

# COMPARING 3 BRANDS OF ASPIRIN PILLS •

Analysis by: Gavin Gunawardena

Using dataset from: LC-MS/MS in the Clinical Laboratory – Where to From Here?

by Stefan KG Greve and Ravinder J Singh and which can be found here:  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3052391/>

# Intro



## Background and Dataset

- Machine generated dataset analyzing the chemical make-up and intensities of three brands of aspirin pills
- Dataset consists of:
  - 3 pills from 3 different brands of aspirin (Bayer, PV, and Walgreens)
  - Dataset consists of 63 tsv files grouping the pills by brand, peak of intensity, replicant, and aliquot
  - There are 3 aliquots for each replicant, 3 replicants for each peak, and 7 peaks for each brand in the dataset
  - Peak: time when the intensity of each pill is at its highest, usually when a chemical is released from the pill, usually at artificially timed intervals
  - Replicant: replicas of the same pill; 3 replicas of 3 different brands have been tested
  - Aliquots: Series of measurements within a timeframe of less than a second

# Intro



## Objective and Assumptions

- Objective is to test the null hypothesis that the 3 brands of aspirin are the same with a 95% confidence interval
  - Plan is to measure the Area Under a Peak (AUP) for each aliquot of each replicant
  - Next, run various ANOVA analyses to compare the AUPs between brands for each peak to confirm or deny the null hypothesis
- Assumptions:
  - Data was gathered using a standardized process, since it was machine generated
  - There is still some possibility for error within the process utilized to attain the data
    - Thus, missing data will be removed, and statistical tests will be run with and without outliers.

