.'= 0.5; 'No.'= 1")

samData$q37440<- car::recode(q37440, "'Yes.'= 0; 'No.'= 1; 'I do not have a p
et.'= NA")

samData$q42406<- car::recode(q42406, "'End it - Avoid getting between parent
and child.'= 1; 'End it - Better to find someone without children.'= 0; \"Con
tinue - It's not about the child anyway.\"= 0.5; \"Discuss - Find a way to ea
rn child's trust.\"= 1.5")

samData$q43261<- car::recode(q43261, "'I am my own person.'= 0; 'I consider t
```

```
heir opinion but go my own way.'= 0.5; 'I almost always do what my parents th
ink is best.'= 1; 'I always do what my parents say.'= 1.5")

samData$q48553<- car::recode(q48553, "'Yes.'= 0; 'No.'= 1")

samData$q49660<- car::recode(q49660, "'10 or more.'= 1; '1 to 10.'= 0.5; 'Non
e.'= 0; \"I don't carry a wallet or purse.\"= NA")

samData$q61705<- car::recode(q61705, "'Yes.'= 1; 'No.'= 0; 'Maybe.'= 0.5")

samData$q85440<- car::recode(q85440, " 'My parents.'= 1.5; 'My children.'= 1;
'My significant other.'= 0.5; 'Myself.'= 0")

samData$q85773<- car::recode(q85773, " 'Very open.'= 0; \"Somewhat open, I'd
be very cautious.\"= 0.5; 'There would be no chance for this person.'= 1")

#Appearance
samData$q122<- car::recode(q122, "'A lot'= 1; 'A little'= 0.5; 'Not at all'=
0")

samData$q124<- car::recode(q124, "'Yes'= 1; 'No'= 0")

samData$q125<- car::recode(q125, "'Yes'= 1; 'No'= 0")

samData$q126<- car::recode(q126, "'Great'= 1; 'Average'= 0.5; 'Shitty'= 0")

samData$q127<- car::recode(q127, "'Yes'= 1; 'No'= 0")

samData$q134<- car::recode(q134, "'Yes'= 1; 'No'= 0")

samData$q298<- car::recode(q298, "'A lot more attractive'= 1; 'About the same
'= 0.5")

samData$q382<- car::recode(q382, "'Yes'= 1; 'No'= 0")

samData$q383<- car::recode(q383, "\"Sure, it's fine\"= 0; 'No way'= 1")

samData$q7079<- car::recode(q7079, "'Yes'= 0; 'No'= 1; \"I don't know\"= 0.5"
)

samData$q13669<- car::recode(q13669, "'Yes'= 0; 'No'= 1")

samData$q14663<- car::recode(q14663, "'Yes'= 1; 'No'= 0")

samData$q15872<- car::recode(q15872, "'Agree'= 1; 'Disagree'= 0")

samData$q18875<- car::recode(q18875, "'Very damn important'= 1.5; 'Important-
ish'= 1; 'Less important than you think'= 0.5; \"I just don't care at all\"=
0")
```

```
samData$q18935<- car::recode(q18935, "'Zits all over the face'= 1; 'Zits all
over the back/bum/legs'= 1; 'Major obesity'= 1; 'None of those things matter
to me'= 0")

samData$q19365<- car::recode(q19365, "'Yes'= 0; 'No'= 1")

samData$q19820<- car::recode(q19820, "'Yes, braces are hot!'= 0; \"I'm not su
re.\"= 0.5; 'No way!  How uncool.'= 1")

samData$q20121<- car::recode(q20121, "'Yes'= 0; 'No'= 1; 'Perhaps if it were
small and difficult to notice.'= 0.5")

samData$q26742<- car::recode(q26742, "'Yes'= 0; 'No'= 1; \"I'm Not Sure\"= 0.
5")

samData$q28537<- car::recode(q28537, "'Yes'= 0; 'No'= 1; 'No Preference'= 0.5
")

samData$q35997<- car::recode(q35997, "'They look wrong together.'= 1; 'I actu
ally like it that way.'= 0; \"Interesting, but it really doesn't matter.\"= 0
.5; 'Who cares?'= 0")

samData$q38186<- car::recode(q38186, "'Yes.'= 0; 'No.'= 1")

samData$q49038<- car::recode(q49038, "'Yes.'= 1; 'No.'= 0")

samData$q52682<- car::recode(q52682, "'Yes, even if they were slightly overwe
ight.'= 1.5; 'Yes, but only if they were obese.'= 1; 'No.'= 0; 'No, in fact I
prefer overweight people.'= 0.5")

samData$q57854<- car::recode(q57854, "'Yes.'= 0; 'No.'= 1")

samData$q64923<- car::recode(q64923, "'Yes.'= 0; 'No.'= 1")

samData$q70218<- car::recode(q70218, "'Favorably.'= 0; 'Unfavorably.'= 1; 'It
makes no difference.'= 0.5")

samData$q128<- car::recode(q128, "'I have 1 or more BIG tattoos'= 0; 'I have
1 or more LITTLE tattoos'= 0.5; 'I have no tattoos'= 1")

samData$q129<- car::recode(q129, "'Yes'= 0; 'No'= 1")

samData$q332<- car::recode(q332, "'Yes'= 0; 'No'= 1")

samData$q36355<- car::recode(q36355, "'Yes.'= 0; 'No.'= 1")

#Education
samData$q1119<- car::recode(q1119, "'Yes'= 1; 'No'= 0")

samData$q6988<- car::recode(q6988, "'Yes'= 1; 'No'= 0")
```

```r
samData$q18955<- car::recode(q18955, "'Yes'= 0; 'No'= 1")

samData$q19566<- car::recode(q19566, "'Yes'= 0; 'Yes, if they are doing well'
= 0.5; 'No'= 1")

samData$q20530<- car::recode(q20530, "'Yes'= 1; 'No'= 0")

samData$q40194<- car::recode(q40194, "'Extremely important.'= 1; 'Somewhat im
portant.'= 0.5; 'Not at all important.'= 0")

samData$q41393<- car::recode(q41393, "'Turn-on. I find scatterbrained people
cute.'= 0; 'Turn-off. I find easily-confused types annoying.'= 1; \"I'm neutr
al / it depends on their other traits.\"= 0.5")

samData$q49016<- car::recode(q49016, "'Yes.'= 0; 'No.'= 1")

samData$q82755<- car::recode(q82755, "'Yes.'= 1; 'No.'= 0.5")

samData$q83547<- car::recode(q83547, "'Very important.'= 1.5; 'Somewhat impor
tant.'= 1; 'Not at all important.'= 0.5; 'Actually, this would be a turn-off.
'= 0")

samData$q80404<- car::recode(q80404, "'Disappointed.'= 1; 'Indifferent.'= 0.5
; 'Pleased.'= 0")

#Moral
samData$q16<- car::recode(q16, "'Yes'= 1; 'No'= 0")

samData$q65<- car::recode(q65, "'Yes'= 1; 'No'= 0")

samData$q70<- car::recode(q70, "'Yes'= 0; 'No'= 1")

samData$q71<- car::recode(q71, "'Yes'= 0; 'No'= 1")

samData$q868<- car::recode(q868, "'Yes, always'= 1; 'No, never'= 0; 'Sometime
s'= 0.5")

samData$q9418<- car::recode(q9418, "'Important'= 1; 'Not very important'= 0.5
; 'I fart on this question'= 0")

samData$q18699<- car::recode(q18699, "'Yes'= 0; 'No'= 1")

samData$q23401<- car::recode(q23401, "'Yes'= 0; 'No'= 1; \"I'm Not Sure\"= 0.
5")

samData$q27915<- car::recode(q27915, "'Yes'= 0; 'No'= 1; 'Not sure / maybe'=
0.5")

samData$q36671<- car::recode(q36671, "'Yes.'= 0; 'No.'= 1")
```

```r
samData$q38072<- car::recode(q38072, "'Yes.'= 1; 'No.'= 0; 'Only if they were inaccurate.'= 0.5")

samData$q57694<- car::recode(q57694, "'Yes.'= 0.5; 'No.'= 1")

samData$q86761<- car::recode(q86761, "'Yes.'= 0.5; 'No.'= 1; 'Maybe, depending upon the specifics.'= 0.5")

samData$q13077<- car::recode(q13077, "'Yes'= 0; 'No'= 1")

samData$q20661<- car::recode(q20661, "'Yes'= 0; 'No'= 1; \"I'm Not Sure / Sometimes / Depends\"=0.5")

samData$q68746<- car::recode(q68746, "'It raises my opinion.'= 1; 'It lowers my opinion.'= 0; 'It has no effect on my opinion.'= 0.5")

samData$q15372<- car::recode(q15372, "'Yes'= 0; 'No'= 1")

samData$q212814<- car::recode(q212814, "'Very important'= 1.5; 'Somewhat important'= 1; 'A little important'= 0.5; 'Not at all important'= 0")

#Religion
samData$q41<- car::recode(q41, "'Extremely important'= 1.5; 'Somewhat important'= 1; 'Not very important'= 0.5; 'Not at all important'= 0")

samData$q42<- car::recode(q42, "'Yes'= 1; 'No'= 0")

samData$q43<- car::recode(q43, "'Yes'= 1; 'No'= 0")

samData$q193<- car::recode(q193, "'Yes'= 1; 'No'= 0")

samData$q210<- car::recode(q210, "'Yes'= 1; 'No'= 0")

samData$q13032<- car::recode(q13032, "'Yes'= 1; 'Hesistant, but willing'= 1; 'No'= 0; 'Only if he/she was non-practicing'= 0.5")

samData$q13033<- car::recode(q13033, "'Yes'= 1; 'Hesistant, but willing'= 1; 'No'= 0; 'Only if he/she was non-practicing'= 0.5")

samData$q15889<- car::recode(q15889, "'Science'= 0; 'Faith'= 1; 'Equally in both'= 0.5")

samData$q19198<- car::recode(q19198, "'Yes'= 0; 'No'= 1; \"I'm Not Sure\"= 0.5")

samData$q19915<- car::recode(q19915, "'Yes, I love the Tao Te Ching.'= 1; 'Sure, why not?'= 1; \"No, they'd be too laid back.\"= 0.5; 'What the hell is Taoism?'= 0")
```

```r
samData$q21729<- car::recode(q21729, "'Sure'= 0; 'Yes, it seems a bit immature though.'= 0.5; 'No, it is wrong.'= 1; \"I don't know/care.\"= 0")

samData$q25107<- car::recode(q25107, "'Yes, openly'= 0;'Yes, to myself/with friends'= 0;'No'= 1")

samData$q43237<- car::recode(q43237, "'Yes.'= 0.5; 'No.'= 0; 'Only if the faith is consistent with my beliefs.'= 1")

samData$q48372<- car::recode(q48372, "'Yes'= 1;'No'= 0")

samData$q61786<- car::recode(q61786, "'Yes'= 1; 'No'= 0")

#samData$q72000<- car::recode(q72000, "'I have a faith, I wish my partner to have one also'= 1; \"I have a faith, I don't mind if my partner doesn't\"= 1; \"I don't have a faith, neither should my partner\"= 0; \"I don't have a faith, I don't mind about my partne\"= 0.5")

samData$q82618<- car::recode(q82618, "'Yes, to each their own.'= 1; 'Yes, but only if consistent with my beliefs.'= 1; 'No.'= 0")

samData$q85583<-car::recode(q85583, "'Yes.'= 1;'No.'= 0")

samData$q85932<- car::recode(q85932, "'Yes.'= 1; 'No.'= 0")

samData$q57886<- car::recode(q57886, "'Their religious beliefs.'= 1; 'Their political beliefs.'= 1; 'Neither is important.'= 0; 'Both are equally important.'= 1.5")

#Stigma
samData$q13<- car::recode(q13, "'Yes'= 1; 'No'= 0")

samData$q14<- car::recode(q14, "'Yes'= 1; 'No'= 0")

samData$q155<- car::recode(q155, "'Yes'= 0; 'No'= 1")

samData$q156<- car::recode(q156, "'Yes'= 0; 'No'= 1")

samData$q395<- car::recode(q395, "'Yes'= 0; 'No'= 1")

samData$q397<- car::recode(q397, "'Yes'= 0; 'No'= 1")

samData$q499<- car::recode(q499, "'Yes'= 0; 'No'= 1")

samData$q501<- car::recode(q501, "'Yes'= 1; 'No'= 0")

samData$q546<- car::recode(q546, "'Yes'= 0; 'No'= 1")

samData$q723<- car::recode(q723, "'Yes'= 0; 'No'= 1")
```

```r
samData$q1117<- car::recode(q1117, "'Yes'= 0; 'Depends on the illness'= 0.5;
\"Only if they're symptom free through treatment\"= 1; 'No'= 1.5")

samData$q1618<- car::recode(q1618, "'Yes'= 0; 'No'= 1")

samData$q2017<- car::recode(q2017, "'Yes'= 0; 'No'= 1")

samData$q5417<- car::recode(q5417, "'Yes'= 0; 'No'= 1")

samData$q6258<- car::recode(q6258, "'Yes'= 0; 'No'= 1")

samData$q7920<- car::recode(q7920, "'Yes'= 0; 'No'= 1")

samData$q8197<- car::recode(q8197, "'Yes'= 0; 'No'= 1")

samData$q8215<- car::recode(q8215, "'Yes'= 0; 'No'= 1")

samData$q12560<- car::recode(q12560, "'Yes'= 0; 'No'= 1")

samData$q12949<- car::recode(q12949, "'Yes'= 0; 'No'= 1; \"I'm Not Sure\"= 0.
5")

samData$q15063<- car::recode(q15063, "'Yes'= 0; 'No'= 1; \"I don't know\"= 0.
5")

samData$q15743<- car::recode(q15743, "'Completely uninterested'= 1.5; 'Extrem
ely hesitant'= 1; 'A little hesitant'= 0.5; 'Totally fine'= 0")

samData$q16317<- car::recode(q16317, "'Yes'= 0; 'No'= 1; \"I don't know\"= 0.
5")

samData$q16807<- car::recode(q16807, "'Yes'= 0; 'No'= 1")

samData$q19148<- car::recode(q19148, "'Yes'= 0; 'No'= 1; \"I'm Not Sure\"= 0.
5")

samData$q19690<- car::recode(q19690, "'Yes'= 0; 'No'= 1; \"I'm Not Sure\"= 0.
5")

samData$q31488<- car::recode(q31488, "'Yes, I would be reluctant.'= 1; 'No, n
ot at all.'= 0; 'A little, but I would be honest with them.'= 0.5")

samData$q33107<- car::recode(q33107, "'Yes, I like that type of polygamy.'= 0
; 'I could be convinced by the right people'= 0.5; 'I am commited to total mo
nogamy'= 1.5; 'I have open relationships only'= 1")

samData$q33625<- car::recode(q33625, "'Sure, if they are fun/interesting in p
rivate.'= 0; 'Yes, but ONLY if they get counseling/medication.'= 1; \"No! The
y couldn't be possibly be fun/interesting.\"= 1.5; \"Depends / Don't know / D
on't care.\"= 0.5")
```

```
samData$q36235<- car::recode(q36235, "'Yes.'= 0; 'No.'= 1; 'What is a commune
?'= NA")

samData$q42190<- car::recode(q42190, "'Yes.'= 0; 'No.'= 1")

samData$q81847<- car::recode(q81847, "'Yes.'= 0; 'No.'= 1")

samData$q6971<- car::recode(q6971, "'Yes'= 1; 'No'= 0")

samData$q9415<- car::recode(q9415, "'Yes'= 0; 'No'= 1")

samData$q48260<- car::recode(q48260, "'Yes.'= 0; 'No.'= 1")

samData$q1287<- car::recode(q1287, "'Yes - severely'= 0; 'Yes - low grade'= 0
.5; 'No'= 1; \"I'm not sure\"= 0")

samData$q14884<- car::recode(q14884, "'Yes'= 0; 'No'= 1; \"I'm Not Sure\"= 0.
5")

samData$q15280<- car::recode(q15280, "'Yes'= 0; 'No'= 1")

samData$q17168<- car::recode(q17168, "'Yes'= 0; 'No'= 1")

samData$q19815<- car::recode(q19815, "\"Yes, I wouldn't be able to tolerate t
hat.\"= 1; \"Somewhat, but it's their life.\"= 0.5; \"No, I don't mind.\"= 0"
)

samData$q252<- car::recode(q252, "'Yes'= 0; 'No'= 1")

samData$q37936<- car::recode(q37936, "'Yes.'= 1; 'No.'= 0; \"I'm not sure.\"=
0.5")

samData$q46301<- car::recode(q46301, "'Yes.'= 0; 'No.'= 1")

samData$q59456<- car::recode(q59456, "'Yes.'= 0; 'No.'= 1")

samData$q81259<- car::recode(q81259, "'Yes.'= 0; 'No.'= 1")

samData$q1138<- car::recode(q1138, "'Yes'= 0; 'No'= 1; 'Just to visit / I was
working'= 0.5")

#Substance Use
samData$q77<- car::recode(q77, "'Very often'= 0; 'Sometimes'= 1.5; 'Rarely'=
1; 'Never'= 0.5")

samData$q78<- car::recode(q78, "'Yes'= 0; 'No'= 1; \"I don't know, because I'
ve never been drunk.\"= NA")

samData$q79<- car::recode(q79, "'I smoke regularly.'= 0; 'I smoke occasionall
```

```
y.'= 0.5; 'I smoked in the past, but no longer.'= 1; 'I never do drugs.'= 1.5
; 'Never.'= 1.5")

samData$q80<- car::recode(q80, "'I do drugs regularly.'= 0; 'I do drugs occas
ionally.'= 0.5; \"I've done drugs in the past, but no longer.\"= 1; 'I never
do drugs.'= 1.5")

samData$q42665<- car::recode(q42665, "'Yes.'= 1; 'No.'= 0")

samData$q59457<- car::recode(q59457, "'Yes.'= 0; 'No.'= 1")

samData$q9688<- car::recode(q9688, "'No'= 1 ; 'Yes, but only soft stuff like
marijuana'= 0.5; 'Yes'= 0")

samData$q13006<- car::recode(q13006, "'Yes'= 0; 'Yes, but only an occasional/
social smoker'= 0.5; 'No'= 1")

samData$q26600<- car::recode(q26600, "'Definitely, their effort counts a lot'
= 1; 'Not sure / depends'= 0.5; 'No'= 0")

samData$q39189<- car::recode(q39189, "'Yes.'= 0; 'No.'= 1")

samData$q46647<- car::recode(q46647, "'Yes.'= 0; 'No.'= 1")
```

Convert all 'factors' to numeric

the car::recodes above still leave the variables as factors. To properly change these to numbers (for correlations, etc.), we use the following (from https://stackoverflow.com/questions/23915131/change-all-columns-from-factor-to-numeric-in-r).

```
#glimpse(samData)

w <- which( sapply( samData,
                   class ) == 'factor' )

samData[w] <- lapply( samData[w],
                  function(x)
                    as.numeric(as.character(x)) )

## Warning in FUN(X[[i]], ...): NAs introduced by coercion

## Warning in FUN(X[[i]], ...): NAs introduced by coercion

write.csv(samData,"samItemsRecodedNR.csv", row.names=FALSE)
```

Appendix B. Alpha Scale Testing

Alpha Testing

Sam DiPiero

December 13, 2017

```
knitr::opts_chunk$set(echo = TRUE)
library(car)

## Warning: package 'car' was built under R version 3.4.4

## Loading required package: carData

## Warning: package 'carData' was built under R version 3.4.4

library(psych)

## Warning: package 'psych' was built under R version 3.4.4

##
## Attaching package: 'psych'

## The following object is masked from 'package:car':
##
##     logit

library(knitr)
library(papeR)

## Loading required package: xtable

##
## Attaching package: 'papeR'

## The following object is masked from 'package:utils':
##
##     toLatex

library(tidyverse)

## ── Attaching packages ──────────────────────────── tidyverse 1.2.1 ──

## ✓ ggplot2 2.2.1     ✓ purrr   0.2.4
## ✓ tibble  1.4.2     ✓ dplyr   0.7.4
## ✓ tidyr   0.8.0     ✓ stringr 1.3.0
## ✓ readr   1.1.1     ✓ forcats 0.3.0

## ── Conflicts ──────────────────────────────── tidyverse_conflicts() ──
## ✗ ggplot2::%+%()    masks psych::%+%()
## ✗ ggplot2::alpha()  masks psych::alpha()
```

```
## ✖ dplyr::filter()    masks stats::filter()
## ✖ dplyr::lag()       masks stats::lag()
## ✖ dplyr::recode()    masks car::recode()
## ✖ purrr::some()      masks car::some()
## ✖ dplyr::summarise() masks papeR::summarise()
## ✖ dplyr::summarize() masks papeR::summarize()

samData<-read.csv('samItemsRecodedNR.csv')
```

Try kable package for cool table

```
x<-describeBy(samData)

## Warning in FUN(newX[, i], ...): no non-missing arguments to min; returning
## Inf

## Warning in FUN(newX[, i], ...): no non-missing arguments to min; returning
## Inf

## Warning in FUN(newX[, i], ...): no non-missing arguments to max; returning
## -Inf

## Warning in FUN(newX[, i], ...): no non-missing arguments to max; returning
## -Inf

## Warning in describeBy(samData): no grouping variable requested

str(samData)

## 'data.frame':    66271 obs. of  166 variables:
##  $ q13        : int  NA NA 0 NA NA 0 0 0 NA NA ...
##  $ q14        : int  NA NA 0 NA NA 0 0 0 NA NA ...
##  $ q16        : int  NA NA NA NA NA 0 NA NA NA NA ...
##  $ q41        : num  NA NA 1.5 NA 0 0 0.5 NA 0 NA ...
##  $ q42        : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q43        : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q65        : int  NA NA NA NA NA 0 1 NA NA NA ...
##  $ q70        : int  1 NA 1 NA 1 1 1 1 1 NA ...
##  $ q71        : int  1 NA 1 NA NA 1 1 1 NA NA ...
##  $ q77        : num  1.5 NA 1 1.5 1.5 1.5 0 1.5 1 1.5 ...
##  $ q78        : int  NA NA NA NA NA 0 0 NA NA NA ...
##  $ q79        : num  1.5 NA 1 NA NA 0.5 NA NA 1 NA ...
##  $ q80        : num  1.5 NA NA NA NA 1.5 NA NA 1.5 NA ...
##  $ q122       : num  1 NA 1 NA NA 1 NA NA NA NA ...
##  $ q124       : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q125       : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q126       : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q127       : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q128       : num  1 NA 1 NA NA 1 1 NA NA NA ...
##  $ q129       : int  NA NA 1 NA NA NA NA NA NA NA ...
##  $ q134       : int  NA NA NA NA NA NA 0 0 NA NA ...
```

```
##  $ q155        : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q156        : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q193        : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q210        : int  NA NA 1 NA 0 0 NA NA NA NA ...
##  $ q252        : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q298        : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q332        : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q350        : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q382        : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q383        : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q395        : int  NA NA 1 NA NA NA 0 NA NA NA ...
##  $ q397        : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q401        : int  1 NA NA NA NA NA NA NA NA NA ...
##  $ q427        : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q499        : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q501        : int  0 NA 1 NA 0 0 1 0 NA 0 ...
##  $ q546        : int  NA NA 1 NA NA NA 0 NA NA NA ...
##  $ q723        : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q868        : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q1117       : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q1119       : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q1138       : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q1287       : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q1495       : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q1618       : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q2017       : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q5417       : int  1 NA 0 NA NA NA NA NA NA NA ...
##  $ q6258       : int  NA NA 1 NA NA 0 0 NA NA NA ...
##  $ q6971       : int  NA NA NA NA NA 0 NA NA NA NA ...
##  $ q6988       : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q7079       : num  NA NA 1 NA NA NA NA NA NA NA ...
##  $ q7920       : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q8197       : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q8215       : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q9415       : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q9418       : num  NA NA 1 NA NA NA NA NA NA NA ...
##  $ q9589       : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q9688       : num  1 NA 0 NA 0 0 NA NA 1 NA ...
##  $ q12560      : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q12949      : num  NA NA 1 NA NA NA NA NA NA NA ...
##  $ q13006      : num  NA NA 0 NA NA 0 0 NA NA NA ...
##  $ q13032      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q13033      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q13077      : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q13669      : int  NA NA 1 NA NA NA NA NA NA NA ...
##  $ q14649      : int  NA NA NA NA NA 0 NA NA NA NA ...
##  $ q14663      : int  NA NA NA NA NA 0 NA NA NA NA ...
##  $ q14884      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q15063      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q15261      : int  NA NA NA NA NA NA NA NA NA NA ...
```

```
##  $ q15280      : int  1 NA NA 1 NA 0 NA NA NA NA ...
##  $ q15372      : int  NA NA 1 NA NA 1 NA NA NA NA ...
##  $ q15742      : num  NA NA 0.5 NA NA NA 0.5 NA NA NA ...
##  $ q15743      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q15778      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q15872      : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q15889      : num  0 NA 0.5 NA NA 0 0 NA NA NA ...
##  $ q16288      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q16317      : num  1 NA NA NA NA 0.5 NA NA NA NA ...
##  $ q16714      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q16807      : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q17168      : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q18699      : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q18875      : num  1.5 NA 1 NA NA 1 1 NA NA NA ...
##  $ q18935      : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q18955      : int  1 NA NA NA NA NA NA NA NA NA ...
##  $ q19125      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q19148      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q19198      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q19365      : int  NA NA 0 NA NA NA NA NA NA NA ...
##  $ q19566      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q19690      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q19815      : num  1 NA 0 NA NA NA NA NA NA NA ...
##  $ q19820      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q19915      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q20121      : num  NA NA NA NA NA NA NA NA NA NA ...
##  $ q20424      : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ q20530      : int  NA NA 1 NA NA NA 1 NA NA NA ...
##   [list output truncated]
```

```
b <-as.data.frame(samData)
summarize(b, type = "numeric")
```

```
##      type
## 1 numeric
```

```
kable(summarize(b,  quantiles = FALSE,
                type = "numeric"))
```

quantiles	type
FALSE	numeric

Family and Friend Scale (myKeys1)

SSalpha=.66 NSS alpha= .63

```
data(samData)
```

```
## Warning in data(samData): data set 'samData' not found
```

```r
myKeys1<-list(familyandfriend=c("q350", "q401", "q427", "q1495", "q9589", "q14649", "q15261", "q15742", "q15778", "q16288", "q16714", "q19125", "q20424", "q20755", "q22991", "q23779", "q35281", "q35555", "q35856", "q37440", "q42406", "q43261", "q48553", "q49660", "q61705", "q85440", "q85773"))

ffScores<-scoreItems(myKeys1,samData, impute = "none", missing = TRUE )

ffScores["alpha"]

## $alpha
##       familyandfriend
## alpha       0.6297046

ffScores["item.cor"]

## $item.cor
##        familyandfriend
## q350        0.43650156
## q401        0.54301905
## q427        0.28048622
## q1495       0.36987803
## q9589      -0.13411313
## q14649      0.44251710
## q15261      0.19615362
## q15742      0.36477675
## q15778      0.33458621
## q16288      0.39608282
## q16714      0.15329725
## q19125      0.31171561
## q20424      0.31004371
## q20755      0.29794820
## q22991      0.57873582
## q23779      0.20499948
## q35281      0.41984526
## q35555     -0.02819357
## q35856      0.47645351
## q37440      0.23336630
## q42406      0.24125630
## q43261      0.43537740
## q48553      0.37581150
## q49660      0.21046308
## q61705      0.23555857
## q85440      0.29475530
## q85773      0.18564997

familyandfriend1<-ffScores[["scores"]]
str(familyandfriend1)

##  num [1:66271, 1] 0.75 NaN 0.5 NaN NaN 0 0.25 NaN NaN NaN ...
##  - attr(*, "dimnames")=List of 2
```

```
##   ..$ : NULL
##   ..$ : chr "familyandfriend"
```

```
sd(familyandfriend1, na.rm =TRUE)
```

```
## [1] 0.243516
```

```
familyandfriend<-as.tibble(familyandfriend1)
```

Appearance Scale (myKeys2)

alpha=.78 NSSalpha= .78

