

Fair Fare: A Deep Dive into Taxi Price Fluctuations and Congestion Pricing in New York City

Understanding Pricing Dynamics and Enhancing Fairness in Urban Mobility
Team 3: Crystal Leatvanich, Courtney Vincent, Jooyeon Lee, Christina Son

Enhancing Fare Transparency for Trust & Confidence

By improving fare transparency, we aim to help riders better anticipate costs, reduce unexpected expenses, and ultimately enhance trust and satisfaction in the yellow taxi system.



Introduction to Taxi Pricing



Fares fluctuate based on:
Time & Distance
Surcharges (tax, toll, congestion)
Time of Day
Ride Characteristics



The lack of transparent pricing in NYC taxis **creates confusion** and **distrust** among passengers, leading many to choose Uber and Lyft for their upfront fare estimates.

Enhancing Fare Transparency for Trust & Confidence

By improving fare transparency, we aim to help riders **better anticipate costs, reduce unexpected expenses, and ultimately enhance trust and satisfaction** in the yellow taxi system.



Exploratory Data Analysis (EDA) & Feature engineering

Data overview

Original Data → Sampled Data
2023 Yellow Taxi Trip Data
NYC open data
~38 million rows

Random sampling
200k rows, 19 columns

→ Cleaned Data



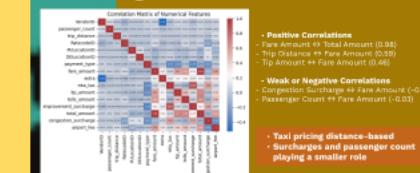
Cleaned Data

- Dropped the missing value
- Filtered negative values
- Updated some columns to correct type

Feature Engineering

- Time**
 - time_of_day: Morning / Afternoon / Night
 - day_type: Weekday(Mon-Fri)/Weekend(Sat-Sun)
- Distance**
 - trip_length: short, long
 - (based on travel duration, with the mean as the threshold)
- Passenger**
 - Number of passengers: Single, Couple, Group
- Location**
 - PU_Borough / DO_Borough: Pickup/Drop-off NYC zones
 - interborough_trip: Same Borough, Different borough
 - distance_category: Close, Moderate, Far
- Surchage fee**
 - congestion_surchage: Free, No Fee

EDA Findings



- Positive Correlations
 - Fare Amount ↔ Total Amount (0.98)
 - Trip Distance ↔ Fare Amount (0.59)
 - Tip Amount ↔ Fare Amount (0.49)
- Weak or Negative Correlation
 - Congestion Surchage ↔ Fare Amount (-0.16)
 - Passenger Count ↔ Fare Amount (-0.03)

- Taxicab pricing distance-based
- Surcharges and passenger count playing a smaller role