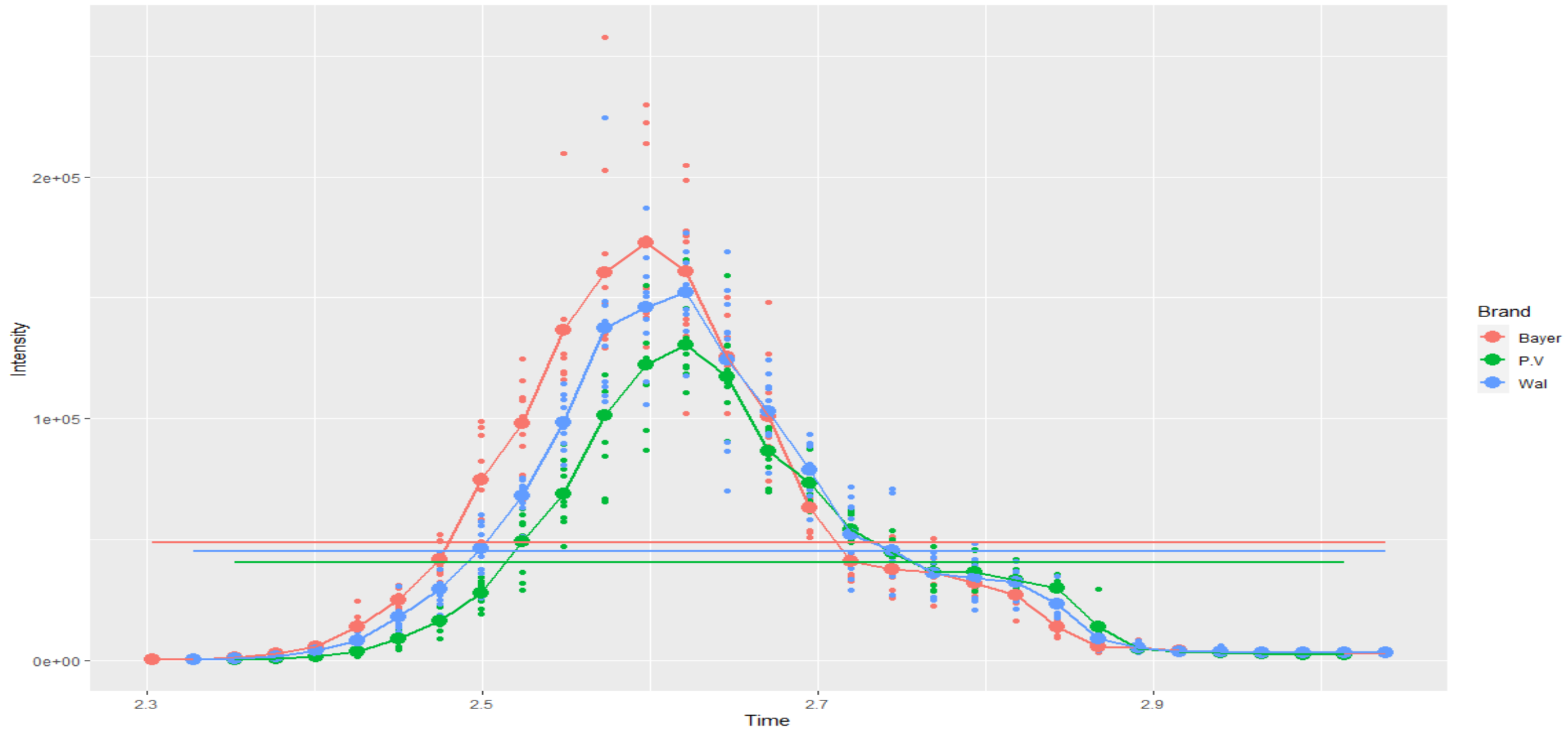
# Exploratory Analysis

## Dataset Summary

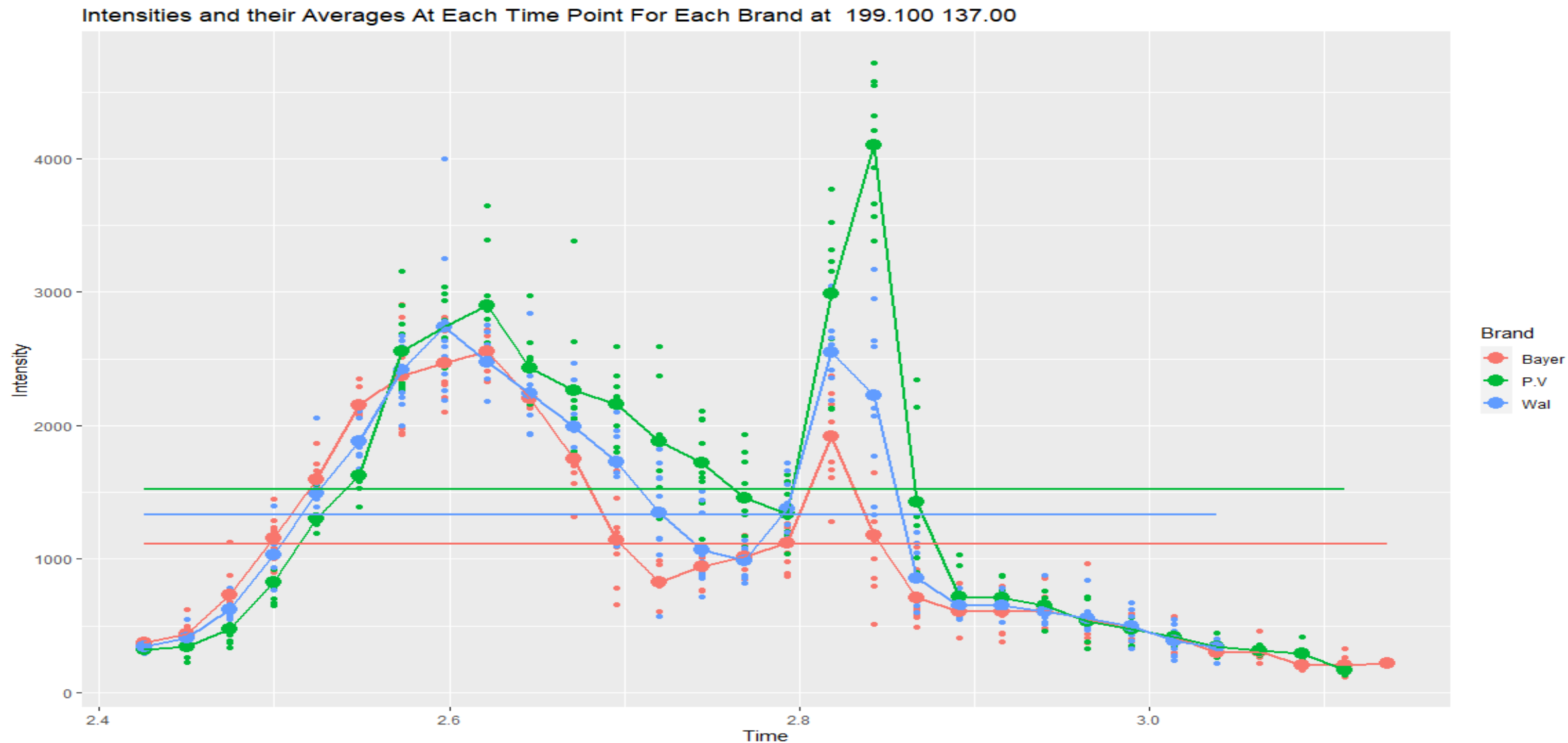
##	Brand	Peak	Replicant	Aliquot
##	Length:4798	Length:4798	Min. :1.00	Min. :1.000
##	Class :character	Class :character	1st Qu.:1.00	1st Qu.:1.000
##	Mode :character	Mode :character	Median :2.00	Median :2.000
##			Mean :1.99	Mean :2.003
##			3rd Qu.:3.00	3rd Qu.:3.000
##			Max. :3.00	Max. :3.000
##	Time	Intensity		
##	Min. :2.303	Min. : 10		
##	1st Qu.:2.548	1st Qu.: 270		
##	Median :2.720	Median : 785		
##	Mean :2.717	Mean : 10071		
##	3rd Qu.:2.867	3rd Qu.: 2590		
##	Max. :3.161	Max. :257870		

