



# Gen3 Model Development Project

Travel Forecasting Subcommittee Meeting

September 18, 2020

IN PARTNERSHIP WITH



Metropolitan Washington  
Council of Governments

# Discussion Topics

- Synthetic Population Creation
- On-board Survey Data Review





# **Synthetic Population Generation**

# Why is Synthetic Population Needed?

- ActivitySim simulates travel choices for individual households and persons
- This requires the generation of a 'synthetic population' that represents the base or future year population to be modeled
- Data available:
  - Population sample (PUMS, HH Survey)
  - Distribution of person and HH attributes (Census Summary File and/or Agency forecasts)

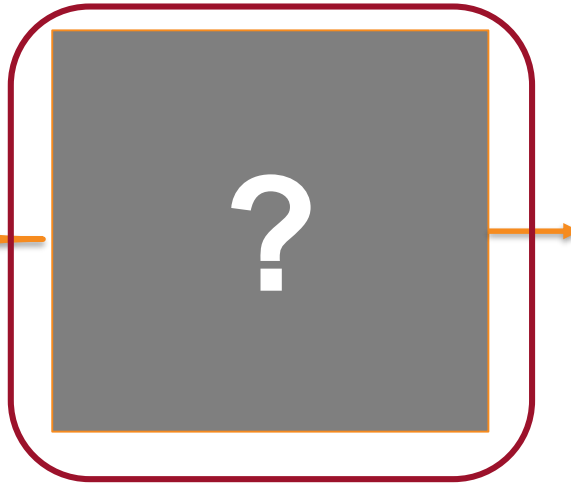


# Problem Statement

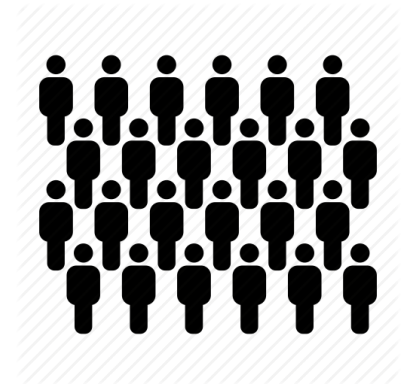
Seed Sample  
(PUMS)



Marginal Controls  
(Census Distributions)



Synthetic Population



Data Fitting Technique  
a.k.a.  
Population Synthesizer

# Terminology

- ***Geography***
  - Geographic resolution or the unit(s) of geography (e.g., MAZ, TAZ, TRACT)
- ***Seed Sample***
  - Sample of household and person records representing a population [e.g., American Community Survey - Public Use Microdata Sample (PUMS)]
- ***Marginal Controls [or Targets]***
  - Aggregate distribution of households or persons for a control variable at a geographic resolution [e.g., Households by Income Quintiles in each TAZ]
- ***Synthetic Population***
  - Set of household and person records representing a region's entire population
- ***Balancing/Fitting***
  - Process of reweighting a disaggregate sample of agents to satisfy aggregate constraints
- ***Integerization***
  - Process of converting the floating point weights to integer weights while satisfying the aggregate constraints



# PopulationSim Software

ActivitySim / populationsim

Unwatch 8 Star 13 Fork 18

Code Issues 17 Pull requests 1 Actions Projects 0 Security 0 Insights

An Open Platform for Population Synthesis <https://activitysim.github.io/populat...>

python data-science population-synthesis activitysim

274 commits 3 branches 0 packages 2 releases 1 environment 7 contributors View license

Branch: master New pull request Create new file Upload files Find file Clone or download

3 authors Develop (#120) Latest commit c9e54f6 5 hours ago

docs	various updates (#111)	3 months ago
example_calm	various updates (#111)	3 months ago
example_calm_repop	various updates (#111)	3 months ago
example_survey_weighting	various updates (#111)	3 months ago
example_test	various updates (#111)	3 months ago
papers	include papers as well	5 months ago
populationsim	various updates (#111)	3 months ago
scripts	Develop (#120)	5 hours ago



# PopulationSim Wiki

PopulationSim  
0.4.1

Search docs

Getting Started

- Installation
- Run Examples

Application & Configuration

Validation of Results

Software Implementation

Resources

Docs » Getting Started [View page source](#)

## Getting Started

This page describes how to install and run PopulationSim with the provided example.

