DRAFT 1 Capstone proposal

**Client/project owner:** Lou Rosenfeld, Rosenfeld Media

**Secondary client:** Abby Covert, Etsy

**Codeup contact:** Sean Oslin

**Summary**

Analyze quantitative and qualitative survey data for a prospective new multiday training/conference on research. The project would replicate analysis already done by other parties but using data science rather than manual analysis techniques.

**Background**

Rosenfeld Media (RM) conducted an online survey as part of the planning process for a new conference on research. RM expected 80 responses. They received more than 800 partial and complete responses.

Analysis of the data was conducted in 2 ways. First, the purely quantitative data was cleaned in Excel and analyzed/presented in Tableau.

After cleaning the qualitative data, a statistician analyzed the data to determine the best way to segment the data. The qualitative data were hand-reviewed/tabulated by a small team with experience in the field of user experience. This analysis was used to create user personas that helped define potential audiences for the conference. Analysis of the quantitative data was then integrated into the personas.

**End products**

* A predictive classification model where the labeled data comes from the qualitative data.
* Comparison of the modeled personas with the previously constructed personas
* Presentation of results to client as proof of concept of data science techniques applied to survey data

**Data**

The survey contained more than 50 qualitative, quantitative and mixed questions. More than 800 people completed the survey. The data is stored in an Excel spreadsheet. Substantial cleaning needs to be done to the data, including deciding what to do with partial responses.

Challenges with the data include cleaning the data, segmenting the responses, creating personas and integrating quantitative and qualitative data.

**Analysis tools**

Analysis/modeling will use Python. Data wrangling will employ MySQL and Python. Data visualization will use Tableau or other software.

**Permissions/restrictions**

Permission has been given to analyze the data and publish the results.