Do Weekend Drive Higher Sales? A Data Analysis Of Q4 2022

Felisha Dita Zafirah

The data used is from Tokopedia (not real data). The explanation of the data set is as follows:

1. **order_detail**
2. **sku_detail**
3. **customer_detail**
4. **payment_detail**

order_detail:

1. id → angka unik dari order / *id_order*
2. customer_id → angka unik dari pelanggan
3. order_date → tanggal saat dilakukan transaksi
4. sku_id → angka unik dari produk (sku adalah *stock keeping unit*)
5. price → harga yang tertera pada tagging harga
6. qty_ordered → jumlah barang yang dibeli oleh pelanggan
7. before_discount → nilai harga total dari produk (*price * qty_ordered*)
8. discount_amount → nilai diskon product total
9. after_discount → nilai harga total produk ketika sudah dikurangi dengan diskon
10. is_gross → menunjukkan pelanggan belum membayar pesanan
11. is_valid → menunjukkan pelanggan sudah melakukan pembayaran
12. is_net → menunjukkan transaksi sudah selesai
13. payment_id → angka unik dari metode pembayaran

sku_detail:

1. id → angka unik dari produk (dapat digunakan untuk *key* saat join)
2. sku_name → nama dari produk
3. base_price → harga barang yang tertera pada tagging harga / *price*
4. cogs → *cost of goods sold* / total biaya untuk menjual 1 produk
5. category → kategori produk

customer_detail:

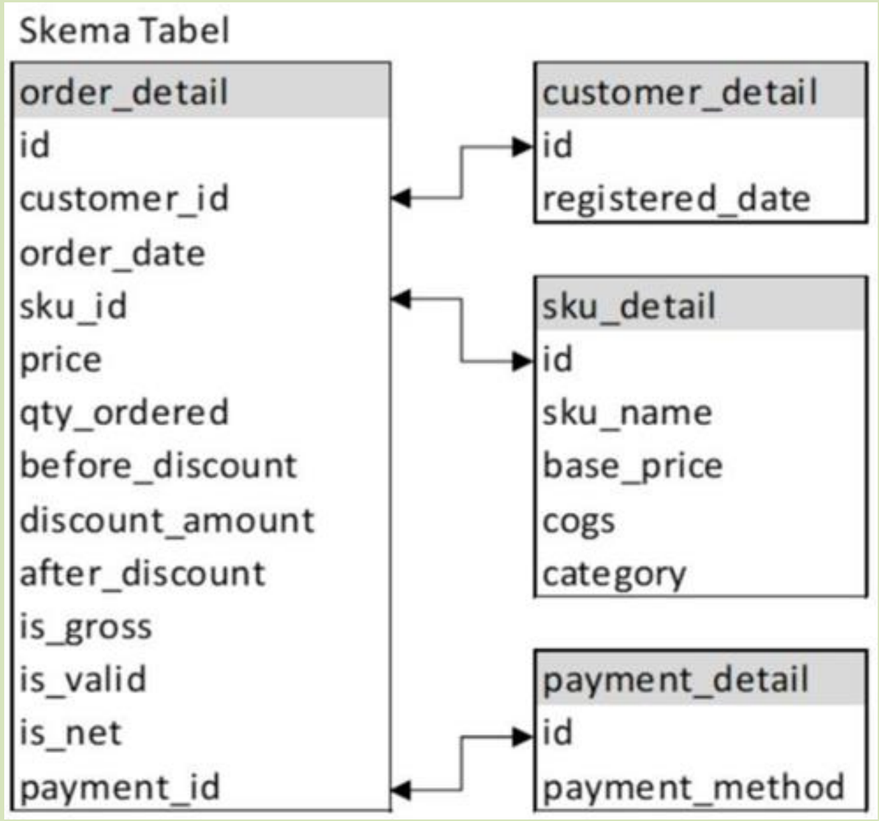
1. id → angka unik dari pelanggan
2. registered_date → tanggal pelanggan mulai mendaftarkan diri sebagai anggota

customer_detail:

