Global Initiative of Academic Networks (GIAN)

BRINGING SYNERGY ACROSS DIFFERENT TRANSIT MODES IN INDIA BY ADDRESSING CHALLENGES FOR SUSTAINABLE TRANSPORT MODES

JUNE 23 - 27, WARANGAL, INDIA

MODULE 2

Instructors

Naveen Eluru, University of Central Florida

Raghuram Kadali, NIT, Warangal

CSRK Prasad, NIT, Warangal

COURSE MODULES

Introduction	• Public Transportation – An Introduction
Public transport data	• Background on data components useful for public transportation system analysis, their compilation and consistency analysis
Modeling approaches for public transit analysis	• Introduce traditional frameworks for public transit analysis – linear regression, discrete choice models (such as multinomial logit, ordered logit, and count models)
Emerging models for public transit data analysis	• Flexible discrete choice models (NL, ML, discrete continuous models) and machine learning models (KNN, RF, SVM, Decision Tress and Gradient Boost)
Integrating emerging modes with public transit	• Bringing it all together to leverage emerging modes and data analytics to improve public transportation across India

LECTURE OUTLINE

- Public Transit Data
- Passenger Data
  - Examples
- What Dimensions To Analyze
- How to Analyze

PUBLIC TRANSIT DATA COMPONENTS

Public transit data includes

- Passenger data
- Vehicle data
  - Fleet data, engine fractions, and occupancy
- Travel Information data
  - Vehicle arrival times, current location, disruptions
- Financial data
  - Revenue, expense, funding
- Workforce data
  - Employees, salary and union

PASSENGER DATA

Ridership data

- At the stop level, route level, system level by various time periods (time of day, weekday, weekend, holidays)

Demographic and socioeconomic data

- In the catchment area, for riders (and non-riders)

