trip-behaviors

suzanne

2024-04-18

Trips per person per day by Purpose by Mode

by Worker/Non-Worker

title: "first-example-mode-summary" author: "suzanne" date: "2024-03-20" output: pdf_document: default html_document: default —

Load in Packages

IF YOU need to download some packages on github: travelSurveyTools, psrcelmer, and psrcplot. For this use the syntax: library(devtools) devtools::install_github('rsgInc/travelSurveyTools') devtools::install_github('psrc/psrcelmer') devtools::install_github('psrc/psrcplot')

Debugging travelSurveyTools

travelSurveyTools still has many cryptic error messages usually related to data or codebook formatting and naming. You may need to debug to find where or why the problem exists.

You can use: devtools::load_all('C:/GitHub/travelSurveyTools/R') and then browser() and put breakpoints in the code to find the problem.

Read in data items from Elmer

The travelSurveyTools package expects everything to be a data.table, so run setDT all over the place.

```r
hh<- get_query(sql= "select  household_id as hh_id, hhincome_broad,survey_year, hh_weight
                from HHSurvey.v_households_labels")


person<- get_query(sql= "select household_id as hh_id,race_category,
person_id, workplace, telecommute_freq, survey_year, person_weight, gender, age, industry,office_availab
                from HHSurvey.v_persons_labels")

day<- get_query(sql= "select person_id, day_id, household_id as hh_id, telework_time, day_weight , surve

trip<- get_query(sql= "select trip_id, household_id as hh_id, day_id,
person_id, mode_1, dest_purpose, travelers_total, survey_year, trip_weight
                from HHSurvey.v_trips_labels")
setDT(hh)
```

```
setDT(person)
setDT(day)
setDT(trip)
```

#race_category needs to be added to the codebook

to do: add Seattle, Bellevue geographies

Read in Codebook

Read in Codebook

Set IDs as characters

I guess for joining?, functionalize convert all ids to charcets

```
hh[, hh_id:=as.character(hh_id)]
person[, hh_id:=as.character(hh_id)]
day[, hh_id:=as.character(hh_id)]
trip[, hh_id := as.character(hh_id)]

person[, person_id := as.character(person_id)]
day[, person_id := as.character(person_id)]
trip[, person_id := as.character(person_id)]

day[, day_id := as.character(day_id)]
trip[, day_id := as.character(day_id)]

trip[, trip_id := as.character(trip_id)]
hh <- hh%>%mutate(survey_year=as.character(survey_year))
person <- person%>%mutate(survey_year=as.character(survey_year))
day <- day%>%mutate(survey_year=as.character(survey_year))
trip <- trip%>%mutate(survey_year=as.character(survey_year))
```

Make a new variable

If your variable is going to be a transformation of existing data in the codebook- for example grouping- you need to add the new variable and its new values to the internal to code, working codebook.

Adding a new variable to the codebook

make a function with bunch of default #to do make this easier

```
variable_list<-rbind(
    variable_list,
    data.table(
      variable = c("mode_simple", "mode_w_sov"),
      is_checkbox = c(0,0),
      hh = c(0,0),
```

```
      person = c(0,0),
      day = c(0,0),
      trip = c(1,1),
      vehicle = c(0,0),
      location = c(0,0),
      description = c("mode aggregation","SOV mode aggregation"),
      logic = c("mode aggregation","SOV mode aggregation"),
      data_type = c("integer/categorical","integer/categorical"),
      shared_name = c("mode_simple", "mode_w_sov")
    )
  )
```

Add associated values

```
group_labels<-get_grouped_labels(group_id='group_1', group_name='mode_simple')
value_labels<-add_values_code(group_name='mode_simple')
trip<-grp_to_tbl(tbl=trip, ungrouped_name='mode_1', grouped_name='mode_simple')
```

```
value<- c(0,1,2,3,4,5)
label=c('SOV', 'HOV', 'Walk', 'Transit', 'Bike', 'Other')
val_labels<-data.frame(value, label)
val_labels<-val_labels%>%mutate(variable='mode_w_sov')%>%
    mutate(group_1_title = NA, group_1_value = NA,
           group_2_title = NA, group_2_value= NA,
           group_3_title = NA, group_3_value = NA)

all_value_labels<-value_labels%>%select(variable, value, label, group_1_title, group_1_value,
                                 group_2_title, group_2_value, group_3_title, group_3_value)
new_value_labels<-rbind(all_value_labels, val_labels)
new_value_labels[, val_order := seq_len(nrow(new_value_labels))]
value_labels<-new_value_labels
```

add mode_w_sov to the trip table

```
trip<- trip%>%mutate(travelers_total_fix= ifelse(travelers_total!='1 traveler', 'More than 1', '1 travel
trip<-trip%>%mutate(mode_simple= replace_na(mode_simple, 'Drive'))%>%
mutate(mode_w_sov=case_when(
  mode_simple=="Drive"& travelers_total=='1 traveler' ~ 'SOV',
  is.na(travelers_total) ~ 'SOV',
  mode_simple=="Drive"& travelers_total!='1 traveler'~  'HOV',
  .default= mode_simple
))
```

