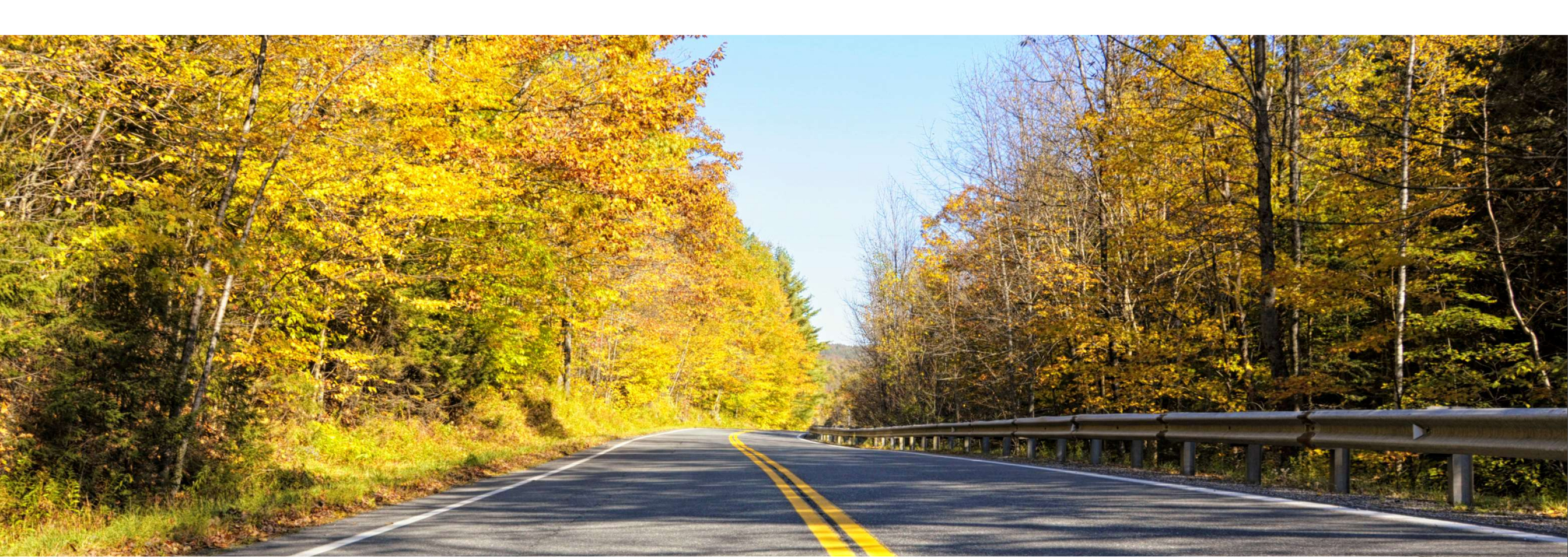




The Intersection of Smartphone Travel Survey and Passive Big Data

PRESENTED BY:

Michelle Lee
Abby Rosenson

2022 Fall NCMUG Meeting

NOVEMBER 7, 2022

Learning Objectives

TRAVEL SURVEYS:

- State of Smartphone GPS Travel Surveys
- **Case Study:** Triangle Travel Survey Program

PASSIVE BIG DATA:

- State of Passive Big Data
- **Case Study:** Treasure Valley Survey & Big Data Program

BRINGING THE TWO TOGETHER:

- Benefits of Each Data Type & Data Applications





Current State of Practice: Smartphone GPS Travel Surveys

Context: How are travel surveys typically conducted?

1

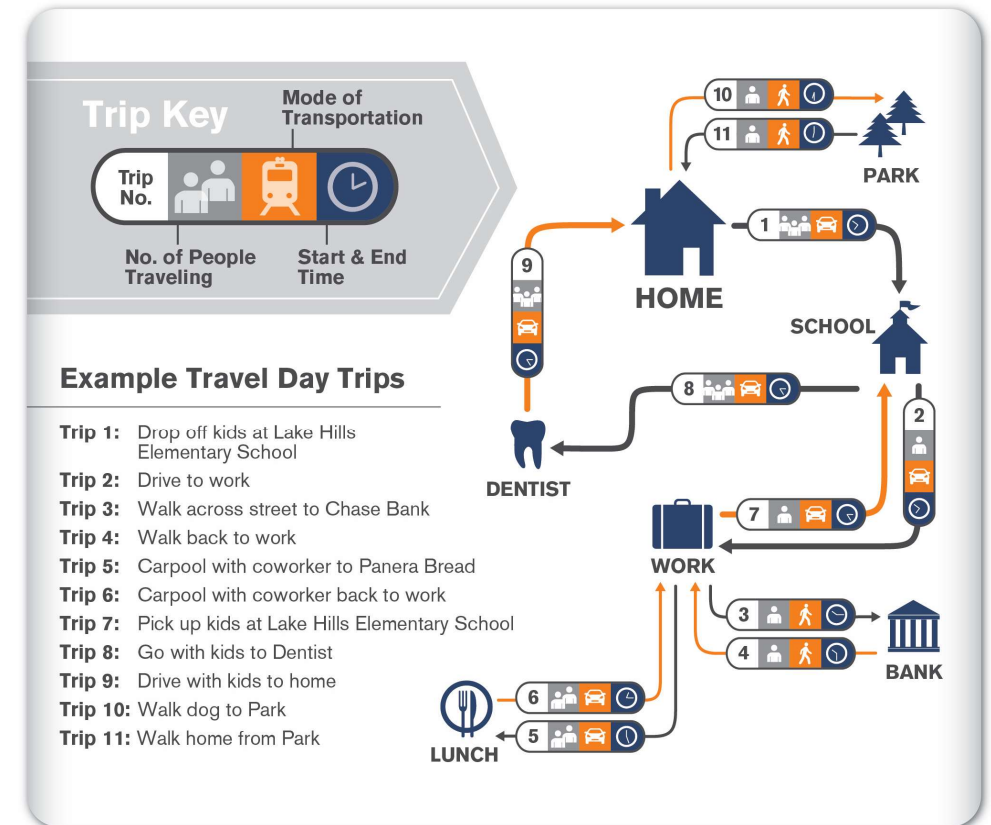
Residents are invited to take survey online, by phone, or via smartphone app

2

Participants complete a demographic survey + trip diary for at least 1 weekday

3

Data results in household-, person-, vehicle-, trip-, day-, and location-level data



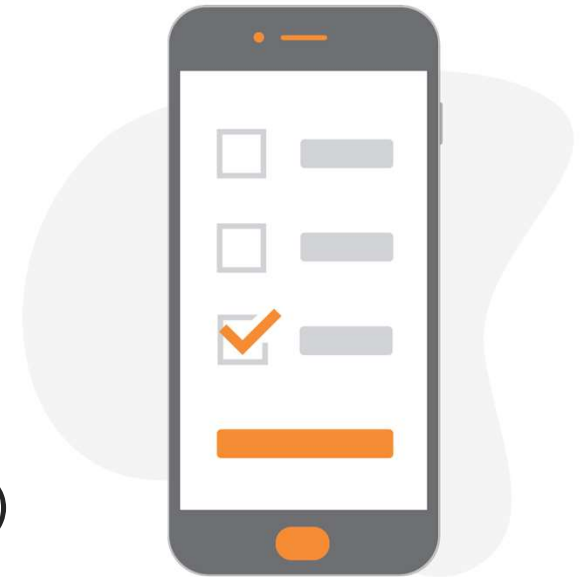
State of Practice: Primarily Smartphone-Based Data Collection

