

Delivery Performance and Environmental Impact Data Summary

Delivery Performance and Environmental Impact Data Summary

About the dataset

This dataset contains information on 1586 delivery records, encompassing various aspects of the delivery process, from pickup to completion. Key columns include `delivery_id`, `customer_id`, `delivery_date`, `pickup_hour`, `time_of_day`, `delivery_status`, `on_time_status`, `actual_delivery_ts`, `driver_name`, `experience_level`, `truck_model`, `emission_standard`, `last_maintenance_date`, `pickup_location`, `delivery_location`, `region`, `traffic_index`, `traffic_condition`, `cargo_type`, `cargo_load_category`, `trip_distance_km`, `delivery_duration_hours`, `avg_speed_kmh`, `fuel_efficiency_km_per_liter`, `total_co2_emitted_kg`, `co2_efficiency_kg_per_km`, and `alert_type`.

Analyzing the first five sample data entries, we observe deliveries spanning from March 26, 2025. The deliveries involve different drivers with varying experience levels (Senior and Mid-Level), truck models (Model-B and Model-C), and emission standards (Euro 6, BS4, and BS6). The pickup hours range from night to evening. Some deliveries were completed on time, while others were delayed. The `trip_distance_km` varies significantly, from 70 to 400 km, impacting `delivery_duration_hours` and `total_co2_emitted_kg`. The `alert_type` column indicates instances of "HighEngineTemp" and "Speeding".

The dataset provides a comprehensive view of delivery operations, including environmental impact metrics such as `fuel_efficiency_km_per_liter`, `total_co2_emitted_kg`, and `co2_efficiency_kg_per_km`. The average `trip_distance_km` is 226.38 km, with an average `delivery_duration_hours` of 4.50. The average `avg_speed_kmh` is 51.63, and the average `fuel_efficiency_km_per_liter` is 0.67. The average `total_co2_emitted_kg` is 1403.12, and the average `co2_efficiency_kg_per_km` is 5.26. These metrics can be used to assess and optimize delivery efficiency and environmental performance.

Relevant Inquiries

Q1.What is the monthly trend of delivery completion rate? Show YoY comparison.

year	month	completion_rate_percent
2025	3	45.0980392157
2025	4	51.7374517375
2025	5	56.106870229
2025	6	54.8148148148
2025	7	53.5460992908
2025	8	51.8382352941
2025	9	51.5789473684

Monthly Completion Rates

- Year:** The data exclusively covers the year **2025**.
- Months Covered:** Delivery completion rates are available for months **March (3) through September (9)**.
- Average Completion Rate:** The average monthly completion rate across these months in 2025 was **52.10%**.
- Rate Fluctuation:** The completion rate varied from a minimum of **45.10%** (March) to a maximum of **56.11%** (May).

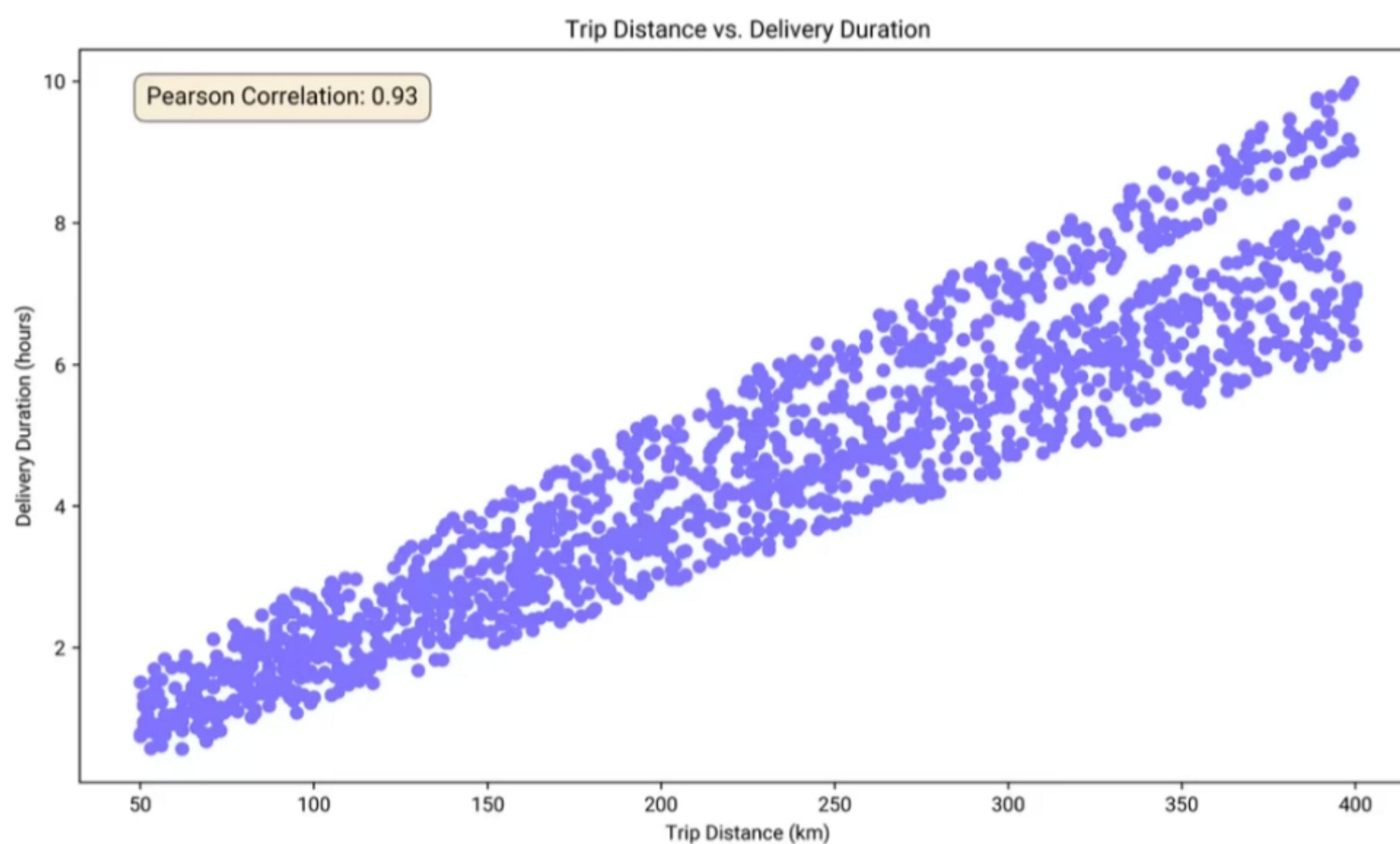
Detailed Monthly Breakdown (2025)

- **March:** The completion rate was **45.10%**.
- **April:** The rate increased to **51.74%**.
- **May:** The rate peaked at **56.11%**.
- **June:** The rate slightly decreased to **54.81%**.
- **July:** A further decrease was observed, with the rate at **53.55%**.
- **August:** The rate continued to decline to **51.84%**.
- **September:** The rate remained relatively stable at **51.58%**.

