NYC Taxi Demand: Insights & 2026 Forecast

A Power BI dashboard + SARIMAX forecasting, written for business stakeholders

Prepared by: Abdelmoutaleb BenLyazid
Tools: Power BI (Power Query + DAX), Python (SARIMAX)
Data source: Open-source public dataset (NYC taxi trips)
Date: December 27, 2025

What this report is: a clear explanation of what I built, why it matters, and what decisions we can take from it.
What this report is not: a purely technical document. Technical choices are explained, but always linked to business value.

Contents

1 Executive Summary (Human Version)

This project was built as if we were advising a mobility operator during a major event season: **measure the impact**, **find the hotspots**, **understand when demand peaks**, and **forecast what comes next**.

Using an open-source NYC taxi dataset, I delivered:

- an interactive **Power BI dashboard** (4 pages) that tells the story from performance to operations,
- a **SARIMAX forecast** to estimate 2026 demand and revenue indicators under an event scenario,
- a client-ready narrative that turns charts into decisions.

In simple terms, the dashboard shows **where the city heats up (zones)**, **when it heats up (hours/days)**, and **how big the 2026 demand could be**, so operations can plan capacity and leadership can plan targets.

2 Project Goal and Why It Matters

2.1 Business Question

If we are a mobility business (or a taxi operations team), we care about three things during major events:

1) **Demand:** How many trips will happen?
2) **Revenue:** What happens to fare totals during event periods?
3) **Operations:** Where and when should we position drivers to avoid shortages and lost revenue?

2.2 Project Objectives

This project answers those questions with four deliverables:

- **Impact analysis (2025 vs 2024):** quantify uplift in Trips, Tips, and Fare.
- **Hotspot analysis:** highlight zones with the strongest activity (including airport-related activity).
- **Peak timing:** identify the hour $\times$ day combinations that consistently produce peaks.
- **Forecast (2026):** produce a planning baseline using SARIMAX to support staffing and budgeting.

3 Data Source and What We Used

3.1 Data Source (Open Source)

The dataset used is an **open-source public NYC taxi trips dataset**. It contains time-based trip activity and monetary fields that can be used as demand and revenue proxies.

3.2 What the Data Represents (In Plain Language)

Each record captures taxi activity with:

- **Time** (date and time that allow daily, monthly, hourly analysis),
- **Location/Zone** (where trips are concentrated),
- **Trips** (demand volume),
- **Fare and Tips** (commercial outcomes).

3.3 Modeling Approach in Power BI

To keep results consistent and easy to filter:

- I used a **star-schema style model**: a central fact table (trips) linked to date and zone dimensions.
- I created measures (DAX) for KPIs like totals and uplift percentages.
- I organized the report into pages that match how a stakeholder thinks: *Overview* $\rightarrow$ *Where* $\rightarrow$ *When* $\rightarrow$ *What's next.*

4 Forecasting Method (SARIMAX) Explained Simply

4.1 What SARIMAX Does

SARIMAX is a time-series forecasting method that:

- learns patterns over time (trends + repeating seasonality),
- uses past values to predict future values,
- can also incorporate external signals (*exogenous variables*), like an event indicator or scenario intensity.

4.2 Why SARIMAX Was a Good Choice Here

Taxi demand behaves like a real-life rhythm: it repeats weekly, reacts to special days, and changes with events. SARIMAX is a good match because it is designed to model exactly that kind of structured pattern.

4.3 Forecast Outputs

The forecast produces 2026 planning baselines for:

- **Trips (demand)**,
- **Fare (revenue proxy)**,
- **Tips (behavioral, more volatile)**.

Important note for clients: forecasts reduce uncertainty, they do not remove it. This report treats forecasting as a planning tool, not as a promise.

5 Dashboard Walkthrough (Page by Page) + Where to Insert Images

5.1 Page 1: Overview — 2025 vs 2024 (Impact Snapshot)

What this page is for: a leadership view. If someone has 30 seconds, this page tells them whether the event period boosted performance and by how much.

