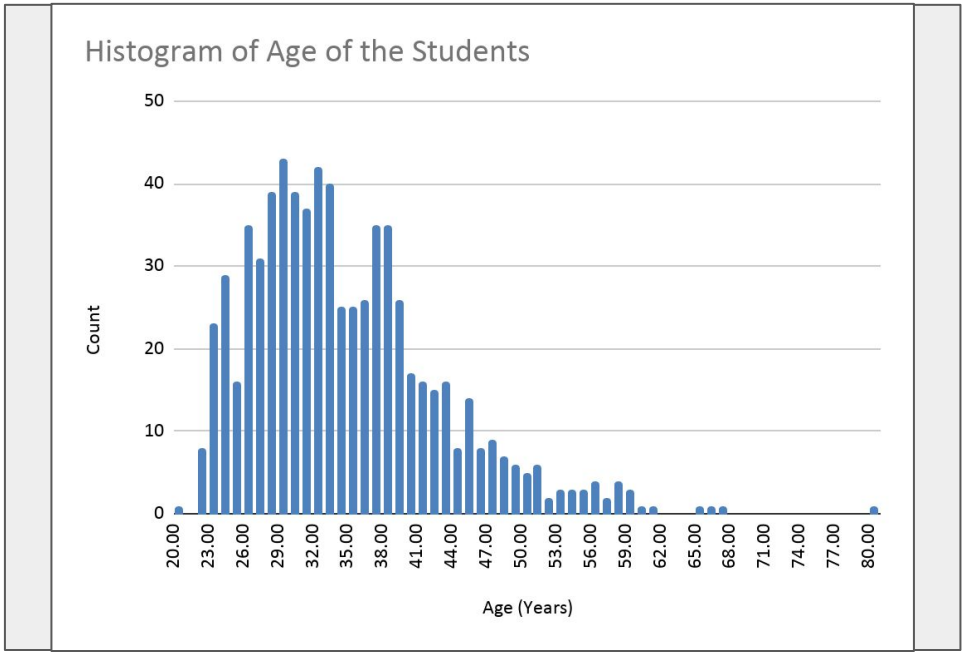
Analyze Survey Data

Marketing Analytics

Submitted by Tonima Ferdous

The data used in this report is obtained from the Survey Respondents only, not from the entire Udacity Student population.

Average Age of Udacity Students

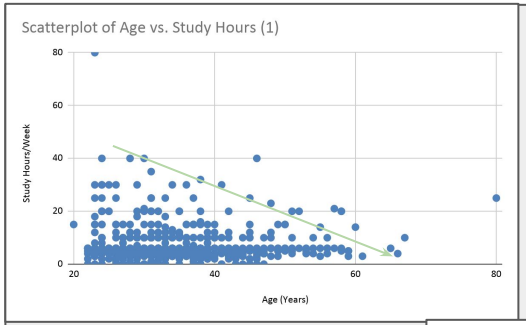


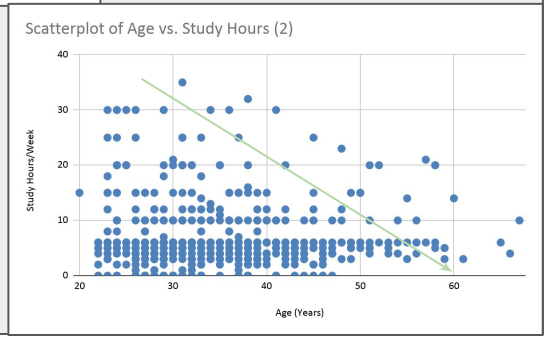
Question 1: **What is the average age of the Udacity students?**

The histogram chart of age is right-skewed with a mean greater than the median and a left tail shorter than the right tail. Here, the mean, median, and mode are 34, 33 and 29, respectively. The mean indicates the average age of the students---34 years. The most commonly occurring age is 29 years, which is both the mode and the peak of the histogram. Also, most of the students age beyond 29 years.

The range is 60 years, which is the difference of the minimum and the maximum values. The standard deviation of the dataset shows that the average distance of each observation from the mean is 8.4 years. Because of these high values, the variability in the data is large and the spread is wide. If the outlier is removed, both of the values become lower, and a less dispersed chart can be obtained.