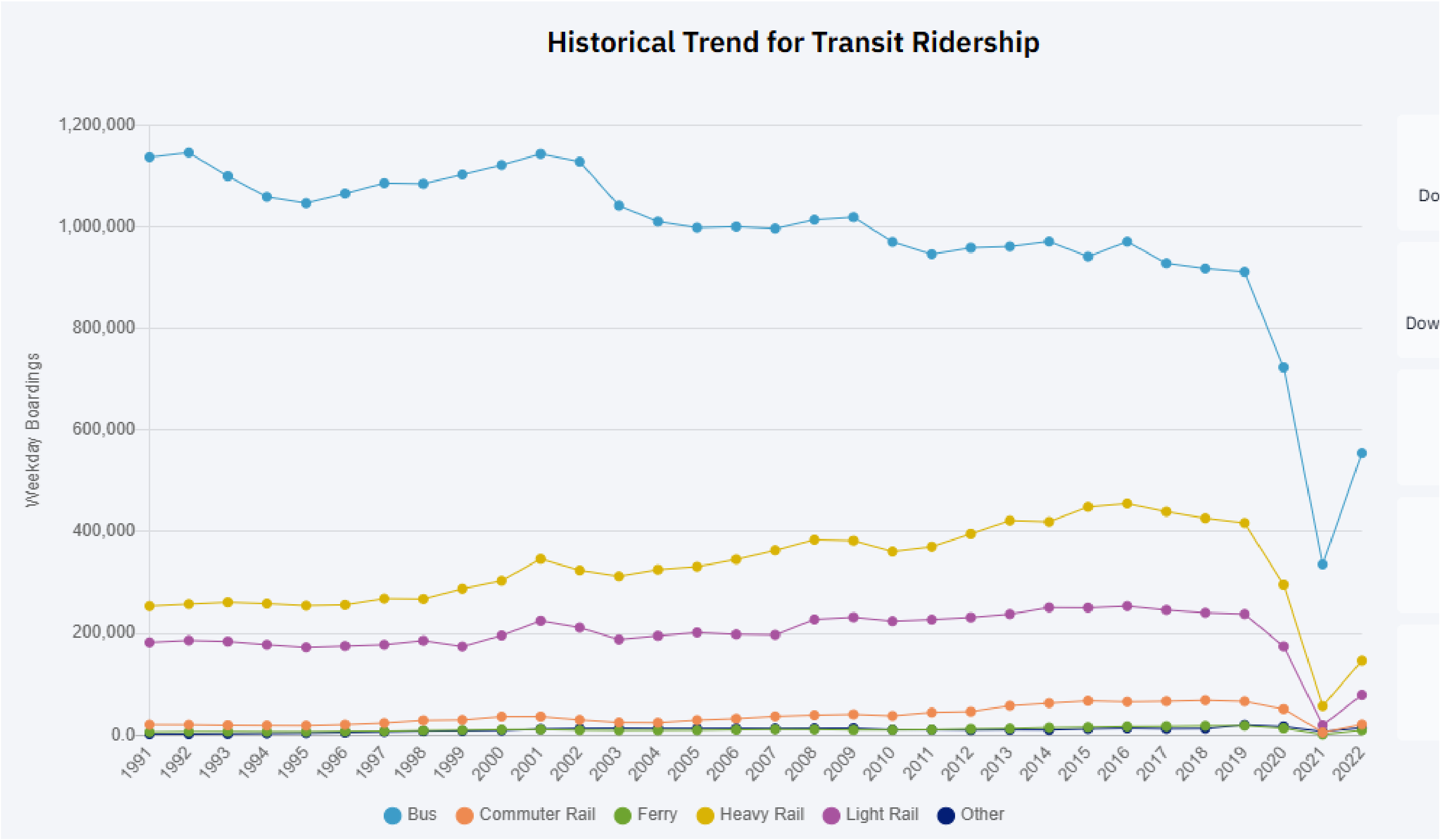
Travel patterns

- High fidelity or aggregate movement patterns

Individual data

- At the individual level (number of trips in a week, trip purposes)

SAN FRANCISCO - RIDERSHIP



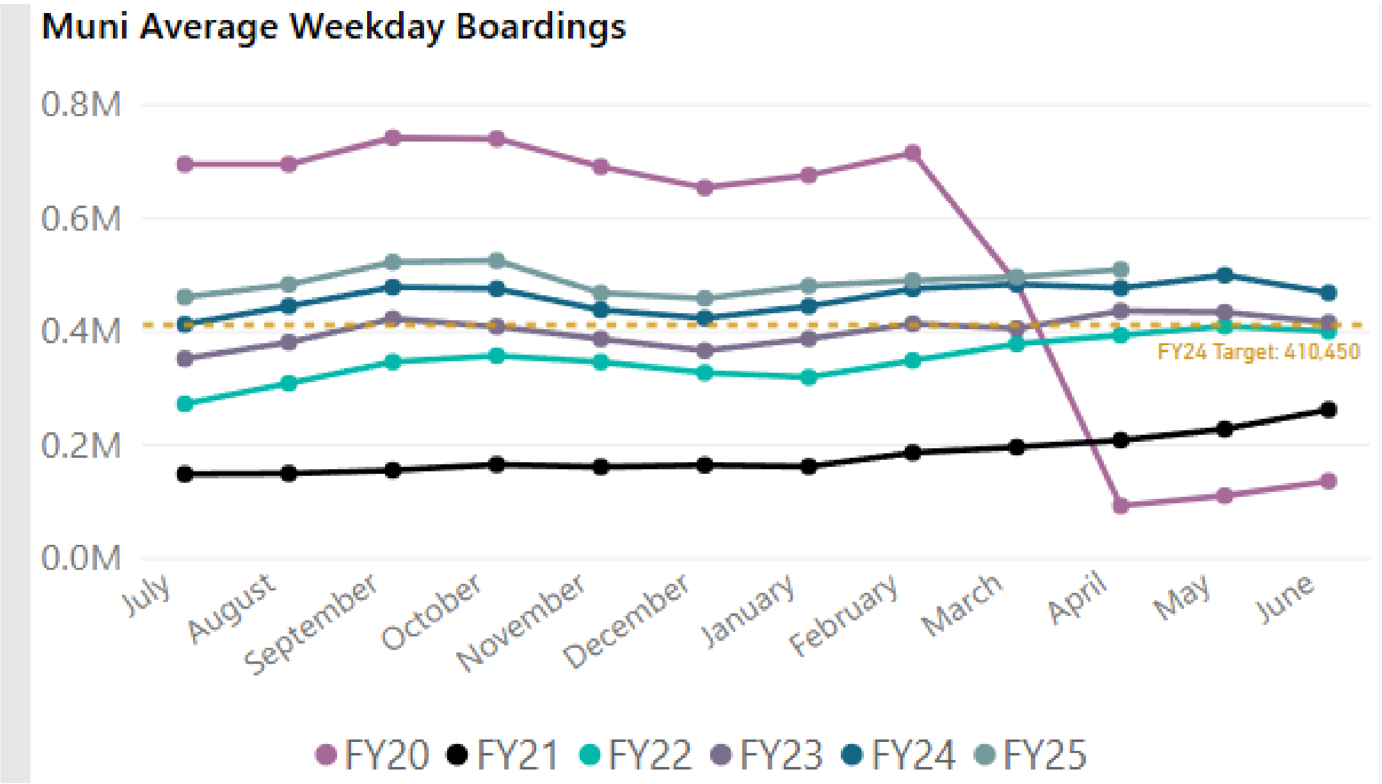
Source: https://vitalsigns.mtc.ca.gov/indicators/transit-ridership