# Exploratory Analysis

Intensities and their Averages At Each Time Point For Each Brand at 136.5 92.90

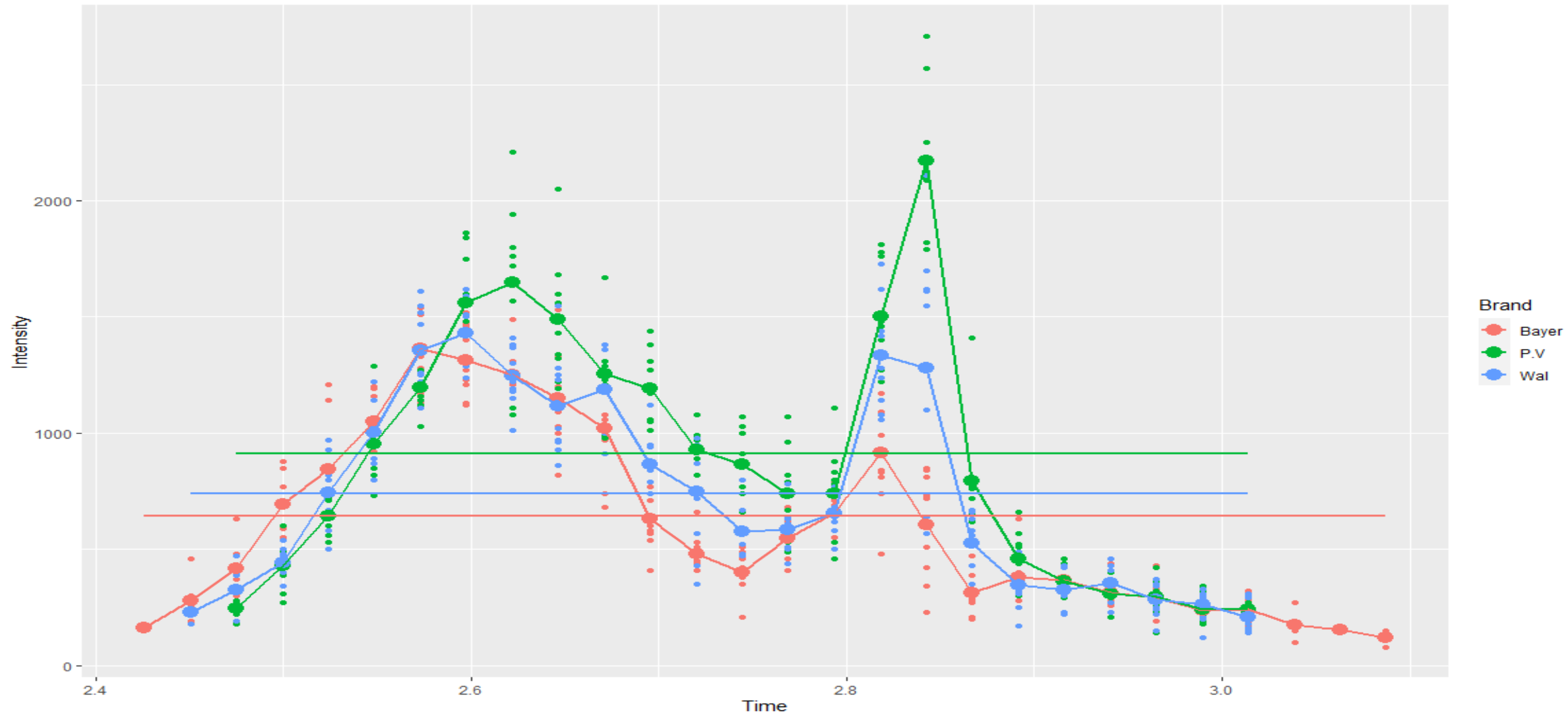


# Exploratory Analysis

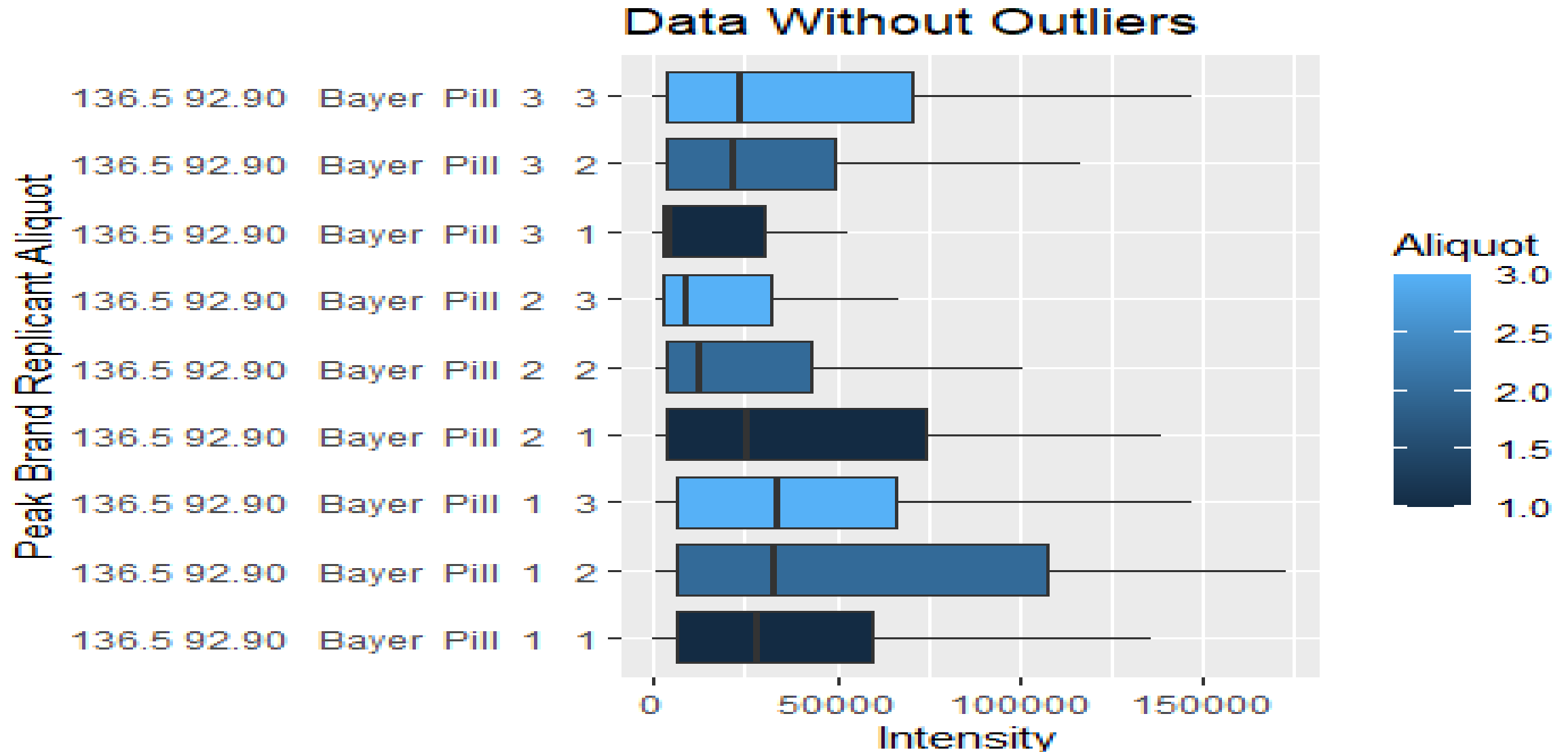


# Exploratory Analysis

Intensities and their Averages At Each Time Point For Each Brand at 199.10 91.90

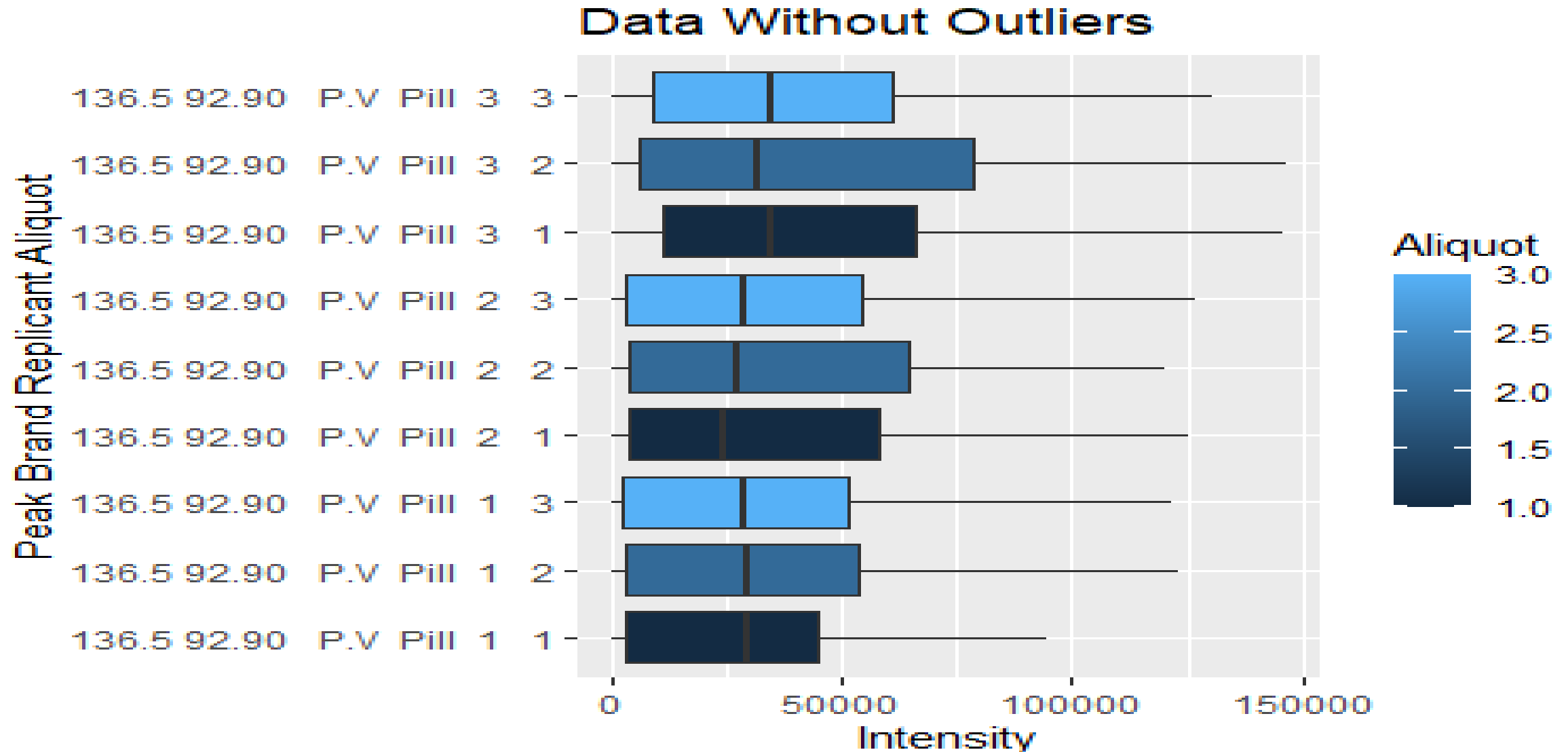


# Exploratory Analysis

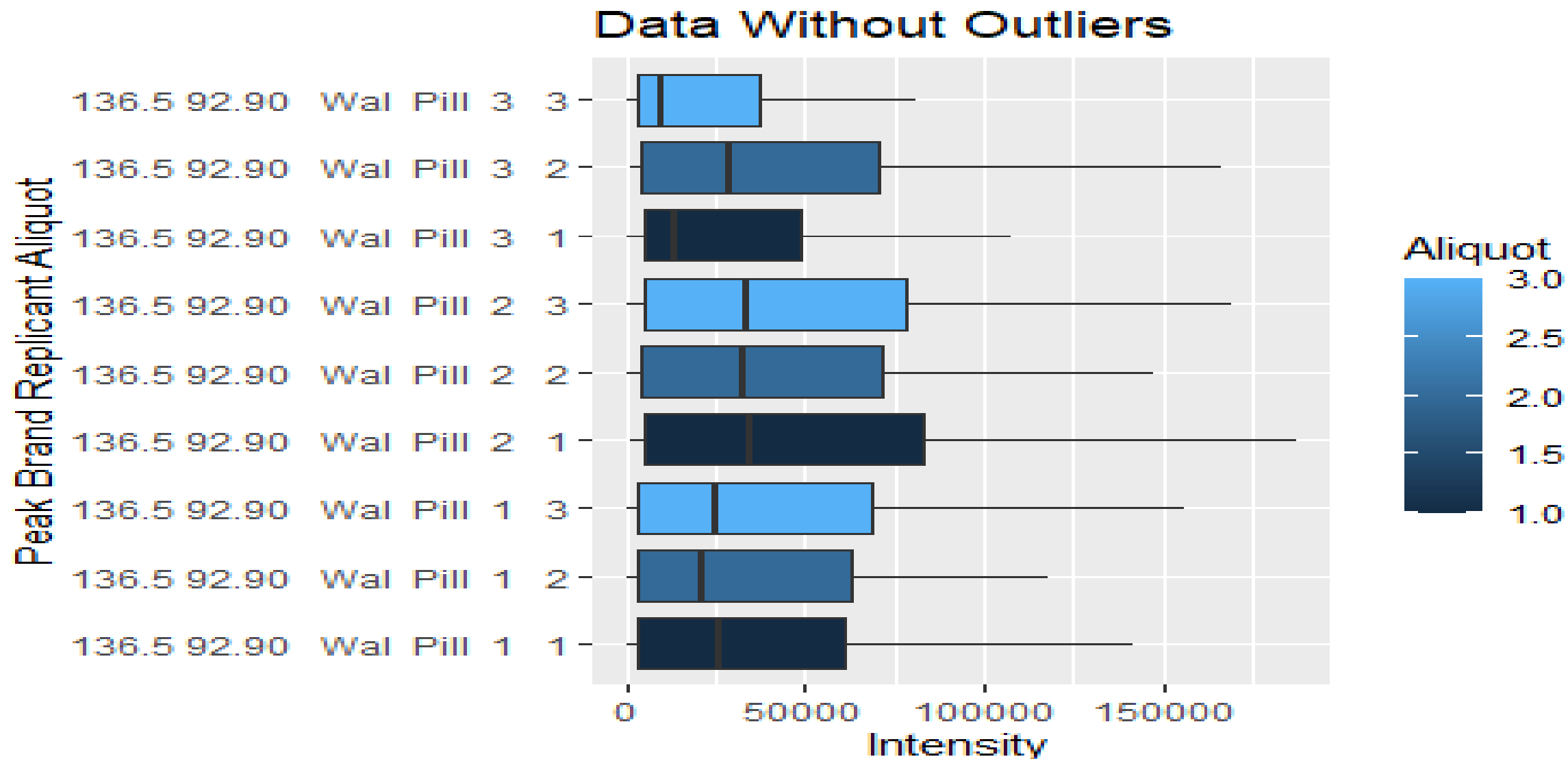




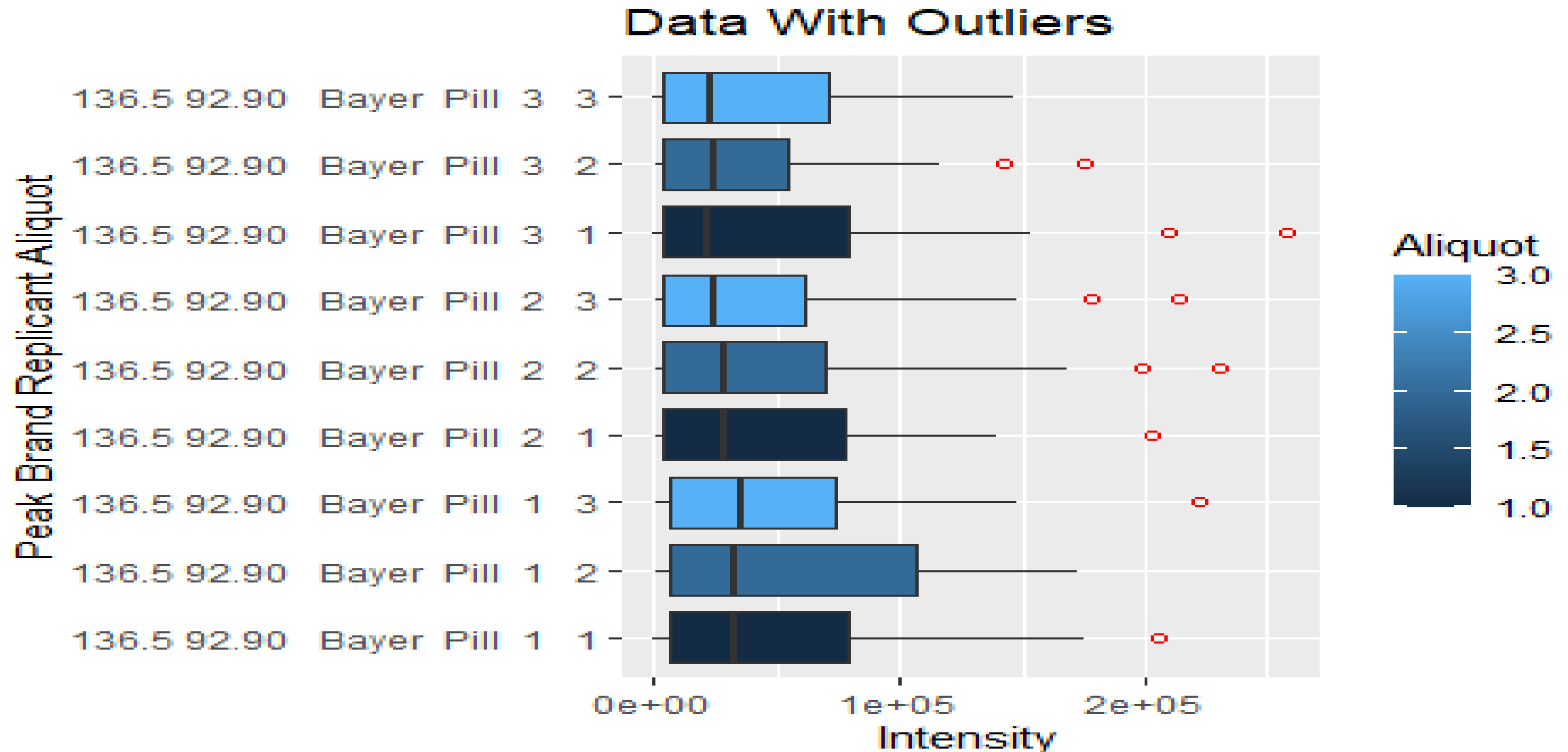
# Exploratory Analysis



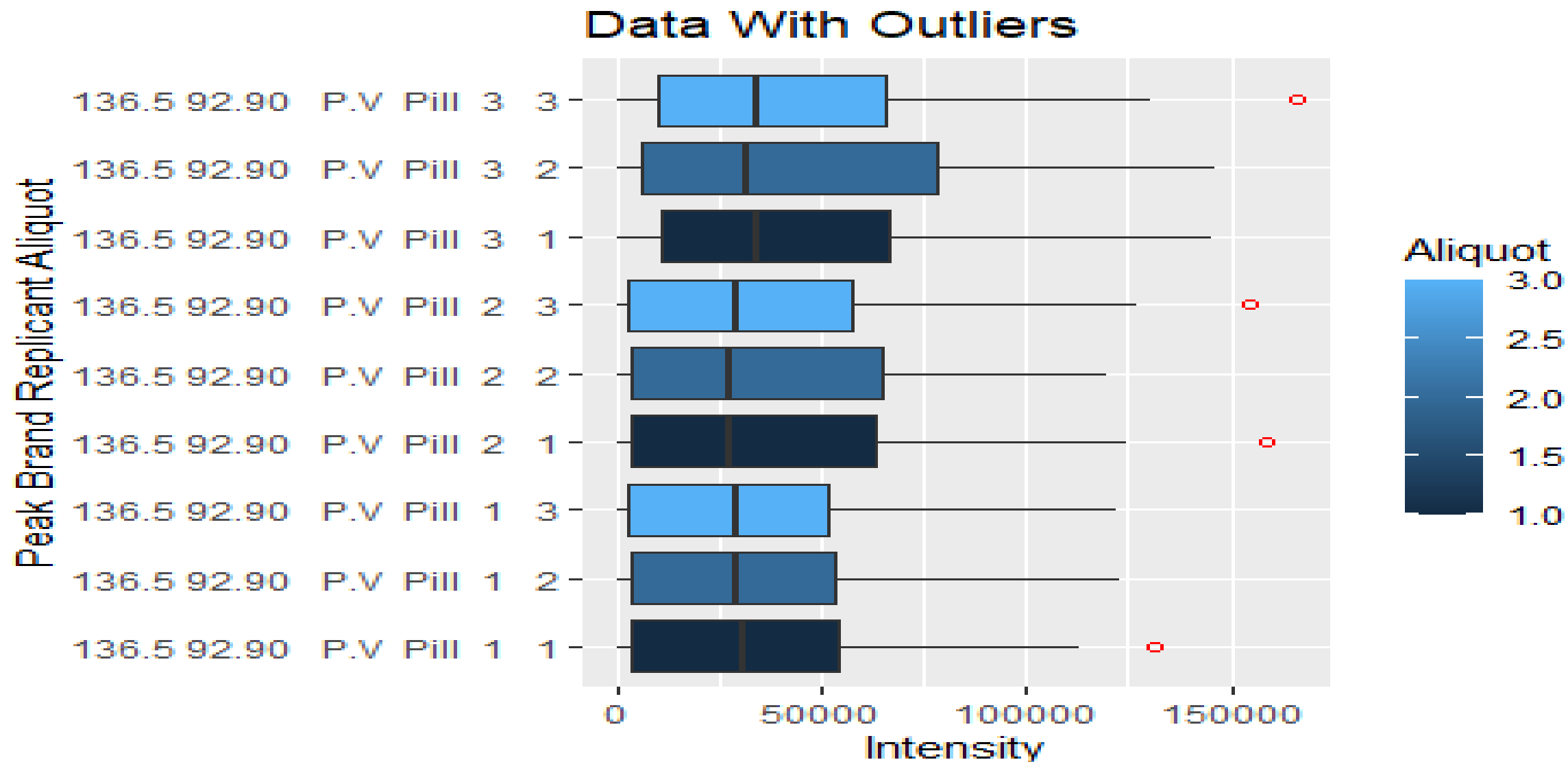
