NYC Taxi Demand & Revenue Forecasting

An End-to-End Analytics, Forecasting, and Business Intelligence Project

1. Introduction

Urban transportation systems generate vast volumes of data that can be leveraged to understand demand patterns, optimize operations, and inform planning decisions. This project analyzes New York City taxi trip data to uncover historical demand and revenue trends, identify behavioral patterns, and generate short-term forecasts of taxi demand.

The project combines data engineering, exploratory analytics, time-series forecasting, and business intelligence visualization into a single end-to-end workflow. Final insights are delivered through an interactive Power BI dashboard supported by a structured analytical report.

2. Project Objectives

The main objectives of this project are to:

- Analyze historical taxi demand and revenue trends
- Understand temporal and behavioral patterns in taxi usage
- Build a short-term demand forecasting model
- Present insights through a professional Power BI dashboard
- Demonstrate end-to-end data engineering and analytics capability

3. Data Description

The dataset used in this project is derived from NYC taxi trip records and aggregated to a daily level for analytical and forecasting purposes.

Key Metrics Included

- Total number of trips per day
- Total daily revenue
- Average fare amount
- Average trip distance
- Average trip duration

• Trip date and derived calendar features

The aggregation strategy ensures that the data is suitable for both trend analysis and time-series modeling while remaining performant for BI consumption.

4. Data Engineering & Architecture

This project follows a Medallion Architecture approach to ensure clear separation between raw data ingestion, cleaned analytical datasets, and business-ready outputs.

Architecture Layers

• Bronze Layer: Raw NYC taxi data as ingested

• Silver Layer: Cleaned, transformed, and aggregated daily metrics

• Gold Layer: Analytics-ready tables and forecast outputs consumed by Power BI

This layered design improves data quality, reproducibility, and scalability while reflecting industry-standard data engineering practices.

5. Demand & Revenue Insights

Exploratory analysis was conducted to understand how taxi demand and revenue evolve over time and how usage patterns differ across temporal dimensions. This section focuses on what was observed, not how the analysis was performed.

5.1 Overall Demand & Revenue Trends

Daily trip volume and revenue exhibit strong co-movement, indicating that revenue is primarily driven by demand rather than large fluctuations in pricing. Periods of increased trip activity are consistently associated with higher total revenue.

Short-term fluctuations are present, reflecting natural variability in daily transportation demand.

