**Case-Study 19: Behavioral Risk Factor Surveillance System (BRFSS)**

***Overview***

This dataset includes data on adult's diet, physical activity, and weight status from the US Behavioral Risk Factor Surveillance System (BRFSS). This data is used for DNPAO's Data, Trends, and Maps database, which provides national and state specific data on obesity, nutrition, physical activity, and breastfeeding. This dataset combines variables including age, education level, gender, income and race/ethnicity with the BRFSS.

***Driving Challenges***

1. Is there a relationship between BRF and various explanatory variables (age, education, gender, income and race-ethnicity)? If so, what and how strong is the relationship?
2. Can we find any tendency in terms of time or regional distribution when controlling for some of the variables mentioned above？
3. Based on the analytical results derived above and the stratification information from the raw data, are there recommendations to improve peoples’ general nutrition or physical condition?
4. Can we use the same estimate model for stratification in different condition or variables, or what adjustment may be necessary?

***Meta-data***

|  |  |
| --- | --- |
| Elements | Corresponding columns in data |
| Years | YearStart, YearEnd |
| Region | LocationAbbr, LocationDesc, GeoLocation |
| Topics & related | Topic, Questions, Response |
| Data Info | DataValueUnit, DataValueType, DataValueFootnoteSymbol, DataValueFootnote |
| Data  | DataValue, DataValueAlt |
| Result & Reference | LowConfidentLimit, HighConfidentLimit, Strtification(category)n |
| Potential variables | Age(years), Education, Gender, Income, Race/Ethnicity |
| ID | Location, Topic, Question, DataValue, Stratification + ID |

This data is stored as categorical questions with response as indicators, like a tree model.

There is unique identifier for the dataset that may also be useful:

<https://chronicdata.cdc.gov/api/views/hn4x-zwk7>

***Data Source:***

<https://chronicdata.cdc.gov/api/views/g4ie-h725/rows.csv?accessType=DOWNLOAD>

***Provenance:***

The dataset is reorganized from its original data source from Behavioral Risk Factor Surveillance System. BRFSS is a state-based system of health surveys that collects information on health risk behaviors, preventive health practices, and health care access primarily related to chronic disease and injury. This database can be accessed here:

<https://catalog.data.gov/dataset/cdc-behavioral-risk-factor-surveillance-system-brfss>.

The BRFSS relating website may be useful: <https://www.cdc.gov/brfss/index.html>.