Idea#2

College basketball prediction

https://www.kaggle.com/datasets/andrewsundberg/college-basketball-dataset?select=cbb23.csv
- Main datasets

https://www.sportico.com/business/commerce/2023/college-sports-finances-database-intercollegiate-1234646029/
- Revenue/cost data

Data tractability:

Explore which variables are associated with a team making the NCAA tournament. Potentially bring in revenue/operating costs from outside source.

Data retrieval:

I downloaded 20-21, 21-22l and 22-23 Men's College Basketball datasets off of Kaggle. I have operating costs and revenue for different schools from Sportico.

Exploratory data analysis:

There are ~350 rows per dataset. Variables include Team, Conference, Wins, In-game performance statistics, and more. I concatenated the datasets so that there are 1068 rows and 23 columns.

Implications for stakeholders:

College athletics generate large amounts of money for their respective universities. Those working in college athletics will be interested to see which factors are associated with making playoffs.

Ethical, legal, societal implications:

This type of data will change how college athletics programs operate and cause teams to focus on key variable that will drive their chances of making the NCAA tournament forwards.

Idea#3

British Airways Reviews

https://www.kaggle.com/datasets/girishchowdary22/british-airways-review-dataset2012-2023

Data tractability:

I want to explore which variables contribute to the numerical (1-10) rating left on British Airways flights by passengers.

Data retrieval:

I got the dataset from Kaggle. It is one dataset and does not require any merging with additional tables.

Exploratory data analysis:

There are 1300 unique reviews in the dataset. There are 19 columns, ranging from author, date, and context to overall rating, seat comfort, cabin staff service rating, and more. Entertainment appears to be right skewed as well as value for money.

Implications for stakeholders:

British Airway investors and executives will be interested to see how their airline is regarded and what people think about the service they have provided for the last 10 years.

Ethical, legal, societal implications:

Ethically, results from modeling this data can impact British Airways business and profitability. That is not something to be taken lightly and can have legal implications as well.