Smartphone app-based data collection provides:

- Detailed demographic information
- Detailed trip paths/traces, locations, and travel times

Multi-day travel periods provide:

- Greater/more **geographic coverage** in trip data
- Greater/more sufficient data for **rare travel modes** (e.g., TNCs)
- **Higher volume of data per person / household**



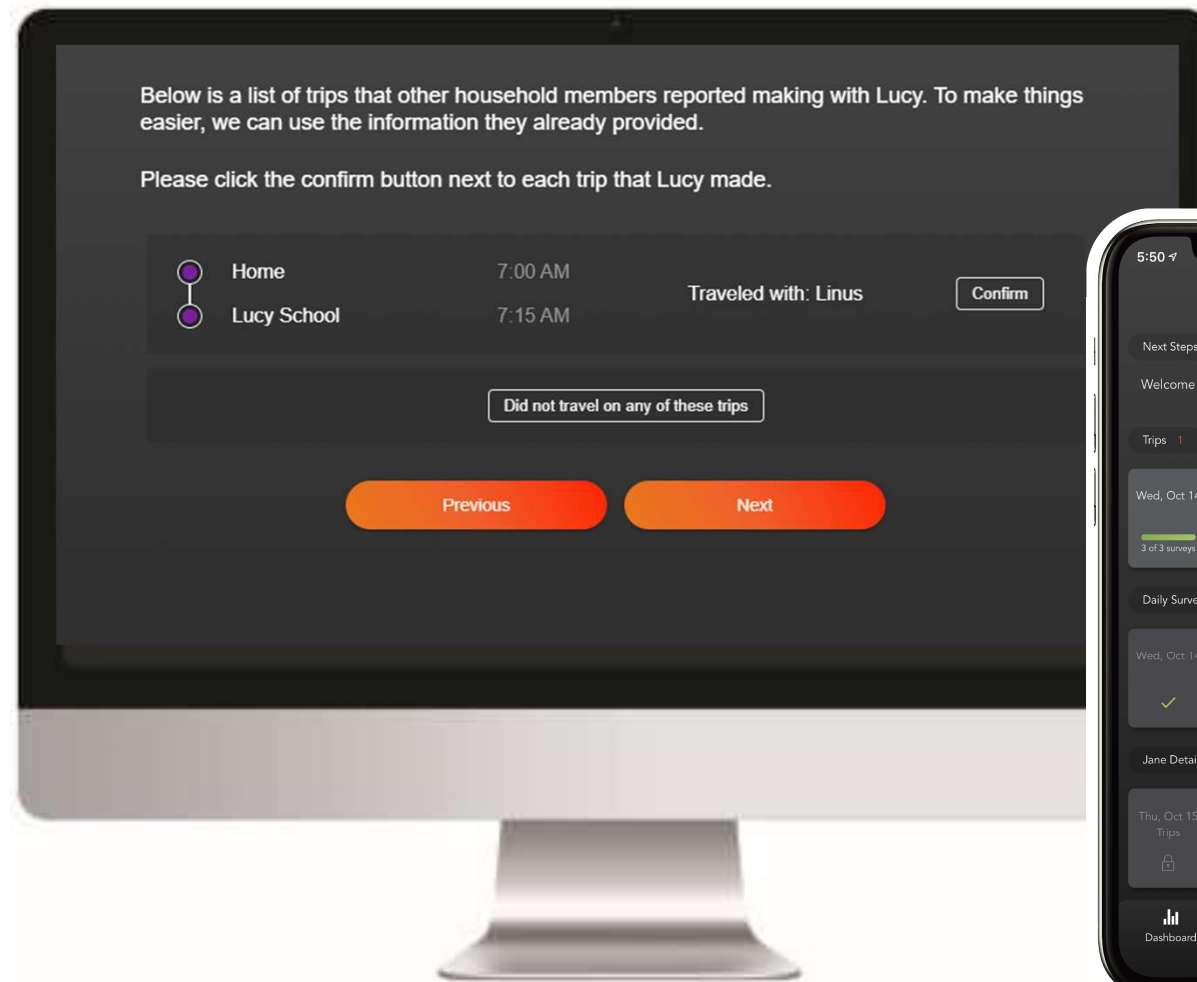
Smartphone Ownership Continues to Rise

- Recent studies show 50-80% smartphone participation rates.
- 80% or more of trip data is collected by smartphone in multi-day smartphone studies.

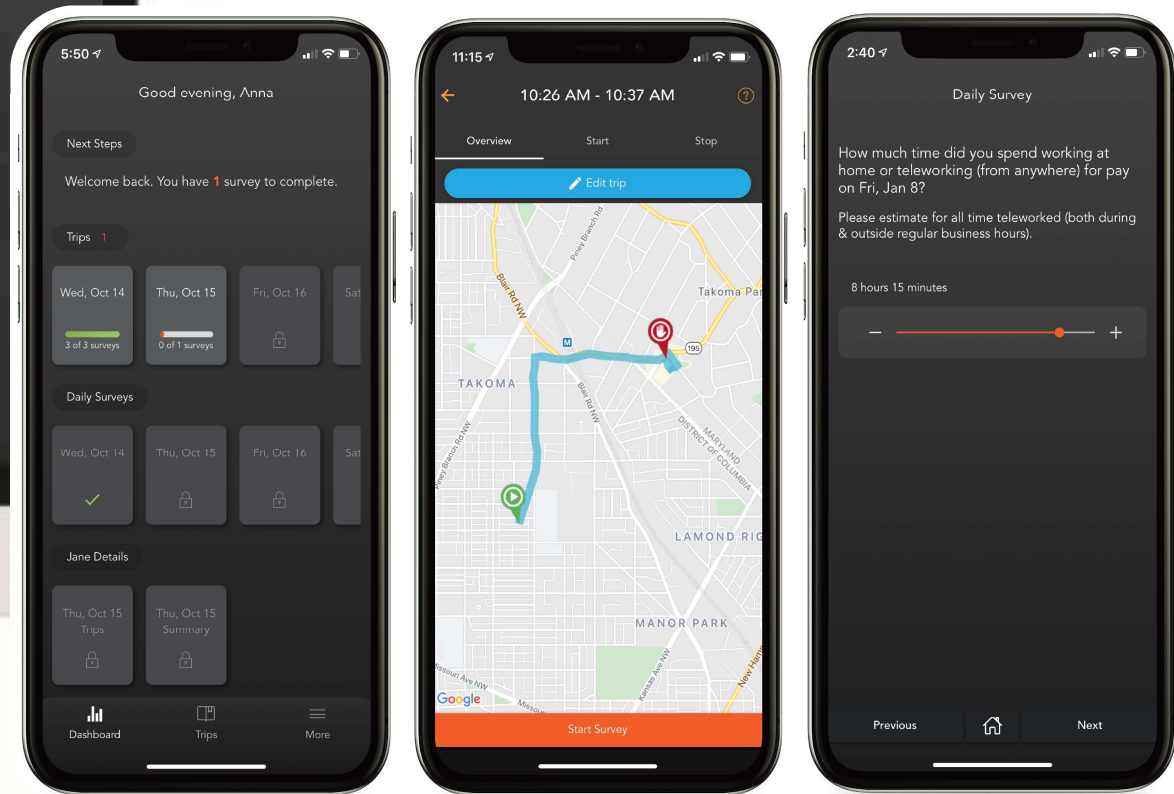
According to Pew Research, 85% of American adults owned smartphones in 2021.



