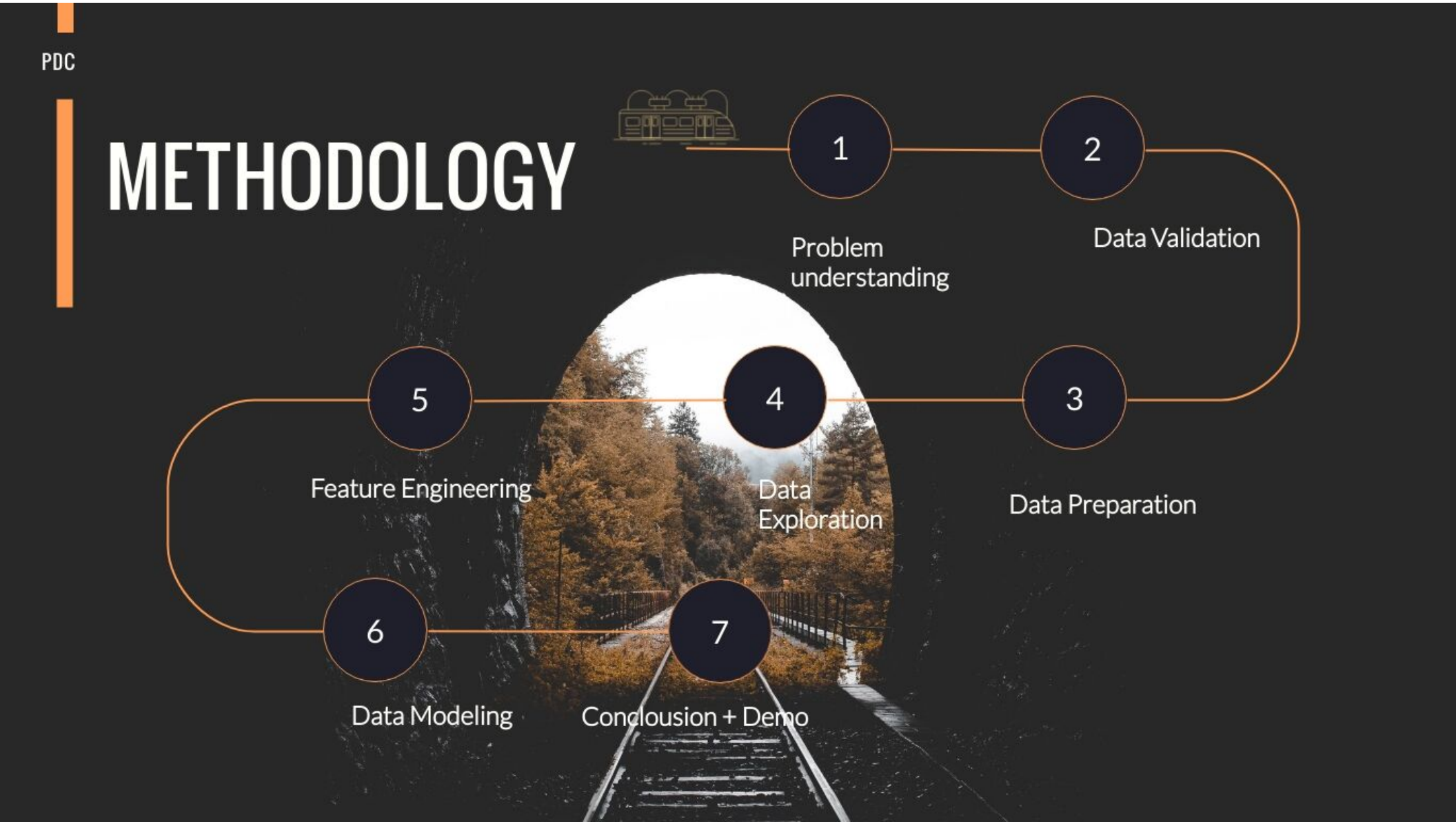

EDA OF MTA TURNSTILE DATA
NEW YORK CITY
PDC
ONE STOP SOLUTION
Abdultawwab Safarji
Oct 7, 2021
The crowds are still there
Really?

