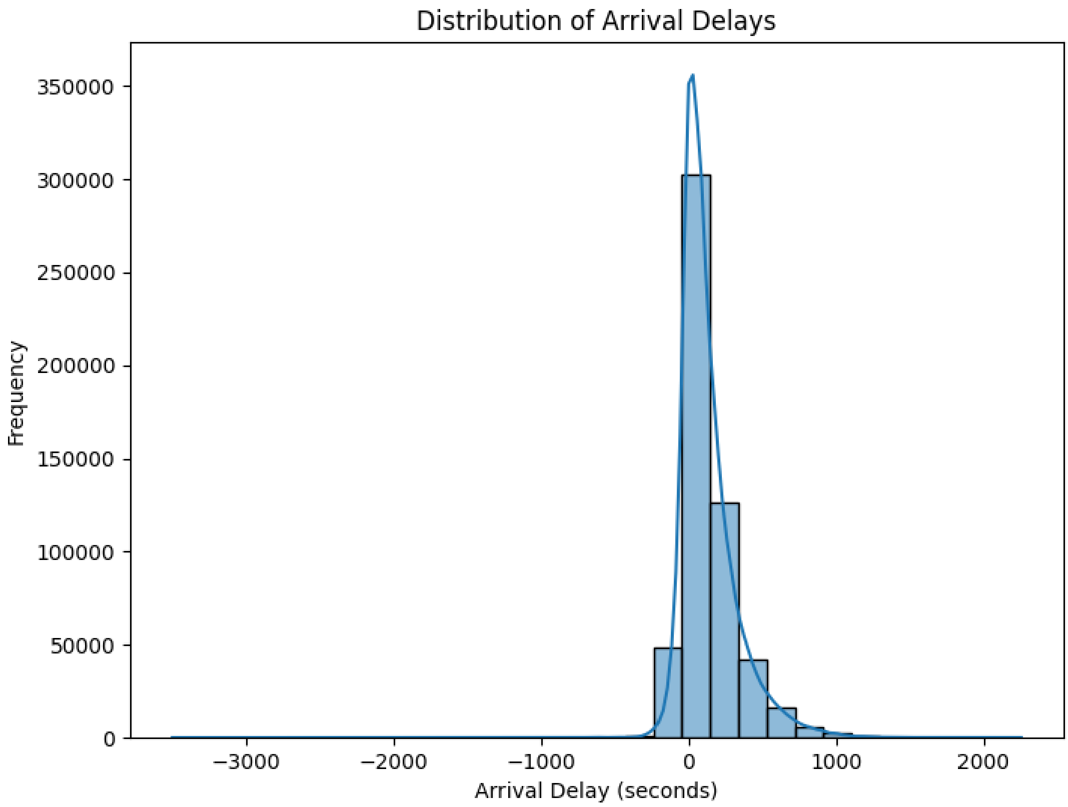
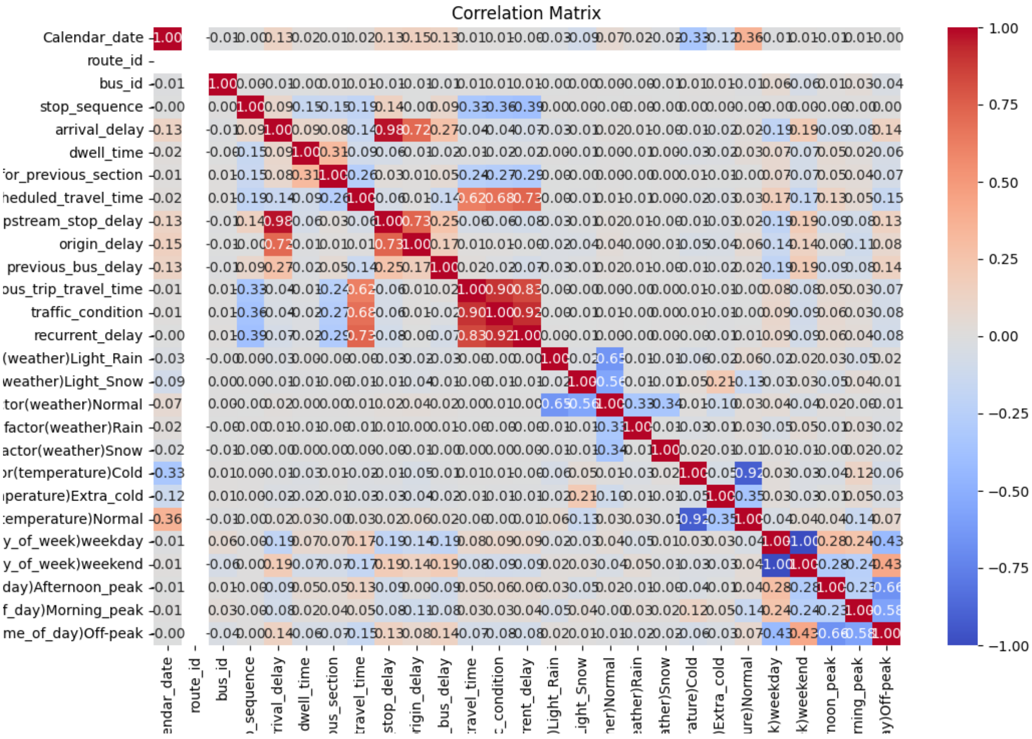
## Data Exploration