Conclusion and Insights

- **Limited Data for YoY Comparison:** A Year-over-Year (YoY) comparison cannot be performed as the provided data only includes delivery completion rates for a single year, 2025. To conduct a YoY analysis, data from multiple years would be required.
- **Monthly Trend in 2025:** For the year 2025, the delivery completion rate showed an initial **increase from March (45.10%) to a peak in May (56.11%)**, followed by a **gradual decline from June (54.81%) to September (51.58%)**. This suggests a potential seasonal pattern or operational changes influencing delivery efficiency within the year.

Q2. What is the correlation coefficient between trip distance and delivery duration?



Pearson Correlation Coefficient

- **Correlation Value:** The Pearson correlation coefficient between 'trip_distance_km' and 'delivery_duration_hours' is **0.933**. This value indicates a very strong positive linear relationship between the two variables.
- **Visual Representation:** The scatter plot visually confirms this strong positive correlation, with data points generally trending upwards from left to right, indicating that as trip distance increases, delivery duration also tends to increase. The correlation coefficient is explicitly displayed on the plot as **0.93**.

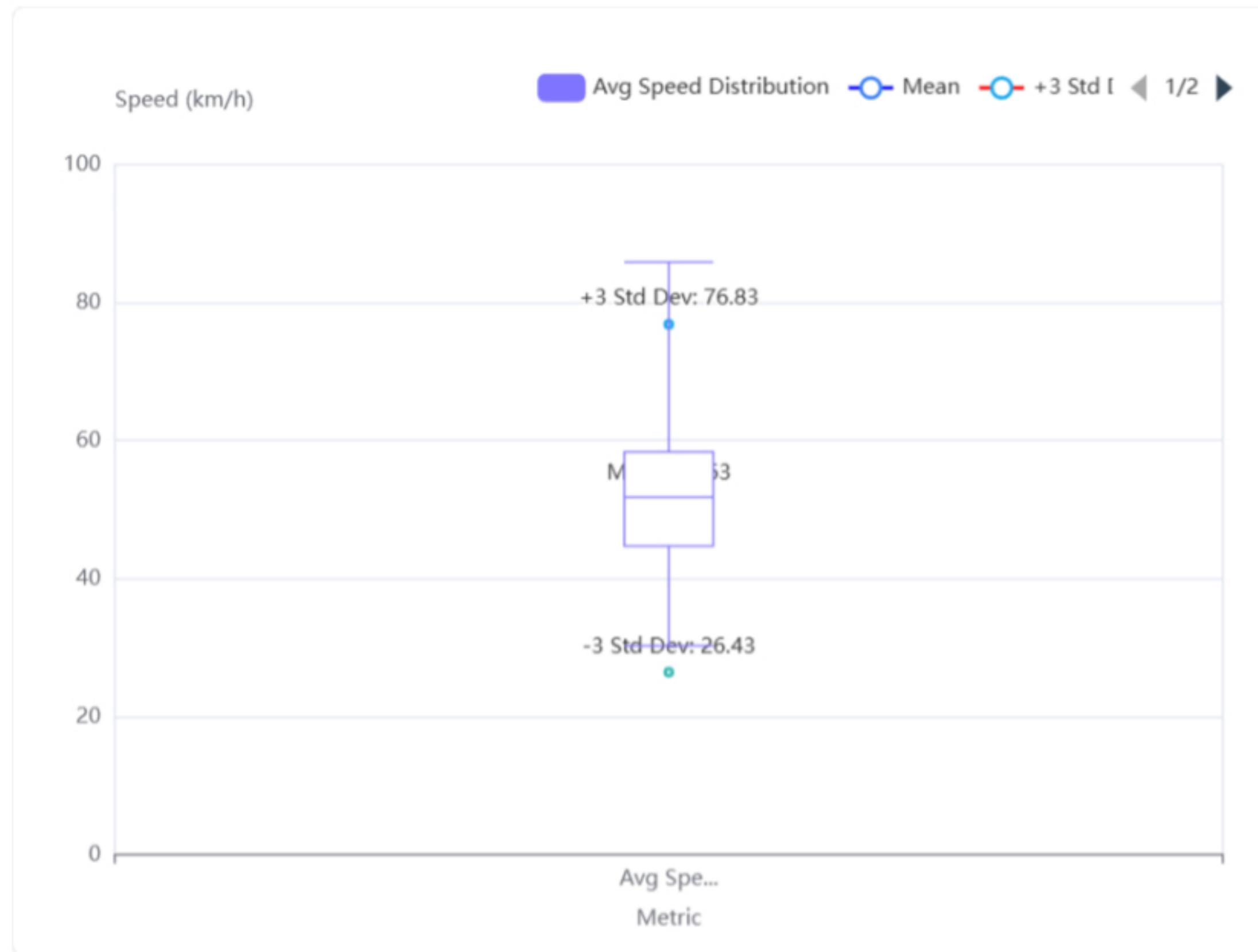
Data Characteristics

- **Trip Distance Range:** The trip distances observed range from **50.00 km to 400.00 km**, with an average of 226.38 km.
- **Delivery Duration Range:** The delivery durations range from **0.57 hours to 9.98 hours**, with an average of 4.50 hours.

Conclusion and Insights

- **Strong Positive Relationship:** There is a **very strong positive linear relationship** between trip distance and delivery duration, as evidenced by a Pearson correlation coefficient of approximately **0.93**. This means that longer trip distances are consistently associated with longer delivery durations.
- **Predictive Power:** The high correlation suggests that trip distance is a **significant predictor** of delivery duration. This insight can be valuable for logistics planning, estimated time of arrival (ETA) calculations, and resource allocation.

Q3. Which deliveries have average speeds outside of 3 standard deviations from the mean?



Outlier Identification Criteria

- **Mean Average Speed:** The average speed across all deliveries is **51.63 km/h**.
- **Standard Deviation of Average Speed:** The variability in average speed is **8.40 km/h**.
- **Outlier Thresholds:** Deliveries are considered outliers if their average speed falls outside the range of **26.43 km/h** (mean - 3 * std dev) and **76.83 km/h** (mean + 3 * std dev).
- **Identification Flag:** The `is_outlier` column in the `outlier_data` dataset is set to `True` for any delivery whose `avg_speed_kmh` is below 26.43 km/h or above 76.83 km/h.

