

# The future of travel surveys

September 11<sup>th</sup> Memorial Program

Niloufar Nouri

City University of New York

September 15, 2021

# Outline

- Research objectives
- Available travel survey data
- Recent surveys by the MPOs
- Alternative travel data collection
- Big data challenges/strength
- Questions to ask from data providers
- Key findings

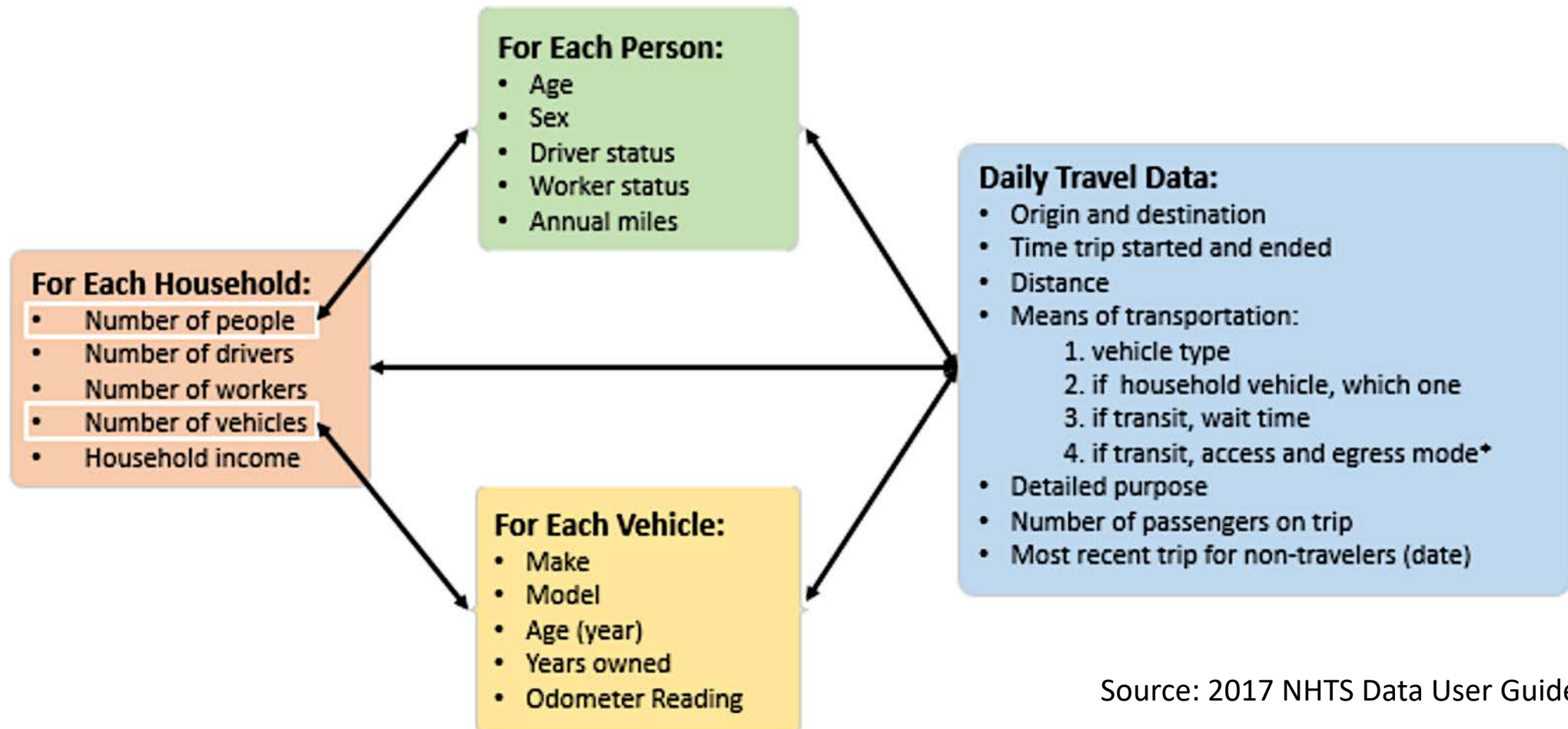


# Research Objectives

- Alternatives for the travel survey data collection
- Evaluating the strength and challenges of big data sources
- What needs to consider for choosing the alternative(s)

