SUPPLEMENTAL REFERENCE DOCUMENT FOR DATASETS

PUGET SOUND REGIONAL TRAVEL STUDY 2015

This document provides additional explanation of the Puget Sound Regional Travel Study datasets beyond the variable labels in the data dictionaries, for the datasets in the March 2016 Release 1. A comprehensive list of data variables is in the data codebooks and tabulations. This document seeks to expand on those descriptions in cases where additional explanation may be helpful.

1.1 | PRIVACY

This dataset guide accompanies versions of the datasets that do not include certain data that can be considered sensitive or confidential. Specifically, home, work, school and trip (location) coordinates are provided. The coordinates, along with detailed geographic layers, can be obtained by special request, for use in academic and policy research.

1.2 | OVERVIEW

2,442 HHs completed the main study in April–June of 2015.

After review of the dataset, 23 HHs were excluded from the final dataset. The final number of HHs in the dataset is 2,419. This includes data from 2015 for two types of HHs:

1. 816 Cross-sectional HHs (newly invited in 2015 based on sample plan criteria)
2. 1,603 Panel HHs (completed HTS in 2014 and re-invited to take HTS in 2015). The 2015 survey data are included in this dataset.

The data deliverable includes four datasets:

1. HH-level data
2. Person-level data
3. Trip-level data
4. Vehicle-level data

All data in these datasets are from HHs that completed the entire survey, meaning they filled out the HH information/recruit survey and every travel diary for HH members age 5 and over. Partially complete HHs are not included in these datasets.

MISSING DATA

Blank cells are intentionally missing data, e.g. a question was not asked, or an answer choice was not shown to the respondent. The HH income questions offered respondents the option to select "Prefer not to answer", and has been coded 98 in the dataset. If data are missing because of a logic error, the value is coded to -9999 (unintentionally missing). This dataset does not have any unintentionally missing data.

1.3 | HH-LEVEL DATASET

The HH-level dataset has 2,419 rows- one row per HH.

Unique identifier: HHID

Example = 15100432. Cross-sectional HHIDs start with '15', marking the year of the study. Panel HHIDs are the same as in 2014, e.g. 14104442.

Sample type (sampletype)

1. 816 Cross-sectional HHs (newly invited in 2015 based on sample plan criteria)
2. 1,603 Panel HHs (completed HTS in 2014 and re-invited to take HTS in 2015). The 2015 survey data are included in this dataset.

Home location variables (h_cnty, h_city, h_zip)

These variables are derived from the final home address provided with the original data. They are the result of reconciling the sample provider address, the address reported by the HH in the recruit survey, and in the case of Panel households, the final 2014 home location.

Version of home address used in derived geographic variables (address_use_flag)
Sample provider or reported home address. See the report for more documentation.

Note: In the **Trip-level** dataset, the reported home address was retained as a default, because the Google time and distance correspond with the reported coordinates. When it was determined that the reported home address was incorrect (some respondents reported a nearby landmark, school, or business as the home location), the distance, Google time, and implied speed were recalculated.

Possible Panel move (Panel_moveinspect)

The distance between 2014 final home address and 2015 reported home address exceeded 0.25 miles. Inspect for possible move or other issues with respondent reporting.

Note: Before mailing invitations, the PSRC 2015 address file was checked against the National Change of Address (NCOA) register. Approximately 5-7% of addresses were flagged as having moved (mail forwarding in effect). Because the sampling plan for Cross-sectional and Panel HHs was based on original addresses, and to ensure most Panel invitations reached the same HH in 2015 as in 2014, only addresses without moves were invited. Therefore, this dataset cannot answer questions about HH moves, e.g. how many HHs move in a year. See instead res_dur ("how long lived in current residence").

Sample segment (sample_segname)

The segment assigned to the HH based on the purchased sample home address.

Home segment (h_segname)

Based on the final home address (sample provider or reported). Both sample segment and final segment are used in weighting, and thus included in this dataset.

Recruit survey duration (hh_info_dur)

The recruit survey duration is calculated as the difference between the timestamp recorded for the last survey page ("participate") and the timestamp for the first survey page ("intro"). Exclude extremely long durations when interpreting survey duration as it is possible the respondent left their web browser open for periods of time.

HH income variables

HHs had the option of reporting income in ten categories or select "prefer not to answer" ("hh_income_detailed"). A follow-up question offered the option of reporting a broad income category.

Foreign language HH (foreign_language)

HHs were flagged as a 'foreign language participant' if one or more of the following was true: HH used a web browser with a non-English language setting, used Google Translate on at least one page, called the PSRC language line, or called the call center and participated in Spanish.

Previous residence address information (prev_home_loc…)

Participants who reported living in their current residence for five years or less (res_dur) were asked whether their previous home was in Washington. If yes, participants were asked to locate their previous home address on a map. If the previous home was outside of Washington, participants had the option to either input the city and state of their previous residence, or both.