Data overview



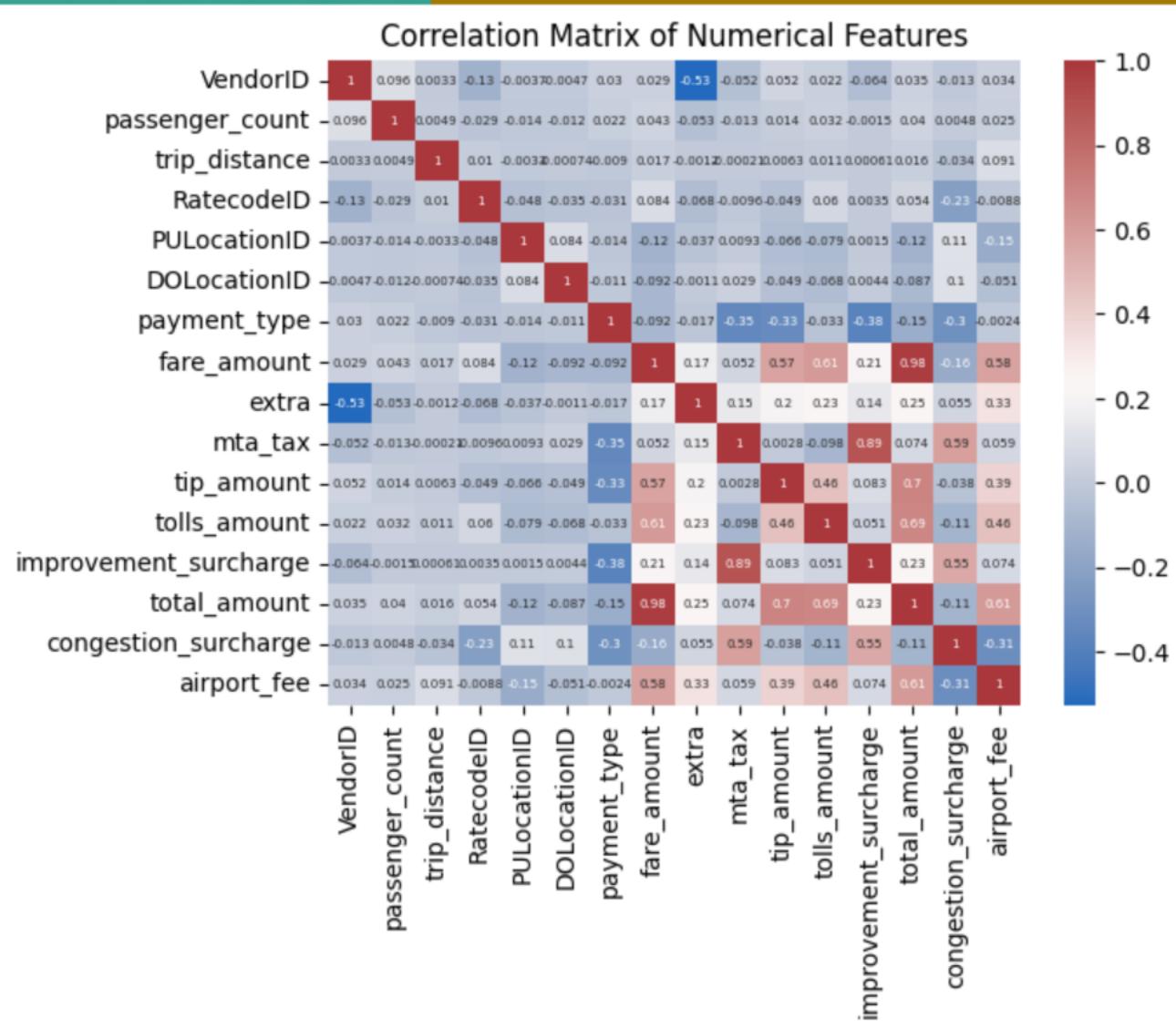
Original Data → Sampled Data → Cleaned Data

2023 Yellow Taxi Trip Data
NYC open data
~38 million rows

Random sampling
200k rows, 19 columns

- Dropped the missing value
- Filtered negative values
- Updated columns to correct type

EDA Findings



- **Positive Correlations**

- Fare Amount \leftrightarrow Total Amount (0.98)
- Trip Distance \leftrightarrow Fare Amount (0.59)
- Tip Amount \leftrightarrow Fare Amount (0.46)

- **Weak or Negative Correlations**

- Congestion Surcharge \leftrightarrow Fare Amount (-0.16)
- Passenger Count \leftrightarrow Fare Amount (-0.03)

- **Taxi pricing distance-based**
- **Surcharges and passenger count playing a smaller role**

Feature Engineering

Time

- time_of_day: Morning / Afternoon / Night
- day_type: Weekday(Mon-Fri)/Weekend(Sat-Sun)

Distance

trip_length: short, long
(Based on travel duration, with the mean as the threshold)

Passenger

Number of passengers: Single, Couple, Group

Location

- PU_Borough / DO_Borough: Pickup/Drop-off NYC zones
- interborough_trip: Same Borough, Different Borough
- distance_category: Close, Moderate, Far