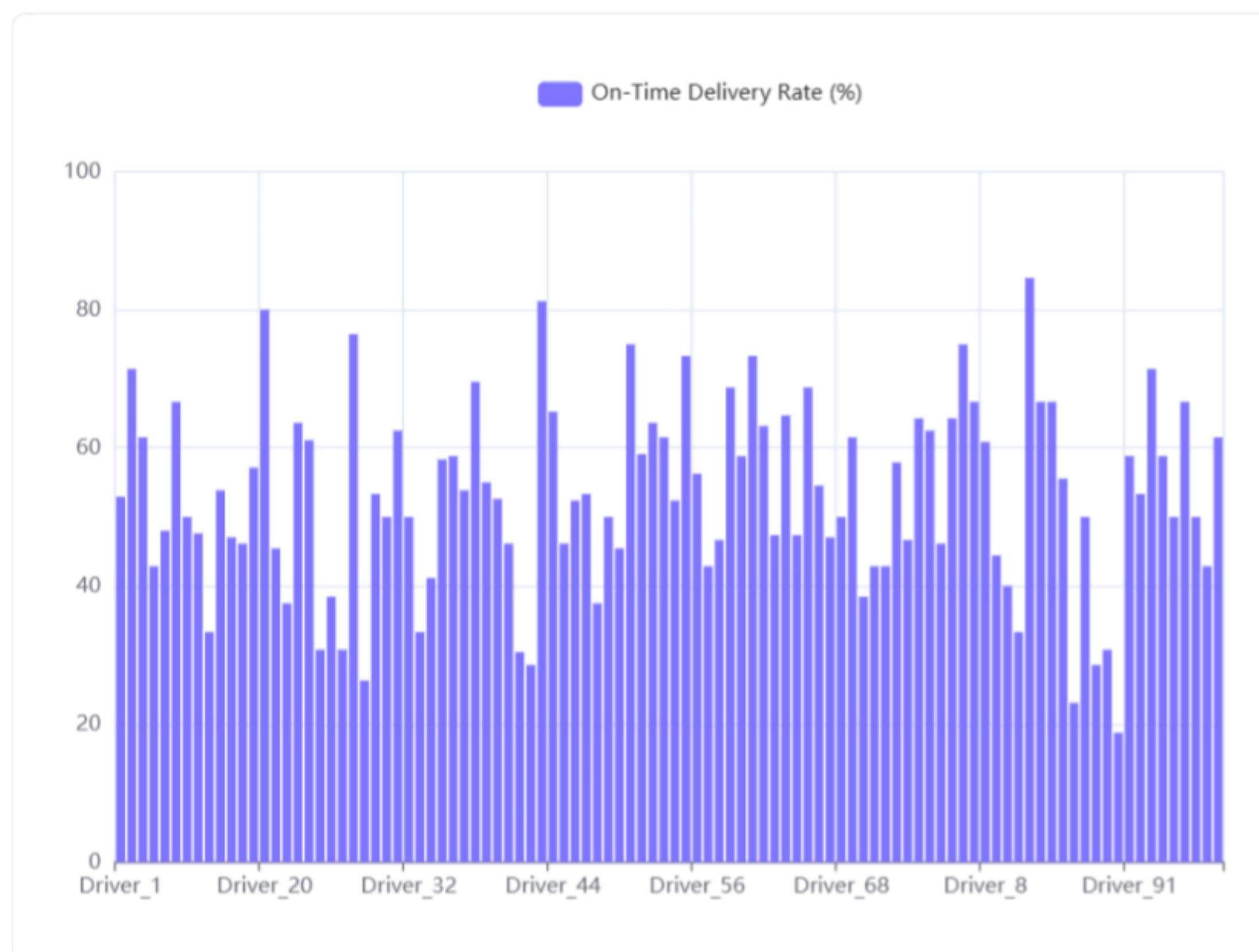
Visual Representation of Outliers

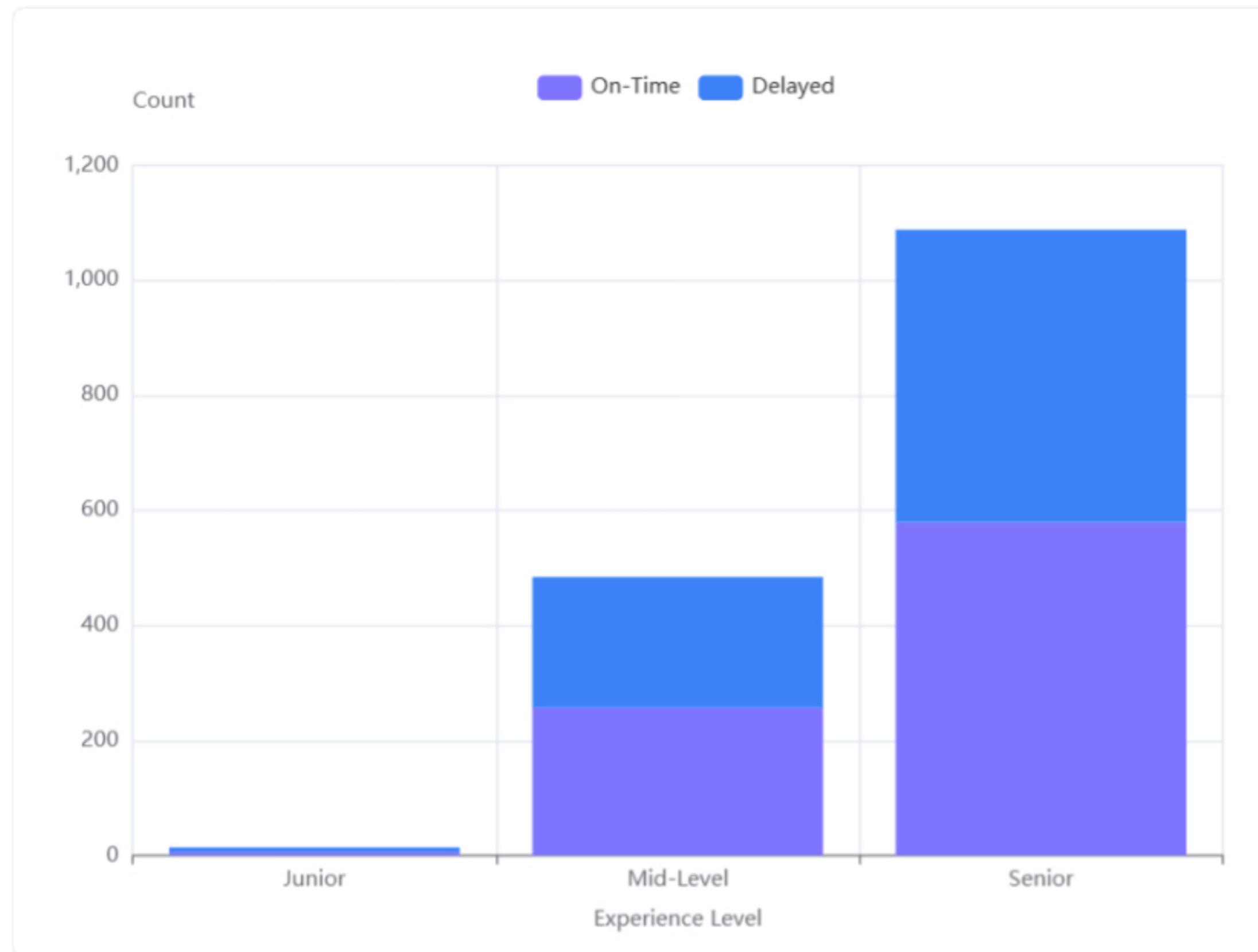
- **Box Plot Visualization:** A box plot of `avg_speed_kmh` visually confirms the calculated mean and the +/- 3 standard deviation thresholds.
- **Threshold Values:** The plot explicitly labels the +3 standard deviation threshold at **76.83 km/h** and the -3 standard deviation threshold at **26.43 km/h**. The mean is indicated at approximately **51.63 km/h**.