SAN FRANCISCO - RIDERSHIP

- Muni

Chart description

- Y-axis: Average number of riders, in millions
- X-axis: Months in the fiscal year



Source: https://www.sf.gov/data--muni-ridership

SAN FRANCISCO - RIDERSHIP

- SamTrans

Total Passengers	Fiscal Year 2015	Fiscal Year 2016	Fiscal Year 2017	Fiscal Year 2018	Fiscal Year 2019	Fiscal Year 2020
Bus	13,158,700	12,802,550	11,825,380	11,134,270	10,670,850	8,788,180
Paratransit	329,040	351,200	361,380	354,680	337,420	256,730

Total Passengers	Fiscal Year 2021	Fiscal Year 2022	Fiscal Year 2023
Bus	4,503,358	6,956,853	8,528,698
Paratransit	121,394	171,130	202,425

Source: https://www.samtrans.com/about-samtrans/bus-operations/ridership

SAN FRANCISCO - RIDERSHIP

- BART

The average weekday BART station exits.

	2019	2020	2021	2022	2023	2024	2025
January	395,860	388,910	43,012	85,463	134,140	151,854	162,938
February	407,337	404,552	47,665	105,374	151,390	162,186	171,856
March	409,515	166,574	51,596	124,094	151,150	162,459	174,538
April	414,397	25,136	57,886	132,181	159,696	163,267	181,466
May	412,165	29,878	64,934	135,824	159,918	168,356	170,293
June	413,521	40,979	75,963	140,564	158,361	164,743	
July	401,465	45,633	85,291	133,858	154,825	159,220	
August	410,854	46,020	92,402	144,008	166,637	165,764	
September	426,755	48,838	105,997	161,902	172,051	184,248	
October	420,277	53,255	109,781	159,099	171,277	180,834	
November	411,183	52,198	112,282	150,242	165,802	166,035	
December	376,551	45,893	102,993	130,283	144,070	156,466	

Source: https://mtc.ca.gov/tools-resources/data-tools/monthly-transportation-statistics

SAN FRANCISCO – DEMOGRAPHIC DATA

Muni

- 91% of respondents live in the Bay Area, 9% said they were visiting.
- The average household size was 2.7 people, of which about 2 are employed.
- Most respondents said they spoke English (74%) or Spanish (15%) at home, but Chinese (6%) and 34 other languages were also mentioned.
- About 7% of respondents say they have a disability that limits their ability to travel.
- The average age of respondents is 39.9 years.
- The average household income of respondents is $91,797

Source: https://mtcdrive.app.box.com/s/9cq49dnbc8i7e257mgqqgy36v5tl2kpv6/folder/285362884103