Surcharge fee

congestion_surcharge: Fee, No Fee



Analysis & Experiments



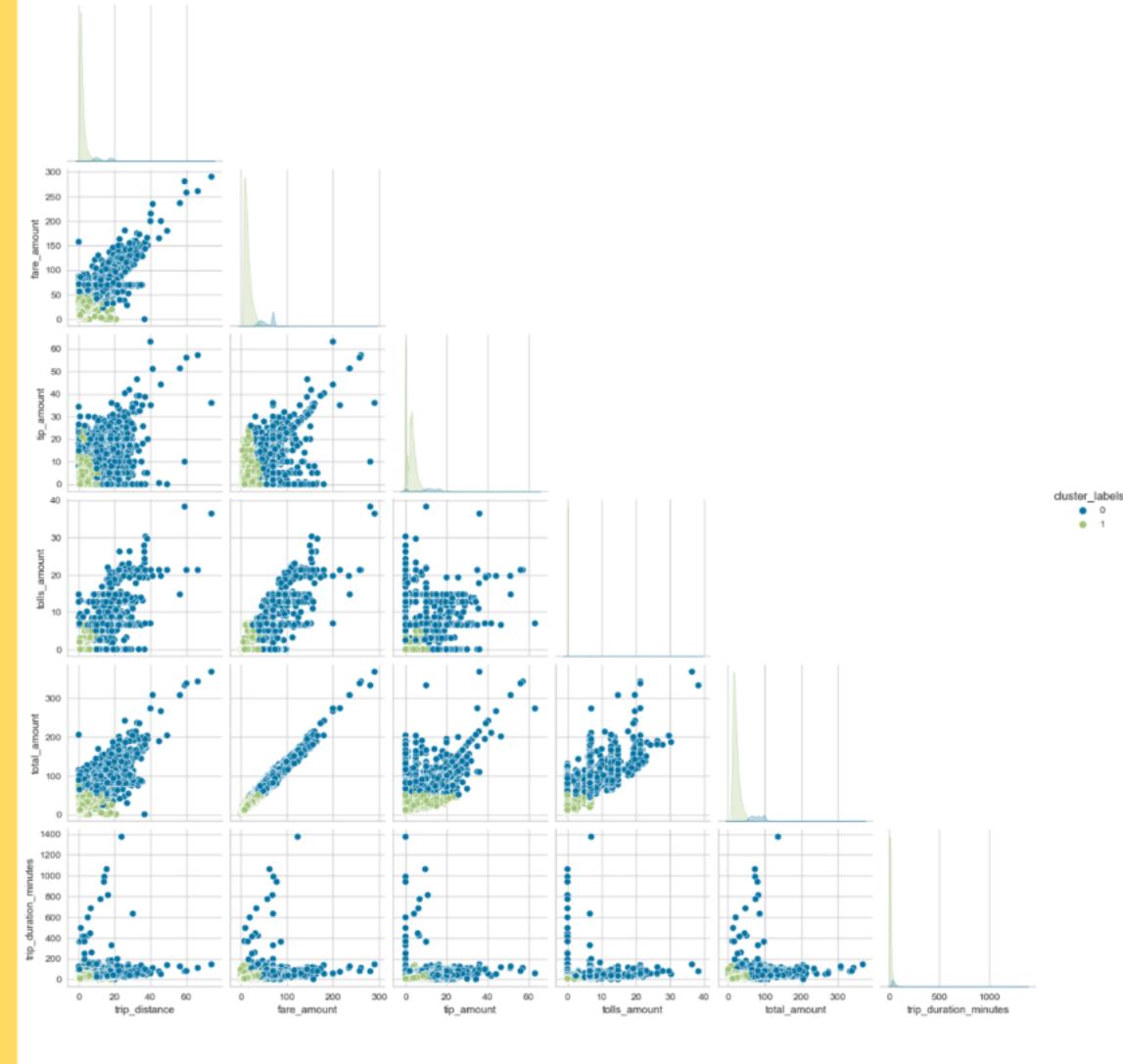
	trip_distance	fare_amount	tip_amount	tolls_amount	total_amount	trip_duration_minutes
PC1	0.952517	0.973621	0.763068	0.793500	0.991510	0.800972
PC2	-0.093186	-0.021533	0.468284	-0.515506	0.012786	0.185738
PC3	0.036635	0.094718	-0.429291	-0.215838	-0.041524	0.515502

