Digital Egypt Pioneers

Dynamic Supply Chain Logistics Analytics 📦

Team Members:

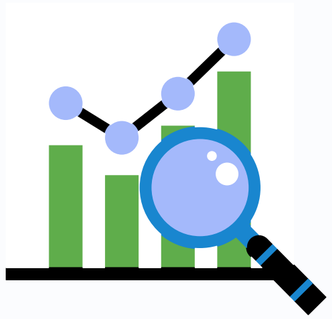
Ziad Khaled
Abdelrhman Ahmed
Sama Mahmoud
Mohamed Rafaat
Seif El-Deen Ahmed
Mohamed Hymeda
Reneh Romany

Digital Egypt Pioneers

Table of Content

Digital Egypt Pioneers

Explore Data with Python



Dataset Overview & Structure

Digital Egypt Pioneers

Dataset Description:

This dataset simulates operations from a **modern logistics network in Southern California**, spanning from **January 2021** to **January 2024.**

It includes **hourly records** of key activities such as **transportation**, **warehouse operations**, **inventory levels**, and **real-time GPS tracking.**

The data was **generated** from systems like **IoT sensors**, **vehicle trackers**, and s**upply chain platforms**, making it ideal for analysis of delivery performance, risk exposure, and operational efficiency.

Dataset Overview & Structure

- **Size**: 32,065 rows × 26 columns
- **Data Types**: 24 numerical, 2 categorical (timestamp, risk_classification)
- **Key Domains**:
  - **Time-based**: timestamp
  - **Location-based**: vehicle GPS coordinates
  - **Performance-based**: delivery time deviation, ETA variation
  - **Risk-related**: disruption score, route risk level, delay probability

```
👉 Dataset Shape: (32065, 26)

👉 Data Types:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 32065 entries, 0 to 32064
Data columns (total 26 columns):
 #   Column                          Non-Null Count  Dtype
---  ------                          --------------  -----
 0   timestamp                       32065 non-null  object
 1   vehicle_gps_latitude            32065 non-null  float64
 2   vehicle_gps_longitude           32065 non-null  float64
 3   fuel_consumption_rate           32065 non-null  float64
 4   eta_variation_hours             32065 non-null  float64
 5   traffic_congestion_level        32065 non-null  float64
 6   warehouse_inventory_level       32065 non-null  float64
 7   loading_unloading_time          32065 non-null  float64
 8   handling_equipment_availability 32065 non-null  float64
 9   order_fulfillment_status        32065 non-null  float64
 10  weather_condition_severity      32065 non-null  float64
 11  port_congestion_level           32065 non-null  float64
 12  shipping_costs                  32065 non-null  float64
 13  supplier_reliability_score      32065 non-null  float64
 14  lead_time_days                  32065 non-null  float64
 15  historical_demand               32065 non-null  float64
 16  iot_temperature                 32065 non-null  float64
...
 24  risk_classification             32065 non-null  object
 25  delivery_time_deviation         32065 non-null  float64
dtypes: float64(24), object(2)
```

Data Cleaning & Validation

Digital Egypt Pioneers

- **Missing Values:** None found across all columns
- **Duplicates:** 0 duplicate rows
- **Timestamp Conversion:** Converted to datetime for time-based analysis
- **Data Types:** Confirmed using .info() and .describe() in pandas

```
df['timestamp'] = pd.to_datetime(df['timestamp'])
```

```
👉 Missing Values Analysis:
Series([], dtype: float64)

👉 Duplicate rows: 0

👉 Numerical Columns Statistics:
       vehicle_gps_latitude  vehicle_gps_longitude  fuel_consumption_rate  \
count          32065.000000           32065.000000           32065.000000
mean              38.023589             -90.116648               8.011735
std                6.917909              17.369244               4.264960
min               30.000000            -119.999998               5.000000
25%               31.280550            -106.253913               5.019984
50%               36.413820             -86.293414               5.636036
75%               44.453655             -73.079367               9.669944
max               50.000000             -70.000000              19.999875

       eta_variation_hours  traffic_congestion_level  \
count         32065.000000              3.206500e+04
mean              2.893068              4.991493e+00
std               2.274044              3.532048e+00
min              -1.999993              1.091633e-09
25%               1.185744              1.474720e+00
50%               3.882059              4.981244e+00
75%               4.884355              8.534902e+00
max               5.000000              9.999999e+00
...
75%                       0.998746          0.982391               9.249206
max                       1.000000          1.000000              10.000000

[8 rows x 24 columns]
```

Feature Engineering & New Columns

- **Risk Metrics:** total_risk_score **(scale: 0–1, Higher = higher overall risk)**
- **Performance Metrics:**
  - **delivery_efficiency:** score out of 100 % **(Higher = better timing accuracy)**
  - **delivery_performance:** score out of 100 % **(Near 100 = perfect delivery, Near 0 = bad)**
- **Efficiency Metric:**
  - **cost_efficiency:** score out of 100 %,**(>100 = more efficient, <100 = less efficient)**
- **Delay Labels:** **(e.g., Very Early, On Time, Very Late)**
- **Distance Feature:** GPS shifts **(approximate km)**

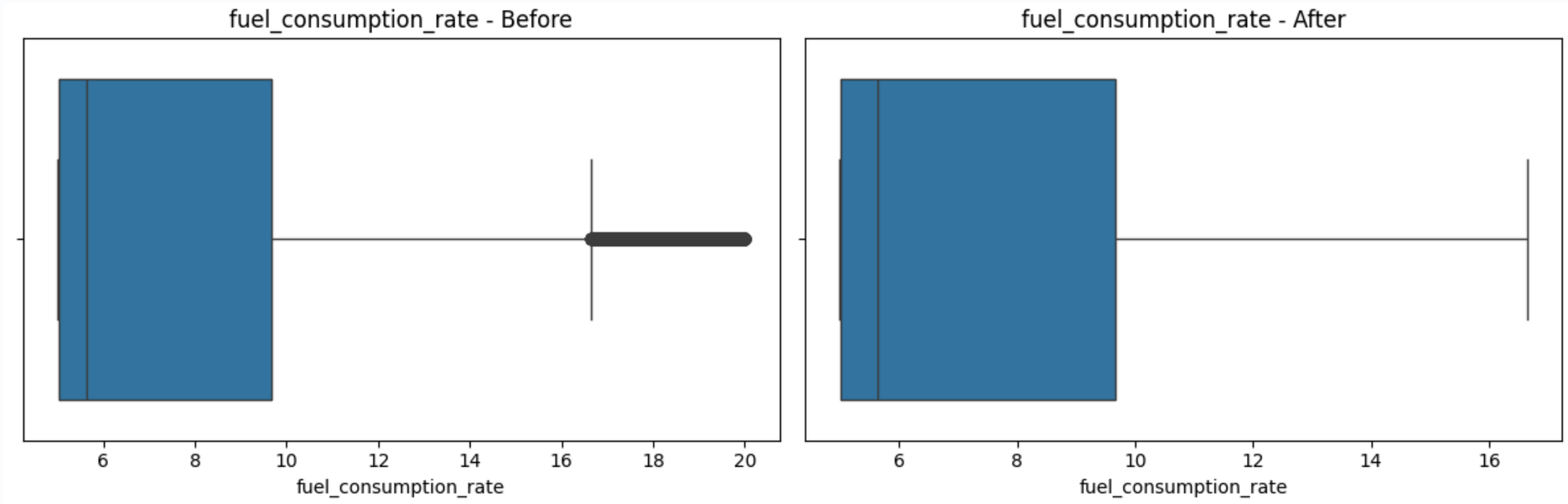
	total_risk_score	delivery_efficiency	delay_category	distance	delivery_performance	cost_efficiency
0	0.86	-16.10	Very Late	30.121	0.50	100.63
1	3.71	13.32	Very Late	40.608	0.16	71.73
2	2.79	62.30	Late	41.913	0.02	294.94
3	0.52	-8.52	Very Late	8.351	0.05	440.36
4	3.59	6.39	Very Late	6.650	0.02	47.01

