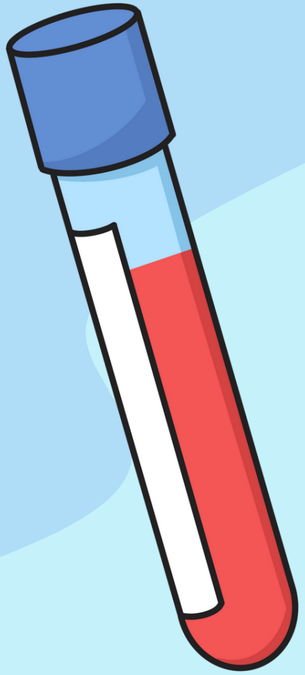
Capstone Project

Predicting Clinical Trial Terminations

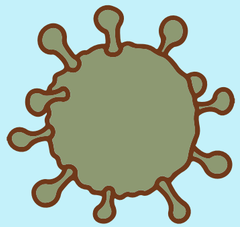
Clement Chan

March 8th 2024



What are Clinical Trials?

- Research studies involving human participants
- Evaluating safety, performance, and effects of drugs
- Ex. Covid Vaccines, diabetic medications, hypertension.



Problem Area

- Out of 8000 trials, 960 trials (12%) will be terminated [1]
- Average costs of clinical trials range from 1.4 million - 52.9 million USD in 2015, and is increasing at a rapid rate. [2]
- Significant resources are dedicated towards the design, analysis, and conclusive determination of clinical trials.

Factors Associated with Trial Terminations

- Terminations are often caused by insufficient patient enrollment
- Lack of funding
- Industry sponsorship cannot guarantee that a clinical trial will not be terminated due to poor business or commercial decisions.

However, it is difficult to pinpoint what factors or combination of factors lead to clinical trial terminations.

Data Science Solution

Goal:

1. Determine the main factors or parameters associated to terminated trials

2. Predict clinical trial terminations based on those factors.

Data Science Solution

Data Wrangling
+
EDA

Statistical Data Analysis

Predictive Modeling

NLP

+

Word Embedding

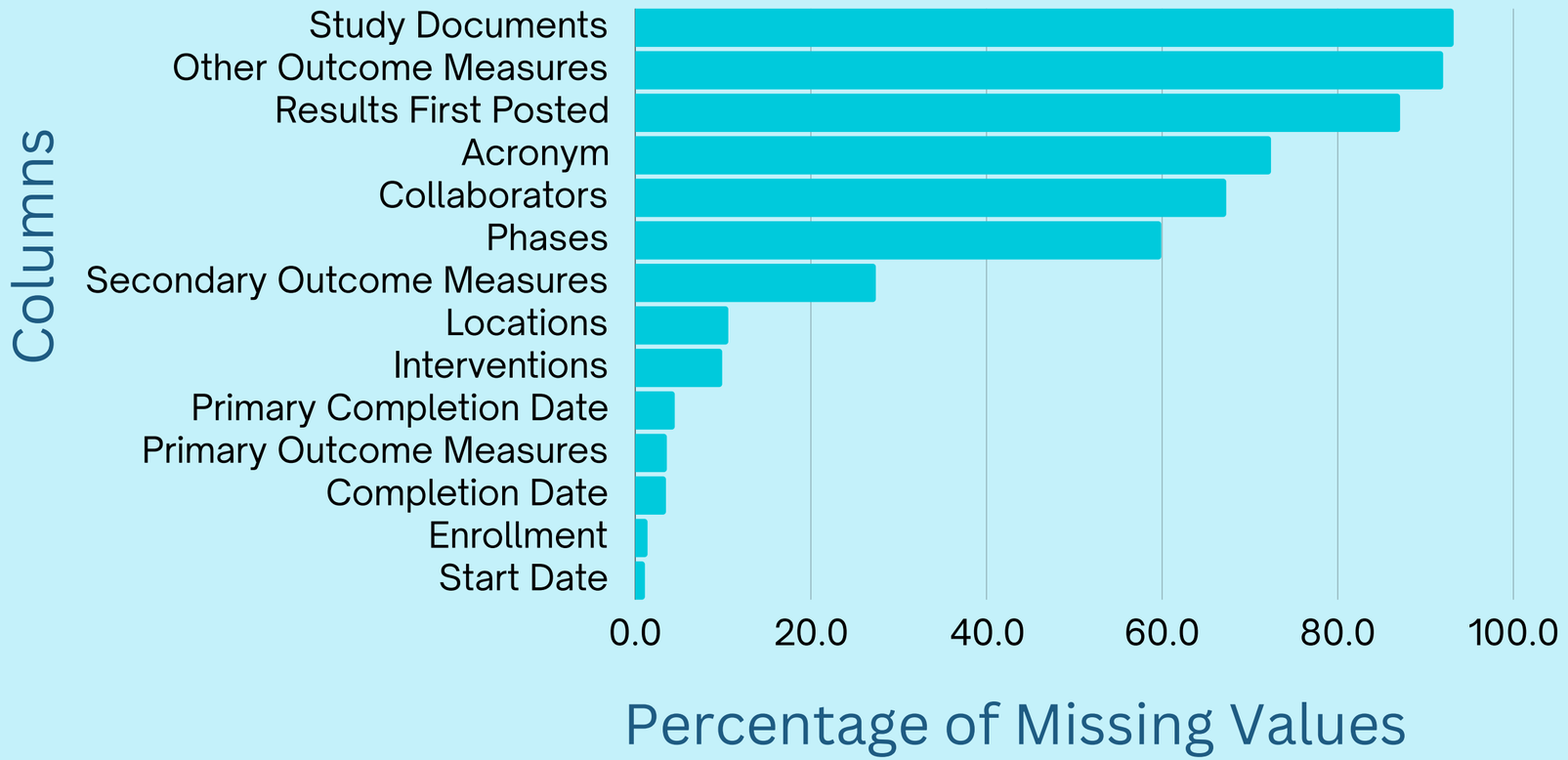
Results

1. Save millions of dollars in funding by sponsors, investors, and government tax money.

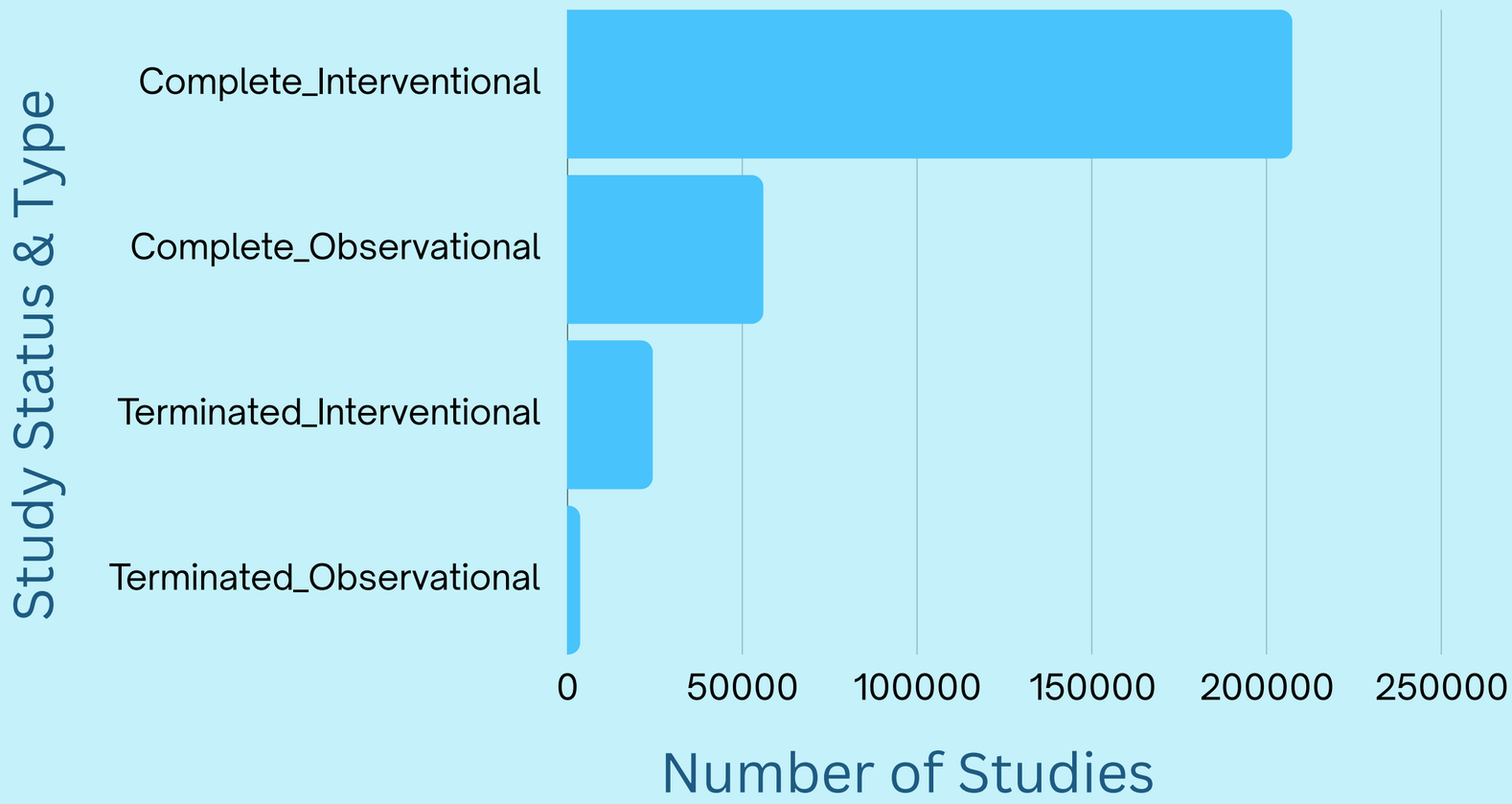
2. Prevent loss of scientific advancements in science

Missing Values in the Dataset

- 483,238 rows and 30 columns of unclean data



Comparing Study Status & Study Type of Clinical Trials



Next Steps?

1. There is still a lot of cleaning to do within the dataset since most of the columns are sentences.
2. We can use NLP and Word Embedding to classify words and group them into smaller categories
3. Try out different classification models: logistic regression, KNN, and decision trees to classify factors that cause clinical trial terminations.

References

1. https://pubmed.ncbi.nlm.nih.gov/26011295/#:~:text=The%20main%20outcomes%20were%20to,905%2F7%2C646)%20were%20terminated.
2. https://journals.sagepub.com/doi/10.1177/1740774515625964
3. https://www.nature.com/articles/s41598-021-82840-x
4.