Analytical Report

for DS by Cipto Bagus Jati Kusumo

TOC

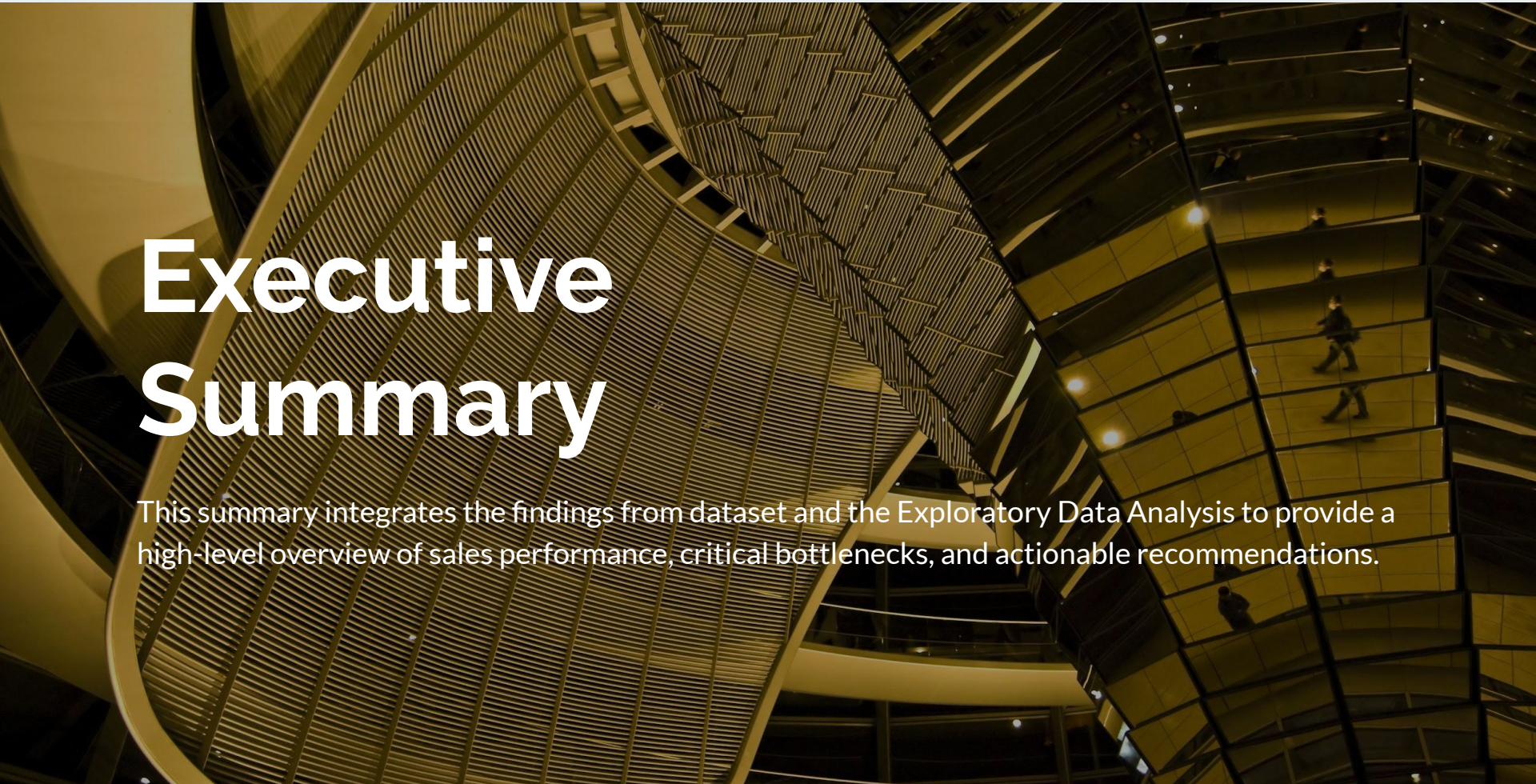
Overview

The candidate is provided with a set of column names and is allowed to create their own dummy dataset. After that, the candidate is expected to analyze the dataset, identify the company's issues based on the data, and provide actionable recommendations that the company should implement.

Here are the set of column names:
ChannelCode, ChannelStoreName, CustomerID, Gender, Age, City, TglReg, ProductCode, ProductName, TotalSales, PaymentMethodType, PaymentMethodName, DoctorID, DoctorName, MedicalJobdesk

The Dataset

1. Dataset consist of 8581 rows and 17 columns
2. Dataset is dummy data generated using Python script
3. Each row represent one transaction by Customer
4. The transaction date ranged from 1 January 2023 until 31 December 2023



Executive Summary

This summary integrates the findings from dataset and the Exploratory Data Analysis to provide a high-level overview of sales performance, critical bottlenecks, and actionable recommendations.

Performance Overview

The year 2023 was characterized by strong performance in the first three quarters, followed by a significant downturn in the final quarter.

1. **Total Revenue** : Rp 5,810,450,000
2. **Total Transactions** : 8,581
3. **Total Unique Customers** : 1,972
4. **Top Performing Store** : Erha Derma Center Kemanggisan
5. **Top Revenue Driver** : Laser Rejuvenation contributed ~43.6% of total revenue (Rp 2.53 Billion)

Critical Findings

From January to September, the business maintained a healthy monthly revenue average of approximately Rp 580 Million. However, a sharp decline began in October, with revenue plunging to an average of Rp 197 Million for the remainder of the year.

1. Monthly Revenue Peak : July 2023 at Rp 637.6M
2. Monthly Revenue Low : November 2023 at Rp 163.6M
3. Revenue Loss : There was a loss of approximately Rp 413M between September and November alone.

Collapse of High-Ticket Treatments

The revenue decline is directly tied to the collapse in volume of Laser Rejuvenation treatments.