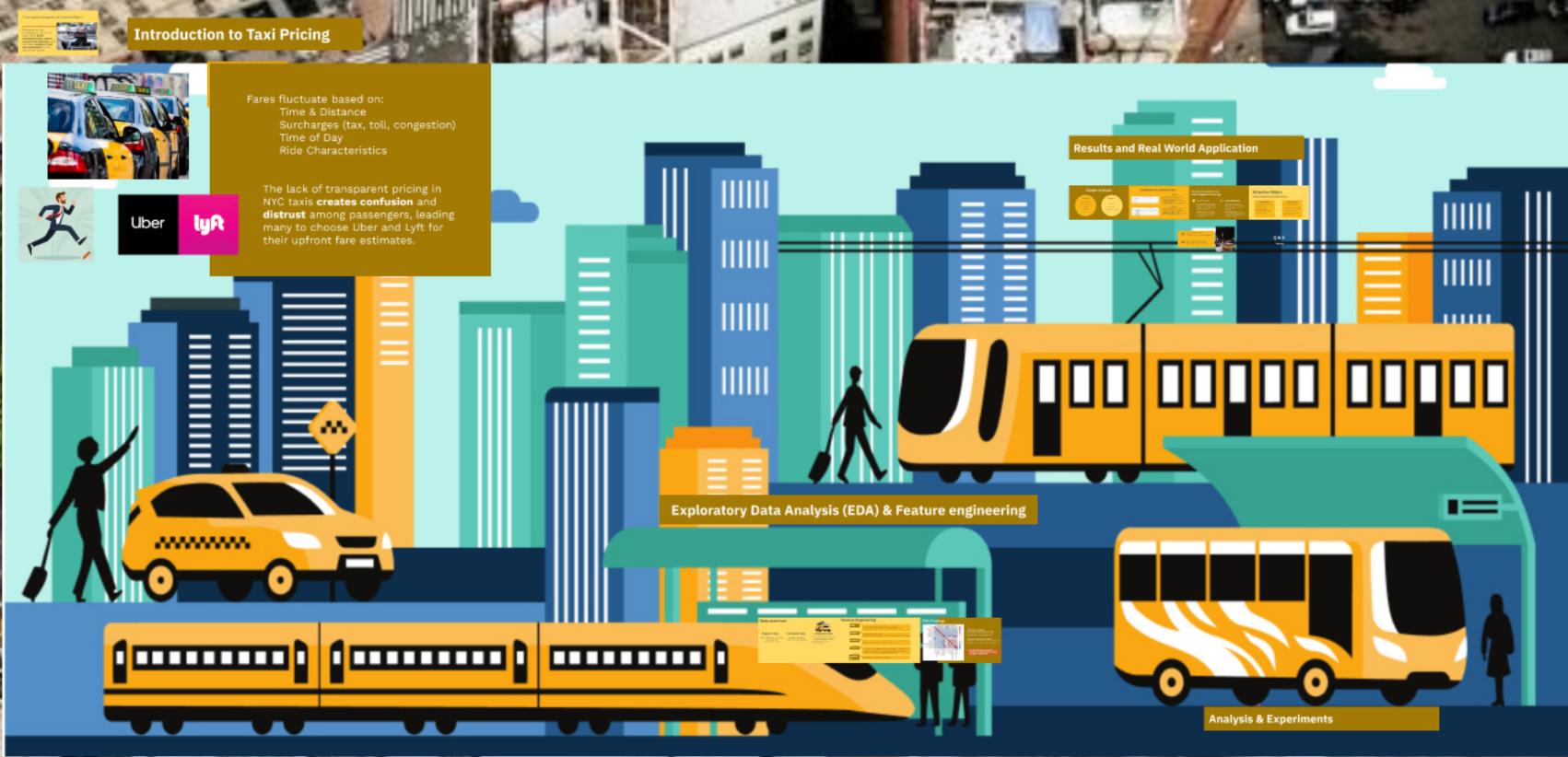
PCA

- Reconstruction Error: Identified 5% of the data as outliers.
- Revised PCA Run:
 - Scree Plot 3 components for a more optimal representation.
- PC 1: Positively contributes to each numerical feature, captures an overall trip cost and length dimension

K-Means Clustering

- Elbow Method: 2 Cluster
- **Cluster 1: Shorter Rides, lower cost**
- **Cluster 0: Longer rides, higher cost**





Fair Fare: A Deep Dive into Taxi Price Fluctuations and Congestion Pricing in New York City

Understanding Pricing Dynamics and Enhancing Fairness in Urban Mobility
Team 3: Crystal Leatvanich, Courtney Vincent, Jooyeon Lee, Christina Son

Results and Real World Application

Cluster Analysis

- Cluster 0
Interborough Trips
Congestion Surcharge
• Passenger
• Walker
• Driver
- Cluster 1
Intra-city
Intra-borough Trips
Congestion Surcharge

Congestion Surcharge Cause



Real World Applications to avoid Congestion Surcharge

1 Smart Routing

- Strategic Pick-up/drop-off Points: Drivers can choose the best route to avoid traffic jams and reach their destination to avoid extra fees.
- Multiple Transportation Methods: Encourage passengers to use public transit and ride bicycles instead of driving during surcharge hours.