Multi-Method Data Collection: Example Survey Instruments



◀ rMove™ and rMove for Web™ Instruments





Case Study: Triangle Travel Survey

Case Study: Triangle Travel Survey Program

2016 HTS: Intro to Recurrent HTS Program

- Primary web- and call-in survey to collect data from **4,194 households**.
- Conducted a separate three-day GPS smartphone app panel to evaluate methods and response. Collected data from 351 households.

2018 HTS: First Wave of Recurrent Program

- Web and call-in survey to collect data from **1,498 households**.
- Households provided data for one weekday.
- Trip rate correction used GPS data from 2016.



Case Study: Triangle Travel Survey Program

2020 (2021) HTS: Second Wave of Recurrent Program

- Web, call-in, and smartphone data collection from **1,120 households**.
- About 40% of the households participated via smartphone using RSG's smartphone app, rMove™.

2022 HTS: Third Wave of Recurrent Program

- In-field through mid-November, using similar methods as 2021.



Recurrent Data Collection Programs

(annually or every 2-3 years)

Benefits of Multi-Year Survey Programs:

- Aids trend analysis as travel evolves quickly
- Allows flexibility as survey methods evolve
- Helps smooth agency funding with smaller sample sizes collected more frequently (after initial “jump start” sample)

RSG has worked with the following agencies on this design:

- ***Puget Sound Regional Council***
(Seattle, WA region)
- ***Metropolitan Council***
(Twin Cities, MN region)
- ***NYC Department of Transportation***
- ***North Carolina State University ITRE***
(Raleigh-Durham, NC region)
- ***City of Calgary***
(Calgary, AB, Canada)





Current State of Practice: Passive Big Data

How is passive “Big Data” collected?

- Collected primarily from technologies (e.g., GPS smartphone apps) without overt interaction with consumer.
- Sequential location “pings” can be turned into trips. Habitual locations (e.g., home and work) can be inferred.



Transportation agencies must weigh the strengths and weaknesses of passively collected data

Advantages

- On a given day, it may be possible to observe 10 – 15% of the population and up to 50% over a month.
- It is straightforward to analyze data spanning multiple weeks, months, or years.
- Large samples of data can be collected passively at relatively low cost.
- Passive data can cover a much greater percent of the population and OD matrix cells than most surveys.
- Includes segments of the population that can be difficult to capture in surveys.

Limitations

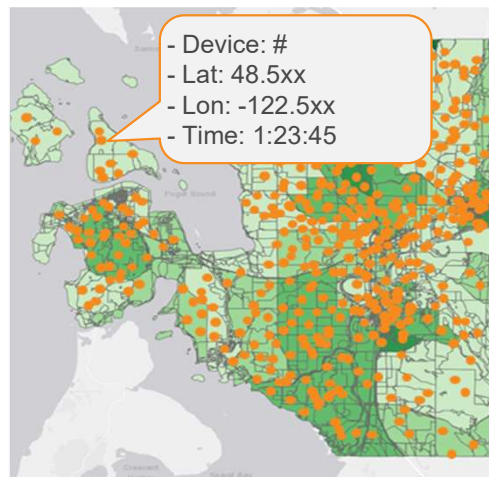
- There is no information on the device owner, their travel purpose, the activity they are engaged in, who may be traveling with them, or mode of travel. Some of these data can be imputed.
- Excludes travelers without mobile devices or vehicle navigation services.
- Short-distance trips or short-duration activities are under-represented.
- The protection of privacy involves some loss of information.
- Continuing concern about accuracy and data privacy.



4-step process for passive OD tables (rMerge)

1

PREPARE INPUT DATA



Millions to billions of individual device location points from commercial LBS data* are extracted, evaluated for basic metrics & cleaned

2

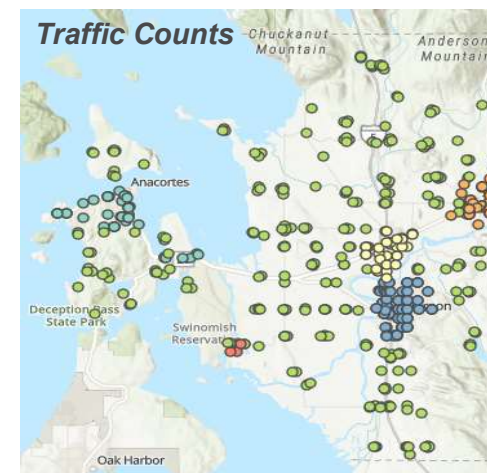
IDENTIFY TRIPS



Points are clustered to identify stop locations, locations are classified (home, work, other) and linked to create trips

3

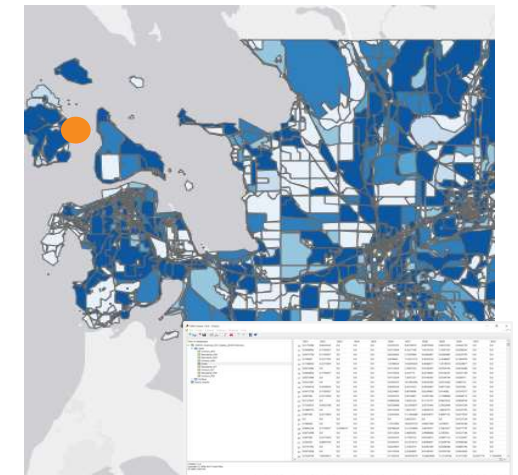
EXPAND TO REGION



Trips are expanded to region based on Census and traffic count data, surveys and other sources to provide representative O-D flows

4

AGGREGATE TO O-D



Trip data aggregated to OD matrices, with key dimensions (such as time period, visitor / resident) broken out



* Typically represents 10-15% of population per day, or 50%+ for one month of observations



Case Study: Bringing the Two Together

Treasure Valley Travel Behavior Survey

2021 Treasure Valley Travel Survey

Two primary goals:

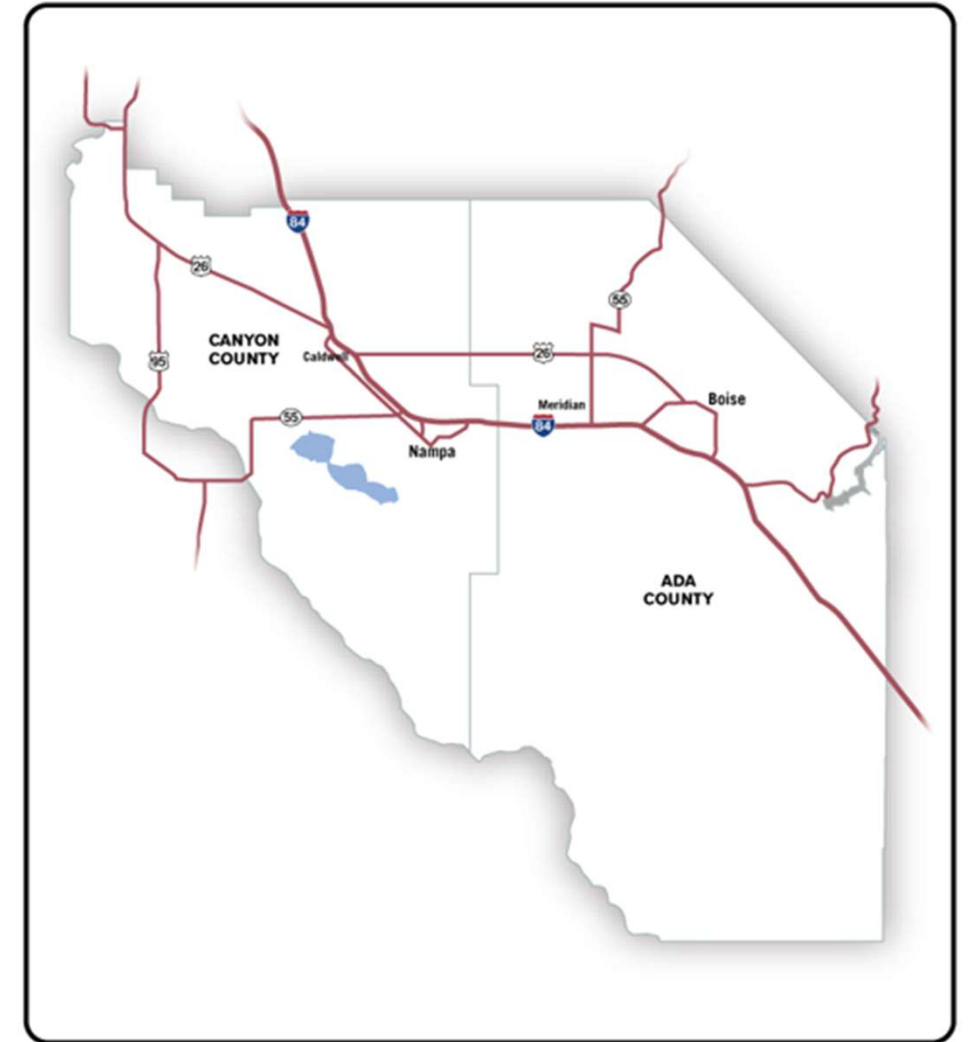
- Gain a better understanding of how local roads, highways, bike lanes, sidewalks, and public transportation are being used, and how the organization can improve upon them moving forward.
- To have recent changes like population growth, new transportation options, and increased teleworking frequency have affected the area.



2021 Treasure Valley Travel Survey

An integrated solution:

- Household travel survey
- Analysis of passively collected data
- Transit onboard survey (not discussed)



Household Travel Survey

SURVEY RECORDS



4,012
HOUSEHOLDS
Unweighted Records



9,066
PERSONS
Unweighted Records



7,753
VEHICLES
Unweighted Records



33,251
TRAVEL DAYS
Unweighted Records



133,511
TRIPS
Unweighted Records



1,804,914
LOCATIONS
Unweighted Records

- Survey fielded from **August 9, 2021, through November 16, 2021.**
- **Smartphone participants completed a 7-day travel diary.**
- **Online and call center participants completed a 1-day travel diary.**
- All children (under age 18) have a complete 1-day travel diary regardless of participation mode that was proxy reported by an adult member of the household.
- Same questionnaire was used for smartphone, online, and call center participants.
- Survey was available in English and Spanish.

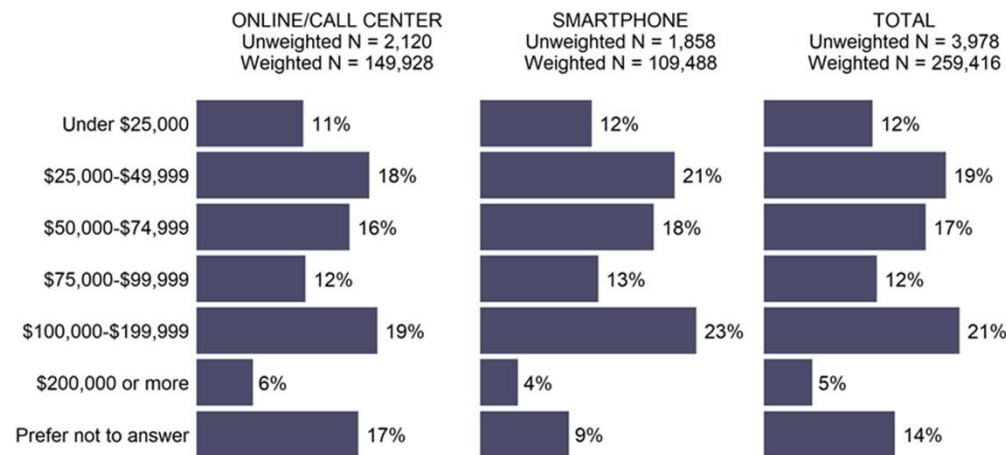


HTS tells us detailed demographic information

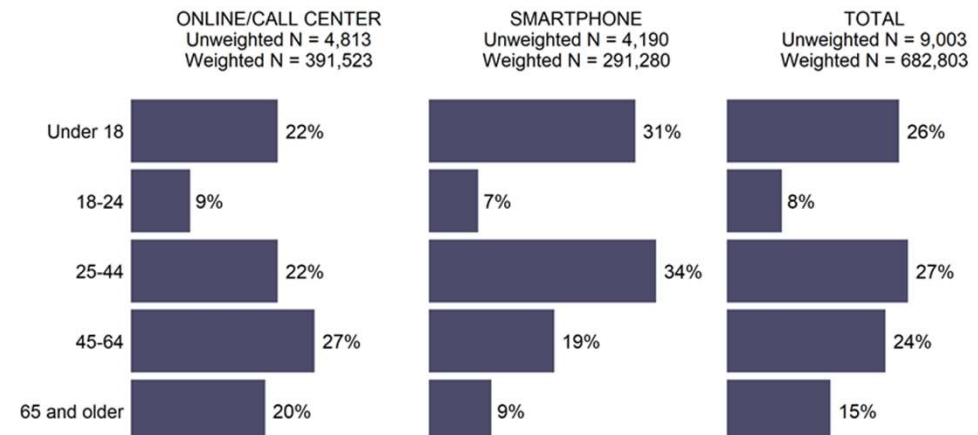
Online and call center participants are more likely to decline reporting household income in comparison to smartphone participants.

