Project Milestone - 2: Student Qualities and Salary

Basic Info

Names: Carsten Schaufert (C61306769,carstes@clemson.edu)

Lindsey Myers (C76524020, lmyers3@g.clemson.edu)

Tucker Nelson (C14669125, tlnelso@clemson.edu)

Link to GitHub: Project

Background and Motivation

As seniors studying computer science, we are thinking about our futures more than ever and the next steps for us after graduation. A typical post graduation route is finding a full time job, and as we are active in our job searches we wonder "What qualities of a student are good predictors for career outcome?" In this project we explore this question with a focus on career outcomes in the technical industry by using salary as a metric for career outcome. We hope to find meaningful correlations between attributes in our dataset and salary. Our research should help us to better understand how to be successful in our field. We are especially drawn to this project because our field has a massive scope of job opportunities. In order to narrow down what jobs we want to pursue in the future, this project will show us what to expect in our field of jobs.

Project Objectives

There are many factors that can influence salary outcomes, our mission is to choose a few and observe their relationships to salary. We have broken our problem down into 3 categories of these factors: gender, academic, and personality.

A curiosity we have is about the relationship between gender and salary. We explore their relationship through two visualizations. One of which explores the average salaries of the two genders between various technology field specializations. The aim of this visualization is to observe if any of the genders are favored among the various specializations. The second visualization aims to convey if personalities affect salary differently among the genders. It would be beneficial to explore these questions in order to observe possible gender biases that may affect salary outcomes.

Academic history will naturally have a strong relationship with salary outcome as students who perform better academically are highly sought after by employers. Using AMCAT scores of individuals, we will measure the effect various section scores have on salary. We are aiming to define which sections are the strongest determinants of salary. The benefit of this is that we will be able to decide which sections covered by the AMCAT are the most significant to perform well in. Then, by these sections define which skills may be the most useful for ourselves to further develop for our job search. Another problem we explore is: how personality affects GPA. Personality has a high influence on our behaviors and it is important to understand which personality traits may encourage more studious behaviors among students. This understanding can guide us into altering our behaviors to be more studious by attempting to adopt personalities that encourage these behaviors.

Personality is a very interesting characteristic for us to look at as it defines who we are to ourselves and the world around us. The broadest problem we look at is how personality affects

salary outcomes. We wonder: "Which personality traits and what range is most advantageous in determining a successful career?" The benefit of exploring this problem is deciding whether our own personality traits may be putting us at a disadvantage in our job hunts. Additionally, we found it interesting to look at which personality traits may be advantageous by technology field specialization. This is to understand which specializations may be more appropriate for us based on our current personalities. If a person is more introverted, they may want to steer away from specializations that require the use of frequent communication skills.

Data

The data has been collected by a user on Kaggle and it contains the salary and various AMCAT scores of around 3000 Indian recent graduates.
Data Location: https://www.kaggle.com/manishkc06/engineering-graduate-salary-prediction

Data Processing

There is no real need for any data cleanup since most of the columns will be utilized. The mean and standard deviation will be found for the various personality scores of the overall data set and for each individual specialization. We will be using Javascript and the D3.js library to find these quantities and properly display them.

Visualization Design

The data will initially be displayed in various visualizations, allowing for the observer to make comparisons based on one of the three categories our problem is split into: academic, gender, and personality. Each of these categories will focus around a characteristic and its influence on salary.

At the end of this document are possible designs for our data visualizations and why we think they will work.

Must-Have Features

Salary, conscientiousness, agreeableness, extraversion, neuroticism, and openness to experience are all must-have features because of their importance in answering the overall question of "What qualities of a student are good predictors for career outcome?". Without these features we can not accurately measure their monetary success or their various personalities.