This section summarizes the results of the exploratory data analysis (EDA) conducted on the dataset concerning bus delays across various factors, including stop sequences, day types, times of day, weather conditions, and temperature. The insights derived from this analysis are essential for understanding the operational dynamics of the public transportation system and identifying areas for improvement.

### Overview of Average Delays

The overall average delay across all observations in the dataset is approximately 134.97 seconds. This average is calculated across all bus routes and timeframes, providing a baseline understanding of the extent of delays experienced in the system. An average delay exceeding two minutes indicates potential challenges in maintaining on-time performance, which can significantly affect passenger satisfaction and overall transit reliability. This metric serves as a critical reference point for further analyses and comparisons.

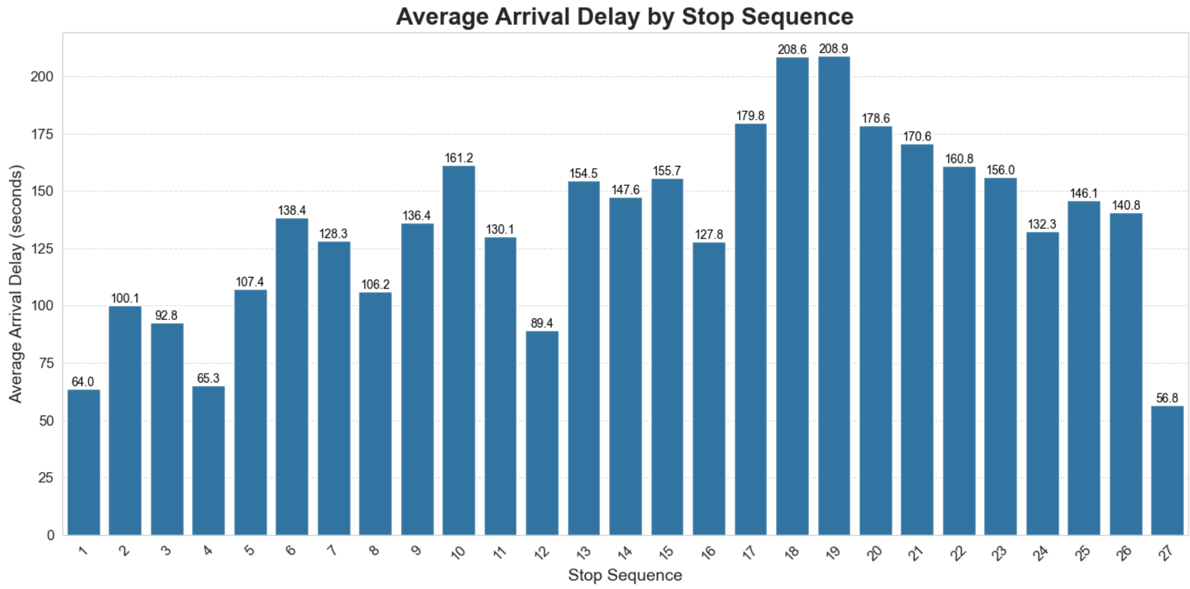




The variables significantly correlated with bus arrival delay provide valuable insights into the factors influencing delays. The most highly correlated variable is upstream_stop_delay (0.975), which indicates that delays at previous stops strongly predict delays at subsequent stops. Origin_delay (0.716) also shows a substantial positive correlation, meaning that delays at the starting point of the bus route contribute to delays throughout the journey. Additionally, previous_bus_delay (0.265) suggests that delays of earlier buses on the same route have a moderate impact on subsequent bus delays. The day_of_week factor reveals that weekends are associated with slightly increased delays (0.190), while off-peak times (0.138) show a smaller positive correlation, meaning buses tend to be delayed during less busy periods. Scheduled_travel_time (-0.142) is negatively correlated with delays, indicating that longer scheduled travel times may allow for more slack, reducing the likelihood of delays. Finally, weekday service is mildly negatively correlated with delays (-0.191), implying that buses are slightly more punctual on weekdays compared to weekends. These variables help in understanding the primary drivers of bus delays and can be useful for predictive modeling

