**Project Proposal and Final Paper Instructions**

**EMSE 6992 Data Science Introduction and Practicum**

**Fall 2018**

## **Proposal Paper** Due: 10/4

No Format (includes Abstract, Introduction - Overview, Motivation, Related Work, Initial Questions, Data, and the Business Case Evaluation portion of the Methodology):

**Abstract**

**Introduction**

- **Overview, and Motivation**: Provide an overview of the project goals and the motivation for it. Consider that this will be read by people who did not see your project proposal.
- **Related Work and Audience**: Anything that inspired you, such as a paper, a web site, or something we discussed in class. Who is your audience? Who will you show your results?
- **Initial Questions**: What questions are you trying to answer? How did these questions evolve over the course of the project? What new questions did you consider in the course of your analysis?
- **Data**: Source, scraping method, cleanup, storage, etc.

**Methodology** (Explain the steps in the process to answer your questions.)

First. Business Case Evaluation

**Goals** (Questions), **Processes** (Analysis - Python Libraries/functions and Visualizations), **Materials** (Data), **Technology** (Processing/Storage/etc.)

---

## **Final Paper** Due: 12/7

**Abstract**

**Introduction**

**Methodology** (Explain the steps in the process to answer your questions.)

In IEEE format (also includes Abstract, Introduction, and first step of Business Case Evaluation):

| | |
|---|---|
| Second. | Data Identification |
| Third. | Data Acquisition & Filtering |
| Fourth. | Data Extraction |
| Fifth. | Data Validation & Cleansing |
| Sixth. | Data Aggregation & Representation |
| Seventh. | Data Analysis |
| Eighth. | Data Visualization |
| Ninth. | Utilization of Analysis Results |

**Analysis** (Explain what you actually did with the data?)

- **Exploratory Data Analysis**: What visualizations did you use to look at your data in different ways? What are the different statistical methods you considered? Justify the decisions you made, and show any major changes to your ideas. How did you reach these conclusions?
- **Final Analysis**: What did you learn about the data? How did you answer the questions? How can you justify your answers?

**Results** (Explain what your results were able to show?)

**Results**: Present your final results in a compelling and engaging way using text, visualizations, images, and videos on the paper and on your project web site.

**Conclusion**

**References**


Datasets

https://www.analyticsvidhya.com/blog/2016/11/25-websites-to-find-datasets-for-data-science-projects/

https://www.dataquest.io/blog/free-datasets-for-projects/

https://www.analyticsvidhya.com/blog/2016/10/17-ultimate-data-science-projects-to-boost-your-knowledge-and-skills/

https://www.springboard.com/blog/free-public-data-sets-data-science-project/

Libraries and Tutorials

https://github.com/rasbt/pattern_classification/blob/master/resources/python_data_libraries.md

https://medium.com/activewizards-machine-learning-company/top-15-python-libraries-for-data-science-in-in-2017-ab61b4f9b4a7

https://www.upwork.com/hiring/data/15-python-libraries-data-science/