Optional Features

The degree, specialization, various academic AMCAT scores and gender features all help further enrich our data by providing some circumstantial information about the various people. This will also help us determine the reasons for the potential outliers being outliers.
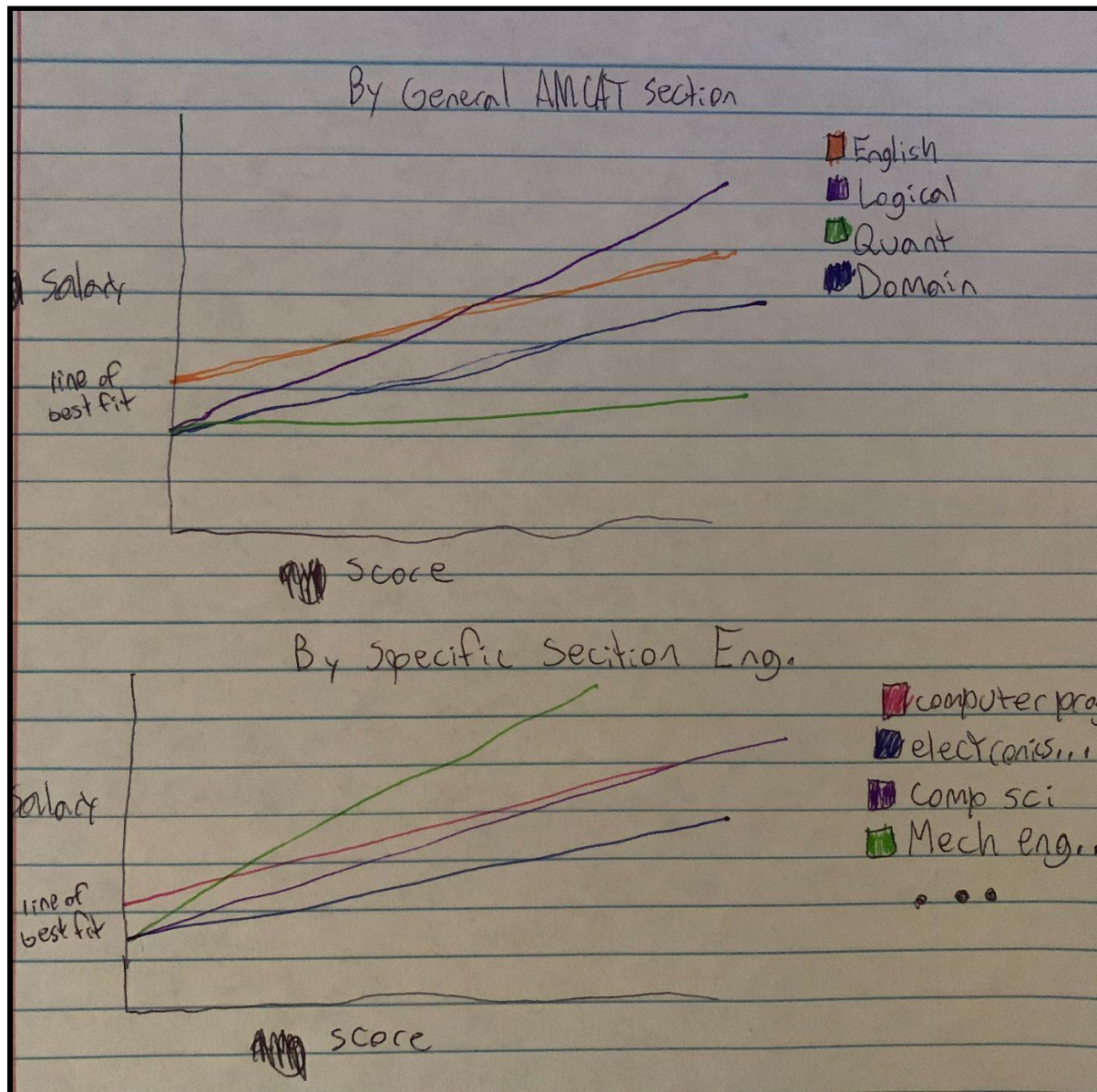
Project Schedule

