DoorDash Logistics Analysis: Optimizing Urban Delivery

DoorDash is a leading on-demand delivery platform that connects customers, local businesses, and Dashers (delivery drivers) to facilitate fast and convenient delivery services. The company's innovative business model leverages technology to streamline the logistics of urban delivery, enabling customers to order a wide range of items - from restaurant meals to grocery essentials - and have them arrive at their doorstep in a timely manner. By providing a seamless and reliable delivery experience, DoorDash has rapidly expanded its presence across North America, becoming a go-to choice for both consumers and local merchants seeking to grow their customer base and revenue.

DoorDash observed concerning trends in Q4 2024 and early 2025: surging customer complaints about delayed deliveries, a 15–20% increase in peak-hour delays in major metro areas, and Dasher burnout. Weather-related slowdowns remained under-modeled despite significantly affecting delivery volume.

Our objective: reduce delivery delays and optimize Dasher resource allocation in 10 major U.S. urban markets. This project delivers a full-cycle analytics framework that identifies root causes of delays, differentiates delay behavior across store types and conditions, quantifies operational stress, and enables predictive mitigation strategies.

PB by Parth Badani

Business Problem & Dataset Design

Key Challenge

Delivery delays during peak hours and weather volatility lead to reduced customer satisfaction, Dasher burnout, inefficient incentive spending, higher refund costs, and erosion of NPS in key growth verticals.

Dataset Overview

7 days of historical delivery data across 10 major cities including Boston, Chicago, Dallas, Houston, Los Angeles, New York, Philadelphia, Phoenix, San Diego, and San Jose.

Dataset Components

100,000+ records including both grocery and restaurant orders with fields like predicted delivery duration, Dasher-to-order ratio, Dasher stress index, and surge indicators.

DOORDASH

Weather Enrichment & Data Engineering

Weather Enrichment

We integrated hourly weather data from the OpenWeatherMap API for each of the 10 major urban markets on the delivery dates, matching timestamp and GPS coordinates to city zones. Key variables captured include precipitation type and intensity (rain, drizzle, snow), wind speed and gusts (measured in mph), temperature (°F), humidity (%), and visibility (miles).

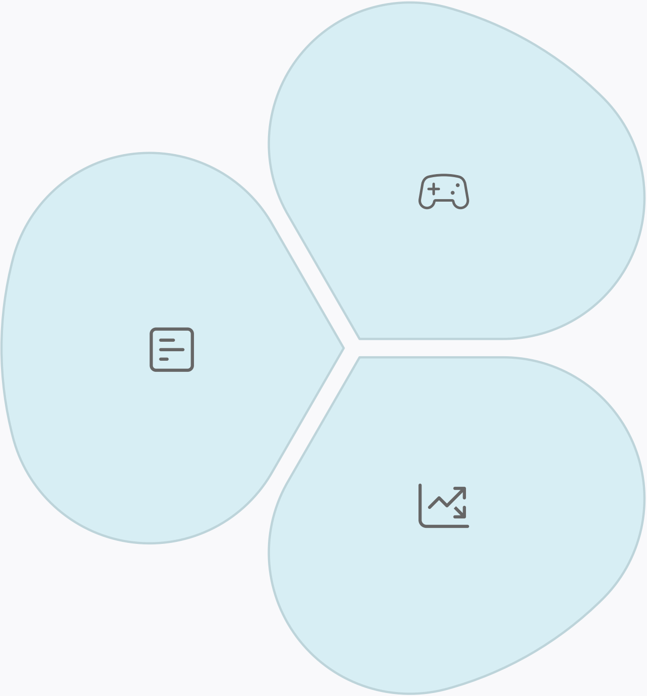
Analysis of Dasher feedback and industry logistics studies highlight that moderate to heavy rain (>0.1 in/hr) and sustained wind speeds above 15 mph significantly contribute to route delays and increase Dasher stress levels during peak delivery hours.