Conclusion and Insights

- **Outlier Definition:** Deliveries with average speeds significantly deviating from the norm, specifically those falling outside **3 standard deviations** from the mean, are flagged as outliers. This corresponds to average speeds **below 26.43 km/h or above 76.83 km/h**.
- **Identification Method:** The `is_outlier` column in the `outlier_data` dataset serves as a direct indicator for these specific deliveries. Any `delivery_id` associated with `is_outlier` being `True` meets the criteria.
- **Data Observation:** Based on the provided sample data, **no specific delivery IDs are shown to be outliers** (all sample records have `is_outlier` as `False`). To identify the exact deliveries, a full scan of the `outlier_data` for `is_outlier = True` would be required.

Q4. What is the on-time delivery rate for each driver? Show the distribution of on-time status by driver experience level.





On-Time Delivery Rates by Individual Driver

- Wide Range of Performance:** The on-time delivery rates among drivers vary significantly, ranging from a **minimum of 18.75% to a maximum of 84.62%**.
- Average Performance:** The average on-time delivery rate across all drivers is approximately **53.05%**, with a standard deviation of 13.69%, indicating a considerable spread in individual driver performance.
- Top Performers:** Drivers like **Driver_83 (84.62%)**, **Driver_43 (81.25%)**, and **Driver_20 (80.0%)** demonstrate exceptionally high on-time delivery rates.
- Underperformers:** Conversely, drivers such as **Driver_90 (18.75%)**, **Driver_87 (23.08%)**, and **Driver_29 (26.32%)** show the lowest on-time delivery rates, indicating potential areas for improvement or investigation.

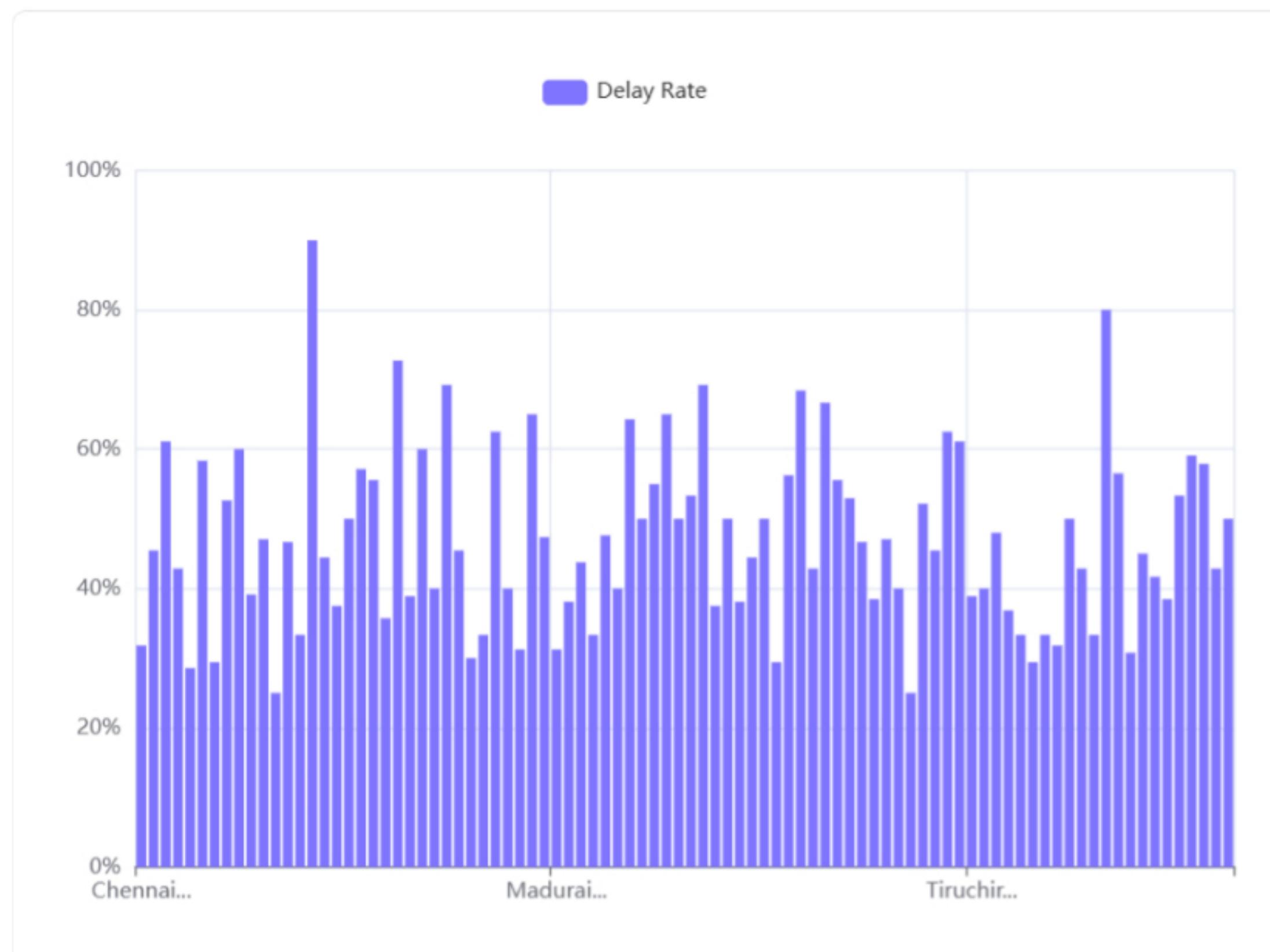
On-Time Status Distribution by Driver Experience Level

- Junior Drivers' Performance:** Junior drivers have the lowest on-time delivery percentage at **42.86%**, with a higher proportion of delayed deliveries (57.14%). This group also has the fewest total deliveries (6 on-time, 8 delayed).
- Mid-Level Drivers' Performance:** Mid-Level drivers show a better on-time performance with **53.10%** of their deliveries being on-time, and 46.90% delayed. They account for a moderate volume of deliveries (257 on-time, 227 delayed).
- Senior Drivers' Performance:** Senior drivers exhibit the highest on-time delivery percentage at **53.22%**, with 46.78% delayed. This group handles the largest volume of deliveries (579 on-time, 509 delayed), suggesting that experience correlates with a slightly higher on-time rate and a greater capacity for deliveries.