1. id → angka unik dari pelanggan
2. registered_date → tanggal pelanggan mulai mendaftarkan diri sebagai anggota

1. **order_detail:** Order data that includes details of the product purchased and the payment method.
2. **sku_detail: Prod**uct data sold on e-commerce
3. **customer_detail:** Buyer information such as ID and registration date
4. **payment_detail:** Payment method ID used for the purchase

TABLE SCHEMA



DATA PREPARATION

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from pandas.tseries.offsets import BDay
```

```
#Sumber data yang digunakan
path_od = "https://raw.githubusercontent.com/dataskillsboost/FinalProjectDA11/main/order_detail.csv"
path_pd = "https://raw.githubusercontent.com/dataskillsboost/FinalProjectDA11/main/payment_detail.csv"
path_cd = "https://raw.githubusercontent.com/dataskillsboost/FinalProjectDA11/main/customer_detail.csv"
path_sd = "https://raw.githubusercontent.com/dataskillsboost/FinalProjectDA11/main/sku_detail.csv"
df_od = pd.read_csv(path_od)
df_pd = pd.read_csv(path_pd)
df_cd = pd.read_csv(path_cd)
df_sd = pd.read_csv(path_sd)
```

variable	class	description
order_detail:		
id	object	angka unik dari order / id_order
customer_id	object	angka unik dari pelanggan
order_date	object	tanggal saat dilakukan transaksi
sku_id	object	angka unik dari produk (sku adalah stock keeping unit)
price	int64	harga yang tertera pada tagging harga
qty_ordered	int64	jumlah barang yang dibeli oleh pelanggan
before_discount	float64	nilai harga total dari produk (price * qty_ordered)
discount_amount	float64	nilai diskon product total
after_discount	float64	nilai harga total produk ketika sudah dikurangi dengan diskon
is_gross	int64	menunjukkan pelanggan belum membayar pesanan
is_valid	int64	menunjukkan pelanggan sudah melakukan pembayaran
is_net	int64	menunjukkan transaksi sudah selesai
payment_id	int64	angka unik dari metode pembayaran
sku_detail:		
id	object	angka unik dari produk (dapat digunakan untuk key saat join)
sku_name	object	nama dari produk
base_price	float64	harga barang yang tertera pada tagging harga / price
cogs	int64	cost of goods sold / total biaya untuk menjual 1 produk
category	object	kategori produk
customer_detail:		
id	object	angka unik dari pelanggan
registered_date	object	tanggal pelanggan mulai mendaftarkan diri sebagai anggota
payment_detail:		
id	int64	angka unik dari metode pembayaran
payment_method	object	metode pembayaran yang digunakan

Question

Number 1

Dear Data Analyst,

At the end of this year, the company will be giving prizes to customers who win the Year-End Festival competition. The Marketing Team needs assistance to determine the estimated prizes that will be awarded to the competition winners. These prizes will be selected from the TOP 5 Products in the Mobiles & Tablets Category during 2022, with the highest quantity of valid sales (valid = 1).

We kindly request your assistance in sending this data to the Marketing Team before the end of this month. We greatly appreciate your support.

Regards

Tim Marketing

ANSWER

NUMBER 1

SYNTAX

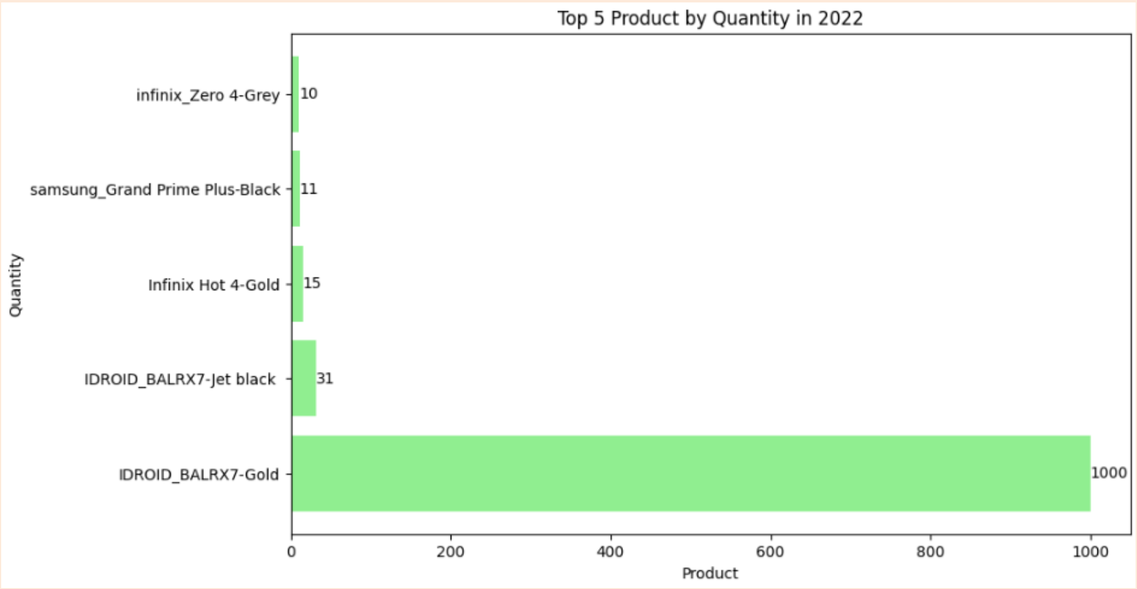
```
# Filter data for valid transactions (is_valid==1)
df_valid = df[df['is_valid']==1].copy()

# Filter data for 'Mobile & Tablets' category
df_mobile = df_valid[df_valid['category']=='Mobiles & Tablets'].copy()

# Filter data for 2022
df_2022 = df_mobile[df_mobile['order_date'].dt.year==2022].copy()

# Group by sku_name and sum the quantity ordered
sku_sales_2022 = df_2022.groupby('sku_name')['qty_ordered'].sum().reset_index()

# Sort by quantity ordered in descending order
top_5_sku= sku_sales_2022.sort_values(by='qty_ordered', ascending=False).head(5)

# Print the top 5 products
print("Top 5 Produk dari Kategori Mobiles & Tablets dengan Kuantitas Penjualan Tertinggi:")
print(top_5_sku)
```

RESULT PYTHON

Number 1

	sku_name	qty_2022
0	IDROID_BALRX7-Gold	1000
1	IDROID_BALRX7-Jet black	31
2	Infinix Hot 4-Gold	15
3	samsung_Grand Prime Plus-Black	11
4	infinix_Zero 4-Grey	10

VISUALISASI



- Insight: There is a significant difference in sales quantity between the top product ("IDROID_BALRX7-Gold" with 1000 units) and the products below it in the top 5. This indicates that one product is particularly popular or may be the focus of a promotion.
- Key Value: Specific sales quantity for each product in the top 5 (1000, 31, 15, 11, and 10 units).
- Key Point: Focus on the product "IDROID_BALRX7-Gold," which dominated sales in this category in 2022. This product is a prime candidate for the "Year-End Festival" competition prize.

RECOMMENDATION

1. Optimize stock and distribution

- Ensure stock availability is maintained to prevent shortages during high demand periods.
- Focus distribution on regions with the highest demand based on geographic sales data.

2. Conduct targeted promotions

- Use data-driven marketing to promote this product more aggressively, such as through social media, online ads, or customer emails.
- Offer bundle deals or discounts for bulk purchases.

3. Expand product variants

- Develop other variants of IDROID_BALRX7 (such as color options or storage capacity) to expand the market and reach new segments.

4. Analyze customer satisfaction

- Collect and analyze customer reviews related to this product to identify popular features or recurring complaints, then use these insights to improve the product or service.

5. Evaluate competitors

- Compare the features and pricing of IDROID_BALRX7-Gold with competitor products to maintain a competitive advantage.

Question

C-6

Number 2

Dear Data Analyst,

Following up on the joint meeting between the Warehouse Team and the Marketing Team, we found that the stock availability for products in the "Others" category at the end of 2022 remained high.

1. We kindly request your assistance in checking the sales data for this category compared to 2021 in terms of sales quantity. Our initial assumption is that there was a decrease in sales quantity in 2022 compared to 2021. (Please also include data for the other 15 categories.)
2. If there was indeed a decrease in sales quantity in the "Others" category, we request your assistance in providing the TOP 20 product names that experienced the most significant decline in 2022 compared to 2021. This data will be used as material for discussion in the next meeting.

We kindly request that you send the data within 4 days from today. Thank you very much for your assistance.

Regards

Tim Werehouse

ANSWER

Number 2.1

SYNTAX

