Big Data Problem Proposal:

Given the family income and expenditure data of the Philippines, what are the spending behaviors of different socioeconomic classes and what are the implications of these spending behavior? Results will be shown in different charts. Information to be found are the following:

- Sorting and counting families by socioeconomic class
- Average numerical values of each category per socioeconomic class
  (ex. Average of educational expenditure of all families)
- Average types of food expenditure per socioeconomic class
- Average meal cost per socioeconomic class
- Correlation of income and food expenditures per socioeconomic class
- Correlation of food types and medical expenses
- Correlation of income versus vices and necessities expenditures
- Total excess/deficit per socioeconomic class

Description of Data Source:

Data source is from a website called Kaggle, a website specializing in datasets and data analysis. By simply searching for Philippine Family Income and Expenditure, one can download the dataset in the links provided. The dataset is available in csv format. The procedures that the team did to obtain the dataset are as follows:

1. Go to Kaggle
2. Search for Filipino family income and expenditures
3. Download the dataset in csv format

The Philippine Family Income and Expenditure dataset contains 41,544 entries which includes data regarding a family's:

1. Total Monthly Income
2. Location/Region
3. Total Food Expenditure
4. Total Income and its Source
5. Food Expenditure Breakdown (Bread, Rice, Meat, Fish, Vegetables, etc.)
6. Other Expenditure (Alcohol, Medical, Education, Communication, Clothing, etc.)
7. Household Information (Family Head, Age, Number, Occupation, Facilities, etc.)

This dataset does not categorize each family entry to its own socioeconomic class. Because of this, the group used data sourced from Rappler where a table was shown of the various socioeconomic status and their characteristics such as the range of monthly income of each income class. Rappler calculated this data from a similar dataset the group used; the 2012 Family Income and Expenditure Survey (FIES), Philippine Statistics Authority.

Specific Link for the Dataset:

https://www.kaggle.com/grosvenpaul/family-income-and-expenditure/data

https://www.rappler.com/thought-leaders/98624-who-are-middle-class