SAN FRANCISCO – DEMOGRAPHIC DATA

SamTrans

- 93% of respondents live in the Bay Area, 7% said they were visiting.
- The average household size was 3.3 people, of which about 2 are employed.
- Most respondents said they spoke English (56%) or Spanish (32%) at home, but Chinese (4%), Tagalog/Filipino (3%), Burmese (1%), Russian (1%), and 16 other languages were also mentioned.
- About 9% of respondents say they have a disability that limits their ability to travel.
- The average age of respondents is 37.9 years.
- The average household income of respondents is $53,401

Source:
https://mtcdrive.app.box.com/s/9cq49dnbc8i7e257mgqgy36v5tl2kpv6/folder/285362884103

SAN FRANCISCO – DEMOGRAPHIC DATA

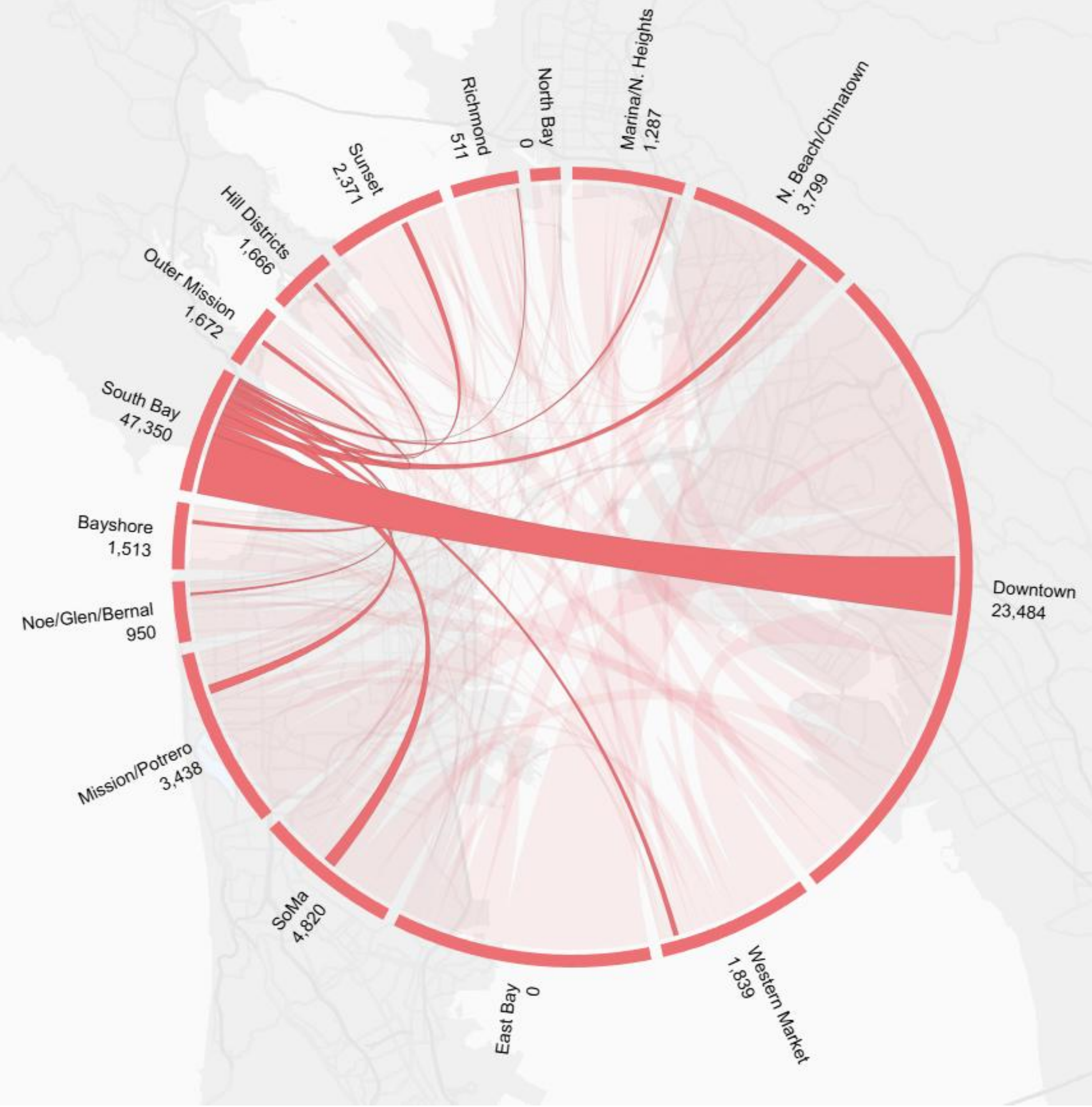
BART

- 92% of respondents live in the Bay Area, 8% said they were visiting.
- The average household size was 2.9 people, of which about 2 are employed.
- Most respondents said they spoke English (82%) or Spanish (11%) at home, but Chinese (3%) and 33 other languages were also mentioned.
