Ryan Aclan

Principles of Data Science

April 3rd, 2022

Progress Report I

Filipino Family Income Project

What Has Been Done:

The data has been obtained and cleaned a bit. The handling of categorical values has only been considered.

- Obtained the data
- Partially cleaned the data
- Look at categorical values

What Has Not Been Done:

The data has not been completely cleaned and I have not yet committed to a way of handling the categorical variables.

- Completely clean the data
- Completely handle the categorical values

The rest of the project also remains, such as:

- Building the linear regression model
- Evaluating the performance of the linear regression model
- Consideration of other regression models
- Looking into principal components
  - Performing PCA
  - Rebuilding the regression model
  - Reevaluating the performance of the regression model

What Will Be Done in the Following Week:

The following week will be used to catch up on this week's work as well as complete next week's agenda.

Tasks from last week to do:

- Finish cleaning the data
- Handle the categorical values

This week's planned tasks:

- Build the linear regression model
- Evaluate the performance of the linear regression model
- Consider other regression models