What the visuals communicate:

- total Trips, Tips, and Fare for 2024 and 2025,
- uplift percentages that translate change into a single clear number,
- a monthly comparison view to see how uplift behaves across the period.

Insights you can say (example script):

- "We see a clear uplift in trips during the event period compared to baseline."
- "Fare uplift is strong, meaning revenue increases with the event effect."
- "Tips move differently and can be more volatile, so we treat them as a secondary signal."

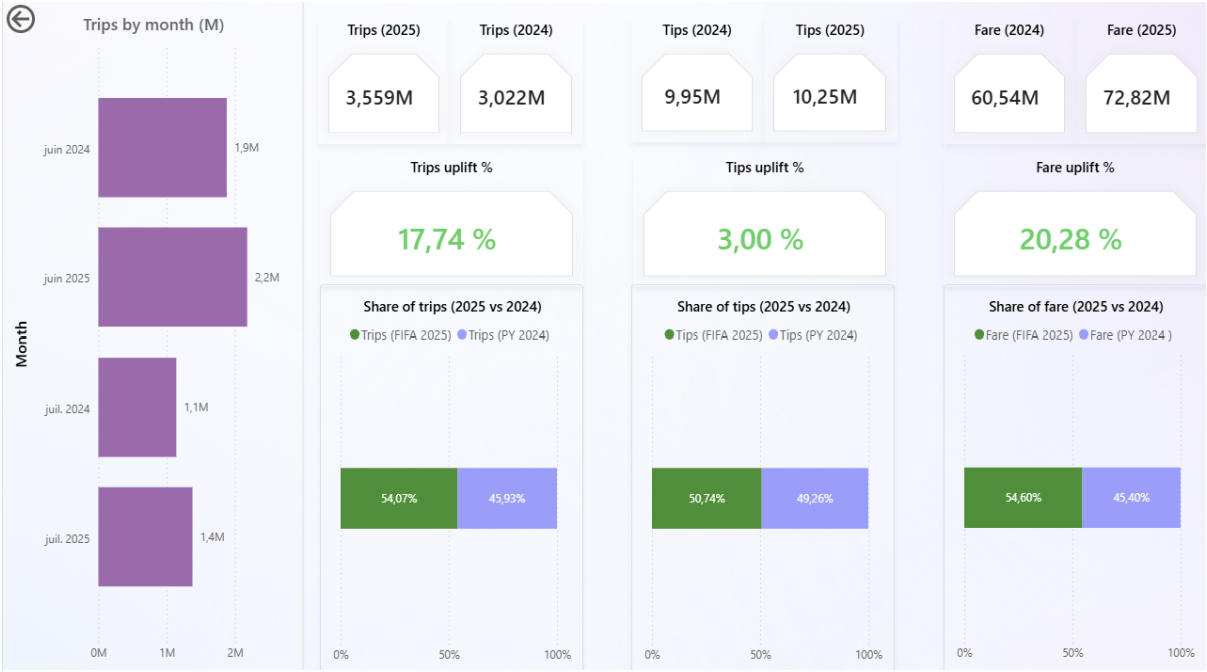


Figure 1: Power BI Page 1 — Overview (2025 vs 2024)