Outlier Detection & Handling

Technique Used: Interquartile Range **(IQR Method)**

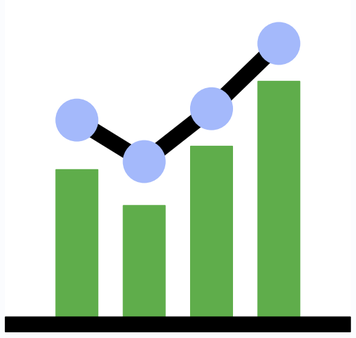
- **Columns Excluded:**
  - **Engineered metrics:** **delivery_performance, delivery_efficiency, cost_efficiency**
  - **Probabilities:** **delay_probability, driver_behavior_score, etc.**
- **Clipping Method:** **.clip(lower, upper)** instead of deleting rows
- **Visualization:** Boxplots used to confirm distributions pre/post handling
- **Result:** **6,363 total** outliers handled safely without harming metric integrity

Ex: fuel_consumption_rate – Outliers: 2536 (7.91%)



Digital Egypt Pioneers

Analysis Questions Phase



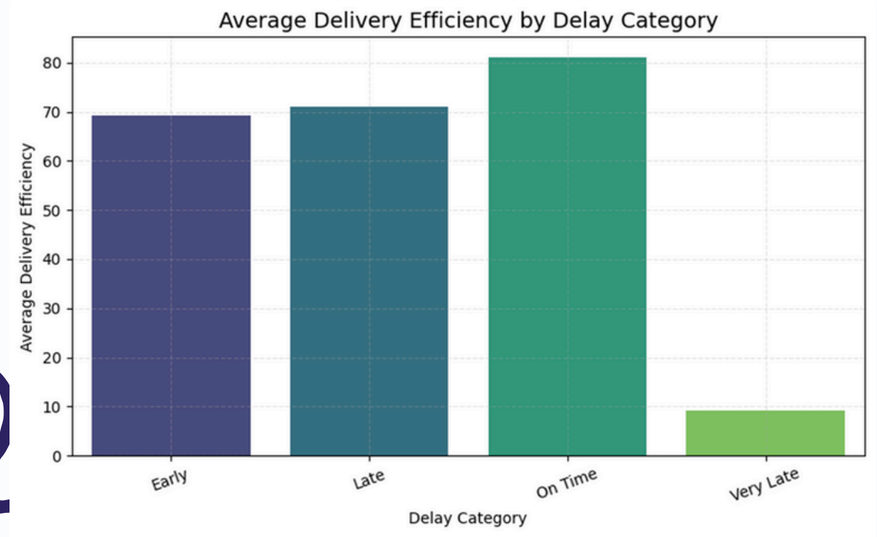
Fuel & Efficiency Analysis

Digital Egypt Pioneers

Average fuel consumption

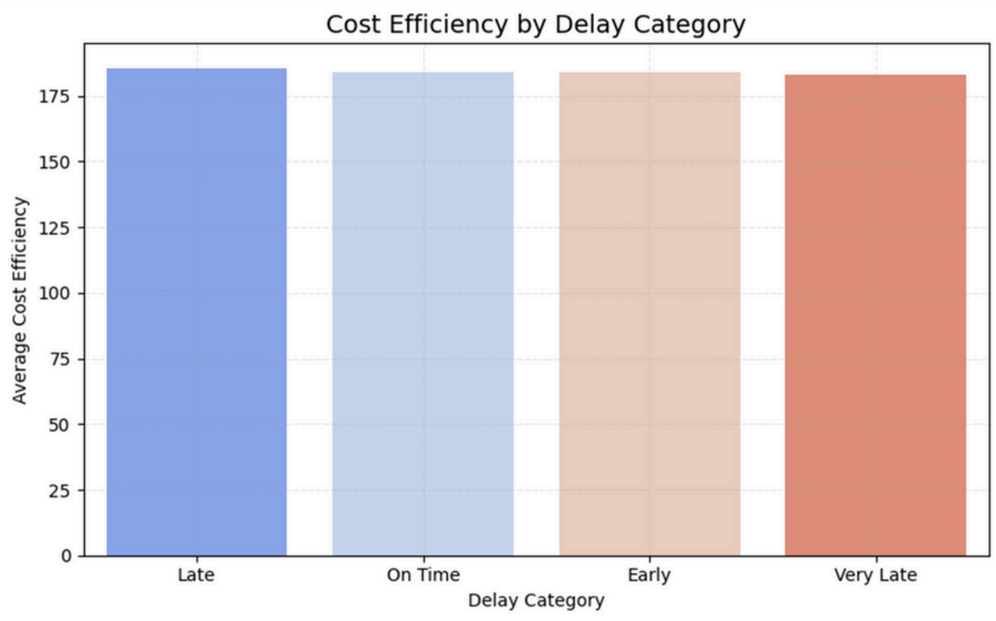
- Insight: This metric highlights overall fuel efficiency. A value of 7.74 liters suggests moderate fuel usage, but outliers (e.g., values like 19.9 liters) indicate potential inefficiencies in specific vehicles or routes
- Project Impact: Identifies opportunities for fuel optimization, such as adopting fuel-efficient vehicles or optimizing routes. It prompts discussions on sustainability and cost reduction strategies.
- Discussion Points: Are high fuel consumption rates tied to specific vehicle types or routes? Can alternative fuels or hybrid vehicles reduce costs?

```
Q1: Average fuel consumption rate: 7.87
```



Cost efficiency vary across different delay categories

- Insight: Costs are highest for Early deliveries, possibly due to expedited efforts, and Very Late deliveries, likely due to penalties or inefficiencies. On-Time and Late categories show slightly lower costs, indicating better cost efficiency.
- Project Impact: Highlights the financial burden of delays, guiding strategies to prioritize on-time deliveries or recover costs for early shipments.
- Discussion Points: Why are early deliveries costly? Can process improvements reduce Very Late delivery costs?

