ANA-522-OL1 Spring 2022
Mod03 week06 HW: Data Munging
Due: Sunday February 20th at midnight

Explore the Dataset of your interests - Continued

The purpose of the data wrangling and munging homework is to utilize what we have learned in pandas to explore data in limited but practical ways. It is the expectation that the same dataset adopted from the Data Prep HW of week 05 is used for this assignment.

From last HW, you had found the dataset of the topic of your interest, understood its data fields, and applied data cleaning and preparation techniques. This week, the work continues to apply data wrangling techniques learned from text reading (Chapter 8), to join, combine, and reshape data with the emphasis on using hierarchical indexing. The deliverable items to be submitted including:

1. (25%) Expand the Abstract from last homework, and add at least two additional problem statements that the dataset could be analyzed to answer them preliminarily, if not fully.
    - Revise the title, if needed, and add the data munging process to the work.
    - Provide metadata of the dataset.
    - A copy of the dataset file is to be submitted alongside with the Jupyter Notebook report.
2. (25%) Create different multiple index levels on rows and/or columns of the dataset, so that they can be used to suggest answers to the problem statements proposed above. Please create at least two different ones to gain insight from various perspectives and make summaries.
    - See Data Munging Lab for examples.
3. (25%) Data Wrangling Playground.
    - Apply any technique covered in the text reading of this week to seek additional support to reorganize the data in shapes and categorical structures, seek benefits in analyzing the data context, and draw conclusions using the the auxiliary tools.
    - Provide informative descriptions and comments to these opertions performaned in the context of the data topics.
4. (25%) Summary and Conclusion
    - Summarize the findings to the problem statements and give conclusions, after applying data munging steps above.

Turn in your Jupyter Notebook file (.ipynb) along with the dataset file to upload on the Blackboard for submission.