Analyzing Traffic Congestion in Metro Manila Using Non-Traditional Data : A Use Case of the Waze Dataset

The Context

- Traffic congestion is a long-standing issue in Metro Manila, with recent estimates of hours spent in jams being the eighth (8th) worst among all cities in the world.
- An average Filipino is estimated to spend around 4 days in traffic jams annually (Ang, 2022).
- Opportunity costs related to traffic jams amount to about P3.5 billion per day, according to the Japan International Cooperation Agency (JICA).
- If status quo is preserved, this amount is expected to triple by 2030 (De Vera, 2018).

The Context

- Traffic congestion also intensifies air pollution due to increased vehicle emissions. This degradation in ambient air quality has been linked to excess morbidity and mortality for drivers, commuters and individuals, especially for those living near major thoroughfares (Zhang and Batterman, 2013).
- The World Bank (2016) provided an estimate of around 5,000 premature deaths (12 percent of all deaths) per year in Metro Manila primarily due to respiratory and cardiovascular diseases arising from exposure to harmful air pollutants.
- Traffic also threatens the commuters' overall well-being as it causes stress, fatigue, and hindrances to pursuing wellness activities.

Leveraging nontraditional data and technology

- Activities of households and businesses have increasingly involved more digital technologies, which gave rise to big data
- Developments in computing and machine learning paved the way for more complex information to be processed and used for creating valuable, evidence-based insights regarding real-world problems
- Mobile applications as a goldmine for traffic and commuting information – one of these being Waze – that is relatively accurate, high in frequency, and more readily available

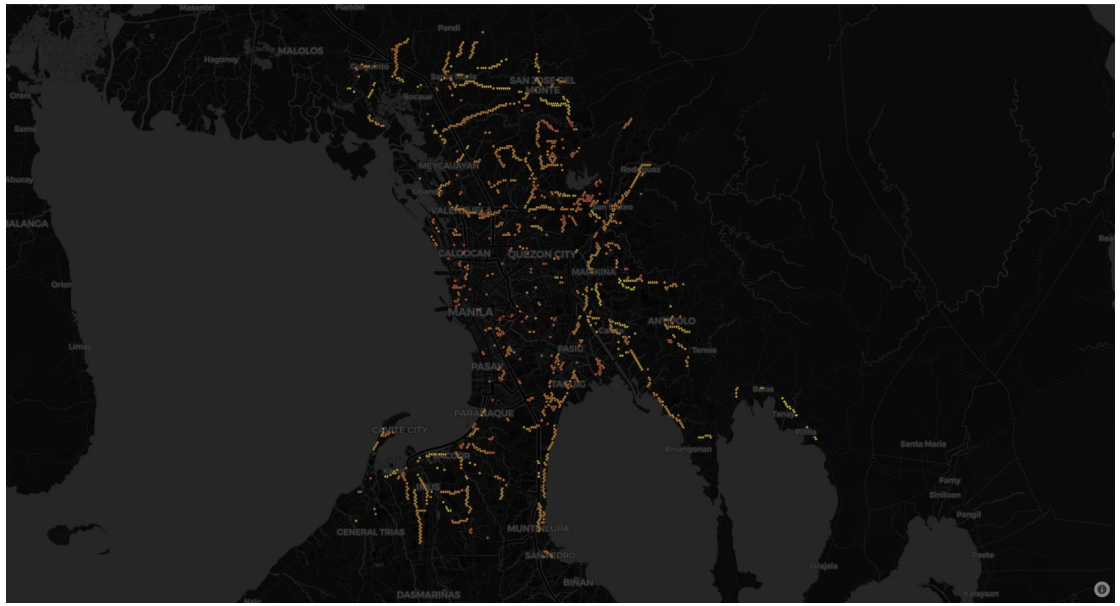
Waze Traffic Congestion Data

- Accessed through the World Bank Development Data Partnership
- 35,729,550 reports from April 2019 – April 2022
- Covers 9,297 streets across 61 cities