Attention Riders

To Avoid Congestion Costs We Recommend:

2 Policy Adjustments

- Time-based Prioritization: Offer discounts or incentives for drivers who avoid peak traffic hours. During the day, drivers can take advantage of lower fares.
- Congestion Exemptions: Consider a congestion exemption for certain vehicles or individuals during peak hours.

Q & A

Thank You

- 01 Granular Analysis of Congestion Surcharge Triggers
- 02 Sentiment Analysis on NYC Yellow Taxi Resources

Cluster Analysis

Cluster 0

Inter-borough Trip

Congestion Surcharge

1 Passenger

Weekday

Queens

Cluster 1

Manhattan

Intra-borough Trip

Congestion Surcharge

Congestion Surcharge Cause

antecedents	consequents	antecedent support	consequent support	support	confidence	lift
(DO_Borough_Queens, PU_Borough_Manhattan)	(congestion_surcharge_type_Fee)	0.208123	0.78585	0.198784	0.955127	1.215406
(DO_Borough_Queens, PU_Borough_Manhattan, interborough_trip_1)	(congestion_surcharge_type_Fee)	0.208123	0.78585	0.198784	0.955127	1.215406
(day_type_Weekday, PU_Borough_Manhattan, DO_Borough_Queens)	(congestion_surcharge_type_Fee)	0.153608	0.78585	0.146386	0.952987	1.212683

Cluster 0

- High Congestion Surcharge Correlation:** Trips between Queens and Manhattan show the strongest link to congestion surcharges.
- Time & Weekday Patterns:** Morning and afternoon trips occasionally appear.
- Possible Commuting Trend:** Interborough trips in Cluster 0 may follow a commuting pattern.

antecedents	consequents	antecedent support	consequent support	support	confidence	lift
(day_type_Weekday, PU_Borough_Manhattan, time_of_day_Night, DO_Borough_Manhattan)	(congestion_surcharge_type_Fee)	0.267504	0.956167	0.264432	0.988518	1.033834
(day_type_Weekday, PU_Borough_Manhattan, time_of_day_Night, interborough_trip_0)	(congestion_surcharge_type_Fee)	0.267504	0.956167	0.264432	0.988518	1.033834
(day_type_Weekday, PU_Borough_Manhattan, interborough_trip_0, DO_Borough_Manhattan, time_of_day_Night)	(congestion_surcharge_type_Fee)	0.267504	0.956167	0.264432	0.988518	1.033834

Cluster 1

- Intra-Manhattan Trips:** Pickup and drop-off locations are within the same borough.
- Weekday Night Pattern:** Strong association with trips occurring on weekday nights.
- Entertainment & Nightlife Link:** Likely used for social outings and nightlife activities.

Real World Applications to avoid Congestion Surcharge

1 Smart Routing 2 Policy Adjustments

- **Strategic PickUp/Drop-off Points:** Instead of a direct pickup/drop-off, riders could go to a slightly outside surcharge zone and walk a short distance to avoid extra fees.
- **Multiple Transportation Methods:** Riders can use e-scooters, bikes, or short distance public transit and reduce reliance on taxis during surcharge hours.

- **Time-Based Promotions:** Offer discounts for taxi trips during peak surcharge hours. (Long trip users: Morning - Afternoon, Short trip users: Night)
- **Surcharge Exemptions:** Consider a pricing model that incorporates trip distance more effectively, adjusting base fares instead of applying additional surcharges.