1. September : 136 transactions
2. November/December : Dropped to only 4 transactions per month. As the most expensive service (averaging ~Rp 1.9M per session), this drop-off severely impacted the bottom line

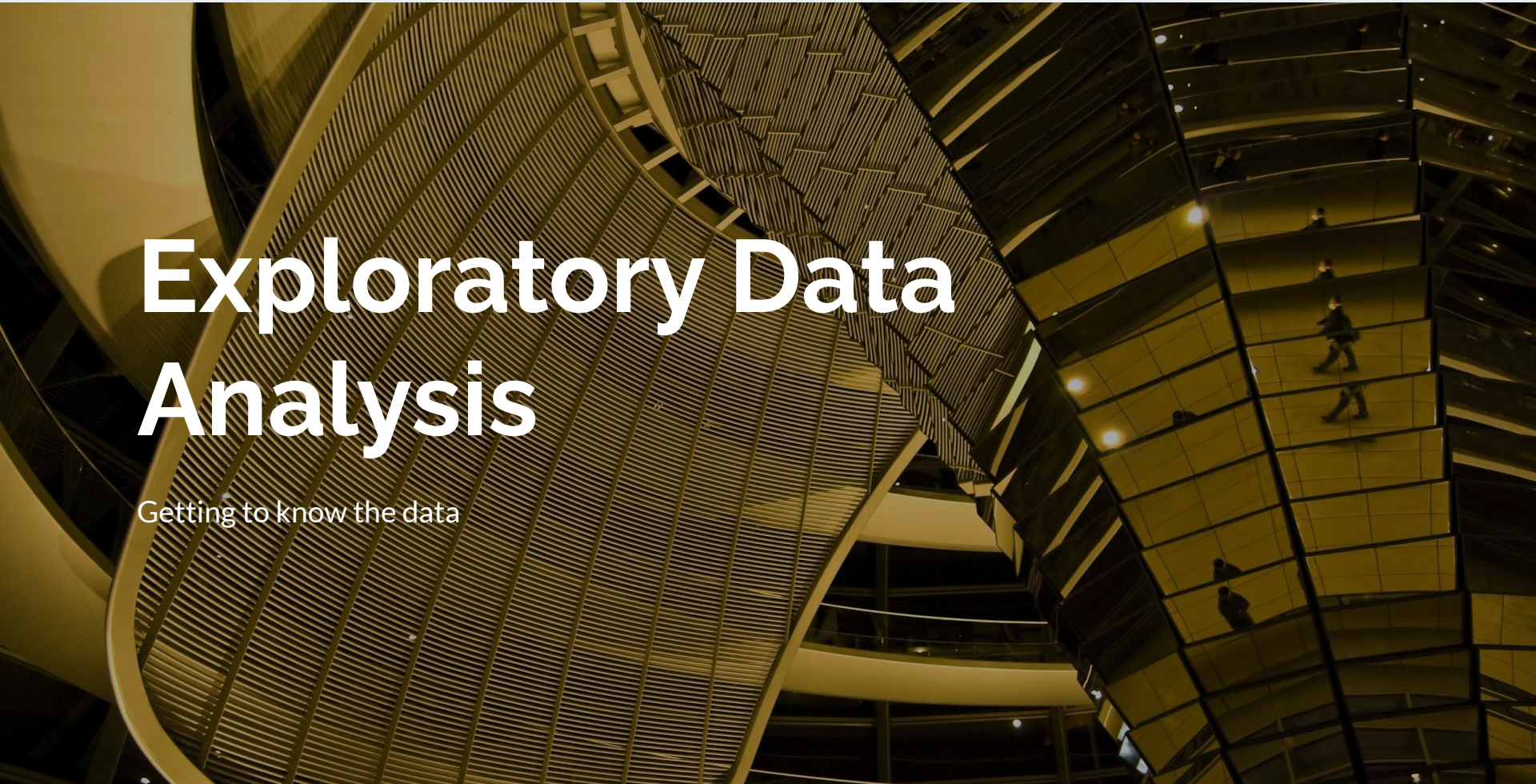
Root Cause: Doctor Availability & Capacity

The EDA identifies Doctor availability as the primary driver of this decline.

1. While five doctors/medical staff were listed in the system, their active transaction volume per head dropped by over 50-60% in Q4.
2. The absence or reduced handling capacity of doctors directly limited the clinics' ability to perform high-margin medical treatments (like Laser Rejuvenation), which cannot be fulfilled by retail nurse alone.

Strategic Recommendations

To recover the lost revenue and stabilize performance for the upcoming year, management should conduct immediate Medical Staff Audit & Interviews with the medical team (Doctors Amanda, Budi, and Citra) to identify the reasons for the Q4 slump (e.g., scheduling conflicts, attrition, or burnout) as suggested in the EDA.



Exploratory Data Analysis

Getting to know the data

Dataset Structure

01

We don’t know the data yet.

Here we see the list of columns, how many rows that has value, and its data type.

Turns out 2 date columns has object data type.