Conclusion and Insights

- Individual Driver Variability:** There is a **substantial difference in on-time delivery performance among individual drivers**, with rates ranging from under 20% to over 80%. This highlights the importance of individual driver performance monitoring and potential targeted training or incentives.
- Experience Level Correlation:** While the difference is not drastic, **more experienced drivers (Mid-Level and Senior) tend to have slightly higher on-time delivery rates** compared to Junior drivers. Senior drivers also handle a significantly larger volume of deliveries while maintaining a good on-time rate.
- Opportunity for Improvement:** The overall average on-time delivery rate is around 53%, indicating that nearly half of all deliveries are delayed. This suggests a systemic opportunity to improve delivery efficiency across the board.
- Targeted Interventions:** Focusing on improving the performance of **Junior drivers** and those individual drivers with **exceptionally low on-time rates** could lead to a significant overall improvement in delivery service.

Q5.What is the delivery delay rate for each route (pickup_location to delivery_location)?



Overall Delay Rate Distribution

- **Average Delay Rate:** The average delivery delay rate across all routes is **47.03%**, with a standard deviation of 12.93%.
- **Range of Delay Rates:** The delay rates for individual routes vary significantly, ranging from a **minimum of 25.00% to a maximum of 90.00%**.

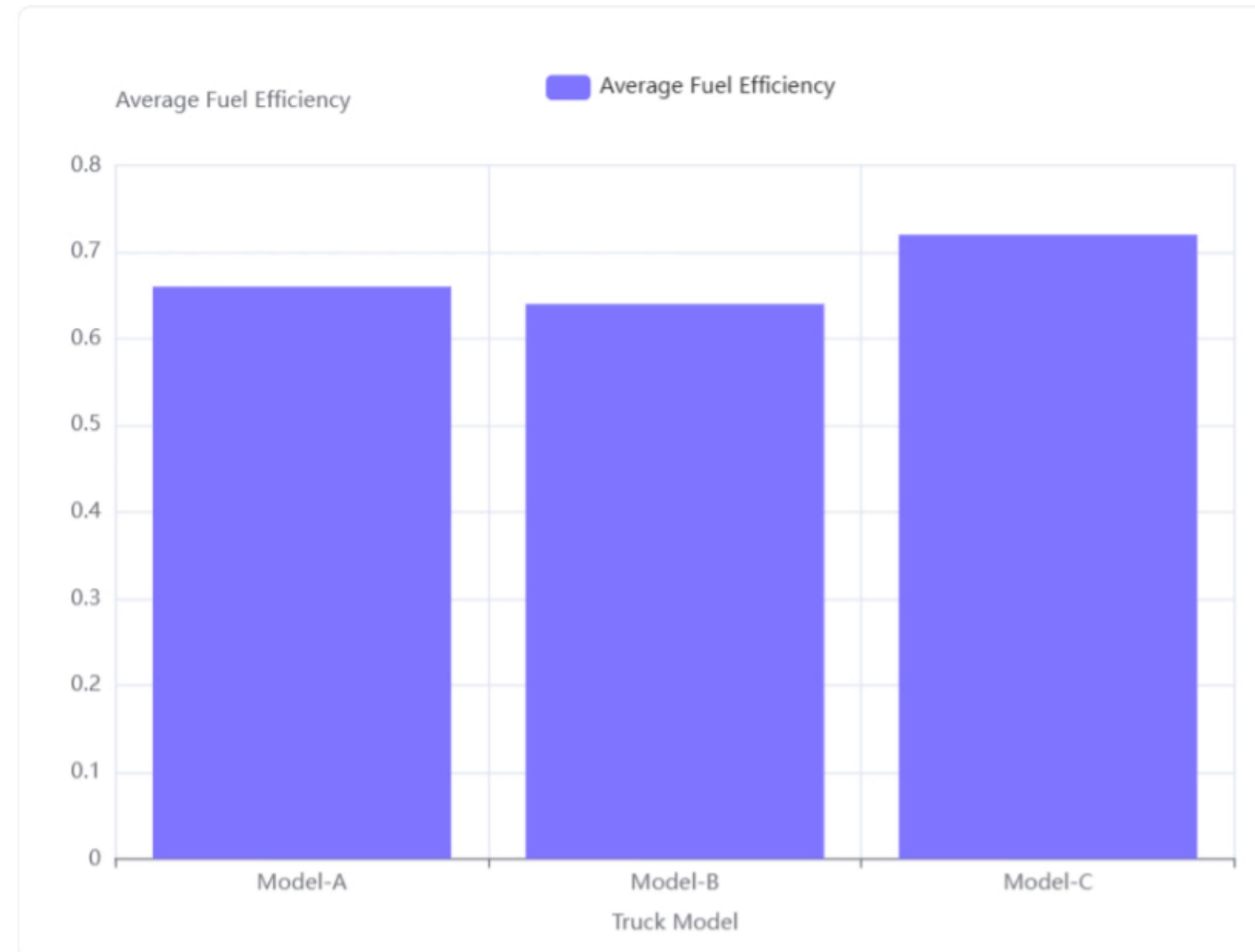
Specific Route Delay Rates

- **Highest Observed Delay Rate:** The highest delay rate identified is **90%** for at least one route (e.g., "Tiruchirappalli to Tiruchirappalli" or "Tiruchirappalli to Nagercoil" based on visual inspection and OCR, though specific route names are hard to definitively link to the 90% bar from the OCR).
- **Lowest Observed Delay Rate:** The lowest delay rate identified is **25%** for at least one route (e.g., "Chennai to Salem" or "Coimbatore to Thoothukudi" based on visual inspection and OCR).
- **Example Routes and Their Delay Rates:**
 - **Chennai to Madurai:** This route experiences a delay rate of **61.11%**.
 - **Chennai to Kanyakumari:** The delay rate for this route is **45.45%**.
 - **Chennai to Nagercoil:** This route has a delay rate of **42.85%**.
 - **Chennai to Coimbatore:** The delay rate for this route is **31.82%**.
 - **Chennai to Salem:** This route shows a delay rate of **28.57%**.
 - Other routes also exhibit varying delay rates, such as **72.72%, 69.23%, 62.5%, and 50%** as seen in the visualization.

Conclusion and Insights

- **Significant Variability in Delay Rates:** There is a **wide disparity in delivery delay rates across different routes**, ranging from 25% to 90%. This suggests that some routes are significantly more prone to delays than others.
- **Identification of High-Risk Routes:** Routes with delay rates approaching 90% represent **critical bottlenecks or operational challenges** that require immediate attention to understand the underlying causes and implement targeted improvements.
- **Identification of Low-Risk Routes:** Routes with lower delay rates (e.g., 25-30%) indicate **relatively efficient operations**, which could serve as benchmarks for best practices.
- **Actionable Insights for Route Optimization:** The detailed delay rates for each route provide **actionable data for logistics and operations teams** to prioritize interventions, such as optimizing routes, improving infrastructure, or adjusting delivery schedules for high-delay routes.

Q6. What is the average fuel efficiency for each truck model?



Average Fuel Efficiency Values

- **Model-A:** The average fuel efficiency for **Model-A** trucks is **0.656 km/liter**.
- **Model-B:** The average fuel efficiency for **Model-B** trucks is **0.644 km/liter**.
- **Model-C:** The average fuel efficiency for **Model-C** trucks is **0.717 km/liter**.

