Project Title: Pateint\_Health\_Care Analysis

Project Sponsor: Arbor Academy

Project Manager: Aniket Borkar

Date: 11/02/2025

Project Purpose:

To analysis patient health records for analysis patient distribution based on age, gender blood group. Find key trends based on patient health records.

Project Objectives:

1. To analysis patient distribution based on age, gender and blood group.
2. Analysis medical conditions and find trends over time.
3. Analysis patient billing amounts based on medical conditions.
4. Analysis insurance provider based on patient medical conditions.
5. Find average admission duration of patient in hospital.
6. Group by the doctor based on medical conditions.
7. Find average billing amounts of patient based on medical condition.

1. Business Understanding

* Objective: Age and gender distribution can reveal which age groups are most frequently admitted. Analysis admission types (Elective, Emergency, Urgent) can help in resource allocation.
* Questions:
  + Which age group have most medical problems?
  + Which periods have peak hospital usage?
  + Which Hospital have most patient admit and most revenue?
  + Calculate doctor\_pateint\_ratio:- Count number of patient per doctor
* Success Criteria: A clear understanding of patient health, room availability, and actionable insights for Room utilization data can be used for better space management optimization.

2. Data Understanding

* Data Sources:
  + *Patient personal information*: To understand patient background.
  + *Hospital information*: Hospital name, Doctor name, Room availability, insurance adviser, patient admit and discharge details.
  + *Medical condition details*: gives information about medical condition and billing details.
* Activities:
  + Explore data to understand data (e.g., patient details, medical condition details ).
  + Visual inspection to identify missing or inconsistent values.
* Challenges: Inconsistent data details & potential missing values`.

3. Data Preparation

* Tasks:
  + *Data Cleaning*: Remove duplicates, handle missing values.
  + Transform categorical data: (e.g., convert ‘Blood Type’ or ‘Admission Type’ into numerical categories for analysis).
  + *Feature Engineering*: Create new features like length of hospital stay, age bracket,

billing Category based on billing amount, Year Column for analysis.

* Output: Cleaned dataset with new derived fields to use for analysis and modeling.

4. Modeling

* Techniques:
  + Age Group Classification: Calculate age bracket based on patient age to analysis which age group have health issues.
  + *Length of Hospital Stay Analysis*: Calculate the patient stay days to manage room availability.
  + *Billing Category based on billing amount analysis: Categorize bill amount to analysis* patient expenses on health.
* Models:
  + Using Regression to predict Disease using input features based on data.
  + Clustering patient by medical condition based on similar medical condition.
  + Using Regression to predict Billing amount using features based on data.
* Considerations: Choose features with the highest influence, use cross-validation to ensure stability.

5. Evaluation

* Performance Metrics:
  + Mean Absolute Percentage Error (MAPE) for Billing amount forecast accuracy.
  + Mean Squared Error (MSE): Measures squared differences to penalize large errors.
* Validation:
  + Validate if the model identifies the room availability based on length of hospital stay.
  + Evaluate model results with domain experts to confirm feasibility and correctness.

6. Deployment

* Deliverables:
  + A dashboard visualizing disease based on age bracket, length of hospital stay and categorize bills based on billing amounts.
  + Reports with key findings and recommended actions (e.g., suggested doctor and insurance provider, suggest which time period most of the patient admit).
* Considerations: Automate regular updates of the patient dashboard using live data integration.

7. Monitoring and Maintenance

* Monitoring:
  + Set up regular checks for model accuracy to ensure consistent results.
  + Monitor the data quality of patient records figures.
* Maintenance:
  + Update models periodically to incorporate patterns patient disease and adapt to changing trends.
* Feedback Loop: Collect feedback from stakeholders on the usefulness of the analysis and adjust methods accordingly.

Timeline: 02/02/2025 to 09/02/2025

* Day 1 & 2: Data exploration and requirement analysis.
* Day 3 & 4: Data cleaning and preparation.
* Day 5 & 6: Analysis and creation of insights.
* Day 7: Report generation and presentation of results.

Member Name:-

Email ID:-

Role and responsibility:-

Signature:-

Date:-

Location:-