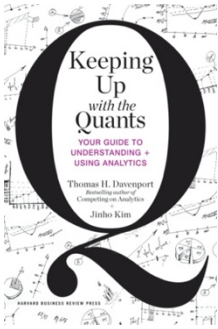

Keeping Up with the Quants

Framing the Problem:

1. **Problem recognition**:
   - What is the Business problem that you can analyze from this dataset? Why is it relevant?
2. **Review of Previous findings**:
   - What does your research guide you into? Are there key insights that you found from your research about the Busines problem? – This will be the area where, as a team, you would look into Business articles (WSJ / Economist / Financial times) to highlight about the business problem that you are trying to explore.

- What is the Testable Hypothesis / Thought process that you established based on your initial research? Your analysis can be predictive or inferable. If your analysis is predictive, there would not be a hypothesis, instead it would report model performance.

Solving the Problem:

1. **Variable Selection**: Introduce your Data using key attributes. What is the data about?
2. **Data collection:** What are the data sources that you collected?
3. **Data Analysis**: Summarization and Visualization (5-7 charts / analyses)
   - What are the key trends and patterns that you find about the data? Each trend /chart should have 3-4 lines about why is that trend /chart important. How does it add value to your Data Analysis project?
   - Are there Outliers in your data? What charts/visualization did you use to identify them? How did you handle your Outliers?
4. What are the updates/ modifications that you did to your initial hypothesis/ thought process after Summarization and Visualization?

Modelling and Communication:

1. **Modelling**: (OLS and / Logistic) to identify relations /connections in the data
2. **Results presentation**:
   - Validate your Hypothesis / thought process. What are your inferences / model performance?
   - Preparing your R markdown for presentation
   - What are your 3# specific insights for the data analysis? Connect your data analysis from Stage 1 and Modelling from Stage 3 to support your findings. It is also expected that you use with domain knowledge (i.e. research from external sources). Make sure to site your sources.

Your Deliverables

- **Checkin# 1:** R markdown file with the below 2# sections and sub-sections for each of the areas. You may choose to have additional sub-sections as relevant to your analysis. You will also submit the datasets.
  - Framing the Problem
  - Solving the Problem
- **Final Submission:** R markdown file with all the 3# sections and sub-sections. You may choose to have additional sub-sections as relevant to your analysis. You will also submit the datasets.
  - Framing the Problem: You can revise this section based on the feedback received for Checkin#1.
  - Solving the Problem: You can revise this section based on the feedback received for Checkin#1.
  - Modelling and Communication