```
# Filter data for valid transactions (is_valid == 1)
df_valid = df[df['is_valid'] == 1].copy()

# Filter data for the year 2021 and 2022
df_2021 = df_valid[df_valid['order_date'].dt.year == 2021].copy()
df_2022 = df_valid[df_valid['order_date'].dt.year == 2022].copy()

# Group by category and sum the quantity ordered for each year
category_sales_2021 = df_2021.groupby('category')['qty_ordered'].sum().reset_index()
category_sales_2021.rename(columns={'qty_ordered': 'qty_2021'}, inplace=True)

category_sales_2022 = df_2022.groupby('category')['qty_ordered'].sum().reset_index()
category_sales_2022.rename(columns={'qty_ordered': 'qty_2022'}, inplace=True)

# Merge the sales data for 2021 and 2022
category_sales_comparison = pd.merge(category_sales_2021, category_sales_2022, on='category', how='outer').fillna(0)

# Calculate the difference in quantity ordered
category_sales_comparison['qty_difference_2022_vs_2021'] = category_sales_comparison['qty_2022'] - category_sales_comparison['qty_2021']

# Sort the categories by quantity difference in ascending order to see the biggest drops
category_sales_comparison_sorted = category_sales_comparison.sort_values(by='qty_difference_2022_vs_2021')

print("Perbandingan Kuantitas Penjualan per Kategori (2022 vs 2021):")
category_sales_comparison_sorted
```

ANSWER

Number 2.1

VISUALIZATION SYNTAX