PDC

METHODOLOGY



PDC

PDC
ONE STOP SOLUTION

PROBLEM UNDERSTANDING

Overview:

- MTA commute plagued by delays as riders gripe over crowds and delays on lines led to dangerously crowded trains.

Problem Statment:

- Find the congestion and detected it on a live map.
- Discover demand of trains across all stations and further optimize the availability of trains to busy stations.
- Identify how to reduce congestion and delays on crowded station that waste commuters' time.

Scope:

- Due to the computational and time constraint as well as the unavailability of complete datasets. The MTA Turnstile and MTA maps will be used.

PDC

DATA EXPLORATION

Data preparation & Data Imputation

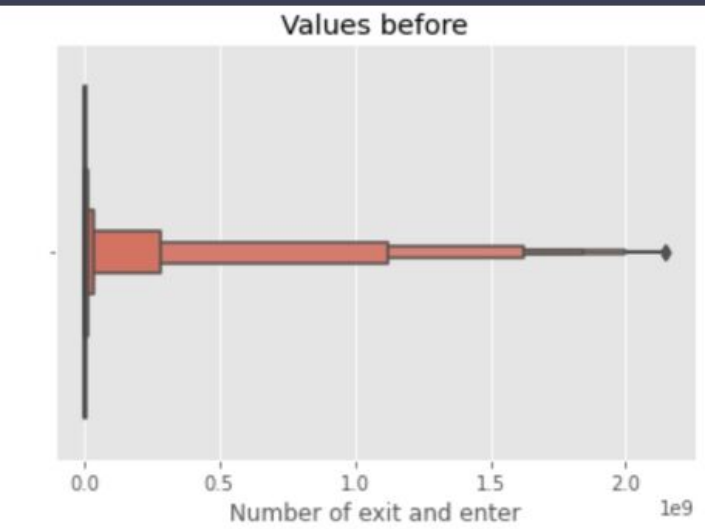
DATA VALIDATION

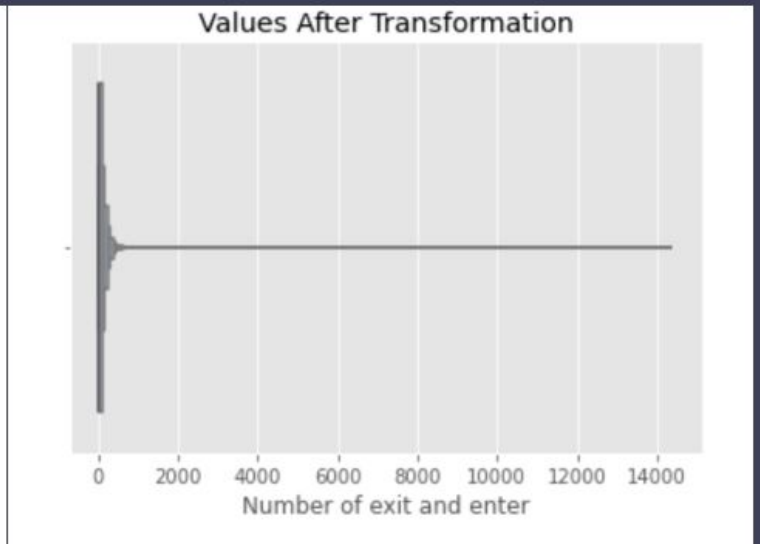
Validate for

Duplicate Data

Missing Values

outlier and Illogical Entries



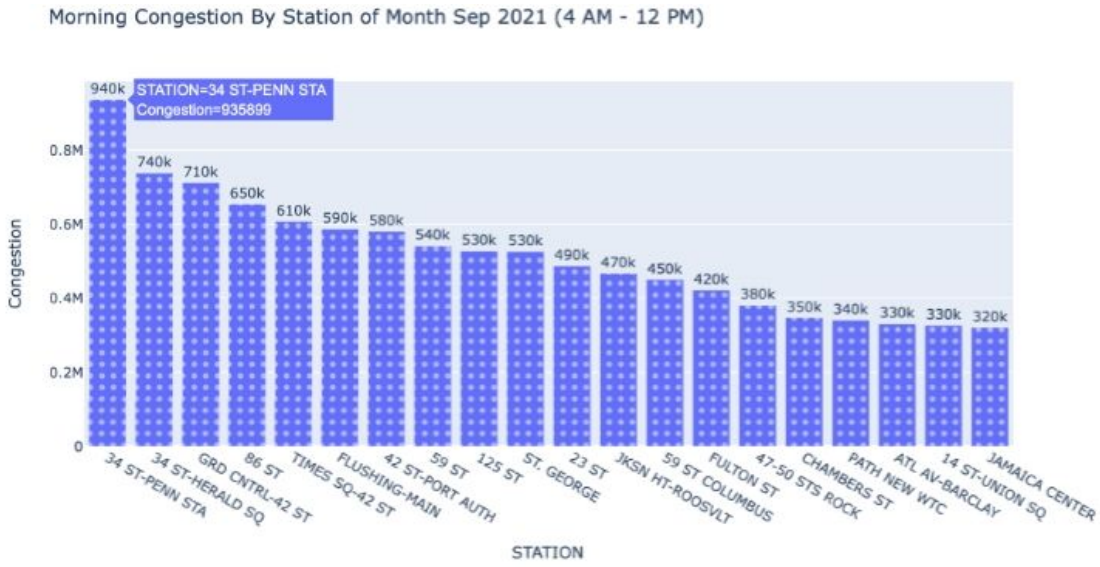


PDC

FEATURE ENGINEERING

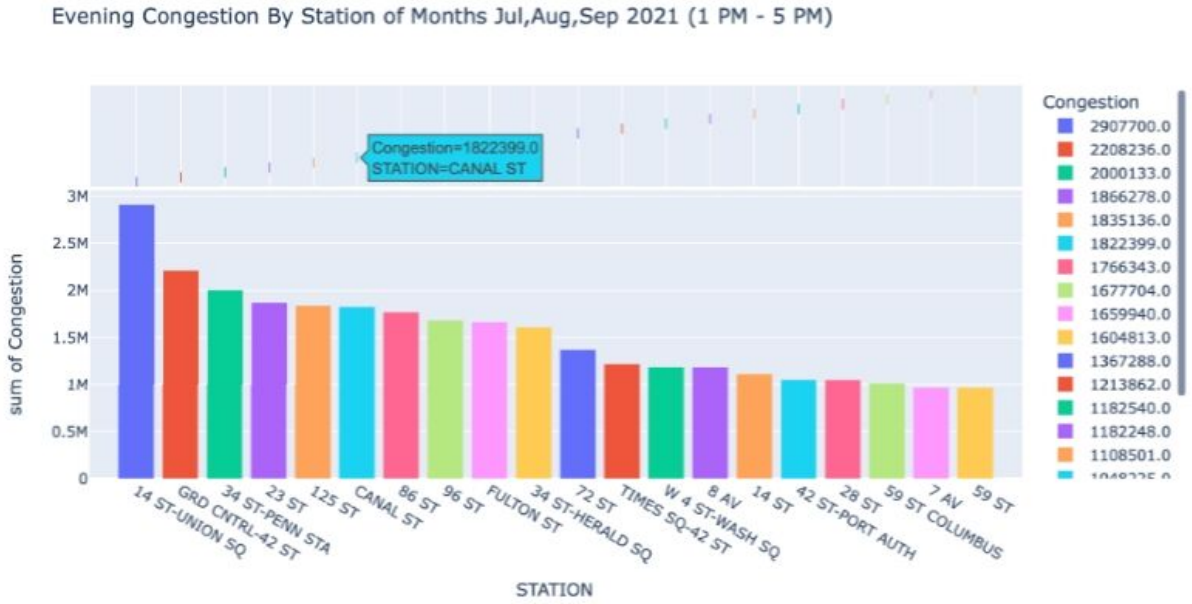
Observations:

