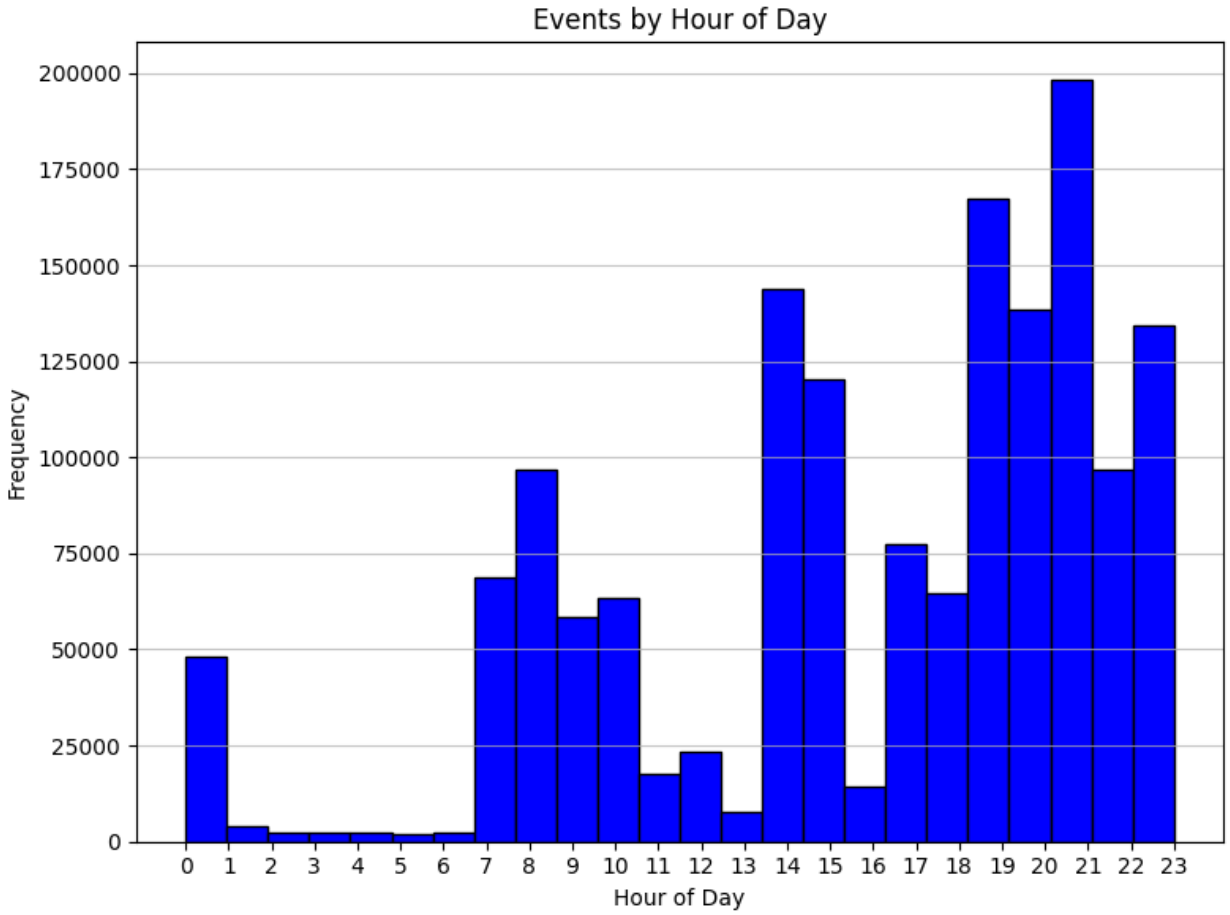
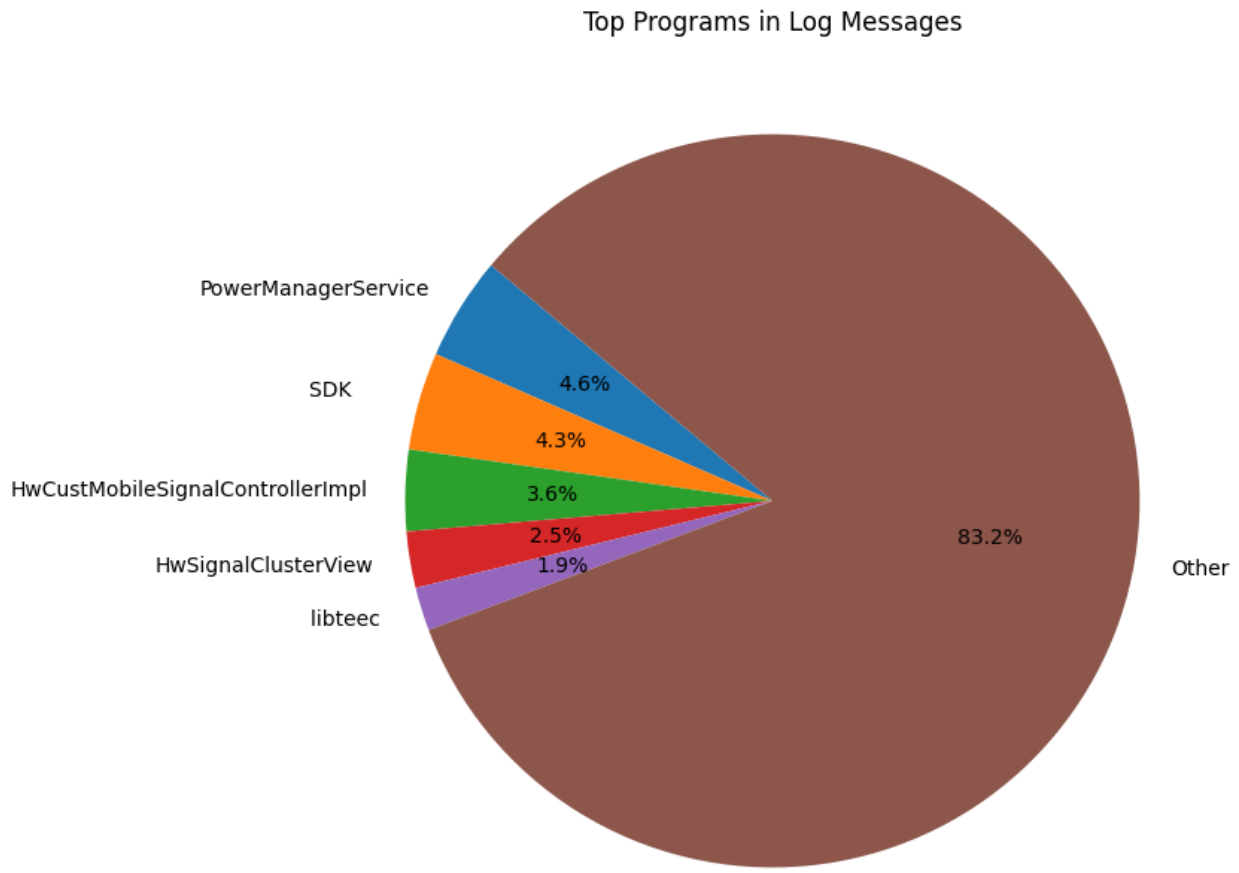
Prof. Subhasis Ray
Introduction to Data Science
Usman Akinyemi
U20220090