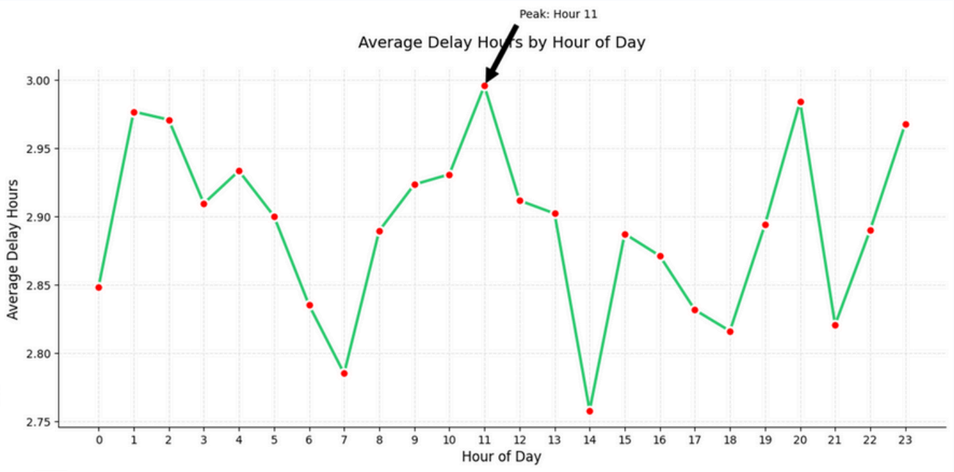


Delays & Time Analysis

Digital Egypt Pioneers

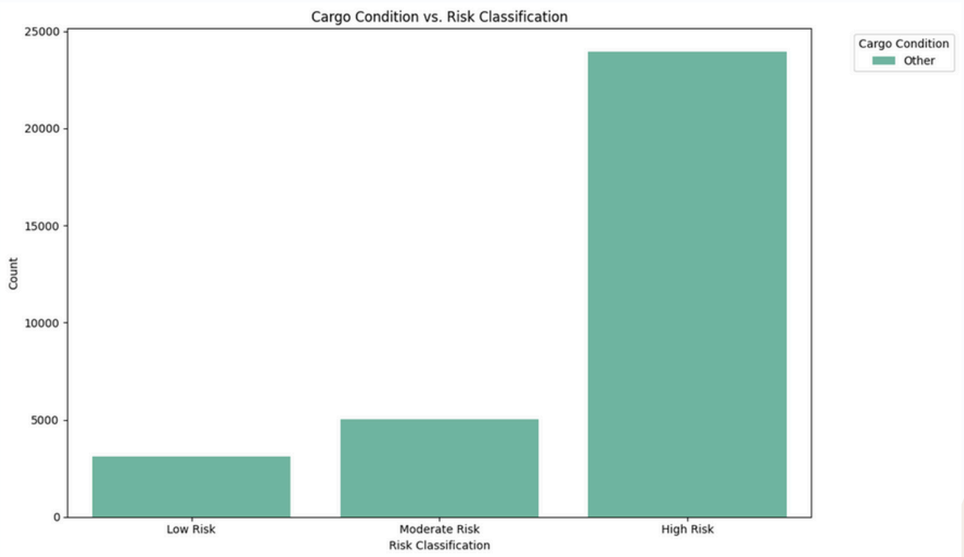
During which hour of the day do deliveries experience the highest average delay

- Insight: Peak delays at 10:00 AM suggest morning traffic or operational bottlenecks, such as warehouse processing delays.
- Project Impact: Informs scheduling adjustments or route planning to avoid peak delay hours, improving delivery reliability.
- Discussion Points: Are 10:00 AM delays due to traffic, staffing, or warehouse issues? Can deliveries be rescheduled to off-peak hours?



ETA Variation by Weekday

- Insight: A significant 78.51% of deliveries are classified as High Risk, indicating substantial risk exposure, possibly due to factors like route risks or cargo conditions.
- Project Impact: Highlights the need for enhanced risk management, such as better route planning or cargo monitoring, to reduce high-risk cases.
- Discussion Points: What factors (e.g., route, weather, cargo) drive high-risk classifications? Can targeted interventions lower risk levels?

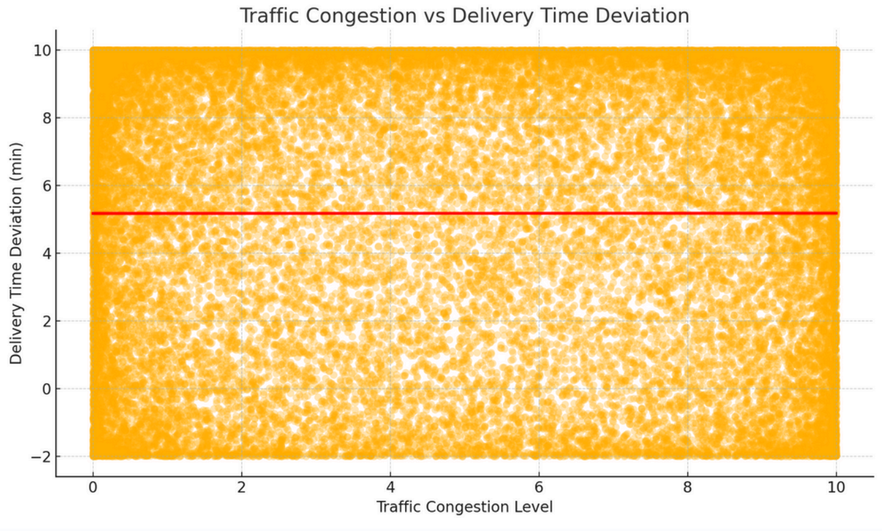


Risk & Performance Analysis

Digital Egypt Pioneers

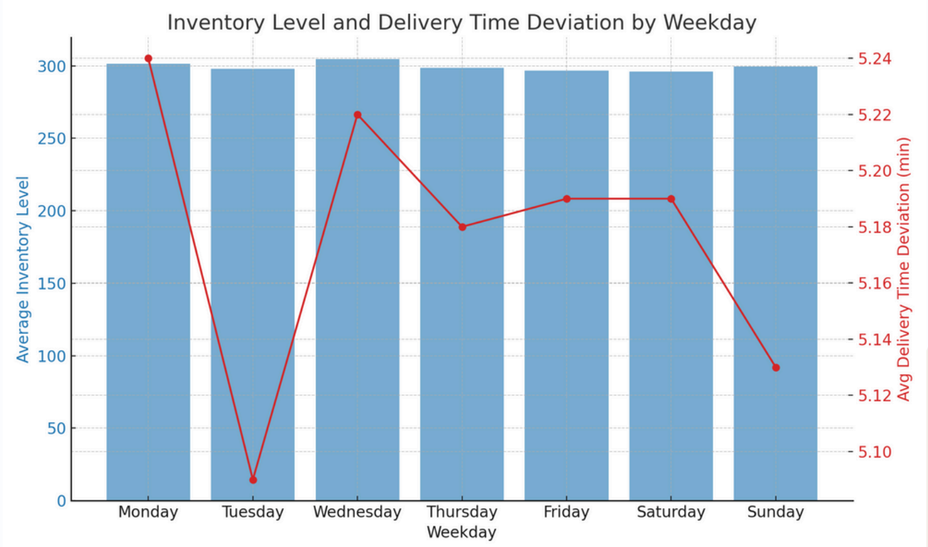
How does traffic congestion correlate with delivery time deviation

- Insight: The weak correlation suggests traffic congestion is not a primary driver of delays in this dataset, possibly due to effective route planning or other dominant factors (e.g., warehouse delays).
- Project Impact: Encourages exploration of other delay causes while maintaining basic traffic monitoring, as congestion has minimal impact.
- Discussion Points: Why is congestion's impact low? Are there specific routes where congestion matters more?



Relationship between warehouse inventory levels and delivery performance across days of the week

- Insight: Inventory levels remain relatively stable (~297 units), but delays peak on Wednesday (4.79 minutes) and are lowest on Tuesday (4.41 minutes), suggesting possible inventory processing inefficiencies on certain days.
- Project Impact: Guides inventory replenishment schedules and warehouse staffing to reduce delays, particularly on high-delay days like Thursday.
- Discussion Points: Do Wednesday delays correlate with inventory processing bottlenecks? Can inventory adjustments improve Tuesday's performance further?