Delays by Stop Sequence

The analysis of average delays at each bus stop reveals significant variability:




Stop 1 (the first stop in the sequence) experiences an average delay of 63.96 seconds, while Stop 27 (the last stop) shows a delay of 56.8 seconds. These lower average delays at the start and end stops indicate relatively efficient operations, likely due to less congestion and reduced passenger loading at these points in the route.

In contrast, the delays fluctuate significantly through the intermediate stops, peaking at 208.89 seconds by Stop 19. This sharp increase in average delay suggests that cumulative delays may be exacerbated by several factors, including:

Increased passenger loading: As buses pick up more passengers at subsequent stops, dwell times can increase, leading to longer delays.

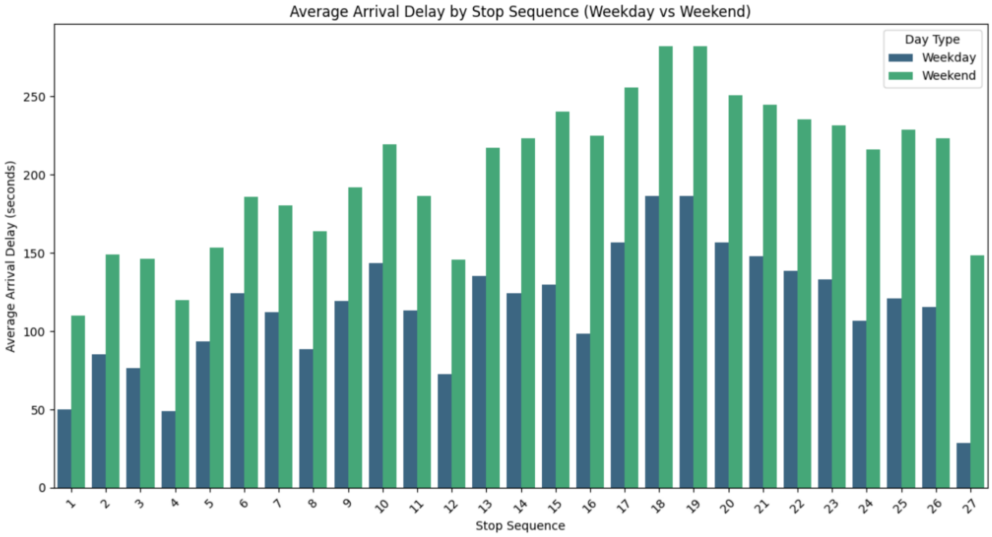
Traffic conditions: Stops located in areas with higher traffic volumes or congestion may contribute to longer delays.

Previous delays: If a bus has encountered delays at earlier stops, this can compound as it progresses through the route.

This variability highlights the need for targeted operational improvements at specific stops to enhance overall service efficiency.

Delays by Day of the Week

The analysis illustrates marked differences in average delays based on the day of the week:




Weekdays (Monday to Friday) have an average delay of 114.59 seconds, reflecting a more predictable and consistent service pattern during high-demand periods when ridership is typically at its peak.

In contrast, weekends report a substantially higher average delay of 202.11 seconds. This difference may suggest several underlying issues:

Reduced service frequency: Weekend schedules often feature fewer buses, which can lead to longer wait times and increased passenger loading at stops.