The package expects the data to be in a list of data.tables.

```
hts_data = list(#hh = hh,
                #person = person,
```

```
                    #day = day,
            trip = trip)
```

```
#some how a duplicate snuck into the variable list not sure how
variable_list<-variable_list%>%distinct(variable, .keep_all=TRUE)
```

##summarize data

```
mode_summary = summarize_weighted(hts_data= hts_data,
                            summarize_var = 'mode_w_sov',
                            summarize_by = 'survey_year',
                            id_cols='trip_id',
                            wt_cols='trip_weight',
                            wtname='trip_weight'
                            )
```

```
## Warning in hts_validate_variable_list(variables_dt, data): Variable(s) do not
## have a location specified: accuracy, age, attend_school_1, attend_school_2,
## attend_school_3, attend_school_998, attend_school_999, bearing, bike_freq,
## broadband, carshare_freq, collect_time, commute_dur, commute_freq,
## commute_subsidy_1, commute_subsidy_2, commute_subsidy_3, commute_subsidy_4,
## commute_subsidy_5, commute_subsidy_6, commute_subsidy_7, commute_subsidy_996,
## commute_subsidy_998, commute_subsidy_use_1, commute_subsidy_use_2,
## commute_subsidy_use_3, commute_subsidy_use_4, commute_subsidy_use_5,
## commute_subsidy_use_6, commute_subsidy_use_7, commute_subsidy_use_996,
## day_weight, deliver_elsewhere, deliver_food, deliver_grocery, deliver_none,
## deliver_office, deliver_other, deliver_package, deliver_work, diary_platform,
## disability_person, education, employment, ethnicity_1, ethnicity_2,
## ethnicity_3, ethnicity_4, ethnicity_997, ethnicity_999, ethnicity_other,
## ev_typical_charge_1, ev_typical_charge_2, ev_typical_charge_3,
## ev_typical_charge_4, ev_typical_charge_5, ev_typical_charge_6,
## ev_typical_charge_997, fuel, gender, hh_day_iscomplete, hh_is_complete,
## hh_weight, hhgroup, hhincome_detailed, hhincome_followup, hhsize,
## home_in_region, hours_work, industry, industry_other, is_participant,
## jobs_count, lat, license, loc_end, loc_start, lon, make, mobility_aides, model,
## model_other, no_school_closed, no_school_dont_know, no_school_no_answer,
## no_school_online_home, no_school_online_other, no_school_other, no_school_sick,
## no_school_vacation, notravel_delivery, notravel_housework, notravel_kidsbreak,
## notravel_kidshomeschool, notravel_madetrips, notravel_not_sure,
## notravel_notransport, notravel_other, notravel_sick, notravel_telecommute,
## notravel_vacation, notravel_weather, num_complete_fri, num_complete_mon,
## num_complete_sat, num_complete_sun, num_complete_thu,
## num_complete_trip_surveys, num_complete_tue, num_complete_wed,
## num_days_complete_weekday, num_days_complete_weekend, num_participants,
## num_students, num_surveyable, num_trips, numadults, numchildren,
## numdayscomplete, numworkers, office_available, participate, person_is_complete,
## person_weight, prev_home_notwa_city, prev_home_notwa_state,
## prev_home_notwa_zip, prev_home_wa, prev_rent_own, prev_res_factors_amenities,
## prev_res_factors_community_change, prev_res_factors_crime,
## prev_res_factors_employment, prev_res_factors_forced, prev_res_factors_hh_size,
## prev_res_factors_housing_cost, prev_res_factors_income_change,
## prev_res_factors_less_space, prev_res_factors_more_space,
## prev_res_factors_no_answer, prev_res_factors_other, prev_res_factors_quality,
```

```
## prev_res_factors_school, prev_res_factors_specify, prev_res_factors_telework,
## prev_res_type, proxy, proxy_complete, proxy_parent, race_afam, race_aiak,
## race_asian, race_hapi, race_noanswer, race_other, race_other_specify,
## race_white, race_category, relationship, remote_class_freq, rent_own,
## reported_lat, reported_lng, res_dur, res_months, res_type, sample_lat,
## sample_lng, sample_segment, school_freq, school_in_region, school_loc_lat,
## school_loc_lng, school_mode_typical, schooltype, second_home,
## second_home_in_region, second_home_lat, second_home_lon, sexuality, share_1,
## share_2, share_3, share_4, share_5, share_996, signup_platform,
## smartphone_type, speed, student, summary_complete, surveyable,
## telecommute_freq, telework_time, tnc_freq, toll_transponder, transit_freq,
## transit_pass, traveldate_end, traveldate_start, trips_yesno, vehicle_count,
## vehicleused, vehid, vehnum, walk_freq, work_in_region, work_lat, work_lng,
## work_mode, workplace, year
```