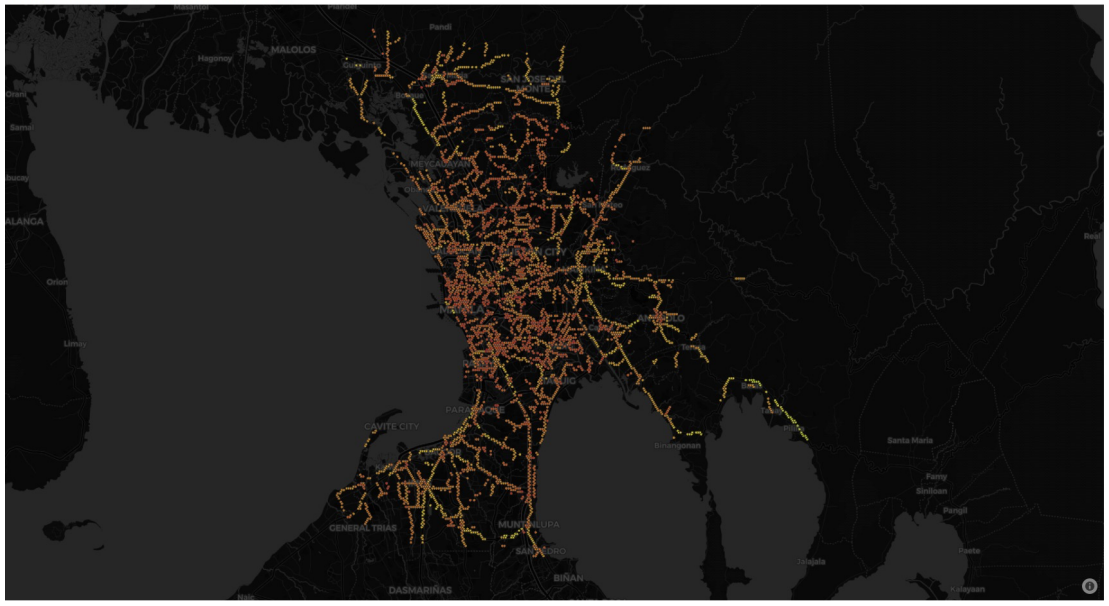
Attribute	Description
id	Unique ID of a report
ts	Date and time (UTC) of the report's occurrence
city	City in which report occurred
street	Street in which report occurred
startNode	Area in which a user starts their navigation
endNode	Destination chosen by a user
speedKMH	Speed in kilometers per hour of a report
length	Length of jam in meters
delay	Difference between jam and free flow speed in seconds
level	Traffic congestion level (categorical: 0 for free flow and 5 for blocked)
geo	Line string of latitude-longitude coordinates of a report

Waze Traffic Congestion Data

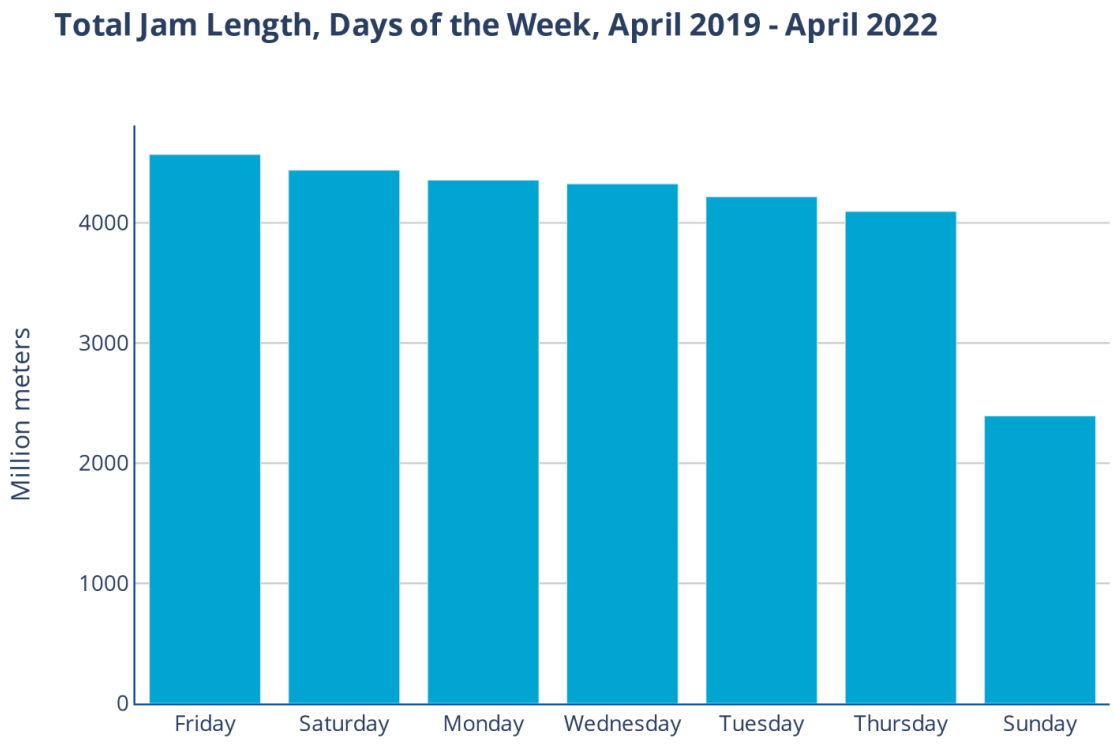
Average Speed (KM/H), *30 August 2021, 6-7PM*



Average Speed (KM/H), *01 December 2021, 6-7PM*

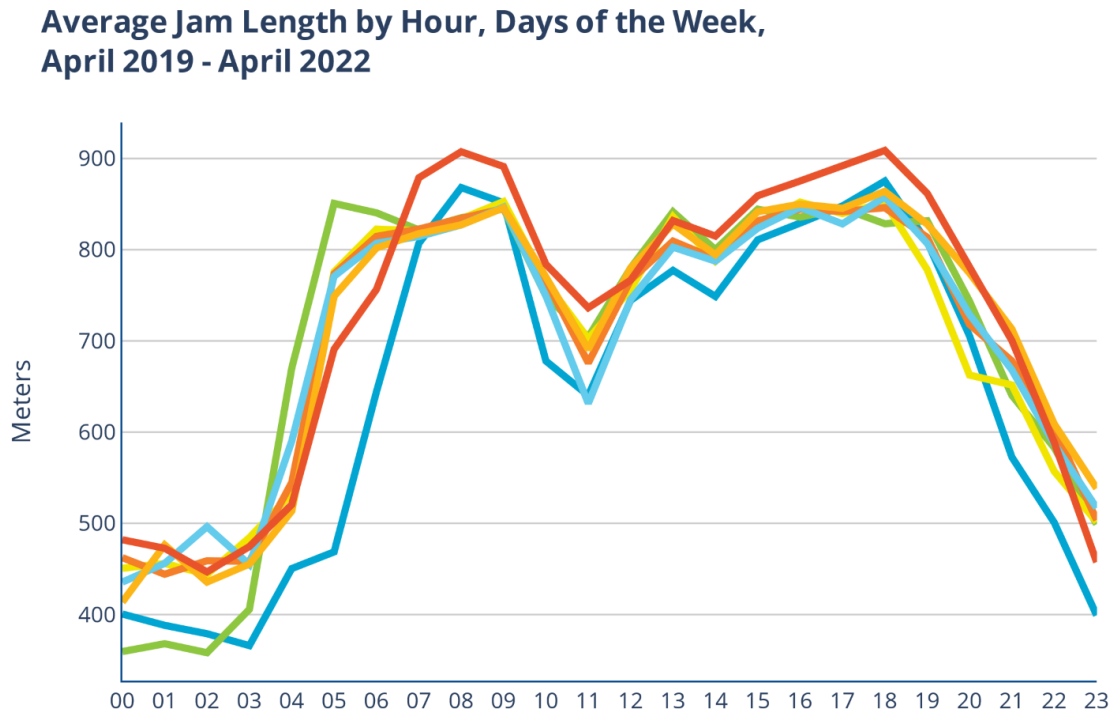


Waze Data validation: Exploratory Data Analysis



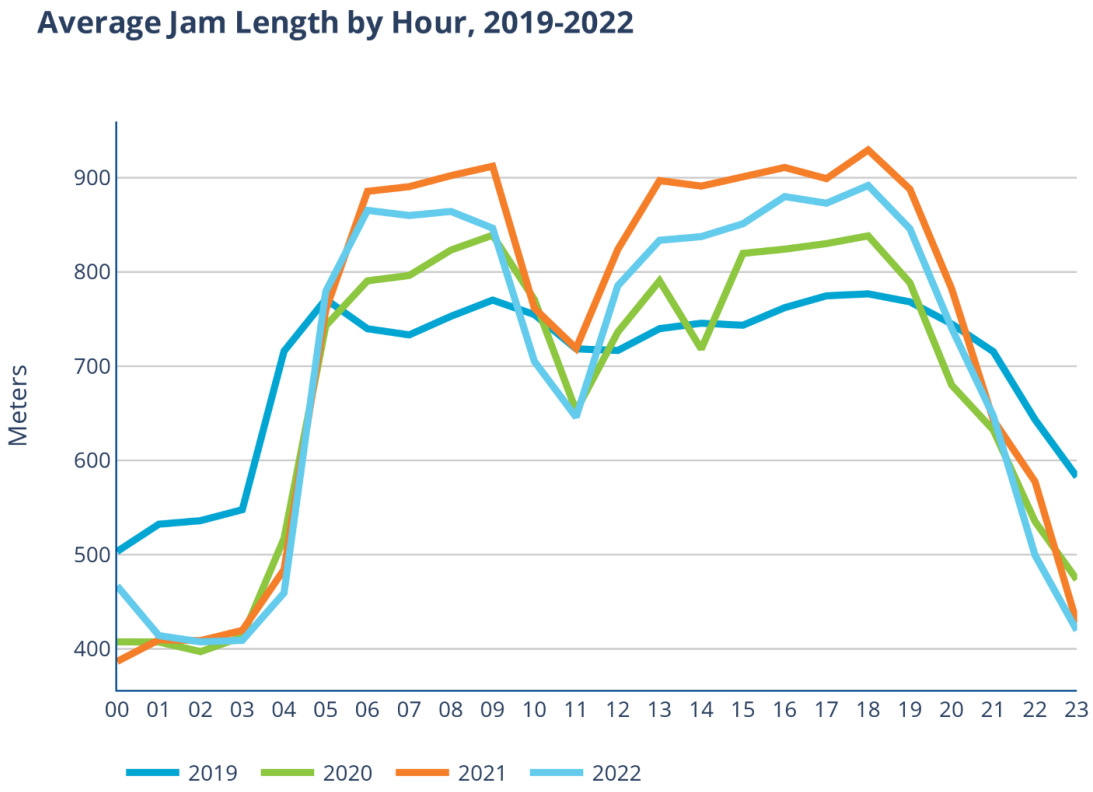
- From April 2019 to April 2022, Friday was the most congested day in terms of total jam length, followed by Saturday and Monday
- On the other hand, Sunday, which is generally recognized in the country as a time for rest and family, was the least congested day

