Project: Analyze Survey Data

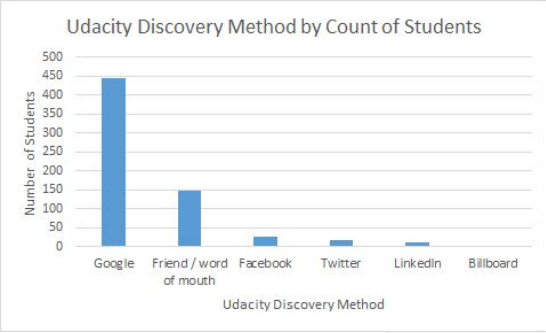
1) What is the most common way students find out about Udacity?

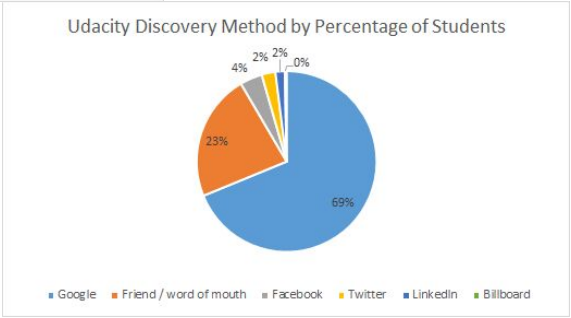
2) What is the most common Nanodegree completed by students?

3) Do US students have shorter commute times?

4) Do older students tend read more books per year than younger students?

What is the most common way students find out about Udacity?



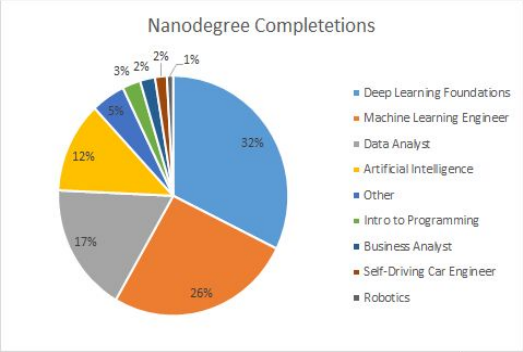


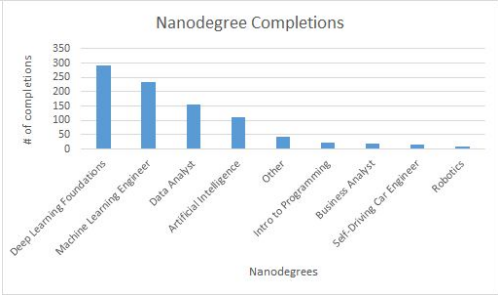
Here is a bar chart and pie chart that shows the most common way students discovered Udacity.

The bar chart shows that Google is the most common way students discovered Udacity. 446 of the 648 students who responded to this question selected Google as the method of finding out about Udacity. The pie chart also shows that Google is the most common method as 69% of students discovered Udacity through Google.

Since Google represents over ⅔ of the way students who completed this survey and who responded to this question, Udacity has lots of room to grow in the social media advertising space as social media represents <10% of the way students found out about Udacity.

What is the most common Nanodegree program completed?

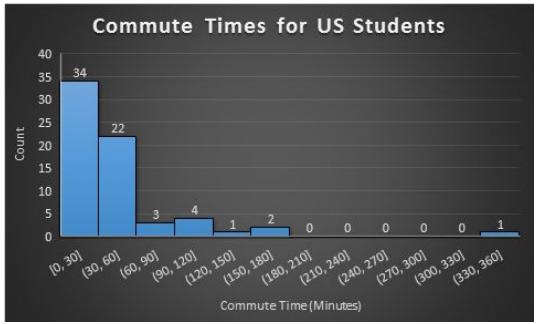


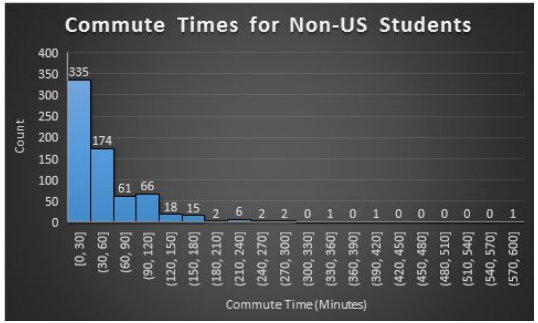


Here is a pie chart that shows the most common Nanodegree program completed by students who took the survey.

