

General Info: BUAN/MKT 6337 Group Presentation

General information

- Please see the first week's lecture slide on the group project.
- Always think about the audience. Will answering this question help a manager/executive in a particular function of a particular industry? Will answering this question help policymakers or government agencies in any particular way? Will answering this question change the way people think about specific business processes?
- I understand that most data sources (e.g., Kaggle) have related notebooks with codes that generate fancy data plots or run various analyses for different reasons. If you simply copy these analyses without thinking about the questions above (believe me, many students have done just that in previous years), you will not do well in your presentation.

Tool

- You can use any software tools or program language for the group project

Deliverables

- Project proposal and the final presentation (see the syllabus and the first lecture for details.)

Data sources

- If you have access to data through individual connections, employers, scraping websites, or cold calls, be my guest. You will get extra points for unique data sets. If not, see the below public data sources for potentially interesting data related to your question.
- For your project, state where your data come from, and that you have the right to use the data for class presentations.
- Public data sources
 - Variety of data repositories
 - KDnuggets ([link](#))
 - Kaggle datasets ([link](#))
 - Amazon Data Exchange ([link](#)). Not all of them are free.
 - Stanford Large Network Datasets ([link](#))
 - data.world ([link](#))
 - Google toolbox ([link](#))
 - Social science-focused data

- National Bureau of Economic Research data ([link](#))
- ICPSR ([link](#))
- Marketing-related data
 - Yelp datasets ([link](#))
 - Airbnb datasets ([link](#))
 - Safegraph foot traffic data ([link](#))
- Health care data
 - CDC website ([link](#))
 - Tobacco and health study by the FDA ([link](#))
 - Healthdata.gov ([link](#))
 - Dartmouth Atlas: Data on variation of costs and outcomes ([link](#))
 - Open payment (pharmaceutical industry payment to doctors) ([link](#))
 - Doctors' prescription data for Medicare patients ([link](#))

Projects you are NOT allowed to choose

- These projects/data sets have been used in previous years many times. You will get ZERO points if you choose any of the projects listed below:
 - IBM HR Analytics Employee Attrition & Performance
(<https://www.kaggle.com/datasets/pavansubhasht/ibm-hr-analytics-attrition-dataset>)
 - Campus Analytics Challenge 2021
(<https://www.mindsumo.com/contests/campus-analytics-challenge-2021>)
 - Stroke Prediction
(<https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset>)