| Date	Carsten	Lindsey	Tucker
10-24 Making reflective changes(all)	1 visualization and description -Personality and Specialization	1 visualization and description -AMCAT section score and salary	1 visualization and description -personality and GPA
10-31 Making reflective changes(all)	1 visualization and description -Personality and Salary	1 visualization and description - gender, salary, among specializations	1 visualization and description -personality and salary by gender
11-7 Milestone 3 Due	QA - look over other visualizations	QA - look over other visualizations	QA - look over other visualizations
11-14 Milestone 4 Due	Evaluate Visual Encodings of peer	Evaluate Interaction Techniques of peer	Evaluate Design Quality of peer
11-21	Make changes according to peer evaluation	Make changes according to peer evaluation	Make changes according to peer evaluation
11-30 Milestone 5 Due	Prepare Slide and practice presentation	Prepare Slide and practice presentation	Prepare Slide and practice presentation
12-5 Final Submission	Create Screencast + Peer Evaluation	Create Screencast + Peer Evaluation	Create Screencast + Peer Evaluation
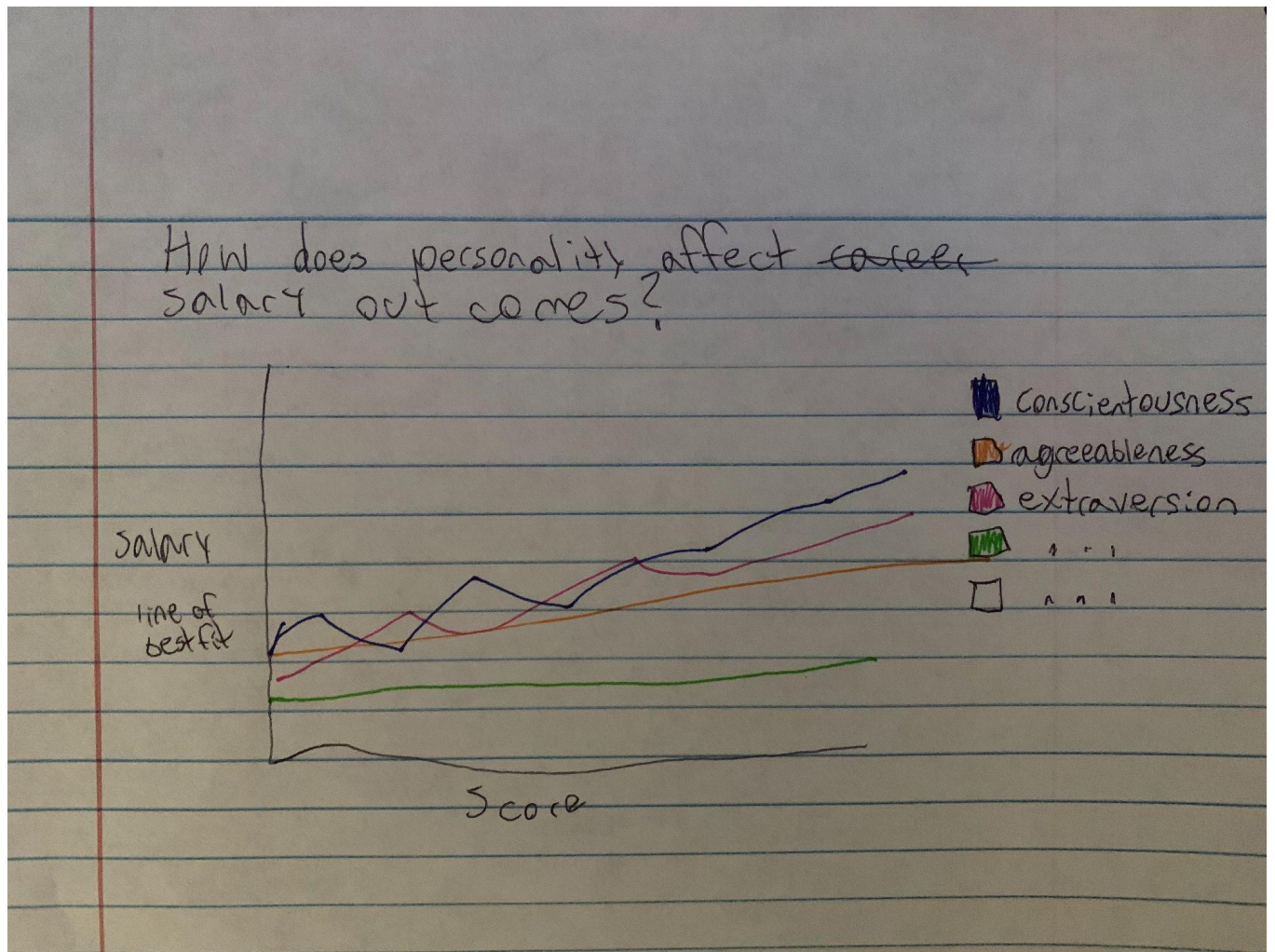
How does personality affect salary outcomes? This design makes use of a line graph in order to show as someone's personality score changes their salary also changes. We can use this visualization to determine the effect personality has on salary. It can easily show what aspects of someone's personality they need to work on to improve their job outcome. The visualization uses color to quickly show different personality traits side by side in the same graph.