Digital Egypt Pioneers

Forcasting Questions Phase

Fuel & Efficiency Analysis

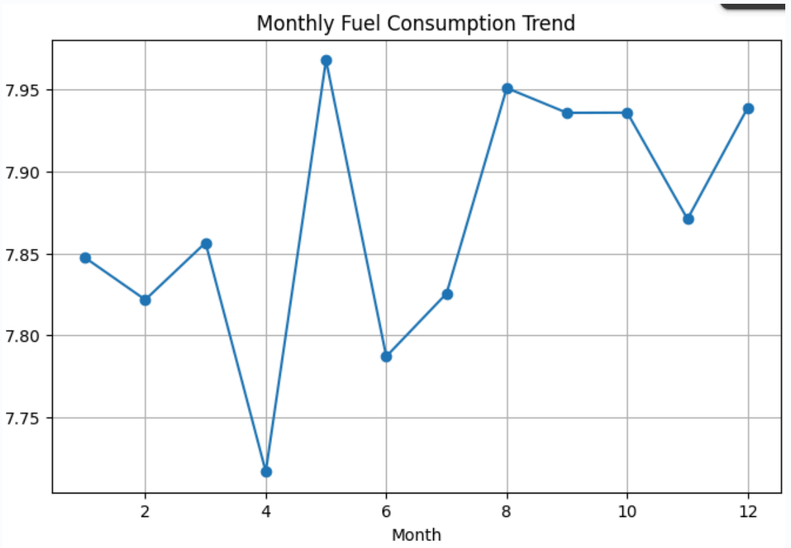
Digital Egypt Pioneers

Fuel consumption prediction

- Analysis: The mean error (MSE = 15.39) shows acceptable accuracy. The model predicted higher fuel consumption in May, August, and December, which helps improve planning and reduce waste

- Project impact: The project helps improve fuel efficiency by predicting critical periods, which contributes to reducing costs, minimizing environmental impact, and supporting operational decisions.

- Discussion topic: How can fuel consumption forecasting be used to improve operational efficiency and reduce environmental costs in organizations?

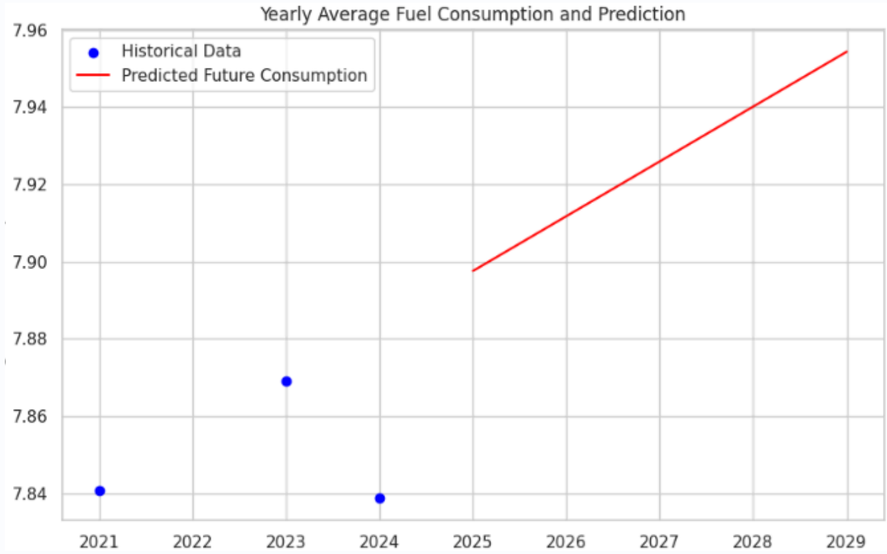


Delivery fuel consumption forecast

- Analysis : The results show a decrease in fuel consumption in deliveries over the past few years, and this is expected to continue in the future. This is due to new combat vehicle technologies that enable more efficient delivery.

- Project Impact : The project helps reduce fuel consumption and operating costs, lowers environmental impact, and enhances the efficiency and effectiveness of delivery operations.

- Discussion topic : How will technological developments and modern delivery methods impact fuel consumption in delivery operations in the coming years?

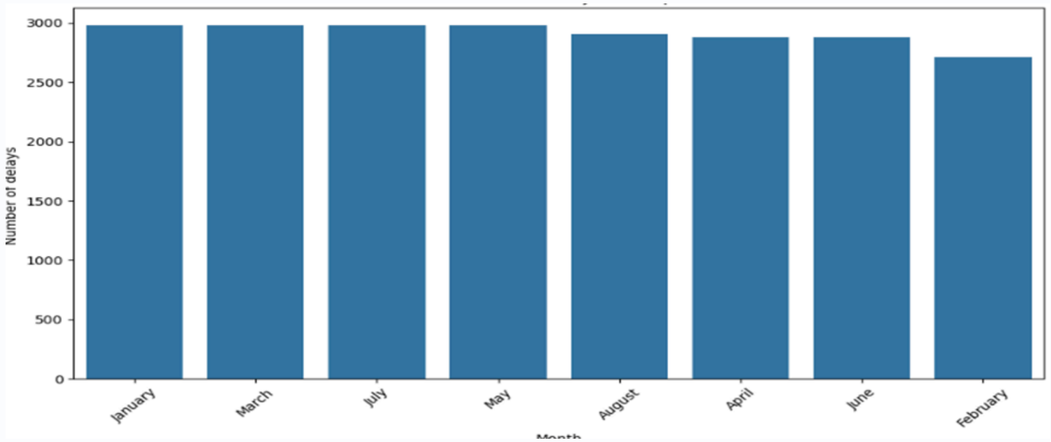


Expected delivery delays

Digital Egypt Pioneers

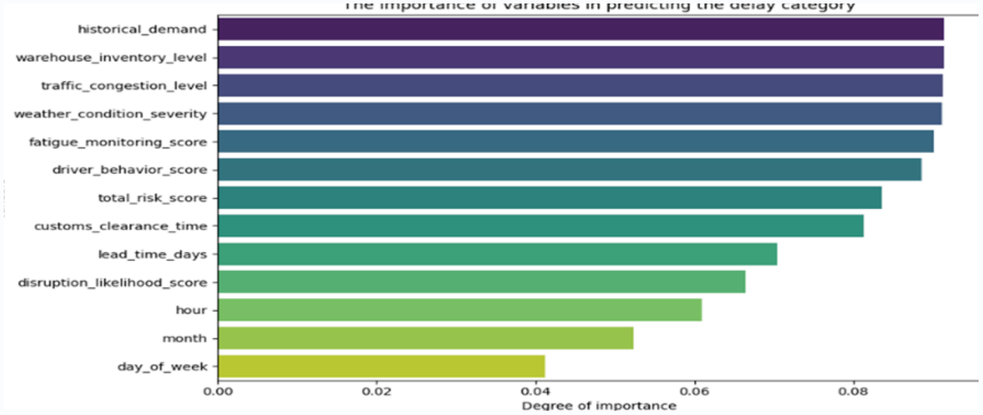
What are the expected delay months?