### Installation

1. Install [Anaconda 64bit Python 3](#). Anaconda Python is required for PopulationSim.
2. If you access the internet from behind a firewall, then you will need to configure your proxy server. To do so, create a `.condarc` file in your Anaconda installation folder (i.e. `C:\ProgramData\Anaconda3`), such as:

```
proxy_servers:  
  http: http://myproxy.org:8080  
  https: https://myproxy.org:8080  
ssl_verify: false
```
3. Create and activate an Anaconda environment (basically a Python install just for this project)

```
conda create -n popsim python=3.7  
  
#Windows  
activate popsim  
  
#Mac  
conda activate popsim
```
4. Get and install the PopulationSim package on the activated conda Python environment:

```
pip install populationsim
```

### Python 2 or 3?





# PopulationSim – Key Advantages

- Algorithmic advantages including
  - Entropy maximization-based list balancing
  - Linear programming-based population generation
- Specify controls at multiple levels of geography
- Implemented in ActivitySim framework and planned for consortium management going forward
- Fast runtime
  - DVRPC setup runs 15x times faster than PopGen
- Re-populate feature
  - Generate synthetic population for traffic impact studies and similar small area applications



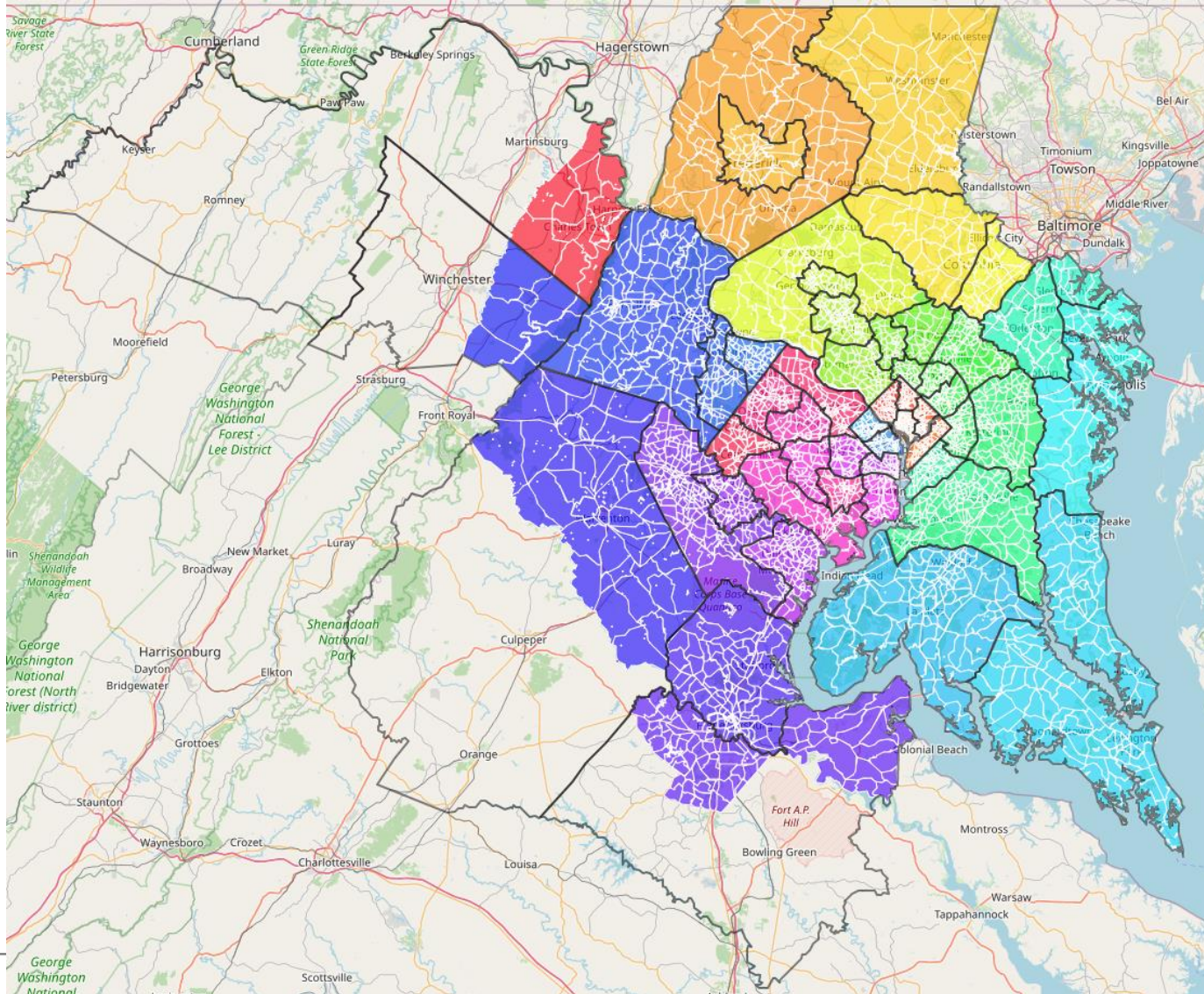
# MWCOG PopulationSim Inputs

Three main inputs:

- Disaggregate household and person samples
  - 2014-18 5-Year ACS PUMS
- TAZ-level marginal controls
  - Tract-level Census distributions
  - MWCOG land use data (Round 9.1a Cooperative Forecasts)
- Geographic crosswalk (TAZ-to-PUMA)
  - TAZs assigned to PUMAs based on maximum overlap



# TAZ – PUMA Crosswalk



# MWCOG PopulationSim Controls

CONTROL VARIABLE	CATEGORIES	PUMS FIELD	CONTROL SOURCE	GEOG
<b>Household Variables</b>				
Total number of households		WGTP	Round 9.1 Cooperative Forecast	TAZ
Household Size	1, 2, 3, 4, 5+	NP	2018 ACS 5-year. Census Tract [Table S2501]	TAZ
Household Income*	0-\$30K, \$30K-\$60K, \$60K-\$100K, \$100k-\$150K, \$150K+	HINCP	2018 ACS 5-year. Census Tract [Table B19001]	TAZ
Number of Workers	0, 1, 2, 3+	ESR	2018 ACS 5-year. Census Tract [Table B08202]	TAZ
Presence of Children	0, 1	HUPAC	2018 ACS 5-year. Census Tract [Table S1101]	TAZ
<b>Person Variables</b>				
Person Age	0-4, 5-19, 20-34, 35-64, 65+	AGEP	2018 ACS 5-year. Census Tract [Table S0101]	TAZ
Person Race/Ethnicity	White, Hispanic, Black, Asian, Other	HISP, RAC1P	2018 ACS 5-year. Census Tract [Table DP05]	TAZ

\* Categories subject to change



# MWCOG PopulationSim Controls (GQ)

CONTROL VARIABLE	CATEGORIES	PUMS FIELD	SOURCE	GEOG
Total GQ units		TYPE	Round 9.1 Cooperative Forecast	TAZ
GQ Type	University, Other Non-Institutional	SCHG & TYPE	2018 ACS 5-year. Census Tract [Table B26101]	TAZ