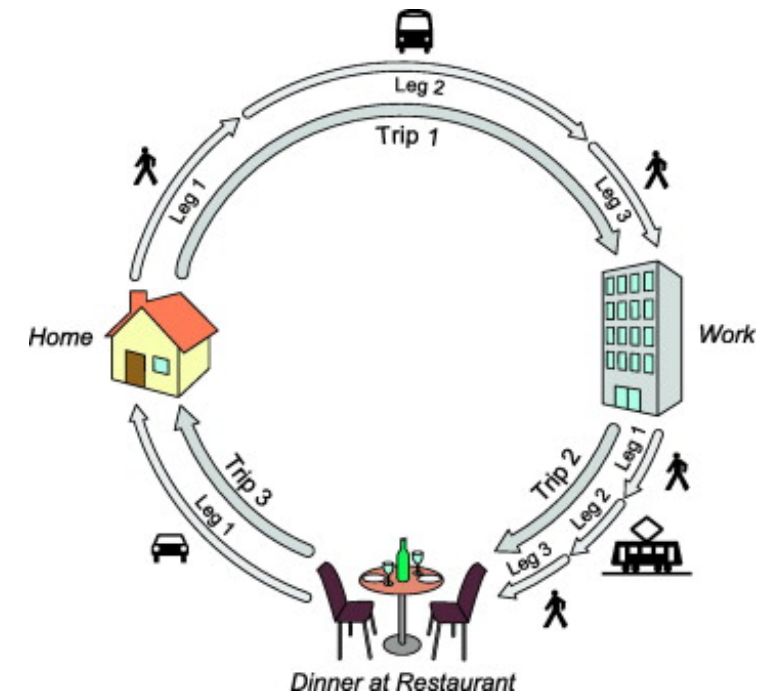
# Travel Survey Data

- Travel survey variables for the transportation purposes



# Available Datasets for Travel Surveys

- National Household Travel Survey
- American Community Survey (ACS)
  - Social, Economic, Housing, and Demographic Characteristics
- Census Transportation Planning Products (CTPP)
  - Based on ACS data
- NYMTC Regional Household Travel Survey



Source: Nitsche et al., Transportation Research, part C

## Surveys by MPOs

---

- More frequent surveys with smaller samples instead of traditional studies every 8 - 10 years.

---

- Combined data collection methods (telephone/mail/web/app)

---

- A tree-stage survey design and collecting demographic information during recruitment phase

---

- Using smartphone app in survey design

---

- Email/phone-call reminders and monetary incentives

# Alternative Data Collection: Passive Data

## Data Definition

- Data collected without explicit or noticeable interaction with a person

## Data Type 1

- Data collected through a device to increase response
- More accurate and cleaner data, lower response burden
- Trip data collected by Smartphone app

## Data Type 2

- Data collected for purpose unrelated to travel
- Location information provided by Smartphone apps

# Data Collection Methods

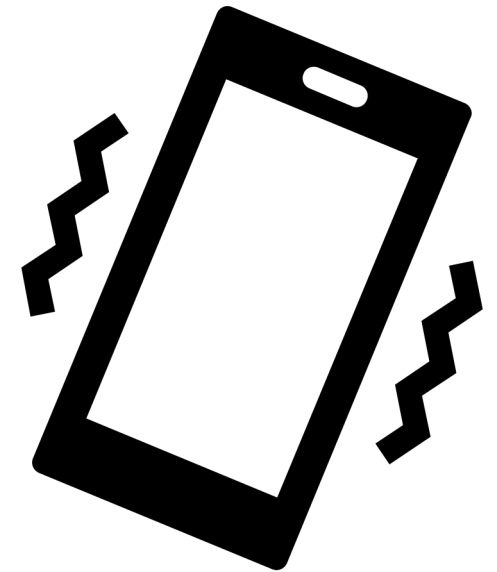
- Cellular Network-based data
  - Call data record (CDR)
- In-vehicle GPS data
- Location-based services
  - Geospatial and communication technologies utilize Internet to provide real-time geographic location
- Smartphone sensor-based data
  - Data from smartphone apps while checking social media such as Twitter, LinkedIn, Facebook, etc.





# Call Data Record (CDR)

- Details of an incoming or outgoing call, text message, or a connection to an app, web browser, or email
- Active: when users make calls, send messages, or use the Internet actively
- Passive: when users receive calls or texts and while apps accessing the phone
- Easily available since most service providers maintain these data for billing purposes
- The amount and frequency of CDR data depends on usage patterns, number of apps installed and the frequency of interaction with towers



# CDR Challenges



- Unreported trips
- Poor spatial precision
- Unable to observe traveler's characteristics
- Unable to provide travel mode and traveling party size
- Lack of socioeconomic data, make the market segmentation infeasible
- Determining an origin or destination is difficult in mixed land-use

# Challenges of Big Data Sources

Data	Strength	Challenges
<p>Cellular Tower: from cellular tower “triangulation” <b>(50-300 m spatial precision)</b></p>	<p><b>Large sample size</b> more than 25M devices Inferred home and work locations</p>	<ul style="list-style-type: none"> <li>• Poor spatial precision</li> <li>• Personal and commercial trips are not separated</li> <li>• Can not reliably infer modes of transportation</li> <li>• Poor coverage in rural areas</li> </ul>
<p>Navigation-GPS: cars and trucks <b>(3- 5m spatial precision)</b></p>	<p><b>Very frequent pings , precise location</b> Personal and commercial trips</p>	<ul style="list-style-type: none"> <li>• lower sample size</li> <li>• Can not infer non-vehicular modes</li> </ul>
<p>Location-Based Services: <b>(5- 25m spatial precision)</b></p>	<p><b>Large sample</b> Frequent ping rate</p>	<ul style="list-style-type: none"> <li>• Less suppliers</li> <li>• Sample size varies across data providers</li> <li>• Ability to infer trip purpose and trip modes on an aggregated basis</li> </ul>

# Important Questions



- The source of raw data in the platform?
- Filtration/cleaning processes are performed on raw data?
- Number and percent of devices captured in urban and rural areas?
- Trip breaking time criteria?
- What normalization techniques are used to correct data?
- The methodology for expanding sample data?
- What modes of travel data represented? How travel modes determined?
- How geographical/demographical biases are controlled?

# Findings and Conclusions

- Uncertainties around travel analytics using big data options
  - Demographic information
  - Travel mode and trip party
- Comparison between the results of previous surveys with the trip estimates from a data platform
- Using smartphone apps besides self-reporting methods
- Advantages and disadvantages of using a typical smartphone app instead of/combined with travel diary
- More frequent surveys with smaller sample size
- Running a pilot survey before the main survey



—

# Thank you!

---