```
# Visualisai The Sales Comparison
import matplotlib.pyplot as plt

# Visualize the sales comparison
category_sales_comparison_sorted.set_index('category')[['qty_2021', 'qty_2022']].plot(kind='bar', figsize=(12, 6))
plt.title('Sales Quantity Comparison by Category (2021 vs 2022)')
plt.xlabel('Category')
plt.ylabel('Total Quantity Ordered')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.show()
```

RESULT PYTHON

Number 2.1

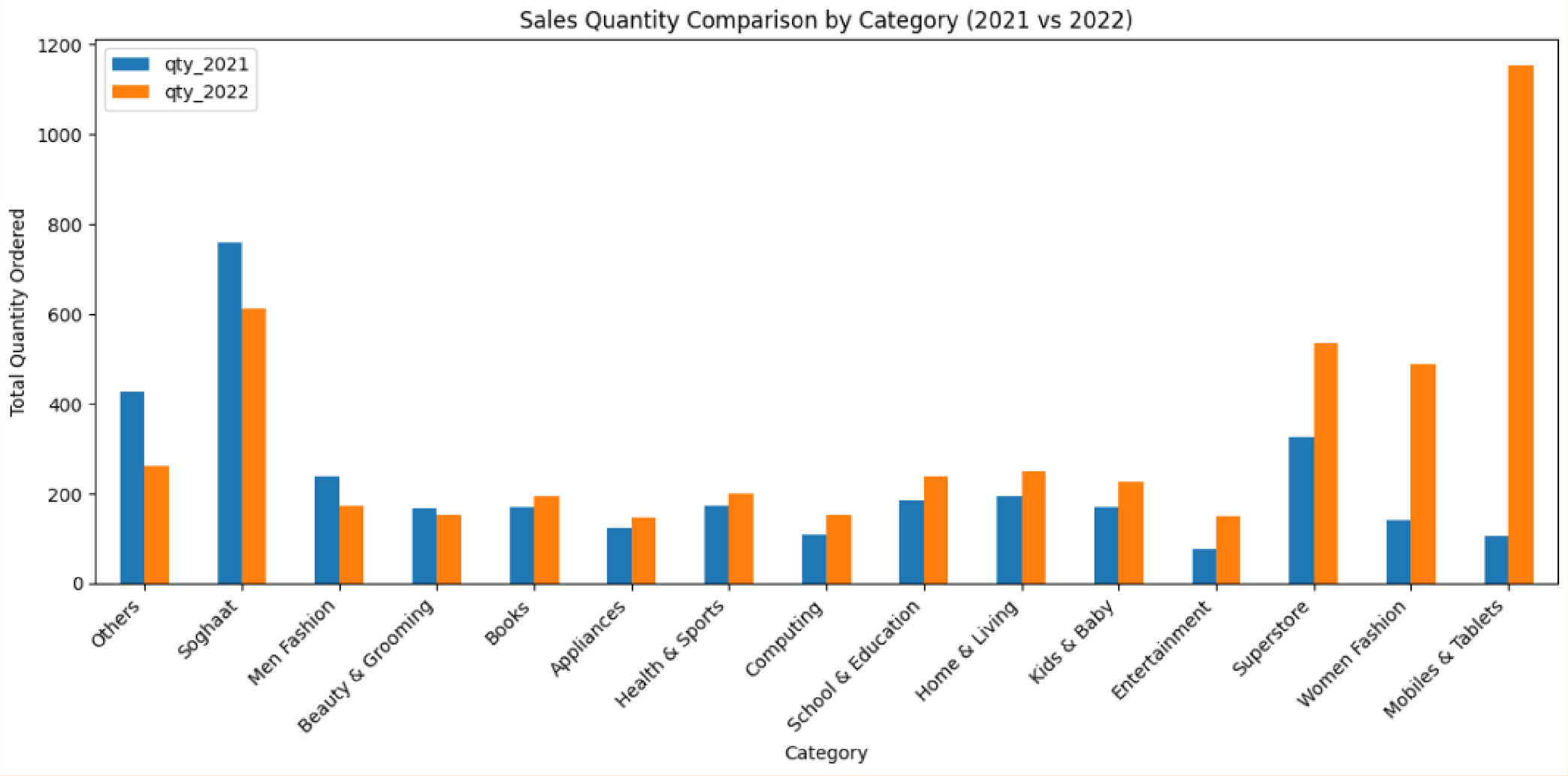
	category	qty_2021	qty_2022	qty_difference
10	Others	426	263	-163
12	Soghaat	759	612	-147
8	Men Fashion	237	175	-62
1	Beauty & Grooming	168	153	-15
2	Books	171	195	24
0	Appliances	124	148	24
5	Health & Sports	173	200	27
3	Computing	109	153	44
11	School & Education	184	237	53
6	Home & Living	193	250	57
7	Kids & Baby	170	227	57
4	Entertainment	77	150	73
13	Superstore	327	536	209
14	Women Fashion	140	489	349
9	Mobiles & Tablets	107	1154	1047

INSIGHT

The image beside shows a table of sales quantity data for 2021 and 2022. The data presents 15 sales categories for both years. It is noted that the highest sales decline occurred in the '**Others**' category, with a decrease of **163 units.**

Number 2.1

VISUALISASI



INSIGHT, KEY VALUE, KEY POINT

It can be seen that the "Others" category experienced a decline in sales quantity in 2022 compared to 2021.

- **Insight:** The Warehouse Team's assumption regarding the decline in sales for the "Others" category is confirmed by the data. This decline is the largest in absolute quantity compared to other categories.
- **Key Value:** The sales quantity of the "Others" category decreased by 163 units.
- **Key Point:** The "Others" category recorded the largest drop in sales quantity from 2021 to 2022.

RECOMMENDATION

- **Investigation of Causes of Decline:**
  Conduct a deeper investigation to understand the factors causing the sales decline in the "Others" category. Is this related to market trend changes, competition, product quality issues, ineffective marketing strategies, or other factors?
- **Evaluation of Products with the Highest Decline:**
  Perform an in-depth analysis of the 20 products with the steepest declines. Consider revisiting stock, pricing, and promotion strategies, or even reducing stock levels or discontinuing products that continue to underperform.
- **Remarketing Strategy:**
  Consider launching targeted marketing campaigns for the "Others" category or for specific products within that category that still have potential.
- **Stock Optimization:**
  Adjust stock availability for the "Others" category based on the latest sales data to avoid future stock accumulation of unsold items. Leverage sales data from growing categories to optimize stock in those areas.
- **Focus on Growing Categories:**
  While focusing on the "Others" category, do not overlook other categories that show positive growth. Continue developing strategies for these categories to maximize sales potential.
- **Team Collaboration:**
  Discuss these findings with the Marketing and Warehouse Teams to jointly formulate strategies to address the sales decline in the "Others" category and optimize overall sales performance.

ANSWER

Number 2.2

SYNTAX