# MWCOG PopulationSim Implementation



## Data Download

- Auto download Census data and shape files
- Use Census API functionality

## Data preparation

- Seed Sample
- Marginal Controls
- Geographic crosswalk

## Run PopulationSim

- User specified controls and settings
- Separate household and GQ runs

## Post-processing

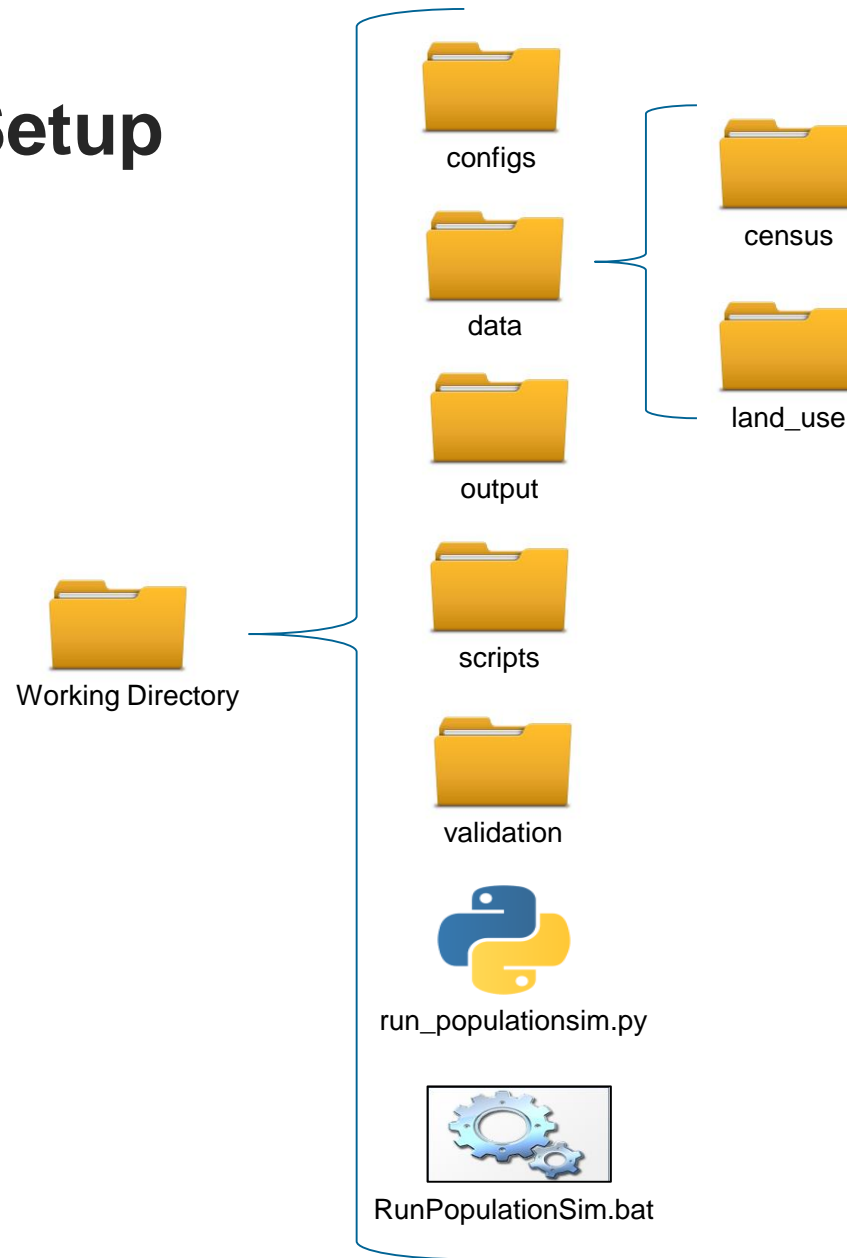
- Combine GQ and residential outputs

## Validation

- Run standard PopulationSim validation script
- Mean percentage difference, STDEV, uniformity distribution



# Directory Setup



# Validation Techniques

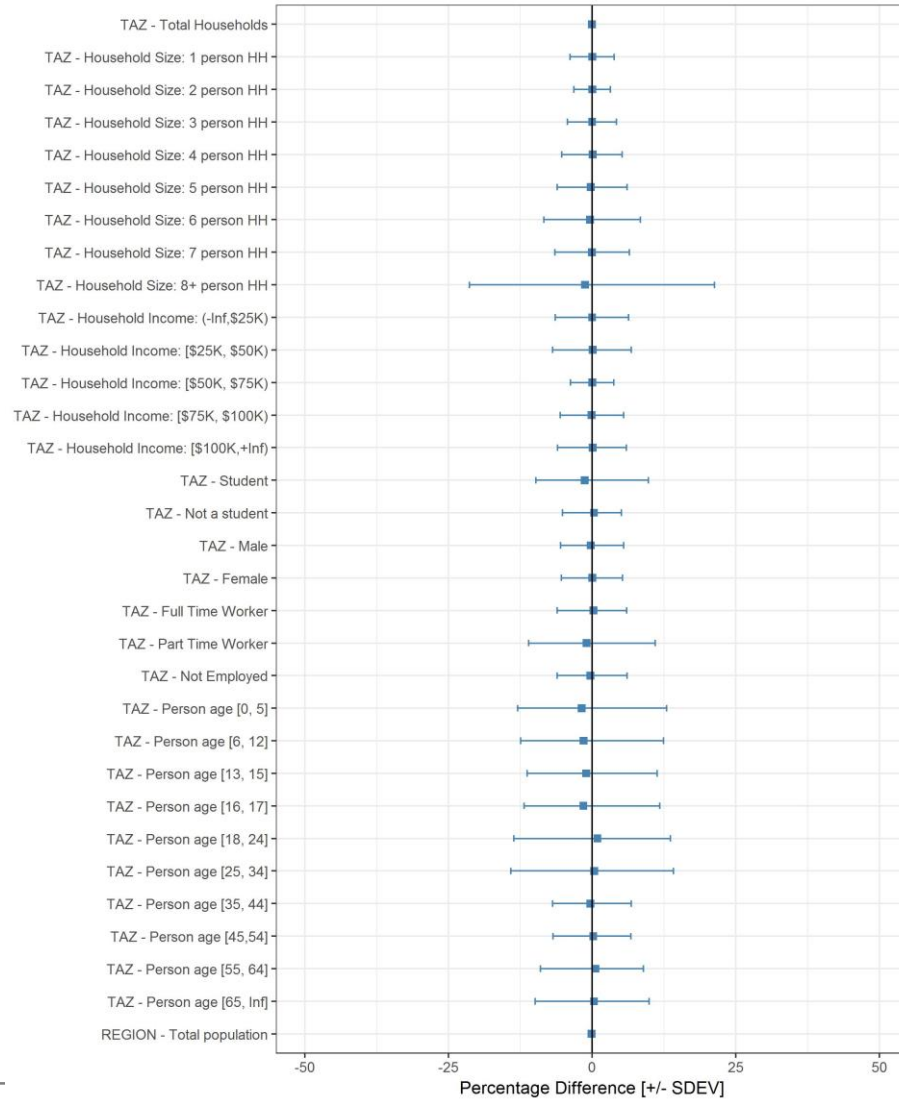
- Synthetic Population vs Marginal Controls
  - Average percentage difference
  - Standard deviation
- Expansion factor distribution
  - Distribution of ratio of initial and final weight
- Uncontrolled variable distributions
- Data checks
  - Sample vs controls
  - Consistency among controls