Converting to date data type make it easier to wrangling later on.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8581 entries, 0 to 8580
Data columns (total 17 columns):
 #   Column             Non-Null Count  Dtype
---  ------             --------------  -----
 0   TransactionDate    8581 non-null   object
 1   ChannelCode        8581 non-null   object
 2   ChannelStoreName   8581 non-null   object
 3   City               8581 non-null   object
 4   CustomerID         8581 non-null   object
 5   Gender             8581 non-null   object
 6   Age                8581 non-null   int64
 7   TglReg             8581 non-null   object
 8   ProductCode        8581 non-null   object
 9   ProductName        8581 non-null   object
 10  TotalSales         8581 non-null   int64
 11  PaymentMethodType  8581 non-null   object
 12  PaymentMethodName  8581 non-null   object
 13  DoctorID           8581 non-null   object
 14  DoctorName         8581 non-null   object
 15  MedicalJobdesk     8581 non-null   object
 16  Month              8581 non-null   object
dtypes: int64(2), object(15)
memory usage: 1.1+ MB
```

```
'''
So the problem only data type, we could fix this issues by executing following
lines.
'''
dataset.TransactionDate = pd.to_datetime(dataset.TransactionDate)
dataset.TglReg = pd.to_datetime(dataset.TglReg)
```

0.0s

Dataset Behaviour

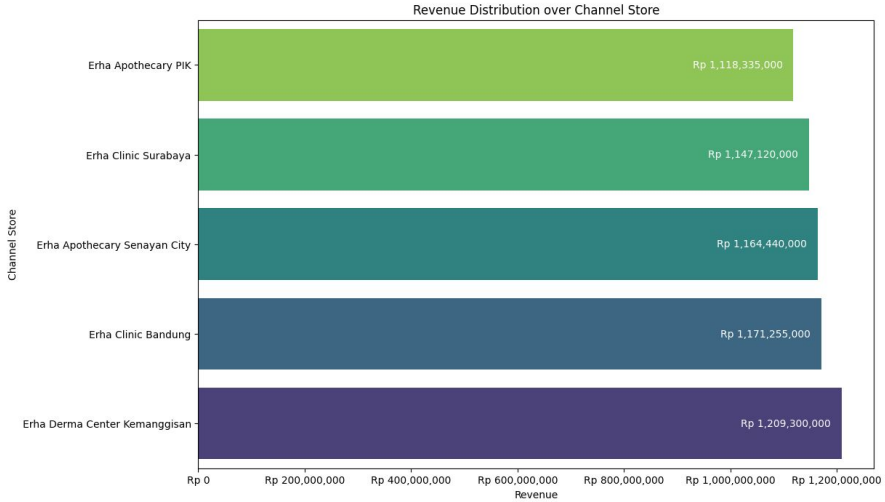
02

My first assumption if revenue declining is plotting distribution of revenue across channel store.

Overall, there is no noticeable difference in term of revenue for each clinic.

Erha Derma Center Kemanggisan has the most revenue generation, followed by Erha Clinic in Bandung.

Other than that, Erha Apothecary PIK has the lowest revenue generation.



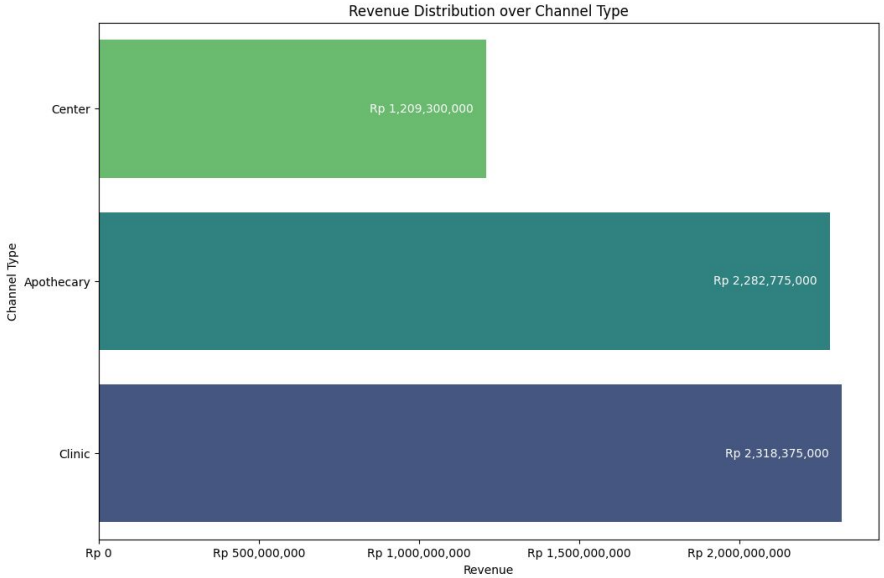
Dataset Behaviour

02

From previous chart, even though Store with type "Center" i.e. Erha Derma Center Kemanggisan has the largest revenue, but overall store type "Clinic" and "Apothecary" has almost twice revenue of store type "Center".

This is because we only have 1 of our "Center" type of store and the rest is 2.

Is there any differences between store type about the sold items or treatments?

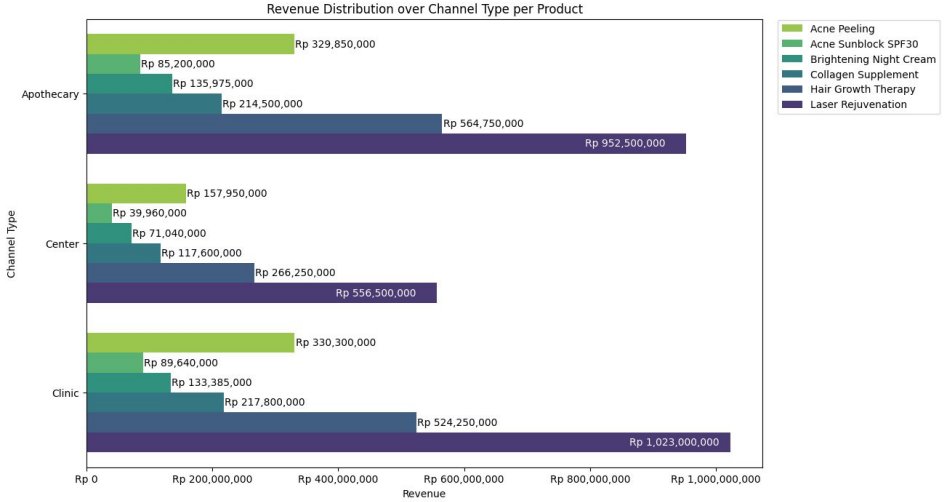


Dataset Behaviour

02

Turns out no, all type of channel type sold all of our items and products, thus there is not because of difference of sold items and products that drives inequality of revenue generation over channel type.

We could see overview of revenue distribution over product name below to get more glance of revenue generation per products.



Dataset Behaviour

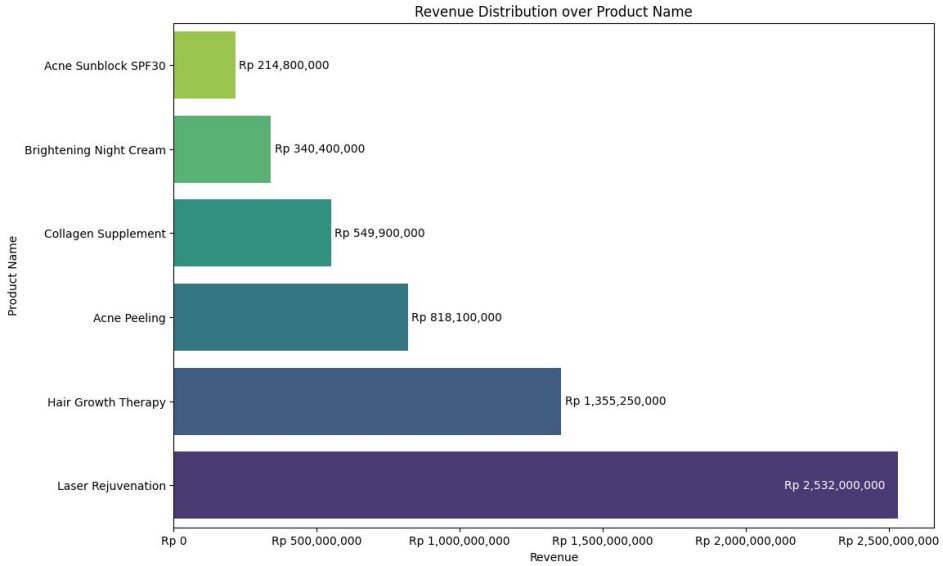
02

Overall, treatment Laser Rejuvenation is the most generated revenue and the least one is product Acne Sunblock SPF30.

Why?

Is it because the price too high?

Is it because no one bought Acne Sunblock SPF 30 and many customer bought Laser Rejuvenation?



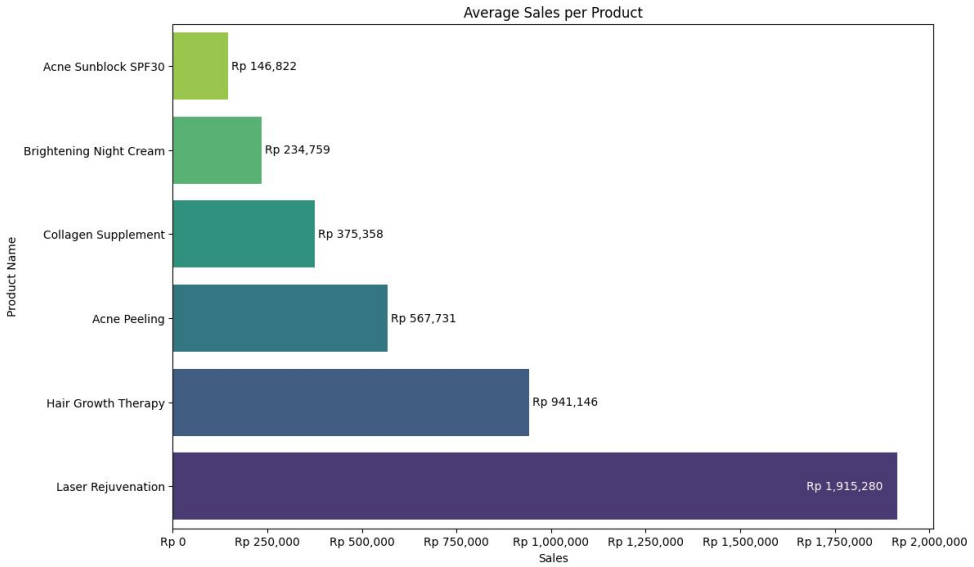
Dataset Behaviour

02

From bar chart on the right, we could see that average Acne Sunblock SPF 30 sales is the cheapest amongst all products.

Unfortunately this is only average, we won't be able to see the spectrum sales price for all the products.

We could plot the spectrum in the following chart.



Dataset Behaviour

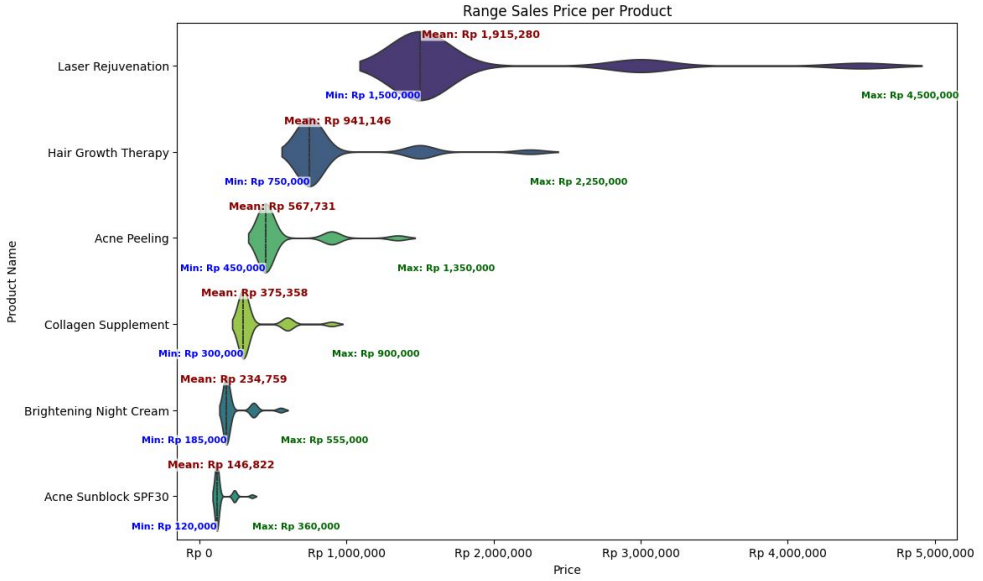
02

The range of sales per product is quite large, even for the cheapest product Acne Sunblock SPF 30 ranged from Rp 120.000 to Rp 360.000, it is 3 times from the lowest one.

And the product that has most generated revenue, the Laser Rejuvenation, ranged from Rp 1.5 mio to Rp 4.5 mio.

So the price for either Acne Sunblock SPF 30 and Laser Rejuvenation has the same quite large range.

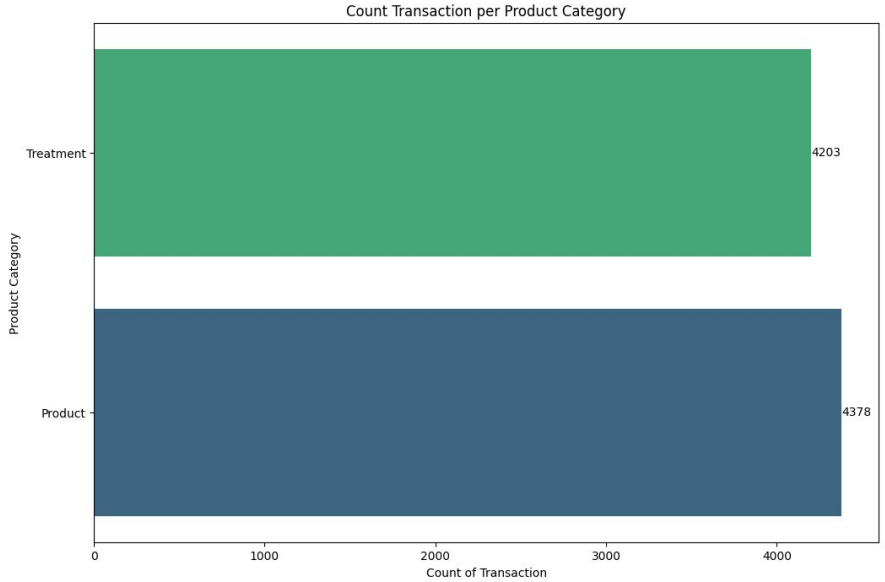
How's about the volume transaction between product and transaction?



Dataset Behaviour

02