- Analysis: January, March, May, and July are expected to have the highest number of delivery delays at 2,976, while February is expected to have the lowest number at 2,712
- Project Impact: The project contributes to reducing delays and increasing delivery efficiency by predicting critical periods in advance and making data-driven decisions.
- Discussion topic: What are the expected delivery delay months and how can this information be used to improve logistics planning?



Main reasons for delivery delays

- Analysis: Customs delays and traffic congestion are the main factors in delivery delays, with disruptions and driver behavior having a moderate impact.
- Project impact: Improved accuracy of forecasting delivery delays, which helps in better planning supply chains, reducing operational losses, increasing customer satisfaction, and enhancing the overall efficiency of logistics operations.
- Discussion topic: What are the main reasons for future delivery delays?

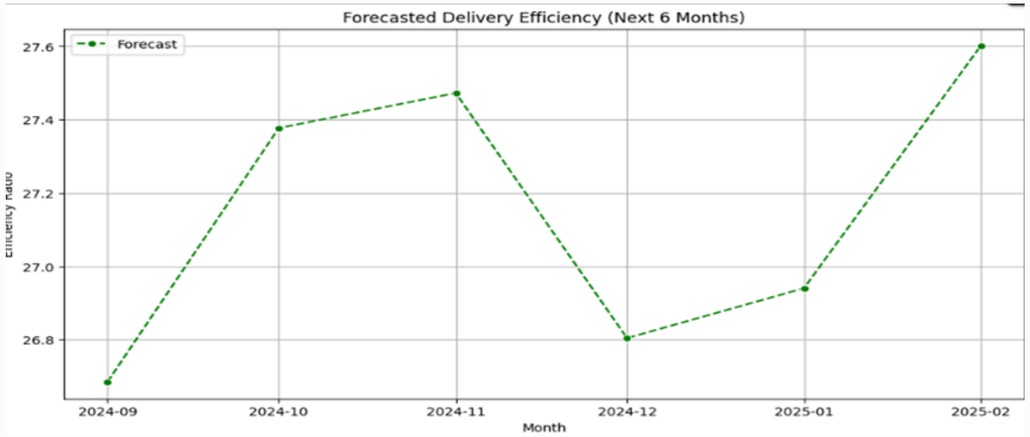


Effect of performance on conductivity

Digital Egypt Pioneers

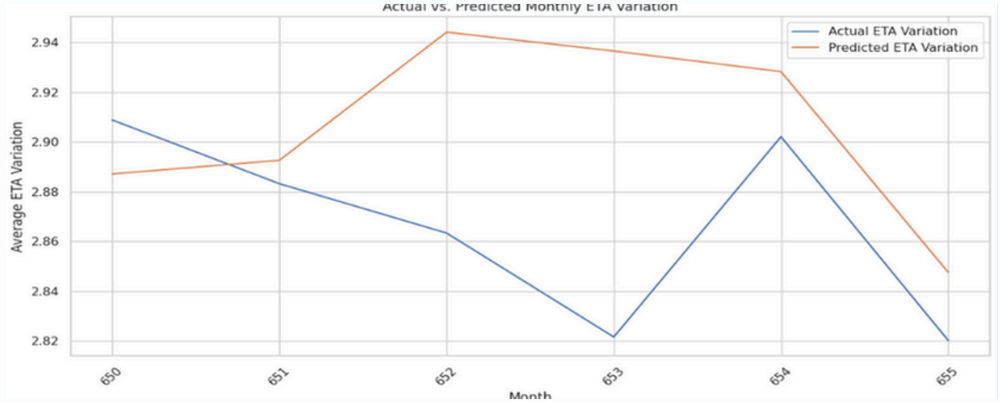
The impact of monthly changes on delivery efficiency

- Analysis: Delivery efficiency is high in October and November 2024, declines in December to 26.8, and then gradually improves to 27.6 in February 2025.
- Project impact: Enables companies to accurately predict periods of reduced delivery efficiency, which helps improve operational planning, better allocate resources, and reduce delays.
- Discussion topic: How do monthly performance changes affect delivery efficiency?



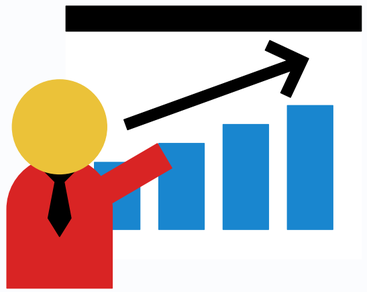
The effect of driver fatigue on ETA accuracy

- Analysis: This indicates the relationship between driver fatigue and its impact on the accuracy of expected time of arrival (ETA) estimates. Using average fatigue scores from previous months, the variance in ETA can be predicted for the current month.
- Project Impact: The project helps improve delivery timeliness by anticipating the impact of driver fatigue, reducing delays and increasing operational efficiency.
- Discussion topic: How does driver fatigue affect the accuracy of expected time of arrival (ETA) estimates?



