**Project Plan For Group3\_Project1**

**Project Name:** "Unraveling the Interplay Between Glycemic Control and Diabetic Microvascular Complications in Shanghai Patients with Diabetes: A Visual Exploration"

**Duration:** 4 days

**Presentation to Stakeholders:** Day 5

------------------------------------------------------------------------------------------------------------------

**Day 1: Project Setup and Data Acquisition**

*La Shawn Sykes*

Task 1: Data Collection- **Complete on 4/09/2024**

Subtask 1: Research and identify relevant datasets for the project

Subtask 2: Download or access the identified datasets, ensuring data relevance and quality

*La Shawn Sykes*

Task 2: Environment Setup- **Complete on 4/09/2024**

Subtask 1: Set up the programming environment with required libraries (e.g., Pandas, NumPy, Matplotlib, Seaborn)

Subtask 2: Ensure all team members have access to the required tools and resources

*Aaron Cranor/Lewis Hill*

Task 3: Initial Data Exploration

Subtask 1: Perform a preliminary review of the acquired datasets to understand their structure and contents

Subtask 2: Document initial observations and potential areas for further analysis

**Day 2: Data Preprocessing and Cleaning**

*La Shawn Sykes*

Task 4: Data Integration- **Complete on 4/10/2024**

Subtask 1: Combine and integrate the acquired datasets into a single, unified dataset for analysis

Subtask 2: Address any inconsistencies or discrepancies in the combined dataset

*La Shawn Sykes*

Task 5: Data Cleaning- **Complete on 4/10/2024**

Subtask 1: Identify and handle missing or null values in the dataset

Subtask 2: Remove duplicate entries and outliers to ensure data integrity

*Aaron Cranor/Lewis Hill*

Task 6: Feature Engineering – **In Progress**

Subtask 1: Derive and create new features or variables from the existing dataset to enhance analysis

Subtask 2: Standardize or normalize relevant features as required for the analysis

**Day 3: Data Analysis and Visualization**

*La Shawn Sykes*

Task 7: Exploratory Data Analysis (EDA)

Subtask 1: Perform in-depth analysis of the preprocessed dataset to identify patterns, trends, and relationships

Subtask 2: Generate descriptive statistics and visualizations to understand the data distribution

*Aaron Cranor*

Task 8: Advanced Data Analysis

Subtask 1: Apply advanced statistical and analytical methods to uncover deeper insights within the dataset

Subtask 2: Conduct correlation analysis and feature prioritization to identify key variables of interest

*Lewis Hill*

Task 9: Visual Representation

Subtask 1: Create visualizations such as scatter plots, bar charts, and heatmaps to represent key findings

Subtask 2: Develop interactive and informative visualizations for effective communication of insights

**Day 4: Final Analysis and Reporting Preparation**

*La Shawn Sykes*

Task 10: Final Data Analysis

Subtask 1: Perform a final review and validation of the analysis results

Subtask 2: Document and summarize the key findings and insights from the analysis

*Aaron Cranor*

Task 11: Presentation Preparation

Subtask 1: Compile the analysis results and visualizations into a cohesive presentation format

Subtask 2: Rehearse and finalize the presentation for the stakeholders' review

*Lewis Hill*

Task 12: Documentation

Subtask 1: Prepare a detailed project report with the analysis methodology and results

Subtask 2: Document the data preprocessing and analysis workflow for future reference

**Day 5: Presentation to Stakeholders**

La Shawn Sykes, Aaron Cranor, and Lewis Hill

Task 13: Stakeholder Presentation

Subtask 1: Deliver a comprehensive presentation showcasing the project's objectives, analysis findings, and visualizations

Subtask 2: Address any questions or feedback from the stakeholders regarding the project and its outcomes