- About 7% of respondents say they have a disability that limits their ability to travel.
- The average age of respondents is 39.5 years.
- The average household income of respondents is $105,458.

Source:
https://mtcdrive.app.box.com/s/9cq49dnbc8i7e257mgqqgy36v5tl2kpv6/folder/285362884103

SAN FRANCISCO – TRAVEL PATTERNS

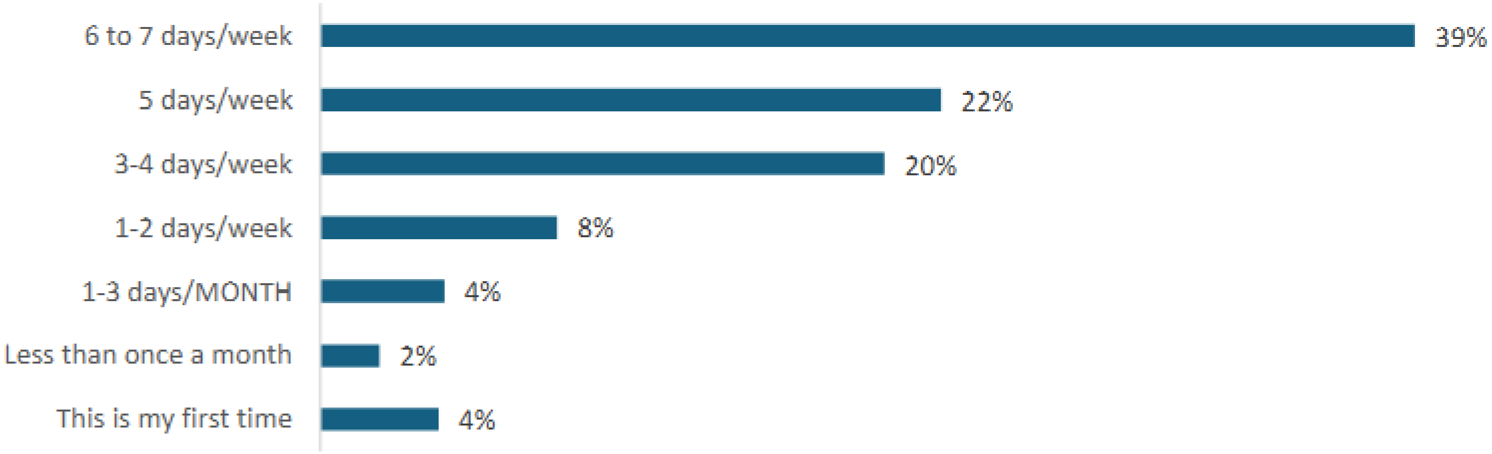
https://connectsf-trippatterns.sfcta.org/



GIAN: Bring

SAN FRANCISCO – INDIVIDUAL DATA

- Muni



Changes to Increase Transit

In response to the question, "What changes would get you to use transit more?," respondents gave the following answers most frequently (multiple responses accepted):

- More frequent service – 47%
- More reliable service – 27%
- Cleaner vehicles/stations – 26%
- Lower fares – 25%

Source: https://mtcdrive.app.box.com/s/9cq49dnbc8i7e257mgqgy36v5tl2kpv6/folder/285362884103

WHAT DIMENSIONS TO ANALYZE

- https://www.mentimeter.com/app/presentation/al7tuwegvhskecycp85osrk5ba7nud1m/edit?source=share-invite-modal