5.2 Page 2: Zones - Performance & Hotspots

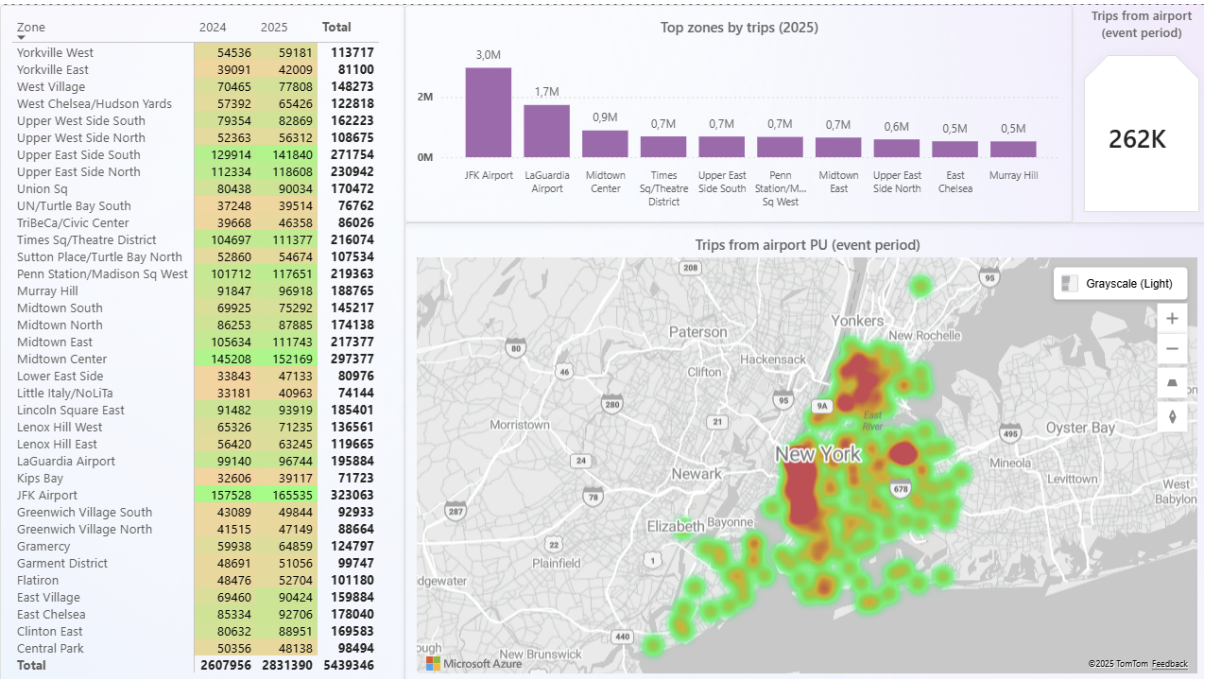
What this page is for: a geographic decision view. It answers: *Where is the demand concentrated?*

What the visuals communicate:

- a ranked view of zones (which zones drive the most trips),
- an airport-related KPI (useful for event visitors and travel peaks),
- a hotspot map that visually reveals concentration.

Insights you can say:

- "Demand is not evenly distributed: a small number of zones contribute a large share of activity."
- "Airport-related activity is significant, which supports targeted staging near airports during peak windows."
- "The heatmap confirms clusters that should be operational priority zones."



Zone	2024	2025	Total
Yorkville West	54536	59181	113717
Yorkville East	39091	42009	81100
West Village	70465	77808	148273
West Chelsea/Hudson Yards	57392	65426	122818
Upper West Side South	79354	82869	162223
Upper West Side North	52363	56312	108675
Upper East Side South	129914	141840	271754
Upper East Side North	112334	118608	230942
Union Sq	80438	90034	170472
UN/Turtle Bay South	37248	39514	76762
TriBeCa/Civic Center	39668	46358	86026
Times Sq/Theatre District	104697	111377	216074
Sutton Place/Turtle Bay North	52860	54674	107534
Penn Station/Madison Sq West	101712	117651	219363
Murray Hill	91847	96918	188765
Midtown South	69925	75292	145217
Midtown North	86253	87885	174138
Midtown East	105634	111743	217377
Midtown Center	145208	152169	297377
Lower East Side	33843	47133	80976
Little Italy/NoLiTa	33181	40963	74144
Lincoln Square East	91482	93919	185401
Lenox Hill West	65326	71235	136561
Lenox Hill East	56420	63245	119665
LaGuardia Airport	99140	96744	195884
Kips Bay	32606	39117	71723
JFK Airport	157528	165535	323063
Greenwich Village South	43089	49844	92933
Greenwich Village North	41515	47149	88664
Gramercy	59938	64859	124797
Garment District	48691	51056	99747
Flatiron	48476	52704	101180
East Village	69460	90424	159884
East Chelsea	85334	92706	178040
Clinton East	80632	88951	169583
Central Park	50356	48138	98494
Total	**2607956**	**2831390**	**5439346**