HH lifecycle stage (lifecycle)

A variable derived for use in weighting. Derived from the presence and ages of children, and the number and ages of adults in the HH. HHs are initially classified by the presence of children, first by the presence of any young children (under 5 years old) and then by the presence of only school children (ages 5-17). If there are no children in the HH, it is then classified by the HH size (either 1-person or 2-or-more-person HH) and the age of the oldest person in the HH (under 35, 35-64, and 65 years and over). Note: This variable uses the same lifecycle definition as PSRC 2006, with the slight difference that the age category is "under 5".

HH number of trips on travel day (hhnumtrips)

The number of trips reported by each HH is the count of trip records associated with each HH's password. This includes trip records for children under 5 (see Trip-level dataset).

NEW HH-LEVEL VARIABLES IN 2015

sampletype "Sample type: Cross-sectional or Panel"

tacoma_addon "Sample purchase group (PSRC or City of Tacoma) (derived)"

1.4 | VEHICLE-LEVEL DATASET

The vehicle-level dataset has 3,705 rows – one row per vehicle reported by the 2,127 HHs (out of 2,430) that own a vehicle.

Unique identifier: hhid + vehnum

To get a unique identifier for the vehicle-level dataset, concatenate hhid and vehnum.

NEW VEHICLE-LEVEL VARIABLES IN 2015

None.

1.5 | PERSON-LEVEL DATASET

These are the person-level variables from the travel diary, for which there is one row per person for a total of 4,814 persons, including adults and children.

Unique identifier: PersonID

Example = 1510043201, where each PersonID is the HH's HHID (15100432) with a unique ID appended for each HH member (01, 02, etc.)

Person number (pernum)

Person number (1 through hhsize). Unique within each HH. Person number 1 completed the recruit survey.

Type of HH respondent (resptype)

This variable identifies a "primary" person (the person who completed the recruit survey), "other" HH members (any HH members aged 18 and up), and children (any HH members under the age of 18).

Number of trips made on travel day (numtrips)

The number of trips reported by each respondent is the count of trip records associated with each person's personID. This includes trip records for children under 5 (see Trip-level dataset).

Travel diary duration (diary_duration_minutes)

The full diary duration, or the amount of time participants spent taking the travel diary survey, is calculated as the difference between the timestamp recorded for the first question in the diary "proxy" and the last question ("comments"). Use care when interpreting survey duration, as it is possible the respondent left their web browser open for long periods of time.

Proxy variable (proxy)

The "proxy" variable indicates if respondents took the diary for themselves, if other people filled out the answers while they were present, or if people filled out the answers and they were not present. The call center was given instructions to answer this question as if they were the respondent. HHsize 1 is coded 1.

Previous workplace

Participants who commute to their current workplace for five years or less were asked whether their previous workplace was in Washington. If yes, participants were asked to locate their previous work address on a map. If the previous workplace was outside of Washington, participants had the option to either input the city and state of their previous workplace, or both.

Had additional trips to report (added_trip_flag)

Indicates if the respondent went back and added more trips in the roster after seeing the prompt asking whether they had made any additional trips not already reported. Also see related variables with prefix "added_" for the types of trips that were added.

NEW PERSON-LEVEL VARIABLES IN 2015

commute_mode "If commutes: Typical commute mode"

wpktyp "Commutes: Where parks when drive to work"

wpk_dur " Travel travel time (minutes) from parking location to workplace "

wpk_dis " Travel distance (miles) from parking location to workplace "

carshare_car2go "Has carshare membership: car2go"

carshare_relayrides "Has carshare membership: RelayRides"

carshare_zipcar "Has carshare membership: Zipcar"

carshare_none "Does not have any carshare membership"

carshare_dontknow "Does not know if has carshare membership"

carshare_other "Has other carshare membership"

carshare_other_specify "Has other carshare membership: Specify"

prontoshare "Age 16+: Has/plans to get Pronto Cycle Share membership"

share_freq_car2go "Past 30 days: How often used car2go"

share_freq_relayrides "Past 30 days: How often used RelayRides"

share_freq_zipcar "Past 30 days: How often used Zipcar"

share_freq_othercar "Past 30 days: How often used other carshare service"

share_freq_lyft "Past 30 days: How often used Lyft"

share_freq_sidecar "Past 30 days: How often used sidecar"

share_freq_uberx "Past 30 days: How often used Uberx"

share_freq_pronto "Past 30 days, heard of Pronto Cycle Share: How often used Pronto Cycle Share"

av_interest_nodriver "Autonomous cars interest: Taking a taxi ride in an autonomous car with no driver present"

av_interest_backupdriver "Autonomous cars interest: Taking a taxi ride in an autonomous car with a back-up driver present"

av_interest_commutesov "Autonomous cars interest: Commuting alone using an autonomous vehicle"

av_interest_commutehov "Autonomous cars interest: Commuting with others (carpool) using a shared autonomous vehicle"

av_interest_own "Autonomous cars interest: Owning an autonomous car"

av_interest_carshare "Autonomous cars interest: Participating in an autonomous car-share system for daily travel"

av_interest_short "Autonomous cars interest: Riding in an autonomous car for a short trip to get to a vehicle (e.g. from airport terminal to parking lot)"

av_concern_safeequip "Autonomous cars concern: Equipment and system safety"

av_concern_legal "Autonomous cars concern: Legal liability for drivers or owners"

av_concern_safeveh "Autonomous cars concern: System and vehicle security"

av_concern_react   "Autonomous cars concern: Capability to react to the environment (other cars, bicyclists, pedestrians, etc.)"

av_concern_perform "Autonomous cars concern: Performance in poor weather or other unexpected conditions"

wbt_more_transitsafety "Walk, bike or ride transit more if: Safer ways to get to transit stops (e.g. more sidewalks, lighting, etc.)"

wbt_more_transitfreq "Walk, bike or ride transit more if: Increased frequency of transit (e.g. how often the bus arrives)"

wbt_more_reliability "Walk, bike or ride transit more if: Increased reliability of transit (e.g. the bus always arrives at exactly the scheduled time)"

bike3mile "How often would make a 3 mile 15 minute bike trip if safe and convenient bike route available"

telecommute_if "If not telecommute already: How often would telework if employer offered telework option"

1.6 | TRIP-LEVEL DATASET

These are the trip-level variables, from the trips made on the assigned travel date. The trip-level dataset has 19,675 trips – one row per (one-way) trip.

The travel date starts at 3 AM on the assigned travel date and ends at 3 AM the following day.

Record ID (recordID)

A unique ID number for each trip record. The ID number comprises the letter "C" following by a 5-digit number, beginning with 00001. The "C" indicates this is the third survey conducted within the 2014 Puget Sound Regional Travel Study, following the 2014 Household Survey and the 2014 College Survey. The ID was assigned to the original, unprocessed data and remained unchanged despite any subsequent additions or deletions. Added records are identified by 5-digit numbers starting with 90001.

Unique identifier (tripID)

Example = 151004320101, where each tripID is the HH's HHID (15100432) with a unique ID appended for each HH member (01, 02, etc.) and a unique ID appended for each unique trip (01, 02, etc.)

Person number (pernum)

Person number (1 through hhsize). Unique within each HH. Person number 1 completed the recruit survey.

Trip number (tripnum)

Values 1 through number of trips for a respondent, sorted by departure time (the first trip of the day is trip number 1). Trip number is unique within each respondent.

Last trip in the diary (last)

The value 1 is assigned to the last identified trip in a person's diary. Otherwise the field is empty.

Origin trip purpose and destination trip purpose (o_purpose and d_purpose)

Respondents report the destination trip purpose. The origin purpose is derived from the destination purpose of the previous trip, except for first trip in the day. For the first trip in the day, origin purpose is instead coded based on the response to where the respondent started the day (home, work or other). If the respondent was branched over that question, the origin-location description in place_start is used. O_purpose for the first trip in the day is typically home.

Start and end time: Minutes after midnight (time_start_mam and time_end_mam)

Format is minutes after midnight, from 3 AM on the travel date to 3 AM on the following day.

Start and end time: HH:MM (time_start_hhmm and time_end_hhmm)

Format is HH:MM (24-hour clock). Included for convenience.

Reported trip duration (trip_duration)

Travel time is derived as the difference between respondent-reported start and end time of the trip.

Driving distance and time (gdist, gtime)

The RSG survey instrument estimated travel time and duration for each trip in addition to the user-reported travel time. The estimates of time and duration were calculated using the Google Maps API Distance Matrix Service. These estimates indicate the distance and duration of a trip for "standard driving directions using the road network". The estimates do not account for traffic, thus representing free flow conditions on the roadway. Google estimated drive time and distance can be used to validate the reported trip durations. Time and distance estimates were not available for ferry and airplane trips.

Implied speed in miles per hour (implied_speed)

Driving distance over reported travel time. Can be used in trip validation to detect trip records with issues (extremely high or low speeds).

Activity duration at destination (a_dur)

Time spent at destination. For the last trip of the day, A_DUR is the time remaining until 3 AM on the following day.

Driver or passenger (driver)

Asked of respondents whose mode was drive with others (HH members or not) or vanpool. Derived as driver for drive-alone trips.

Trip copied from other HH member (prepop)

Flag to indicate that a respondent copied this trip from another HH member who had already reported them on their trip. This option is available to all HH members regardless of age, and reduces the respondent burden of repeating trip details.

Trip record derived from trip reported by HH member age 5 or over (child_under5)

Children under the age of 5 were not required to complete a diary, but could be reported on trips made by HH members 5 or older. During data processing, trip records were created for the children under 5 by copying trip records from other HH members and editing relevant details such as creating unique trip ids and recoding instances of 'driver' to 'passenger' for vehicle trips. The variable "child_under5_originaltripid" indicates the trip the record was created from.

HH member on trip (member1 through member8)

PersonIDs of HH members on the trip.

NEW TRIP-LEVEL VARIABLES IN 2015

park_loc_ dur " Travel time (minutes) from parking location to destination"

park_loc_ dis " Travel distance (miles) from parking location to destination"