Seems like the treatment has less transactions compared to product, let's see in the more details by dividing it per products/treatments.

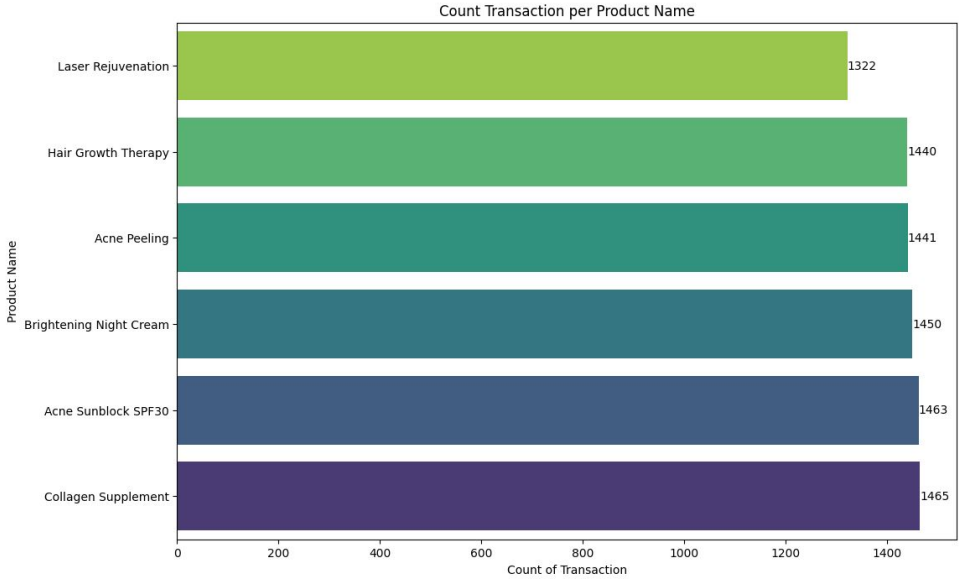


Dataset Behaviour

02

The product that makes treatment category transaction declining is Laser Rejuvenation, but even so, Laser Rejuvenation has the highest sales price so it is contribute to most revenue generation.

Until now, we see no problem at all, let's see distribution of transaction over time.

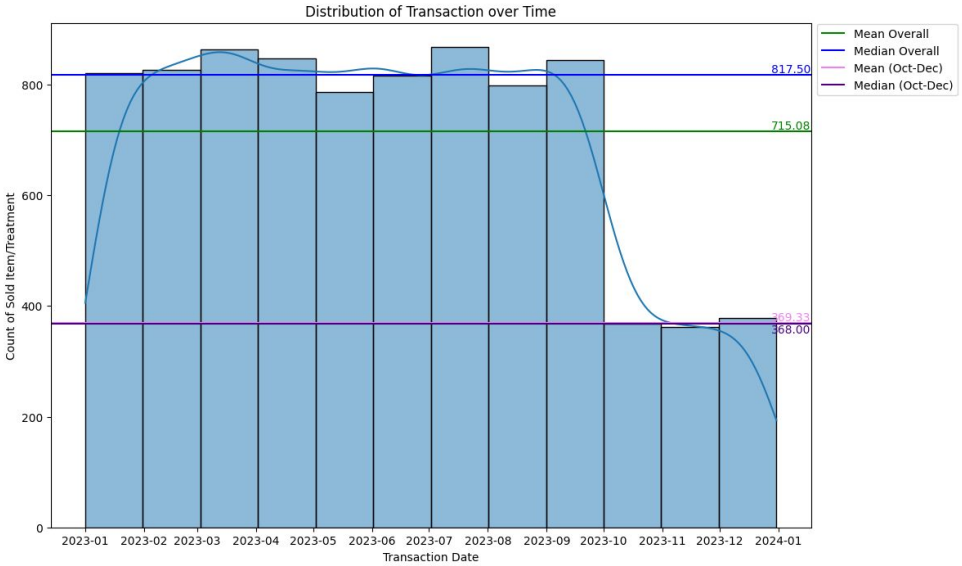


Dataset Behaviour

02

Overall, we've median around 715 transactions every month and from histogram we also knew that there are declining transactions from October to December 2023 to around 368 transactions.

How is the trend monthly revenue?



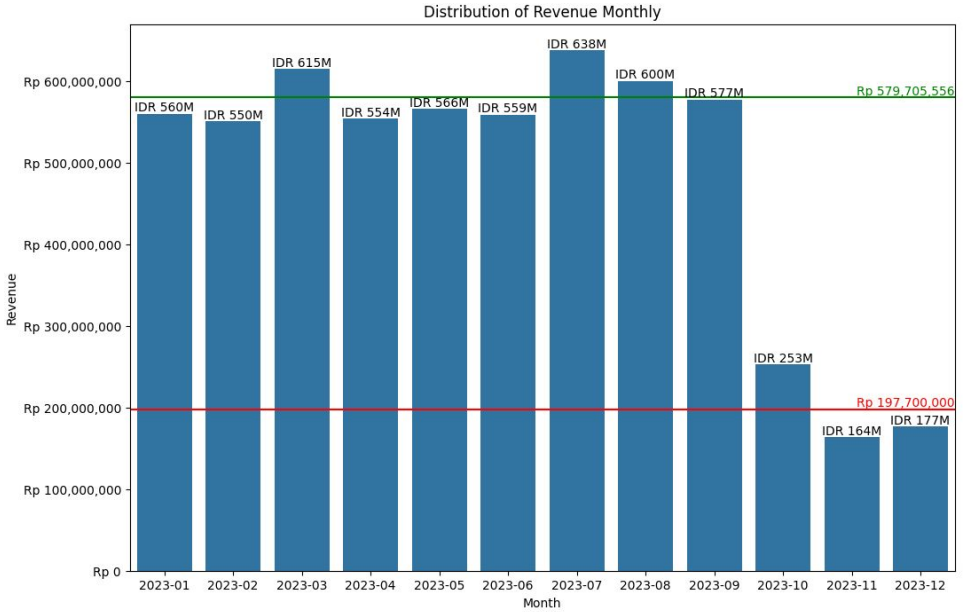
Dataset Behaviour

02

Transactions declining, revenue also declining.

So the transactions affecting revenue.

How it looks like when in the month to month growth?



Dataset Behaviour

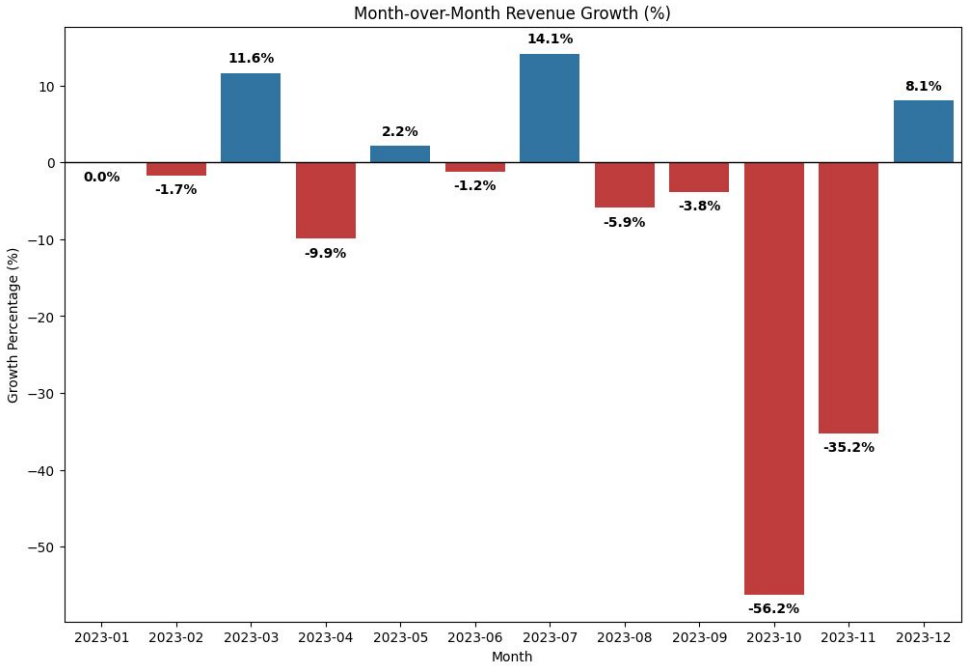
02

Even though it seems like from Jan-Sep we generate much revenue, but in the term growth, it is no so growing,.

This is proven by month over month growth from Jan-Sep has 5 red bars and 3 blue bars. This is getting worse in the Oct - Dec where Oct and Nov has the lowest revenue growth.