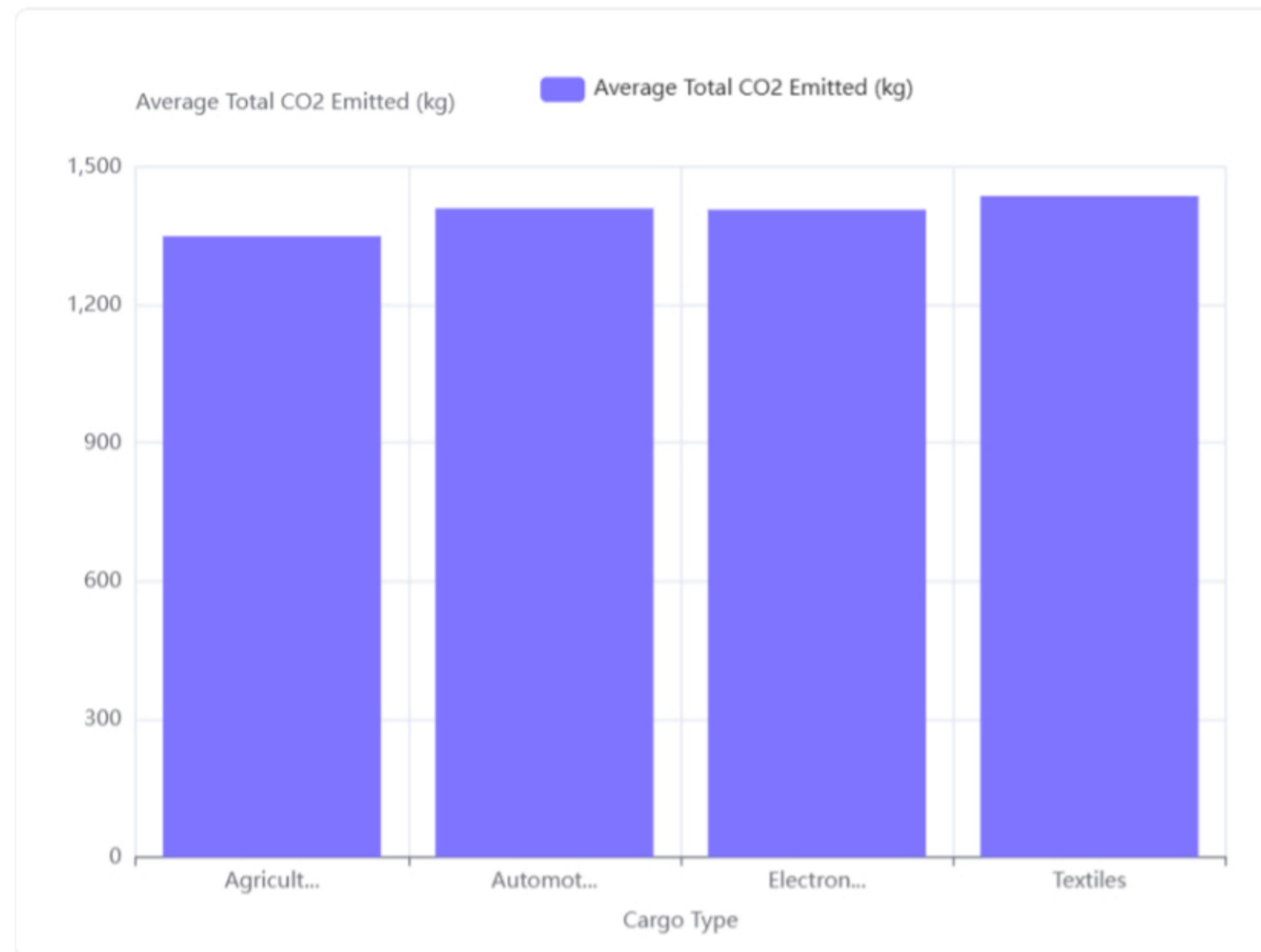
Visual Representation of Fuel Efficiency

- **Comparative Overview:** A bar chart visually represents the average fuel efficiency for each truck model, showing **Model-C** as the most efficient and **Model-B** as the least efficient among the three.
- **Model-A Efficiency:** The bar for Model-A indicates an average fuel efficiency of approximately **0.66 km/liter**.
- **Model-B Efficiency:** The bar for Model-B indicates an average fuel efficiency of approximately **0.64 km/liter**.
- **Model-C Efficiency:** The bar for Model-C indicates an average fuel efficiency of approximately **0.72 km/liter**.

Conclusion and Insights

- **Model-C is the most fuel-efficient:** **Model-C** demonstrates the highest average fuel efficiency at **0.717 km/liter**, making it the most efficient truck model among the analyzed group.
- **Model-B is the least fuel-efficient:** **Model-B** has the lowest average fuel efficiency at **0.644 km/liter**, indicating it is the least efficient model.
- **Model-A has moderate efficiency:** **Model-A** falls in the middle with an average fuel efficiency of **0.656 km/liter**, performing better than Model-B but less efficiently than Model-C.

Q7. What is the average total CO2 emitted per delivery for each cargo type?