# Exploratory Analysis



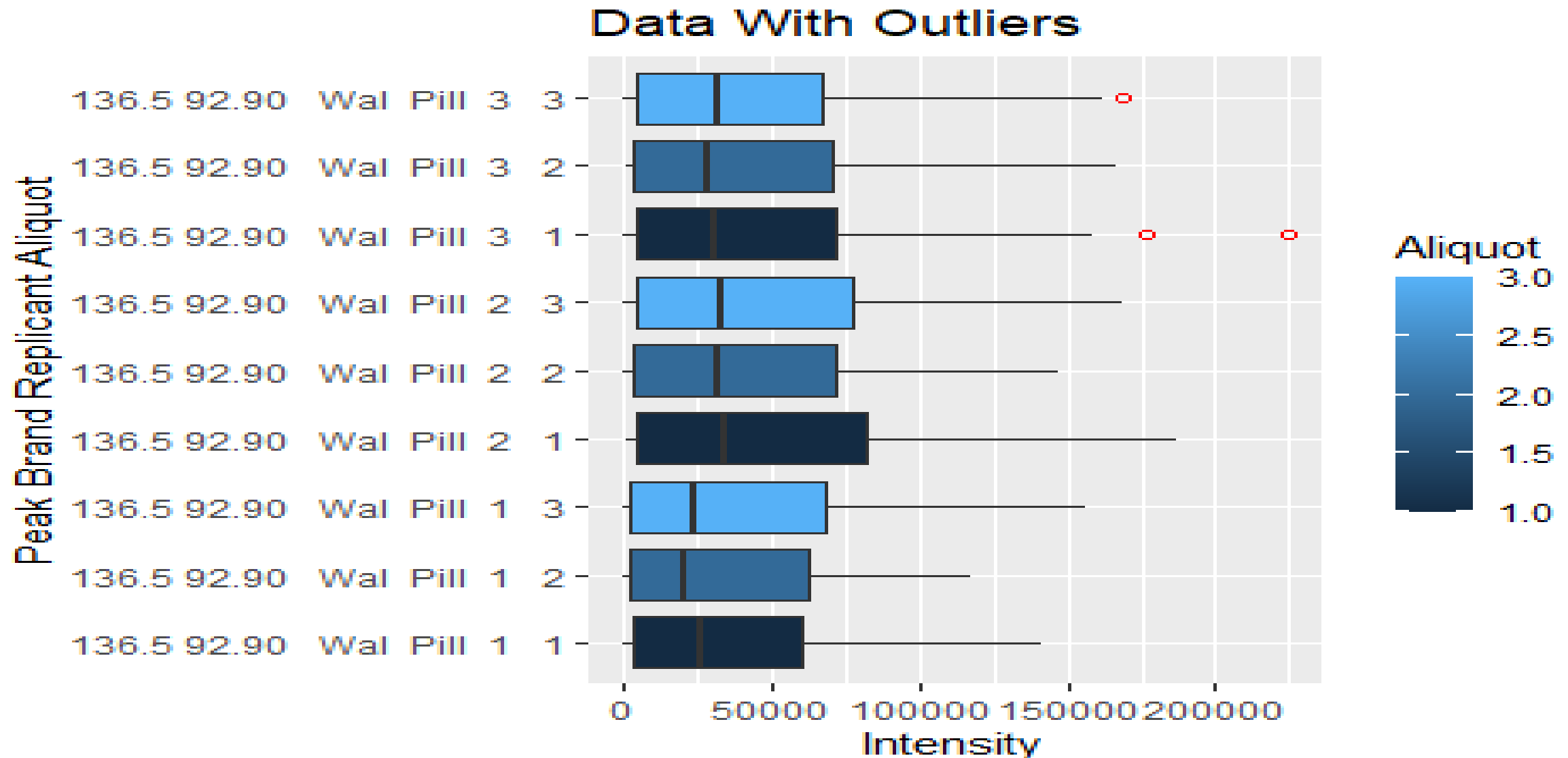
# Exploratory Analysis



# Exploratory Analysis



# Exploratory Analysis



# Analysis



## One Way ANOVA and Pairwise comparisons of AUPs for each Brand

- Complete with an alpha of .5 to get a 95% confidence interval statistical significance comparison
- All 3 at once and pairwise comparisons using Student's t-distribution tests

Brand	AoV P Value	Pairwise T Test P Value
Brand - With Outliers	0.973	
Brand - w/o Outliers	0.729	
P.V. vs Bayer - With Outliers		0.85
