Capstone 1: Inferential Statistics Write Up

To examine the effectiveness of MELA’s google keywords data I performed some inferential statistics. The first thing I did to examine the date further was to run correlations on all variables in the dataset to look for relationships between variables. Then I used these correlations to create a heatmap to better visualize these relationships. The heatmap I created showed that there were large positive correlations between number of clicks and the cost of the keyword, sessions, transactions, and revenue. This indicates that as the number of clicks the keyword receives increases so does that cost, sessions, transactions, and revenue from that keyword. Additionally, there were large positive correlations between cost and sessions, transactions, and revenue. This indicates that as the cost of the keyword increases so does the sessions, transactions, and revenue. Lastly, large positive correlations between number of sessions and the transactions as well as revenue. This indicates that as the number of sessions increased the number of transactions and revenue also increased. There were moderate positive correlations between the cost per click (CPC) and the cost per revenue, bounce rate and cost per revenue, pages per session and ROI, pages per session and ecommerce conversion rate, and ROI and sessions per cost.

Large negative correlations were found between CPC and sessions per cost indicating that as CPC increases the sessions per cost decreases. Additionally, large negative correlations were found between pages per session and bounce rate indicating that as the bounce rate increased the pages per session decreased.

Next I conducted some pearson r correlations to examine the statistical significance of these correlations. I found that the correlation of .86 between cost and revenue was statistically significant at the .01 alpha level. Clicks and revenue had a correlation of .87 which was also statistically significant at the .01 alpha level. Finally, I calculated the overall pearson r correlation between the bounce rate and ROI which lead to a statistically significant negative correlation of -.31. To further examine these 3 correlations I broke out the data by device category. The correlation between cost and revenue was positive and significant for all three device categories. The correlation between clicks and revenue was also positive and significant for all three device categories. For the correlations of bounce rate and ROI the tablet keywords did not have a statistically significant correlation however both desktop and mobile devices did have a negative and significant correlation.

The final set of inferential statistics that I calculated where one way ANOVAs to see if there were statistically significant differences in the clicks, cost per click, revenue, and ROI between the device categories. There was not a statistically significant difference of the mean clicks, revenue, or ROI for any of the device categories. There was a statistically significant difference between the mean cost per click between mobile and desktop keywords with desktop keywords having a higher cost per click than those on mobile devices.