

Group Name: Data Force

Name	Email	Company	Country	Specialization
Anshi Mathur	anshimathur0325@gmail.com	Data Glacier	United States of America	Data Science
Lujain Saad	ljainsaadcs@gmail.com	Data Glacier	Saudi Arabia	Data Science
Prince Kumar Lat	princek.iitk@gmail.com	Data Glacier	Canada	Data Science
Mohamed Amine Kina	kinaamine@gmail.com	Data Glacier	Germany	Data Science

Problem Description: Pharmaceutical companies like ABC Pharmacy want to understand the persistency of a drug based on a doctor's prescription and hire an analytics company to automate the process of identification.

Business Understanding: Using various different models, we will be able to create a machine learning application that is the most accurate to successfully predict the persistency of a drug based on a dataset we have been provided.

Project Life Cycle Along With Deadline:

- Week 1 deliverables (8/19): Establishing a group and project.
- Week 2 deliverables (8/26): Take a close look at the data and understand what the group will need to do to remove outliers and missing values to create the most accurate model.
- Week 3 deliverables (9/2): Clean up the data set to remove outliers and missing information.
- Week 4 deliverables (9/9): Create an EDA ipynb file to contain the EDA performed on the data.
- Week 5 deliverables (9/16): Create an EDA presentation to the business users.
- Week 6 deliverables (9/23): Choose a model and build it based on the business requirements.
- Week 7 deliverables (9/30): Create a final presentation on the model.

Data Intake Report:

Name: Healthcare project – Drug Persistency

Internship Batch: LISUM11

Version: 1.0

Data intake by: Data Force team

Data intake reviewer: Data Force team

Data storage location: local

Tabular data details:

Total number of observations	3424
Total number of files	1
Total number of features	70
Base format of the file	xlsx
Size of the data	920 KB

Github Repo Link:

<https://github.com/anshimathur0325/Data-Science-Persistency-of-a-Drug-Project>