By General AMCAT Section

- English
- Logical
- Quant
- Domain

Salary

line of best fit

Score

By Specific Section Eng.

- computer pro
- electronics...
- Comp sci
- Mech eng...

Salary

line of best fit

Score

How do specific test scores affect salary? In this visualization we go with a line of best fit formulated from an unshown scatter plot. We look at how salary changes based on the score achieved in each test. Since we are interested in the STEM section of jobs we made another visualization that looks specifically at test scores that have to do with STEM. We can see what sections are most helpful to earning a higher salary in the end. Again we used color to quickly show different attributes in the same graph so we save space.
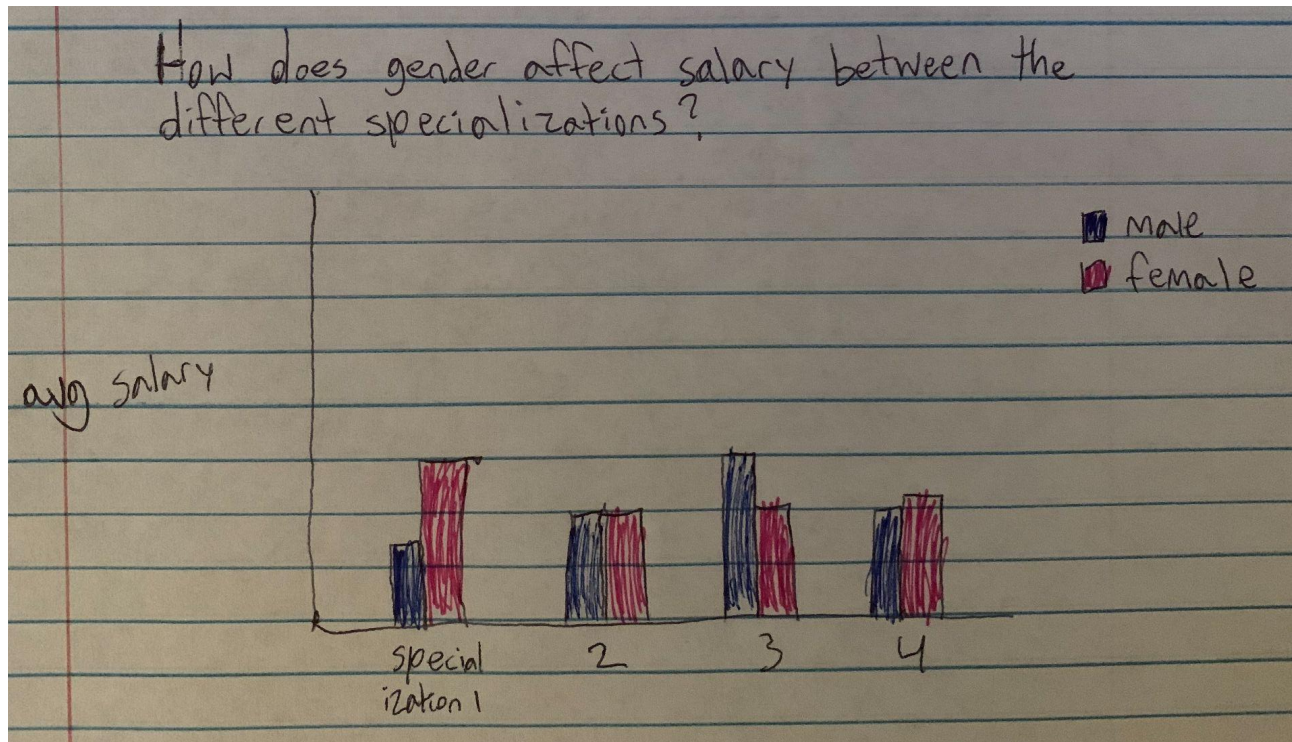
How does gender affect salary between the different specializations? Here we use a dual bar chart, where we can see the average salary of a male or female based on their specialization. This can help us see a gender bias per specialization and across the board. This use of a standard starting point for each bar allows us to quickly compare all bars across the char. Our two colors are distinctly different and consistent. This color allows the viewer to know which bar is for male or female.