# Matching Controls

MetCouncil PopulationSim Controls Validation





## **On-Board Surveys**

# Uses of on-board survey for model development

- Understand transit markets:
  - Origin and destination location and purpose (code trip purpose, impute tour purpose)
  - Access and egress mode (walk, bike, PNR, KNR, other)
  - Socio-demographic: autos owned/auto sufficiency, income, person type (worker/student), etc.
- Understand transfer rates
  - Within and between modes/operators
  - Convert boardings to ‘linked trips’
- Test transit network & path-building parameters
- Model estimation, calibration and validation targets
- Other
  - Understand “external” transit trips, fare, visitor trips, etc.

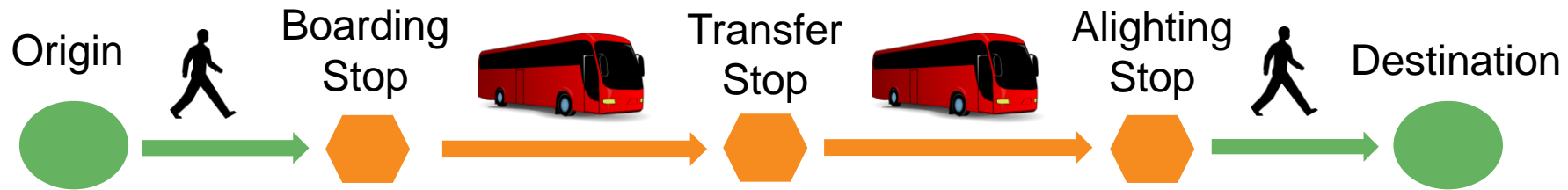


# The ideal on-board survey...

- Has a sample plan based on observed boardings by route, direction, and time-of-day
- Collects the following data
  - Origin and destination location, first boarding location
  - Origin and destination purpose
  - Access and egress mode
  - Route sequence from first boarding to final alighting
  - Relevant household and person information
  - Fare paid
- Implemented with in-person, random interviewing
  - using tablet computers with integrated mapping functionality
- Is carefully expanded to boardings
  - by station on/off for fixed guideway, boarding stop group for high volume routes, total boardings for other routes



# Transit Trip Data



Origin purpose  
Origin place name  
Origin location  
Origin departure time  
Access mode  
Access distance  
First boarding stop location  
Route\run\direction surveyed on  
Boarding stop on surveyed route  
Alighting stop off surveyed route

Route number transferred from  
Route number transferred to  
Total number of routes  
Last alighting stop location  
Egress mode  
Egress distance  
Destination purpose  
Destination place name  
Destination location



# RVTD On-Board Survey (2011)

Please provide any additional comments.

Return the completed survey to the surveyor, or drop it in any mailbox (no postage required).

Thank you!

If you have additional customer comments or questions about this organization and its services, please visit the survey website at [www.webaddress.org/titleofsurvey](http://www.webaddress.org/titleofsurvey), email: [emailaddress@organization.com](mailto:emailaddress@organization.com), or call toll-free 1-800-123-4567.

8 return to mail, please do not open


PTV NUSTATS LLC  
1340 WONDER WORLD DR STE 2212  
SAN MARCOS TX 78666-9805

POSTAGE WILL BE PAID BY ADDRESSEE

BUSINESS REPLY MAIL

FIRST CLASS MAIL PERMIT NO. 25 SAN MARCOS TX

NO POSTAGE NECESSARY IF MAILED IN THE UNITED STATES



## Transit Survey

Please complete this survey to help us improve YOUR transit system.

**1. Register to win one of five \$100 cash prizes when you answer all questions. Provide your contact information below.**

Fill in answers clearly as shown in this example: A B C 1 2 3

Name: \_\_\_\_\_

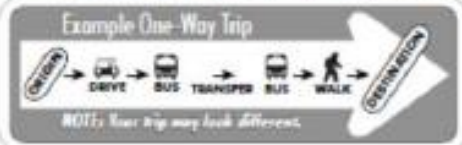
Phone Number: (\_\_\_\_) \_\_\_\_\_ - \_\_\_\_\_ All information is confidential and will not be shared or sold.