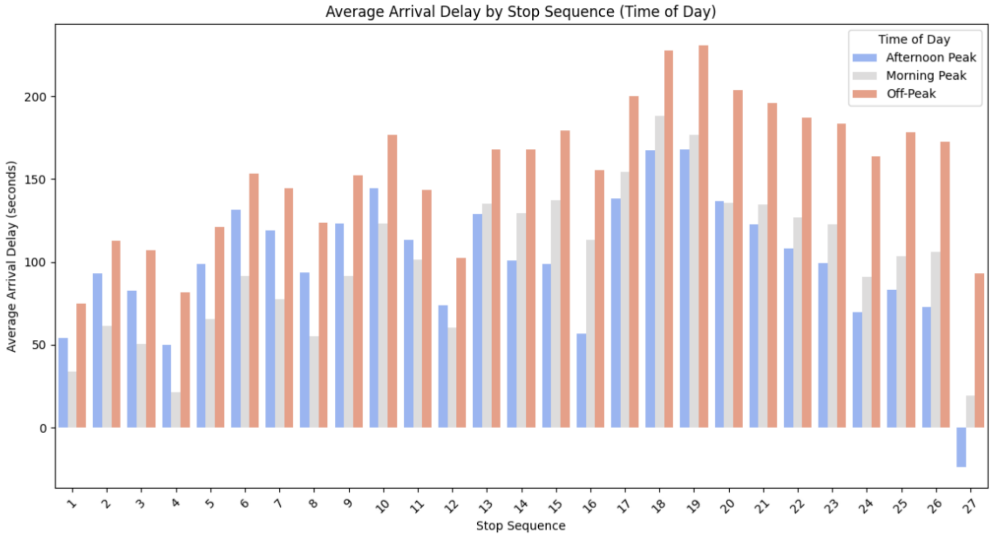
Traffic conditions: Weekend traffic patterns may differ, with increased congestion in popular areas, contributing to delays.

Operational challenges: Staffing and resource allocation might differ on weekends, impacting service reliability.

The pronounced disparity between weekday and weekend delays warrants further investigation into weekend operations to identify potential improvements and strategies to enhance service reliability.

Delays by Time of Day

When examining delays based on the time of day, distinct patterns emerge:




The Afternoon Peak (AP) period experiences an average delay of 100.25 seconds, while the Morning Peak (MP) records a slightly higher average delay of 100.35 seconds. These delays are relatively consistent with expectations for peak travel times, where high ridership can lead to congestion and operational challenges.

Conversely, the Off-Peak (OP) hours show a considerable average delay of 155.65 seconds, indicating potential inefficiencies during these less busy times. Possible factors contributing to these delays include:

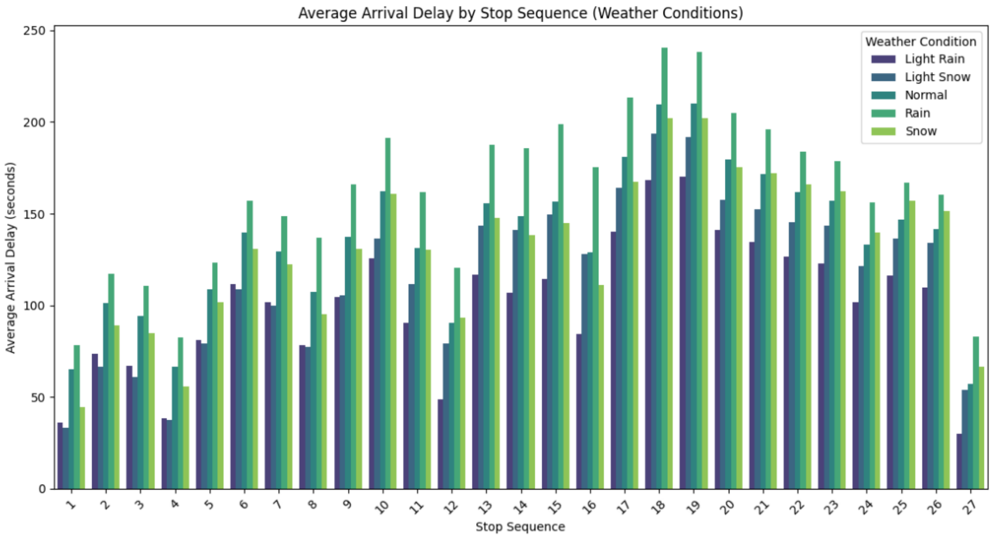
Reduced service frequency: Off-peak services may operate with fewer buses, leading to longer wait times and increased delays.