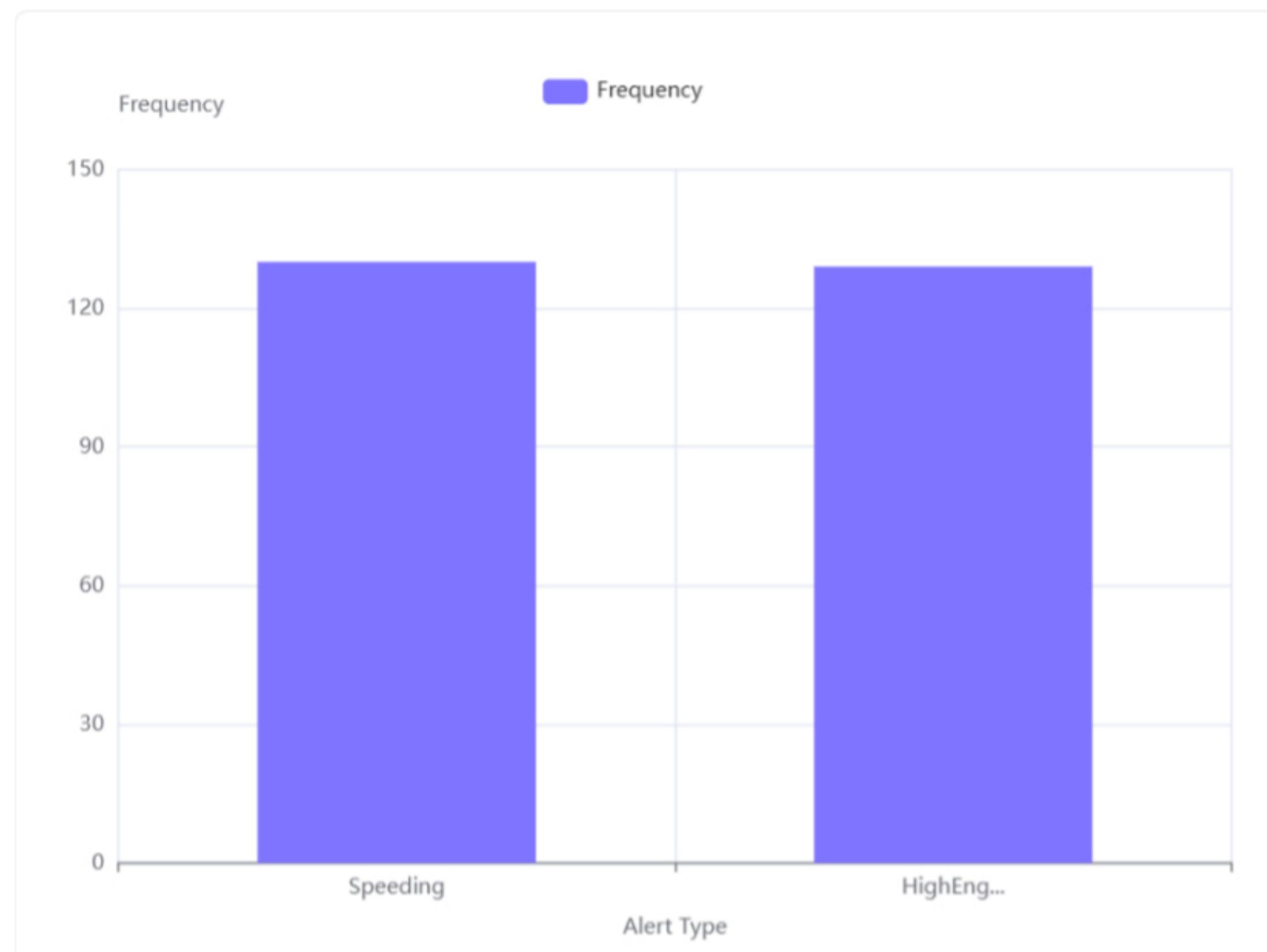
Average CO2 Emissions by Cargo Type

- Agricultural Produce:** The average total CO2 emitted for Agricultural Produce deliveries is **1349.36 kg**.
- Automotive Parts:** The average total CO2 emitted for Automotive Parts deliveries is **1409.75 kg**.
- Electronics:** The average total CO2 emitted for Electronics deliveries is **1406.95 kg**.
- Textiles:** The average total CO2 emitted for Textiles deliveries is **1436.67 kg**.

Conclusion and Insights

- Varying Emissions:** The average total CO2 emitted per delivery varies across different cargo types, ranging from approximately 1349 kg to 1437 kg.
- Highest Emissions:** **Textiles** cargo type has the highest average total CO2 emissions per delivery, at **1436.67 kg**.
- Lowest Emissions:** **Agricultural Produce** cargo type has the lowest average total CO2 emissions per delivery, at **1349.36 kg**.
- Similar Emissions:** Automotive Parts and Electronics cargo types show relatively similar average CO2 emissions, at 1409.75 kg and 1406.95 kg respectively, placing them in the mid-range compared to other cargo types.

Q8.What is the frequency of each alert type?



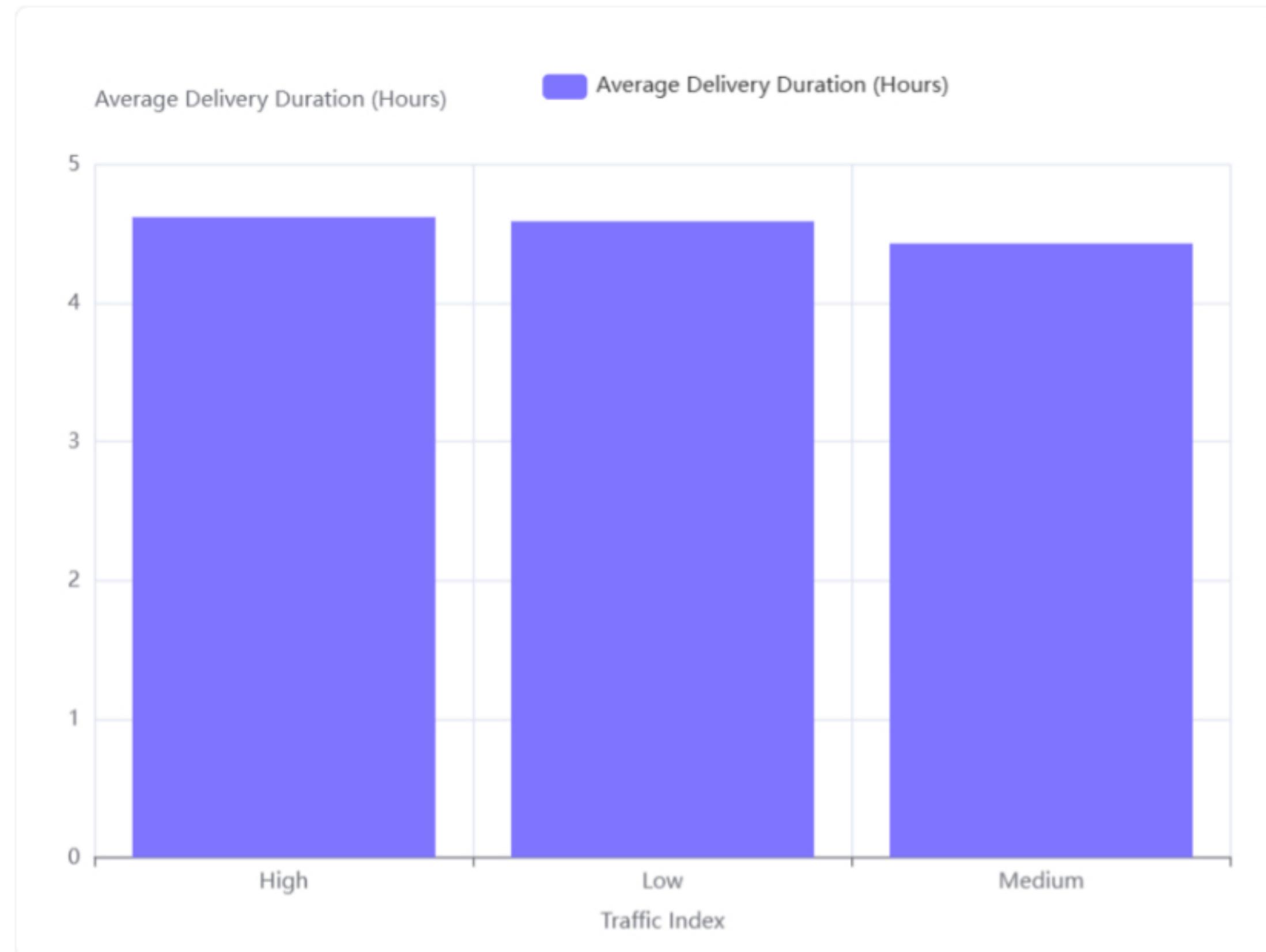
Frequency Distribution of Alert Types

- Speeding Alerts:** The frequency for 'Speeding' alerts is **130**.
- HighEngineTemp Alerts:** The frequency for 'HighEngineTemp' alerts is **129**.