```
data(samData)
```

```
## Warning in data(samData): data set 'samData' not found
```

```
myKeys2<- list(appearance=c("q122", "q124", "q125", "q126", "q127", "q134", "
q298", "q382", "q383", "q7079", "q13669", "q14663", "q15872", "q18875", "q189
35", "q19365", "q19820", "q20121", "q26742", "q28537", "q35997", "q38186", "q
49038", "q52682", "q57854", "q64923", "q70218", "q128", "q129","q332", "q3635
5"))

appearanceScores<-scoreItems(myKeys2,samData, impute = "none", missing = TRUE
)

appearanceScores["alpha"]
```

```
## $alpha
##       appearance
## alpha  0.7845039
```

```
appearanceScores["item.cor"]
```

```
## $item.cor
##        appearance
## q122    0.4696350
## q124    0.4107087
## q125    0.3513569
## q126    0.4144456
## q127    0.3269512
## q134    0.4248457
## q298    0.1176584
## q382    0.2865421
## q383    0.3853954
## q7079   0.5171858
## q13669  0.2453145
## q14663  0.3783243
## q15872  0.3761762
## q18875  0.4023827
## q18935  0.4028369
## q19365  0.3691837
```

```
## q19820  0.3351552
## q20121  0.4219406
## q26742  0.5293778
## q28537  0.1633515
## q35997  0.2624082
## q38186  0.2728084
## q49038  0.4248064
## q52682  0.5126284
## q57854  0.5963557
## q64923  0.5459937
## q70218  0.2407504
## q128    0.1565513
## q129    0.1658505
## q332    0.1965070
## q36355  0.4963285
```

```
appearance1<-appearanceScores[["scores"]]
str(appearance1)
```

```
##  num [1:66271, 1] 1.125 NaN 0.722 NaN NaN ...
##  - attr(*, "dimnames")=List of 2
##   ..$ : NULL
##   ..$ : chr "appearance"
```

```
sd(appearance1, na.rm=TRUE)
```

```
## [1] 0.2781008
```

```
appearance<- as.tibble(appearance1)
```

Education Scale (myKeys3)

alpha=.69 NSS= .71

```
data(samData)
```

```
## Warning in data(samData): data set 'samData' not found
```

```
myKeys3<- list(education=c("q1119", "q6988", "q18955", "q19566", "q20530", "q
40194", "q41393", "q49016", "q82755", "q83547", "q80404"))
```

```
educationScores<-scoreItems(myKeys3,samData, impute = "none", missing = TRUE
)
```

```
educationScores["alpha"]
```

```
## $alpha
##       education
## alpha 0.7056414
```

```
educationScores["item.cor"]
```

```
## $item.cor
##        education
## q1119  0.5490009
## q6988  0.5990369
## q18955 0.4513283
## q19566 0.5710736
## q20530 0.5435207
## q40194 0.4497049
## q41393 0.4753614
## q49016 0.5844454
## q82755 0.3761275
## q83547 0.5199009
## q80404 0.4662705
```

```
education1<-educationScores[["scores"]]
str(education1)
```

```
##  num [1:66271, 1] 1 NaN 1 NaN NaN 0.5 0.75 0.5 NaN NaN ...
##  - attr(*, "dimnames")=List of 2
##   ..$ : NULL
##   ..$ : chr "education"
```

```
sd(education1, na.rm=TRUE)
```

```
## [1] 0.2629919
```

```
education<- as.tibble(education1)
```

Moral Scale (myKeys4)

alpha=.62 NSS= .64

```
data(samData)
```

```
## Warning in data(samData): data set 'samData' not found
```

```
myKeys4<- list(moral=c("q16", "q65", "q70", "q71", "q868", "q9418", "q18699",
"q23401", "q27915", "q36671", "q38072", "q57694", "q86761", "q13077", "q20661
", "q68746", "q15372", "q212814"))

moralScores<-scoreItems(myKeys4,samData, impute = "none", missing = TRUE )

moralScores["alpha"]
```

```
## $alpha
##          moral
## alpha 0.6407302
```

```
moralScores["item.cor"]
```

```
## $item.cor
##             moral
```

```
## q16     0.3547909
## q65     0.5218608
## q70     0.1201825
## q71     0.1367997
## q868    0.2375461
## q9418   0.2001492
## q18699  0.4227171
## q23401  0.2918772
## q27915  0.4289356
## q36671  0.4816988
## q38072  0.4952432
## q57694  0.3848238
## q86761  0.3699284
## q13077  0.4023725
## q20661  0.5678510
## q68746  0.3995444
## q15372  0.4452225
## q212814 0.3589955
```

```
moral1<-moralScores[["scores"]]
str(moral1)
```

```
##  num [1:66271, 1] 1 NaN 1.1 NaN 1 ...
##  - attr(*, "dimnames")=List of 2
##   ..$ : NULL
##   ..$ : chr "moral"
```

```
sd(moral1, na.rm= TRUE)
```

```
## [1] 0.2483765
```

```
moral<- as.tibble(moral1)
```

Religion Scale (myKeys5)

alpha=.82 NSS alpha= .8

```
data(samData)
```

```
## Warning in data(samData): data set 'samData' not found
```

```
myKeys5<- list(religion=c("q41", "q42", "q43", "q193", "q210", "q13032", "q13
033", "q15889", "q19198", "q19915", "q21729", "q25107", "q43237","q48372", "q
61786", "q82618", "q85583", "q85932", "q57886"))
```

```
religionScores<-scoreItems(myKeys5,samData, impute = "none", missing = TRUE )
```

```
religionScores["alpha"]
```

```
## $alpha
##        religion
## alpha 0.7950727
```

```
religionScores["item.cor"]

## $item.cor
##           religion
## q41     0.83117705
## q42     0.60053032
## q43     0.41808257
## q193    0.58888746
## q210    0.77055446
## q13032  0.39324102
## q13033 -0.23456966
## q15889  0.68779810
## q19198  0.57202412
## q19915  0.01154776
## q21729  0.47322114
## q25107  0.53835001
## q43237  0.63237991
## q48372  0.77652438
## q61786  0.65293455
## q82618  0.43179867
## q85583  0.32933052
## q85932  0.27102259
## q57886  0.08071302

religion1<-religionScores[["scores"]]
str(religion1)

##  num [1:66271, 1] 0 NaN 1 NaN 0 0 0.625 NaN 0 NaN ...
##  - attr(*, "dimnames")=List of 2
##   ..$ : NULL
##   ..$ : chr "religion"

sd(religion1, na.rm=TRUE)

## [1] 0.3772593

religion<-as.tibble(religion1)
```

Stigma Scale (myKeys6)

alpha=.88 NSS alpha= .89

```
data(samData)

## Warning in data(samData): data set 'samData' not found

myKeys6<- list(stigma=c("q13", "q14", "q155", "q156", "q395", "q397", "q499",
"q501", "q546", "q723", "q1117", "q1618", "q2017", "q5417", "q6258", "q7920",
"q8197", "q8215", "q12560", "q12949", "q15063", "q15743", "q16317", "q16807",
"q19148", "q19690", "q31488", "q33107", "q33625", "q36235", "q42190", "q81847
", "q6971", "q9415", "q48260", "q1287", "q14884", "q15280", "q17168", "q19815
", "q252", "q37936", "q46301", "q59456", "q81259", "q1138"))
```

```r
stigmaScores<-scoreItems(myKeys6,samData, impute = "none", missing = TRUE )

stigmaScores["alpha"]
```

```
## $alpha
##          stigma
## alpha 0.887174
```

```r
stigmaScores["item.cor"]
```

```
## $item.cor
##              stigma
## q13     0.3555609
## q14     0.3465303
## q155    0.3051308
## q156    0.2181580
## q395    0.3440554
## q397    0.3303029
## q499    0.4614012
## q501   -0.1047862
## q546    0.5059308
## q723    0.6027177
## q1117   0.5608308
## q1618   0.4050575
## q2017   0.4201837
## q5417   0.3380484
## q6258   0.5555362
## q7920   0.6212715
## q8197   0.4977229
## q8215   0.4694336
## q12560  0.4942057
## q12949  0.5554086
## q15063  0.4284830
## q15743  0.4945757
## q16317  0.4636543
## q16807  0.5951879
## q19148  0.7037379
## q19690  0.5309547
## q31488  0.1374365
## q33107  0.5182309
## q33625  0.4680843
## q36235  0.5417444
## q42190  0.2746511
## q81847  0.4064794
## q6971   0.2064201
## q9415   0.4784658
## q48260  0.5202764
## q1287   0.3737331
## q14884  0.3265539
```

```
## q15280  0.4253876
## q17168  0.4686060
## q19815  0.4546004
## q252    0.1986663
## q37936  0.1354812
## q46301  0.3930682
## q59456  0.2419769
## q81259  0.2928964
## q1138   0.1219777
```

```
stigma1<-stigmaScores[["scores"]]
str(stigma1)
```

```
##  num [1:66271, 1] 0.75 NaN 0.556 1 0 ...
##  - attr(*, "dimnames")=List of 2
##   ..$ : NULL
##   ..$ : chr "stigma"
```

```
sd(stigma1, na.rm=TRUE)
```

```
## [1] 0.3091575
```

```
stigma<- as.tibble(stigma1)
```

Substance Use Scale (myKeys7)

alpha=.64 NSSalpha= .63

```
data(samData)
```

```
## Warning in data(samData): data set 'samData' not found
```

```
myKeys7<- list(substanceUse=c("q77", "q78", "q79", "q80", "q42665", "q59457",
"q9688", "q13006", "q26600", "q39189", "q46647"))

suScores<-scoreItems(myKeys7,samData, impute = "none", missing = TRUE )

suScores["alpha"]
```

```
## $alpha
##       substanceUse
## alpha    0.6342986
```

```
suScores["item.cor"]
```

```
## $item.cor
##        substanceUse
## q77       0.3335579
## q78       0.5488605
## q79       0.6936787
## q80       0.6348803
## q42665    0.2762141
```

```
## q59457    0.5868158
## q9688     0.7225529
## q13006    0.5447619
## q26600   -0.0176915
## q39189    0.2305750
## q46647    0.5164296
```

```r
substanceUse1<-suScores[["scores"]]
str(substanceUse1)
```

```
##  num [1:66271, 1] 1.375 NaN 0.333 1.5 0.75 ...
##  - attr(*, "dimnames")=List of 2
##   ..$ : NULL
##   ..$ : chr "substanceUse"
```

```r
sd(substanceUse1, na.rm=TRUE)
```

```
## [1] 0.3057025
```

```r
substanceUse<- as.tibble(substanceUse1)

allSSscores<-cbind(familyandfriend,appearance,education,moral,religion,stigma
,substanceUse)

write.csv(allSSscores,"samScalescores1.csv", row.names=FALSE)
```

Appendix C. Grouped Averages

Grouped Averages

Sam DiPiero

December 13, 2017

```r
knitr::opts_chunk$set(echo = TRUE)
library(psych)
```

```
## Warning: package 'psych' was built under R version 3.4.4
```

```r
library(knitr)
library(papeR)
```

```
## Loading required package: car
```

```
## Warning: package 'car' was built under R version 3.4.4

## Loading required package: carData

## Warning: package 'carData' was built under R version 3.4.4

##
## Attaching package: 'car'

## The following object is masked from 'package:psych':
##
##     logit

## Loading required package: xtable

##
## Attaching package: 'papeR'

## The following object is masked from 'package:utils':
##
##     toLatex

suppressMessages(library(tidyverse))
library(car)
library(dplyr)
library(sjPlot)

## Warning in checkMatrixPackageVersion(): Package version inconsistency dete
cted.
## TMB was built with Matrix version 1.2.12
## Current Matrix version is 1.2.14
## Please re-install 'TMB' from source using install.packages('TMB', type = '
source') or ask CRAN for a binary version of 'TMB' matching CRAN's 'Matrix' p
ackage

library(GPArotation)

allSSscores<-read.csv('samScalescores1.csv')
demo <- read.csv('demographicData1.csv')

all <-cbind(allSSscores,demo)

library(dplyr)

#GENDER
all %>%
  group_by(d_gender) %>%
  summarise_at(vars(familyandfriend), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 3 x 3
##   d_gender  mean     n
##   <fct>    <dbl> <int>
## 1 1_Man    0.448 38995
```

```
## 2 2_Woman  0.471 25170
## 3 Other    0.498  2106
```

```r
all %>%
  group_by(d_gender, d_age) %>%
  summarise_at(vars(familyandfriend), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 12 x 4
## # Groups:   d_gender [?]
##    d_gender d_age            mean     n
##    <fct>    <fct>           <dbl> <int>
##  1 1_Man    1_Millennials   0.450 24069
##  2 1_Man    2_Generation X  0.446 13673
##  3 1_Man    3_Baby Boomers  0.434  1227
##  4 1_Man    6_Other         0.443    26
##  5 2_Woman  1_Millennials   0.474 20442
##  6 2_Woman  2_Generation X  0.458  4558
##  7 2_Woman  3_Baby Boomers  0.439   163
##  8 2_Woman  6_Other         0.479     7
##  9 Other    1_Millennials   0.420   122
## 10 Other    2_Generation X  0.377    23
## 11 Other    3_Baby Boomers  0.250     2
## 12 Other    6_Other         0.506  1959
```

```r
all %>%
  group_by(d_gender, d_ethnicity) %>%
  summarise_at(vars(familyandfriend), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 24 x 4
## # Groups:   d_gender [?]
##    d_gender d_ethnicity       mean     n
##    <fct>    <fct>            <dbl> <int>
##  1 1_Man    Asian            0.484  1383
##  2 1_Man    Black            0.435  1636
##  3 1_Man    Hispanic/Latin   0.453  1765
##  4 1_Man    Indian           0.481   549
##  5 1_Man    Middle Eastern   0.429   266
##  6 1_Man    Native American  0.428    85
##  7 1_Man    Other            0.423 10233
##  8 1_Man    Pacific Islander 0.528    78
##  9 1_Man    White            0.453 23000
## 10 2_Woman  Asian            0.505  1707
## # ... with 14 more rows
```

```r
#AGE
all %>%
  group_by(d_age) %>%
  summarise_at(vars(familyandfriend), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 4 x 3
##   d_age           mean     n
```

```
##   <fct>          <dbl> <int>
## 1 1_Millennials  0.461 44633
## 2 2_Generation X 0.449 18254
## 3 3_Baby Boomers 0.434  1392
## 4 6_Other        0.505  1992
```

```
all %>%
  group_by(d_age, d_ethnicity) %>%
  summarise_at(vars(familyandfriend), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 33 x 4
## # Groups:   d_age [?]
##    d_age          d_ethnicity       mean     n
##    <fct>          <fct>            <dbl> <int>
##  1 1_Millennials  Asian            0.502  2434
##  2 1_Millennials  Black            0.436  1933
##  3 1_Millennials  Hispanic/Latin   0.452  1971
##  4 1_Millennials  Indian           0.486   581
##  5 1_Millennials  Middle Eastern   0.414   294
##  6 1_Millennials  Native American  0.435    86
##  7 1_Millennials  Other            0.434 12842
##  8 1_Millennials  Pacific Islander 0.489    83
##  9 1_Millennials  White            0.469 24409
## 10 2_Generation X Asian            0.474   644
## # ... with 23 more rows
```

```
#ETHNICITY
all %>%
  group_by(d_ethnicity) %>%
  summarise_at(vars(familyandfriend), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 9 x 3
##   d_ethnicity       mean     n
##   <fct>            <dbl> <int>
## 1 Asian            0.495  3091
## 2 Black            0.436  2610
## 3 Hispanic/Latin   0.453  2715
## 4 Indian           0.489   697
## 5 Middle Eastern   0.432   376
## 6 Native American  0.432   135
## 7 Other            0.441 19099
## 8 Pacific Islander 0.498   124
## 9 White            0.463 37424
```

```
#ALL THREE
all %>%
  group_by(d_gender, d_ethnicity, d_age) %>%
  summarise_at(vars(familyandfriend), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 71 x 5
## # Groups:   d_gender, d_ethnicity [?]
```

```
##    d_gender d_ethnicity    d_age           mean     n
##    <fct>    <fct>          <fct>          <dbl> <int>
##  1 1_Man    Asian          1_Millennials  0.484   964
##  2 1_Man    Asian          2_Generation X 0.483   408
##  3 1_Man    Asian          3_Baby Boomers 0.537    11
##  4 1_Man    Black          1_Millennials  0.428  1137
##  5 1_Man    Black          2_Generation X 0.448   464
##  6 1_Man    Black          3_Baby Boomers 0.440    35
##  7 1_Man    Hispanic/Latin 1_Millennials  0.455  1203
##  8 1_Man    Hispanic/Latin 2_Generation X 0.453   535
##  9 1_Man    Hispanic/Latin 3_Baby Boomers 0.400    27
## 10 1_Man    Indian         1_Millennials  0.476   451
## # ... with 61 more rows

#GENDER
all %>%
  group_by(d_gender) %>%
  summarise_at(vars(appearance), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 3 x 3
##   d_gender  mean     n
##   <fct>    <dbl> <int>
## 1 1_Man    0.560 38995
## 2 2_Woman  0.591 25170
## 3 Other    0.498  2106

all %>%
  group_by(d_gender, d_age) %>%
  summarise_at(vars(appearance), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 12 x 4
## # Groups:   d_gender [?]
##    d_gender d_age           mean     n
##    <fct>    <fct>          <dbl> <int>
##  1 1_Man    1_Millennials  0.569 24069
##  2 1_Man    2_Generation X 0.547 13673
##  3 1_Man    3_Baby Boomers 0.564  1227
##  4 1_Man    6_Other        0.519    26
##  5 2_Woman  1_Millennials  0.596 20442
##  6 2_Woman  2_Generation X 0.570  4558
##  7 2_Woman  3_Baby Boomers 0.620   163
##  8 2_Woman  6_Other        0.439     7
##  9 Other    1_Millennials  0.328   122
## 10 Other    2_Generation X 0.381    23
## 11 Other    3_Baby Boomers 0.447     2
## 12 Other    6_Other        0.511  1959

all %>%
  group_by(d_gender, d_ethnicity) %>%
  summarise_at(vars(appearance), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 24 x 4
## # Groups:   d_gender [?]
##    d_gender d_ethnicity       mean     n
##    <fct>    <fct>            <dbl> <int>
##  1 1_Man    Asian            0.617  1383
##  2 1_Man    Black            0.622  1636
##  3 1_Man    Hispanic/Latin   0.603  1765
##  4 1_Man    Indian           0.630   549
##  5 1_Man    Middle Eastern   0.667   266
##  6 1_Man    Native American  0.552    85
##  7 1_Man    Other            0.586 10233
##  8 1_Man    Pacific Islander 0.557    78
##  9 1_Man    White            0.542 23000
## 10 2_Woman  Asian            0.703  1707
## # ... with 14 more rows

#AGE
all %>%
  group_by(d_age) %>%
  summarise_at(vars(appearance), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 4 x 3
##   d_age            mean     n
##   <fct>           <dbl> <int>
## 1 1_Millennials   0.580 44633
## 2 2_Generation X 0.552 18254
## 3 3_Baby Boomers 0.570  1392
## 4 6_Other         0.511  1992

all %>%
  group_by(d_age, d_ethnicity) %>%
  summarise_at(vars(appearance), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 33 x 4
## # Groups:   d_age [?]
##    d_age          d_ethnicity       mean     n
##    <fct>          <fct>            <dbl> <int>
##  1 1_Millennials  Asian            0.671  2434
##  2 1_Millennials  Black            0.611  1933
##  3 1_Millennials  Hispanic/Latin   0.630  1971
##  4 1_Millennials  Indian           0.663   581
##  5 1_Millennials  Middle Eastern   0.721   294
##  6 1_Millennials  Native American  0.578    86
##  7 1_Millennials  Other            0.599 12842
##  8 1_Millennials  Pacific Islander 0.628    83
##  9 1_Millennials  White            0.559 24409
## 10 2_Generation X Asian            0.636   644
## # ... with 23 more rows

#ETHNICITY
all %>%
```

```r
  group_by(d_ethnicity) %>%
  summarise_at(vars(appearance), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 9 x 3
##   d_ethnicity       mean     n
##   <fct>            <dbl> <int>
## 1 Asian            0.662  3091
## 2 Black            0.612  2610
## 3 Hispanic/Latin   0.616  2715
## 4 Indian           0.640   697
## 5 Middle Eastern   0.690   376
## 6 Native American  0.537   135
## 7 Other            0.577 19099
## 8 Pacific Islander 0.609   124
## 9 White            0.553 37424

#ALL THREE
all %>%
  group_by(d_gender, d_ethnicity, d_age) %>%
  summarise_at(vars(appearance), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 71 x 5
## # Groups:   d_gender, d_ethnicity [?]
##    d_gender d_ethnicity    d_age           mean     n
##    <fct>    <fct>          <fct>          <dbl> <int>
##  1 1_Man    Asian          1_Millennials  0.627   964
##  2 1_Man    Asian          2_Generation X 0.601   408
##  3 1_Man    Asian          3_Baby Boomers 0.576    11
##  4 1_Man    Black          1_Millennials  0.631  1137
##  5 1_Man    Black          2_Generation X 0.607   464
##  6 1_Man    Black          3_Baby Boomers 0.612    35
##  7 1_Man    Hispanic/Latin 1_Millennials  0.619  1203
##  8 1_Man    Hispanic/Latin 2_Generation X 0.575   535
##  9 1_Man    Hispanic/Latin 3_Baby Boomers 0.543    27
## 10 1_Man    Indian         1_Millennials  0.651   451
## # ... with 61 more rows

#GENDER
all %>%
  group_by(d_gender) %>%
  summarise_at(vars(education), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 3 x 3
##   d_gender  mean     n
##   <fct>    <dbl> <int>
## 1 1_Man    0.601 38995
## 2 2_Woman  0.721 25170
## 3 Other    0.616  2106
```

```r
all %>%
  group_by(d_gender, d_age) %>%
  summarise_at(vars(education), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 12 x 4
## # Groups:   d_gender [?]
##    d_gender d_age           mean     n
##    <fct>    <fct>          <dbl> <int>
##  1 1_Man    1_Millennials  0.600 24069
##  2 1_Man    2_Generation X 0.606 13673
##  3 1_Man    3_Baby Boomers 0.574  1227
##  4 1_Man    6_Other        0.664    26
##  5 2_Woman  1_Millennials  0.724 20442
##  6 2_Woman  2_Generation X 0.711  4558
##  7 2_Woman  3_Baby Boomers 0.684   163
##  8 2_Woman  6_Other        0.744     7
##  9 Other    1_Millennials  0.671   122
## 10 Other    2_Generation X 0.776    23
## 11 Other    3_Baby Boomers 0.438     2
## 12 Other    6_Other        0.610  1959
```

```r
all %>%
  group_by(d_gender, d_ethnicity) %>%
  summarise_at(vars(education), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 24 x 4
## # Groups:   d_gender [?]
##    d_gender d_ethnicity       mean     n
##    <fct>    <fct>            <dbl> <int>
##  1 1_Man    Asian            0.576  1383
##  2 1_Man    Black            0.573  1636
##  3 1_Man    Hispanic/Latin   0.544  1765
##  4 1_Man    Indian           0.615   549
##  5 1_Man    Middle Eastern   0.559   266
##  6 1_Man    Native American  0.506    85
##  7 1_Man    Other            0.604 10233
##  8 1_Man    Pacific Islander 0.507    78
##  9 1_Man    White            0.607 23000
## 10 2_Woman  Asian            0.678  1707
## # ... with 14 more rows
```

```r
#AGE
all %>%
  group_by(d_age) %>%
  summarise_at(vars(education), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 4 x 3
##   d_age           mean     n
##   <fct>          <dbl> <int>
## 1 1_Millennials  0.656 44633
## 2 2_Generation X 0.630 18254
```

```
## 3 3_Baby Boomers 0.584  1392
## 4 6_Other         0.611  1992

all %>%
  group_by(d_age, d_ethnicity) %>%
  summarise_at(vars(education), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 33 x 4
## # Groups:   d_age [?]
##    d_age          d_ethnicity       mean     n
##    <fct>          <fct>            <dbl> <int>
##  1 1_Millennials  Asian            0.629  2434
##  2 1_Millennials  Black            0.637  1933
##  3 1_Millennials  Hispanic/Latin   0.603  1971
##  4 1_Millennials  Indian           0.643   581
##  5 1_Millennials  Middle Eastern   0.622   294
##  6 1_Millennials  Native American  0.613    86
##  7 1_Millennials  Other            0.670 12842
##  8 1_Millennials  Pacific Islander 0.616    83
##  9 1_Millennials  White            0.659 24409
## 10 2_Generation X Asian            0.626   644
## # ... with 23 more rows

#ETHNICITY
all %>%
  group_by(d_ethnicity) %>%
  summarise_at(vars(education), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 9 x 3
##   d_ethnicity       mean     n
##   <fct>            <dbl> <int>
## 1 Asian            0.628  3091
## 2 Black            0.626  2610
## 3 Hispanic/Latin   0.589  2715
## 4 Indian           0.645   697
## 5 Middle Eastern   0.607   376
## 6 Native American  0.571   135
## 7 Other            0.650 19099
## 8 Pacific Islander 0.576   124
## 9 White            0.648 37424

#ALL THREE
all %>%
  group_by(d_gender, d_ethnicity, d_age) %>%
  summarise_at(vars(education), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 71 x 5
## # Groups:   d_gender, d_ethnicity [?]
##    d_gender d_ethnicity    d_age           mean     n
##    <fct>    <fct>          <fct>          <dbl> <int>
##  1 1_Man    Asian          1_Millennials  0.564   964
```

```
## 2 1_Man Asian 2_Generation X 0.596 408
## 3 1_Man Asian 3_Baby Boomers 0.570 11
## 4 1_Man Black 1_Millennials 0.566 1137
## 5 1_Man Black 2_Generation X 0.585 464
## 6 1_Man Black 3_Baby Boomers 0.556 35
## 7 1_Man Hispanic/Latin 1_Millennials 0.553 1203
## 8 1_Man Hispanic/Latin 2_Generation X 0.530 535
## 9 1_Man Hispanic/Latin 3_Baby Boomers 0.551 27
## 10 1_Man Indian 1_Millennials 0.608 451
## # ... with 61 more rows
```

```
#GENDER
all %>%
  group_by(d_gender) %>%
  summarise_at(vars(moral), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 3 x 3
##   d_gender  mean     n
##   <fct>    <dbl> <int>
## 1 1_Man    0.735 38995
## 2 2_Woman  0.845 25170
## 3 Other    0.712  2106
```

```
all %>%
  group_by(d_gender, d_age) %>%
  summarise_at(vars(moral), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 12 x 4
## # Groups:   d_gender [?]
##    d_gender d_age             mean     n
##    <fct>    <fct>            <dbl> <int>
##  1 1_Man    1_Millennials    0.721 24069
##  2 1_Man    2_Generation X   0.754 13673
##  3 1_Man    3_Baby Boomers   0.762  1227
##  4 1_Man    6_Other          0.803    26
##  5 2_Woman  1_Millennials    0.846 20442
##  6 2_Woman  2_Generation X   0.842  4558
##  7 2_Woman  3_Baby Boomers   0.822   163
##  8 2_Woman  6_Other          0.823     7
##  9 Other    1_Millennials    0.859   122
## 10 Other    2_Generation X   0.877    23
## 11 Other    3_Baby Boomers   0.833     2
## 12 Other    6_Other          0.699  1959
```

```
all %>%
  group_by(d_gender, d_ethnicity) %>%
  summarise_at(vars(moral), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 24 x 4
## # Groups:   d_gender [?]
##    d_gender d_ethnicity       mean     n
```

```
##    <fct>    <fct>            <dbl> <int>
##  1 1_Man    Asian            0.699  1383
##  2 1_Man    Black            0.674  1636
##  3 1_Man    Hispanic/Latin   0.693  1765
##  4 1_Man    Indian           0.737   549
##  5 1_Man    Middle Eastern   0.663   266
##  6 1_Man    Native American  0.616    85
##  7 1_Man    Other            0.725 10233
##  8 1_Man    Pacific Islander 0.695    78
##  9 1_Man    White            0.747 23000
## 10 2_Woman  Asian            0.769  1707
## # ... with 14 more rows