Age vs. Hours of Study

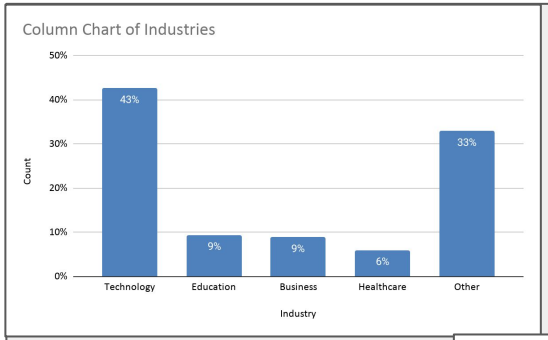


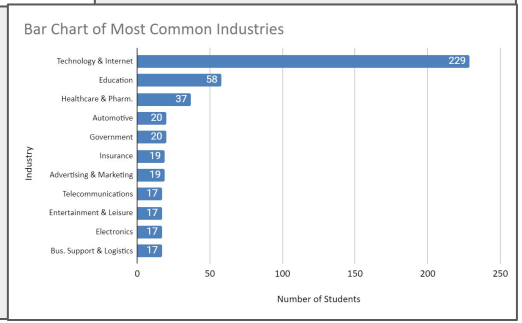


Question 2: **Is there any relationship between age and hours of study?**

Scatterplot (1) investigates the relationship between two quantitative variables, age of the Udacity students and the number of hours of study per week. It shows a strong, negative and linear association between the variables, and the presence of outliers in the data. The negative direction means study hours decrease with age. With the outliers removed, the scatterplot (2) appears weak, still a negative direction can be observed.

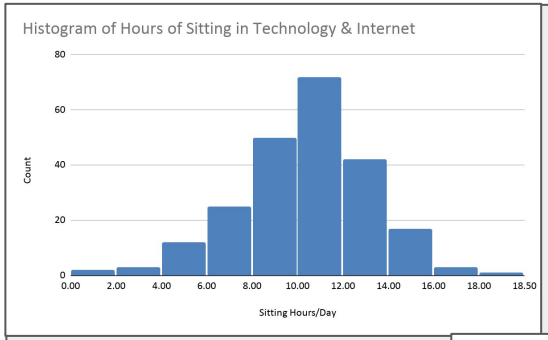
Job Industries of Udacity Students

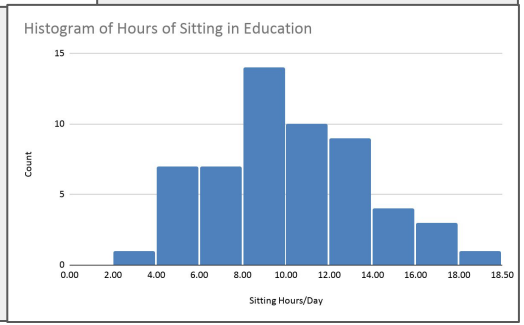




Question 3: **What industries do the students work in?**
The column chart displays the percentages of various job industries students are engaged in. The industries are grouped into five broad categories for the ease of presentation. Of the survey respondents, 43% work in Technology, which is the largest share in the dataset. Education and Business both constitute 9% each, while Healthcare closely follows with 6%.
A bar chart is also developed to show the top ten most common individual industries. Around 230 students out of 750, work in Technology & Internet. This number is 74% higher than the second most common job industry---Education. From the results, it can be assumed that the professionals of Technology & Internet are most likely to enroll in the Udacity Nanodegree programs.

Hours of Sitting by Job Industry





Question 4: **Do the average hours of sitting per day vary by job industries?**

Both of the histograms appear to be symmetric---the mode, the meidan, and the mean are equal. Respondents are found to sit approximately 10 hours per day. It is assumed that these hours are spent both at workplaces and outside of workplaces. Hence, the average sitting hours reported in this study do not vary by job industries. The range of the datasets are found to be 17 hours and 15 hours, respectively. However, the standard deviation is considered to be more accurate while determining the spread. The low values of the standard deviations (2.8 hours and 3.4 hours, respectively) indicate that all the data lay close to the mean. However, the variability in sitting hours are larger in Education compared to Technology & Internet. Note that outliers are not considered while developing the histograms.