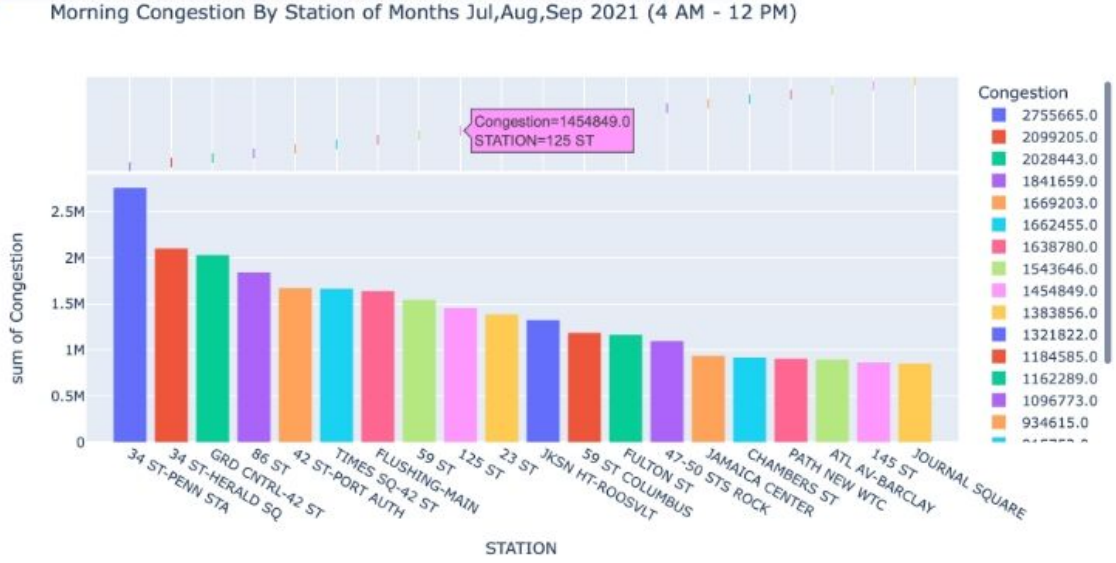
- Predicting most congestion stations of 1 mon.
- Split the time between evening and morning.
- Observing where to dis-congestion during a specific time frame.



PDC FEATURE IMPORTANCE

predicting the whole 3 months





Congestions change between times!

Morning	Evening	High to low
34 ST-PENN	14-St-Union	Highest
34 ST - HERALD Sq	Grand Center-42	Second
Grand Center-42	34 ST-PENN	Third

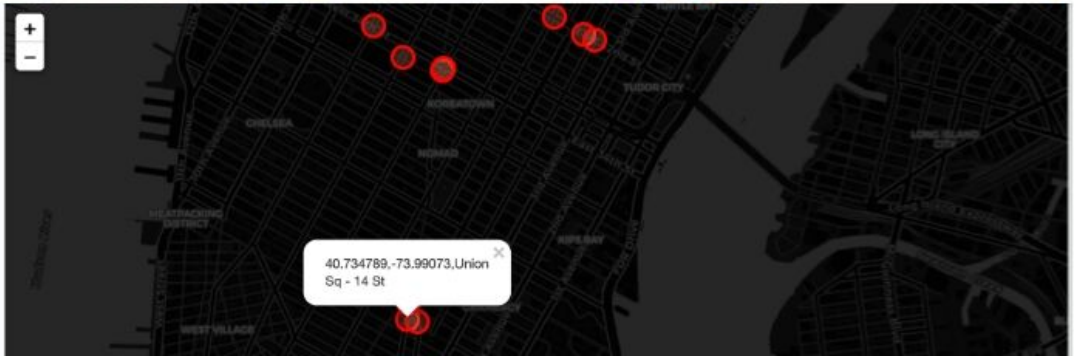
PDC

DATA MODELING

- Congestions are used to detect on the map between morning and evening.
- Fetching external transportations by map (denoted with blue sqares) ■
- A railroad (tracks) switch management.

More details will be on PDC dashboard online!

PDC
ONE STOP SOLUTION



PDC

CONCLOUSION & RECOMMENDATION

PDC
ONE STOP SOLUTION

- After observing data, most congestions were found between morning and evening time.
- Better targeting congestions by the map between morning and evening time for faster operations support to these congestions.
- MTA should fetch for buses from the map to help reducing congestions on time that has been detected.
- MTA should use map for tracks switch management and act upon morning and evening congestion (recommended).

DEMO


PDC
Thank you for listening!
PSYCHIC