Increased traffic: Even during off-peak hours, traffic conditions may still contribute to delays, particularly in congested urban areas.

Addressing the causes of off-peak delays could enhance service reliability and improve passenger satisfaction, particularly for those traveling during these hours.

Delays by Weather Conditions

Weather conditions significantly impact bus delays, as evidenced by the analysis:



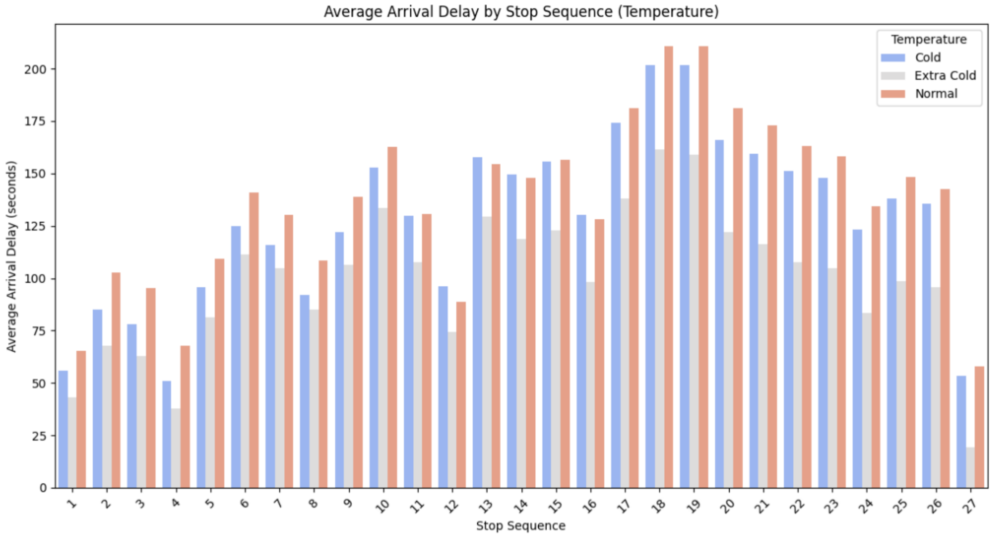

Light Rain conditions correlate with an average delay of 101.52 seconds, while Light Snow records an average delay of 116.78 seconds. These weather-related delays may arise from both traffic disruptions and operational challenges that buses face in adverse conditions.

Notably, Rain results in an average delay of 161.58 seconds, indicating a substantial impact of adverse weather on transit performance. The average delay for Snow conditions stands at 131.27 seconds, further underscoring the challenges posed by winter weather on bus operations.

These findings suggest that strategies to improve service during adverse weather—such as enhancing communication with passengers, adjusting schedules, or implementing contingency plans—could be beneficial in mitigating delays and maintaining service reliability.

Delays by Temperature

Temperature also plays a role in shaping bus delays, revealing distinct patterns:




Cold conditions are associated with an average delay of 127.58 seconds, which may reflect the impact of winter weather on transit operations, including slower travel speeds and potential safety concerns for both passengers and drivers.

Interestingly, Extra Cold conditions report a lower average delay of 99.67 seconds. This trend may suggest that extremely cold temperatures lead to reduced transit operations, fewer riders, or more efficient service due to lower passenger volumes.

Understanding how temperature impacts delays can guide transit authorities in making informed decisions regarding service adjustments during extreme weather conditions.

The exploratory analysis of the dataset reveals critical insights into the factors affecting bus delays, highlighting the complexity of public transportation operations. Variations based on stop sequences, day types, time of day, weather, and temperature underscore the challenges that transit authorities face in delivering timely and efficient service.

These findings are essential for informing strategies aimed at improving bus service efficiency, particularly during peak demand times and adverse weather conditions. Future analyses should delve deeper into these relationships, focusing on operational strategies to mitigate delays and enhance overall service reliability. By addressing the factors contributing to delays, transit authorities can better serve their passengers, optimize resource allocation, and improve overall satisfaction with public transportation services.