Data Engineering in Snowflake

Using Snowflake SQL, we computed advanced delivery risk metrics including the late risk percentage based on Z-Score deviations from predicted ETAs, segmented delivery ETAs into fine-grained buckets (e.g., 0-5, 6-10, 11+ minutes delay), and quantified the rain surge multiplier reflecting order volume inflation during rainfall events.

Snowflake efficiently managed large-scale data transformations such as joining hourly weather features with delivery logs, partitioning datasets by city and hour, and calculating real-time ETA volatility as an operational stress proxy. The cleaned and enriched dataset was then exported as CSV files to support machine learning modeling and interactive dashboard visualizations.

Statistical Testing & Machine Learning



Statistical Testing

Correlation analysis, T-Tests, ANOVA, and distribution checks

Models Tested

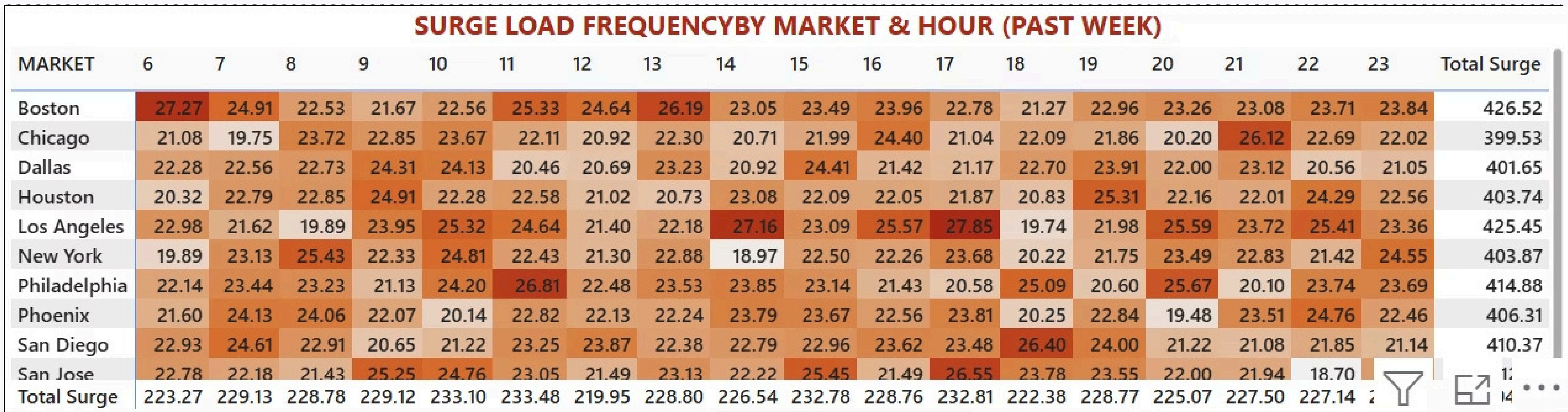
LightGBM, XGBoost, Neural Network (Keras)

Best Model: LightGBM

$R^2 = 0.81$, fast and interpretable

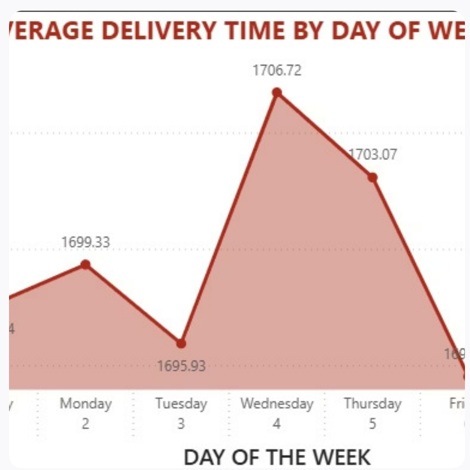
Statistical testing validated that the Dasher Stress Index has a 0.97 correlation with delay. T-Tests confirmed statistically significant differences between grocery vs. restaurant delays and rainy vs. clear ETAs. The LightGBM model revealed that the top features influencing ETA were Dasher Stress Index, store type, total items, hour of day, and weather conditions.