WHAT DIMENSIONS TO ANALYZE

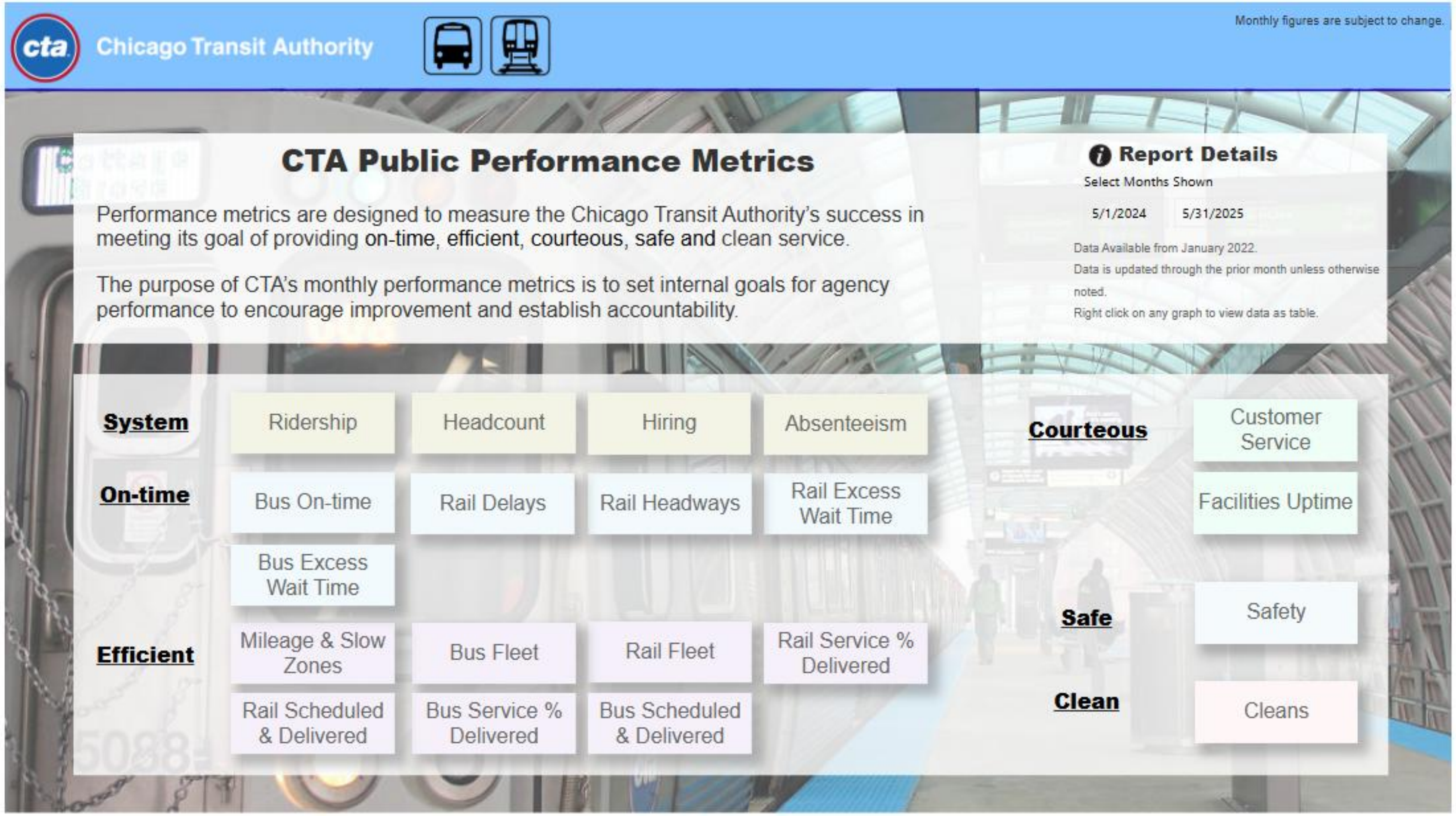
Ridership

- Boarding and alighting (stop level, route level, system level)
- Passenger miles or hours

Accessibility

- What does transit allow individuals to reach – jobs, grocery, health care and leisure
- How is transit accessed [mode, station and]

EXAMPLE



https://www.transitchicago.com/performance/

RIDERSHIP

In urban regions all stops are not the same

- For example, across the 8000 stops in Montreal the ridership (boardings + alightings) varies significantly from 0 to 8000 [Chakour and Eluru, 2016]
- We can't develop one model for all stops!