Question 3-1

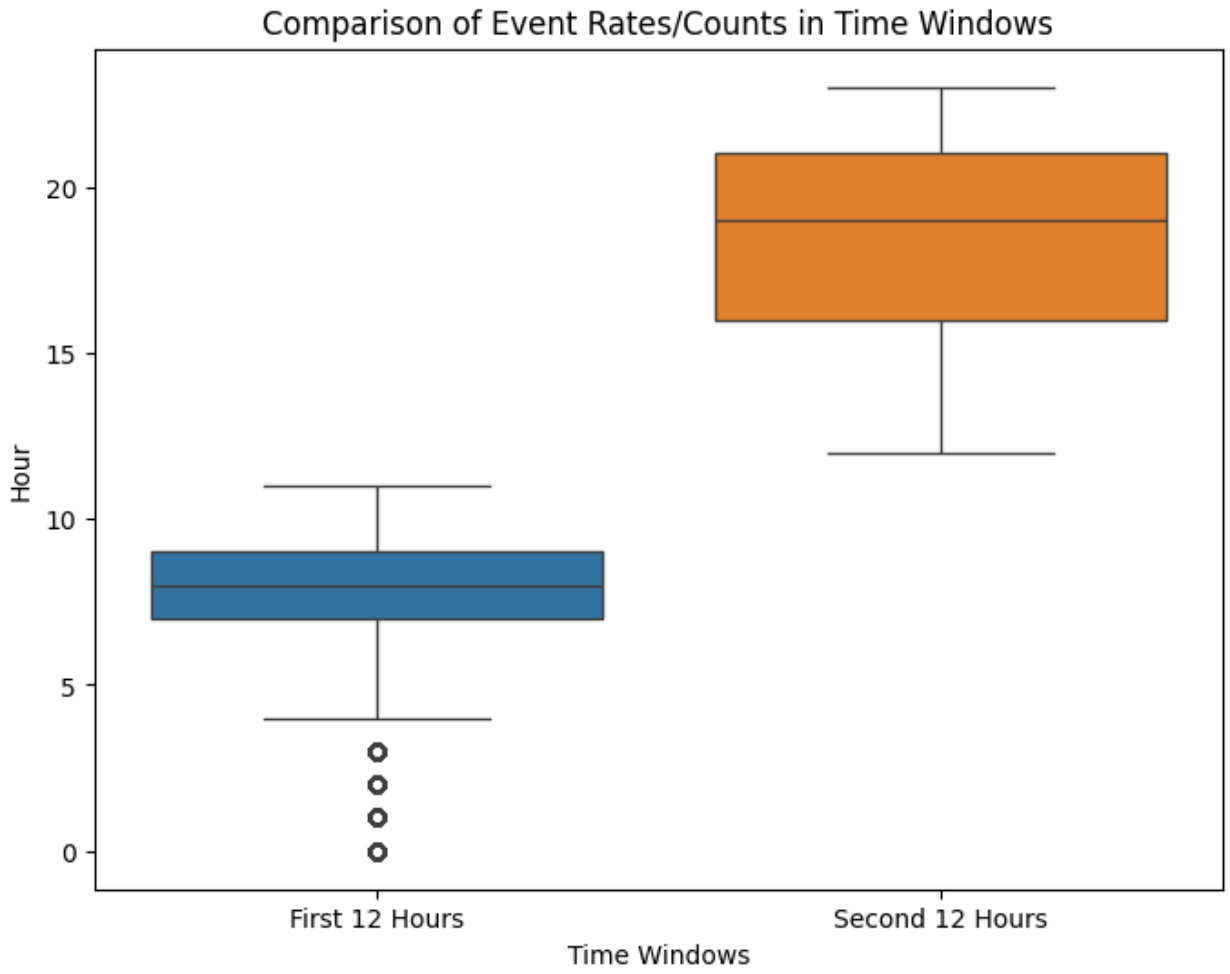


Question 3 - 2



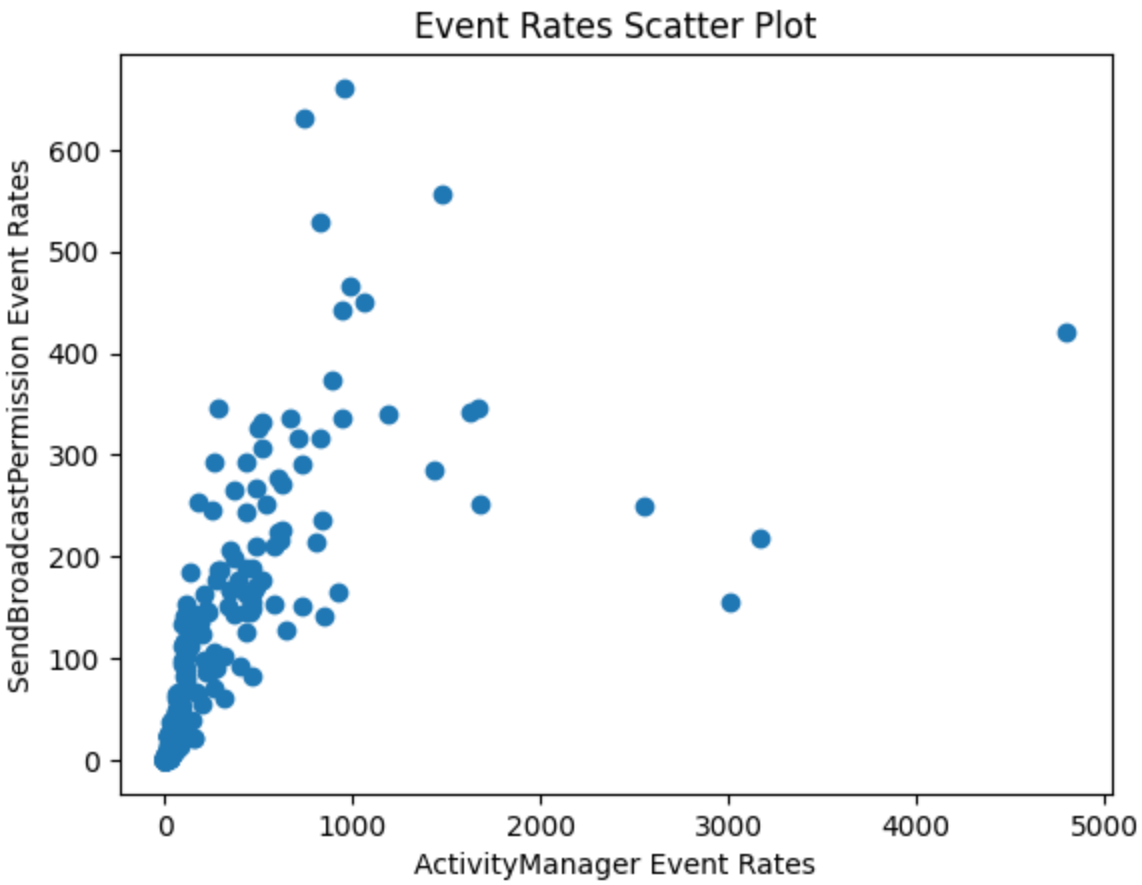
Question 4 - 1
- Null Hypothesis H0: There is no difference in the number of events/mean of event between the first 12 hours and the second 12 hours.
- Alternative Hypothesis H1: There is a significant difference in the number of events/mean of event between the first 12 hours and the second 12 hours.

Resultss:-
T-Statistic: -1920.688925730464
P-Value: 0.0
Reject the null hypothesis. There is a significant difference in event counts.

