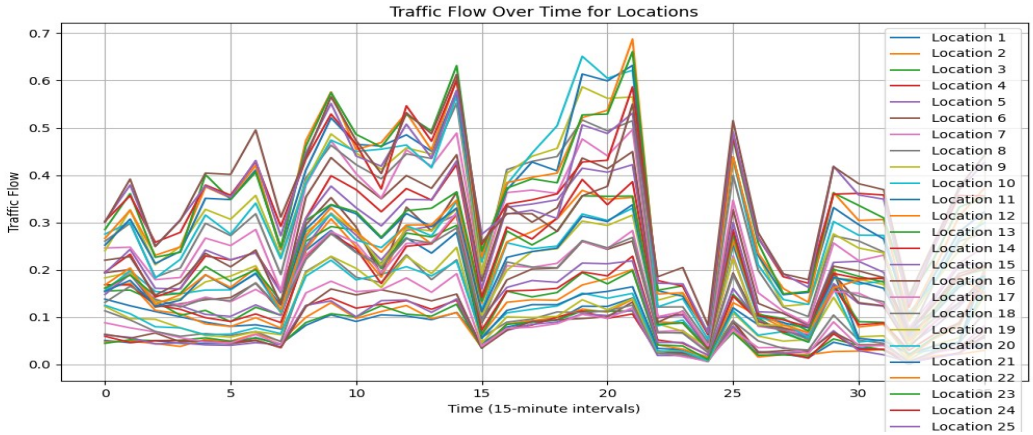
Peter Muller

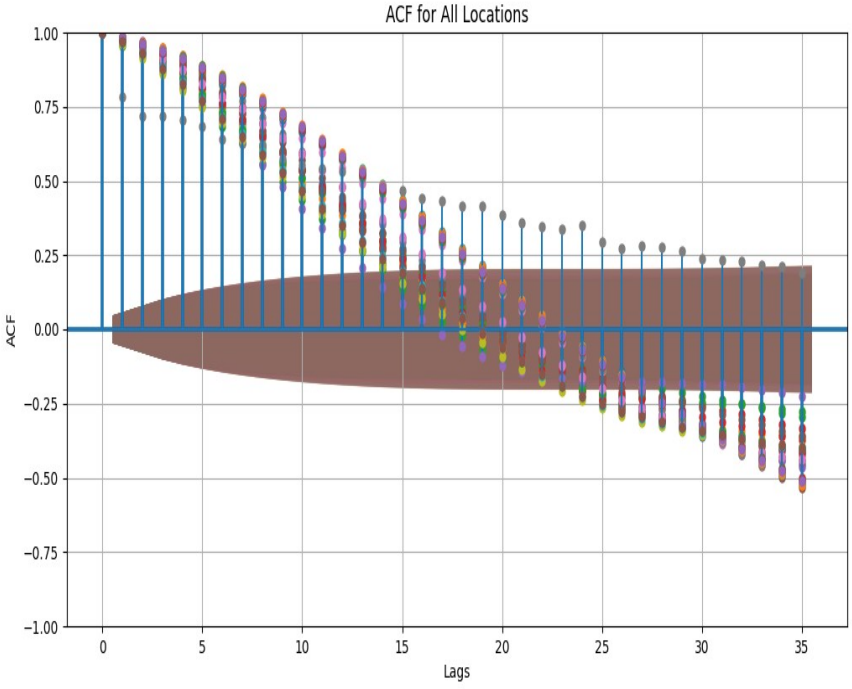
Pmuller

Fundamentals of Operationalizing AI

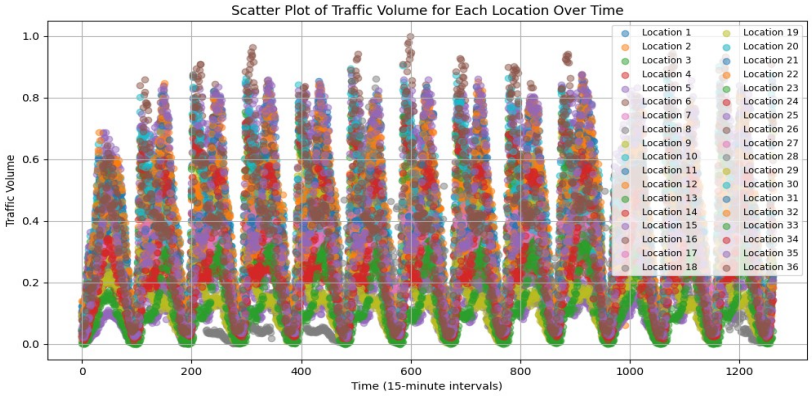
HW1, PR1

Phase 2 Visualizations:

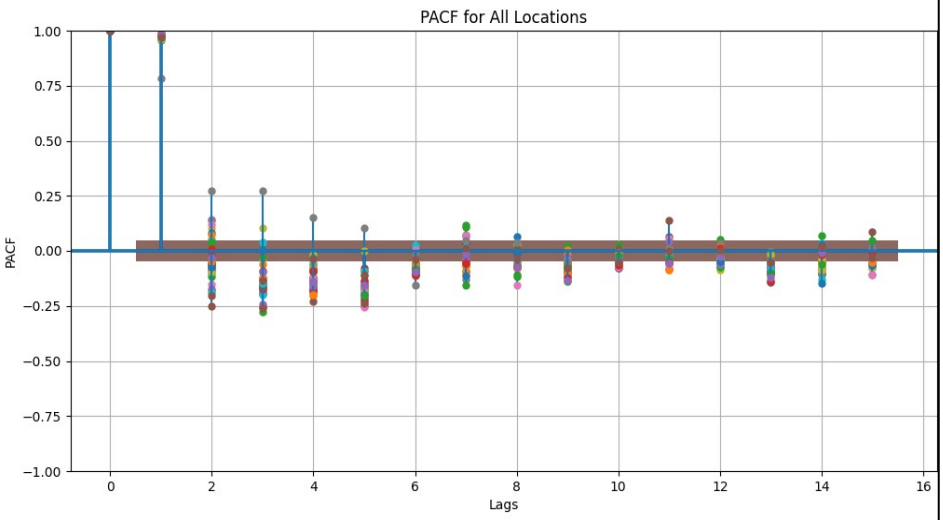



**Negative Autocorrelation**: A negative slope in the ACF indicates that as the lag increases, the correlation between the current value and the value at that lag decreases. This means that high values tend to be followed by low values and vice versa. In other words, there's a tendency for the series to oscillate.

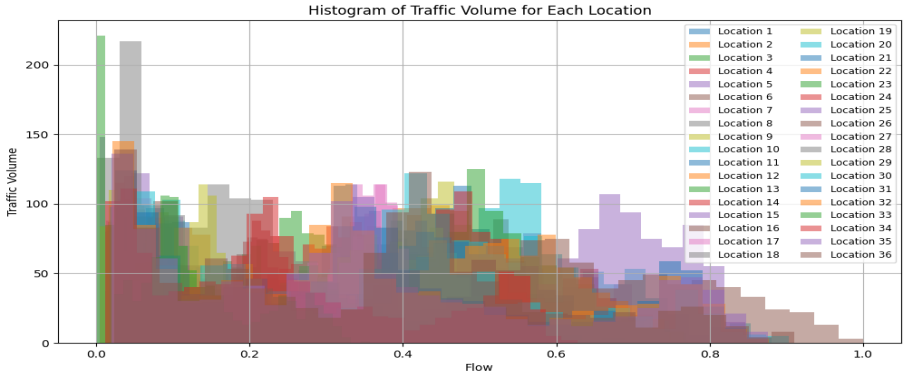



**Scatter plot proves cyclical data that supports ACF/PACF results.**




**Significant Autocorrelation at Short Lags:** This suggests that the relationship between the current observation and its first few predecessors is significant.

**Cut-off After Initial Lags**: If the PACF quickly drops to near zero after a few lags, it implies that there are no significant correlations beyond those initial lags.




The linear regression model is a good model for this dataset, as evidenced by a high R-squared value and statistically significant coefficients. Residual analysis indicates that the model appropriately captures the data structure without systematic errors, and metrics such as MAE and RMSE confirm its accuracy. Overall, the model's performance and the visual comparisons of actual versus predicted values suggest that linear regression is a suitable choice for this dataset.

According to the data, traffic flow is variable to specific hour increments with 3-4 hour breaks between, showing that there is a rush hour at the 15th and 20-23rd 15 minute increment. We believe this corresponds to people going to work and then lunch rushes.