Figure 2: Power BI Page 2 - Zones & Hotspots

5.3 Page 3: When Demand Peaks (Hour × Day)

What this page is for: an operations scheduling view. It answers: *When do we need more drivers?*

What the visuals communicate:

- the heatmap shows exactly which **hour** and **day** combinations produce peaks,
- max/min tiles provide quick operational reference,
- a day-of-week chart helps explain consistent weekly behavior.

Insights you can say:

- "Peak demand concentrates on specific time blocks, making scheduling decisions data-driven."
- "Weekends and late hours often show consistent pressure, which suggests a clear staffing strategy."
- "Knowing the minimum window also helps reduce overstaffing and idle time."

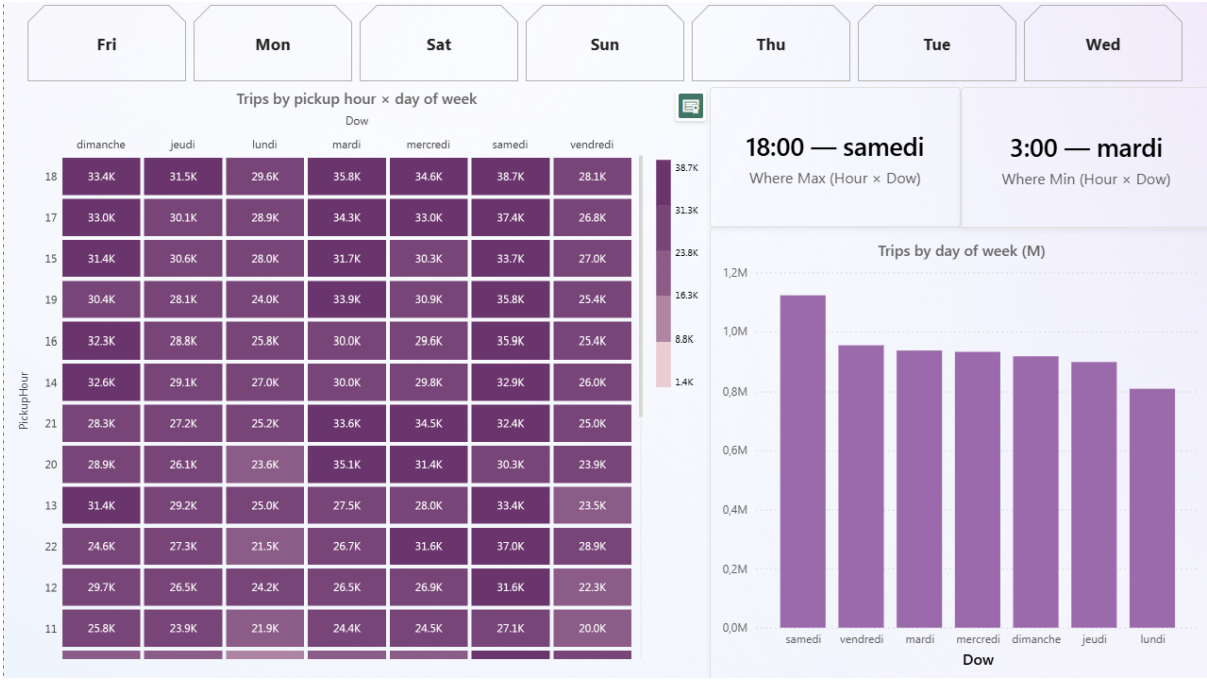


Figure 3: Power BI Page 3 - Peak Timing (Hour × Day

5.4 Page 4: 2026 Forecast & Scenarios

What this page is for: planning. It answers: *What should we expect next year, and what does it mean for capacity and revenue?*

What the visuals communicate:

- forecast KPIs for 2026 (Trips, Fare, Tips),
- uplift comparisons vs historical baselines,
- a clear side-by-side view of 2024 actual, 2025 actual, and 2026 forecast.