Conclusion and Insights

- Dominant Alert Type: Speeding** is the most frequent alert type, occurring 130 times, slightly more often than 'HighEngineTemp'.
- Near-Equal Distribution:** The two alert types, 'Speeding' and 'HighEngineTemp', have a **very similar frequency**, with only a difference of one occurrence, indicating a relatively balanced distribution of these specific alerts.

Q9.How does delivery duration vary across different traffic index levels?



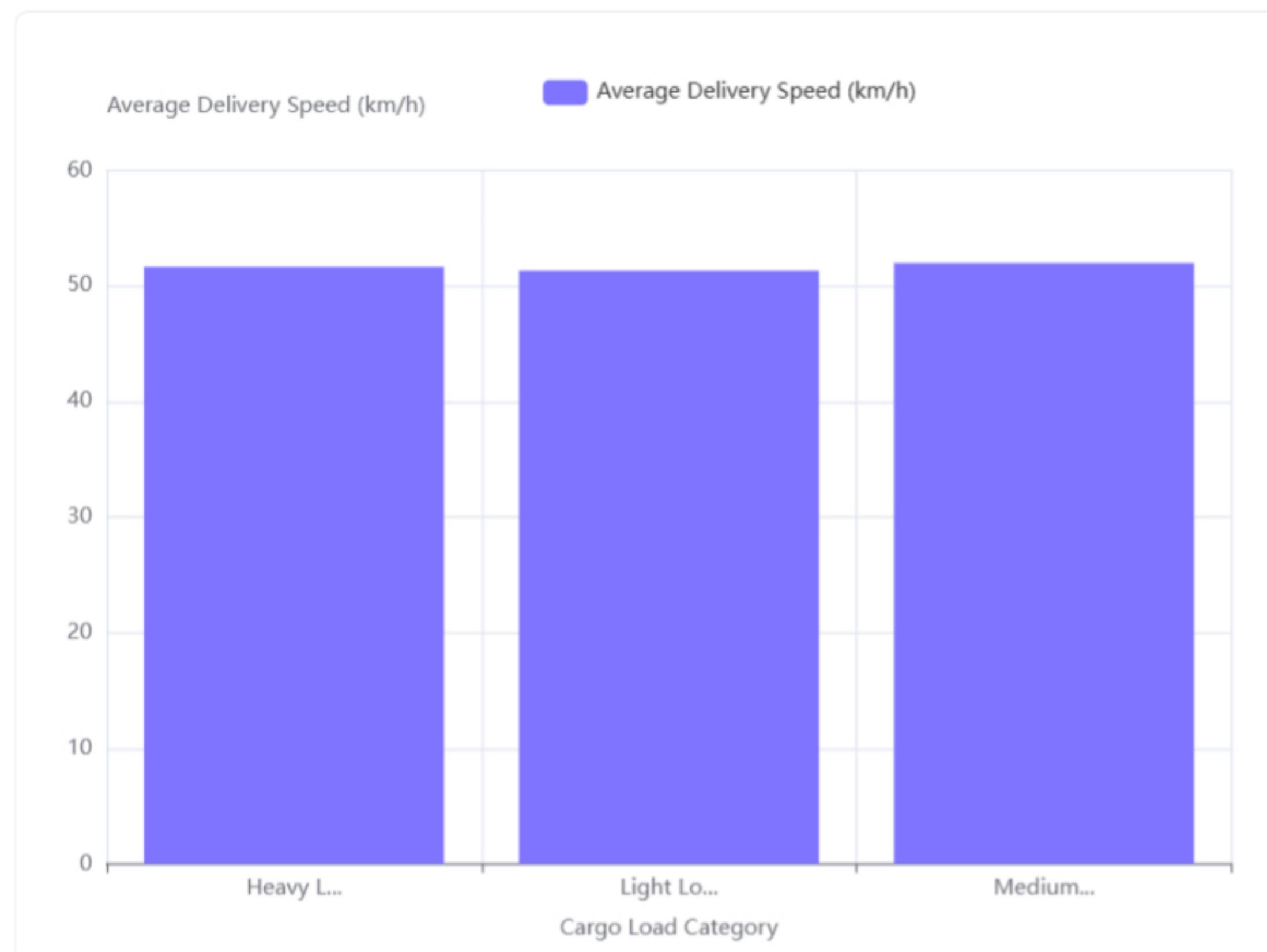
Average Delivery Duration by Traffic Index

- **High Traffic Index:** The average delivery duration is approximately **4.62 hours**.
- **Low Traffic Index:** The average delivery duration is approximately **4.59 hours**.
- **Medium Traffic Index:** The average delivery duration is approximately **4.43 hours**.

Conclusion and Insights

- **Impact of Traffic Index:** The traffic index significantly influences delivery duration. **Medium traffic conditions result in the shortest average delivery duration (4.43 hours)**, while **high traffic conditions lead to the longest average delivery duration (4.62 hours)**.
- **Counter-intuitive Observation:** Interestingly, **low traffic conditions (4.59 hours)** show a slightly longer average delivery duration than medium traffic conditions (4.43 hours), and are very close to high traffic conditions. This suggests that factors beyond just traffic volume, such as route optimization, driver behavior, or specific delivery zones, might play a role, or that "low" traffic might not always equate to optimal speed (e.g., due to fewer drivers, longer routes in less dense areas, or different types of deliveries).
- **Optimization Potential:** Understanding these variations can help in optimizing delivery routes and scheduling, especially by focusing on strategies that mitigate delays during high traffic and investigating the reasons for slightly longer durations in low traffic compared to medium traffic.

Q10. What is the average delivery speed for each cargo load category?



Average Delivery Speeds

- **Heavy Load:** The average delivery speed for **Heavy Load** is approximately **51.65 km/h**.
- **Light Load:** The average delivery speed for **Light Load** is approximately **51.32 km/h**.
- **Medium Load:** The average delivery speed for **Medium Load** is approximately **51.99 km/h**.

Conclusion and Insights

- **Varying Speeds:** The average delivery speeds show slight variations across different cargo load categories, ranging from approximately 51.32 km/h to 51.99 km/h.
- **Medium Load is Fastest:** **Medium Load** deliveries have the highest average speed at **51.99 km/h**.
- **Light Load is Slowest:** **Light Load** deliveries have the lowest average speed at **51.32 km/h**, which is marginally slower than Heavy Load deliveries.
- **Minimal Difference:** The difference in average delivery speed between the fastest (Medium Load) and slowest (Light Load) categories is **less than 1 km/h (0.67 km/h)**, indicating that cargo load category does not significantly impact average delivery speed in this dataset.