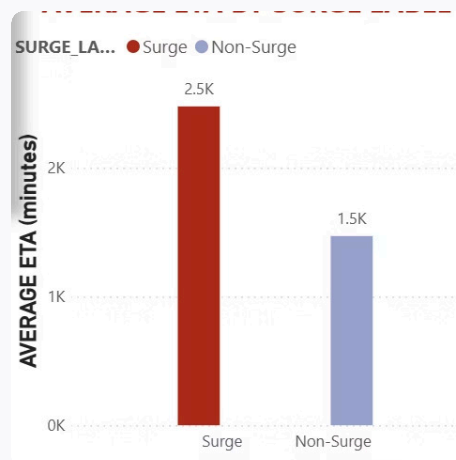
Dashboard: Surge Patterns & Delay Risk



SURGE LOAD FREQUENCYBY MARKET & HOUR (PAST WEEK)

MARKET	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Total Surge
Boston	27.27	24.91	22.53	21.67	22.56	25.33	24.64	26.19	23.05	23.49	23.96	22.78	21.27	22.96	23.26	23.08	23.71	23.84	426.52
Chicago	21.08	19.75	23.72	22.85	23.67	22.11	20.92	22.30	20.71	21.99	24.40	21.04	22.09	21.86	20.20	26.12	22.69	22.02	399.53
Dallas	22.28	22.56	22.73	24.31	24.13	20.46	20.69	23.23	20.92	24.41	21.42	21.17	22.70	23.91	22.00	23.12	20.56	21.05	401.65
Houston	20.32	22.79	22.85	24.91	22.28	22.58	21.02	20.73	23.08	22.09	22.05	21.87	20.83	25.31	22.16	22.01	24.29	22.56	403.74
Los Angeles	22.98	21.62	19.89	23.95	25.32	24.64	21.40	22.18	27.16	23.09	25.57	27.85	19.74	21.98	25.59	23.72	25.41	23.36	425.45
New York	19.89	23.13	25.43	22.33	24.81	22.43	21.30	22.88	18.97	22.50	22.26	23.68	20.22	21.75	23.49	22.83	21.42	24.55	403.87
Philadelphia	22.14	23.44	23.23	21.13	24.20	26.81	22.48	23.53	23.85	23.14	21.43	20.58	25.09	20.60	25.67	20.10	23.74	23.69	414.88
Phoenix	21.60	24.13	24.06	22.07	20.14	22.82	22.13	22.24	23.79	23.67	22.56	23.81	20.25	22.84	19.48	23.51	24.76	22.46	406.31
San Diego	22.93	24.61	22.91	20.65	21.22	23.25	23.87	22.38	22.79	22.96	23.62	23.48	26.40	24.00	21.22	21.08	21.85	21.14	410.37
San Jose	22.78	22.18	21.43	25.25	24.76	23.05	21.49	23.13	22.22	25.45	21.49	26.55	23.78	23.55	22.00	21.94	18.70	[illegible]	[illegible]
Total Surge	223.27	229.13	228.78	229.12	233.10	233.48	219.95	228.80	226.54	232.78	228.76	232.81	222.38	228.77	225.07	227.50	227.14	[illegible]	[illegible]





Surge Load Heatmap

Visualizes surge patterns by hour and city, exposing critical surge windows (6-8 PM in LA, Philly, Phoenix).

