GLOBALCERT

DRUG CLASSIFICATION

PROBLEM DEFINITION

Pharmaquick is a pharmaceutical company/hospital that has approached us with a problem statement that asks us to forecast which drug type is best for each patient based on the information provided.

Our goal is to create a machine learning model that can accurately perform this task while also providing insights into the data and advice on how to use the model's output.

AGENDA

1. Study the Dataset
2. Data Exploration
3. Exploratory Data Analysis- EDA
4. Feature Engineering
5. Data Modeling and Comparing

1. UNDERSTANDING THE DATASET

200 **rows** and 6 **columns**

No **null** values

2 **Numeric** columns
- Age
-Na_to_K)

4 **Categorical** Columns
- Sex
- BP
- Cholesterol
-Drug

2. DATA EXPLORATION

Unique elements per column

Number of Unique elements in **Age** column: **57**
Number of Unique elements in **Sex** column: **2**
Number of Unique elements in **BP** column: **3**
Number of Unique elements in **cholesterol** column: **2**
Number of Unique elements in **Na_to_K**: **198**
Number of Unique elements in **Drug** column: **5**

Skewness graph

1. Age graph: **Normal skewed**, symmetrical
Age skewness: **0.0303083570300607**
2. Na_To_K graph: **right skewed**
Na to K skewness: **1.039341186028881**