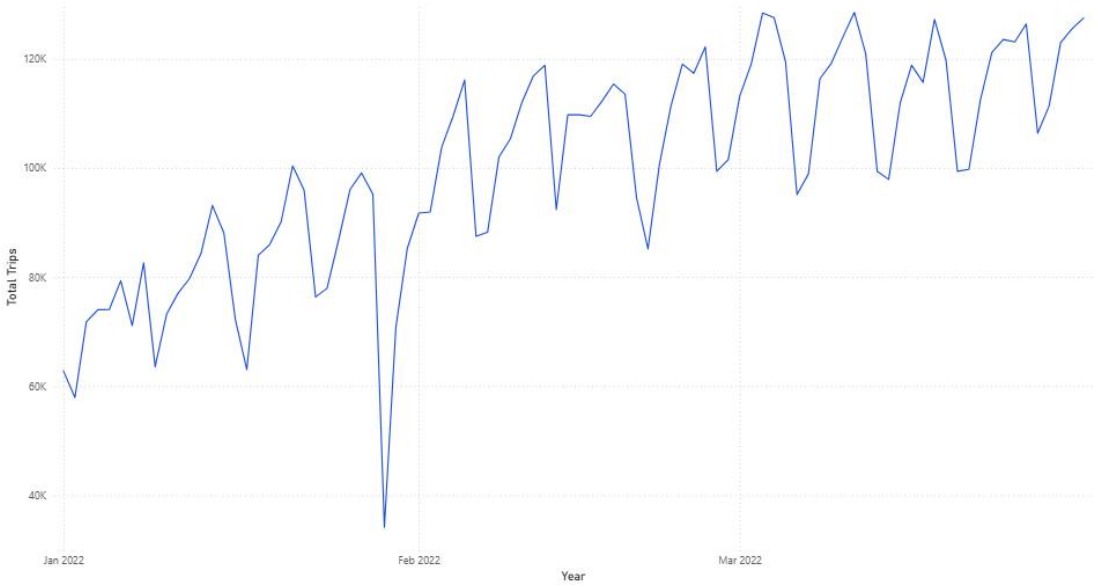

Fig 1: Daily Trips Trend

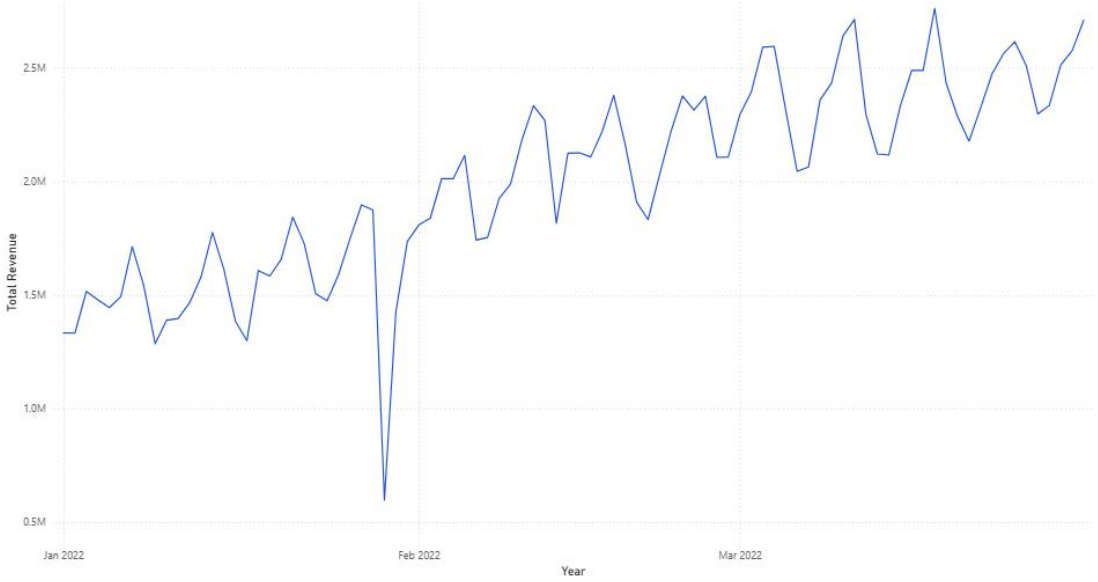

Fig 2: Daily Revenue Trend

5.2 Weekly Demand Patterns

Demand varies systematically across days of the week. Certain weekdays consistently record higher trip volumes, while others show reduced activity. This indicates predictable weekly behavior in taxi usage.

Such patterns are valuable for operational planning and capacity allocation.

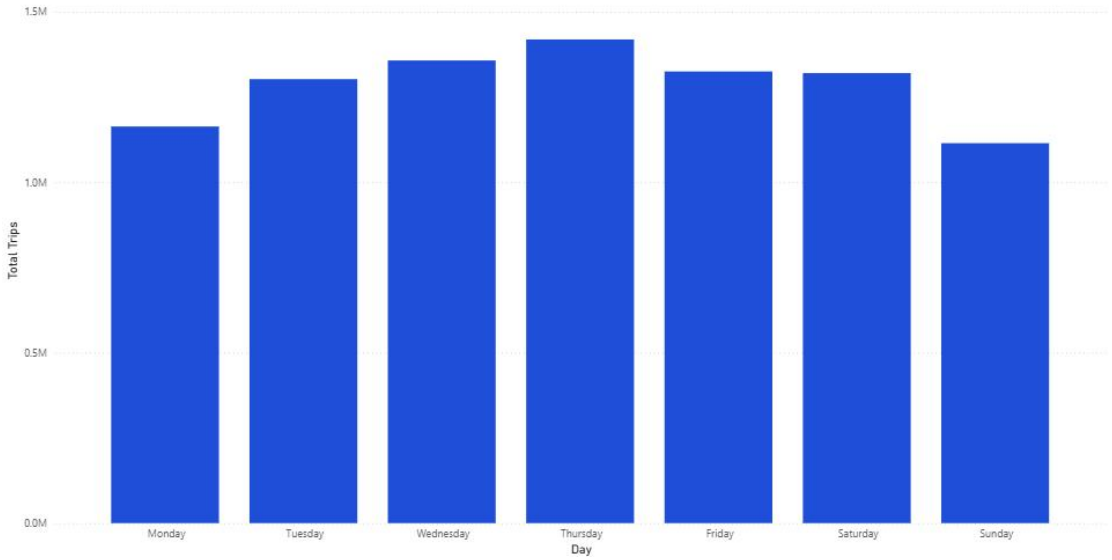

Fig 3: Trips by Day of Week

5.3 Intra-Month Demand Behavior

Demand distribution across days within a month reveals moderate variation, with no single day dominating activity. This suggests that taxi usage remains relatively stable throughout the month, with minor peaks and dips rather than extreme concentration.