Why is it happen?

What kind of products that has declining transactions?

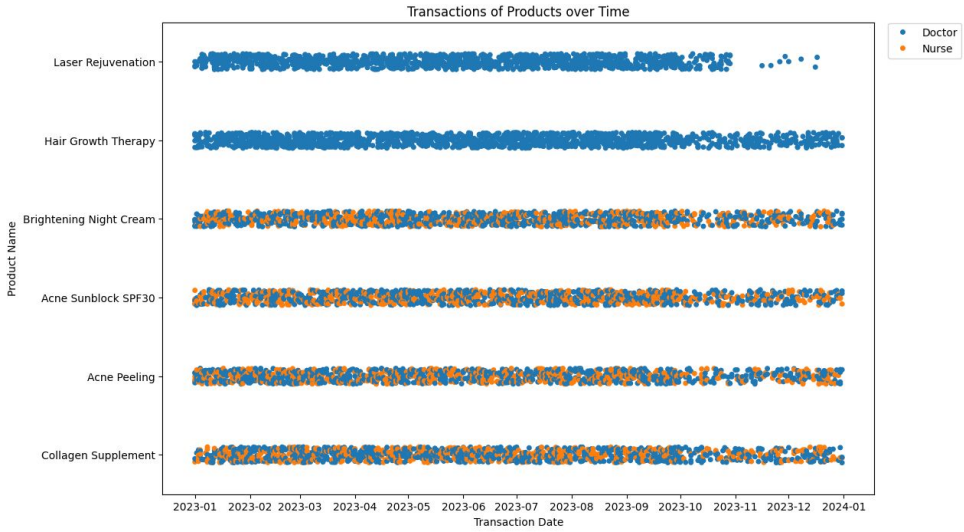


Dataset Behaviour

02

Visualize transactions per products we got what kind of product that has declining in transaction, in this case the Laser Rejuvenation, proven by lesser dot in the Laser Rejuvenation in Oct - Dec compared to Jan - Sep 2023.

How much revenue loss due to this declining transaction of Laser Rejuvenation?

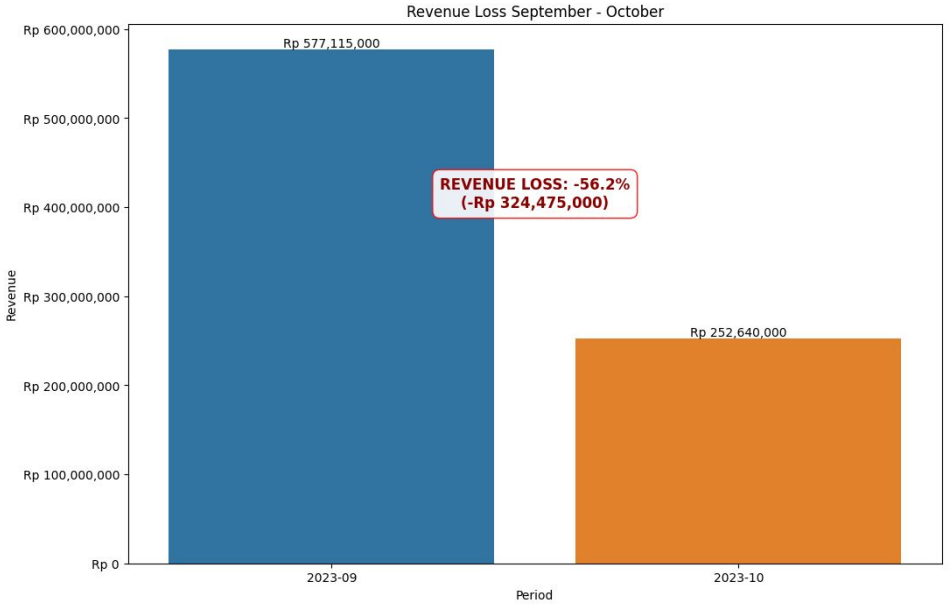


Dataset Behaviour

02

From Sep to Nov, there is total revenue loss Rp 324,4 mio.

Is it happened on every channel store?



Dataset Behaviour

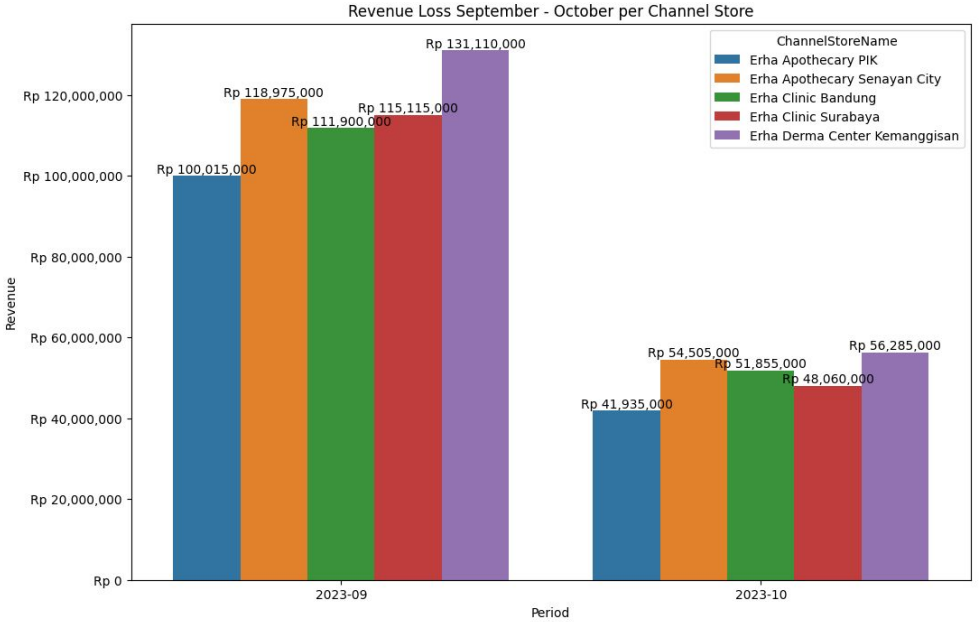
02

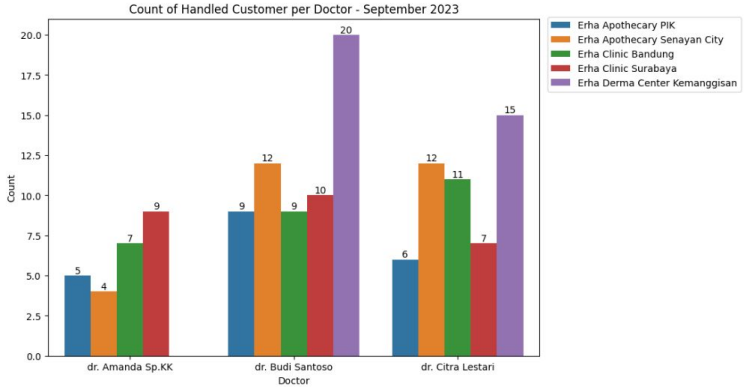
All channel store has the same phenomena, declining revenue.

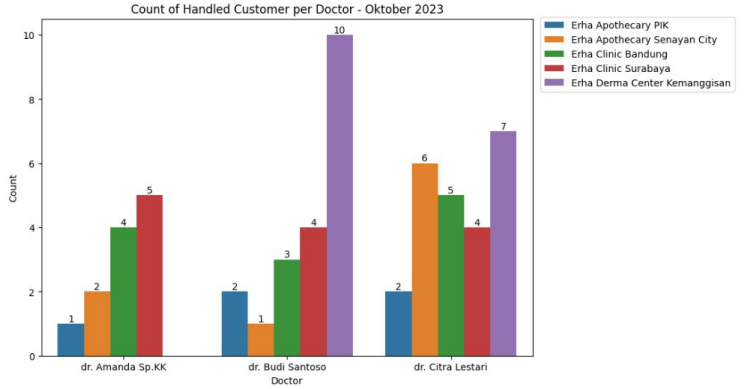
Why?

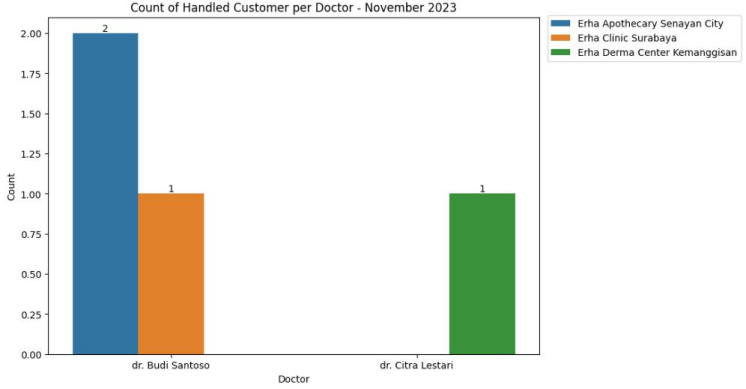
Is it because the Doctor?

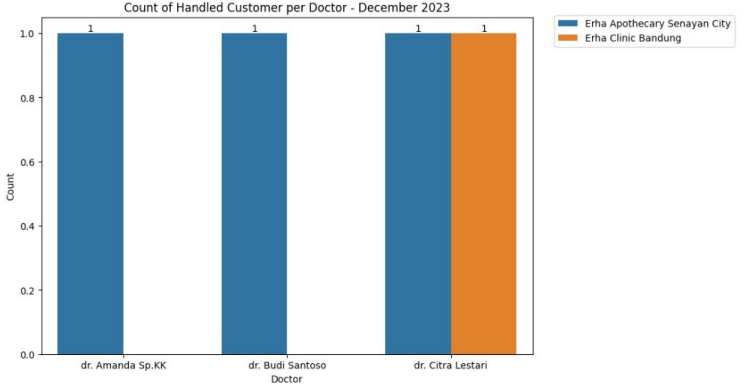
We need to know who is the doctor that handles Laser Rejuvenation in each channel store.




Count of Handled Customer per Doctor - September 2023
Erha Apothecary PIK
Erha Apothecary Senayan City
Erha Clinic Bandung
Erha Clinic Surabaya
Erha Derma Center Kemanggisan
Count
Doctor
dr. Amanda Sp.KK
dr. Budi Santoso
dr. Citra Lestari
5
4
7
9
9
12
9
10
20