**2. Would you be interested in participating in an additional survey to improve transportation in Rogue Valley?**

No  
 Yes → Please provide your contact information in Question 1. We will call you to provide information about how to participate and the rewards you can earn.

The following questions are about the ONE-WAY TRIP you are on now!

Example One-Way Trip



*NOTE: Your trip may look different.*

**3. Did you transfer to this bus? If yes, from which route?**

No  
 RV Mall / Poplar Square  
 Main St / W. Medford  
 Ashland  
 E. Barnett / RVMC  
 Jacksonville  
 Central Point  
 White City / V.A. Dam

**4. Will you transfer after this bus? If yes, to which route?**

No  
 RV Mall / Poplar Square  
 Main St / W. Medford  
 Ashland  
 E. Barnett / RVMC  
 Jacksonville  
 Central Point  
 White City / V.A. Dam

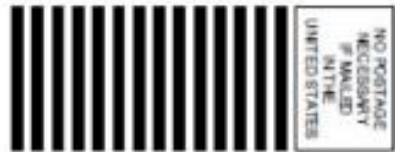
**5. How did you pay for this bus trip?**

Cash  
 Monthly pass - Full Fare  
 Monthly pass - Reduced Fare  
 35 Ride Punch Card Pass  
 Explorer Pass  
 Token  
 Valley Lift ID  
 Ashland Fare 5.50  
 Employer Bus Pass  
 Other, specify: \_\_\_\_\_

**6. Did your employer or another organization pay for your fare?**

Yes  
 No

Continue Inside →



# RVTD On-Board Survey (2011)

## ORIGIN LOCATION

7a. What kind of place are you coming from?

- Home  
 University / College (student only)  
 Shopping  
 Social, Eat Out, Recreational, Religious, Community or Personal Business: \_\_\_\_\_  
 Work or Work-Related  
 High School / Middle School (student only)  
 Medical Services  
 Other, specify: \_\_\_\_\_

7b. What is the exact street address?

(Provide the nearest cross streets if you don't know the exact address)

Address: \_\_\_\_\_

Address: \_\_\_\_\_

Cross Street #1: \_\_\_\_\_

and Cross Street #2: \_\_\_\_\_

City: \_\_\_\_\_ Zip: \_\_\_\_\_

7c. Please record the name of the place or building.  Coming from Home

\_\_\_\_\_

7d. How did you get from your origin location to the first bus stop on your trip?

- Walked / Wheelchair → [ ] # blocks (please answer 0 if less than 1 block)  
 Dropped off  
 Drove alone  
 Carpoled  
 Bicycled  
 Other, specify: \_\_\_\_\_  
 Parking Lot Name / Cross Streets: \_\_\_\_\_

8a. Where did you get on this bus?

- Front Street Station  
 Other, specify cross streets below: \_\_\_\_\_

Cross Street #1: \_\_\_\_\_

and Cross Street #2: \_\_\_\_\_

8b. How long did it take you to get to the bus stop / Transfer Station?

- Less than 5 minutes  
 5-9 minutes  
 10-14 minutes  
 15-19 minutes  
 20-30 minutes  
 More than 30 minutes

## DESTINATION LOCATION

9a. What kind of place are you going to?

- Home  
 University / College (student only)  
 Shopping  
 Social, Eat Out, Recreational, Religious, Community or Personal Business: \_\_\_\_\_  
 Work or Work-Related  
 High School / Middle School (student only)  
 Medical Services  
 Other, specify: \_\_\_\_\_

9b. What is the exact street address?

(Provide the nearest cross streets if you don't know the exact address)

Address: \_\_\_\_\_

Address: \_\_\_\_\_

Cross Street #1: \_\_\_\_\_

and Cross Street #2: \_\_\_\_\_

City: \_\_\_\_\_ Zip: \_\_\_\_\_

9c. Please record the name of the place or building.  Going Home

\_\_\_\_\_

9d. How will you get to your destination location from the last bus stop on your trip?

