***Final Project, Part 1: Proposal***

**Problem Statement:** How can you maximize revenue as an app developer in any given category (games, health, books, finance, etc.)?

**Hypothesis/Assumptions**: Free and ad-based models generate the most revenue for apps in most categories, while gaming app can generate the most revenue by charging a fixed price upfront. I assume that there is an optimal business model for each category of apps.

**Goals/ Success Metrics:** I aim to define a revenue-maximizing model on two variables: number of active users and category. These two variables will be used to determine the best practice for maximizing revenue, whether it is charging for the app upfront or allowing free downloads then supporting the operations with advertising sales. For the apps driven by advertising sales, I would like to conduct a regression analysis to see how much ad revenue can be generated for a given number of users. For apps that are best run by upfront payment, I would like to create an elasticity equation to determine the optimal price.

**Risks/Limitations:** A major risk in this project lies in the intangibles that cannot be measured, but those that would affect consumers’ willingness to spend on an app. We also don’t have the data of the failed apps based on the way data is collected. This create a strong selection bias. Lastly, the hypothesis that free and ad-based models would make the most money for most apps may run into the issue of multicollinearity, as well as reverse causality. The apps with a large user base are able to attract advertisers. Advertising is may not be the key driver of revenue generation; it may be the large user base that drove revenue growth.

**Relevant Dataset:** I have access to SensorTower, which provides data on app sales and their downloads, ad-support status and number of active users in a time series by category of apps and by the store type of top-ranked apps on Android and Apple app stores. I am granted access through work.