#AGE
all %>%
  group_by(d_age) %>%
  summarise_at(vars(moral), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 4 x 3
##   d_age           mean     n
##   <fct>          <dbl> <int>
## 1 1_Millennials  0.779 44633
## 2 2_Generation X 0.775 18254
## 3 3_Baby Boomers 0.768  1392
## 4 6_Other        0.701  1992

all %>%
  group_by(d_age, d_ethnicity) %>%
  summarise_at(vars(moral), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 33 x 4
## # Groups:   d_age [?]
##    d_age          d_ethnicity       mean     n
##    <fct>          <fct>            <dbl> <int>
##  1 1_Millennials  Asian            0.730  2434
##  2 1_Millennials  Black            0.717  1933
##  3 1_Millennials  Hispanic/Latin   0.725  1971
##  4 1_Millennials  Indian           0.763   581
##  5 1_Millennials  Middle Eastern   0.718   294
##  6 1_Millennials  Native American  0.661    86
##  7 1_Millennials  Other            0.787 12842
##  8 1_Millennials  Pacific Islander 0.681    83
##  9 1_Millennials  White            0.789 24409
## 10 2_Generation X Asian            0.755   644
## # ... with 23 more rows

#ETHNICITY
all %>%
  group_by(d_ethnicity) %>%
  summarise_at(vars(moral), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 9 x 3
##   d_ethnicity       mean     n
##   <fct>            <dbl> <int>
## 1 Asian            0.736  3091
## 2 Black            0.727  2610
## 3 Hispanic/Latin   0.731  2715
## 4 Indian           0.767   697
## 5 Middle Eastern   0.719   376
## 6 Native American  0.669   135
## 7 Other            0.770 19099
## 8 Pacific Islander 0.711   124
## 9 White            0.786 37424

#ALL THREE
all %>%
  group_by(d_gender, d_ethnicity, d_age) %>%
  summarise_at(vars(moral), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 71 x 5
## # Groups:   d_gender, d_ethnicity [?]
##    d_gender d_ethnicity    d_age           mean     n
##    <fct>    <fct>          <fct>          <dbl> <int>
##  1 1_Man    Asian          1_Millennials  0.680   964
##  2 1_Man    Asian          2_Generation X 0.734   408
##  3 1_Man    Asian          3_Baby Boomers 0.868    11
##  4 1_Man    Black          1_Millennials  0.644  1137
##  5 1_Man    Black          2_Generation X 0.728   464
##  6 1_Man    Black          3_Baby Boomers 0.749    35
##  7 1_Man    Hispanic/Latin 1_Millennials  0.682  1203
##  8 1_Man    Hispanic/Latin 2_Generation X 0.715   535
##  9 1_Man    Hispanic/Latin 3_Baby Boomers 0.747    27
## 10 1_Man    Indian         1_Millennials  0.727   451
## # ... with 61 more rows

#GENDER
all %>%
  group_by(d_gender) %>%
  summarise_at(vars(religion), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 3 x 3
##   d_gender  mean     n
##   <fct>    <dbl> <int>
## 1 1_Man    0.443 38995
## 2 2_Woman  0.435 25170
## 3 Other    0.418  2106

all %>%
  group_by(d_gender, d_age) %>%
  summarise_at(vars(religion), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 12 x 4
## # Groups:   d_gender [?]
##    d_gender d_age           mean     n
##    <fct>    <fct>          <dbl> <int>
##  1 1_Man    1_Millennials  0.423 24069
##  2 1_Man    2_Generation X 0.465 13673
##  3 1_Man    3_Baby Boomers 0.582  1227
##  4 1_Man    6_Other        0.512    26
##  5 2_Woman  1_Millennials  0.415 20442
##  6 2_Woman  2_Generation X 0.516  4558
##  7 2_Woman  3_Baby Boomers 0.584   163
##  8 2_Woman  6_Other        0.502     7
##  9 Other    1_Millennials  0.289   122
## 10 Other    2_Generation X 0.343    23
## 11 Other    3_Baby Boomers 0.361     2
## 12 Other    6_Other        0.428  1959
```

```
all %>%
  group_by(d_gender, d_ethnicity) %>%
  summarise_at(vars(religion), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 24 x 4
## # Groups:   d_gender [?]
##    d_gender d_ethnicity       mean     n
##    <fct>    <fct>            <dbl> <int>
##  1 1_Man    Asian            0.497  1383
##  2 1_Man    Black            0.745  1636
##  3 1_Man    Hispanic/Latin   0.590  1765
##  4 1_Man    Indian           0.525   549
##  5 1_Man    Middle Eastern   0.675   266
##  6 1_Man    Native American  0.593    85
##  7 1_Man    Other            0.492 10233
##  8 1_Man    Pacific Islander 0.603    78
##  9 1_Man    White            0.389 23000
## 10 2_Woman  Asian            0.562  1707
## # ... with 14 more rows
```

```
#AGE
all %>%
  group_by(d_age) %>%
  summarise_at(vars(religion), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 4 x 3
##   d_age           mean     n
##   <fct>          <dbl> <int>
## 1 1_Millennials  0.419 44633
## 2 2_Generation X 0.477 18254
## 3 3_Baby Boomers 0.582  1392
## 4 6_Other        0.430  1992
```

```
all %>%
  group_by(d_age, d_ethnicity) %>%
  summarise_at(vars(religion), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 33 x 4
## # Groups:   d_age [?]
##    d_age          d_ethnicity       mean     n
##    <fct>          <fct>            <dbl> <int>
##  1 1_Millennials  Asian            0.534  2434
##  2 1_Millennials  Black            0.765  1933
##  3 1_Millennials  Hispanic/Latin   0.589  1971
##  4 1_Millennials  Indian           0.535   581
##  5 1_Millennials  Middle Eastern   0.645   294
##  6 1_Millennials  Native American  0.533    86
##  7 1_Millennials  Other            0.453 12842
##  8 1_Millennials  Pacific Islander 0.657    83
##  9 1_Millennials  White            0.351 24409
## 10 2_Generation X Asian            0.522   644
## # ... with 23 more rows

#ETHNICITY
all %>%
  group_by(d_ethnicity) %>%
  summarise_at(vars(religion), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 9 x 3
##   d_ethnicity       mean     n
##   <fct>            <dbl> <int>
## 1 Asian            0.532  3091
## 2 Black            0.766  2610
## 3 Hispanic/Latin   0.608  2715
## 4 Indian           0.529   697
## 5 Middle Eastern   0.630   376
## 6 Native American  0.538   135
## 7 Other            0.465 19099
## 8 Pacific Islander 0.635   124
## 9 White            0.385 37424

#ALL THREE
all %>%
  group_by(d_gender, d_ethnicity, d_age) %>%
  summarise_at(vars(religion), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 71 x 5
## # Groups:   d_gender, d_ethnicity [?]
##    d_gender d_ethnicity    d_age           mean     n
##    <fct>    <fct>          <fct>          <dbl> <int>
##  1 1_Man    Asian          1_Millennials  0.494   964
##  2 1_Man    Asian          2_Generation X 0.499   408
##  3 1_Man    Asian          3_Baby Boomers 0.703    11
##  4 1_Man    Black          1_Millennials  0.745  1137
```

```
##   5 1_Man     Black          2_Generation X 0.738   464
##   6 1_Man     Black          3_Baby Boomers 0.820    35
##   7 1_Man     Hispanic/Latin 1_Millennials  0.569  1203
##   8 1_Man     Hispanic/Latin 2_Generation X 0.624   535
##   9 1_Man     Hispanic/Latin 3_Baby Boomers 0.799    27
##  10 1_Man     Indian         1_Millennials  0.535   451
## # ... with 61 more rows

#GENDER
all %>%
  group_by(d_gender) %>%
  summarise_at(vars(stigma), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 3 x 3
##   d_gender  mean     n
##   <fct>    <dbl> <int>
## 1 1_Man    0.392 38995
## 2 2_Woman  0.423 25170
## 3 Other    0.441  2106

all %>%
  group_by(d_gender, d_age) %>%
  summarise_at(vars(stigma), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 12 x 4
## # Groups:   d_gender [?]
##    d_gender d_age           mean     n
##    <fct>    <fct>          <dbl> <int>
##  1 1_Man    1_Millennials  0.386 24069
##  2 1_Man    2_Generation X 0.402 13673
##  3 1_Man    3_Baby Boomers 0.389  1227
##  4 1_Man    6_Other        0.404    26
##  5 2_Woman  1_Millennials  0.415 20442
##  6 2_Woman  2_Generation X 0.456  4558
##  7 2_Woman  3_Baby Boomers 0.403   163
##  8 2_Woman  6_Other        0.513     7
##  9 Other    1_Millennials  0.235   122
## 10 Other    2_Generation X 0.301    23
## 11 Other    3_Baby Boomers 0.266     2
## 12 Other    6_Other        0.456  1959

all %>%
  group_by(d_gender, d_ethnicity) %>%
  summarise_at(vars(stigma), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 24 x 4
## # Groups:   d_gender [?]
##    d_gender d_ethnicity       mean     n
##    <fct>    <fct>            <dbl> <int>
##  1 1_Man    Asian            0.413  1383
##  2 1_Man    Black            0.345  1636
```

```
##   3 1_Man    Hispanic/Latin   0.371  1765
##   4 1_Man    Indian           0.393   549
##   5 1_Man    Middle Eastern   0.412   266
##   6 1_Man    Native American  0.447    85
##   7 1_Man    Other            0.398 10233
##   8 1_Man    Pacific Islander 0.341    78
##   9 1_Man    White            0.393 23000
##  10 2_Woman  Asian            0.397  1707
## # ... with 14 more rows
```

```
#AGE
all %>%
  group_by(d_age) %>%
  summarise_at(vars(stigma), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 4 x 3
##   d_age            mean     n
##   <fct>           <dbl> <int>
## 1 1_Millennials   0.399 44633
## 2 2_Generation X  0.415 18254
## 3 3_Baby Boomers  0.390  1392
## 4 6_Other         0.455  1992
```

```
all %>%
  group_by(d_age, d_gender) %>%
  summarise_at(vars(stigma), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 12 x 4
## # Groups:   d_age [?]
##    d_age          d_gender  mean     n
##    <fct>          <fct>    <dbl> <int>
##  1 1_Millennials  1_Man    0.386 24069
##  2 1_Millennials  2_Woman  0.415 20442
##  3 1_Millennials  Other    0.235   122
##  4 2_Generation X 1_Man    0.402 13673
##  5 2_Generation X 2_Woman  0.456  4558
##  6 2_Generation X Other    0.301    23
##  7 3_Baby Boomers 1_Man    0.389  1227
##  8 3_Baby Boomers 2_Woman  0.403   163
##  9 3_Baby Boomers Other    0.266     2
## 10 6_Other        1_Man    0.404    26
## 11 6_Other        2_Woman  0.513     7
## 12 6_Other        Other    0.456  1959
```

```
all %>%
  group_by(d_age, d_ethnicity) %>%
  summarise_at(vars(stigma), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 33 x 4
## # Groups:   d_age [?]
##    d_age          d_ethnicity       mean     n
```

```
##    <fct>           <fct>             <dbl> <int>
##  1 1_Millennials  Asian             0.391  2434
##  2 1_Millennials  Black             0.349  1933
##  3 1_Millennials  Hispanic/Latin    0.377  1971
##  4 1_Millennials  Indian            0.377   581
##  5 1_Millennials  Middle Eastern    0.425   294
##  6 1_Millennials  Native American   0.452    86
##  7 1_Millennials  Other             0.407 12842
##  8 1_Millennials  Pacific Islander 0.416    83
##  9 1_Millennials  White             0.402 24409
## 10 2_Generation X Asian             0.452   644
## # ... with 23 more rows

#ETHNICITY
all %>%
  group_by(d_ethnicity) %>%
  summarise_at(vars(stigma), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 9 x 3
##   d_ethnicity       mean     n
##   <fct>            <dbl> <int>
## 1 Asian            0.404  3091
## 2 Black            0.362  2610
## 3 Hispanic/Latin   0.385  2715
## 4 Indian           0.386   697
## 5 Middle Eastern   0.426   376
## 6 Native American  0.458   135
## 7 Other            0.411 19099
## 8 Pacific Islander 0.420   124
## 9 White            0.407 37424

#ALL THREE
all %>%
  group_by(d_gender, d_ethnicity, d_age) %>%
  summarise_at(vars(stigma), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 71 x 5
## # Groups:   d_gender, d_ethnicity [?]
##    d_gender d_ethnicity    d_age           mean     n
##    <fct>    <fct>          <fct>          <dbl> <int>
##  1 1_Man    Asian          1_Millennials  0.398   964
##  2 1_Man    Asian          2_Generation X 0.443   408
##  3 1_Man    Asian          3_Baby Boomers 0.517    11
##  4 1_Man    Black          1_Millennials  0.331  1137
##  5 1_Man    Black          2_Generation X 0.368   464
##  6 1_Man    Black          3_Baby Boomers 0.499    35
##  7 1_Man    Hispanic/Latin 1_Millennials  0.364  1203
##  8 1_Man    Hispanic/Latin 2_Generation X 0.390   535
##  9 1_Man    Hispanic/Latin 3_Baby Boomers 0.299    27
## 10 1_Man    Indian         1_Millennials  0.385   451
## # ... with 61 more rows
```

```r
#GENDER
all %>%
  group_by(d_gender) %>%
  summarise_at(vars(substanceUse), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 3 x 3
##   d_gender  mean     n
##   <fct>    <dbl> <int>
## 1 1_Man    0.789 38995
## 2 2_Woman  0.844 25170
## 3 Other    0.743  2106
```

```r
all %>%
  group_by(d_gender, d_age) %>%
  summarise_at(vars(substanceUse), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 12 x 4
## # Groups:   d_gender [?]
##    d_gender d_age            mean     n
##    <fct>    <fct>           <dbl> <int>
##  1 1_Man    1_Millennials   0.775 24069
##  2 1_Man    2_Generation X 0.812 13673
##  3 1_Man    3_Baby Boomers 0.821  1227
##  4 1_Man    6_Other         0.805    26
##  5 2_Woman  1_Millennials   0.835 20442
##  6 2_Woman  2_Generation X 0.885  4558
##  7 2_Woman  3_Baby Boomers 0.865   163
##  8 2_Woman  6_Other         0.986     7
##  9 Other    1_Millennials   0.604   122
## 10 Other    2_Generation X 0.716    23
## 11 Other    3_Baby Boomers 0.689     2
## 12 Other    6_Other         0.754  1959
```

```r
all %>%
  group_by(d_gender, d_ethnicity) %>%
  summarise_at(vars(substanceUse), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 24 x 4
## # Groups:   d_gender [?]
##    d_gender d_ethnicity       mean     n
##    <fct>    <fct>            <dbl> <int>
##  1 1_Man    Asian            0.896  1383
##  2 1_Man    Black            0.841  1636
##  3 1_Man    Hispanic/Latin   0.840  1765
##  4 1_Man    Indian           0.854   549
##  5 1_Man    Middle Eastern   0.843   266
##  6 1_Man    Native American  0.751    85
##  7 1_Man    Other            0.778 10233
##  8 1_Man    Pacific Islander 0.784    78
##  9 1_Man    White            0.778 23000
```

```
## 10 2_Woman   Asian             0.996  1707
## # ... with 14 more rows
```

```
#AGE
all %>%
  group_by(d_age) %>%
  summarise_at(vars(substanceUse), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 4 x 3
##   d_age           mean     n
##   <fct>          <dbl> <int>
## 1 1_Millennials  0.802 44633
## 2 2_Generation X 0.829 18254
## 3 3_Baby Boomers 0.825  1392
## 4 6_Other        0.755  1992
```

```
all %>%
  group_by(d_age, d_ethnicity) %>%
  summarise_at(vars(substanceUse), funs(mean(., na.rm=TRUE), n=n()))
```

```
## # A tibble: 33 x 4
## # Groups:   d_age [?]
##    d_age          d_ethnicity       mean     n
##    <fct>          <fct>            <dbl> <int>
##  1 1_Millennials  Asian            0.951  2434
##  2 1_Millennials  Black            0.859  1933
##  3 1_Millennials  Hispanic/Latin   0.833  1971
##  4 1_Millennials  Indian           0.853   581
##  5 1_Millennials  Middle Eastern   0.829   294
##  6 1_Millennials  Native American  0.747    86
##  7 1_Millennials  Other            0.786 12842
##  8 1_Millennials  Pacific Islander 0.767    83
##  9 1_Millennials  White            0.787 24409
## 10 2_Generation X Asian            0.946   644
## # ... with 23 more rows
```

```
#ETHNICITY
all %>%
  group_by(d_ethnicity) %>%
  summarise_at(vars(substanceUse), funs(mean(., na.rm=TRUE), n=n()))
```

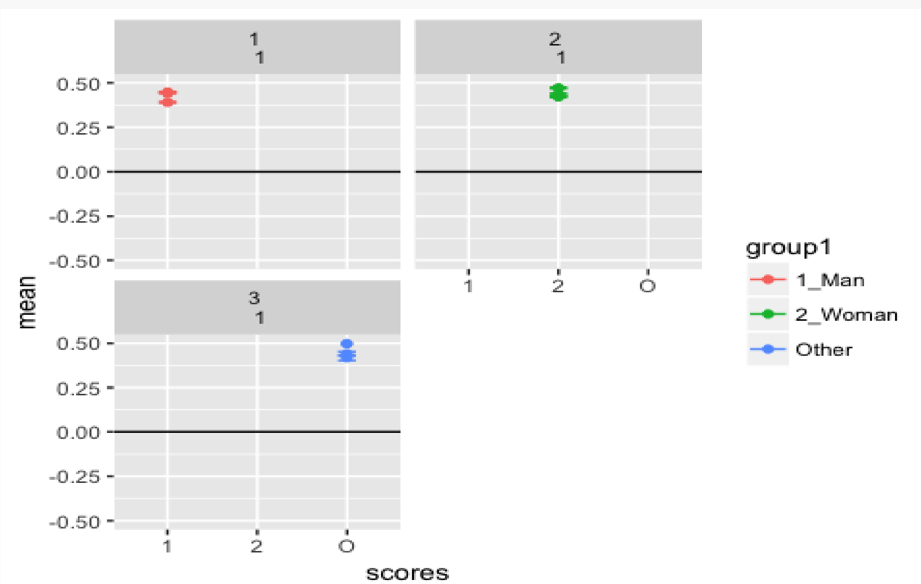
```
## # A tibble: 9 x 3
##   d_ethnicity       mean     n
##   <fct>            <dbl> <int>
## 1 Asian            0.950  3091
## 2 Black            0.873  2610
## 3 Hispanic/Latin   0.854  2715
## 4 Indian           0.864   697
## 5 Middle Eastern   0.854   376
## 6 Native American  0.762   135
## 7 Other            0.789 19099
```

```
## 8 Pacific Islander 0.799   124
## 9 White            0.797 37424

#ALL THREE
all %>%
  group_by(d_gender, d_ethnicity, d_age) %>%
  summarise_at(vars(substanceUse), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 71 x 5
## # Groups:   d_gender, d_ethnicity [?]
##    d_gender d_ethnicity    d_age           mean     n
##    <fct>    <fct>          <fct>          <dbl> <int>
##  1 1_Man    Asian          1_Millennials  0.889   964
##  2 1_Man    Asian          2_Generation X 0.911   408
##  3 1_Man    Asian          3_Baby Boomers 0.927    11
##  4 1_Man    Black          1_Millennials  0.816  1137
##  5 1_Man    Black          2_Generation X 0.899   464
##  6 1_Man    Black          3_Baby Boomers 0.870    35
##  7 1_Man    Hispanic/Latin 1_Millennials  0.820  1203
##  8 1_Man    Hispanic/Latin 2_Generation X 0.883   535
##  9 1_Man    Hispanic/Latin 3_Baby Boomers 0.888    27
## 10 1_Man    Indian         1_Millennials  0.843   451
## # ... with 61 more rows

#GENDER AND SCALES
all %>%
  group_by(d_gender) %>%
  summarise_at(vars(familyandfriend, appearance, education, moral, religion,
stigma, substanceUse), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 3 x 15
##   d_gender familyandfriend_mean appearance_mean education_mean moral_mean
##   <fct>                   <dbl>           <dbl>          <dbl>      <dbl>
## 1 1_Man                   0.448           0.560          0.601      0.735
## 2 2_Woman                 0.471           0.591          0.721      0.845
## 3 Other                   0.498           0.498          0.616      0.712
## # ... with 10 more variables: religion_mean <dbl>, stigma_mean <dbl>,
## #   substanceUse_mean <dbl>, familyandfriend_n <int>, appearance_n <int>,
## #   education_n <int>, moral_n <int>, religion_n <int>, stigma_n <int>,
## #   substanceUse_n <int>

#AGE AND SCALES
all %>%
  group_by(d_age) %>%
  summarise_at(vars(familyandfriend, appearance, education, moral, religion,
stigma, substanceUse), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 4 x 15
##   d_age        familyandfriend_m… appearance_mean education_mean moral_mean
##   <fct>                     <dbl>           <dbl>          <dbl>      <dbl>
## 1 1_Millenni…               0.461           0.580          0.656      0.779
```

```
## 2 2_Generati…                0.449           0.552          0.630      0.775
## 3 3_Baby Boo…                0.434           0.570          0.584      0.768
## 4 6_Other                    0.505           0.511          0.611      0.701
## # ... with 10 more variables: religion_mean <dbl>, stigma_mean <dbl>,
## #   substanceUse_mean <dbl>, familyandfriend_n <int>, appearance_n <int>,
## #   education_n <int>, moral_n <int>, religion_n <int>, stigma_n <int>,
## #   substanceUse_n <int>

#ETHNICITY AND SCALES
all %>%
  group_by(d_ethnicity) %>%
  summarise_at(vars(familyandfriend, appearance, education, moral, religion,
stigma, substanceUse), funs(mean(., na.rm=TRUE), n=n()))

## # A tibble: 9 x 15
##   d_ethnicity   familyandfriend… appearance_mean education_mean moral_mean
##   <fct>                    <dbl>           <dbl>          <dbl>      <dbl>
## 1 Asian                    0.495           0.662          0.628      0.736
## 2 Black                    0.436           0.612          0.626      0.727
## 3 Hispanic/Lat…            0.453           0.616          0.589      0.731
## 4 Indian                   0.489           0.640          0.645      0.767
## 5 Middle Easte…            0.432           0.690          0.607      0.719
## 6 Native Ameri…            0.432           0.537          0.571      0.669
## 7 Other                    0.441           0.577          0.650      0.770
## 8 Pacific Isla…            0.498           0.609          0.576      0.711
## 9 White                    0.463           0.553          0.648      0.786
## # ... with 10 more variables: religion_mean <dbl>, stigma_mean <dbl>,
## #   substanceUse_mean <dbl>, familyandfriend_n <int>, appearance_n <int>,
## #   education_n <int>, moral_n <int>, religion_n <int>, stigma_n <int>,
## #   substanceUse_n <int>

y1<-describeBy(all$familyandfriend, all$d_gender,
            skew=FALSE,IQR=FALSE,mat=TRUE,digits=3)
y1$cimin<-y1$mean-(1.96*y1$se)
y1$cimax<-y1$mean+(1.96*y1$se)

y2<-describeBy(all$appearance, all$d_gender,
            skew=FALSE,IQR=FALSE,mat=TRUE,digits=3)
y2$cimin<-y2$mean-(1.96*y2$se)
y2$cimax<-y2$mean+(1.96*y2$se)

y3<-
describeBy(all$education, all$d_gender,
            skew=FALSE,IQR=FALSE,mat=TRUE,digits=3)
y3$cimin<-y3$mean-(1.96*y3$se)
y3$cimax<-y3$mean+(1.96*y3$se)

y4<-
describeBy(all$moral, all$d_gender,
            skew=FALSE,IQR=FALSE,mat=TRUE,digits=3)
```

```
y4$cimin<-y4$mean-(1.96*y4$se)
y4$cimax<-y4$mean+(1.96*y4$se)

y5<-
describeBy(all$religion, all$d_gender,
            skew=FALSE,IQR=FALSE,mat=TRUE,digits=3)
y5$cimin<-y5$mean-(1.96*y5$se)
y5$cimax<-y5$mean+(1.96*y5$se)

y6<-
describeBy(all$stigma, all$d_gender,
            skew=FALSE,IQR=FALSE,mat=TRUE,digits=3)
y6$cimin<-y6$mean-(1.96*y6$se)
y6$cimax<-y6$mean+(1.96*y6$se)

y7<-
describeBy(all$substanceUse, all$d_gender,
            skew=FALSE,IQR=FALSE,mat=TRUE,digits=3)
y7$cimin<-y7$mean-(1.96*y7$se)
y7$cimax<-y7$mean+(1.96*y7$se)




plotme<-rbind(y1,y2,y3,y4,y5,y6, y7)
plotme$allSSscores<-substr(plotme$group1,1,1)
plotme$demo <-substr(plotme$group1,3,20)
plotme$allSSscores[plotme$demo==8] <- "8/9" #\t can be appended to 8/9 to avo
id interp as date

plotme$allSSscores[plotme$allSSscores=="9-Auton/Integ"] <- "Auton/Integ"
write.csv(plotme,"valuesForPlot.csv")

plotme<-read.csv("valuesForPlot.csv")
# default ordering of factor levels is alphabetical
plotme$item<-as.factor(paste(plotme$item,"\n",plotme$vars))
levels(plotme$item)

## [1] "1 \n 1" "2 \n 1" "3 \n 1"

# so these are reordered
plotme$item<-factor(plotme$item,levels(plotme$item))
levels(plotme$item)

## [1] "1 \n 1" "2 \n 1" "3 \n 1"

# ego level includes label and N to simplify graph
plotme$label<- paste0(plotme[,1],
                      " ", plotme[,2],
                      " (",plotme[,3],")")
```

```
levels(plotme$group1)

## [1] "1_Man"   "2_Woman" "Other"

plotme1 <- plotme[,c(2,3,6,12,13,14)]
a <- ggplot(plotme1) +
# add error bars
        geom_errorbar(aes(x=allSSscores, ymin=cimin,
                         ymax=cimax,color=group1),
                      width=.2, size=.5) +
# add means
        geom_point(aes(allSSscores, mean,color=group1),
                   size=1.5) +
# add axis titles
        labs(x="scores",
             y="mean") +
# set limits to y axis
            scale_y_continuous(limits=c(-.5,.5)) +
# add solid line at z=0
       geom_hline(aes(yintercept=0)) +
# make into multi-panel plot
       facet_wrap(~ item, nrow = 2)
a # display plot

