1. Michael is trying to see if there is a seasonality of the amount of student loans discussed on a reddit subgroup.

He chose this project since he has studied many degrees and while he was doing this, he found helpful tips and advice on reddit.

His hypothesis is that the number of discussions and the amount of student loans that are posted on reddit changes throughout the year depending on the month.

He decided to scrape reddit however there were some problems he encountered since reddit is page based and depending on which page you’re on the data categories change, so what Michael did to solve this was to use an old version of reddit that has basically the same information as the current reddit version and this made it easier for him to scrape. Another two problems he encountered were the dollar format since this could be written in different ways as well as the date format. He also did not use any other outside source for his project.

The title is consistent with what Michael was trying to answer since the information gathered reflects the project’s scope.

1. I liked the idea of choosing that topic since it can be very helpful to people looking for student loans, scraping a site like Reddit seems to be a difficult task specially getting the information you require since as Michael mentioned what you are looking for can be written very differently, and also that the data categories change depending on the page you are on makes it extremely difficult to gather the data. The only thing that I would change in his presentation is the choosing of colours for the line graph he created, I did not know the line changed colour until he mentioned it and had a hard time differentiating between them.