Those age 45 or older participated in the survey online or through the call center more frequently. When the age profile of call center participants alone is compared to online and smartphone the sample tends to include more of the elderly.

HOUSEHOLD INCOME BY DIARY MODE



AGE BY DIARY MODE

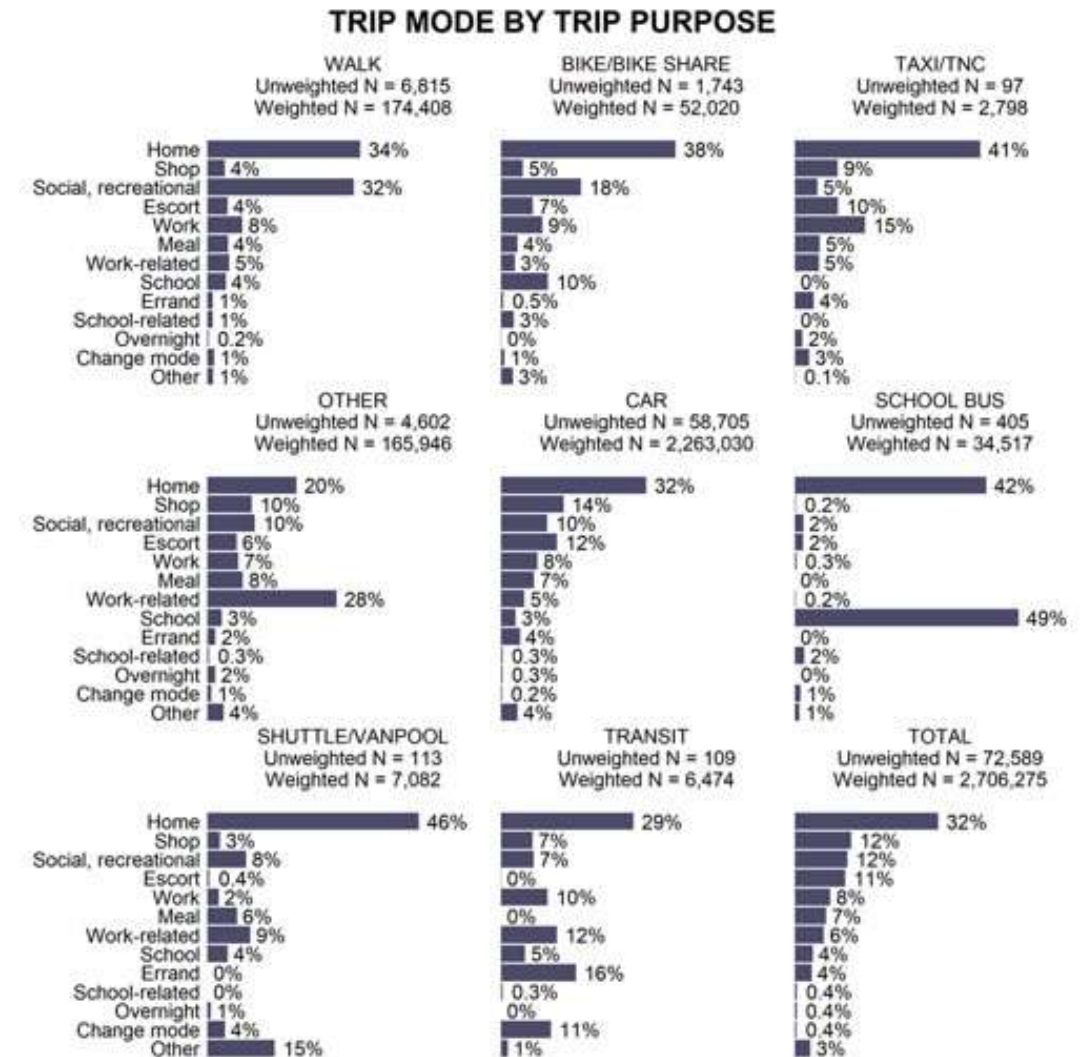


**Note these charts represent the weighted shares and thus only include households that were weighted.*



Provides detailed purpose and mode information

- 32% of walk trips have a social or recreational purpose.
- Transit trips are primarily made to go home, for work/work-related purposes, or errands.
- 18% of bike trips are for social/recreational purposes and 13% are made for school/school-related purposes.

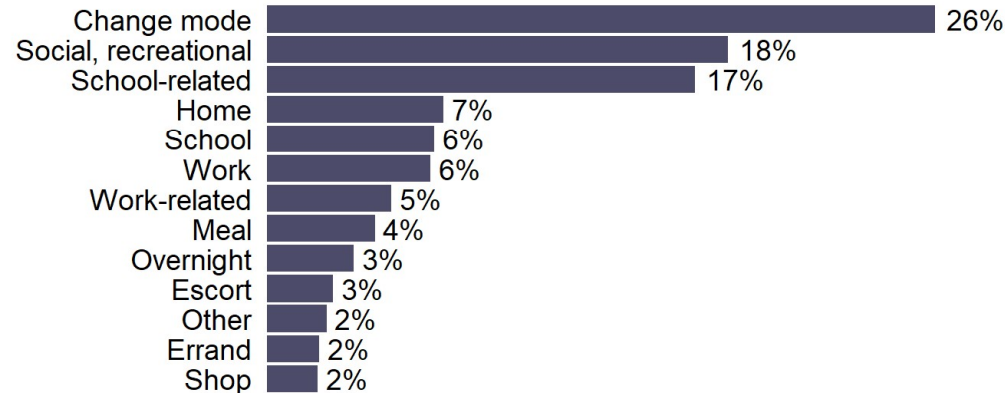


And lets us dig deeper in targeted ways

- A high share of school and school-related trips (23%) as well as social or recreation trips (18%) are made by walking.
- 22% of trips that are less than 1 mile, were walking trips.