Test Score By Specilization

Test Score

0

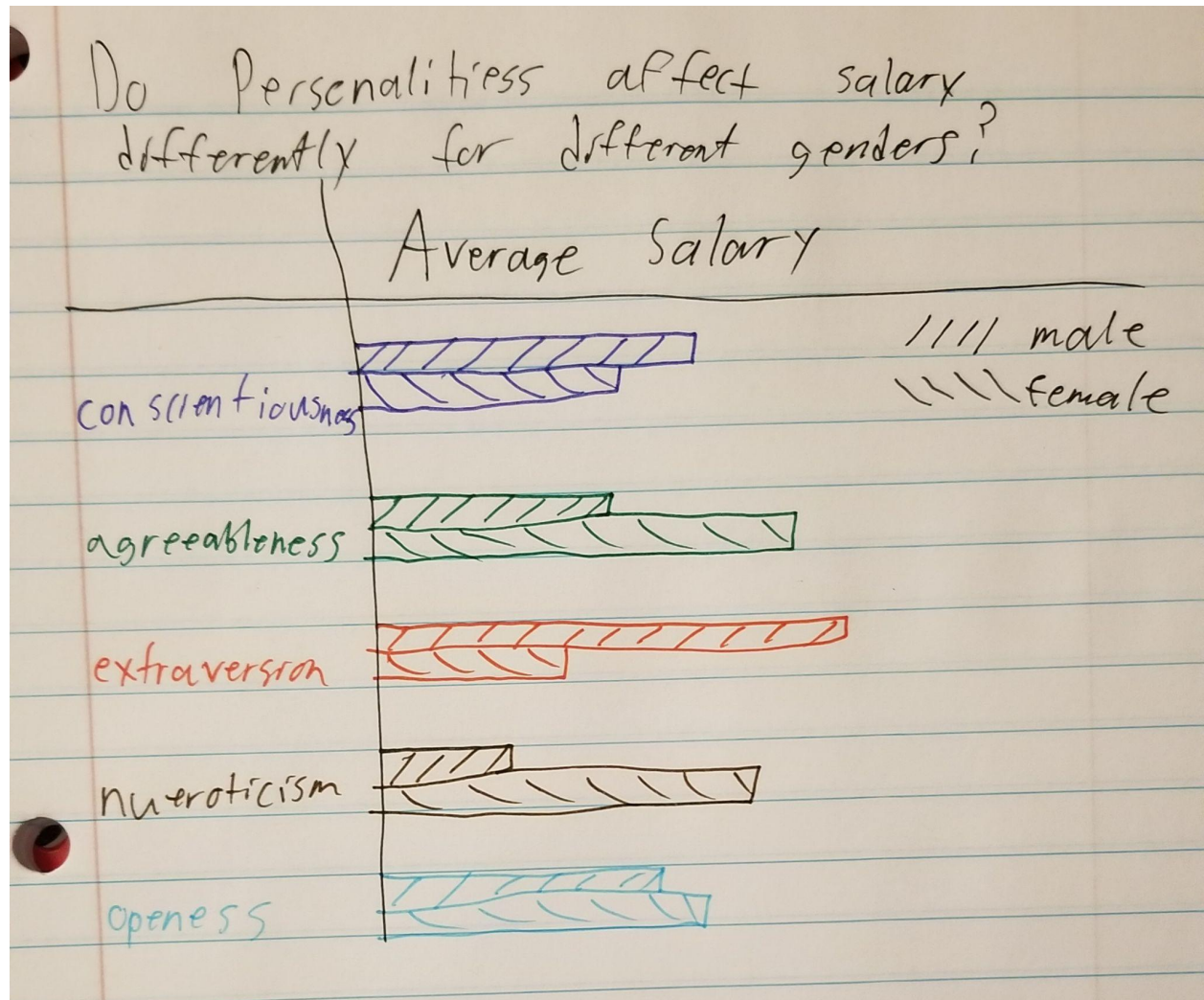electronics and communication engineering
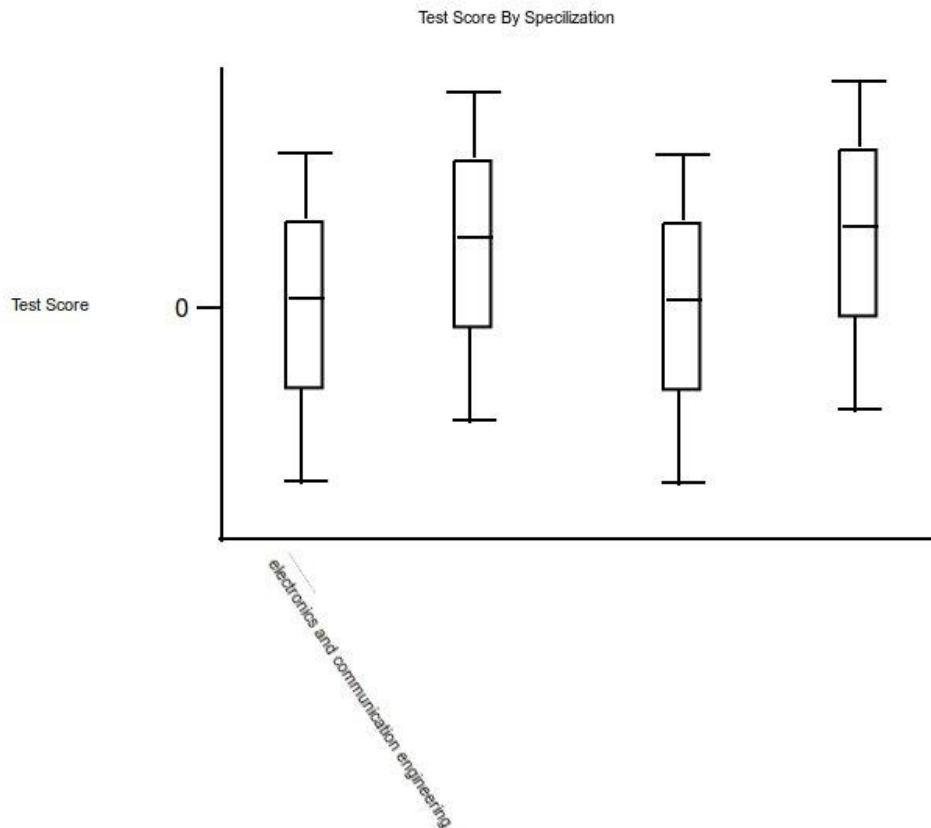
Repeated for each test with each specialization included