Of the 902 Nanodegrees represented in the survey, 291 of the completed Nanodegrees were for the Deep Learning Foundations which is 32% of the Nanodegrees represented in the data.

Commute times by US and Non-US Students



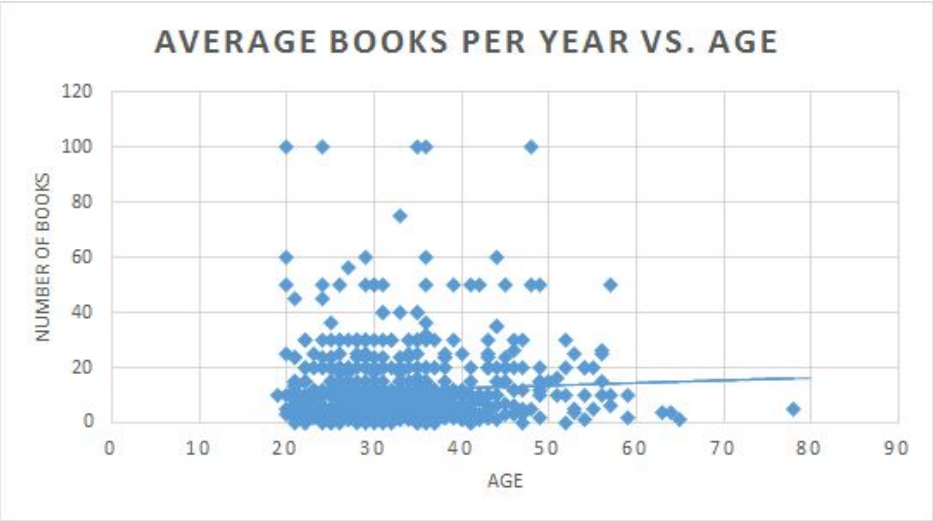


Here are histograms for US and Non-US countries. Both distributions appear to be right skewed. Therefore, the mean for each is higher than the median.

At first glance the mean commute time for US and Non-US students is similar at ~46 and ~51 minutes, respectively. The standard deviation between the two is even more similar with the Non-US standard deviation being slightly higher at ~59 minutes vs. US standard deviation of ~57 minutes. This shows that the average commute time between US and Non-US students is relatively the same with similar variability.

However when removing outliers (any commute time >300 minutes), we start to see a little bit of difference between the two data sets. The US students now have a slightly lower average commute time vs. Non-US students with the US now having a mean of ~42 minutes and Non-US having a mean of ~50 minutes. After removing the outliers, the standard deviations for Non-US students is ~52 minutes , while for US students the standard deviation is ~42 minutes we can see that Non-US students have a greater variability in commute times than US students as the standard deviations

Do older students tend to read more books per year than younger students?



Here is a scatter plot that shows Average Books read per year vs. Age. The age of Udacity students in the survey ranges from 19 to 78. The trend drawn through the scatter plot shows a slight positive correlation between age and the number of books read per year.

After removing age-related outliers (students that showed an age of 0 and any blanks) and book-related outliers (any number of books greater than 100), we can see that the older generation of Udacity students tends to read more books per year than the younger generation. This is could be for a variety of reasons such as the younger generation being more involved with digital content (videos) while the older generation is used to obtaining information through reading/listening to books. Additionally, the older generation may have more free time to read/listen to books due to retirement. The median number of books is 8 and the mode number of books is 10. The bulk of the data points that fall around the median and mode number of books are represented by the younger generation of students (<40). As the student's age increases, we start to see less data points around the median and mode and more data points above the trendline.