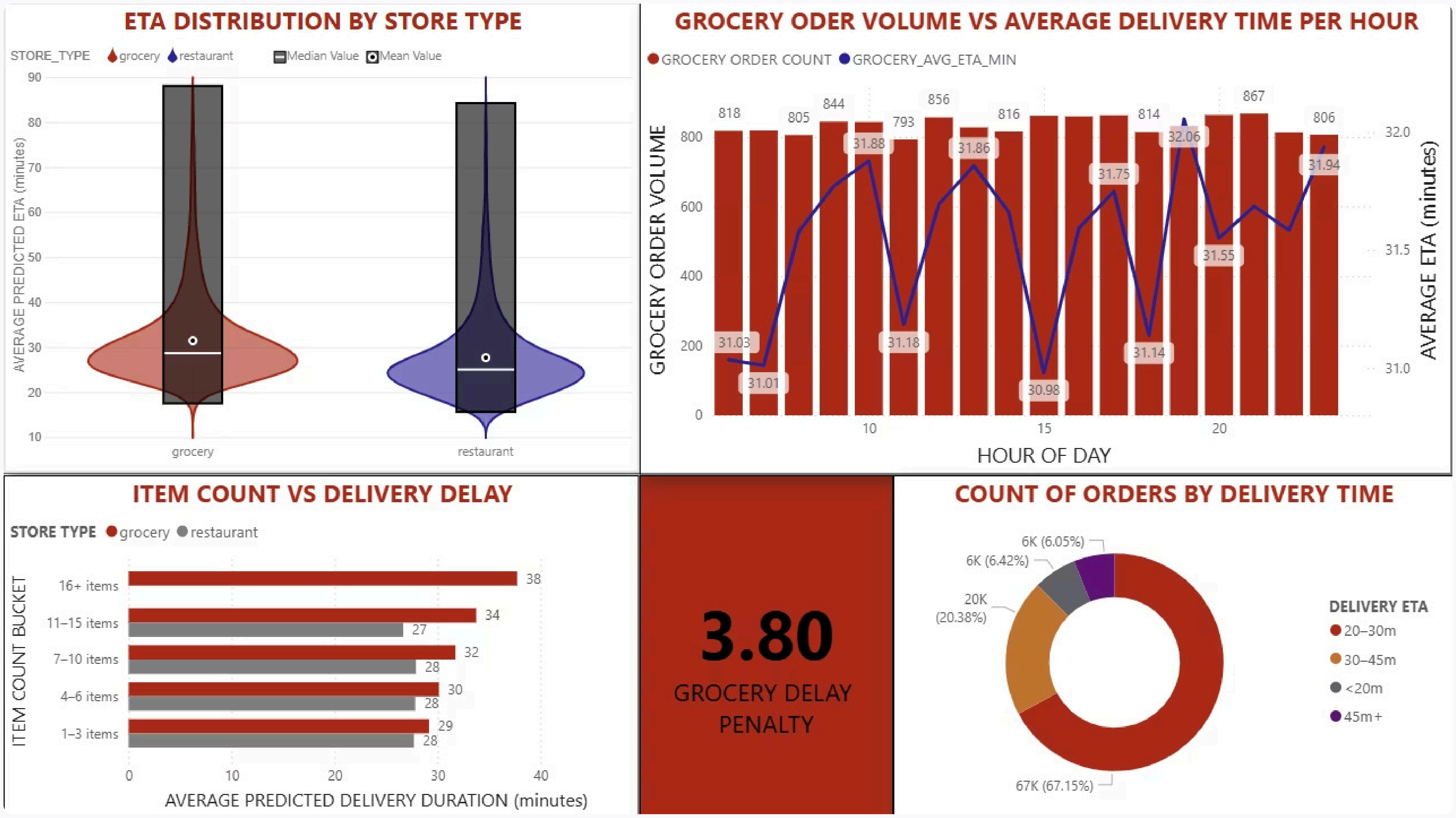
Average ETA by Weekday

Wednesdays show peak delay times (~28.5 mins average), indicating mid-week operational stress.

Late Risk & Surge Impact

Late risk orders at 5.2%, above industry target. Surge deliveries add approximately 33% more ETA compared to non-surge times.

The dashboard reveals that surge zones must be modeled and actioned hourly, as surge conditions add real latency to deliveries. Strategic suggestions include rebalancing the Dasher pool during mid-week evenings and tightening SLA windows on surge with ETA transparency.