## Warning: Removed 13 rows containing missing values (geom_errorbar).

## Warning: Removed 11 rows containing missing values (geom_point).
```



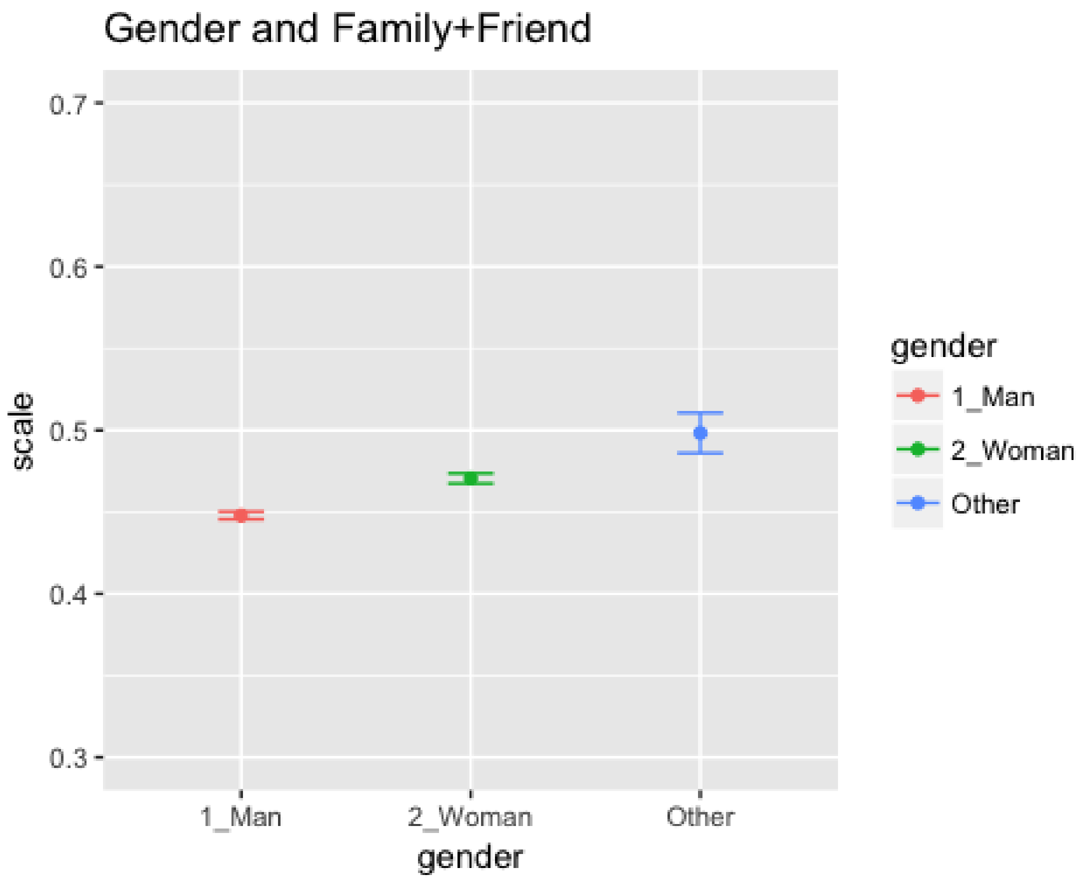
Appendix D. Gender Data Visualizations

Gender Data Visualizations

```
knitr::opts_chunk$set(echo = TRUE)
library(psych)

## Warning: package 'psych' was built under R version 3.4.4

library(knitr)
library(papeR)

## Loading required package: car

## Warning: package 'car' was built under R version 3.4.4

## Loading required package: carData

## Warning: package 'carData' was built under R version 3.4.4

##
## Attaching package: 'car'

## The following object is masked from 'package:psych':
##
##     logit

## Loading required package: xtable

##
## Attaching package: 'papeR'

## The following object is masked from 'package:utils':
##
##     toLatex

suppressMessages(library(tidyverse))
library(car)
library(dplyr)
library(sjPlot)

## Warning in checkMatrixPackageVersion(): Package version inconsistency dete
cted.
## TMB was built with Matrix version 1.2.12
## Current Matrix version is 1.2.14
## Please re-install 'TMB' from source using install.packages('TMB', type = '
source') or ask CRAN for a binary version of 'TMB' matching CRAN's 'Matrix' p
ackage

library(GPArotation)
library(ggthemes)
```

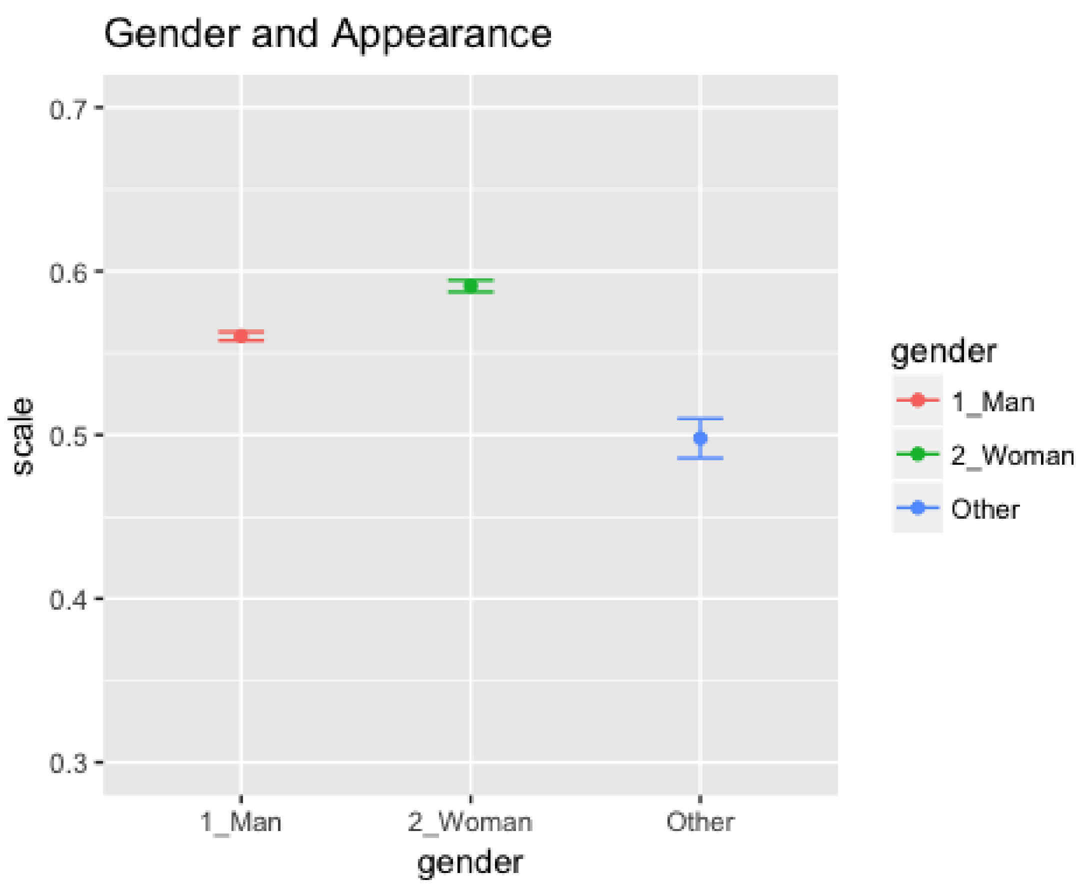
```
## Warning: package 'ggthemes' was built under R version 3.4.4
```

```
allSSscores<-read.csv('samScalescores1.csv')
demo <- read.csv('demographicData.csv')

all <-cbind(allSSscores,demo)

library(dplyr)


FFG <- all %>%
  group_by(d_gender) %>%
  summarise_at(vars(familyandfriend), funs(mean(., na.rm=TRUE),
                                          n=n(),
                                          sd(., na.rm = TRUE)
                                          ))
FFG[,5] <- FFG[,2] + (1.96 * FFG [,4]/sqrt(FFG[,3]))
FFG[,6] <- FFG[,2] - (1.96 * FFG [,4]/sqrt(FFG[,3]))
names(FFG) <- c("gender", "mean", "n", "sd", "ci.max", "ci.min")

plotmeffg<-FFG

plotmeffg$label<- plotmeffg[,1]

GendFam <- ggplot(plotmeffg) +
# add error bars
        geom_errorbar(aes(x=gender, ymin=ci.min,
                         ymax=ci.max,color=gender),
                      width=.2, size=.5, na.rm=TRUE) +
# add means
        geom_point(aes(gender, mean, color=gender),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="gender",
             y="scale",
             title= "Gender and Family+Friend") +
# set limits to y axis

            scale_y_continuous(limits=c(.3,.7))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

GendFam # display plot
```



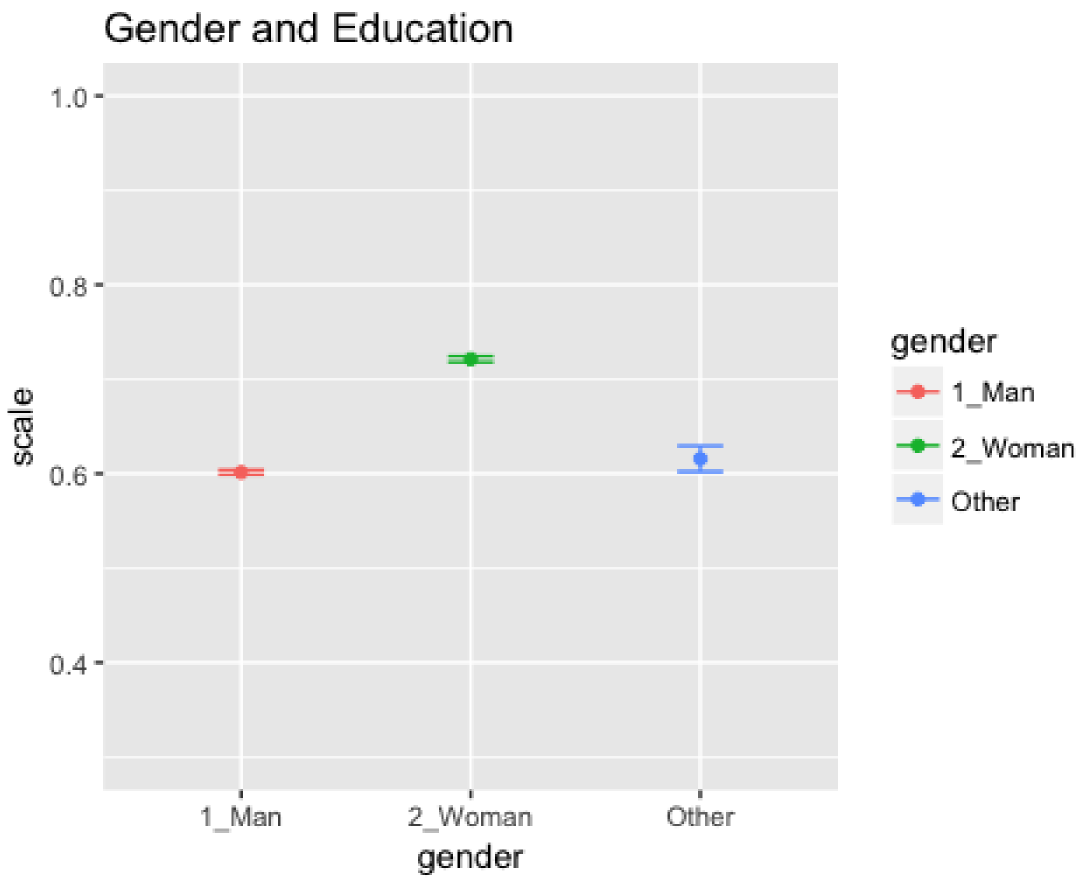

```
library(dplyr)


APG <- all %>%
  group_by(d_gender) %>%
  summarise_at(vars(appearance), funs(mean(., na.rm=TRUE),
                                          n=n(),
                                          sd(., na.rm = TRUE)
                                          ))
APG[,5] <- APG[,2] + (1.96 * APG [,4]/sqrt(APG[,3]))
APG[,6] <- APG[,2] - (1.96 * APG [,4]/sqrt(APG[,3]))
names(APG) <- c("gender", "mean", "n", "sd", "ci.max", "ci.min")

plotmeapg<-APG

plotmeapg$label<- plotmeapg[,1]

GendApp <- ggplot(plotmeapg) +
# add error bars
        geom_errorbar(aes(x=gender, ymin=ci.min,
                         ymax=ci.max,color=gender),
                      width=.2, size=.5, na.rm=TRUE) +
# add means
```

```
        geom_point(aes(gender, mean, color=gender),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="gender",
             y="scale",
             title= "Gender and Appearance") +
# set limits to y axis

            scale_y_continuous(limits=c(.3,.7))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

GendApp # display plot
```



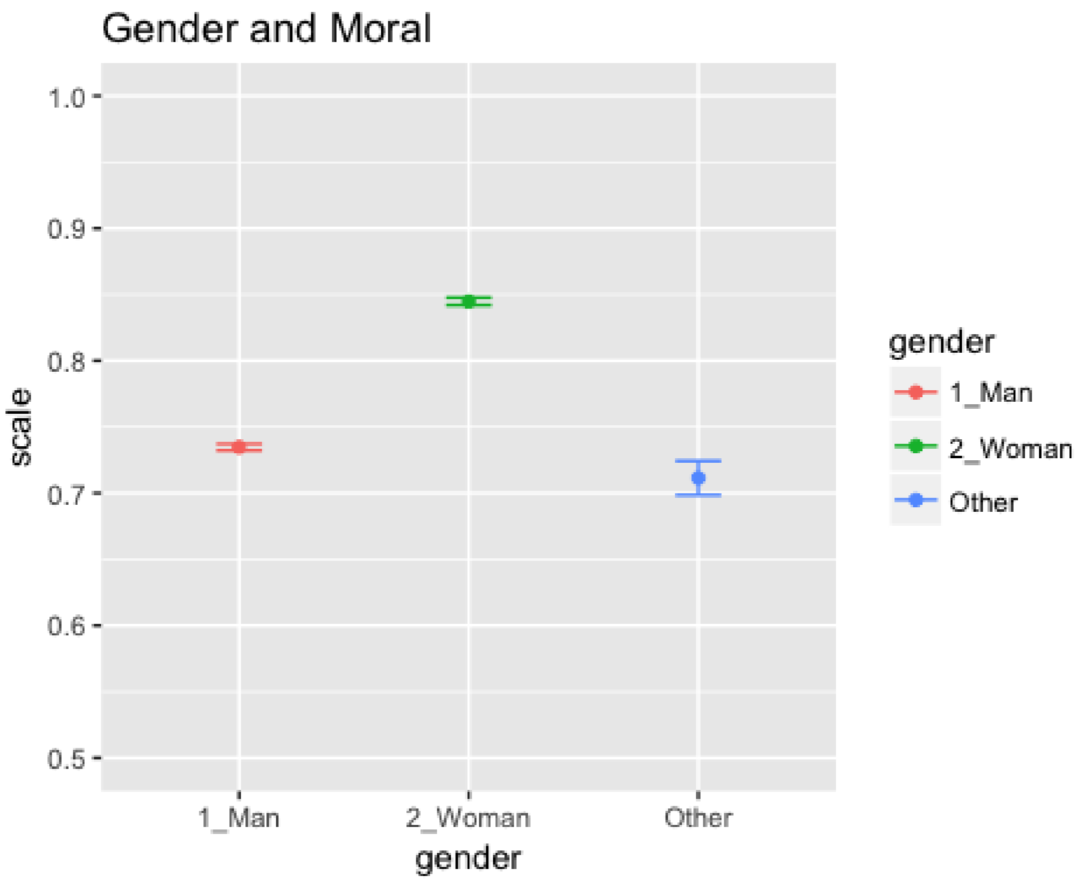

```
library(dplyr)


EDG <- all %>%
  group_by(d_gender) %>%
  summarise_at(vars(education), funs(mean(., na.rm=TRUE),
                                          n=n(),
```

```
                                             sd(., na.rm = TRUE)
                                             ))
EDG[,5] <- EDG[,2] + (1.96 * EDG [,4]/sqrt(EDG[,3]))
EDG[,6] <- EDG[,2] - (1.96 * EDG [,4]/sqrt(EDG[,3]))
names(EDG) <- c("gender", "mean", "n", "sd", "ci.max", "ci.min")

plotmeedg<-EDG

plotmeedg$label<- plotmeedg[,1]

GendEd <- ggplot(plotmeedg) +
# add error bars
        geom_errorbar(aes(x=gender, ymin=ci.min,
                         ymax=ci.max,color=gender),
                      width=.2, size=.5, na.rm=TRUE) +
# add means
        geom_point(aes(gender, mean, color=gender),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="gender",
             y="scale",
             title= "Gender and Education") +
# set limits to y axis

            scale_y_continuous(limits=c(.3,1))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

GendEd # display plot
```



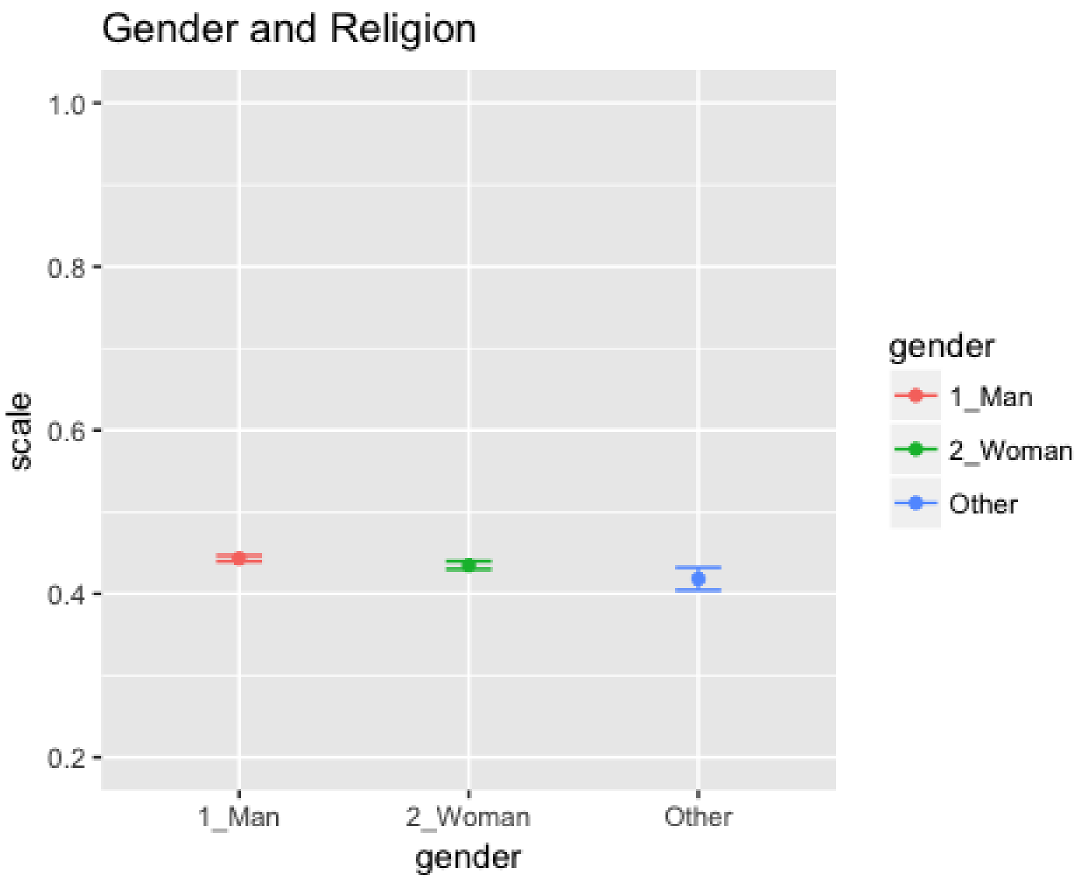

```
library(dplyr)


MG <- all %>%
  group_by(d_gender) %>%
  summarise_at(vars(moral), funs(mean(., na.rm=TRUE),
                                              n=n(),
                                              sd(., na.rm = TRUE)
                                              ))
MG[,5] <- MG[,2] + (1.96 * MG [,4]/sqrt(MG[,3]))
MG[,6] <- MG[,2] - (1.96 * MG [,4]/sqrt(MG[,3]))
names(MG) <- c("gender", "mean", "n", "sd", "ci.max", "ci.min")

plotmemg<-MG

plotmemg$label<- plotmemg[,1]

GendMo <- ggplot(plotmemg) +
# add error bars
        geom_errorbar(aes(x=gender, ymin=ci.min,
                         ymax=ci.max,color=gender),
                      width=.2, size=.5, na.rm=TRUE) +
# add means
```

```r
        geom_point(aes(gender, mean, color=gender),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="gender",
             y="scale",
             title= "Gender and Moral") +
# set limits to y axis

            scale_y_continuous(limits=c(.5,1))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

GendMo # display plot
```



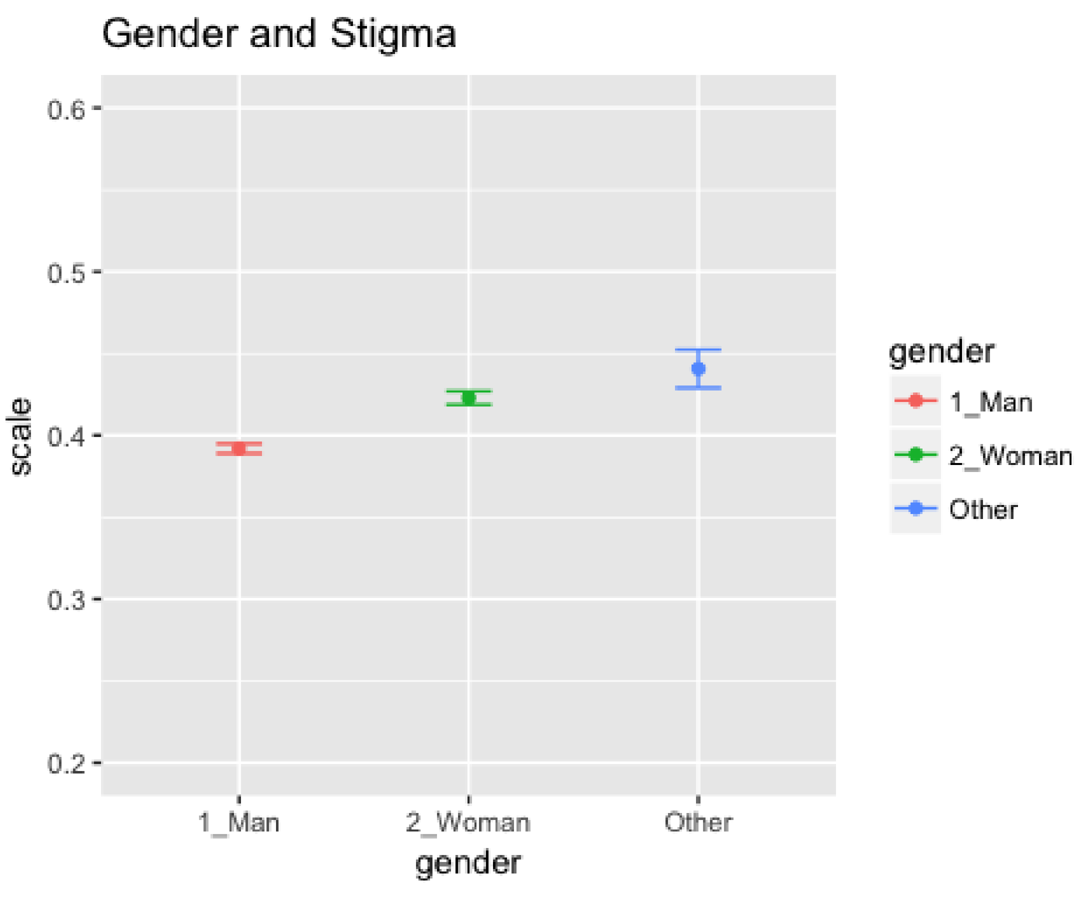

```r
library(dplyr)


RG <- all %>%
  group_by(d_gender) %>%
  summarise_at(vars(religion), funs(mean(., na.rm=TRUE),
                                        n=n(),
```

```
                                              sd(., na.rm = TRUE)
                                              ))
RG[,5] <- RG[,2] + (1.96 * RG [,4]/sqrt(RG[,3]))
RG[,6] <- RG[,2] - (1.96 * RG [,4]/sqrt(RG[,3]))
names(RG) <- c("gender", "mean", "n", "sd", "ci.max", "ci.min")

plotmerg<-RG

plotmerg$label<- plotmerg[,1]

GendRel <- ggplot(plotmerg) +
# add error bars
        geom_errorbar(aes(x=gender, ymin=ci.min,
                         ymax=ci.max,color=gender),
                      width=.2, size=.5, na.rm=TRUE) +
# add means
        geom_point(aes(gender, mean, color=gender),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="gender",
             y="scale",
             title= "Gender and Religion") +
# set limits to y axis

            scale_y_continuous(limits=c(.2,1))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

GendRel # display plot
```



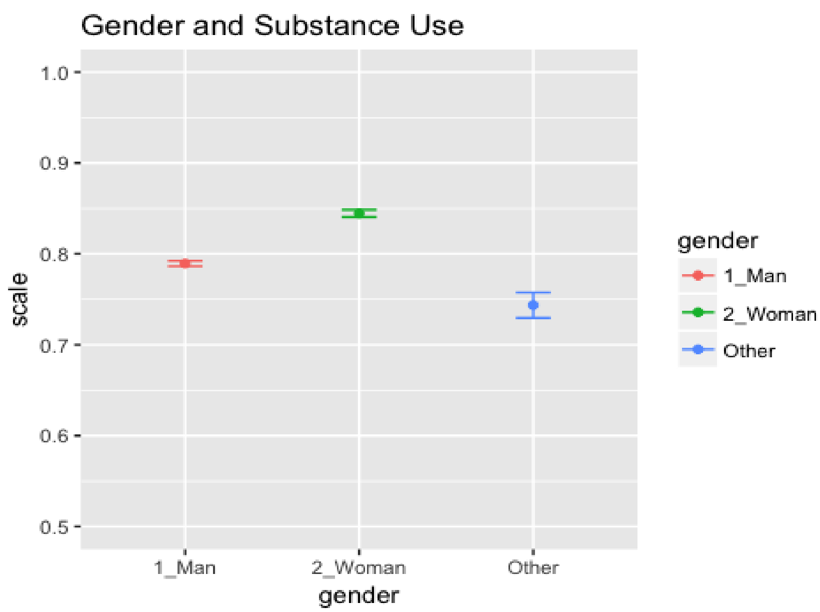

```
library(dplyr)


SG <- all %>%
  group_by(d_gender) %>%
  summarise_at(vars(stigma), funs(mean(., na.rm=TRUE),
                                  n=n(),
                                  sd(., na.rm = TRUE)
                                  ))
SG[,5] <- SG[,2] + (1.96 * SG [,4]/sqrt(SG[,3]))
SG[,6] <- SG[,2] - (1.96 * SG [,4]/sqrt(SG[,3]))
names(SG) <- c("gender", "mean", "n", "sd", "ci.max", "ci.min")

plotmesg<-SG

plotmesg$label<- plotmesg[,1]

GendStig <- ggplot(plotmesg) +
# add error bars
        geom_errorbar(aes(x=gender, ymin=ci.min,
                          ymax=ci.max,color=gender),
                      width=.2, size=.5, na.rm=TRUE) +
# add means
```

```
        geom_point(aes(gender, mean, color=gender),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="gender",
             y="scale",
             title= "Gender and Stigma") +
# set limits to y axis

            scale_y_continuous(limits=c(.2,.6))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

GendStig # display plot
```




```
library(dplyr)


SUG <- all %>%
  group_by(d_gender) %>%
  summarise_at(vars(substanceUse), funs(mean(., na.rm=TRUE),
                                         n=n(),
```

```
                                                  sd(., na.rm = TRUE)
                                                  ))
SUG[,5] <- SUG[,2] + (1.96 * SUG [,4]/sqrt(SUG[,3]))
SUG[,6] <- SUG[,2] - (1.96 * SUG [,4]/sqrt(SUG[,3]))
names(SUG) <- c("gender", "mean", "n", "sd", "ci.max", "ci.min")

plotmesug<-SUG

plotmesug$label<- plotmesug[,1]

GendSU <- ggplot(plotmesug) +
# add error bars
        geom_errorbar(aes(x=gender, ymin=ci.min,
                         ymax=ci.max,color=gender),
                      width=.2, size=.5, na.rm=TRUE) +
# add means
        geom_point(aes(gender, mean, color=gender),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="gender",
             y="scale",
             title= "Gender and Substance Use") +
# set limits to y axis

            scale_y_continuous(limits=c(.5,1))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

