EDA / Descriptive Statistics

## Introduction:

Pharmaceutical industry wants to develops, demand, producing their own drugs. The global pharmaceuticals market produced treatments worth over $1.2 trillion, and this industry serves hospitals and medical facilities worldwide. For demand of the drugs in hospital and maintain an sufficient drug inventory, hospital chains must developed their inventory management strategies. For solving this problem first we collect and analyze the large amount of data related to drug and patient demand. So hospital management ensure that they have correct drugs quantity for the patients.

## Overall design strategy

Based on the problem and objective formal definition, the data sales transactions information system are cleaned, feature engineering approach was defined, and all data are transformed to monthly time series, consisting of aggregate sales among different classes of pharmaceutical drugs in monthly time periods.

### **Objectives**

* Data Cleaning: Identify errors, inconsistencies, and missing values in the dataset.
* Data Transformation: Convert data into an appropriate format or scale for analysis or modeling.
* Feature Engineering: Create new relevant features or variables from the existing data to improve the performance of machine learning models.

8 data preprocessing methods were implemented:

1. Type Casting
2. Handling missing values
3. Removing duplicates
4. One-Hot encoding
5. Data manipulation

After data preprocessing and descriptive statistics, our next step involves developing EDA to analyze the trends and patterns of the datasets.

## Data Overview

The healthcare management they provide high quality care to the patients. In this data set our main value is drug purchases, sales, patient details, and medical histories. The healthcare database is a central repository for all healthcare-related data, making it easily accessible to various departments within the healthcare organization. This includes clinical departments, such as radiology and pathology, as well as administrative departments, such as billing and finance.

## Users

Users of the visualization dashboard will be:

* Quantity by DrugName
* Sales Trend Over Time
* Sales Performance by Specialization
* Sales by Drug Subcategory

## Questions

## 