ISM 4641: Python for Business Analytics

Group Project: Applying Python for Business Analytics (30%)

Instructions

The group project will involve an analysis of a dataset that you and your peer consider interesting. The assignment requires applying data wrangling, visualization, analytical, and presentation skills. The project may involve a modest amount of research on available datasets and experimentation with the data. Examples of publicly available datasets can be found at data.gov (more than 234,134 datasets available).

Your final presentation must have (but is not limited to) the following issues:

- Introduction: Business problem (objective of the analysis)
- Data (source, size, main variables, time frame, etc.)
    - Data wrangling
    - Dealing with missing data/outliers/ unnecessary or unwanted chunks of data
- Preliminary Analysis
    - Summary statistics
    - Correlations
- Visual inspection of the data
    - Examples:
        - Yearly trends (for panel data)
        - Comparative analysis (for cross-sectional data)
        - Distribution of the main variables (e.g., histograms)
        - X-Y plots to show the relationship between variables
- Main analysis (e.g., Regression, Clustering, Choice Model)
- Conclusions

Each group **MUST** have two members and deliver:

- 30-minute presentation (25 min presentation + 5 min Q&A)
- A WELL ORGANIZED (with headers, markdowns, and comments) Jupyter Notebook with the codes used in the project.

If you haven't added your group participants to the list that circulated in class, send me an email with the group participants by Monday (Oct-2) at 5:00 pm. If you don't choose your group partner, one will be randomly assigned to you.

The date that each group will present and the order of the presentation will be randomly picked.