GendSU # display plot
```

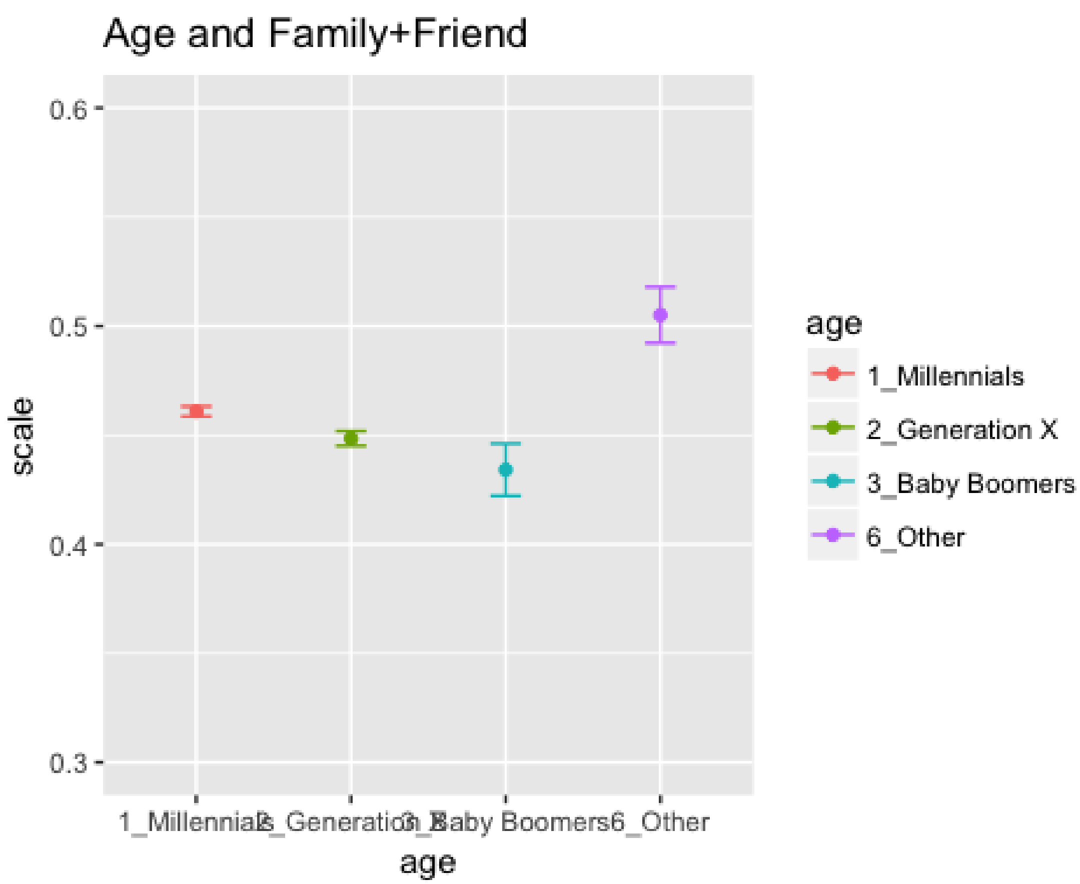


Appendix E. Age Data Visualizations

Age Data Visualizations

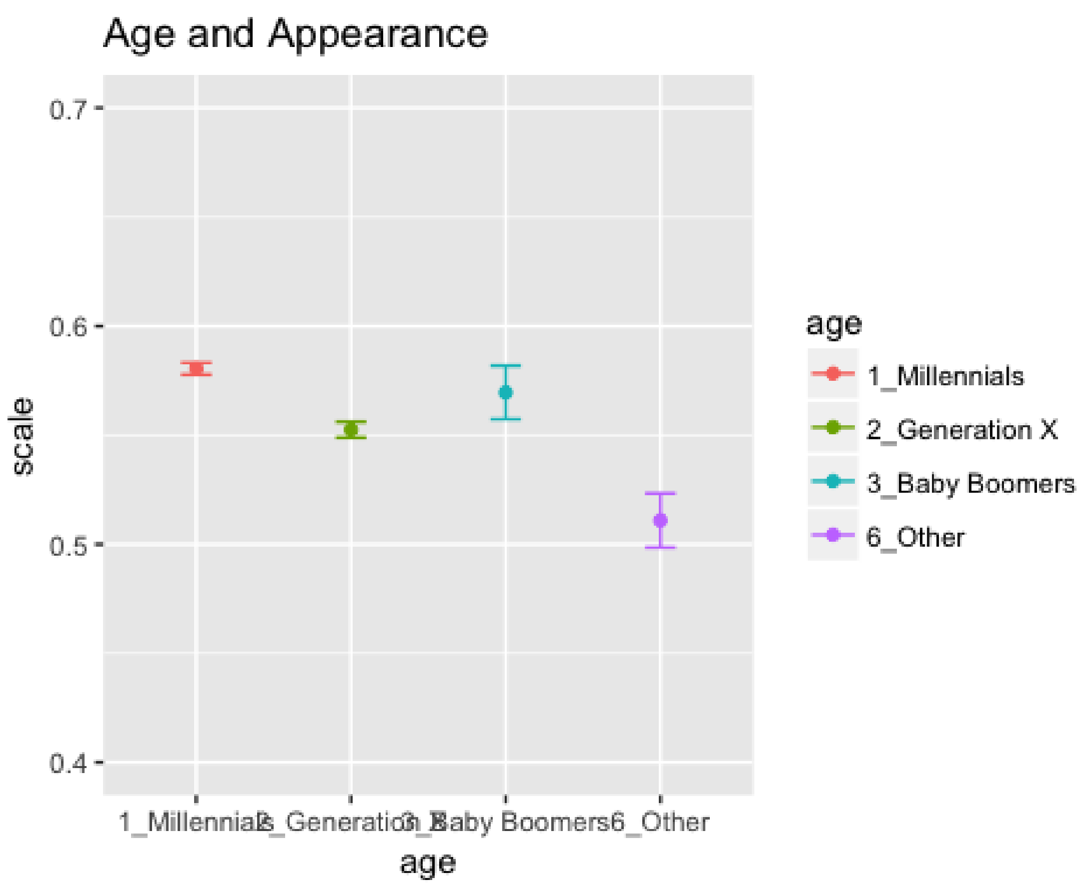
```
knitr::opts_chunk$set(echo = TRUE)
library(psych)

## Warning: package 'psych' was built under R version 3.4.4

library(knitr)
library(papeR)

## Loading required package: car

## Warning: package 'car' was built under R version 3.4.4

## Loading required package: carData

## Warning: package 'carData' was built under R version 3.4.4

##
## Attaching package: 'car'

## The following object is masked from 'package:psych':
##
##     logit

## Loading required package: xtable

##
## Attaching package: 'papeR'

## The following object is masked from 'package:utils':
##
##     toLatex

suppressMessages(library(tidyverse))
library(car)
library(dplyr)
library(sjPlot)

## Warning in checkMatrixPackageVersion(): Package version inconsistency dete
cted.
## TMB was built with Matrix version 1.2.12
## Current Matrix version is 1.2.14
## Please re-install 'TMB' from source using install.packages('TMB', type = '
source') or ask CRAN for a binary version of 'TMB' matching CRAN's 'Matrix' p
ackage

## Learn more about sjPlot with 'browseVignettes("sjPlot")'.
```

```
library(GPArotation)
library(ggthemes)

## Warning: package 'ggthemes' was built under R version 3.4.4

allSSscores<-read.csv('samScalescores1.csv')
demo <- read.csv('demographicData.csv')

all <-cbind(allSSscores,demo)

library(dplyr)


FFA <- all %>%
  group_by(d_age) %>%
  summarise_at(vars(familyandfriend), funs(mean(., na.rm=TRUE),
                                             n=n(),
                                             sd(., na.rm = TRUE)
                                             ))
FFA[,5] <- FFA[,2] + (1.96 * FFA [,4]/sqrt(FFA[,3]))
FFA[,6] <- FFA[,2] - (1.96 * FFA [,4]/sqrt(FFA[,3]))
names(FFA) <- c("age", "mean", "n", "sd", "ci.max", "ci.min")

plotmeffa<-FFA

plotmeffa$label<- plotmeffa[,1]

AgeFam <- ggplot(plotmeffa) +
# add error bars
        geom_errorbar(aes(x=age, ymin=ci.min,
                          ymax=ci.max,color=age),
                       width=.2, size=.5, na.rm=TRUE) +
# add means
        geom_point(aes(age, mean, color=age),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="age",
             y="scale",
             title= "Age and Family+Friend") +
# set limits to y axis

            scale_y_continuous(limits=c(.3,.6))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

AgeFam# display plot
```



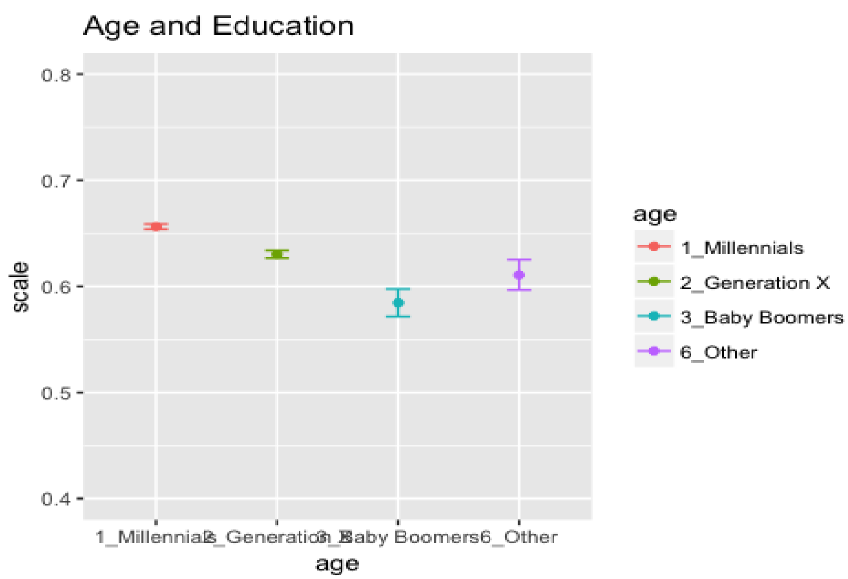

```
library(dplyr)


APA <- all %>%
  group_by(d_age) %>%
  summarise_at(vars(appearance), funs(mean(., na.rm=TRUE),
                                          n=n(),
                                          sd(., na.rm = TRUE)
                                          ))
APA[,5] <- APA[,2] + (1.96 * APA [,4]/sqrt(APA[,3]))
APA[,6] <- APA[,2] - (1.96 * APA [,4]/sqrt(APA[,3]))
names(APA) <- c("age", "mean", "n", "sd", "ci.max", "ci.min")

plotmeapa<-APA

plotmeapa$label<- plotmeapa[,1]

AgeApp <- ggplot(plotmeapa) +
# add error bars
        geom_errorbar(aes(x=age, ymin=ci.min,
                          ymax=ci.max,color=age),
                       width=.2, size=.5, na.rm=TRUE) +
# add means
```

```r
        geom_point(aes(age, mean, color=age),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="age",
             y="scale",
             title= "Age and Appearance") +
# set limits to y axis

            scale_y_continuous(limits=c(.4,.7))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

AgeApp # display plot
```

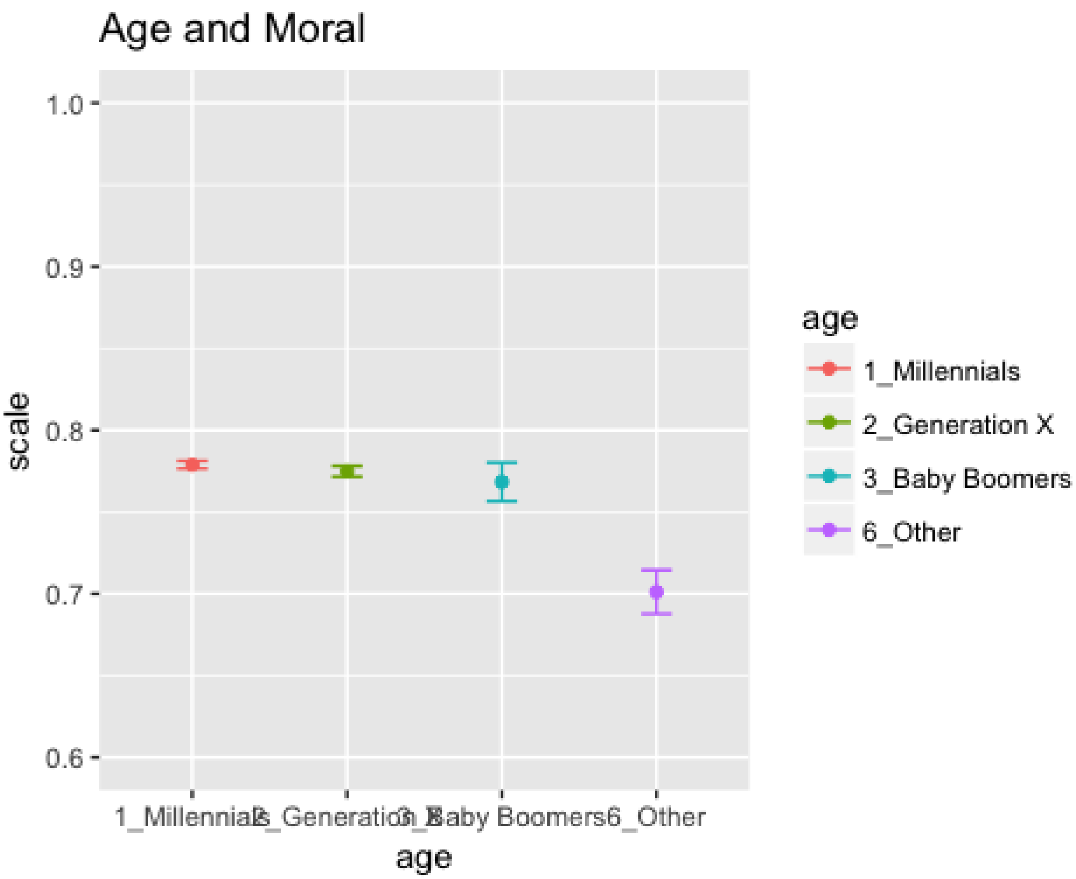



```r
library(dplyr)


EDA <- all %>%
  group_by(d_age) %>%
  summarise_at(vars(education), funs(mean(., na.rm=TRUE),
                                          n=n(),
```

```
                                         sd(., na.rm = TRUE)
                                         ))
EDA[,5] <- EDA[,2] + (1.96 * EDA [,4]/sqrt(EDA[,3]))
EDA[,6] <- EDA[,2] - (1.96 * EDA [,4]/sqrt(EDA[,3]))
names(EDA) <- c("age", "mean", "n", "sd", "ci.max", "ci.min")

plotmeeda<-EDA

plotmeeda$label<- plotmeeda[,1]

AgeEd <- ggplot(plotmeeda) +
# add error bars
        geom_errorbar(aes(x=age, ymin=ci.min,
                         ymax=ci.max,color=age),
                      width=.2, size=.5, na.rm=TRUE) +
# add means
        geom_point(aes(age, mean, color=age),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="age",
             y="scale",
             title= "Age and Education") +
# set limits to y axis

            scale_y_continuous(limits=c(.4,.8))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

AgeEd # display plot
```

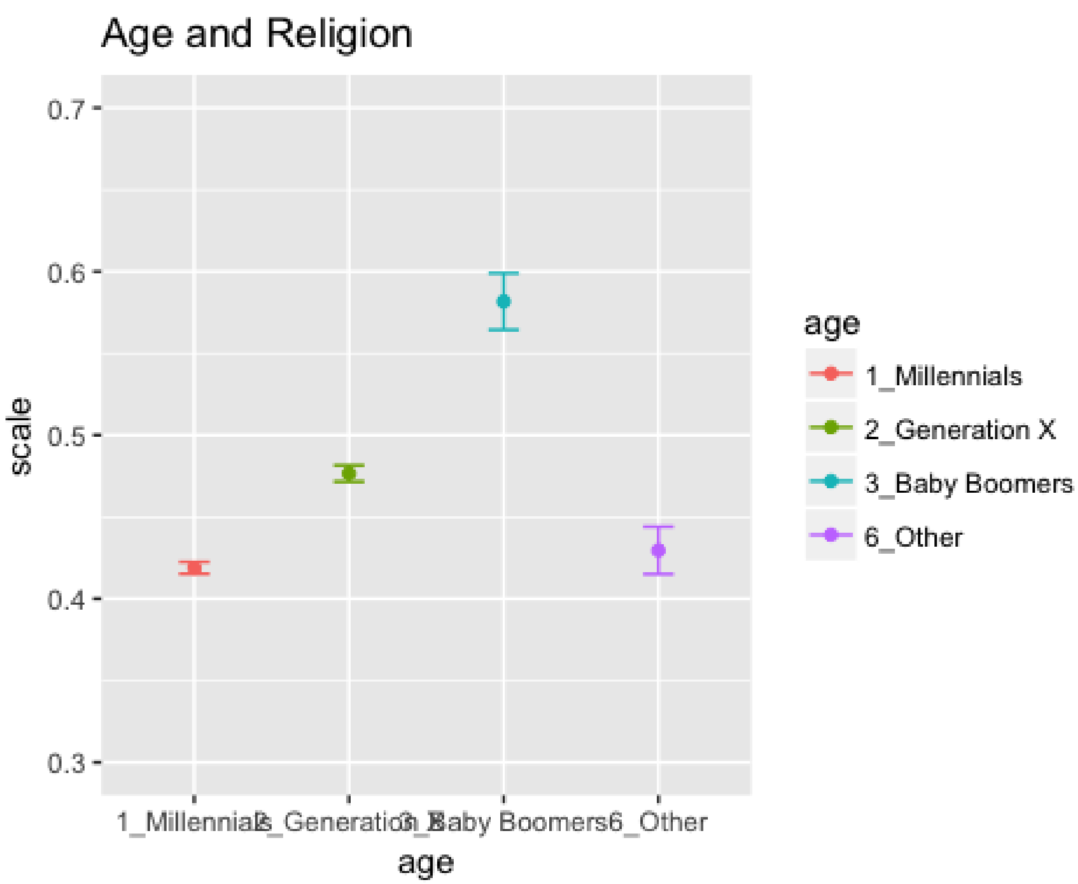


```r
library(dplyr)


MA <- all %>%
  group_by(d_age) %>%
  summarise_at(vars(moral), funs(mean(., na.rm=TRUE),
                                           n=n(),
                                           sd(., na.rm = TRUE)
                                           ))
MA[,5] <- MA[,2] + (1.96 * MA [,4]/sqrt(MA[,3]))
MA[,6] <- MA[,2] - (1.96 * MA [,4]/sqrt(MA[,3]))
names(MA) <- c("age", "mean", "n", "sd", "ci.max", "ci.min")

plotmema<-MA

plotmema$label<- plotmema[,1]

AgeMo <- ggplot(plotmema) +
# add error bars
        geom_errorbar(aes(x=age, ymin=ci.min,
                         ymax=ci.max,color=age),
                      width=.2, size=.5, na.rm=TRUE) +
# add means
        geom_point(aes(age, mean, color=age),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="age",
             y="scale",
             title= "Age and Moral") +
# set limits to y axis

            scale_y_continuous(limits=c(.6,1))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

AgeMo # display plot
```



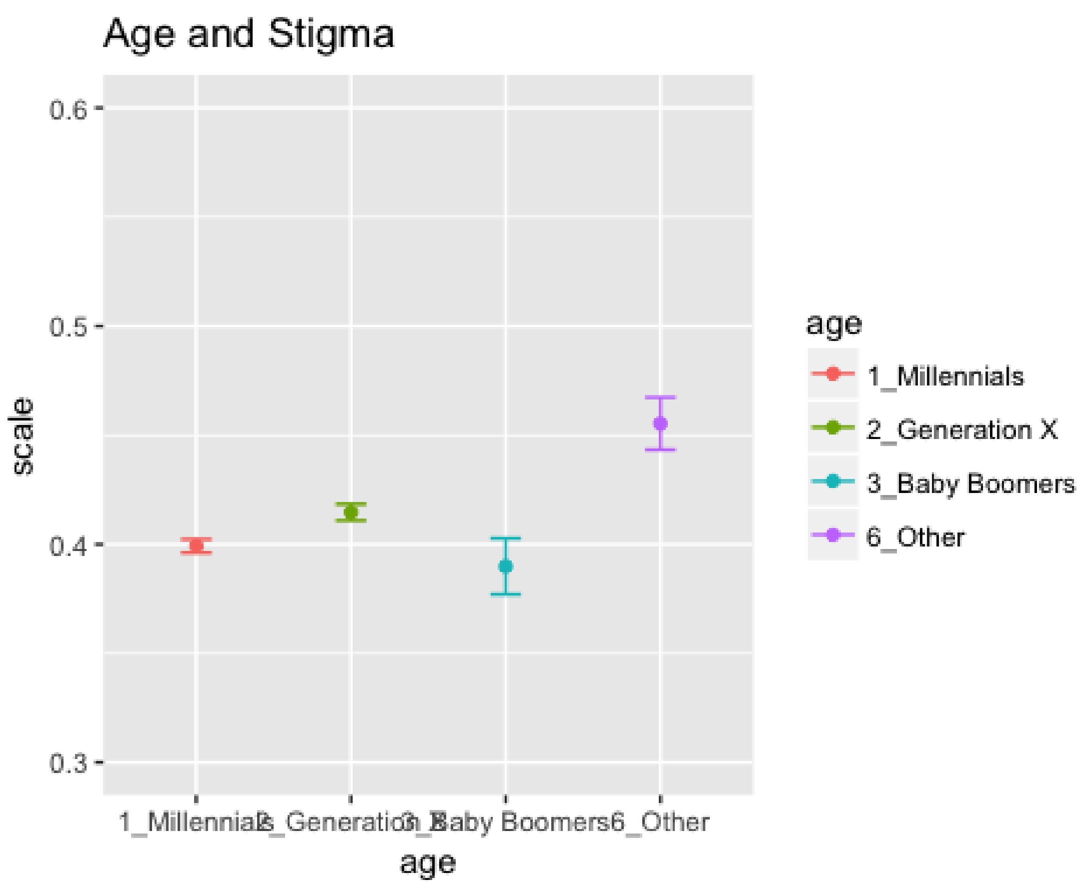

```
library(dplyr)


RA <- all %>%
  group_by(d_age) %>%
  summarise_at(vars(religion), funs(mean(., na.rm=TRUE),
                                          n=n(),
                                          sd(., na.rm = TRUE)
                                          ))
RA[,5] <- RA[,2] + (1.96 * RA [,4]/sqrt(RA[,3]))
RA[,6] <- RA[,2] - (1.96 * RA [,4]/sqrt(RA[,3]))
names(RA) <- c("age", "mean", "n", "sd", "ci.max", "ci.min")

plotmera<-RA

plotmera$label<- plotmera[,1]

AgeRel <- ggplot(plotmera) +
# add error bars
        geom_errorbar(aes(x=age, ymin=ci.min,
                         ymax=ci.max,color=age),
                      width=.2, size=.5, na.rm=TRUE) +
# add means
```

```r
        geom_point(aes(age, mean, color=age),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="age",
             y="scale",
             title= "Age and Religion") +
# set limits to y axis

            scale_y_continuous(limits=c(.3,.7))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

AgeRel # display plot
```



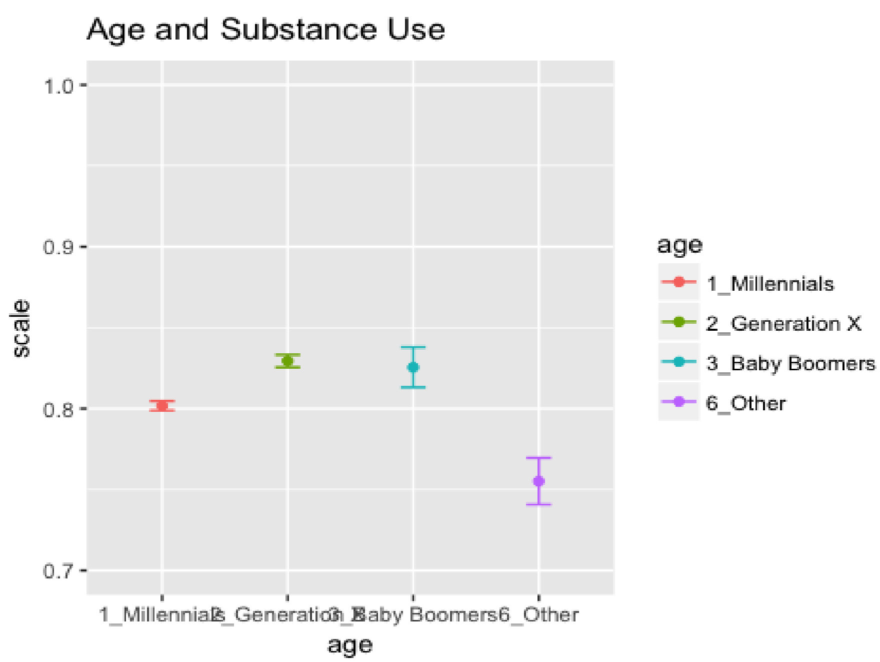

```r
library(dplyr)


SA <- all %>%
  group_by(d_age) %>%
  summarise_at(vars(stigma), funs(mean(., na.rm=TRUE),
                                    n=n(),
```

```
                                                 sd(., na.rm = TRUE)
                                                 ))
SA[,5] <- SA[,2] + (1.96 * SA [,4]/sqrt(SA[,3]))
SA[,6] <- SA[,2] - (1.96 * SA [,4]/sqrt(SA[,3]))
names(SA) <- c("age", "mean", "n", "sd", "ci.max", "ci.min")

plotmesa<-SA

plotmesa$label<- plotmesa[,1]

AgeStig <- ggplot(plotmesa) +
# add error bars
        geom_errorbar(aes(x=age, ymin=ci.min,
                         ymax=ci.max,color=age),
                      width=.2, size=.5, na.rm=TRUE) +
# add means
        geom_point(aes(age, mean, color=age),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="age",
             y="scale",
             title= "Age and Stigma") +
# set limits to y axis

            scale_y_continuous(limits=c(.3,.6))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

AgeStig # display plot
```




```r
library(dplyr)


SUA <- all %>%
  group_by(d_age) %>%
  summarise_at(vars(substanceUse), funs(mean(., na.rm=TRUE),
                                         n=n(),
                                         sd(., na.rm = TRUE)
                                         ))
SUA[,5] <- SUA[,2] + (1.96 * SUA [,4]/sqrt(SUA[,3]))
SUA[,6] <- SUA[,2] - (1.96 * SUA [,4]/sqrt(SUA[,3]))
names(SUA) <- c("age", "mean", "n", "sd", "ci.max", "ci.min")

plotmesua<-SUA

plotmesua$label<- plotmesua[,1]

AgeSU <- ggplot(plotmesua) +
# add error bars
        geom_errorbar(aes(x=age, ymin=ci.min,
                          ymax=ci.max,color=age),
                       width=.2, size=.5, na.rm=TRUE) +
# add means
```

```
        geom_point(aes(age, mean, color=age),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="age",
             y="scale",
             title= "Age and Substance Use") +
# set limits to y axis

            scale_y_continuous(limits=c(.7,1))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

