Project #9 and 10

Portfolio Analytics

General Requirements

This project will use recent market data to support a decision of your choosing. Examples could include investing in a new market, managing working capital, allocating expenditures, revenue, and investments, capital budgeting, and risk management. Analysis would include regression, simulation, and optimization. Visualization of results will reside in an online interactive financial application.

Set A

1. Select a decision. Develop business questions that, once answered, would support the decision. For each question posed construct a set of work steps to include data collection, hypothesis formation, analysis, and visualization.
2. Collect data to support the decision that answers the posed business questions. Clean the data. Organize the data into a data frame(s). Explore the data using univariate and multivariate techniques. Interpret findings.
3. Construct a dashboard with interactive elements and visuals to assist the decision maker explore the data.

Set B

4. Using the results of the exploratory data analysis, generate hypotheses about the data, and methods to help formulate, measure, and test these hypotheses.
5. Using simulation, regression, and optimization techniques, as needed, generate results that will help answer the business questions to support the selected decision. Document all code and approaches used.
6. Develop tables and plots to help summarize the results of the analysis. Situate the tables and plots in the dashboard application.

Set C

7. Interpret and present results. Produce an `R Markdown` document with code chunks to document and interpret our results.

- The first section of the document will summarize observations and recommendations driven by the analysis, disclose assumptions, limitations, and issues within the analysis.
- Subsequent sections that discuss background and context for decisions, the data to be used, and the work flow we have defined.

8. Complete the dashboard application and present the work products to the decision makers.

Rubric

General

Set A is due *24 hours prior to the Week 9 Live Session.* The framework for Set B is due *24 hours prior to the Week 10 Live Session.* The remainder of Set B and all of Set C is due by end of the course term.

You will only receive credit for this assignment if you attempt to answer all questions and address all sections with substantive answers, relevant code, and graphics as needed.

Specific

Grades for assignments will follow this rubric:

- **Words:** The text is laid out cleanly, with clear divisions and transitions between sections and subsections. The writing itself is well-organized, free of grammatical and other mechanical errors, divided into complete sentences, logically grouped into paragraphs and sections, and easy to follow from the presumed level of knowledge.
- **Numbers:** All numerical results or summaries are reported to suitable precision, and with appropriate measures of uncertainty attached when applicable.
- **Pictures:** All figures and tables shown are relevant to the argument for ultimate conclusions. Figures and tables are easy to read, with informative captions, titles, axis labels, and legends, and are placed near the relevant pieces of text.
- **Code:** The code is formatted and organized so that it is easy for others to read and understand. It is indented, commented, and uses meaningful names. It only includes computations which are actually needed to answer the analytical questions, and avoids redundancy. Code borrowed from the notes, from books, or from resources found online is explicitly acknowledged and sourced in the comments. Functions or procedures not directly taken from the notes have accompanying tests which check whether the code does what it is supposed to. All code runs, and the `R Markdown` file `knits` to `pdf_document` output, or other output agreed with the instructor.
- **Modeling:** Model specifications are described clearly and in appropriate detail. There are clear explanations of how estimating the model helps to answer the analytical questions, and rationales for all modeling choices. If multiple models are compared, they are all clearly described, along with the rationale for considering multiple models, and the reasons for selecting one model over another, or for using multiple models simultaneously.
- **Inference:** The actual estimation and simulation of model parameters or estimated functions is technically correct. All calculations based on estimates are clearly explained, and also technically correct. All estimates or derived quantities are accompanied with appropriate measures of uncertainty.
- **Conclusions:** The substantive, analytical questions are all answered as precisely as the data and the model allow. The chain of reasoning from estimation results about the model, or derived quantities, to substantive conclusions is both clear and convincing. Contingent answers (for example, "if X, then Y, but if A, then B, else C") are likewise described as warranted by the model and data. If uncertainties in the data and model mean the answers to some questions must be imprecise, this too is reflected in the conclusions.
- **Sources:** All sources used, whether in conversation, print, online, or otherwise, are listed and acknowledged where they used in code, words, pictures, and any other components of the analysis.