Title: Predicting Hospital Readmissions Within 30 Days Using Patient and Clinical Data.

Overview: This project aims to develop a machine learning model that predicts the likelihood of a patient being readmitted to the hospital within 30 days after discharge. Unplanned readmissions are not only costly but also reflect the quality of patient care. By identifying high-risk patients using factors such as age, primary diagnosis, length of stay, discharge disposition, and comorbidities, hospitals can implement targeted interventions to reduce readmission rates.

Significance: High readmission rates can indicate poor care coordination and inadequate follow-up procedures. Addressing this issue can help hospitals meet regulatory standards, avoid penalties, and, most importantly, improve patient outcomes. The insights generated from this project could guide policy decisions and resource allocation for hospitals.

Data Source: I will use the publicly available "Diabetes 130-US hospitals for years 1999–2008" dataset from the UCI Machine Learning Repository which includes detailed patient visit records such as demographics, diagnoses, lab results, procedures, and discharge outcomes. The link for the data source:
https://archive.ics.uci.edu/dataset/296/diabetes+130-us+hospitals+for+years+1999-2008

The steps that I will be following to complete this project are below:

1. Data acquisition and preprocessing (Data Wrangling).
2. Exploratory data analysis to identify patterns and correlations
3. Feature engineering to enhance model performance
4. Model training using classification algorithms (e.g., logistic regression, random forest)
5. Model evaluation
6. Interpretation and visualization of key findings.

Deliverables:

- A GitHub repository containing:
  - Complete code and documentation
- A slide deck summarizing project objectives, approach, findings, and recommendations
- A comprehensive project report detailing methodology, data processing, model performance, and conclusions

Tools & Technologies:

- Python (Pandas, Scikit-learn, Matplotlib, Seaborn, data visualization tools, machine learning tools)

- Jupyter Notebooks
- Git and GitHub for version control
- Google Slides or PowerPoint for presentation

Conclusion: This project addresses a critical healthcare issue by leveraging machine learning to provide early warnings for patient readmissions. With clear deliverables and a practical application, it holds potential value for hospitals aiming to improve care and reduce unnecessary costs.