AgeSU # display plot
```



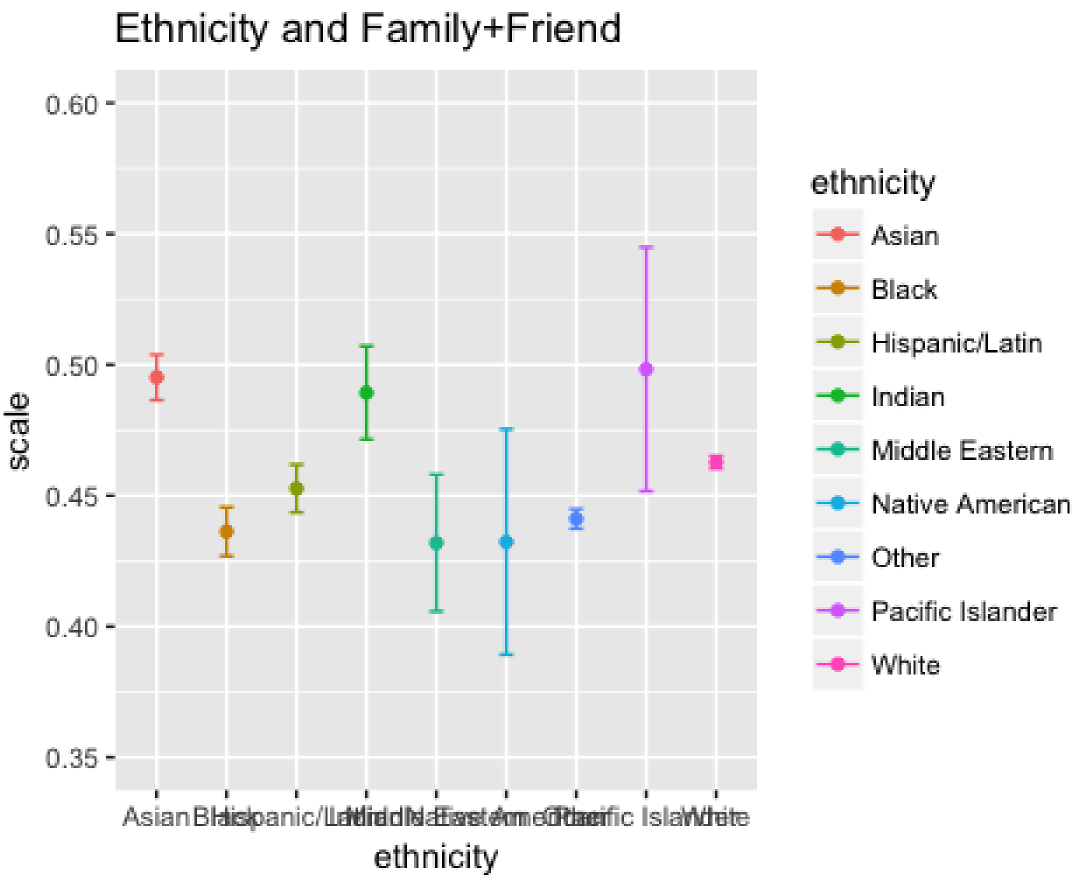
Appendix F. Ethnicity Data Visualizations

Ethnicity Data Visualizations

```
knitr::opts_chunk$set(echo = TRUE)
library(psych)

## Warning: package 'psych' was built under R version 3.4.4
```

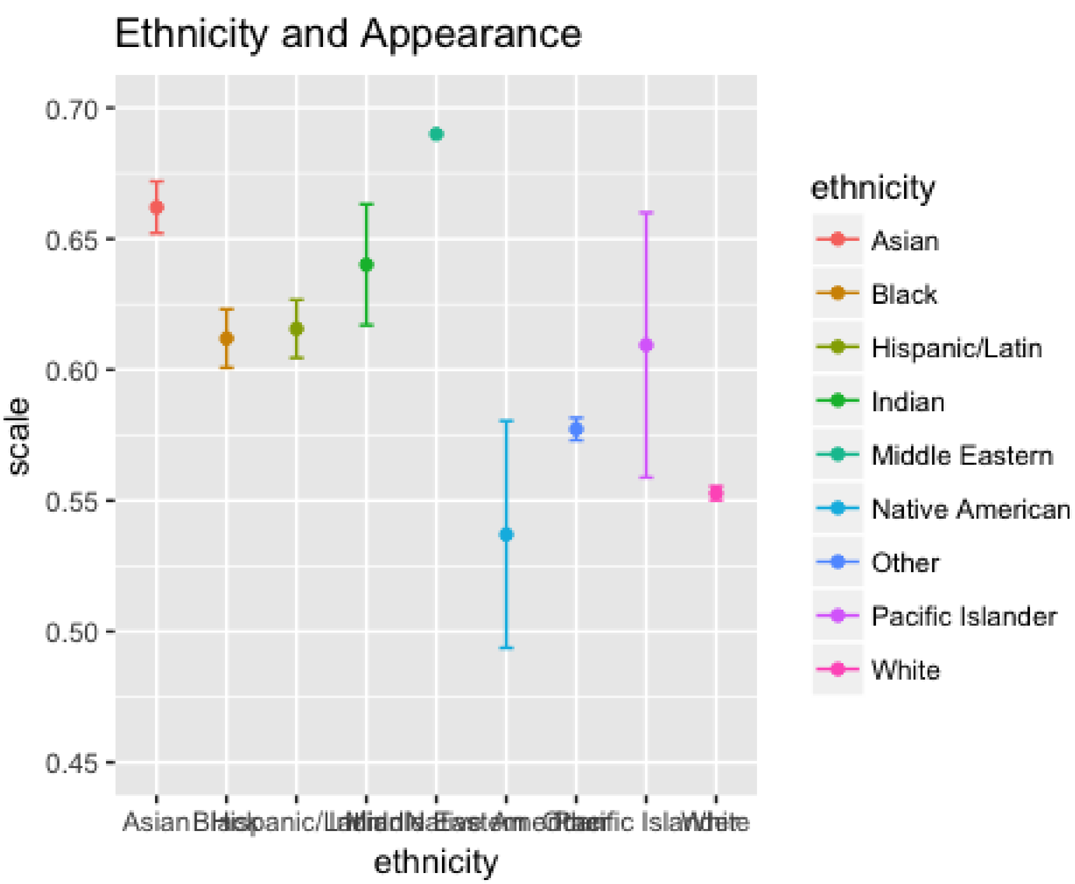
```r
library(knitr)
library(papeR)

## Loading required package: car

## Warning: package 'car' was built under R version 3.4.4

## Loading required package: carData

## Warning: package 'carData' was built under R version 3.4.4

##
## Attaching package: 'car'

## The following object is masked from 'package:psych':
##
##     logit

## Loading required package: xtable

##
## Attaching package: 'papeR'

## The following object is masked from 'package:utils':
##
##     toLatex

suppressMessages(library(tidyverse))
library(car)
library(dplyr)
library(sjPlot)

## Warning in checkMatrixPackageVersion(): Package version inconsistency dete
cted.
## TMB was built with Matrix version 1.2.12
## Current Matrix version is 1.2.14
## Please re-install 'TMB' from source using install.packages('TMB', type = '
source') or ask CRAN for a binary version of 'TMB' matching CRAN's 'Matrix' p
ackage

## #refugeeswelcome

library(GPArotation)
library(ggthemes)

## Warning: package 'ggthemes' was built under R version 3.4.4

allSSscores<-read.csv('samScalescores1.csv')
demo <- read.csv('demographicData.csv')

all <-cbind(allSSscores,demo)

library(dplyr)
```

```
FFE <- all %>%
  group_by(d_ethnicity) %>%
  summarise_at(vars(familyandfriend), funs(mean(., na.rm=TRUE),
                                              n=n(),
                                              sd(., na.rm = TRUE)
                                              ))
FFE[,5] <- FFE[,2] + (1.96 * FFE [,4]/sqrt(FFE[,3]))
FFE[,6] <- FFE[,2] - (1.96 * FFE [,4]/sqrt(FFE[,3]))
names(FFE) <- c("ethnicity", "mean", "n", "sd", "ci.max", "ci.min")

plotmeffe<-FFE

plotmeffe$label<- plotmeffe[,1]

EthFam <- ggplot(plotmeffe) +
# add error bars
        geom_errorbar(aes(x=ethnicity, ymin=ci.min,
                          ymax=ci.max,color=ethnicity),
                      width=.2, size=.5, na.rm=TRUE) +
# add means
        geom_point(aes(ethnicity, mean, color=ethnicity),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="ethnicity",
             y="scale",
             title= "Ethnicity and Family+Friend") +
# set limits to y axis

            scale_y_continuous(limits=c(.35,.6))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

EthFam# display plot
```



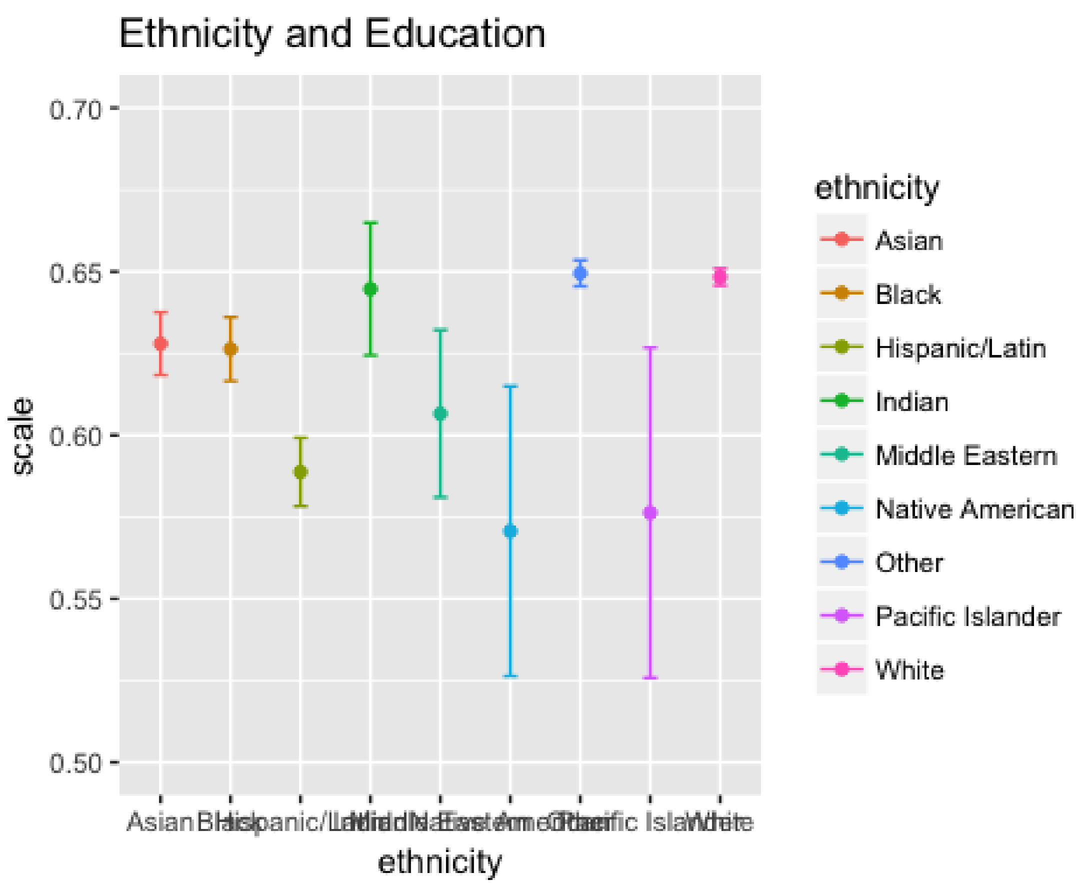

```
library(dplyr)


APE <- all %>%
  group_by(d_ethnicity) %>%
  summarise_at(vars(appearance), funs(mean(., na.rm=TRUE),
                                          n=n(),
                                          sd(., na.rm = TRUE)
                                          ))
APE[,5] <- APE[,2] + (1.96 * APE [,4]/sqrt(APE[,3]))
APE[,6] <- APE[,2] - (1.96 * APE [,4]/sqrt(APE[,3]))
names(APE) <- c("ethnicity", "mean", "n", "sd", "ci.max", "ci.min")

plotmeape<-APE

plotmeape$label<- plotmeape[,1]

EthApp <- ggplot(plotmeape) +
# add error bars
        geom_errorbar(aes(x=ethnicity, ymin=ci.min,
                         ymax=ci.max,color=ethnicity),
                      width=.2, size=.5, na.rm=TRUE) +
# add means
```

```
        geom_point(aes(ethnicity, mean, color=ethnicity),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="ethnicity",
             y="scale",
             title= "Ethnicity and Appearance") +
# set limits to y axis

            scale_y_continuous(limits=c(.45,.7))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

EthApp # display plot
```



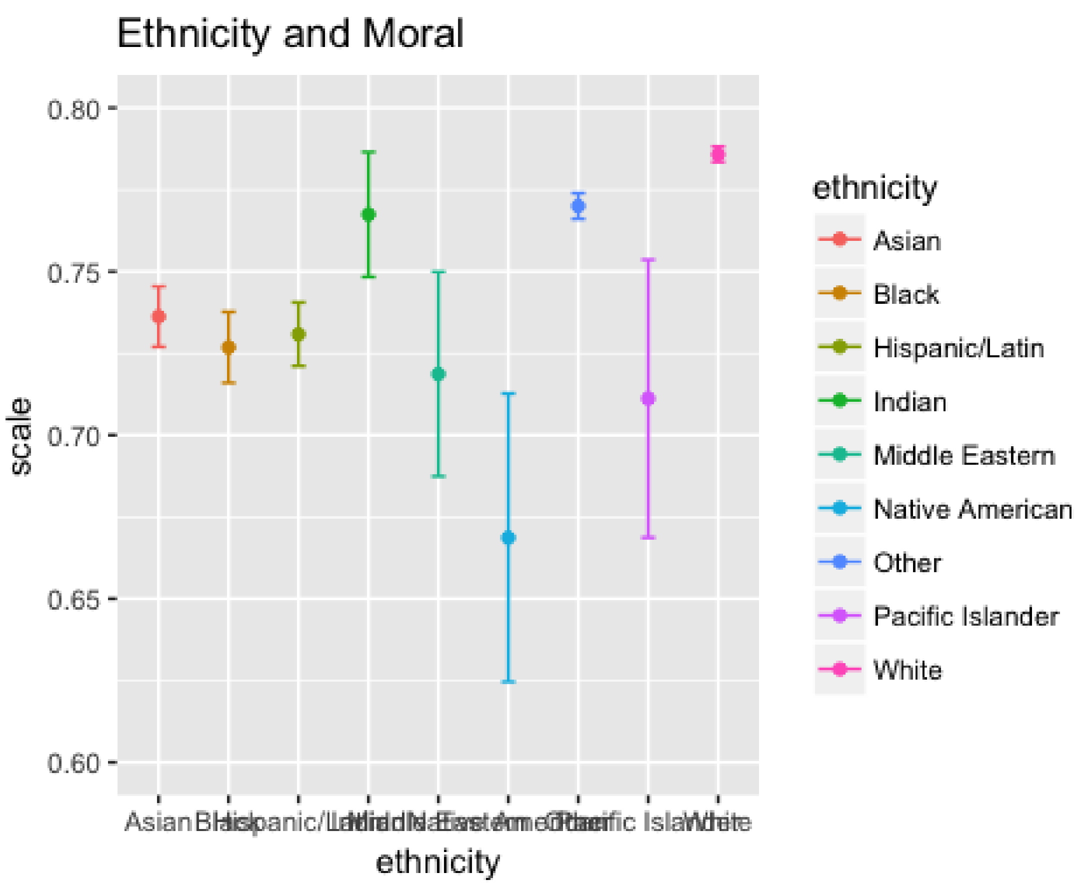

```
library(dplyr)


EDE <- all %>%
  group_by(d_ethnicity) %>%
  summarise_at(vars(education), funs(mean(., na.rm=TRUE),
                                         n=n(),
```

```
                                            sd(., na.rm = TRUE)
                                            ))
EDE[,5] <- EDE[,2] + (1.96 * EDE [,4]/sqrt(EDE[,3]))
EDE[,6] <- EDE[,2] - (1.96 * EDE [,4]/sqrt(EDE[,3]))
names(EDE) <- c("ethnicity", "mean", "n", "sd", "ci.max", "ci.min")

plotmeede<-EDE

plotmeede$label<- plotmeede[,1]

EthEd <- ggplot(plotmeede) +
# add error bars
        geom_errorbar(aes(x=ethnicity, ymin=ci.min,
                         ymax=ci.max,color=ethnicity),
                      width=.2, size=.5, na.rm=TRUE) +
# add means
        geom_point(aes(ethnicity, mean, color=ethnicity),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="ethnicity",
             y="scale",
             title= "Ethnicity and Education") +
# set limits to y axis

            scale_y_continuous(limits=c(.5,.7))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

EthEd # display plot
```



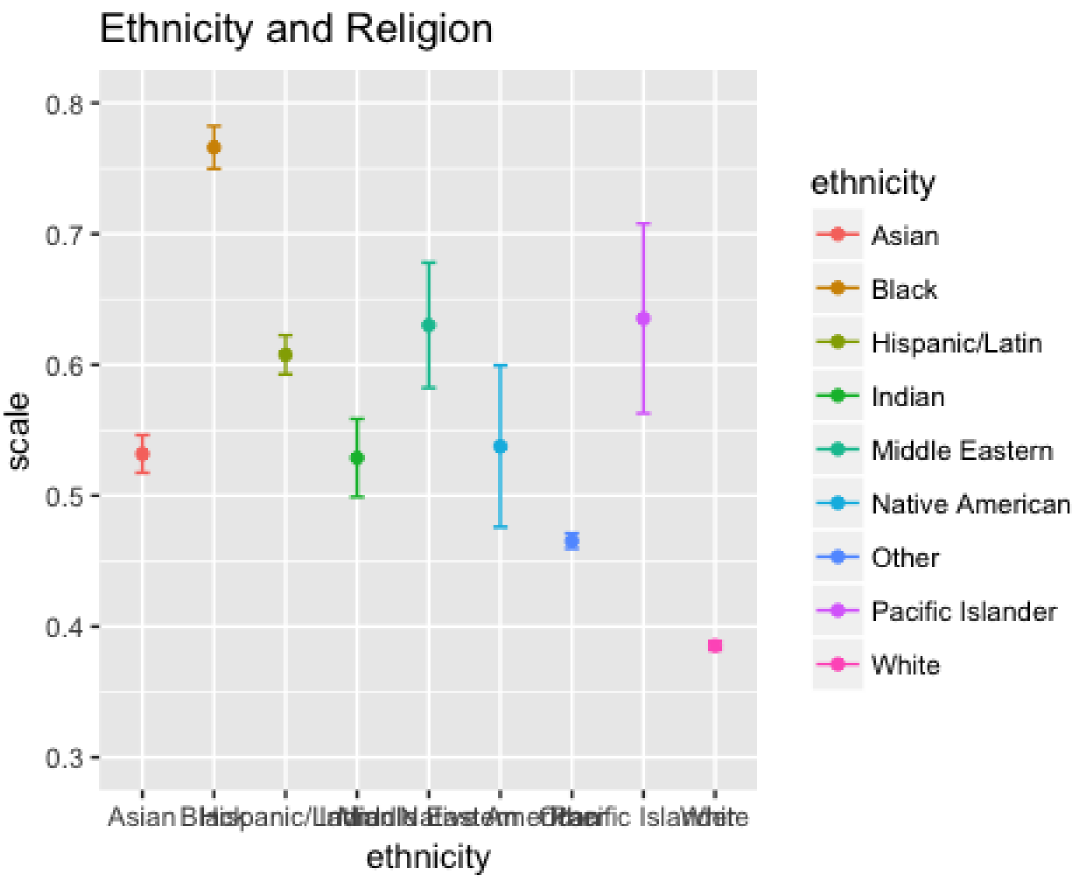

```
library(dplyr)


ME <- all %>%
  group_by(d_ethnicity) %>%
  summarise_at(vars(moral), funs(mean(., na.rm=TRUE),
                                      n=n(),
                                      sd(., na.rm = TRUE)
                                      ))
ME[,5] <- ME[,2] + (1.96 * ME [,4]/sqrt(ME[,3]))
ME[,6] <- ME[,2] - (1.96 * ME [,4]/sqrt(ME[,3]))
names(ME) <- c("ethnicity", "mean", "n", "sd", "ci.max", "ci.min")

plotmeme<-ME

plotmeme$label<- plotmeme[,1]

EthMo <- ggplot(plotmeme) +
# add error bars
        geom_errorbar(aes(x=ethnicity, ymin=ci.min,
                         ymax=ci.max,color=ethnicity),
                      width=.2, size=.5, na.rm=TRUE) +
# add means
```

```r
        geom_point(aes(ethnicity, mean, color=ethnicity),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="ethnicity",
             y="scale",
             title= "Ethnicity and Moral") +
# set limits to y axis

            scale_y_continuous(limits=c(.6,.8))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

EthMo # display plot
```



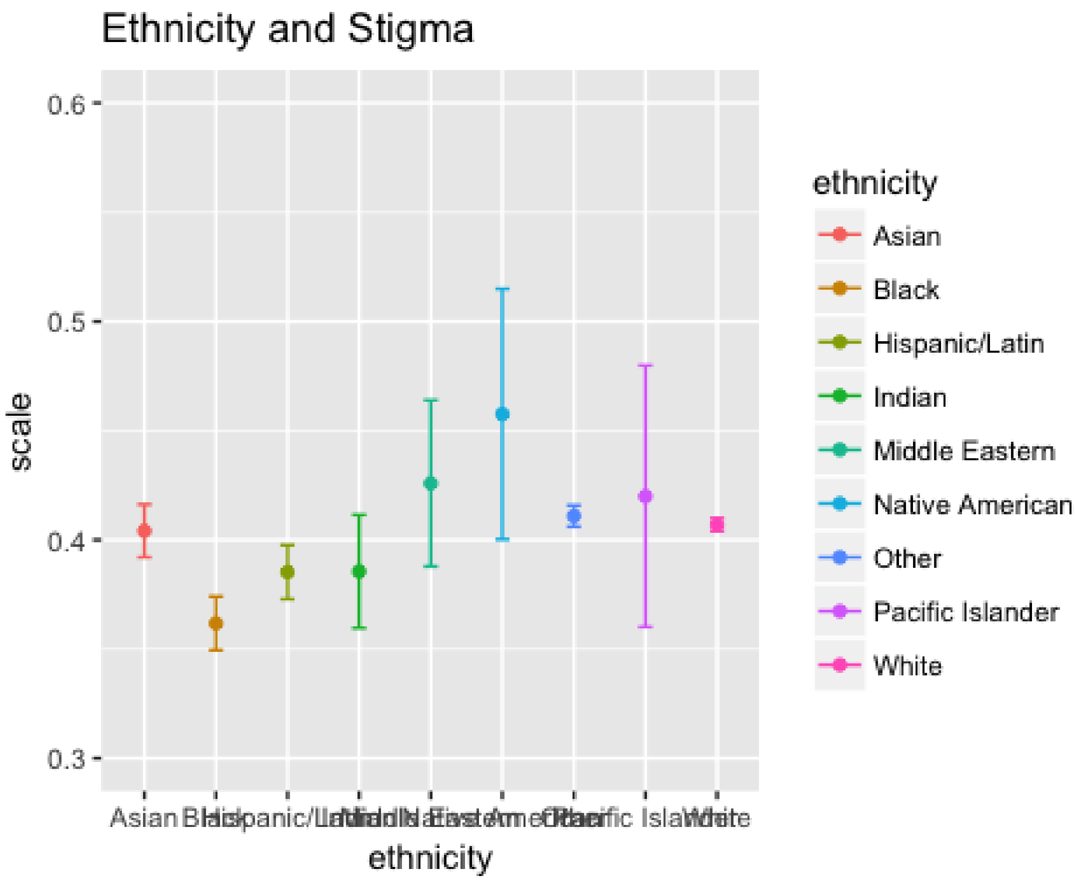

```r
library(dplyr)


RE <- all %>%
  group_by(d_ethnicity) %>%
  summarise_at(vars(religion), funs(mean(., na.rm=TRUE),
                                         n=n(),
```

```
                                             sd(., na.rm = TRUE)
                                             ))
RE[,5] <- RE[,2] + (1.96 * RE [,4]/sqrt(RE[,3]))
RE[,6] <- RE[,2] - (1.96 * RE [,4]/sqrt(RE[,3]))
names(RE) <- c("ethnicity", "mean", "n", "sd", "ci.max", "ci.min")

plotmere<-RE

plotmere$label<- plotmere[,1]

EthRel <- ggplot(plotmere) +
# add error bars
        geom_errorbar(aes(x=ethnicity, ymin=ci.min,
                         ymax=ci.max,color=ethnicity),
                      width=.2, size=.5, na.rm=TRUE) +
# add means
        geom_point(aes(ethnicity, mean, color=ethnicity),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="ethnicity",
             y="scale",
             title= "Ethnicity and Religion") +
# set limits to y axis

            scale_y_continuous(limits=c(.3,.8))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

EthRel # display plot
```



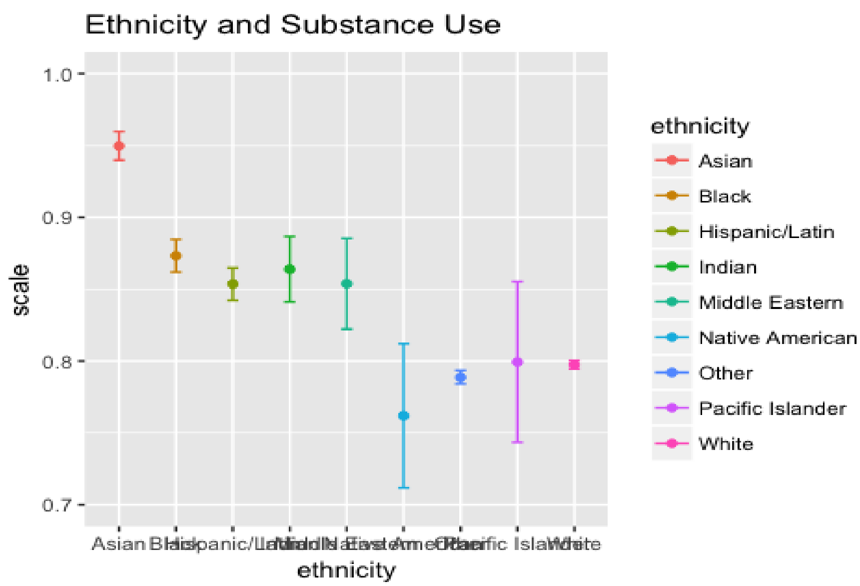

```
library(dplyr)


SE <- all %>%
  group_by(d_ethnicity) %>%
  summarise_at(vars(stigma), funs(mean(., na.rm=TRUE),
                                            n=n(),
                                            sd(., na.rm = TRUE)
                                            ))
SE[,5] <- SE[,2] + (1.96 * SE [,4]/sqrt(SE[,3]))
SE[,6] <- SE[,2] - (1.96 * SE [,4]/sqrt(SE[,3]))
names(SE) <- c("ethnicity", "mean", "n", "sd", "ci.max", "ci.min")

plotmese<-SE

plotmese$label<- plotmese[,1]

EthStig <- ggplot(plotmese) +
# add error bars
        geom_errorbar(aes(x=ethnicity, ymin=ci.min,
                         ymax=ci.max,color=ethnicity),
                      width=.2, size=.5, na.rm=TRUE) +
# add means
```

```
        geom_point(aes(ethnicity, mean, color=ethnicity),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="ethnicity",
             y="scale",
             title= "Ethnicity and Stigma") +
# set limits to y axis

            scale_y_continuous(limits=c(.3,.6))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

EthStig # display plot
```




```
library(dplyr)


SUE <- all %>%
  group_by(d_ethnicity) %>%
  summarise_at(vars(substanceUse), funs(mean(., na.rm=TRUE),
                                        n=n(),
```

```r
                                        sd(., na.rm = TRUE)
                                        ))
SUE[,5] <- SUE[,2] + (1.96 * SUE [,4]/sqrt(SUE[,3]))
SUE[,6] <- SUE[,2] - (1.96 * SUE [,4]/sqrt(SUE[,3]))
names(SUE) <- c("ethnicity", "mean", "n", "sd", "ci.max", "ci.min")

plotmesue<-SUE

plotmesue$label<- plotmesue[,1]

EthSU <- ggplot(plotmesue) +
# add error bars
        geom_errorbar(aes(x=ethnicity, ymin=ci.min,
                          ymax=ci.max,color=ethnicity),
                       width=.2, size=.5, na.rm=TRUE) +
# add means
        geom_point(aes(ethnicity, mean, color=ethnicity),
                   size=1.5, na.rm=TRUE) +
# add axis titles
        labs(x="ethnicity",
             y="scale",
             title= "Ethnicity and Substance Use") +
# set limits to y axis

            scale_y_continuous(limits=c(.7,1))
# add solid line at z=0
       #geom_hline(aes(yintercept=0)) #+
# make into multi-panel plot
       #facet_wrap(~ item, nrow = 2)