Does specialization affect overall test scores? This visualization shows the common range of test scores based on specialization. We think this is a cool thing to look at because you can easily see how the different specializations compare to each other. We can determine which specializations have a better chance to do well on the tests. Here the viewer can determine which specialization does best by looking at the height of the line compared to the others.
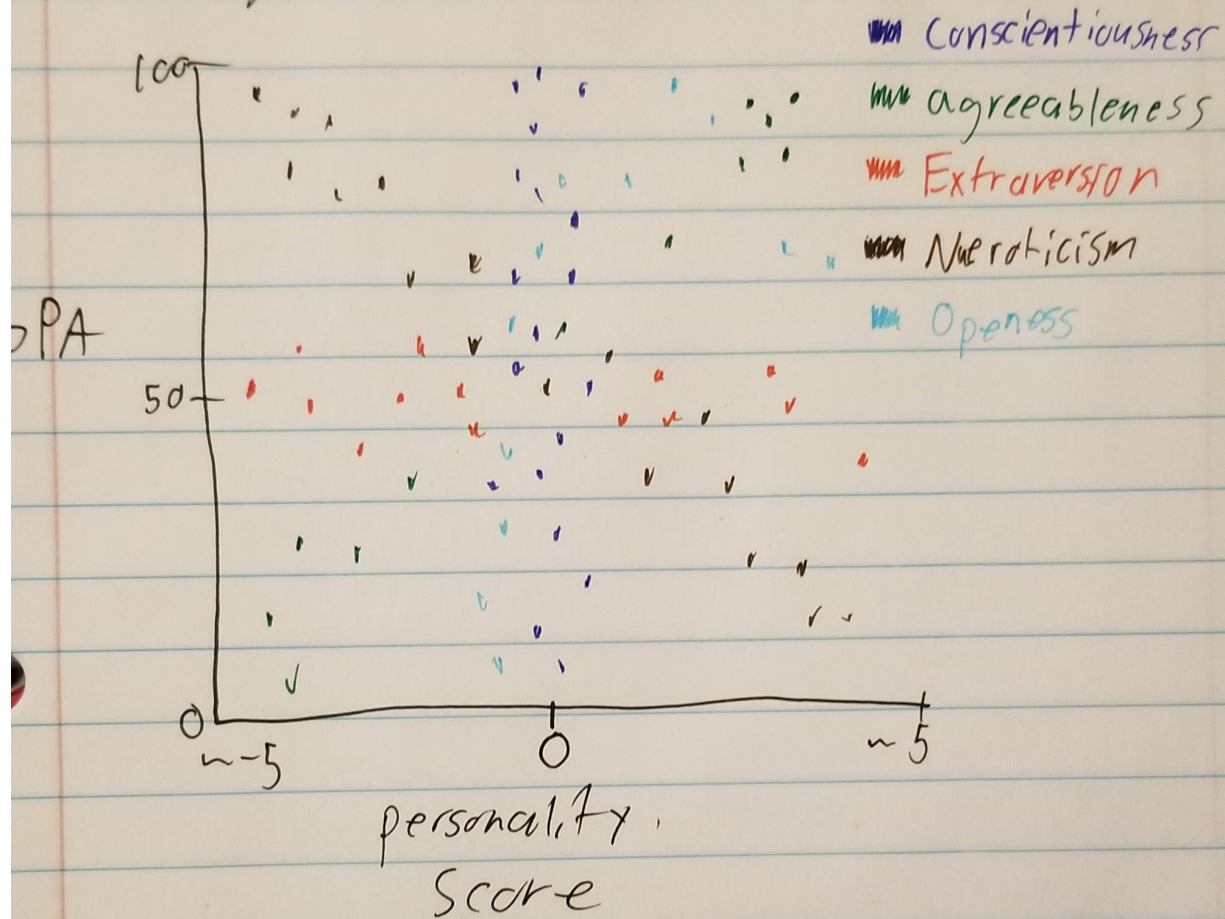
Income

