Project Summary

A chemist is studying the chemical make-up of three tablets from three distinct brands of aspirin using liquid chromatography-tandem mass spectrometry. In our project, we were provided 7 peaks of mass spectrometry data. In each peak there are three tablets with 3 different intensities. Our prime goal is to check whether there is any difference in each peak between the three kinds of tablets.

We have started exploratory data analysis for each pill of each brand under each peak. We used the Time versus IntesnityPlot, QQPlot, Boxplots, and SmoothScattered plots for our exploratory data analysis. We can see the distribution of numerical data and skewness by displaying the data quartiles (or percentiles) and averages using the Intensity vs Time Plot; QQplots shows the type of distribution; and Boxplots shows the distribution of numerical data and skewness by displaying the data quartiles (or percentiles) and averages. After that, we used Smoothscatterd plot to create a smoothed color density representation of a scatterplot using a (2D) kernel density estimate. We then used the mean of the intensity values for a specific aliqout for each tablet of each brand to compute the area under a peak (AUP). Then we averaged the means of three aliqout intensities for each pill type, yielding nine total averages of intensities for three pills from each brand. Then, for each peak, we ran a one-way anova. ANOVA shows us whether there are differences in group means, but it doesn't tell us what they are. Later, we used a Tukey's Honestly Significant Difference (Tukey's HSD) post-hoc test for pairwise comparisons to see whether groups were statistically different from one another.

First Peak was 136.5 92.90. For combined Bayer Pills we have seen from the boxplot that the data is right skewed and for combined PV pills the data was left skewed and for the Walgreen Pills the median was shifted to the right side. From our one way Anova(Intensity~Pills), we had F-statistic of 1.73 with 2 degrees of freedom and the p-value was 0.255 which is greater than the alpha=0.05. For this peak location, it appears that there are **no substantial differences between the tablets**. According to the findings of the post-hoc test, there are no statistically significant differences ($p > 0.05$) between pill groups.

Next Peak was 178.7 93.00. For combined Bayer Pills we have seen from the boxplot that the median shifted to the left and for combined PV pills the median shifted to the right and for the Walgreen Pills the data seems to be uniformly distributed. From our one way Anova(Intensity~Pills), we had F-statistic of 0.559 with (3-1) = 2 degrees of freedom and the p-value was 0.60 which is greater than the alpha=0.05. For this peak location, it appears that there are **no substantial differences between the tablets**. According to the findings of the post-hoc test, there are no statistically significant differences ($p > 0.05$) between pill groups.

Our next Peak was 199.10 91.90. For combined Bayer Pills we have seen from the boxplot that the median shifted to the left and for combined PV pills the data is uniformly distributed with an outlier on the right side and for the Walgreen Pills the median shifted to the right. From our one way Anova(Intensity~Pills), we had F-statistic of 6.54 with (3-1) = 2 degrees of freedom and the p-value was 0.03 which is less than the alpha=0.05. For this peak location, it appears that there are **present a substantial difference between the tablets**. According to the findings of the post-hoc test, we can see that pill groups 2 and 1 have statistically significant differences ($p<0.05$).

Our 4th Peak was 199.10 92.90. For combined Bayer Pills we have seen from the boxplot that the median shifted to the left and for combined PV pills the data is uniformly distributed and for the Walgreen Pills the the data is uniformly distributed. From our one way Anova(Intensity~Pills), we had F-statistic of 13.79 with (3-1) = 2 degrees of freedom and the p-value was 0.006 which is less than the alpha=0.05. For this peak location, it appears that there are **present a substantial difference between the tablets**. According to the findings of the post-hoc test, we can see that pill groups 2 and 1 have statistically significant differences ($p<0.05$).

Our 5th Peak was 199.10 137.00. For combined Bayer Pills we have seen from the boxplot that the median shifted to the left and for combined PV pills the data is uniformly distributed and for the Walgreen Pills the the median shifted to the left with skewness and an outlier on the left side. From our one way Anova(Intensity~Pills), we had F-statistic of 18.52 with (3-1) = 2 degrees of freedom and the p-value was 0.003 which is less than the alpha=0.05. For this peak location, it appears that there are **present a substantial difference between the tablets**. According to the findings of the post-hoc test, we can see that pill groups 2 and 1 have statistically significant differences ($p<0.05$).

Our 6th Peak was 240.00 92.900. For combined Bayer Pills we have seen from the boxplot that the median shifted to the left and for combined PV pills the data is uniformly distributed and for the Walgreen Pills the data is uniformly distributed. From our one way Anova(Intensity~Pills), we had F-statistic of 0.611 with (3-1) = 2 degrees of freedom and the p-value was 0.574 which is greater than the alpha=0.05. For this peak location, it appears that there is **no substantial difference between the tablets**. According to the findings of the post-hoc test, we can see no statistically significant differences between the tablets.

Our last Peak was 240.00 137.100. For combined Bayer Pills we have seen from the boxplot that the data is uniformly distributed and for combined PV pills the data is uniformly distributed and for the Walgreen Pills the data is uniformly distributed. From our one way Anova(Intensity~Pills), we had F-statistic of 1.467 with (3-1) = 2 degrees of freedom and the p-value was 0.303 which is greater than the alpha=0.05. For this peak location, it appears that there is **no substantial difference between the tablets**. According to the findings of the post-hoc test, we can see no statistically significant differences between the tablets.

From Analysis of Variance (ANOVA), we have seen significant differences for the peaks of (199.10 92.90) and (199.10 137.00). For other peaks, we can see that the p-value of the pills variable was higher indicating no significant difference between the pills.