

Data Science Intern at Data Glacier

Project: Healthcare - Persistency of a drug (Data Science)

Week 7: Deliverables

| | | | | |
|-----------------------|----------------------------|---------------------|------------------------|--|
| NAMES | Taiwo Akingbesote | Kerr Tan | Farzana Chowhury | Ayah Ibrahim |
| UNIVERSITY | Montclair state University | New York University | Mount Holyoke College | University of North Carolina at Greensboro |
| EMAIL | Akingbesote1@montclair.edu | St4153@nyu.edu | Chowd23f@mtholyoke.edu | ayariyadh9@gmail.com |
| COUNTRY | USA | USA | USA | USA |
| SPECIALIZATION | Data Science | Data Science | Data Science | Data Science |
| BATCH CODE | LISUM22 | LISUM22 | LISUM22 | LISUM22 |
| DATE | 19 July, 2023 | 19 July, 2023 | 19 July, 2023 | 19 July, 2023 |

ALL SUBMITTED TO DATA GLACIERS

Table of Contents:

| | |
|---------------------------|---|
| 1. Project Plan..... | 3 |
| 2. Problem Statement..... | 3 |

1. Project Plan

| WEEKS | DATE | PROGRESS |
|---------|---------------|---|
| Week 7 | July 19, 2023 | Problem Statement and Data Processing |
| Week 8 | July 26, 2023 | Data Understanding |
| Week 9 | Aug 2, 2023 | Data Cleansing and Transformation |
| Week 10 | Aug 9, 2023 | EDA Analysis and Recommendation |
| Week 11 | Aug 16, 2023 | EDA Presentation and proposed model technique |

| | | |
|---------|--------------|-------------------------------|
| Week 12 | Aug 23, 2023 | Model Selection |
| Week 13 | Aug 30, 2023 | Final project report and code |

2. Problem Statement

The problem faced by ABC pharma company is to understand the persistency of a drug based on physician prescriptions. Persistency refers to the extent to which patients continue to take a drug as prescribed by their physicians. The goal is to identify the factors that influence the persistency of a drug for a given dataset. The company wants to automate this process by building a classification model that can predict the persistency of a drug based on certain features.

To solve this problem, we need to understand the dataset provided by ABC pharma company. This dataset contains information about various patients. We will explore and analyze this dataset to identify patterns and relationships between variables. Then build a model that classifies the dataset.