Attention Riders

To Avoid Congestion Costs We Recommend:

Longer Rides

Avoid inter-borough travel
during rush hour

Combine multiple
transportation methods
(train, bus, taxi)

Shorter Rides

Avoid night time rides

Share rides when possible

Combine multiple
transportation methods
(train, bus, taxi)

01

Granular Analysis of
Congestion Surcharge Triggers

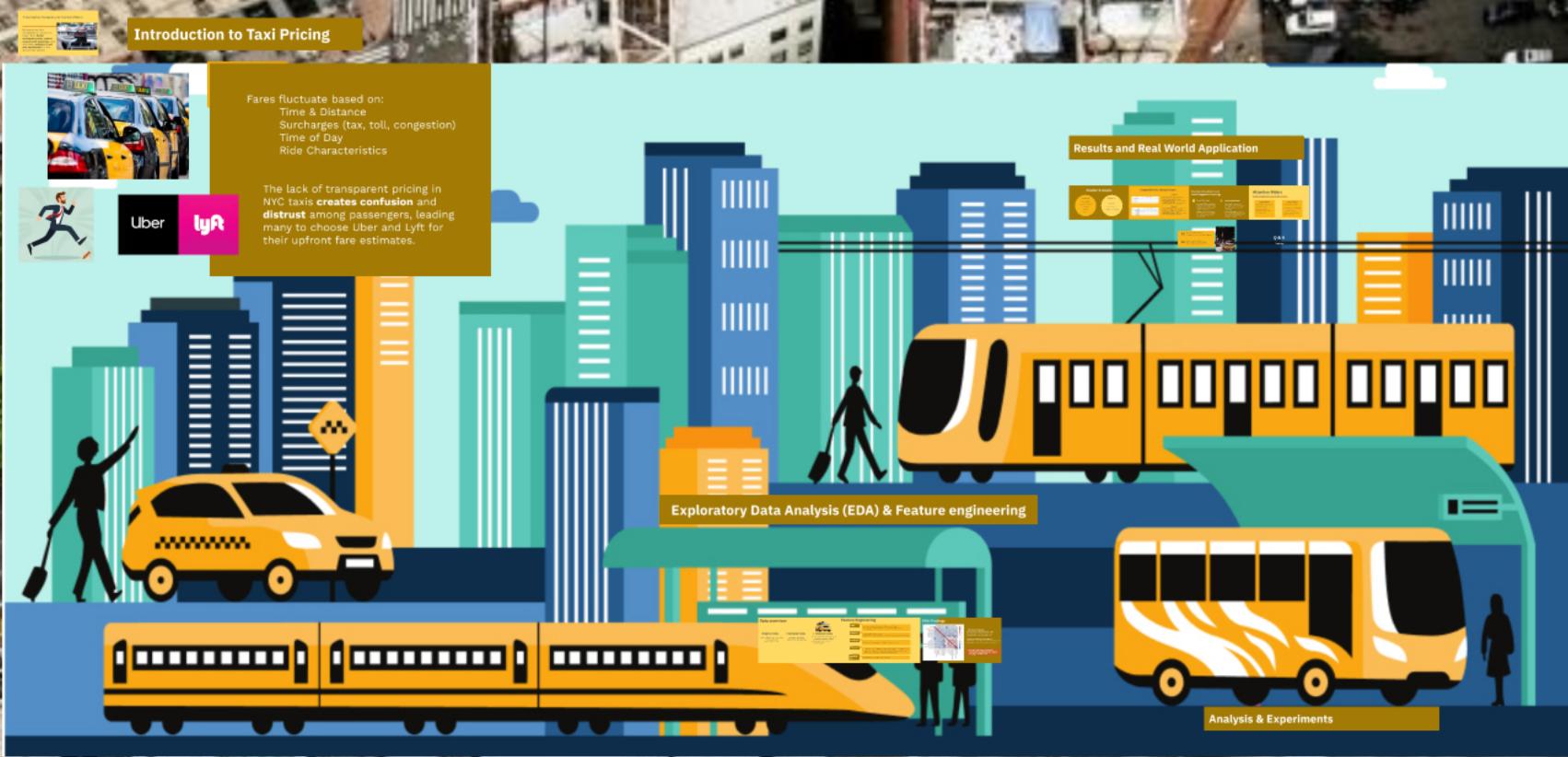
02

Sentiment Analysis on
NYC Yellow Taxi Reviews



Q & A

Thank You



Fair Fare: A Deep Dive into Taxi Price Fluctuations and Congestion Pricing in New York City

Understanding Pricing Dynamics and Enhancing Fairness in Urban Mobility
Team 3: Crystal Leatvanich, Courtney Vincent, Jooyeon Lee, Christina Son