3. EXPLORATORY DATA ANALYSIS

15 EDA graphs in 3 categories

Distribution Graphs

1. Gender
2. BP
3. Cholesterol
4. Drugs

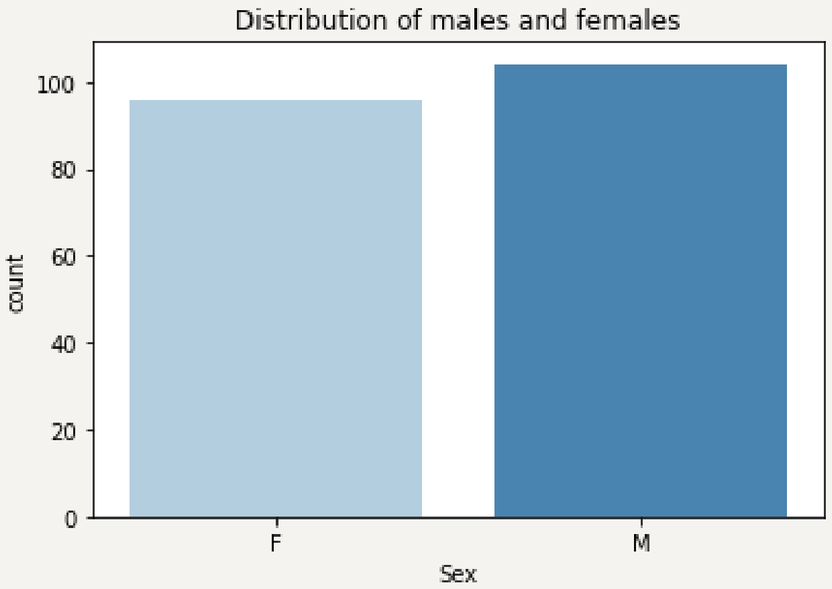
Frequency Graphs

1. Gender vs drugs
2. Gender vs Cholesterol
3. Gender vs BP
4. BP vs Cholesterol
5. Na to K Vs Gender Vs Age

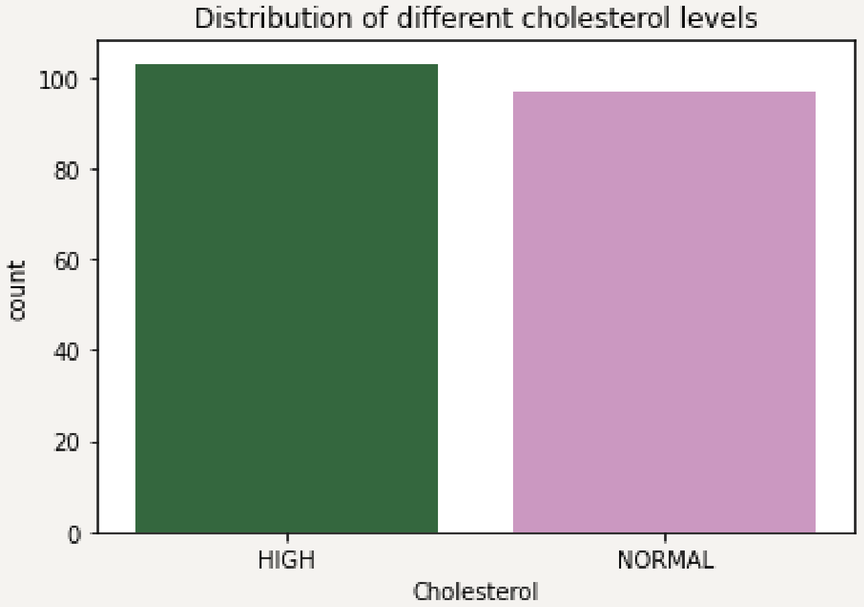
Prescription Graphs

1. SEX
2. BP
3. cholesterol
4. Na-to-K
5. Age
6. Special- Na vs BP vs Drug

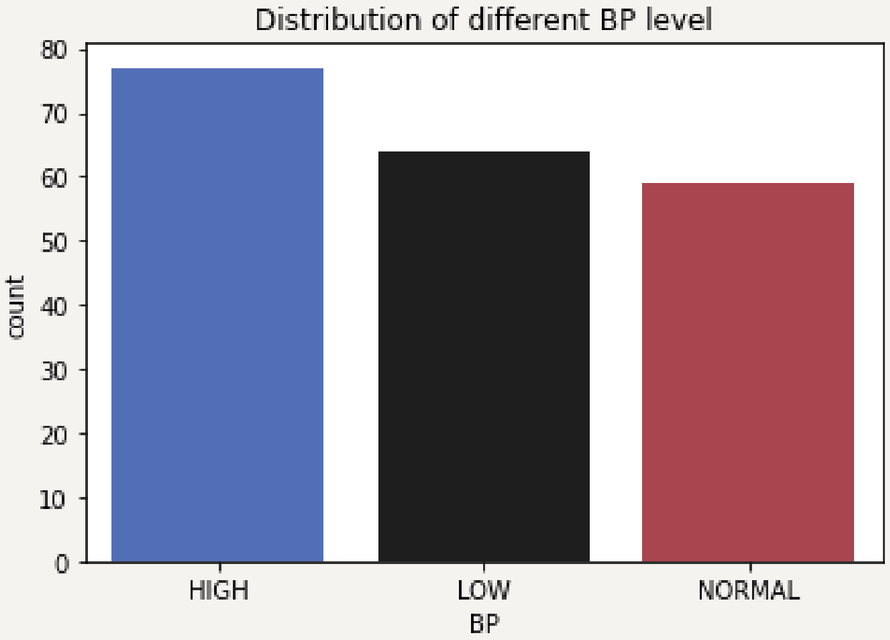
DISTRIBUTION GRAPHS



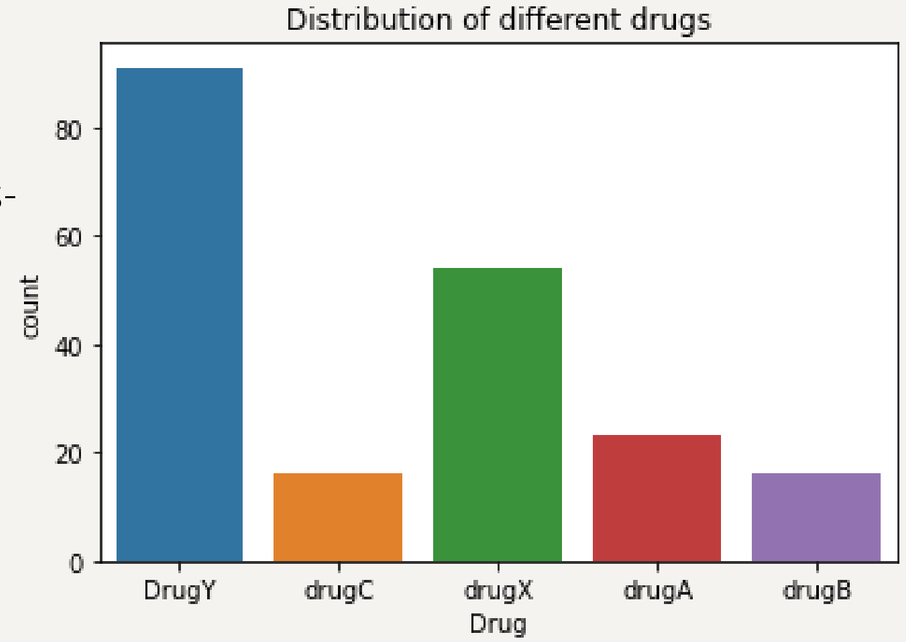
There were **104 M** and **96 F** records



The records were more with **high cholesterol(103)** as compared to **normal (97)**



Blood Pressure recording includes- **77 High** Bp, **64 Low** Bp and **59 Normal** Bp



The maximum amount of drugs available is **DrugY -91** and the least is **DrugC** and **DrugB-16**

CONCLUSION

qqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqq
qqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqq
qqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqq
qqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqqq
qqqqqqqqqqqq

THANK YOU