Exploratory Data Analysis Using MySQL

In a healthcare facility, ensuring the availability of essential medical supplies and drugs is critical for providing timely and effective patient care. However, the increasing bounce rate in medical inventory poses significant challenges to inventory management and patient satisfaction.

**Background:**

* The bounce rate refers to the frequency of stockouts or instances where medical items are not available when needed.
* Stockouts can lead to delays in patient treatment, compromise patient safety, and result in revenue loss for the healthcare facility.
* Managing medical inventory effectively requires balancing the need to maintain adequate stock levels with the goal of minimizing excess inventory and associated costs.

To perform exploratory data analysis using MySQL we first load the dataset into MySQL server.

* Create a new database named ‘medical inventory optimization’.
* Create a new table named ‘medical\_inventory’ to import raw data into SQL
* Import the CSV file into MySQL.
* Create a new table named ‘medical\_inventory\_stagging’ and copy all the data from the ‘medical\_inventory’ table to a new table. ( It is a good practice to make copies of raw data to perform EDA in case of data loss.)

CREATE TABLE medical\_inventory\_stagging

LIKE medical\_inventory\_optimaizatio\_dataset;

* Total number of records is 70270.

SELECT COUNT(\*) AS total\_rows

FROM medical\_inventory\_stagging;

|  |  |
| --- | --- |
| **VARIABLE NAME** | **VARIABLE DESCRIPTION** |
| Typeofsales | Type of sale of the drug. Either the drug is sold or returned. |
| Patient\_ID | ID of a patient |
| Specialisation | Name of Specialisation (eg. Cardiology) |
| Dept | Pharmacy, the formulation is related with. |
| Dateofbill | Date of purchase of medicine |
| Quantity | Quantity of the drug |
| ReturnQuantity | Quantity of drug returned by patient to the pharmacy |
| Final\_Cost | Final Cost of the drug (Quantity included) |
| Final\_Sales | Final sales of drug |
| RtnMRP | MRP of returned drug (Quantity included) |
| Formulation | Type of formulation |
| DrugName | Generic name of the drug |
| SubCat | Subcategory (Type) to the category of drugs. |
| SubCat1 | Subcategory (condition) to the category of drugs |

* After Carefully observing the data we can see that there are blank values in the dataset.

The columns Formulation, DrugName, SubCat, SubCat1 contain Blank rows. Replace these with ‘Unknown’.

* The data type of the Dateofbill column is text. Update it to date format
* Perform primary analysis of the database
* Frequency of Typeofsales:

Sale 🡪 62685

Return 🡪 7585

* Frequency of Department:

Department1 contains 61385 records. Department2 contains 1060 records and Department3 contains 7825 records out of the total number of rows.

* Frequency of Formulation:

Form1 – 58110

Form2 – 6625

Unknown – 2445

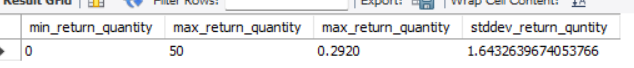
Patent – 2695

Form3 - 395

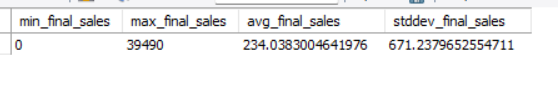
* Let's perform summary statistics on numeric columns.
* Summary statistics of Quantity:



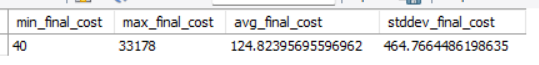
* Summary statistics of Return quantity



* Summary Statistics of Final\_Sales:



* Summary statistics for Final\_Cost



* Questionnaire:

1. What is the statistical breakdown(percentage) of sales versus returns in the dataset?
   * + 1. The Typeofsale column contains two types sales and return representing drugs sold and returned respectively.
       2. The sales percentage is 88.18% and the return rate is 11.82%.
2. What is the statistical distribution of sales versus return within different specialisations?
3. There are 13 records of Specialisations which have no return rate.
4. Specialisation54 has the highest return rate with more than 30%
5. Provide statistical analysis of sales and returns across different departments of the pharmacy.
6. Department 1 has the highest number of Sales count (10760) and highest number of Return count(1680).
7. Department3 100% sales having no drug type returned.
8. Statistically compare the average quantity involved in sales transactions versus returns.

-Average quantity sold per transaction: 2.53

-Average quantity returned per transaction: 2.47