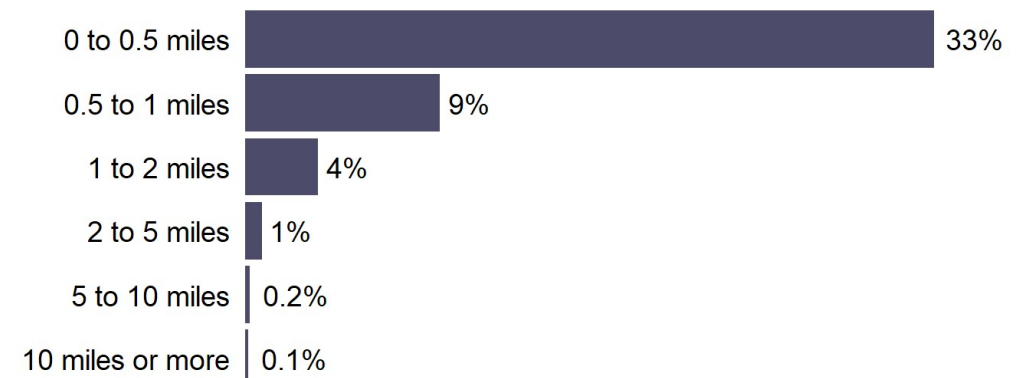
WALKING MODE SHARE BY TRIP PURPOSE

UNWEIGHTED N = 129,573, WEIGHTED N = 2,706,275



WALKING MODE SHARE BY TRIP DISTANCE

UNWEIGHTED N = 129,573, WEIGHTED N = 2,706,275

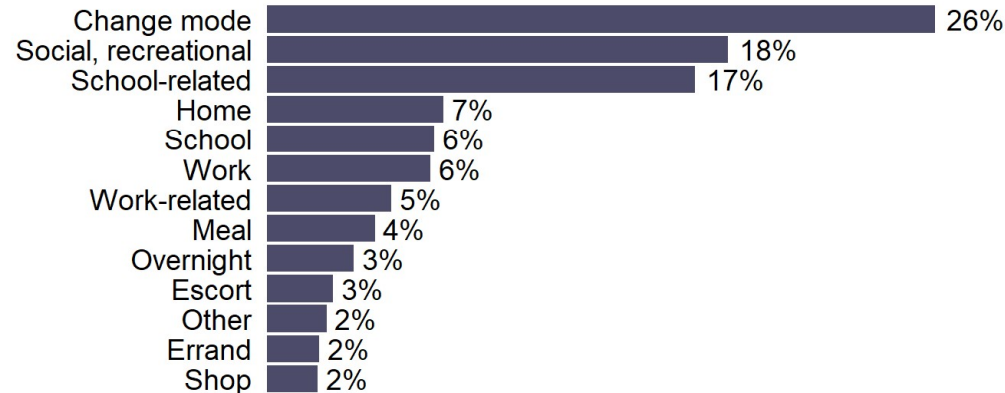


But, it doesn't tell us anything about visitors

- A high share of school and school-related trips (23%) as well as social or recreation trips (18%) are made by walking.
- 22% of trips that are less than 1 mile, were walking trips.

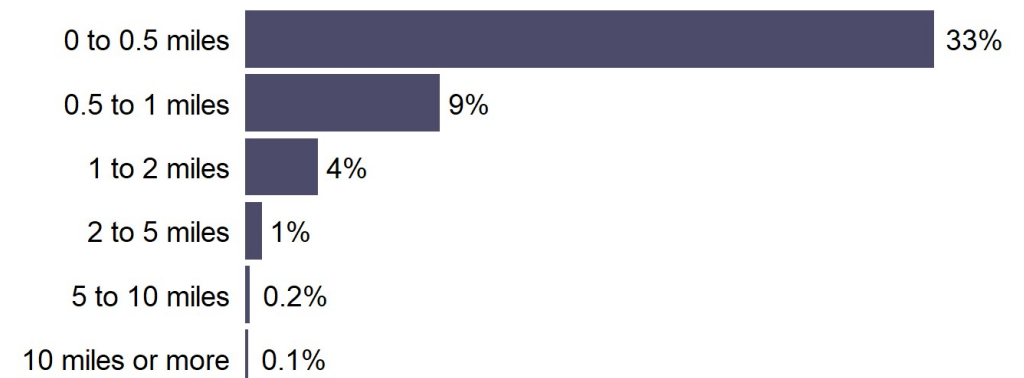
WALKING MODE SHARE BY TRIP PURPOSE

UNWEIGHTED N = 129,573, WEIGHTED N = 2,706,275



WALKING MODE SHARE BY TRIP DISTANCE

UNWEIGHTED N = 129,573, WEIGHTED N = 2,706,275



Passively collected data focused on external travel

OCTOBER
2021



**29,198
RESIDENT
DEVICES**
Unweighted Records



**27,245
VISITOR DEVICES**
Unweighted Records



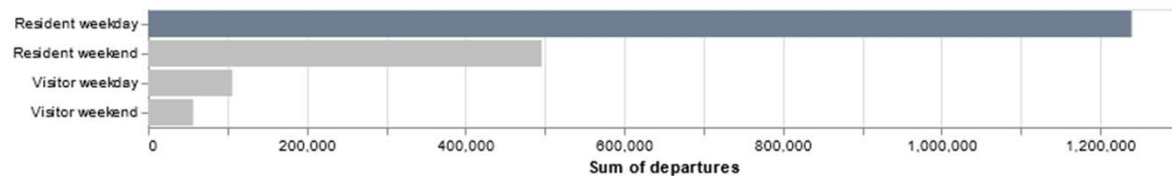
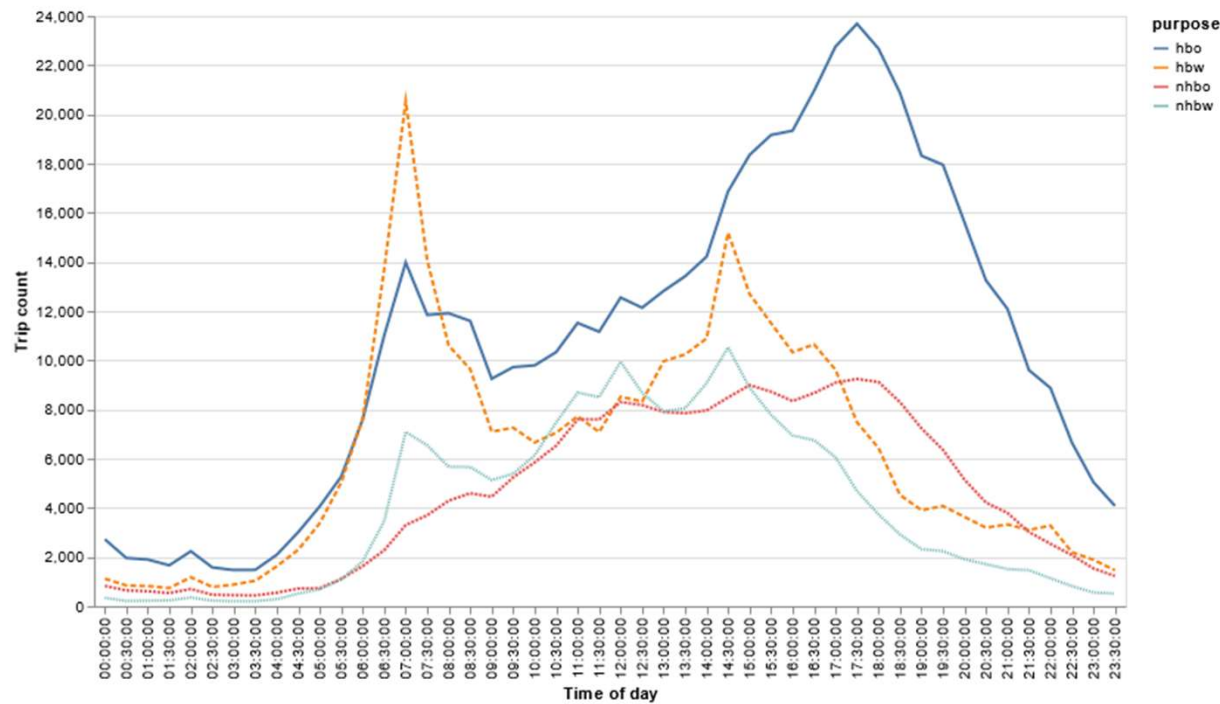
**1,269,738
CLUSTERS**
Unweighted Records



