**Project Proposal**

**Analyzing Medical Procedure Costs in New York**

**Introduction:**

Our project aims to conduct advanced analysis of medical procedures to optimize healthcare costs predictively in New York. Leveraging various datasets related to providers such as doctors and clinicians, healthcare facilities, medical procedures, quality scores, and physician visit costs across different branches of medicine by zip codes. We will develop predictive models to identify cost-effective procedures.

**Motivation:**

In the face of rising healthcare costs, it's crucial to analyse the cost-effectiveness of medical procedures proactively. By predicting procedure costs, we can help healthcare organizations make informed decisions to improve efficiency while minimizing expenses.

**Evaluation:**

The successful outcome of our project would be the development of accurate predictive models capable of forecasting the cost-effectiveness of medical procedures. We will measure success by evaluating the predictive accuracy and utility of our models in optimizing resource allocation and improving cost-effectiveness. Metrics such as accuracy, precision, and recall will guide our evaluation, while the methodology for calculating cost savings will be determined as the project progresses.

**Resources:**

We will utilize several datasets for our analysis:

* Providers: Information about individual healthcare providers, their specialties, practice locations, and affiliations.
* Facilities: Details about healthcare facilities, including type, certifications, and affiliations with providers.
* Procedures: Information about medical procedures performed by providers, including counts and cost data.
* Quality Scores: Quality scores for healthcare providers and facilities, possibly related to performance in various categories.
* Physician Office Visit Costs: Information about physician visit costs across different branches of medicine by zip codes.

The project will utilize the following methods and tools for execution:

* Data Collection
* Data Cleaning and Transformation (Excel, SQL, Alteryx)
* Exploratory Data Analysis (Python)
* Data Visualization (Tableau)
* Develop and implement appropriate algorithms (Python)

**Contributors:**

Our team consists of five members, each responsible for specific tasks:

1. Shubham Kharbanda: Data Preprocessing and Cleaning
   * Handle missing values, standardize data formats, and merge datasets.
2. Pratinav Jinwal: Exploratory Data Analysis (EDA)
   * Gain insights into the data distribution, identify patterns, and visualize relationships between variables.
3. Karan Tejraj Kotian: Data Visualization
   * Design interactive visualizations using charts and plots to effectively communicate key findings.
4. Iqbaldeep Singh Bhullar: Model Selection, Implementation and Evaluation
   * Select appropriate model, assess the performance, tune hyperparameters, and interpret outputs to provide actionable insights for healthcare stakeholders.
5. Divyajot Singh Mankan: Project Management and Documentation
   * Coordinate team efforts, manage timelines, and ensure thorough documentation of the project progress and outcomes.

By distributing tasks among team members, we aim to ensure the successful completion of the project and maximize our collective learning experience.

**References:**

PQDC. <https://data.cms.gov/provider-data/topics/doctors-clinicians>

PQDC. <https://data.cms.gov/provider-data/topics/physician-office-visit-costs>