- Walk / Wheelchair → [ ] # blocks (please answer 0 if less than 1 block)  
 Picked up  
 Drive alone  
 Carpool  
 Bicycle  
 Other, specify: \_\_\_\_\_  
 Parking Lot Name / Cross Streets: \_\_\_\_\_

10. Where do you plan to get off this bus?

- Front Street Station  
 Other, specify cross streets below: \_\_\_\_\_

Cross Street #1: \_\_\_\_\_

and Cross Street #2: \_\_\_\_\_

11. How many days a week do you usually make this trip?

- 0 (irregularly / rarely)  
 1  
 2  
 3  
 4  
 5

12. If bus service was not available, how would you make this trip?

- Walk / Wheelchair  
 Drive  
 Ride with someone else  
 Taxi  
 Bicycle  
 Would not make this trip

13. When you get off this bus, how long will it take you to get where you need to go?

- Less than 5 minutes  
 5-9 minutes  
 10-14 minutes  
 15-19 minutes  
 20-30 minutes  
 More than 30 minutes

14. On a scale of 1-5, how courteous are RVTD's drivers?

- Very Discourteous  
 1 .....  2 .....  3 .....  4 .....  5  
 Very Courteous

15. Do you have a valid driver's license?

- Yes  
 No

16. What is your gender?

- Male  
 Female

17. What is your age?

- 16-18  
 19-64  
 65+

18. What is your ethnicity?

- Asian  
 Black/African American  
 Hispanic  
 Native American  
 White  
 Other

19. Are you ... (fill in all that apply)

- Full-time worker  
 Part-time worker  
 Homemaker  
 Unemployed, but seeking work  
 Unemployed, not seeking work  
 University / College student  
 Middle / High school student  
 Other student  
 Retired

20. How many working vehicles are available to your household?

- None  
 1  
 2  
 3 or more

21. Including yourself, how many people live in your household?

- None  
 1  
 2  
 3 or more

22. Including yourself, how many of the people in your household are employed full-time or part-time?

- None  
 1  
 2  
 3 or more

23. What was your total household income in 2010 before taxes?

- Less than \$10,000  
 \$10,000-\$14,999  
 \$15,000-\$24,999  
 \$25,000-\$34,999  
 \$35,000-\$49,999  
 \$50,000-\$74,999  
 \$75,000-\$99,999  
 \$100,000 or more

Continue →



# Tablet Survey Screenshots

The screenshot shows a web browser window with the URL `sandag.etcsurvey.com/index.php/survey/index`. The page title is "What is the EXACT STREET ADDRESS of this place?". Below the title are four input fields: "Hotel or Place Name", "City", "Zip Code", and "Longitude". Below these are two more input fields: "Street Address" and "State". The "Street Address" field contains the text "100 Park". Below the input fields are two radio buttons labeled "Address" (selected) and "Place", and two buttons labeled "Clear" and "Me".

A map of San Diego is displayed below the input fields. The map shows various neighborhoods and streets. A red pin is placed on the map, and a list of address suggestions is shown on the left side of the map:

- 100 Park Boulevard, San Diego, CA 92101, USA
- 100 Park Boulevard, El Cajon, CA 92021, USA
- 100 Park Way, San Diego, CA 92101, USA
- 2876 El Cajon Boulevard #100, San Diego, CA 92104, USA

The bottom of the screenshot shows the Windows taskbar with various application icons and the system tray showing the time as 12:54 PM on 3/4/2015.





# Tablet Survey Screenshots

The screenshot displays a web browser window with the URL `sandag.etsurvey.com/index.php/survey/index`. The page title is "What is the EXACT STREET ADDRESS of this place?". Below the title is a form with the following fields:

Hotel or Place Name	City	Zip Code	Longitude
<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Street Address	State	Latitude	
<input type="text"/>	<input type="text"/>	<input type="text"/>	

Below the form, a search bar contains the text "petco park". To the right of the search bar are two radio buttons: "Address" (unselected) and "Place" (selected). Further right are "Clear" and "Me" buttons.

The main content area shows a Google Map of San Diego. A red location pin is placed on the map, and a green information box is overlaid on the map with the text: "PETCO Park, 100 Park Boulevard, San Diego, CA 92101, United States". The map includes various neighborhood labels such as Mission Valley, Mid-City, and National City, as well as major highways like I-805 and I-15. Navigation controls for the map are visible on the right side.

The bottom of the screen shows a Windows taskbar with various application icons, including Internet Explorer, Outlook, and Word. The system clock in the bottom right corner indicates the time is 12:53 PM on 3/4/2015.



# Tablet Survey Screenshots

The screenshot shows a web browser window with the URL `sandag.etcsurvey.com/index.php/survey/index`. The survey question is: "Where did you GET ON this bus or train (27 Pacific Beach - Kearny Mesa Transit Center [WB] MTS) for the current one-way trip?".

The form includes the following fields and controls:

- Street Address:** A text input field.
- Latitude:** A text input field.
- Stop ID:** A text input field containing a hyphen (-).
- Longitude:** A text input field.
- User Stop:** A checkbox that is currently unchecked.
- Location Search:** A text input field with the placeholder "Enter a location here".
- Location Type:** Radio buttons for "Address" (selected) and "Place".
- Buttons:** "Clear" and "Me".