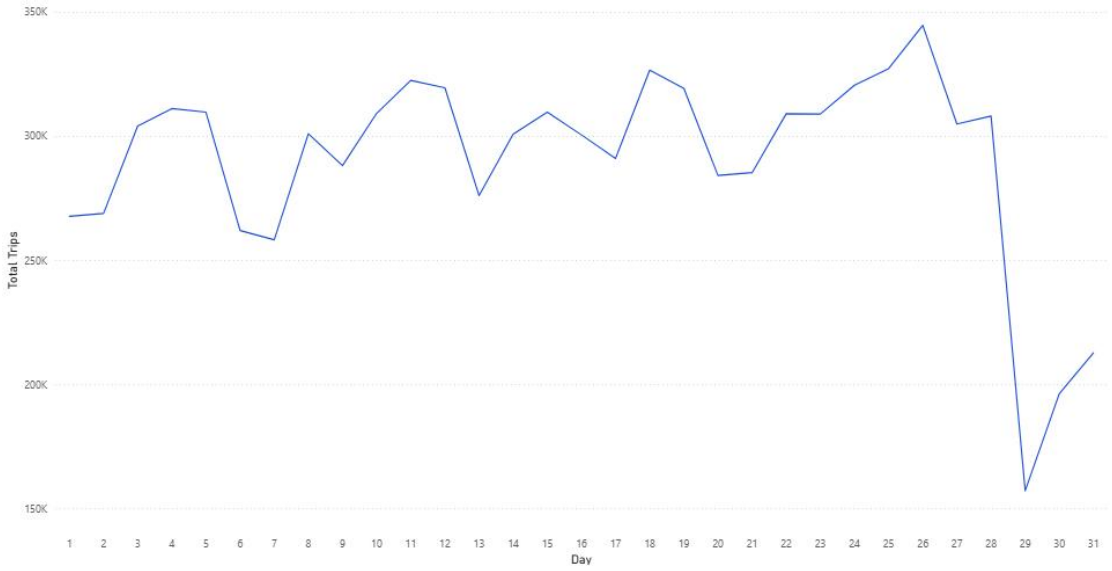

Fig 4: Average Trips by Day of Month

5.4 Trip Characteristics & User Behavior

Analysis of trip characteristics shows that most taxi trips cluster within a narrow duration range, indicating a predominance of short to medium-length journeys. This aligns with typical urban transportation use cases.

Fare levels increase proportionally with trip distance, reinforcing expected pricing behavior.

