

Assignment 1: Familiarize yourself with travel survey data

The purpose of this assignment is to familiarize yourself with household travel survey data. This site gives you access to a national household travel survey data: <https://nhts.ornl.gov/>

A typical household travel survey data comes in 4 files

1. Household file: Each row in the file corresponds to a household
2. Person file: Each row in the file represents a person of the household
3. Trip file: Each row represents a trip taken by the persons of the households
4. Vehicle file: Each row in the file represents a vehicle owned by the households

You can download the national level data file from the following location – Once you download you should get access to the 4 .csv files.

National Household Travel Survey

Conducted by the Federal Highway Administration, the NHTS is the authoritative source on the travel behavior of the American public. It is the only source of national data that allows one to analyze trends in personal and household travel. It includes daily non-commercial travel by all modes, including characteristics of the people traveling, their household, and their vehicles.

[Download 2022 V2.0 Data Now!](#) [Download/Explore National OD Data](#)

2022 Frequently Used National Statistics

Vehicles	Vehicle Trips	Vehicle Miles	Persons	Person Trips
Person Miles	Households	Workers	Drivers	

[2022 NHTS Summary Statistics](#) ▼

However, due to the large size of the data I have curated 4 small size files for this assignment. The smaller files have been uploaded in the Assignment folder in teams: [Assignment 01](#)

Logistics

- Your task for this assignment would be to **familiarize yourself with the type of data collected in typical household interview travel surveys and prepare a 1-to-2-page** (Font size: 10 to 12 point, single spacing) **report in Microsoft Word.**
- Please note that you can work in a group of 2 members for the assignment.
- You can divide the work among yourselves. *But the assignment should clearly identify which portion has been written and analyzed by whom. Without a clear statement about the member contribution the assignment will not be graded.*
- For analysis I highly recommend that you use some programming language and not rely on Excel, I will provide some help on R – but you will need to work on your own if you choose other languages such as Python or Matlab.

Submission

1. Please include the following information in the report
 - a. How many rows are there in the hh, person, vehicle and trip data file
 - b. Please calculate the average number of person per household
 - c. Please calculate the average number of trips taken by the persons of the household
 - d. Please calculate the average number of vehicles per household

Hint: you just need the row number information from the 4 files and do some simple arithmetic to complete part 1

2. Identify 3 to 4 variables from each of the 4 data files that you think would be useful/necessary for transportation planning or travel demand analysis

Hint: The data dictionaries are uploaded in the assignment folder. Identify a column name, go over the dictionary to figure out the use of the variable for transport planning/modeling

3. Report simple statistics on the 9 to 12 variables selected in question 2 above such as the mean, mode, median, frequency distribution, histograms and so on. Basically, I want an idea about the range and the distribution of your variable.

Hint: a cheat sheet has been provided to help you with R coding if you decide to use R coding for the assignment