Wal vs Bayer - With Outliers		0.98
Wal vs P.V. - With Outliers		0.83
P.V. vs Bayer - w/o Outliers		0.53
Wal vs Bayer - w/o Outliers		0.46
Wal vs P.V - w/o Outliers		0.91

# Analysis

## Ad-Hoc Pairwise comparisons

- Tukey-Kramer comparison which has an error correction that assists when comparing pairs within a group with sample sizes that aren't exactly the same.
- Bonferroni comparison which has an error correction that, similarly to the Tukey-Kramer test, also assists when comparing pairs within a group with sample sizes that aren't exactly the same.

Brand	Tukey_Kramer_P_Values	Bonferroni_P_Values
P.V. vs Bayer - With Outliers	0.489	0.818
Wal vs Bayer - With Outliers	0.984	0.887
Wal vs P.V. - With Outliers	0.389	0.118
P.V. vs Bayer - w/o Outliers	0.078	0.140
Wal vs Bayer - w/o Outliers	0.030	0.132
Wal vs P.V - w/o Outliers	0.924	0.624

# Conclusions

## **Null Hypothesis that the Aspirin Pills Are the Same**

- Failed to reject null hypothesis that the aspirin pills are the same
  - Initial ANOVA test revealed that at a 95% confidence interval, there is no difference between the brands with or without the outliers being removed
  - This result was confirmed when running pairwise t-tests between the 3 brands
- Ad-Hoc Analysis
  - When outliers are not removed and when comparing the individual brands between each other using either Tukey-Kramer or Bonferroni analyses, the null hypothesis is not rejected at a 95% confidence interval
  - When outliers are removed and the Tukey-Kramer analysis but not the Bonferroni analysis is used, the Walgreens and Bayers brand aspirin pills are found to be different at a 95% confidence interval, rejecting the null hypothesis
- Standard One-Way ANOVA is being prioritized and thus the null hypothesis fails to be rejected



# Sources

- <https://www.vedantu.com/formula/anova-formula>
- <https://www.educba.com/t-test-formula/>
- <https://www.statology.org/tukey-vs-bonferroni-vs-scheffe/>
- <https://aaronshlegel.me/tukeys-test-post-hoc-analysis.html>
- <https://www.itl.nist.gov/div898/handbook/prc/section4/prc473.htm>
- <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3052391/>