0

(Personality) test score
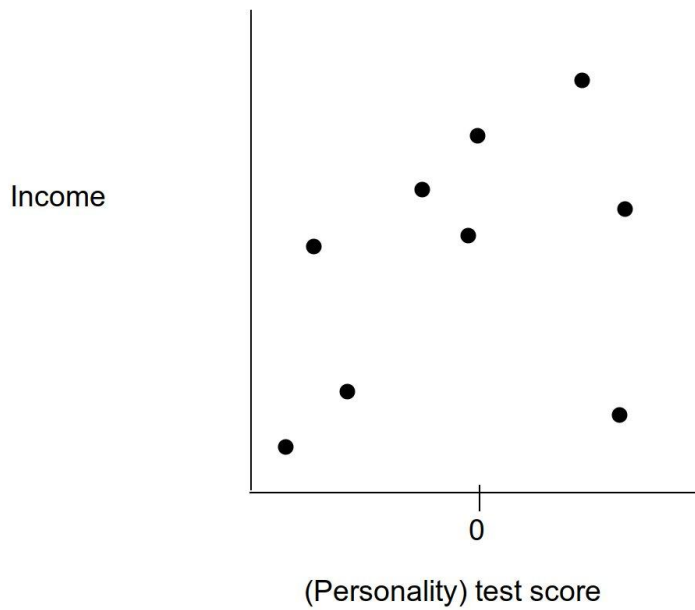
Does personality affect salary? This simple scatter plot is very interesting. We can see a correlation between personality and salary(if one exists). This is the main question we have with our dataset. We hope that we can draw conclusions based on the placement of these dots compared to other dots.

Do Personalitiess affect salary differently for different genders?

Average Salary

//// male
\\\\ female

conscientiousness

agreeableness

extraversion

nueroticism

openess

Do personalities affect salary differently for different genders? Here we have a horizontal dual bar chart that compares salary between genders based on their highest personality score. We can see the overall salary difference between the genders easily if there is one, because they are side by side. This will show us what personalities are more important to people based on their gender. This can possibly show some gender biases in the workplace. We hope the different colors and bar placement help the viewer know which bars go to which personality trait and gender. The viewer can also easily compare all bars because the bars all start justified to the left.

Does personality lead to a higher or lower GPA? This scatterplot might hold a lot of important information. We can determine which personality traits affect how well someone does at school. This information can help someone find out what personality traits they should work on if they want a better grade. Here the viewer can determine what dots show what personality trait based on color, the same colors as the other visualizations.