We cluster them

- Stops with daily ridership of less than 50 are characterized as low stops
- Stops with daily ridership between 50 and 250 are characterized as medium stops
- Stops with daily ridership more than 250 are classified as high stops.
- The largest sample of stops in the low category (3574), and the lowest sample of stops in the high category (1813)

WHAT DIMENSIONS TO ANALYZE

System Operation

- Travel times by route (On-time arrival, delays)
- Waiting and walking times by route
- Customer satisfaction based on customer feedback

Societal impact

- Air quality impact – reducing private vehicle emissions (lowering car /motorbike ownership)
- Community impact – improving communities though connections, positive improvements to property values

HOW TO ANALYZE?

Ridership

- Typically modeled as boardings or alightings
- Continuous in nature
- Linear regression might be appropriate

Travel time

- Continuous in nature

Impact of transit on community

- Property values – traditionally modeled as a linear regression model referred to hedonic price models

HOW TO ANALYZE?

Mode choice, station choice or route choice

- Categorical variables such as Car, Bus, and Walk
- Employ discrete choice models such as MNL, NL and Mixed MNL
- Public transit riders can select bus stop or metro station

Customer satisfaction

- Survey data on customer satisfaction compiled as Likert scale variables [Very Good, Good, Fair, Poor, Very Poor]
- Suitable for ordered response models

SUPPLEMENTARY DATA

Transit system operational attributes

- Average headway for time period, number of lines passing through the stop, night bus passes through stop

Public transit accessibility indices

- Number of bus/metro/train stops around each stop, length of bus/metro/train lines, length of exclusive bus lanes

Transportation infrastructure attributes

- Road length by functional classification, bike lane lengths, distance to central business district),

Built environment attributes

- Number of parks and their areas, residential area, number of commercial enterprises and their area, government and institutional area, resource and industrial area, employment density, walkscore

SAMPLE LIST (CHAKOUR AND ELURU, 2016)

Independent variables
– Stop level variables
Headway AM
Headway PM
Headway OPD
Headway OPN
Number of lines passing through stop
Night bus passes through stop

– Transit around the stop[a]
Number of bus stops in a
200 m buffer
Number of metro stops in a
200 m buffer
Number of train stations in a
200 m buffer
Bus line length in a
600 m buffer
Metro line length in the TAZ
Train line length in the TAZ
Reserved bus lane length in a
200 m buffer
– Infrastructure around the stop
Major roads length in a
400 m buffer
Highway length in a
800 m buffer

– Land use around the stop
Park area in a
200 m buffer
600 m buffer
Number of parks in a
200 m buffer
600 m buffer
Number of commercial enterprises in a
200 m buffer
600 m buffer
800 m buffer
Commercial area in the TAZ
Governmental and institutional area in the TAZ
Residential area in the TAZ
Park and recreational area in the TAZ
Resources and industrial area in the TAZ

DATA PREPARATION PROCESS

HOW TO ANALYZE? COMPLEX SCENARIOS

The analysis needs to be customized by

- Mode or modes of interest
- Study region

Role of headway

- It is important to recognize that headway is determined based on expected demand
- As headway is closely correlated with demand, we need to develop methods that can address the endogeneity

New Infrastructure Additions

- How transit performance and usage changes with new infrastructure changes (bus line or Light rail)
- The new additions and associated connections can affect ridership and needs to be carefully considered

REFERENCES

- Chakour V., and N. Eluru (2016), "Examining the influence of stop level infrastructure and built environment on bus ridership in Montreal," (2016) Journal of Transport Geography Volume 51, February 2016, Pages 205-217
- Slides for the YouTube video - https://people.cecs.ucf.edu/neluru/wp-content/uploads/2022/04/DataPreparation_Webinar_1.pdf