6
12
11
7
15


Count of Handled Customer per Doctor - Oktober 2023
Erha Apothecary PIK
Erha Apothecary Senayan City
Erha Clinic Bandung
Erha Clinic Surabaya
Erha Derma Center Kemanggisan
Count
Doctor
dr. Amanda Sp.KK
dr. Budi Santoso
dr. Citra Lestari
1
2
4
5
2
1
3
4
10
2
6
5
4
7


Count of Handled Customer per Doctor - November 2023
Erha Apothecary Senayan City
Erha Clinic Surabaya
Erha Derma Center Kemanggisan
Count
Doctor
dr. Budi Santoso
dr. Citra Lestari
2
1
1


Count of Handled Customer per Doctor - December 2023
Erha Apothecary Senayan City
Erha Clinic Bandung
Count
Doctor
dr. Amanda Sp.KK
dr. Budi Santoso
dr. Citra Lestari
1
1
1
1

Dataset Behaviour

02

Turns out there is declining number of doctor that handling customer from Sep - Oct and it's getting worse from Oct - Nov which is only 2 Doctors that present and two of them handling customer only in 3 out of 5 our channel store with minimum count of handles.

With lower number of Doctors available, less customer handled, thus affecting sales items and treatment.

Actionable recommendation:
Conduct interview with Doctors to find potential problem about their absent in Nov and Dec and also why in Oct they're handling less customer compared to Sep.

Thank you.

Jupyter notebook and all codes are available on [github](#).