**1,898,715
TRIPS**
Unweighted Records



**213,995,335
SIGHTINGS**
Unweighted Records



- From location-aware smartphone apps
- Records include only unique device id, timestamp, and location
- rMerge processes via four steps:
 - Prepares and filters data
 - Identifies trips
 - Expands to region
 - Summarized and visualized

Passively collected data focused on external travel



1. Preparing data for billions of location points.



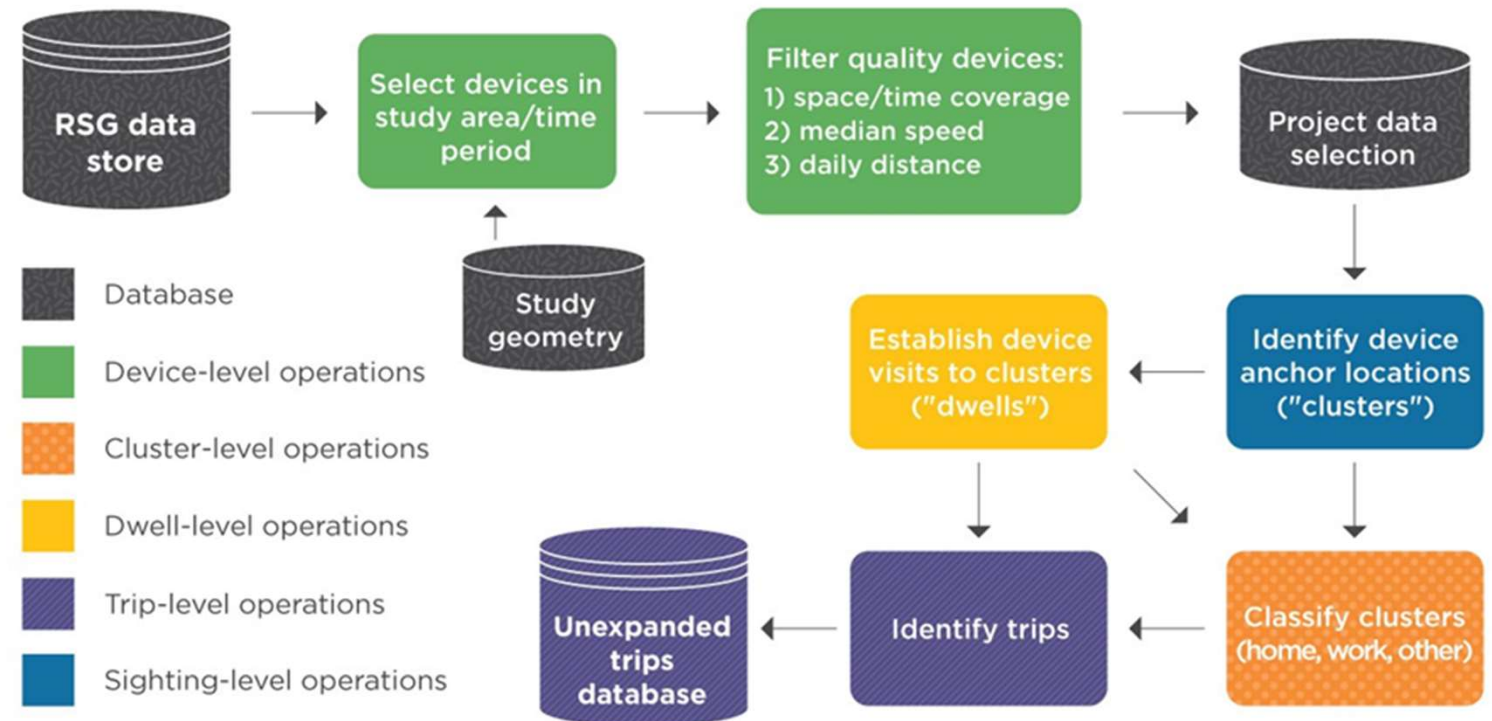
2. Identifying, classifying, and linking stops to create trips.



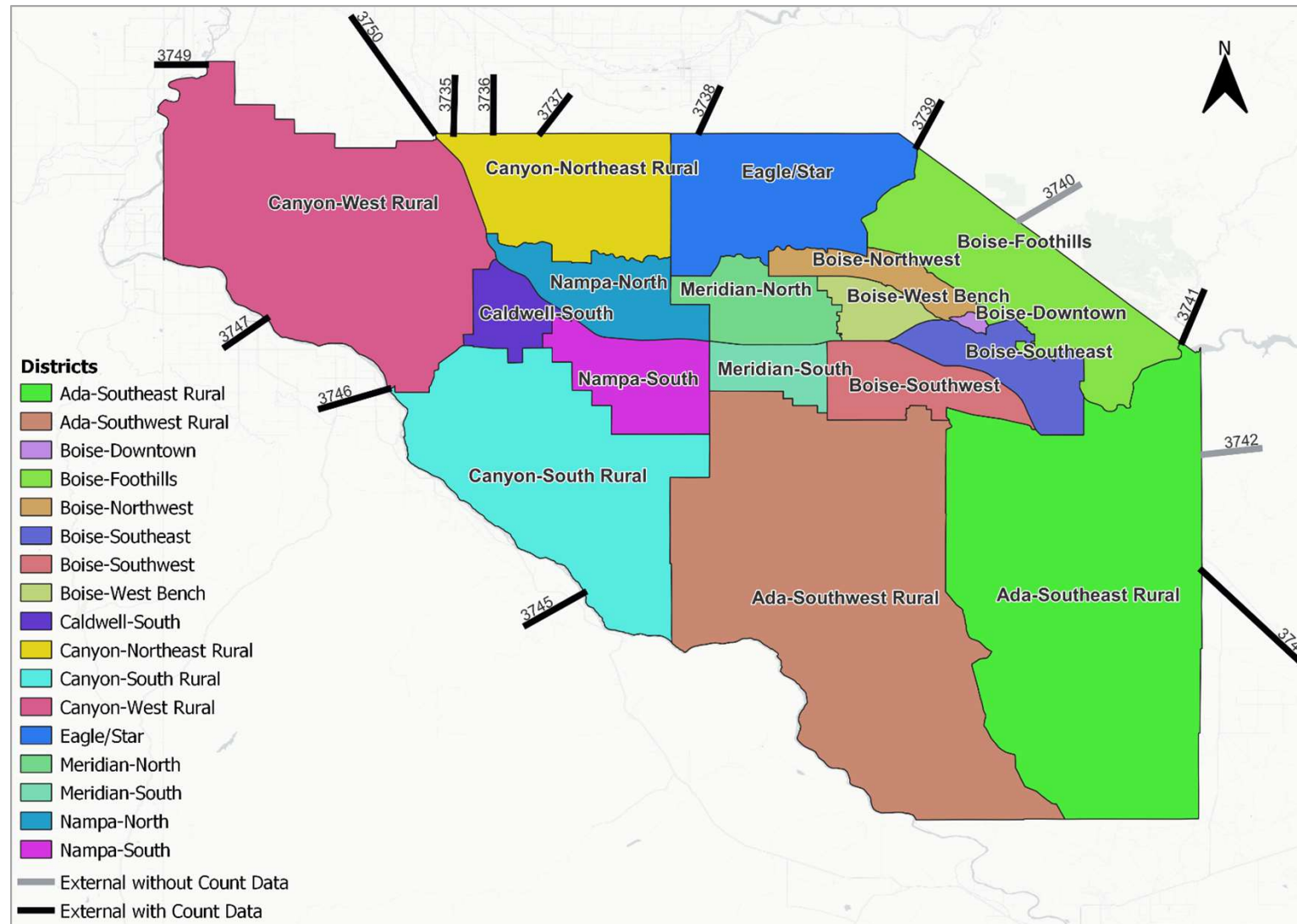
3. Expanding the region based on trip information and other data sources.