Digital Egypt Pioneers

Dashboard Overview



Dashboard Overview

On-Time Delivery Rate
5.9%

Total Shipments
32K

Avg Delay Probability
70%

Avg Risk Score
2.83

Year
All

Quarter Selection
All

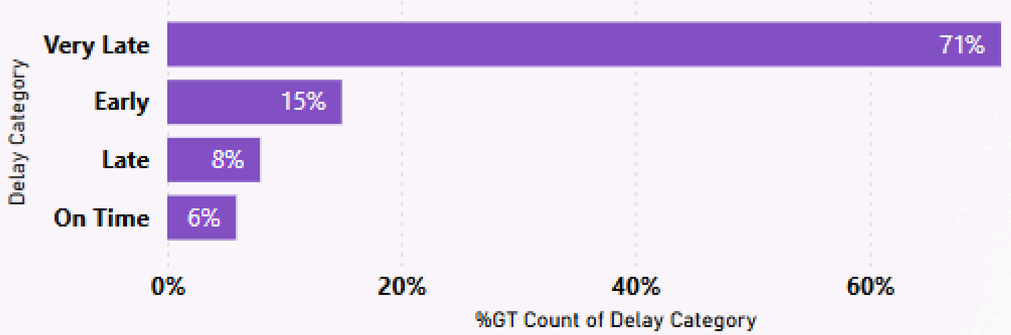
Months
All

Delay Status
All

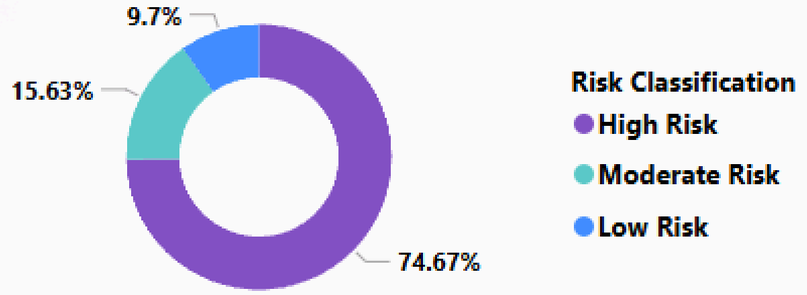
Risk Classification
All

Clear all slicers

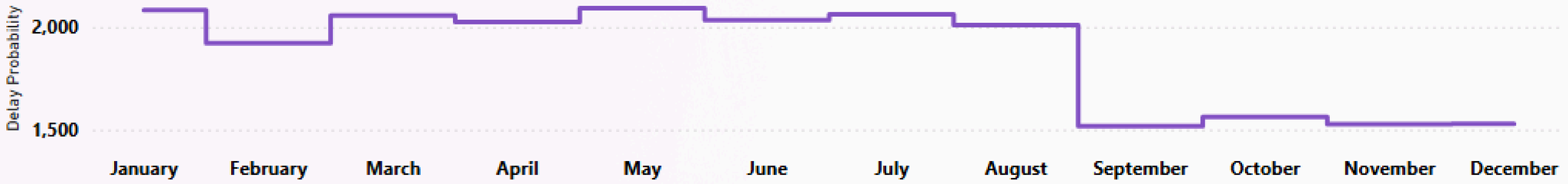
Delay Categories Breakdown



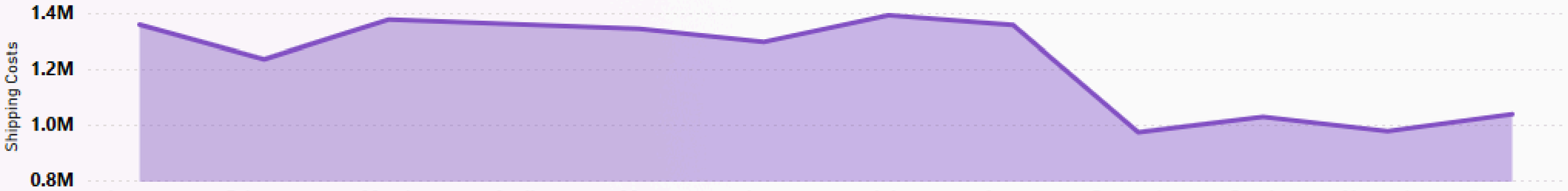
Risk Levels Across All Shipments



Delay Probability Trend



Monthly Shipping Cost Trend



Delivery & Delay Analysis

Total Shipments
32K

On-Time Delivery Rate
5.9%

Avg Delay Probability
70%

Avg Risk Score
2.83

Year
All

Quarter Selection
All

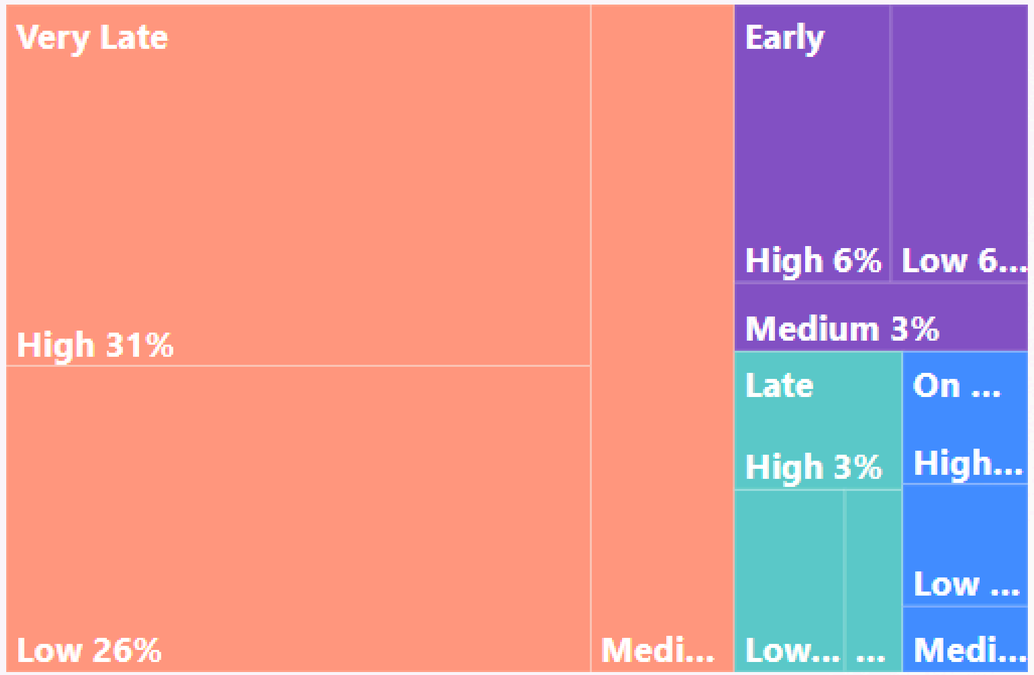
Months
All

Delay Status
All

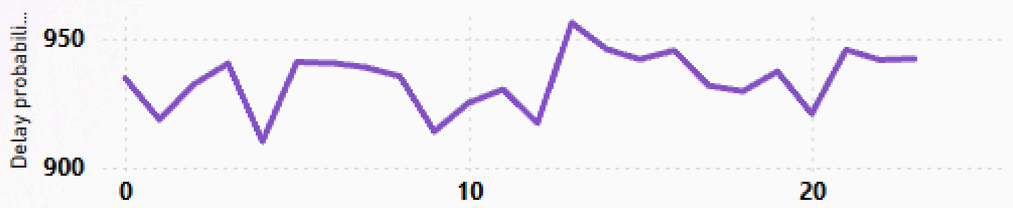
Risk Classification
All

Clear all slicers