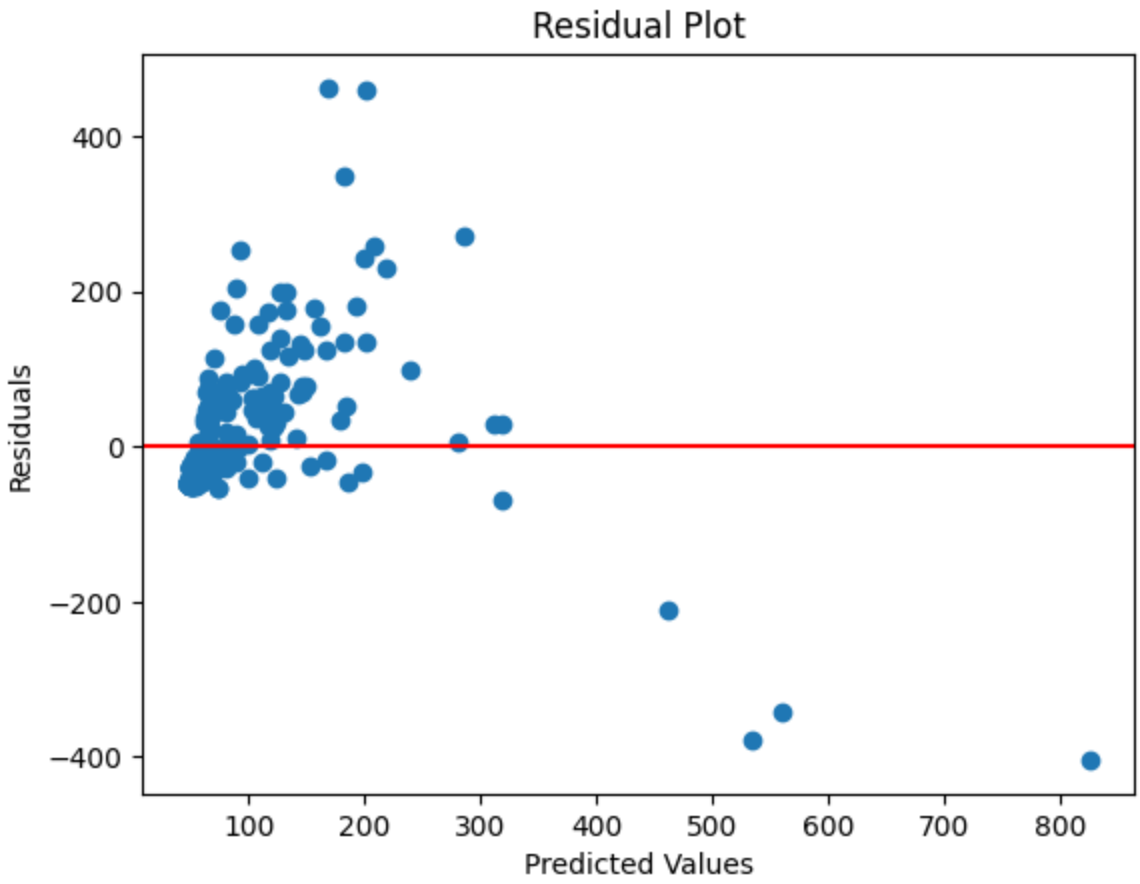


From the box plot, we can see that the first 12 hrs has lower Quartiles values compared to the last 12 hrs thereby showing that the rejection of the hypothesis is valid.

Question 4 - 2


Event Rates Scatter Plot
SendBroadcastPermission Event Rates
ActivityManager Event Rates
0
100
200
300
400
500
600
0
1000
2000
3000
4000
5000

Question 4 2 -



Intercept: 47.52347466337382
Coefficient: [0.16209725]
R-squared: 0.43655017449341504

Personal Comment :-
The intercept of approximately 47.52 indicates the expected value of SendBroadcastPermission event rate when the ActivityManager event rate is zero. In this context, it means that when ActivityManager events do not occur, SendBroadcastPermission events are expected to happen at a rate of around 47.52.

The coefficient of approximately 0.162 suggests that for every one-unit increase in ActivityManager event rate, the SendBroadcastPermission event rate increases by approximately 0.162 units.

From the R-squared value which I got, I will say that there is a moderate level of predictive power in explaining the relationship between the event rates of SendBroadcastPermission and ActivityManager using a linear model.

Also, the residual plot, the residuals should appear random, scattered around the horizontal line at zero. A clustered or non-random pattern may indicate problems with the model. I will say the model is not perfect in this case someone might want to check some other regression or models.