Below the form is a map showing a bus route in San Diego. The route is marked with a blue line and red bus icons. The route starts in the La Jolla area, passes through Bayside, and ends at the Kearny Mesa Transit Center. The map includes street names like Soledad Mountain Rd, Clairemont Mesa Blvd, and Genesee Ave, as well as landmarks like Tecolote Canyon Natural Park. The browser's taskbar at the bottom shows various application icons and the system clock indicating 12:56 PM on 3/4/2015.



# Tablet Survey Screenshots

Where did you GET ON this bus or train (27 Pacific Beach - Kearny Mesa Transit Center [WB] MTS) for the current one-way trip?

Street Address: Garnet Av & Mission Bay Dr

Latitude: 32.8064449652

Stop ID: 146406

Longitude: -117.21688900999999

Enter a location here

Address  Place

User Stop

Map Satellite

Map showing bus stop locations along Garnet Ave and Mission Bay Dr. The selected stop is highlighted with a red bus icon and a red location pin.

Windows taskbar: 12:57 PM 3/4/2015



# CTA OD (Routes Used)



Which other CTA buses or trains did you ride on your trip?

First, I rode the... 12 - Roosevelt

Then, I rode the... Red Line

Then, I rode the...

- 66 - Chicago
- 63 - 63rd
- 63W - West 63rd
- 65 - Grand
- 66 - Chicago
- 67 - 67th-69th-71st
- 68 - Northwest Highway
- 70 - Division
- 71 - 71st/South Shore
- 72 - North
- 73 - Armitage

« Previous

© 2017, RSG | [Privacy Policy](#)

Questions or comments? Contact us at [CTAtravelsurvey@rsginc.com](mailto:CTAtravelsurvey@rsginc.com)



# Fare Type Question (SamTrans Survey)

Which fare type do you use?

Adult (age 19 through 64)

Youth (age 5 through 18)

Senior (65 years or older)

Disabled

Medicare cardholder

« Previous

Next »

© 2018, RSG | [Privacy Policy](#)

Questions or comments? Contact us at [samtrans@rsginc.com](mailto:samtrans@rsginc.com)

25%



# Access Mode (Caltrain Survey)

How did you get to the Belmont station for your most recent trip?

Please select all that apply.

Walked
Drove a private vehicle and parked
Got dropped off in a private vehicle by family or friends
Employer shuttle
Transit (BART, Muni, SamTrans, VTA, etc.)
Taxi
Smartphone-enabled car service (Lyft, Uber, Lyft Line, etc.)
Chariot
Carpool / Vanpool
Bicycle
Bike Share
Other

« Previous Next »

© 2017, RSG | [Privacy Policy](#)  Questions or comments? Contact us at [caltrain@rsgsurvey.com](mailto:caltrain@rsgsurvey.com)

62%



# Survey Instruments/Documentation Reviewed

- 2016 Metrobus and Metrorail Surveys
- 2016 and 2018 Virginia Railways Express Surveys
- 2016 Maryland Area Regional Commuter (MARC) Rail Survey
- 2018 Ride On Title VI Survey
- 2018 DC Circulator Customer Satisfaction Survey
- OmniRide Westerly Subsidy Study
- 2017 Potomac and Rappahannock Transportation Commission (PRTC) On-Board Survey
- 2016 TransIT customer satisfaction survey
- Carrol Transit customer satisfaction survey
- Calvert County Public Transportation survey



# Summary of Survey Data

- Many surveys collected mode of access and mode of egress, fare paid, and limited socio-demographic variables
  - Minimum information required for Title VI reporting
- Most surveys did not collect origin and destination location
  - MARC collected most of the data required for modeling
  - Metrorail collected origin/destination station
  - VRE collected home address and destination address
- Route sequence unknown
- Most surveys used simple trip purposes
- Smaller operators tend to have less useful data for modeling





# Conclusions

- May be able to glean some useful ridership characteristics from surveys and compare to transit trips from household survey
  - Data could be used to create or adjust calibration targets
- Not possible to assign trips to network or attach skims to trips except in limited cases
  - Rail trips – boarding/alighting station
  - Trip segment between home and boarding/alighting
- Datasets may require re-expansion
- Suggest MWCOCG initiate regional coordination on transit on-board survey design and implementation





the science of insight



---

**Joel Freedman**

SENIOR DIRECTOR

[Joel.Freedman@rsginc.com](mailto:Joel.Freedman@rsginc.com)