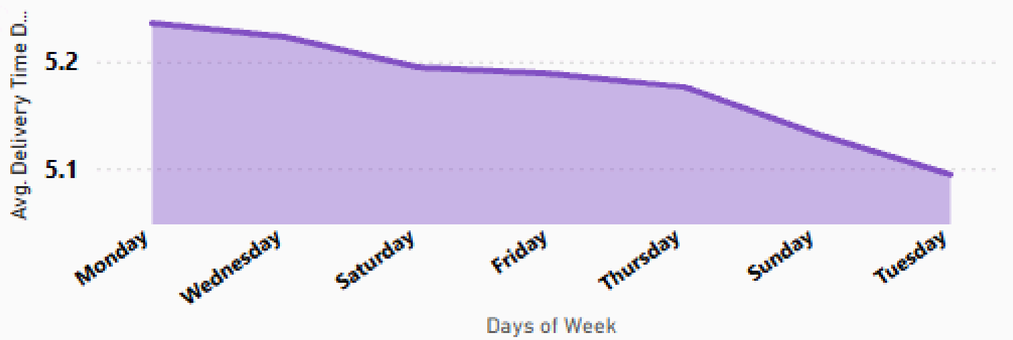
Impact of Traffic on Delivery Times



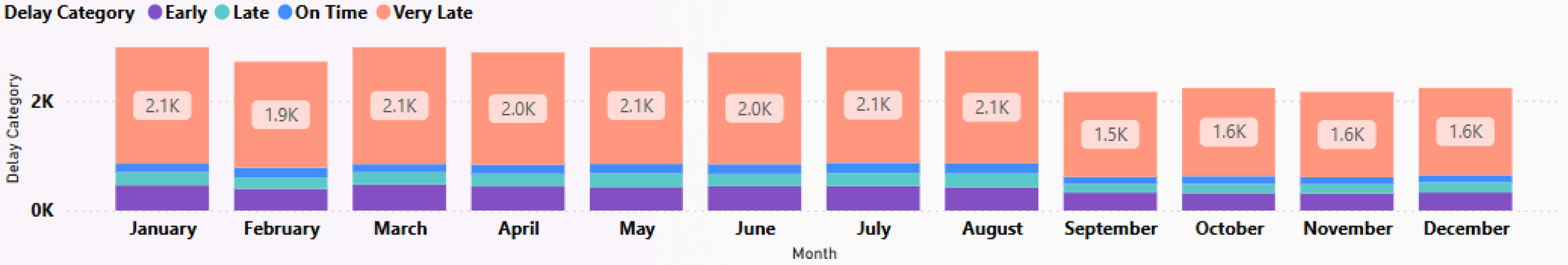
Delay Chances by Time of Day



Delivery Performance by Day of Week



Monthly Evolution of Delay Categories



Risk & Driver Behavior

Total Shipments
32K

On-Time Delivery Rate
5.9%

Avg Delay Probability
70%

Avg Risk Score
2.83

Year
All

Quarter Selection
All

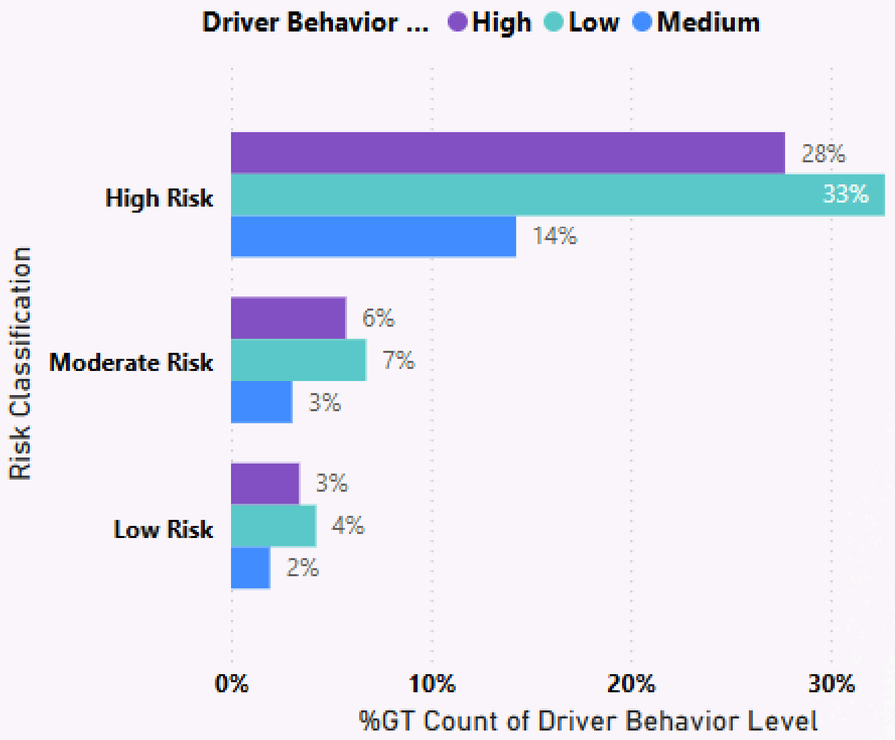
Months
All

Delay Status
All

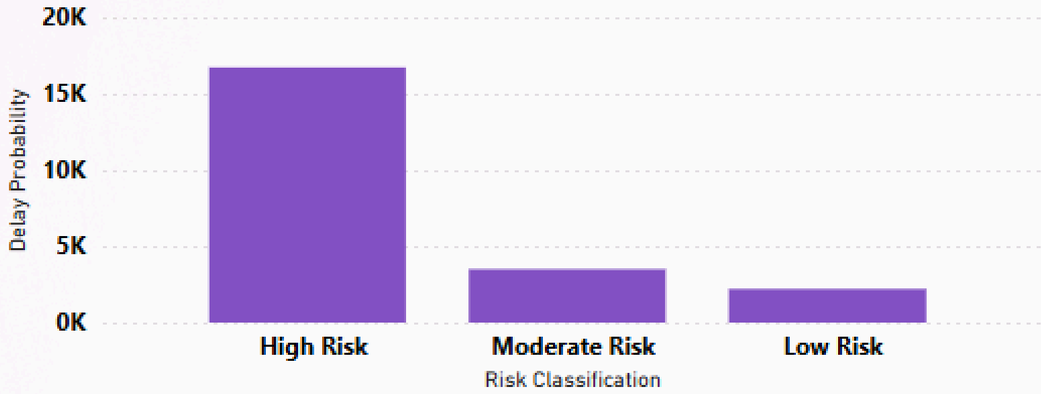
Risk Classification
All

Clear all slicers

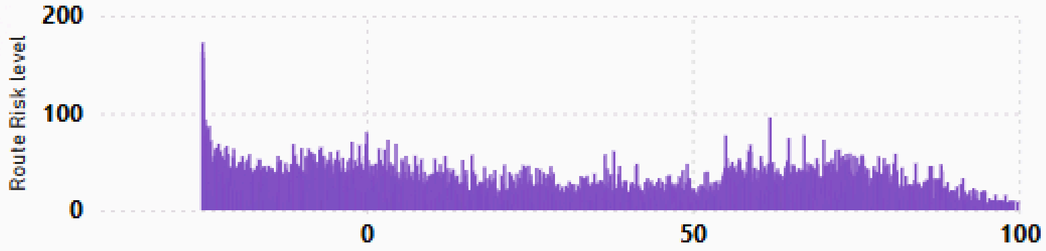
How Driver Behavior Affects Shipment Risk



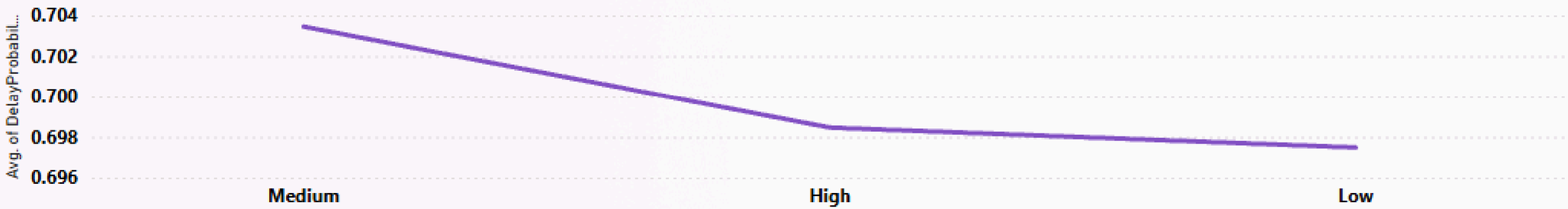
Impact of Risk Levels on Delivery Time



Route Risk vs Delivery Efficiency



Fatigue Impact on Delay



Cost & Fuel Insights

Fuel Used / Distance	Total Delevery Cost	Distance / Fuel Used	Delivery Cost / Distance
7.87	15M	2.90	20.16