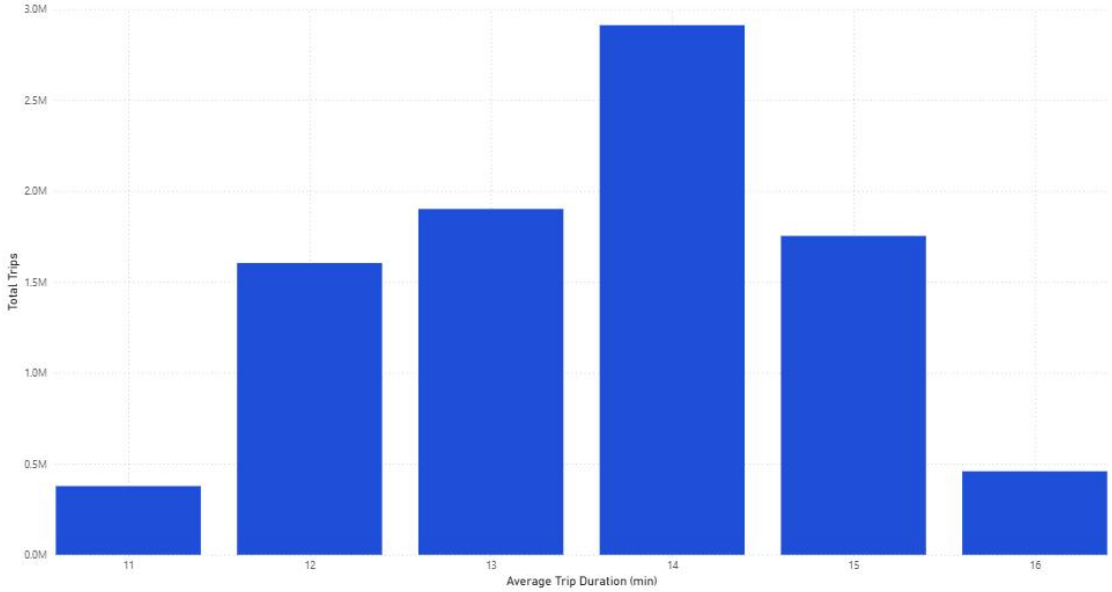

Fig 5: Trip Duration Distribution

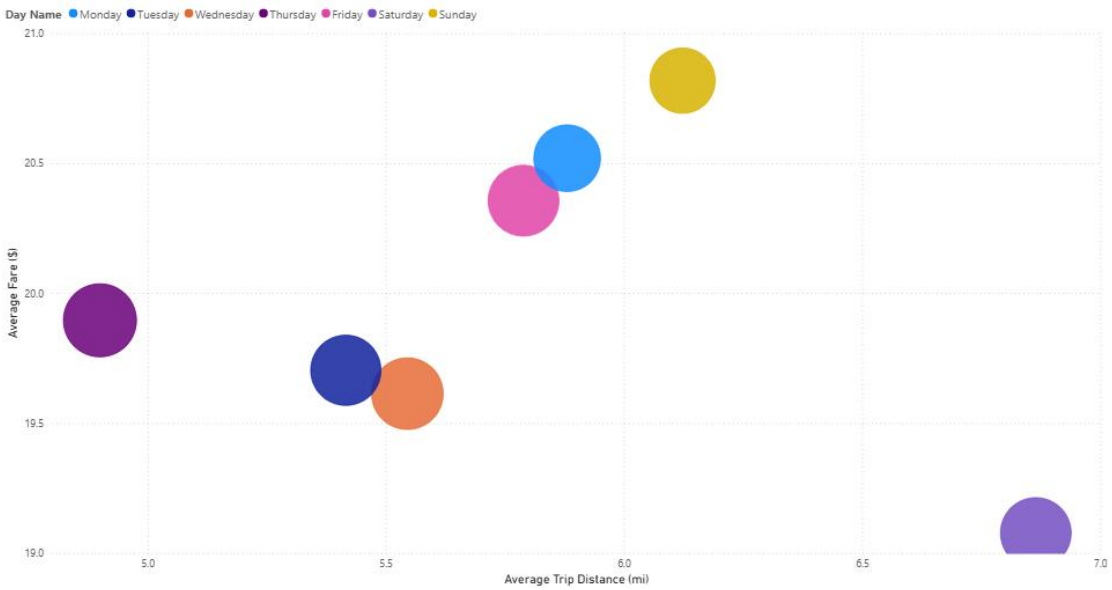

Fig 6: Fare vs Distance Relationship

5.5 Summary of Exploratory Insights

Key observations from the exploratory phase include:

- Taxi demand and revenue move closely together
- Weekly demand patterns are consistent and predictable
- Intra-month demand is relatively stable
- Most trips are short to medium in duration
- Fare levels scale logically with trip distance

6. Demand Forecasting

To extend insights beyond historical analysis, a time-series forecasting approach was applied to estimate short-term future taxi demand. Forecasting focuses on total daily trips, which serves as a direct proxy for demand intensity.

The forecasting step demonstrates how historical patterns can be leveraged to anticipate near-term operational needs.

6.1 Forecasting Approach

A time-series model was trained using historical daily trip data. The model captures:

- Trend components in taxi demand
- Short-term temporal dependencies
- Recent demand behavior

The resulting forecasts provide expected daily demand values for upcoming periods.

6.2 Forecast Interpretation

The forecast indicates that near-term taxi demand remains broadly consistent with recent historical levels, with moderate fluctuations rather than abrupt shifts. This suggests stable demand conditions in the short term.