EthSU # display plot
```



Appendix G. Grouped Data Visualizations

Data Visualizations: Grouped

Sam DiPiero

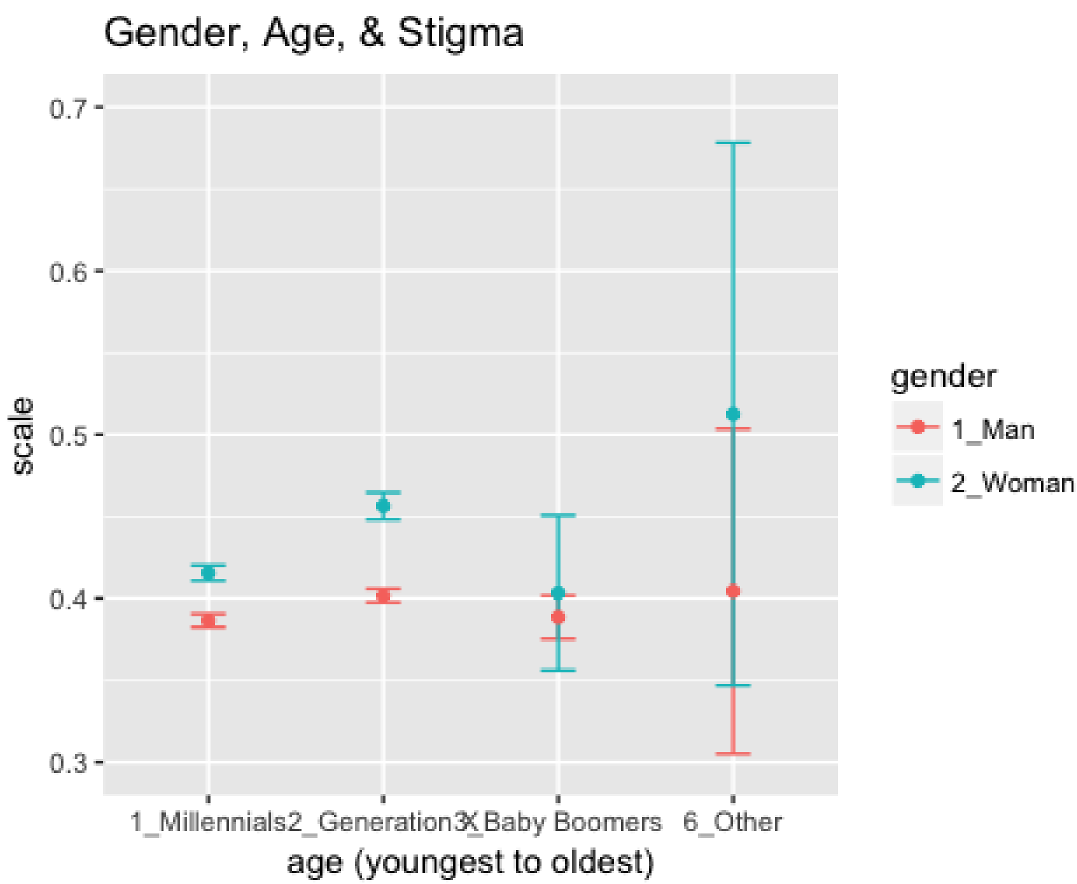
December 13, 2017

install.packages("car") library(tidyverse)

```
knitr::opts_chunk$set(echo = TRUE)
library(psych)

## Warning: package 'psych' was built under R version 3.4.4

library(knitr)
library(papeR)

## Loading required package: car

## Warning: package 'car' was built under R version 3.4.4

## Loading required package: carData

## Warning: package 'carData' was built under R version 3.4.4

##
## Attaching package: 'car'

## The following object is masked from 'package:psych':
##
##     logit

## Loading required package: xtable

##
## Attaching package: 'papeR'

## The following object is masked from 'package:utils':
##
##     toLatex

suppressMessages(library(tidyverse))
library(car)
library(dplyr)
library(sjPlot)

## Warning in checkMatrixPackageVersion(): Package version inconsistency dete
cted.
## TMB was built with Matrix version 1.2.12
## Current Matrix version is 1.2.14
```

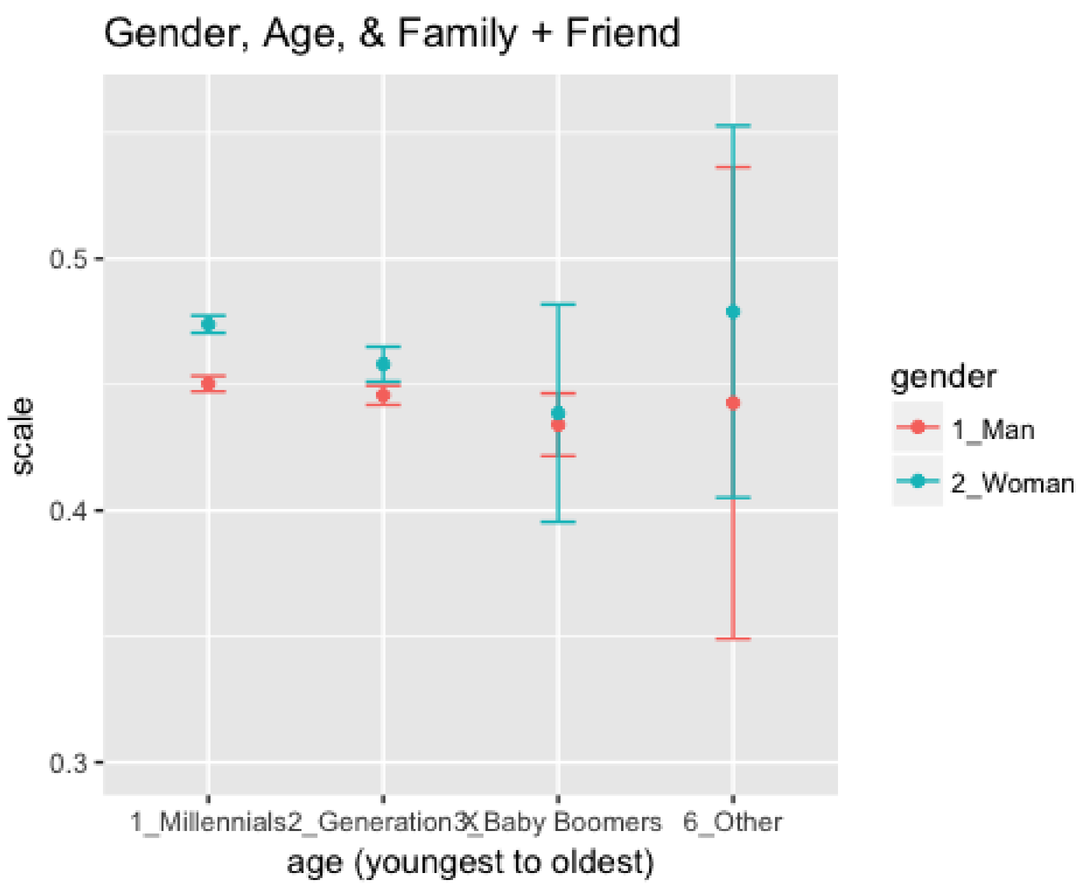
```
## Please re-install 'TMB' from source using install.packages('TMB', type = '
source') or ask CRAN for a binary version of 'TMB' matching CRAN's 'Matrix' p
ackage

## #refugeeswelcome

library(GPArotation)
library(ggthemes)

## Warning: package 'ggthemes' was built under R version 3.4.4

allSSscores<-read.csv('samScalescores1.csv')
demo <- read.csv('demographicData1.csv')

all <-cbind(allSSscores,demo)

library(dplyr)


b <- all %>%
  group_by(d_gender, d_age) %>%
  summarise_at(vars(stigma), funs(mean(., na.rm=TRUE),
                                   n=n(),
                                   sd(., na.rm = TRUE)
                                   ))

b[,6] <- b[,3] + (1.96 * b [,5]/sqrt(b[,4]))
b[,7] <- b[,3] - (1.96 * b [,5]/sqrt(b[,4]))
names(b) <- c("gender", "age", "mean", "n", "sd", "ci.max", "ci.min")

plotme <- b
#plotme <-as.data.frame (b)
#attach(plotme)
plotme<-plotme[which (plotme$gender != "Other"), ]

a <- ggplot(plotme) +
# add error bars
        geom_errorbar(aes(x=age, ymin=ci.min,
                          ymax=ci.max, color= gender),
                      width=.2, size=.5, na.rm=TRUE)  +
# add means
        geom_point(aes(age, mean, color= gender),
                   size=1.5, na.rm=TRUE) +

# add axis titles
        labs(x="age (youngest to oldest)",
             y="scale",
             title= "Gender, Age, & Stigma") +
# set limits to y axis

            scale_y_continuous(limits=c(.3, .7))
```

```
a # display plot
```



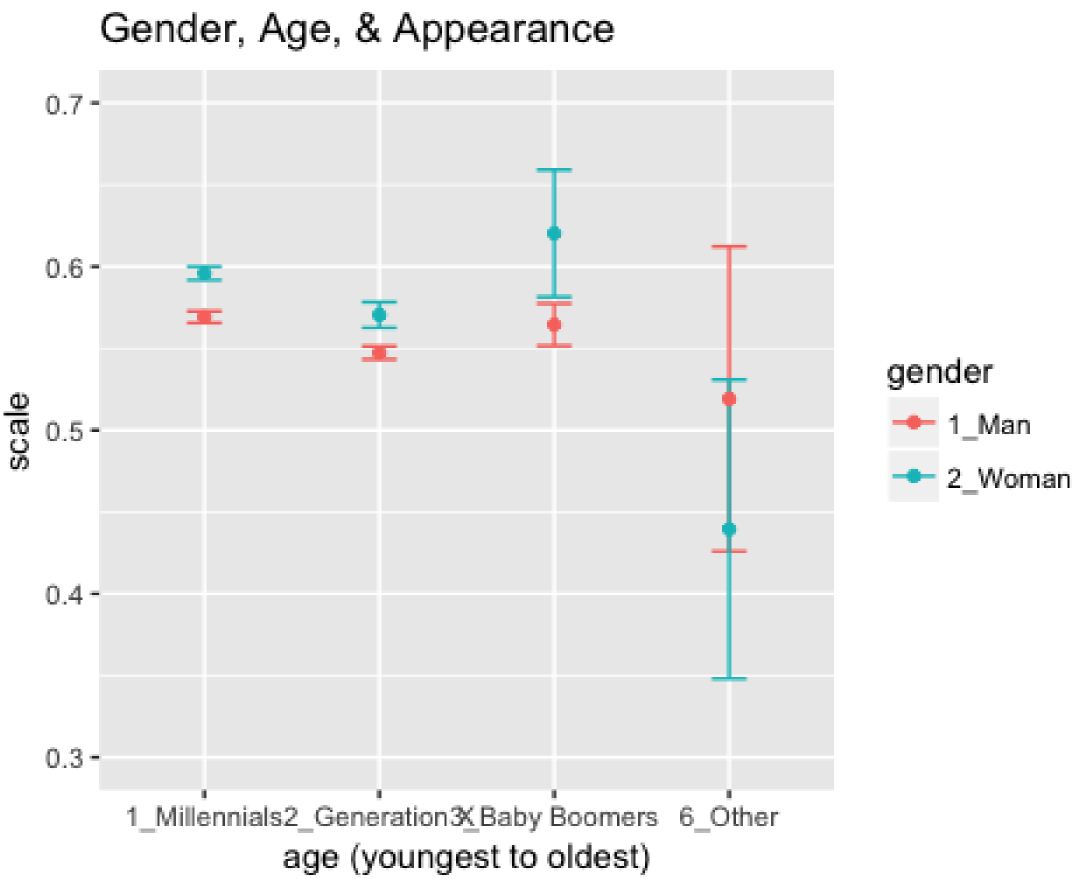

```
library(dplyr)


GAFF <- all %>%
  group_by(d_gender, d_age) %>%
  summarise_at(vars(familyandfriend), funs(mean(., na.rm=TRUE),
                                            n=n(),
                                            sd(., na.rm = TRUE)
                                            ))
GAFF[,6] <- GAFF[,3] + (1.96 * GAFF [,5]/sqrt(GAFF[,4]))
GAFF[,7] <- GAFF[,3] - (1.96 * GAFF [,5]/sqrt(GAFF[,4]))
names(GAFF) <- c("gender", "age", "mean", "n", "sd", "ci.max", "ci.min")

plotmegaff<-GAFF
plotmegaff<-plotmegaff[which (plotmegaff$gender != "Other"), ]
plotmegaff$label<- plotmegaff[1&2]

gaf <- ggplot(plotmegaff) +
# add error bars
        geom_errorbar(aes(x=age, ymin=ci.min,
                          ymax=ci.max, color= gender),
```

```
                              width=.2, size=.5, na.rm=TRUE)  +
# add means
        geom_point(aes(age, mean, color= gender),
                         size=1.5, na.rm=TRUE) +

# add axis titles
        labs(x="age (youngest to oldest)",
             y="scale",
             title= "Gender, Age, & Family + Friend") +
# set limits to y axis

            scale_y_continuous(limits=c(.3, .56))

gaf # display plot
```



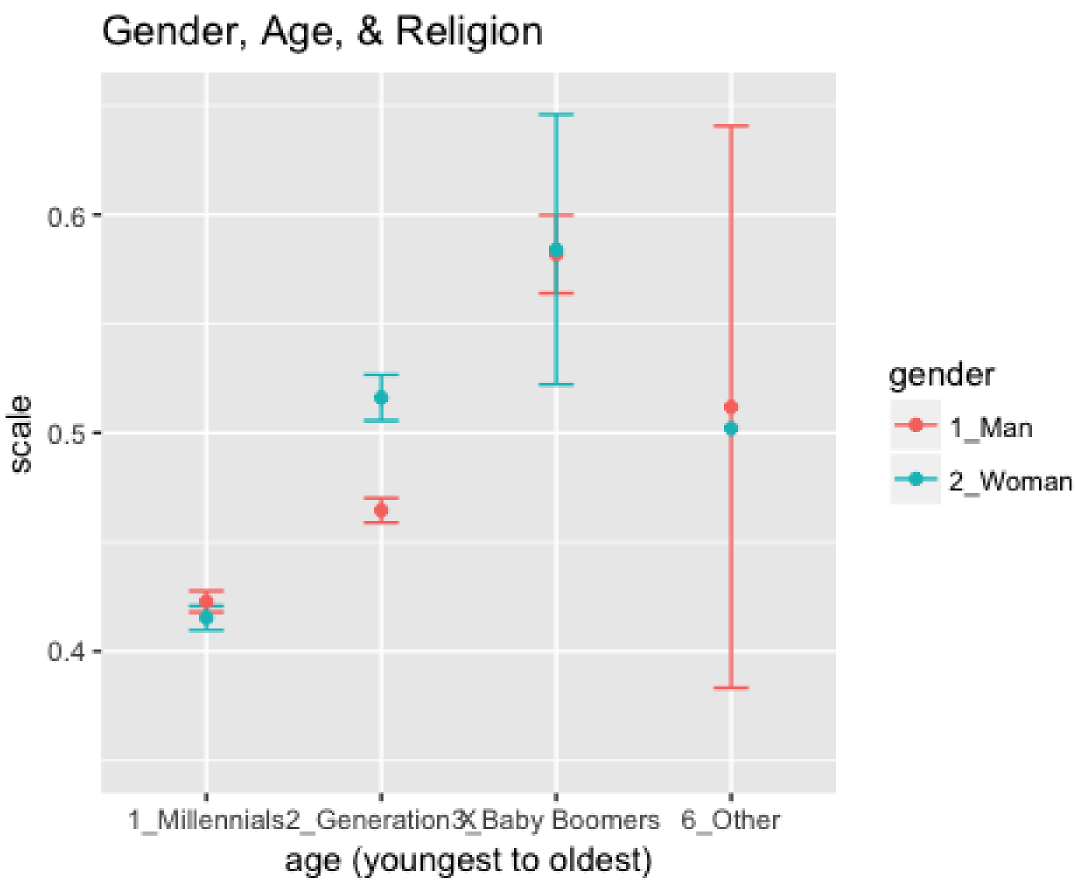

```
library(dplyr)


GAA <- all %>%
  group_by(d_gender, d_age) %>%
  summarise_at(vars(appearance), funs(mean(., na.rm=TRUE),
                                        n=n(),
                                        sd(., na.rm = TRUE)
```

```
                                          ))
GAA[,6] <- GAA[,3] + (1.96 * GAA [,5]/sqrt(GAA[,4]))
GAA[,7] <- GAA[,3] - (1.96 * GAA [,5]/sqrt(GAA[,4]))
names(GAA) <- c("gender", "age", "mean", "n", "sd", "ci.max", "ci.min")

plotmegaa<-GAA
plotmegaa<-plotmegaa[which (plotmegaa$gender != "Other"), ]
plotmegaa$label<- plotmegaa[1&2]

gaap <- ggplot(plotmegaa) +
# add error bars
        geom_errorbar(aes(x=age, ymin=ci.min,
                         ymax=ci.max, color= gender),
                      width=.2, size=.5, na.rm=TRUE)  +
# add means
        geom_point(aes(age, mean, color= gender),
                   size=1.5, na.rm=TRUE) +

# add axis titles
        labs(x="age (youngest to oldest)",
             y="scale",
             title= "Gender, Age, & Appearance") +
# set limits to y axis

            scale_y_continuous(limits=c(.3, .7))

gaap # display plot
```



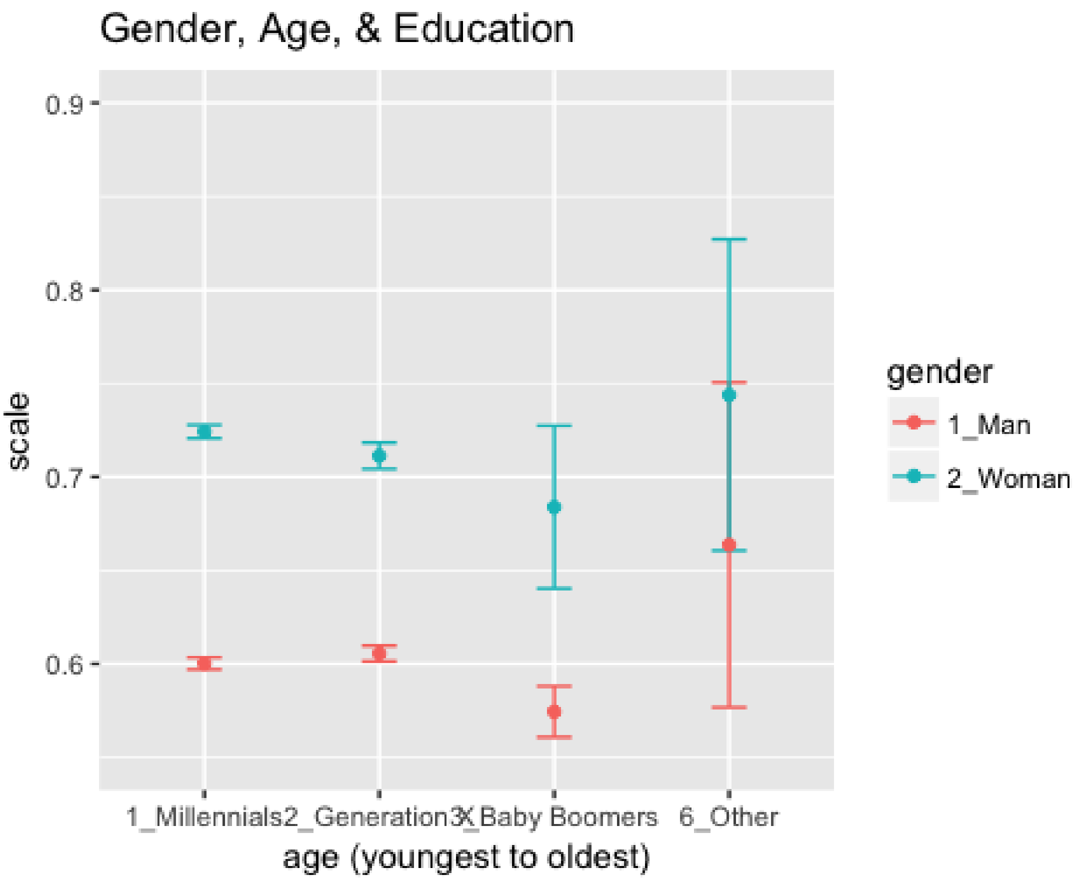

```
library(dplyr)


GAR <- all %>%
  group_by(d_gender, d_age) %>%
  summarise_at(vars(religion), funs(mean(., na.rm=TRUE),
                                         n=n(),
                                         sd(., na.rm = TRUE)
                                         ))
GAR[,6] <- GAR[,3] + (1.96 * GAR [,5]/sqrt(GAR[,4]))
GAR[,7] <- GAR[,3] - (1.96 * GAR [,5]/sqrt(GAR[,4]))
names(GAR) <- c("gender", "age", "mean", "n", "sd", "ci.max", "ci.min")

plotmegar<-GAR
plotmegar<-plotmegar[which (plotmegar$gender != "Other"), ]
plotmegar$label<- plotmegar[1&2]

gar <- ggplot(plotmegar) +
# add error bars
        geom_errorbar(aes(x=age, ymin=ci.min,
                         ymax=ci.max, color= gender),
                      width=.2, size=.5, na.rm=TRUE)  +
# add means
```

```
        geom_point(aes(age, mean, color= gender),
                   size=1.5, na.rm=TRUE) +

# add axis titles
        labs(x="age (youngest to oldest)",
             y="scale",
             title= "Gender, Age, & Religion") +
# set limits to y axis

            scale_y_continuous(limits=c(.35, .65))

gar # display plot
```




```
library(dplyr)


GAE <- all %>%
  group_by(d_gender, d_age) %>%
  summarise_at(vars(education), funs(mean(., na.rm=TRUE),
                                        n=n(),
                                        sd(., na.rm = TRUE)
                                        ))
GAE[,6] <- GAE[,3] + (1.96 * GAE [,5]/sqrt(GAE[,4]))
```

```
GAE[,7] <- GAE[,3] - (1.96 * GAE [,5]/sqrt(GAE[,4]))
names(GAE) <- c("gender", "age", "mean", "n", "sd", "ci.max", "ci.min")

plotmegae<-GAE
plotmegae<-plotmegae[which (plotmegae$gender != "Other"), ]
plotmegae$label<- plotmegae[1&2]

gae <- ggplot(plotmegae) +
# add error bars
        geom_errorbar(aes(x=age, ymin=ci.min,
                         ymax=ci.max, color= gender),
                      width=.2, size=.5, na.rm=TRUE)  +
# add means
        geom_point(aes(age, mean, color= gender),
                   size=1.5, na.rm=TRUE) +

# add axis titles
        labs(x="age (youngest to oldest)",
             y="scale",
             title= "Gender, Age, & Education") +
# set limits to y axis

            scale_y_continuous(limits=c( .55, .9))

gae # display plot
```

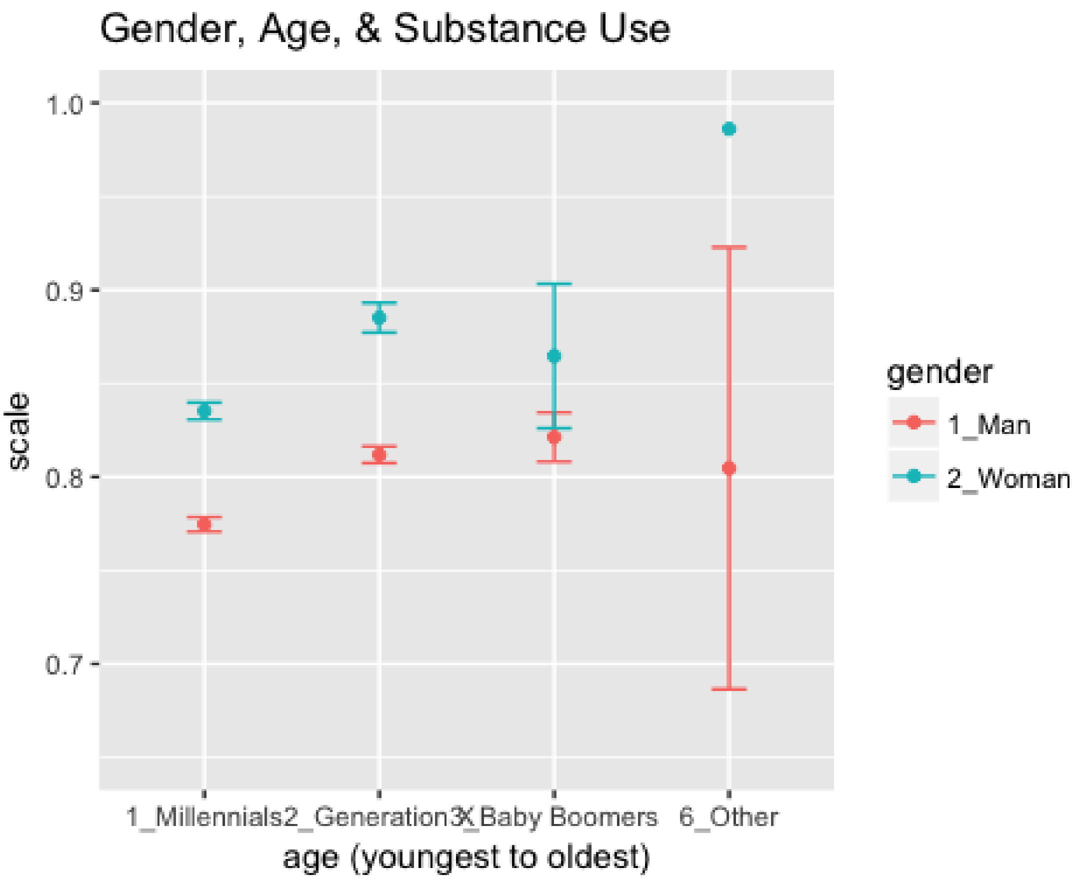



```
library(dplyr)


GAM <- all %>%
  group_by(d_gender, d_age) %>%
  summarise_at(vars(moral), funs(mean(., na.rm=TRUE),
                                         n=n(),
                                         sd(., na.rm = TRUE)
                                         ))
GAM[,6] <- GAM[,3] + (1.96 * GAM [,5]/sqrt(GAM[,4]))
GAM[,7] <- GAM[,3] - (1.96 * GAM [,5]/sqrt(GAM[,4]))
names(GAM) <- c("gender", "age", "mean", "n", "sd", "ci.max", "ci.min")

plotmegam<-GAM
plotmegam<-plotmegam[which (plotmegam$gender != "Other"), ]
plotmegam$label<- plotmegam[1&2]

gam <- ggplot(plotmegam) +
# add error bars
        geom_errorbar(aes(x=age, ymin=ci.min,
                          ymax=ci.max, color= gender),
                       width=.2, size=.5, na.rm=TRUE)  +
# add means
```

```
        geom_point(aes(age, mean, color= gender),
                   size=1.5, na.rm=TRUE) +