Year: All

Quarter Selection: All

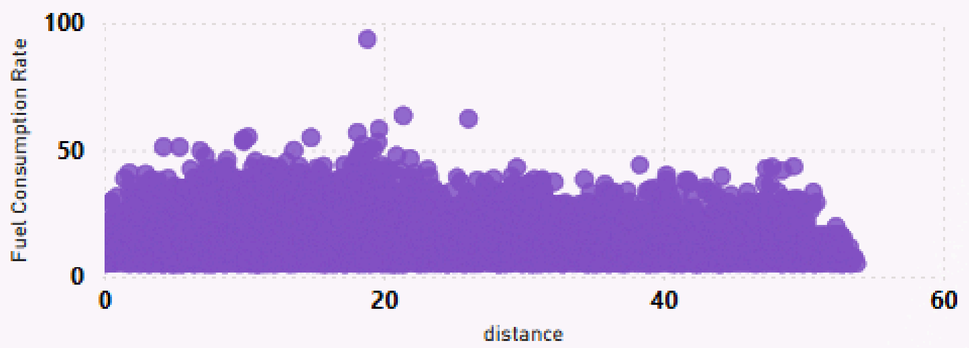
Months: All

Delay Status: All

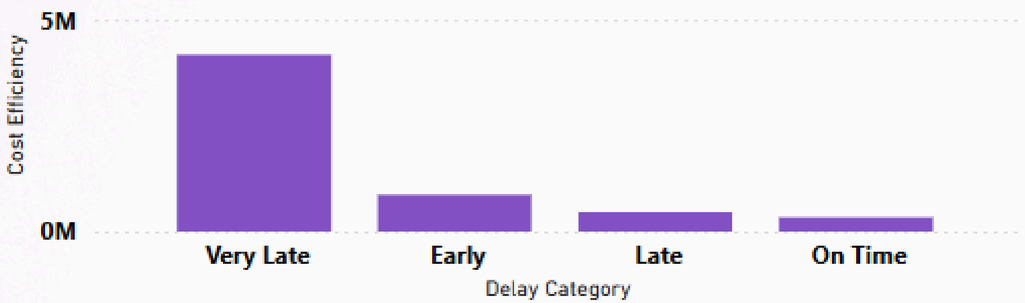
Risk Classification: All

Clear all slicers

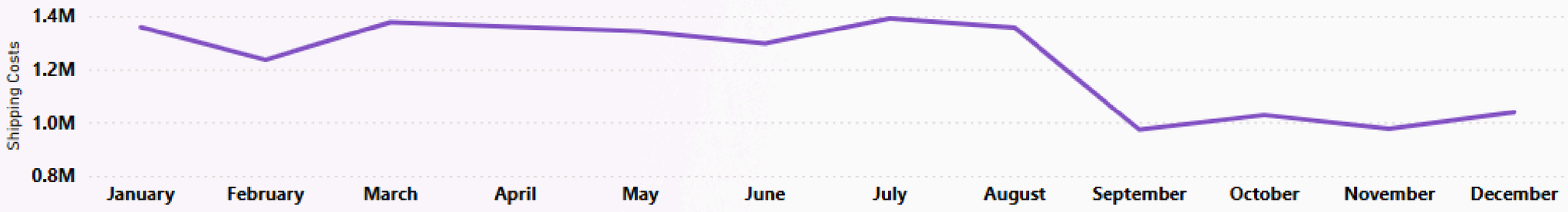
Distance Impact on Fuel Consumption



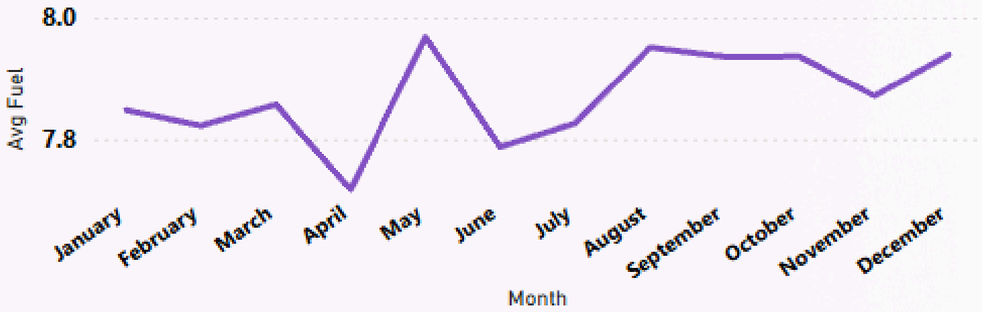
How do costs differ across delay categories?



Shipping Cost Trend Over Months



Monthly Average Fuel Usage



Hourly Average Fuel Consumption

hour	1	2	3	4	5	6	7	8	9	10	11	12
0	7.19	7.68	7.68	7.09	7.97	8.56	7.05	7.99	8.32	7.16	7.88	7.22
1	7.82	7.56	8.18	7.94	7.19	6.23	7.98	7.83	8.68	7.95	6.97	8.89
2	7.93	8.51	8.76	7.02	7.51	8.04	8.14	8.01	7.71	7.04	7.04	7.59
3	8.49	7.37	7.94	8.25	7.55	8.07	8.16	6.93	6.60	7.05	8.10	7.37
4	7.10	8.37	7.24	7.79	8.57	7.74	9.02	7.66	8.75	8.55	8.84	7.32
Total	**7.94**	**8.04**	**7.91**	**7.58**	**7.70**	**7.76**	**8.03**	**7.79**	**7.89**	**7.64**	**7.76**	**7.93**

Key Insights

- **Delivery Performance Overview**
  - **Alarmingly Low On-Time Delivery Rate:** Only **5.9% of shipments** are delivered on time, indicating a critical performance issue.
  - **High Average Delay Probability:** **70% average delay probability** suggests that a large majority of shipments are expected to be late.
  - **Dominant "Very Late" Deliveries:** **71% of delayed shipments** fall into the "Very Late" category, highlighting the severity of delays.
- **Risk and Driver Behavior Analysis**
  - **Significant "High Risk" Shipments:** **15.6% of shipments** are classified as "High Risk," posing potential operational challenges.
  - **Strong Link:** High-Risk Driver Behavior & Shipment Risk: Shipments associated with "**High Risk**" driver behavior are more likely to be classified as "**High Risk.**"

Key Insights

Digital Egypt Pioneers

- **Cost and Fuel Insights**
  - **Average Fuel Consumption & Efficiency:** Average Fuel Used / Distance is **7.87**, and Distance / Fuel Used is **2.90**.
  - **Total Delivery Cost:** Total Delivery Cost is **15 M.**
  - **Cost per Distance:** Average Delivery Cost / Distance is **20.16.**
  - **Costly "Early" Deliveries:** Surprisingly, "**Early**" deliveries contribute to **higher** costs, similar to "**Very Late**" deliveries.
  - **Temporal Fuel Consumption Patterns:** Fuel usage varies by **month** and **hour** of the day.
- **Temporal Delivery Trends**
  - **Monthly Delay Probability Fluctuations:** Delay probability varies throughout the year, with a notable dip in September.
  - **Day-of-Week Delivery Performance Variation:** Delivery performance appears slightly lower at the beginning of the week (Monday & Tuesday).
  - **Monthly Shipping Cost Trends:** Shipping costs fluctuate throughout the year, with a notable decrease in September.

Recommendations

- Conduct a **detailed investigation** into the primary causes of "**Very Late**" deliveries and implement targeted strategies to improve on-time performance.
- Develop specific handling protocols for "High Risk" shipments aimed at minimizing potential delays and analyze the factors contributing to "High Risk" classifications.
- Implement targeted training and coaching programs for drivers exhibiting "High Risk" behaviors.
- implement strategies such as route optimization and driver training to reduce overall fuel consumption and costs.
- explore the potential of transitioning to alternative fuel vehicles or electric vehicles to reduce fuel costs and environmental impact.

Digital Egypt Pioneers

THANK YOU