While forecasts do not represent exact outcomes, they provide a useful planning signal for anticipating demand volume and resource allocation.

7. Dashboard Integration & Interpretation

All analytical outputs were consolidated into a single-page Power BI dashboard designed for executive-level consumption.

The dashboard enables users to:

• Monitor key performance indicators at a glance

• Identify demand and revenue trends

• Understand behavioral usage patterns

• Explore insights interactively across time

Rather than presenting raw analysis, the dashboard focuses on decision-ready information.

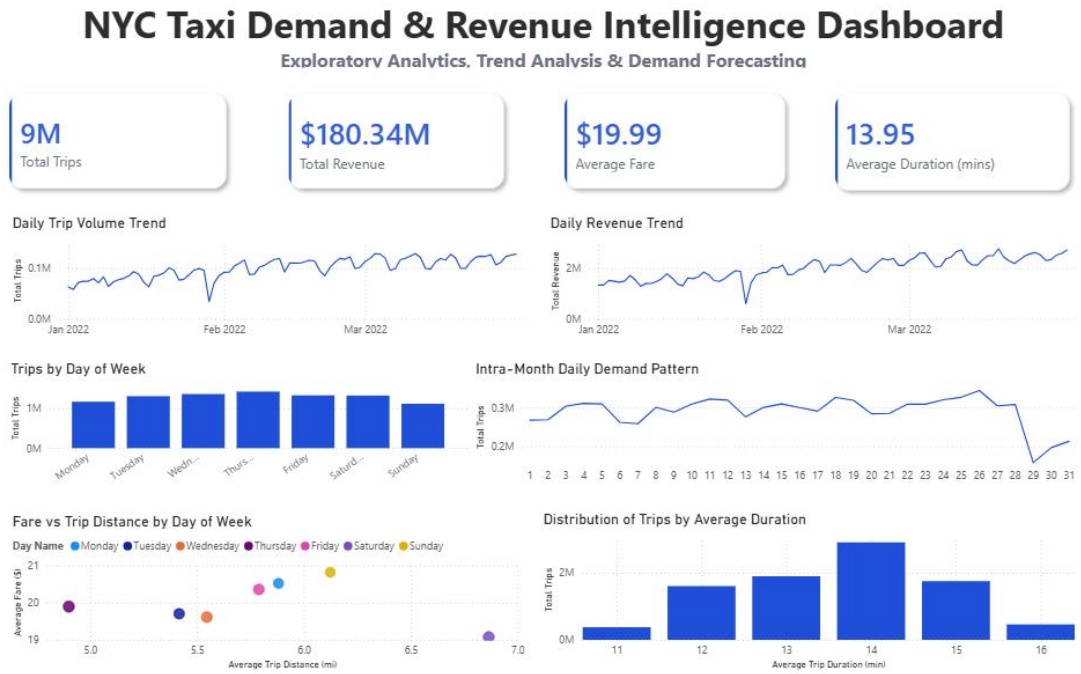

Fig 7: Final Power BI Dashboard Overview

8. Business Value & Practical Implications

The analysis and forecasting outputs can support:

• Transportation demand planning

• Short-term operational decision-making

• Revenue trend monitoring

• Capacity and workforce allocation strategies

By combining historical insights with predictive signals, stakeholders gain both retrospective understanding and forward-looking awareness.

9. Limitations

• Forecasts rely on historical patterns and do not account for external shocks

• The analysis is based on aggregated daily data rather than individual trips

• Long-term forecasting accuracy may decrease beyond the short horizon

These limitations are typical of time-series forecasting and should be considered when interpreting results.

10. Conclusion

The data indicates that NYC taxi demand is generally stable over time, with daily fluctuations remaining within a predictable range. Trip volume does not exhibit extreme volatility, suggesting that demand during the observed period is largely driven by routine urban travel patterns rather than irregular or disruptive events.

Revenue closely tracks changes in trip volume, implying that overall financial performance is primarily demand-driven. Fare levels scale consistently with trip distance, and most journeys fall within short-to-medium duration ranges, reinforcing the dominance of typical city travel over long-distance trips.

Clear temporal patterns are present in the data. Taxi usage varies systematically across days of the week, reflecting consistent behavioral routines, while demand across days within a month is relatively evenly distributed. Short-term forecasts align with these historical patterns, indicating that near-term demand is likely to remain consistent with established trends rather than undergoing abrupt changes.