

## Wells Fargo Campus Analytics Challenge: Live Green and Live Happy

**The motivation** of this project is to accelerate the transition to a low-carbon economy. Taking individual actions can encourage collective responsibility to help achieve environmental sustainability.

**The aim** of the project is to use machine learning to create a data product to help individuals optimize the balance between their carbon footprint and quality of life. Using the data set, creating a machine learning algorithm that minimizes carbon footprint for each customer while maintaining their total quality of life. This project wants to analyze data to recommend environmentally-friendly changes to everyday actions without lessening individuals' quality of life. The aim of the project to find quality substitutes for activities that are high carbon emitters without reducing the happiness and utility that the individuals in the data obtain from these activities.

**The data** gives a peak into the lives of 1,000 individuals who rated several everyday activities (taking a long shower, driving a car, etc.) on a scale of 1-100 based on how important those activities are to their daily lives.

**Data Cleaning:** at this stage we loaded data from the spreadsheet containing the data about the carbon footprints of each individual person from the given .xlsx data file given as part of the Wells Fargo Campus Analytics Challenge.

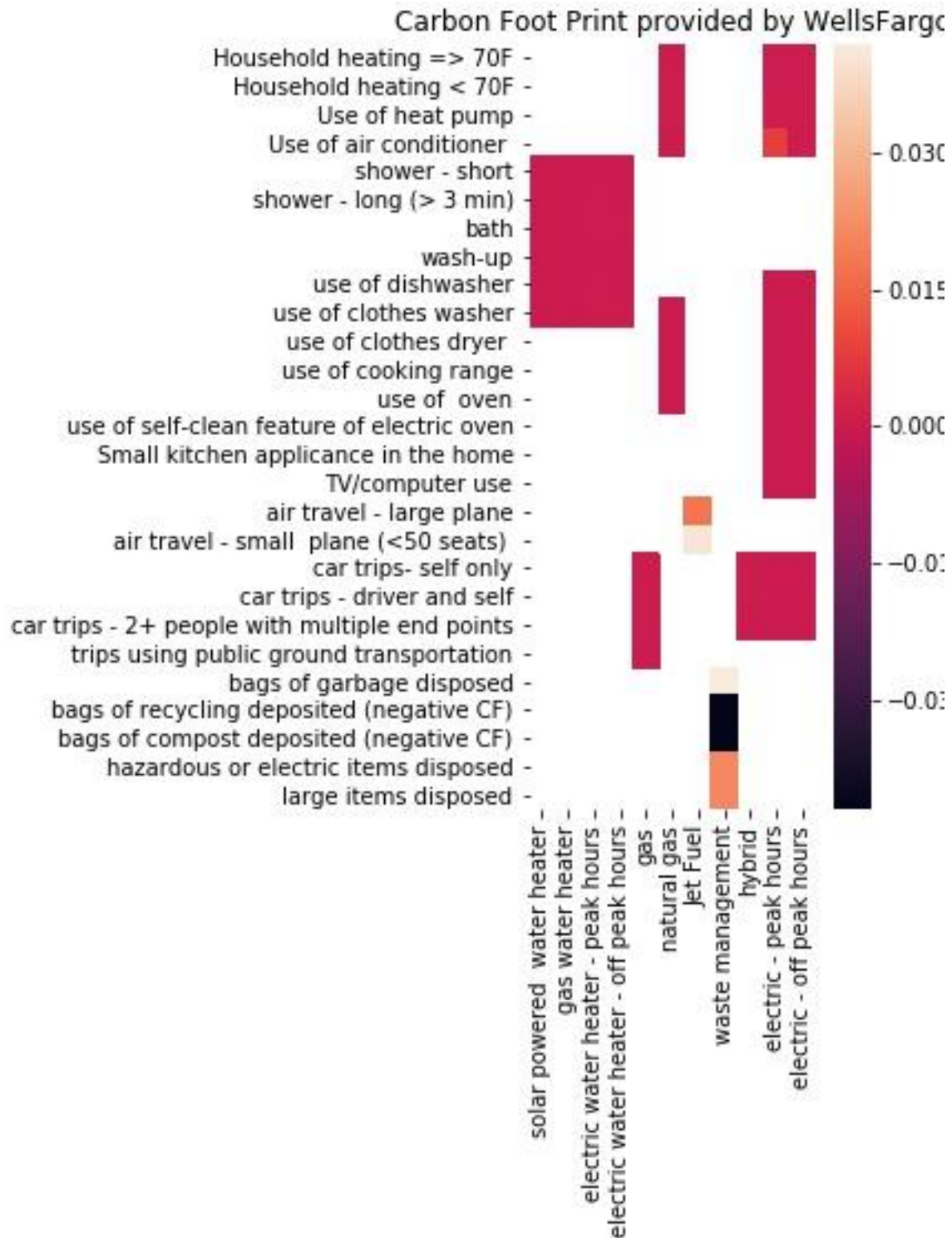
We converted the data from 2-dimensional table form which compares data of every individual v/s the type of resource used to a 1-dimensional table that lists out the data per individual. this process is like the reverse of pivoting a table.

We handled/filtered the various NaN values present in the Amount column of the table when the dataframe has been melted (unpivoted).

We saved the tables generated into csv files as part of saving data as we cleaned and filtered it.

**Data Wrangling:** at this stage we looked at the Carbon Footprint of each activity derived from the data and compared it to the Carbon Footprint provided by the challenge.

First, let's look at the Carbon Footprint provided by the challenge in a heatmap manner:



Then we calculated the cumulative Carbon Footprint of all users for all the activities to compare the ratio of Carbon Footprint of each activity with the one provided by the challenge. Here is the cumulative Carbon Footprint in a heatmap manner:

