

Dataset

The data in this project is divided into 4 files

Reference Data

Contains information that maps devices to a household.

- Device ID - A unique identifier of the device.
- DMA - The Nielsen defined designated market area.
- DMA Code - The DMA code as defined by Nielsen for each DMA.
- Household ID - A unique identifier for each household.
- Household type - A descriptor that maps each household to the company that generated the household id.
- System Type - The system type for the data collection application.

Daily Program Data

Contains information on programs.

- prog_code - Unique program identifier.
- prog_title - Program title of a movie, show, episode, or sports event.
- prog_genres - Comma separated words or groups of words that classifies a show, episode, movie, or sports event.
- air_date - Air date of program (Local) YYYYMMDD.
- air_time - Air time of program (Local) HHMMSS.
- duration - Duration of program in minutes.

Program Viewing Data

contains the information to map the tuning to a program.

- `mso_code` - Unique 5-digit code assigned by TMS to MSO.
- `device_id` - Globally unique device identifier.
- `event_date` - Date viewing event occurred (Local) YYYYMMDD.
- `event_time` - Time viewing event occurred (Local) HHMMSS.
- `station_num` - Station ID number
- `prog_code` - Unique program identifier.

Demographic Data

Contains a wide variety of household characteristics information

- `household_id` - The unique and persistent identifier for the household.
- `household_size` - Number of people residing in the household.
- `num_adults` - Number of adults residing in the household.
- `num_generations` - Number of Generations in Household.
- `adult_range` - A set of flags to indicate the age/genders residing in the household.
- `marital_status` - Code indicating marital status of the head of household.
- `race_code` - Code to indicate the race of the head of household.
- `presence_children` - Presence of Children in the household flag.
- `num_children` - Number of children residing in the household.
- `age_children` - Age of each child residing in the household with one year granularity.
- `age_range_children` - A set of flags to indicate the age/genders of the children residing in the household.
- `dwelling_type` - Dwelling type indicator code.
- `home_owner_status` - Home owner/renter code.
- `length_residence` - Length of time in years that the head of household has lived in the household.
- `home_market_value` - Value of the household.
- `num_vehicles` - Number of vehicles owned by an adult residing in the household.
- `vehicle_make` - Make code of the first (newest) vehicle in the household.
- `vehicle_model` - Model code of the first (newest) vehicle in the household.
- `vehicle_year` - Year of the first (newest) vehicle in the household.
- `net_worth` - Net worth of the household.
- `income` - Household income.
- `gender_individual` - Gender of the individual used to identify the household.

- `age_individual` - Age household member used to identify the household in 2 year increments.
- `education_highest` - Code that indicates the highest level of education.
- `occupation_highest` - Occupation that indicates the highest level in the household.
- `education_1` - Code that indicates the highest level of education of the head of household (1st individual).
- `occupation_1` - Occupation of the head of household.
- `age_2` - Age in two year increments of the second adult residing in the household.
- `education_2` - Code that indicates the highest level of education of the second adult in the household.
- `occupation_2` - Occupation of the second adult in the household.
- `age_3` - Age in two year increments of the third person residing in the household.
- `education_3` - Code that indicates the highest level of education of the 3rd individual in the household.
- `occupation_3` - Occupation of the 3rd Individual residing in the household.
- `age_4` - Age in two year increments of the 4th person in the household.
- `education_4` - Code that indicates the highest level of education of the 4th individual in the household.
- `occupation_4` - Occupation of the 4th Individual residing in the household.
- `age_5` - Age in two year increments of the 5th person in the household.
- `education_5` - Code that indicates the highest level of education of the 5th individual in the household.
- `occupation_5` - Occupation of the 5th Individual residing in the household.
- `polit_party_regist` - Political party indicator for the head of household from registration information.
- `polit_party_input` - Political party indicator for the head of household from individual input.
- `household_clusters` - Personix cluster codes.
- `insurance_groups` - Insurance classification via Personix.
- `financial_groups` - Financial classification via Personix.
- `green_living` - Green Living is an indicator of environmentally friendly households.