Visualize

```
common_modes<-mode_summary$summary$wtd%>%
  mutate(mode_w_sov= fct_reorder(mode_w_sov,-prop))

static_column_chart(common_modes, x='mode_w_sov', y='prop', fill='survey_year')
```
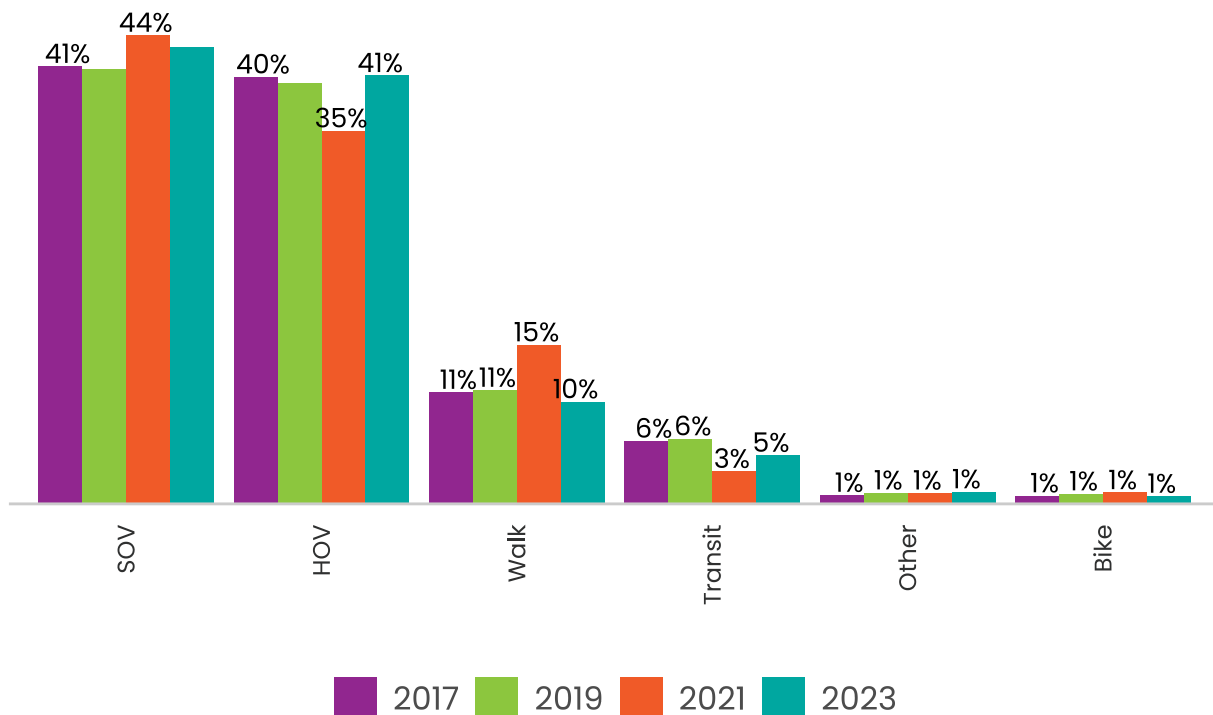

Trip rate overall

```
ids = c('hh_id', 'person_id', 'day_id', 'trip_id')
wts = c('hh_weight', 'person_weight', 'day_weight', 'trip_weight')
```

```
hts_data = list(hh = hh,
                person = person,
                day = day,
                trip = trip)
```

```
triprate_prep<-hts_prep_triprate(summarize_by = c('survey_year'),
                    variables_dt = variable_list,
                    hts_data = hts_data,
                    ids = ids,
                    wts = wts,
                    weighted = TRUE,
                    remove_outliers=FALSE)
```

```
## Warning in hts_validate_variable_list(variables_dt, hts_data): Variable(s) do
## not have a location specified: accuracy, bearing, collect_time, fuel,
## home_in_region, lat, lon, make, model, model_other, speed, toll_transponder,
## vehid, vehnum, year
```

```
output <- hts_summary(triprate_prep$num,
                    summarize_var = 'num_trips_wtd',
                    summarize_by = c('survey_year'),
                    summarize_vartype = 'numeric',
                    id_cols = ids,
                    wtname = 'day_weight')
```

```
wtd_triprate = output$summary$wtd[,
                            .(
                              summarize_by = c(get('survey_year')),
                              'Weighted Trip Rate' = round(mean, 2)
                            )
  ]
```

```
wtd_triprate
```

```
##    summarize_by Weighted Trip Rate
##          <char>             <num>
## 1:         2017              0.00
## 2:         2019              0.00
## 3:         2021              0.00
## 4:         2023              3.84
```

Visualize