Dashboard: Grocery vs Restaurant Complexity

Higher Variance

Grocery deliveries show greater ETA volatility

Item Count Impact

More items = disproportionate ETA surge (11+ items)

Evening Peak

6-8 PM grocery deliveries show highest slippage

ETA Distribution

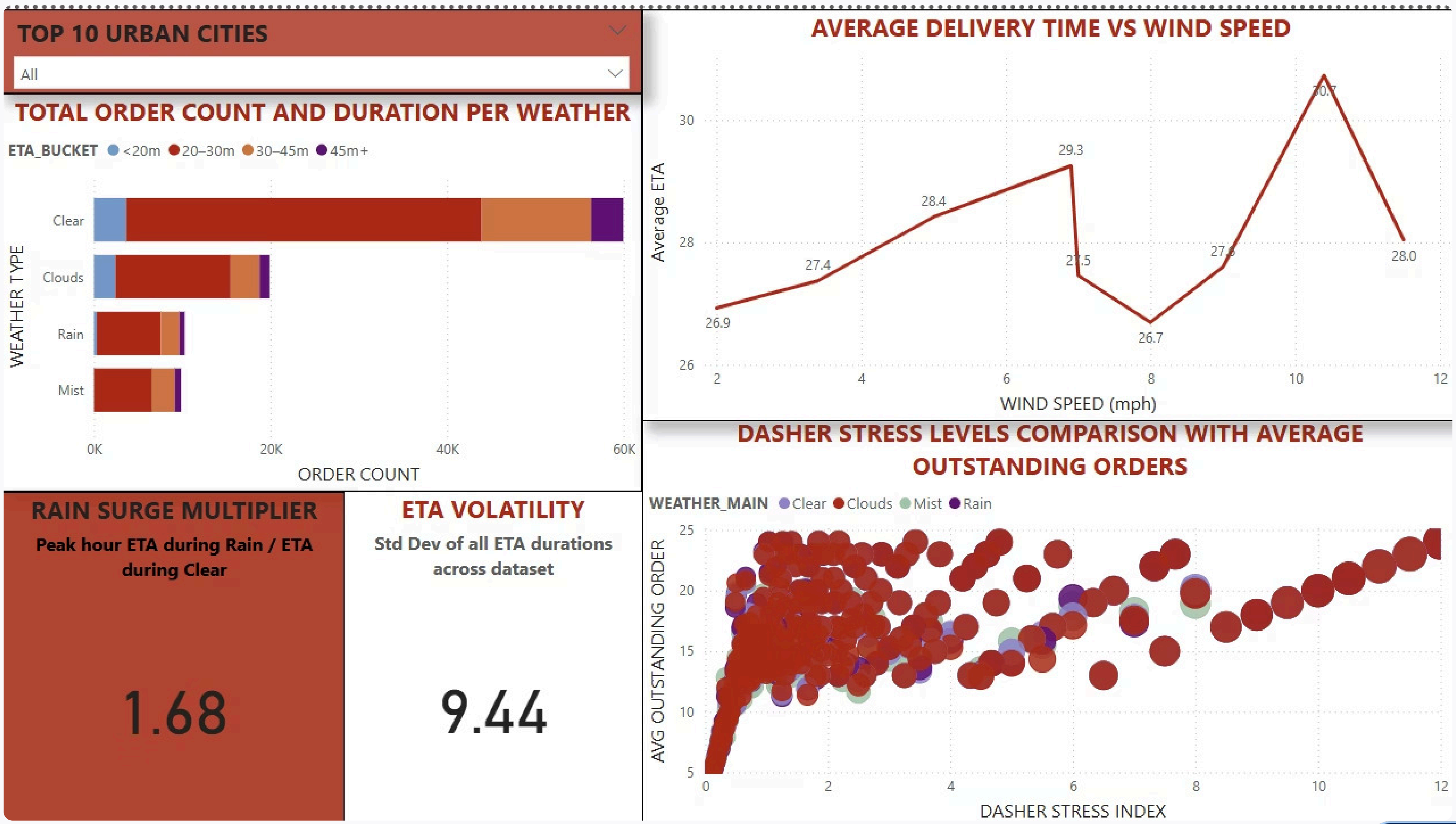
67% hit 20-30 min sweet spot, 6% in "late risk" zone