Insights you can say:

- "This gives us a baseline to plan 2026 staffing and revenue expectations."
- "Fare forecast supports budgeting; Trips forecast supports operational capacity."
- "Tips forecast is useful but less stable, so it is interpreted cautiously."

6 Forecast Visual Analysis (What to Say About the Graphs)

Below is a client-friendly way to interpret your forecasting charts (including baseline vs event vs forecast visuals):

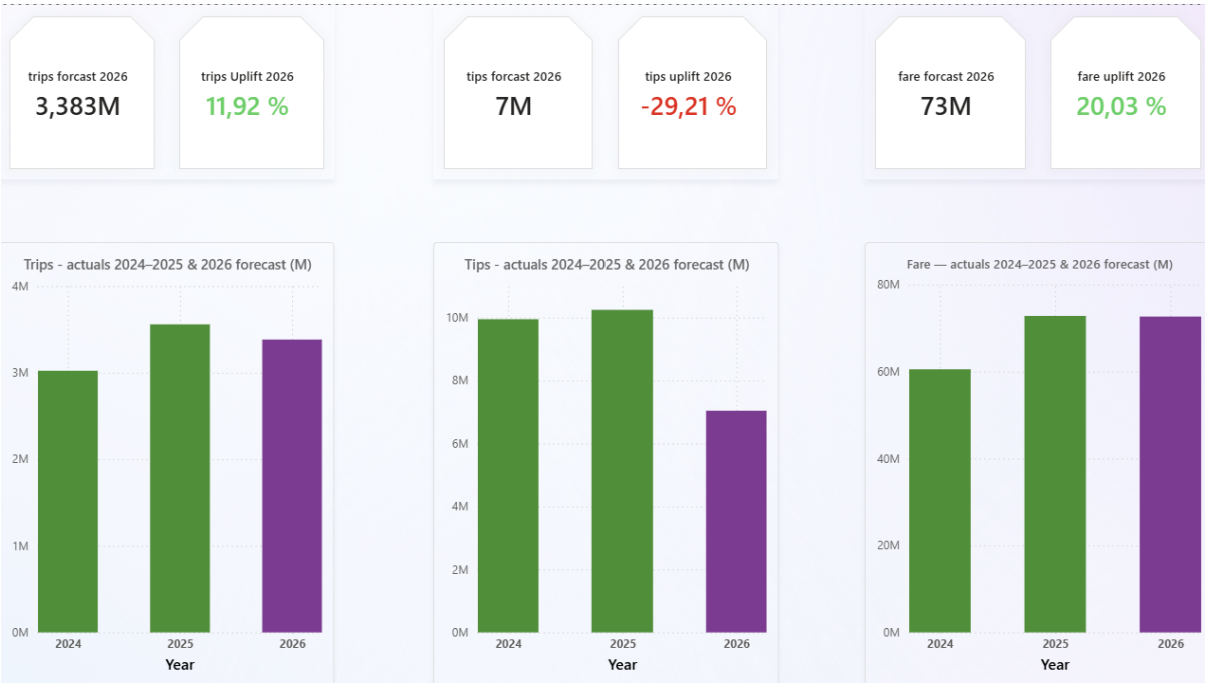


Figure 4: Power BI Page 4 — 2026 Forecast & Scenarios

6.1 Baseline vs Event vs Forecast (2024–2026)

- **2024 baseline** provides the "normal demand" reference.
- **2025 event period** shows how demand behaves when the city experiences an extra shock (tournament/event effect).
- **2026 forecast** estimates what could happen next under a scenario, using learned seasonality and the event signal.

6.2 Trips vs Fare vs Tips (Key Interpretation)

- **Trips** is the best operational KPI (capacity planning).
- **Fare** is the best business KPI (revenue planning).
- **Tips** is valuable, but more dependent on behavior and context, so it is naturally noisier.