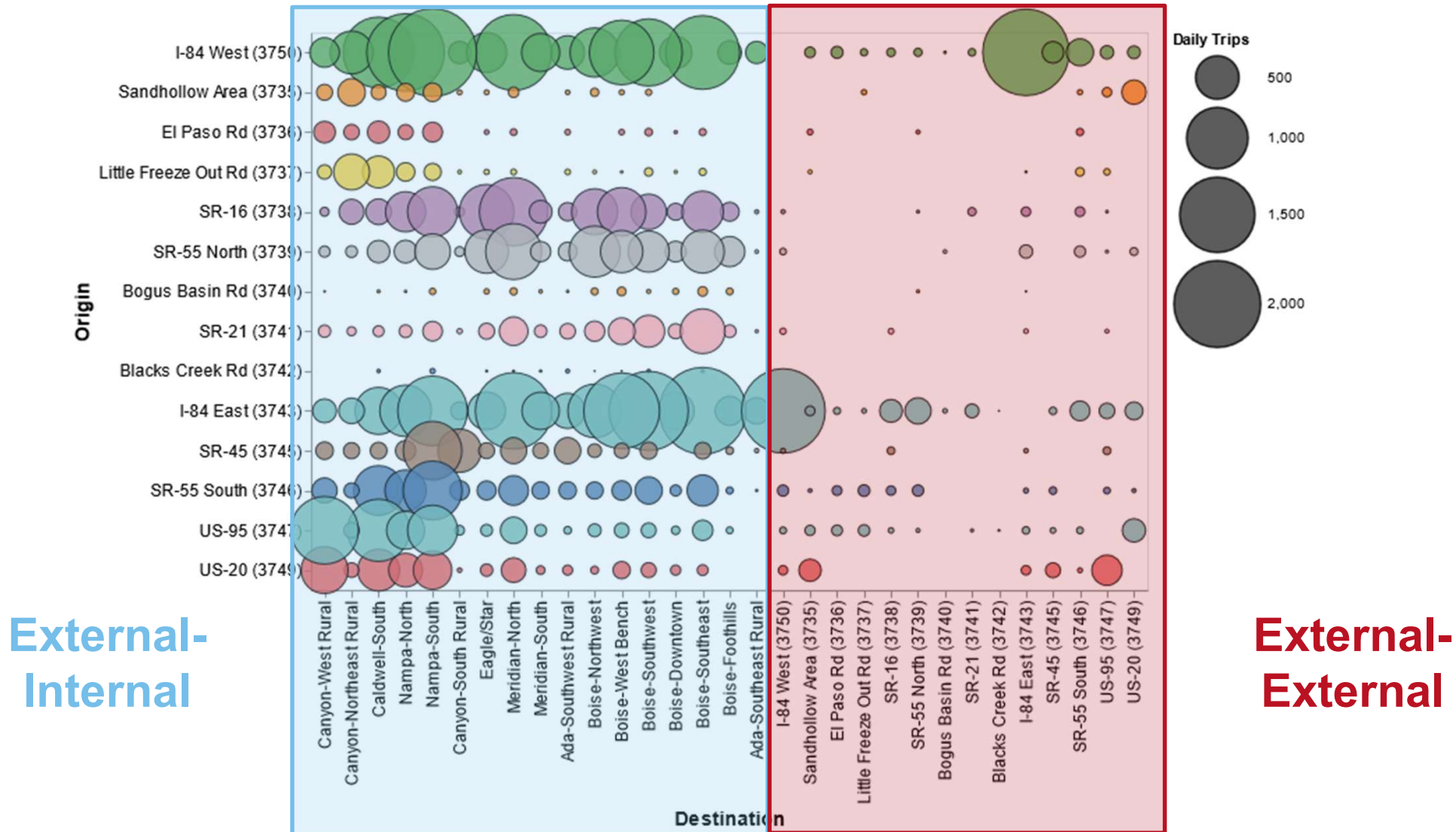
4. Aggregating and visualizing trip data to discover key insights.



Resulting in population-scaled external travel ODs



Resulting in population-scaled external travel ODs



A framework for combining travel surveys and passively collected data

- rMove and rMerge for the Treasure Valley Travel Survey helped COMPASS increase the depth and breadth of data collected.
 - rMove survey data provides an incredible level of detail on travel behavior
 - rMerge passively collected data provides a sample orders of magnitude larger than rMove
- Our Role? **Recognize** that each have their own strengths and weaknesses—and **help clients navigate** these strengths and weaknesses





Summary

The Best of Both: High Detail and High Volume

- Travel surveys are advantageous in their specificity and level of detail (e.g., demographics, trip details, the “why” behind movement in a region).
- Passive big data is advantageous in its comparatively higher volume and greater geographic coverage.



Why use Both Survey and Big Data?

SURVEYS: TRAVELERS, TRIP-MAKING, MODE & TOLL CHOICES

- Surveys provide rich, deep information on who is traveling and why
- Surveys provide insight into how different people choose modes of travel (and tolls)
- Big data is anonymized and contains no direct information on demographics, modes or purposes

BIG DATA: TRIPS (SPATIAL) DISTRIBUTION

- Big data provides information on non-residents
- Big data provides broad information on where people are traveling to and from
- Correctly predicting mode and toll choices requires correctly predicting OD flows
- Surveys do not provide adequate sample size (<2%) to understand OD flows beyond average trip length frequency distributions





Michelle Lee

SENIOR DIRECTOR

Michelle.Lee@rsginc.com

Abby Rosenson

SENIOR CONSULTANT

Abigail.Rosenson@rsginc.com