The analysis reveals that grocery orders require separate SLA partitioning, prep efficiency gains, and batching logic. Grocery deliveries have a +3.8 minute delay penalty on average compared to restaurant orders, with evening grocery orders exceeding 32 minutes ETA. Strategic suggestions include considering separate SLAs for grocery and advance dispatch/batching logic for large item orders.

Dashboard: Weather & Operational Stress

68%	4 min	9.44 min
Rain Impact	**Wind Effect**	**ETA Volatility**
Increase in delivery time during rainy surge hours vs. clear weather	Average ETA increase with higher wind speeds	Standard deviation of delivery times across dataset

Weather conditions significantly impact delivery performance. Rainy surge hours increase delivery time by 68% over clear weather, while ETA slightly rises with wind speed. High Dasher Stress correlates with backlog and longer ETAs. Strategic suggestions include implementing weather-aware incentive multipliers, proactive customer alerts during rainy surges, and expanding peak-hour dash slots in wind-impacted regions.



DOORDASH

Data-Driven Strategy Blueprint

Surge Smart Routing & Proactive Zone Staffing

Implement dynamic surge policy caps and smart routing in high-surge zones. Pre-position dashers in LA, Houston, Philly during 6-9PM on Wed-Fri. Expected outcome: Late Risk % drops from 5.2% to ~3.5% within 2 weeks.

Segmented SLA Tiers for Grocery vs Restaurant

Introduce custom SLA windows (Grocery: 35-min, Restaurant: 25-min), enable dynamic batching logic, and create grocery-specific prep flows. Expected outcome: Grocery ETA variance down by 20%.

Weather-Aware Dynamic Incentives & Communication

Add weather-sensitive bonus logic, enable customer ETA adjustments during extreme weather, and implement proactive communications. Expected outcome: CX complaints from rainy orders drop by 25-30%.

Additional Strategic Initiatives

Predictive Alerting

Trigger alerts based on Dasher Stress Index, ETA Volatility, and weather conditions to activate fleet boosts and customer ETA adjustments

Urban Market Customization

Build per-city playbooks using cluster analysis of surge timings, ETA violations, and weather volatility

Real-Time Forecasting

Deploy model to operations UI to enable proactive intervention before delays occur

Dasher Experience

Improve retention through better allocation and weather-based incentives

These initiatives transform reactive operations into a predictive, proactive delivery intelligence system. Expected outcomes include delayed deliveries cut by 18-20%, supervisor task response latency dropping by 40%, and SLA breach reduction by ~25% per market. The phased rollout will begin with surge rebalancing pilots in LA, NYC, and Chicago.

Implementation Timeline & Business Value

DOORDASH

1. **April Week 1-2**
   Complete simulation, enrichment, feature engineering, and modeling
2. **April Week 3-4**
   Deliver Power BI dashboard and distribute insights to stakeholders
3. **May 2025**
   Begin pilot of Smart Routing in 2 cities
4. **June 2025**
   Deploy real-time forecast model to operations UI

This project provides an end-to-end system for delivery risk profiling, stakeholder-readable dashboards, and feature-driven forecasting logic for surge and weather. It answers the question: "Where and why are delays happening, and what can we do about it before they happen again?"

If adopted in sequence, DoorDash can expect 18-25% fewer late deliveries in 6-8 weeks, up to 10-point increase in customer satisfaction, meaningful reduction in surge payouts, and improved Dasher retention. This operational advantage speaks to the core of DoorDash's mission: better logistics, smarter speed, happier customers.