Project Proposal

1. What is the problem you want to solve?

· Determine the relationship between various demographic factors and hospital readmission rates.

2. Who is your client and why do they care about this problem? What will your client do or decide based on this analysis that they wouldn't have done otherwise?

· **Client:** Any hospital or integrated delivery network across the United States

· **Why do they care about this problem?**

Diabetes is a metabolic disorder in which the body is incapable of producing any or enough of the hormone insulin to absorb blood glucose efficiently.

Though not always preventable, readmission can be considered a marker of negligent care and is a waste of both hospital resources and spending. Patients with diabetes are known to have a high risk of 30-day readmission following initial hospitalization for hyperglycemia (American Diabetes Association 2019). To avoid readmission, hospitals must work to implement a structured protocol for the general admission and discharge of said patients, with focus on treating hyperglycemia and avoiding hypoglycemia (low blood sugar).

· **What will your client do or decide based on this analysis?**

In conjunction with a cost-benefit analysis, this analysis could help to determine which decision is more costly - for a patient to receive inpatient services for longer than necessary in an effort to prevent readmission, or to allow an at-risk patient to be discharged, with the possibility of readmission.

3. What data are you using? How will you acquire the data?

https://archive.ics.uci.edu/ml/datasets/Diabetes+130-US+hospitals+for+years+1999-2008

4. Briefly outline how you'll solve this problem.

- Clean data and wrangle for missing values/outliers
- Visual and statistical EDA to determine which features are best at classifying readmission
- Logistic regression model

5. What are your deliverables?

Paper + slide deck + Jupyter Notebook