Waze Data validation: Exploratory Data Analysis



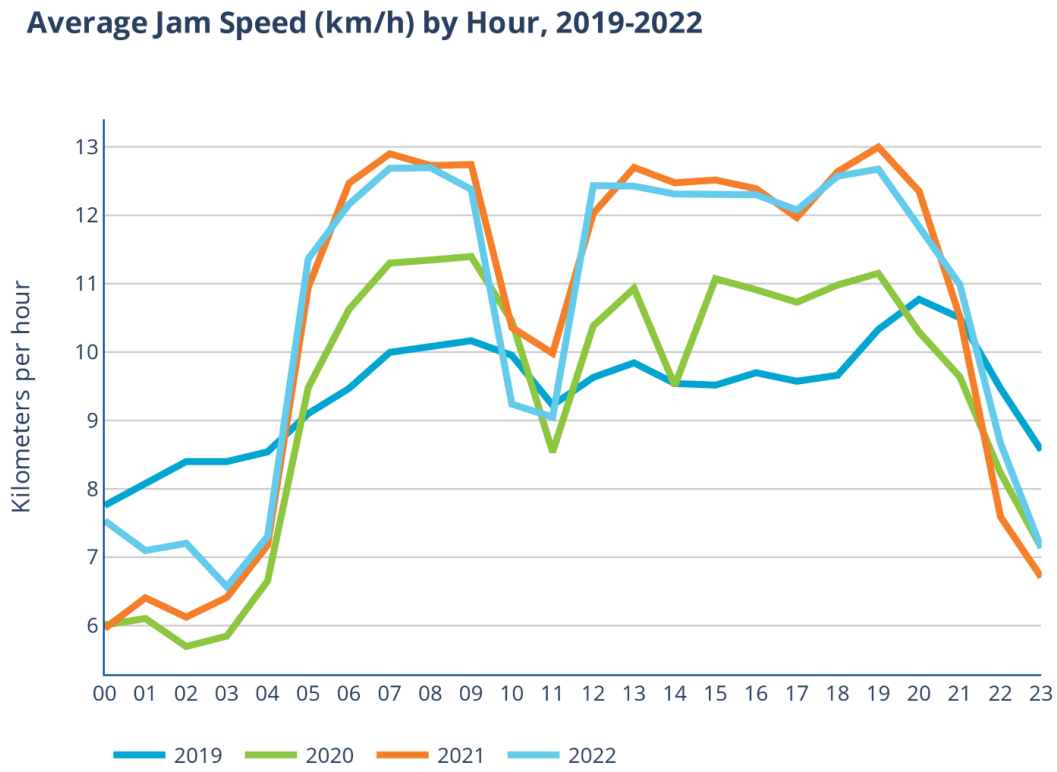
- There are two peak hours for all days of the week – one in the morning and one at night; In general, these peak hours are between 6 to 9 AM as well as 6PM
- Saturday recorded the highest average jam lengths while Sunday also had high peaks