# add axis titles
        labs(x="age (youngest to oldest)",
             y="scale",
             title= "Gender, Age, & Moral") +
# set limits to y axis

            scale_y_continuous(limits=c( .65, 1))

gam # display plot
```

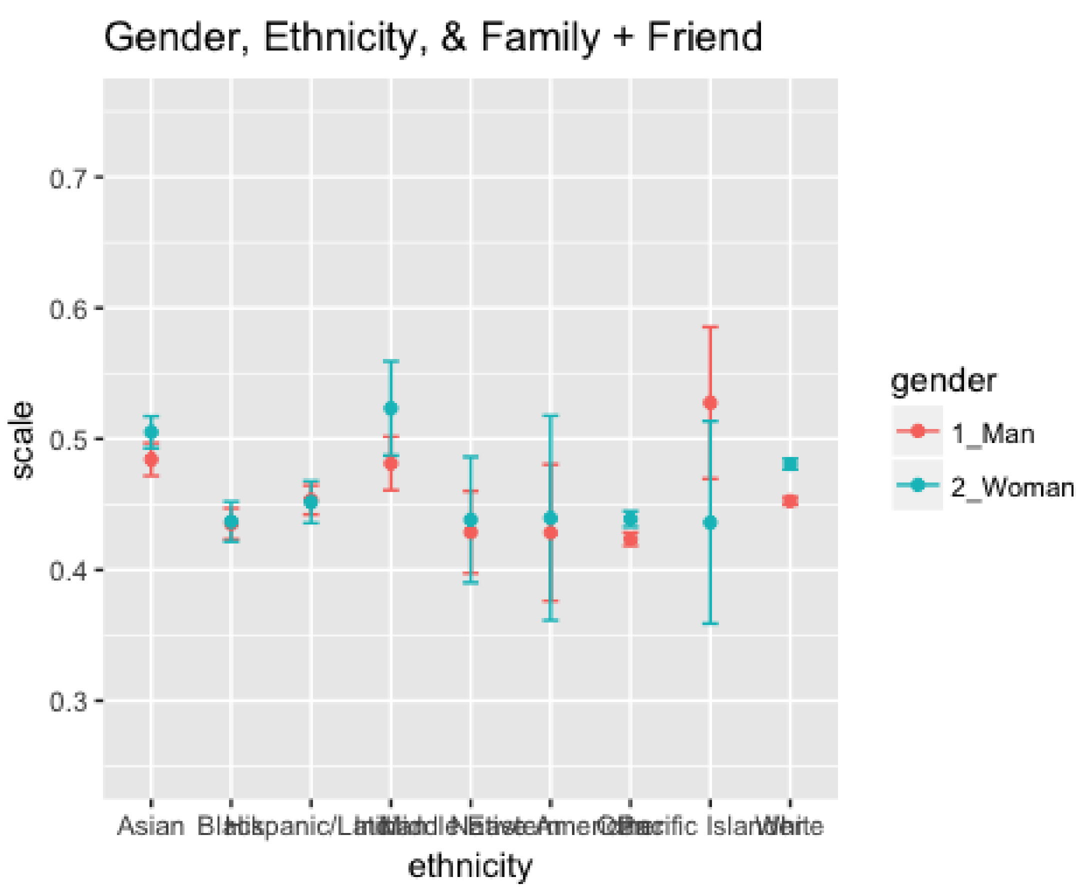

```
library(dplyr)


GASU <- all %>%
  group_by(d_gender, d_age) %>%
  summarise_at(vars(substanceUse), funs(mean(., na.rm=TRUE),
                                          n=n(),
                                          sd(., na.rm = TRUE)
                                          ))
GASU[,6] <- GASU[,3] + (1.96 * GASU [,5]/sqrt(GASU[,4]))
```

```
GASU[,7] <- GASU[,3] - (1.96 * GASU [,5]/sqrt(GASU[,4]))
names(GASU) <- c("gender", "age", "mean", "n", "sd", "ci.max", "ci.min")

plotmegasu<-GASU
plotmegasu<-plotmegasu[which (plotmegasu$gender != "Other"), ]
plotmegasu$label<- plotmegasu[1&2]

gasu <- ggplot(plotmegasu) +
# add error bars
        geom_errorbar(aes(x=age, ymin=ci.min,
                         ymax=ci.max, color= gender),
                      width=.2, size=.5, na.rm=TRUE)  +
# add means
        geom_point(aes(age, mean, color= gender),
                   size=1.5, na.rm=TRUE) +

# add axis titles
        labs(x="age (youngest to oldest)",
             y="scale",
             title= "Gender, Age, & Substance Use") +
# set limits to y axis

            scale_y_continuous(limits=c( .65, 1))

gasu # display plot
```



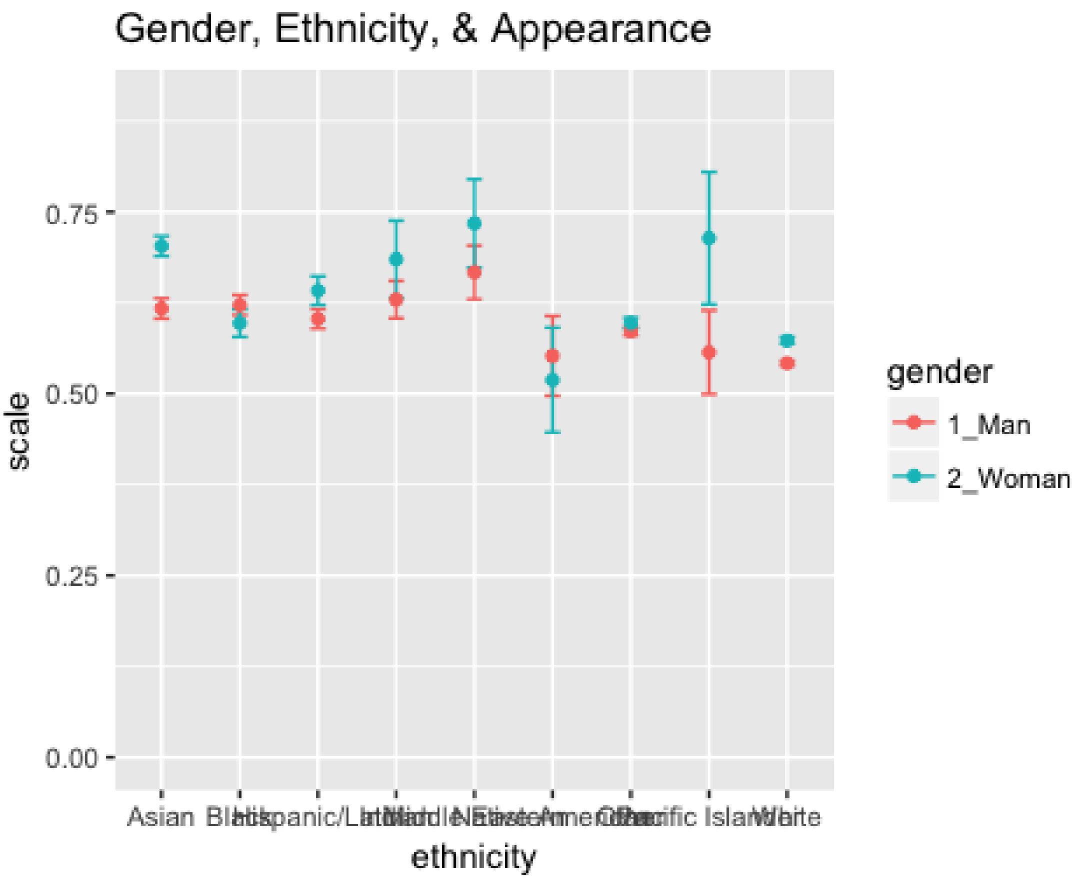

```
library(dplyr)


GEFF <- all %>%
  group_by(d_gender, d_ethnicity) %>%
  summarise_at(vars(familyandfriend), funs(mean(., na.rm=TRUE),
                                              n=n(),
                                              sd(., na.rm = TRUE)
                                              ))
GEFF[,6] <- GEFF[,3] + (1.96 * GEFF [,5]/sqrt(GEFF[,4]))
GEFF[,7] <- GEFF[,3] - (1.96 * GEFF [,5]/sqrt(GEFF[,4]))
names(GEFF) <- c("gender", "ethnicity", "mean", "n", "sd", "ci.max", "ci.min"
)

plotmegeff<-GEFF
plotmegeff<-plotmegeff[which (plotmegeff$gender != "Other"), ]
plotmegeff$label<- plotmegeff[1&2]

gef <- ggplot(plotmegeff) +
# add error bars
        geom_errorbar(aes(x=ethnicity, ymin=ci.min,
                          ymax=ci.max, color= gender),
                      width=.2, size=.5, na.rm=TRUE)  +
```

```
# add means
        geom_point(aes(ethnicity, mean, color= gender),
                   size=1.5, na.rm=TRUE) +

# add axis titles
        labs(x="ethnicity",
             y="scale",
             title= "Gender, Ethnicity, & Family + Friend") +
# set limits to y axis

            scale_y_continuous(limits=c(.25, .75))

gef # display plot
```



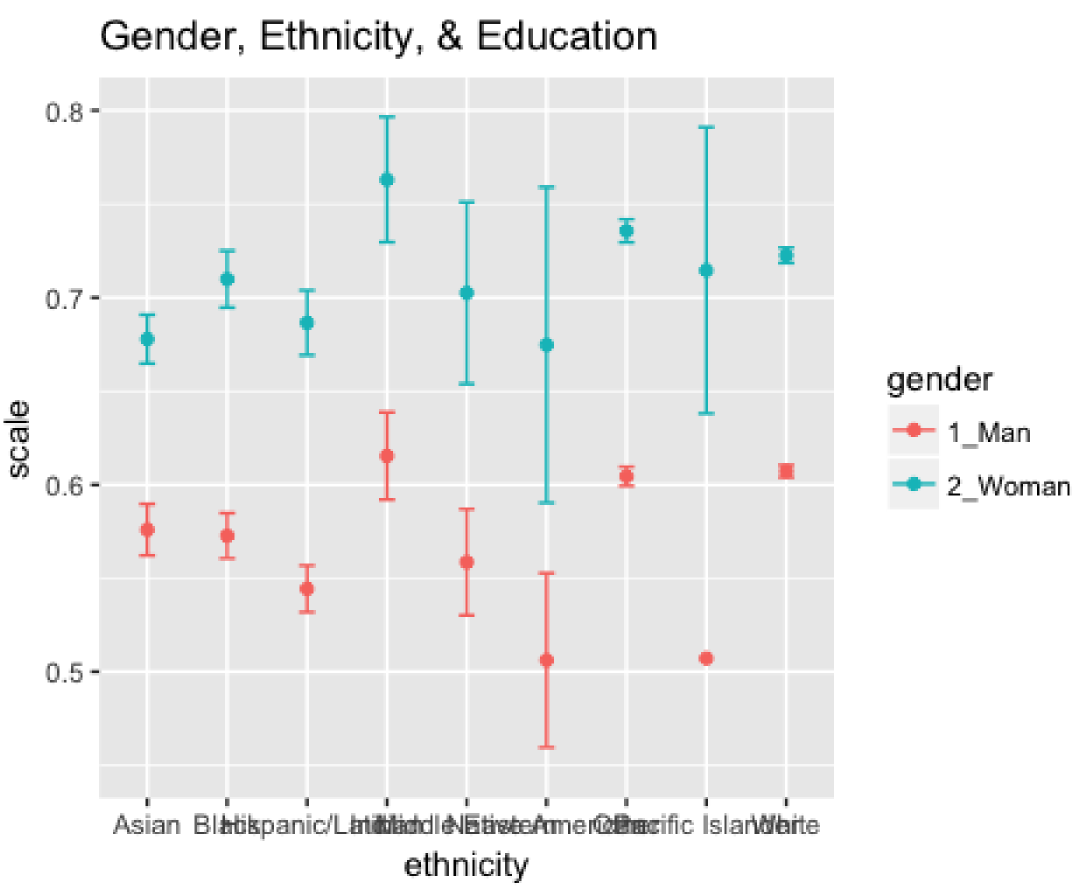

```
library(dplyr)


GEA <- all %>%
  group_by(d_gender, d_ethnicity) %>%
  summarise_at(vars(appearance), funs(mean(., na.rm=TRUE),
                                      n=n(),
                                      sd(., na.rm = TRUE)
                                      ))
```

```r
GEA[,6] <- GEA[,3] + (1.96 * GEA [,5]/sqrt(GEA[,4]))
GEA[,7] <- GEA[,3] - (1.96 * GEA [,5]/sqrt(GEA[,4]))
names(GEA) <- c("gender", "ethnicity", "mean", "n", "sd", "ci.max", "ci.min")

plotmegea<-GEA
plotmegea<-plotmegea[which (plotmegea$gender != "Other"), ]
plotmegea$label<- plotmegea[1&2]

geap <- ggplot(plotmegea) +
# add error bars
        geom_errorbar(aes(x=ethnicity, ymin=ci.min,
                         ymax=ci.max, color= gender),
                      width=.2, size=.5, na.rm=TRUE)  +
# add means
        geom_point(aes(ethnicity, mean, color= gender),
                   size=1.5, na.rm=TRUE) +

# add axis titles
        labs(x="ethnicity",
             y="scale",
             title= "Gender, Ethnicity, & Appearance") +
# set limits to y axis

            scale_y_continuous(limits=c(0, .9))

geap # display plot
```



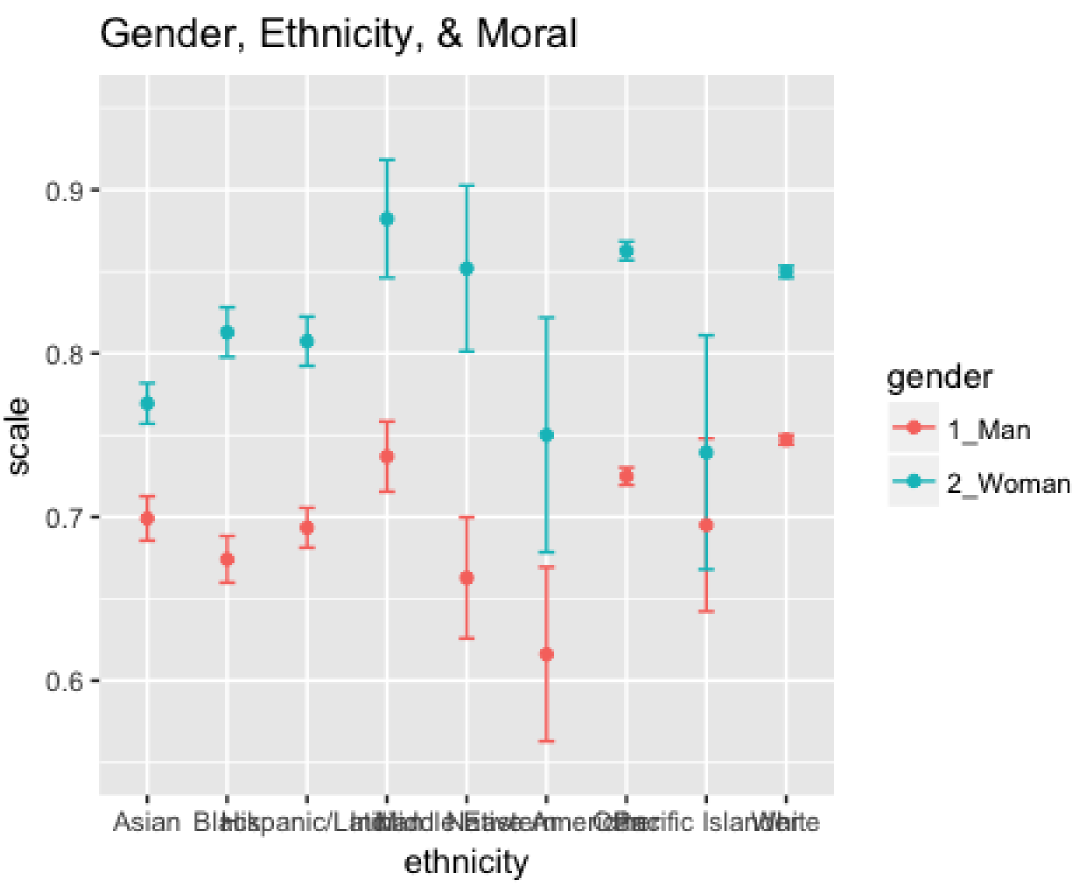

```
library(dplyr)


GEE <- all %>%
  group_by(d_gender, d_ethnicity) %>%
  summarise_at(vars(education), funs(mean(., na.rm=TRUE),
                                    n=n(),
                                    sd(., na.rm = TRUE)
                                    ))
GEE[,6] <- GEE[,3] + (1.96 * GEE [,5]/sqrt(GEE[,4]))
GEE[,7] <- GEE[,3] - (1.96 * GEE [,5]/sqrt(GEE[,4]))
names(GEE) <- c("gender", "ethnicity", "mean", "n", "sd", "ci.max", "ci.min")

plotmegee<-GEE
plotmegee<-plotmegee[which (plotmegee$gender != "Other"), ]
plotmegee$label<- plotmegee[1&2]

gee <- ggplot(plotmegee) +
# add error bars
        geom_errorbar(aes(x=ethnicity, ymin=ci.min,
                          ymax=ci.max, color= gender),
                       width=.2, size=.5, na.rm=TRUE)  +
# add means
```

```
        geom_point(aes(ethnicity, mean, color= gender),
                   size=1.5, na.rm=TRUE) +

# add axis titles
        labs(x="ethnicity",
             y="scale",
             title= "Gender, Ethnicity, & Education") +
# set limits to y axis

            scale_y_continuous(limits=c( .45, .8))

gee # display plot
```



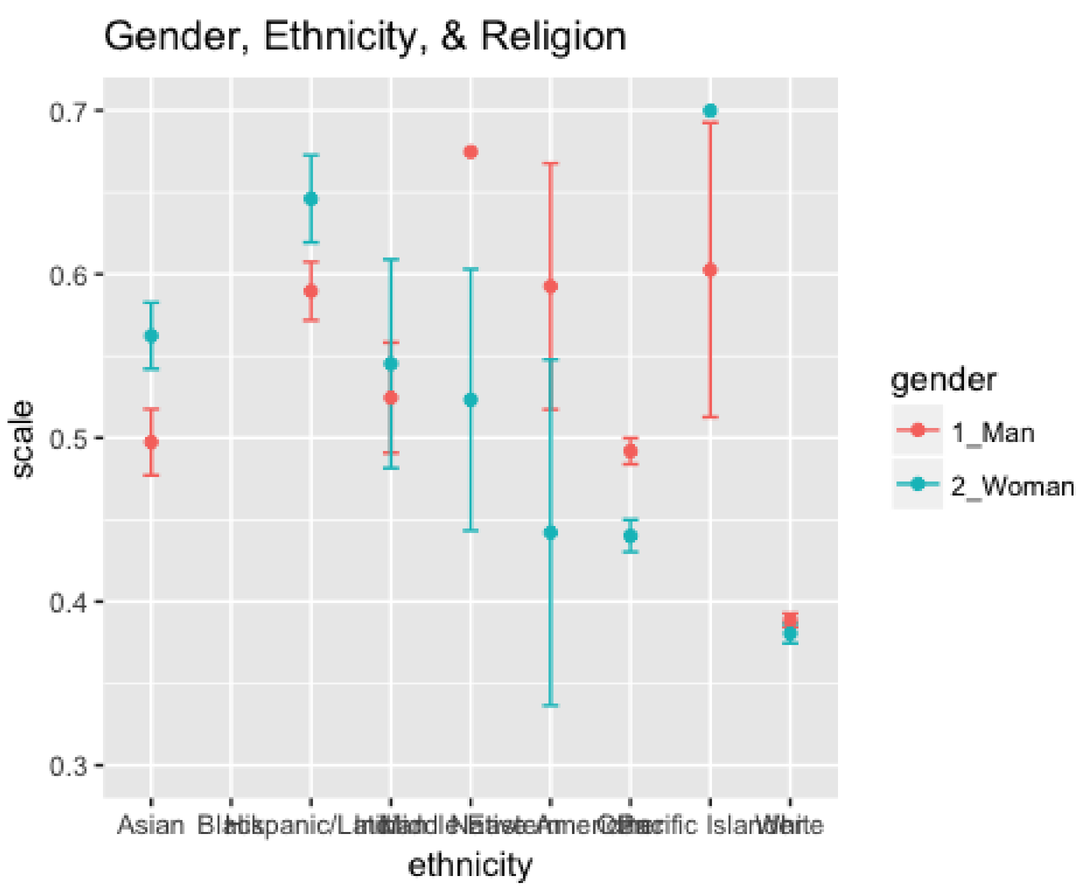

```
library(dplyr)


GEM <- all %>%
  group_by(d_gender, d_ethnicity) %>%
  summarise_at(vars(moral), funs(mean(., na.rm=TRUE),
                                          n=n(),
                                          sd(., na.rm = TRUE)
                                          ))
GEM[,6] <- GEM[,3] + (1.96 * GEM [,5]/sqrt(GEM[,4]))
```

```r
GEM[,7] <- GEM[,3] - (1.96 * GEM [,5]/sqrt(GEM[,4]))
names(GEM) <- c("gender", "ethnicity", "mean", "n", "sd", "ci.max", "ci.min")

plotmegem<-GEM
plotmegem<-plotmegem[which (plotmegem$gender != "Other"), ]
plotmegem$label<- plotmegem[1&2]

gem <- ggplot(plotmegem) +
# add error bars
        geom_errorbar(aes(x=ethnicity, ymin=ci.min,
                         ymax=ci.max, color= gender),
                      width=.2, size=.5, na.rm=TRUE)  +
# add means
        geom_point(aes(ethnicity, mean, color= gender),
                   size=1.5, na.rm=TRUE) +

# add axis titles
        labs(x="ethnicity",
             y="scale",
             title= "Gender, Ethnicity, & Moral") +
# set limits to y axis

            scale_y_continuous(limits=c( .55, .95))

gem # display plot
```



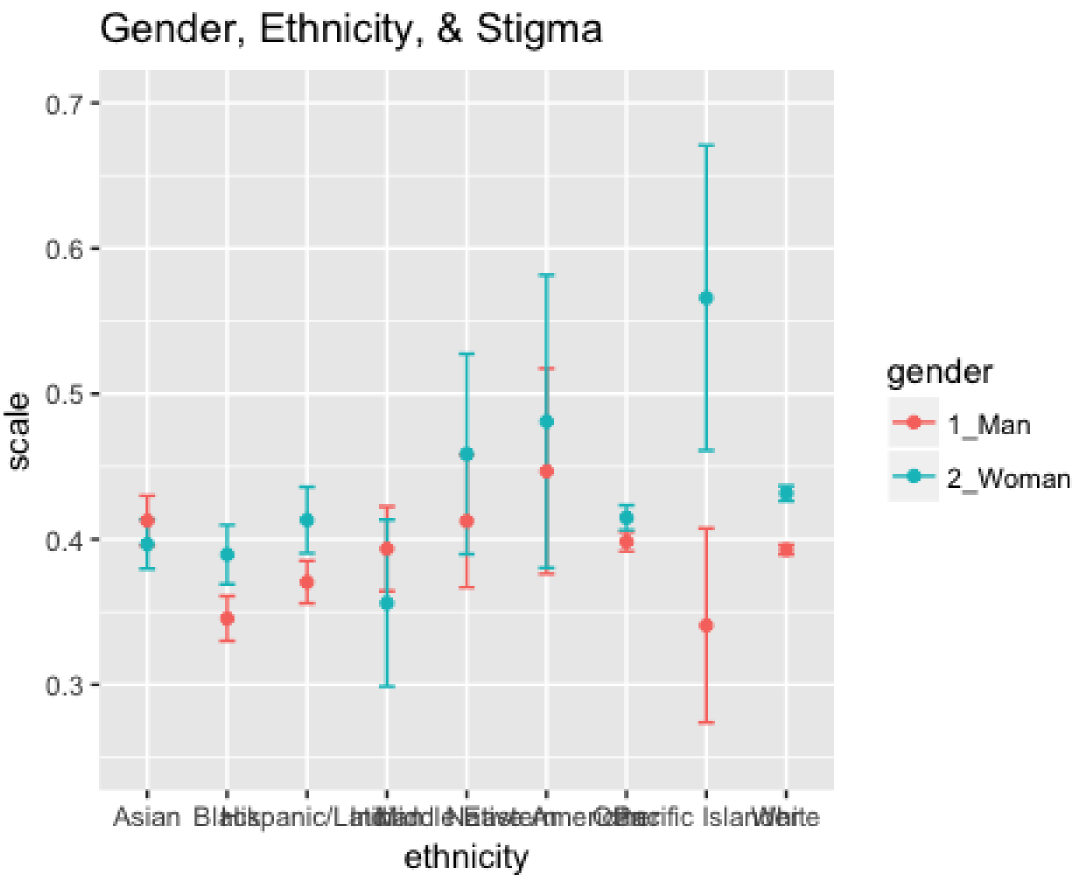

```
library(dplyr)


GER <- all %>%
  group_by(d_gender, d_ethnicity) %>%
  summarise_at(vars(religion), funs(mean(., na.rm=TRUE),
                                     n=n(),
                                     sd(., na.rm = TRUE)
                                     ))
GER[,6] <- GER[,3] + (1.96 * GER [,5]/sqrt(GER[,4]))
GER[,7] <- GER[,3] - (1.96 * GER [,5]/sqrt(GER[,4]))
names(GER) <- c("gender", "ethnicity", "mean", "n", "sd", "ci.max", "ci.min")

plotmeger<-GER
plotmeger<-plotmeger[which (plotmeger$gender != "Other"), ]
plotmeger$label<- plotmeger[1&2]

ger <- ggplot(plotmeger) +
# add error bars
        geom_errorbar(aes(x=ethnicity, ymin=ci.min,
                          ymax=ci.max, color= gender),
                       width=.2, size=.5, na.rm=TRUE)  +
# add means
```

```
        geom_point(aes(ethnicity, mean, color= gender),
                   size=1.5, na.rm=TRUE) +

# add axis titles
        labs(x="ethnicity",
             y="scale",
             title= "Gender, Ethnicity, & Religion") +
# set limits to y axis

            scale_y_continuous(limits=c(.3, .7))

ger # display plot
```




```
library(dplyr)


GES <- all %>%
  group_by(d_gender, d_ethnicity) %>%
  summarise_at(vars(stigma), funs(mean(., na.rm=TRUE),
                                          n=n(),
                                          sd(., na.rm = TRUE)
                                          ))
GES[,6] <- GES[,3] + (1.96 * GES [,5]/sqrt(GES[,4]))
```

```
GES[,7] <- GES[,3] - (1.96 * GES [,5]/sqrt(GES[,4]))
names(GES) <- c("gender", "ethnicity", "mean", "n", "sd", "ci.max", "ci.min")

plotmeges<-GES
plotmeges<-plotmeges[which (plotmeges$gender != "Other"), ]
plotmeges$label<- plotmeges[1&2]

ges<- ggplot(plotmeges) +
# add error bars
        geom_errorbar(aes(x=ethnicity, ymin=ci.min,
                         ymax=ci.max, color= gender),
                      width=.2, size=.5, na.rm=TRUE)  +
# add means
        geom_point(aes(ethnicity, mean, color= gender),
                   size=1.5, na.rm=TRUE) +

# add axis titles
        labs(x="ethnicity",
             y="scale",
             title= "Gender, Ethnicity, & Stigma") +
# set limits to y axis

            scale_y_continuous(limits=c(.25, .7))

ges # display plot
```




```
library(dplyr)


GESU <- all %>%
  group_by(d_gender, d_ethnicity) %>%
  summarise_at(vars(substanceUse), funs(mean(., na.rm=TRUE),
                                          n=n(),
                                          sd(., na.rm = TRUE)
                                          ))
GESU[,6] <- GESU[,3] + (1.96 * GESU [,5]/sqrt(GESU[,4]))
GESU[,7] <- GESU[,3] - (1.96 * GESU [,5]/sqrt(GESU[,4]))
names(GESU) <- c("gender", "ethnicity", "mean", "n", "sd", "ci.max", "ci.min"
)

plotmegesu<-GESU
plotmegesu<-plotmegesu[which (plotmegesu$gender != "Other"), ]
plotmegesu$label<- plotmegesu[1&2]

gesu <- ggplot(plotmegesu) +
# add error bars
        geom_errorbar(aes(x=ethnicity, ymin=ci.min,
                         ymax=ci.max, color= gender),
                      width=.2, size=.5, na.rm=TRUE)  +
```

```r
# add means
        geom_point(aes(ethnicity, mean, color= gender),
                   size=1.5, na.rm=TRUE) +

# add axis titles
        labs(x="ethnicity",
             y="scale",
             title= "Gender, Ethnicity, & Substance Use") +
# set limits to y axis

            scale_y_continuous(limits=c( .5, 1))

gesu # display plot
```