Capstone Project Proposal

Jake Schroeder

**1. What is the problem you want to solve?**

The main problem this project aims to solve is *identifying shifts in digital marketing keyword strategy based on demographic data*.

Changing demographics have an impact on near every other aspect of marketing - how will it impact digital marketing at its most foundational level? By exploring the relationship between the datasets used, we can project future changes and align digital strategy accordingly.

**2a. Who is your client and why do they care about this problem?**

My client is anyone using search engine optimization to increase ROI for his or her business. Keyword research is a critical and fundamental facet of search engine optimization. This solution will directly impact keyword research by showing how keyword trends follow the trends of demographics.

**2b. In other words, what will your client DO or DECIDE based on your analysis that they wouldn’t have otherwise.**

By using this solution alongside a more traditional keyword research tool, the client will have the information they need to build a keyword strategy that will last the test of time. If the client did not have this information, they would be “flying blind” with regard to how changing demographics will impact their keyword strategy in the future.

**3. What data are you going to use for this? How will you acquire this data?**

The fundamental data required for this analysis will be:

- Historical keyword data from Google Trends

- Annual demographic projections / recordings from the US and CA databases

I will acquire the Google Trends data from the pytrends library. The demographic information will be acquired through government websites in a .csv format.

**4. In brief, outline your approach to solving this problem (knowing that this might change later).**

My approach will be to first build a “straw man” analysis of the relationship between changing demographics and a broad set of service-related keywords. Based on the findings/performance of this first iteration, I will then experiment with more granular / localized data to see how far the relationship between the two variables can be quantified.

In more detail, I will use “topic modeling” to measure the relationships between demographic data and keyword trends.

**5. What are your deliverables? Typically, this would include code, along with a paper and/or a slide deck.**

The deliverables for this project will be:

- **Code:** available on GitHub

- **Paper:** analyzing the process, results and insights generated

- **Deck:** displaying the analysis and insights with graphical representations