Waze Data validation: Exploratory Data Analysis



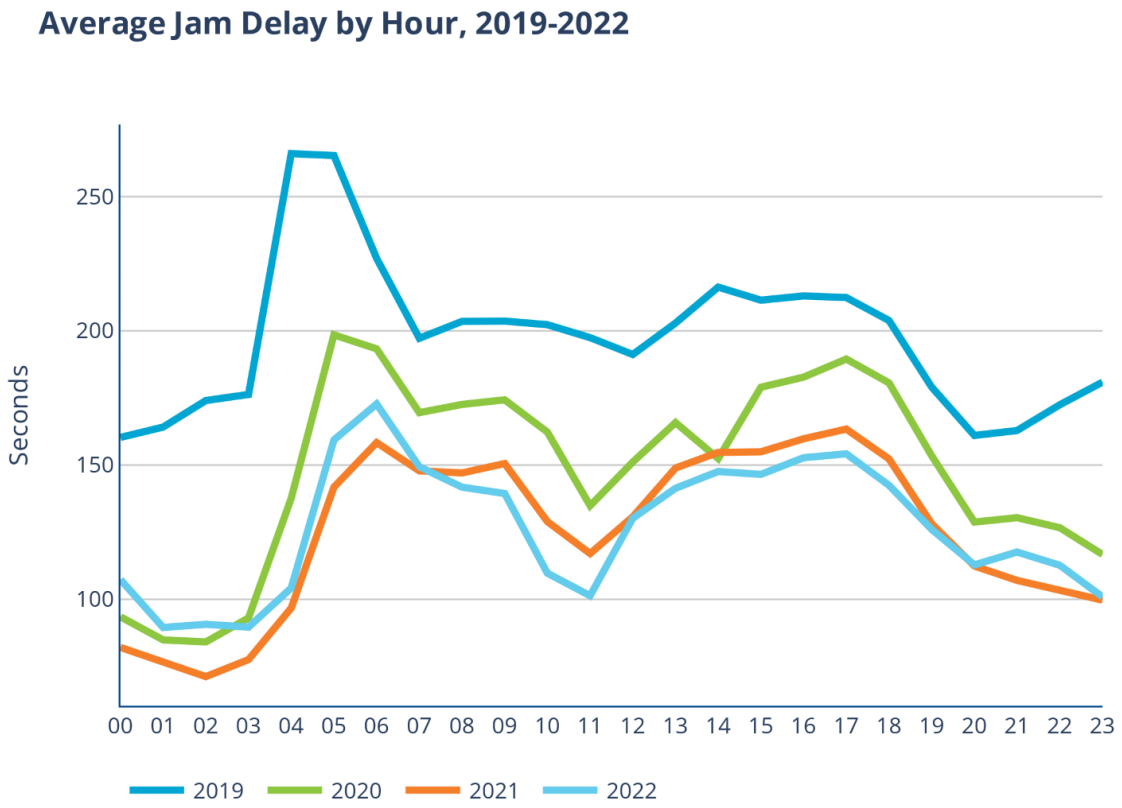
- Similar movement of trends in hourly jam length for 2020, 2021, and 2022
- Average jam speeds and lengths were generally less in 2020 compared to 2021 and 2022

Waze Data validation: Exploratory Data Analysis



- Similar movement of trends in hourly jam length for 2020, 2021, and 2022
- Average jam speeds and lengths were generally less in 2020 compared to 2021 and 2022

Waze Data validation: Exploratory Data Analysis



- Similar movement of trends in hourly jam length for 2020, 2021, and 2022
- Average jam delays (diff from free flow speed) were higher in 2020 compared to 2021 and 2022

Potential Applications

- The Waze dataset can be used as an input for assessing the viability and/or effectiveness of innovations related to traffic such as:
  - Changing the nature of work (WFH, hybrid)
  - Pilot restrictions such as high-occupancy vehicle lanes + route and street traffic flow modifications implemented by LGUs
  - Imposition of modified number coding schemes by the MMDA
  - Measures incentivizing carpooling / use of public transportation by the LTFRB
  - Traffic incentives (e.g., congestion charge) with revenues earmarked for improving public transportation
- It can also be used to inquire into people's behavior towards commuting (e.g., is distance from public transport stations related to a higher propensity to drive a private vehicle?)
- Event planning

Moving Forward

- From a policy perspective, traffic congestion in and of itself is not the main issue
- Economic, health, and environmental costs associated with it create concerns for the government and society in general
- Key role of traffic jams in forming these wide-ranging problems > spatial and temporal analysis of traffic jams a valuable use case for studying development issues
- Merging of traffic congestion data with datasets on health, environment, weather and climate, economy needed for more nuanced analysis, sensible predictions, and impact measurement
- If found useful, regular updates from DDP can be requested to keep track of more current trends and emerging patterns on the subject matter