

Final Project

Choose an interesting dataset from Kaggle or Open Government Data of Thailand (data.go.th) to analyze

1. Use one or more techniques to find interesting insights or create model from chosen dataset:

- a. Data Exploration/ Analysis
- b. Clustering
- c. Classification

Use Python programming and useful libraries (such as Pandas, Matplotlib, Sci-kit Learn, etc.) to develop your project

A result of this project could be the form of graphs, dashboard, model, or others which an interesting insight from the data can be explained.

2. Apply some data preparation technique or other techniques to improve performance/correction of analytics such as:

- a. Data Preparation (such as replacing missing value, removing noise/abnormal value)
- b. Feature Selection
- c. Correlation analysis
- d. Other useful techniques

3. A project presentation slide must include an answers for these questions:

- a. Explain the steps of exploring, analyzing, or creating models from the dataset.
- b. How could this project improve the business? Who benefits from this project?
- c. How to evaluate the performance of the result or the created model?
- d. Are there any solutions to improve this project (future works)?

Some useful examples:

<https://medium.com/softnix/วิเคราะห์ข้อมูล-ผู้บาดเจ็บและเสียชีวิต-ในช่วงเทศกาล-ปี-51-58-b989f9c44a1a>

<https://towardsdatascience.com/suicide-prevention-insights-with-data-science-2fb2636a6d82>

<https://towardsdatascience.com/mall-customers-segmentation-using-machine-learning-274ddf5575d5>