Group Name: Data Force

Name	Email	Company	Country	Specialization
Anshi Mathur	anshimathur0325@gmail.com	Data Glacier	United States of America	Data Science
Lujain Saad	lujainsaadcs@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.

Data Understanding: In the given dataset, the variable persistency is followed by multiple columns of patient's data which all have an affect on the persistency of an individual. In order to create the most effective model, we were able to experiment with different machine learning modules and evaluate for their accuracy.

Type of Data: The dataset contains data of various different types including integers and objects. In order to create the best analysis, we divided the objects into separate columns and converted them into numbers, which makes a more accurate model.

Problems of the Dataset: When searched through, the dataset did not contain null values. To add on, we were able to identify the unique values of the dataset, where we did not encounter any values that seem out of the ordinary and would require us to remove them.

Approach: In order for us to clean the data, we first made sure to print out information about the data with both the data type and if there were null values. After we determined that there were no null values, we identified if there was data in specific columns that stood out, where we discovered that all of the unique values are nothing out of the ordinary.

Github Repo link:
https://github.com/anshimathur0325/Data-Science-Persistency-of-a-Drug-Project