```
# Filter data for the year 2021 and 2022
df_others_2021 = df_others_valid[df_others_valid['order_date'].dt.year == 2021].copy()
df_others_2022 = df_others_valid[df_others_valid['order_date'].dt.year == 2022].copy()

# Group by sku_name and sum the quantity ordered for each year
sku_sales_others_2021 = df_others_2021.groupby('sku_name')['qty_ordered'].sum().reset_index()
sku_sales_others_2021.rename(columns={'qty_ordered': 'qty_2021'}, inplace=True)

sku_sales_others_2022 = df_others_2022.groupby('sku_name')['qty_ordered'].sum().reset_index()
sku_sales_others_2022.rename(columns={'qty_ordered': 'qty_2022'}, inplace=True)

# Merge the sales data for 2021 and 2022
sku_sales_others_comparison = pd.merge(sku_sales_others_2021, sku_sales_others_2022,
                                       on='sku_name', how='outer').fillna(0)

# Calculate the difference in quantity ordered
sku_sales_others_comparison['qty_difference_2022_vs_2021'] = (
    sku_sales_others_comparison['qty_2022'] - sku_sales_others_comparison['qty_2021']
)

# Filter for products with a decrease in sales (difference < 0)
sku_sales_others_decrease = sku_sales_others_comparison[
    sku_sales_others_comparison['qty_difference_2022_vs_2021'] < 0
].copy()

# Sort by the magnitude of the decrease (most negative difference)
top_20_sku_decrease_others = sku_sales_others_decrease.sort_values(
    by='qty_difference_2022_vs_2021'
).head(20)

print("\nTOP 20 Produk Kategori Others dengan Penurunan Penjualan Tertinggi (2022 vs 2021):")
top_20_sku_decrease_others
```

ANSWER

Number 2.2

SYNTAX

```
# Visualisasi Top 20 Prouct (2022 & 2021)
import matplotlib.pyplot as plt

# Visualize the top 20 products with the highest decrease
top_20_sku_decrease_others.set_index('sku_name')['qty_difference_2022_vs_2021'].plot(kind='bar', rot=45, figsize=(10, 6))
plt.title('Top 20 Products in "Others" Category with Highest Sales Decrease (2022 vs 2021)')
plt.xlabel('Product Name')
plt.ylabel('Quantity Difference (2022 - 2021)')
plt.xticks(rotation=45, ha='right')
plt.tight_layout()
plt.show()
```

RESULT PYTHON

Number 2.2

	sku_name	qty_2021	qty_2022	qty_difference
69	RB_Dettol Germ Busting Kit-bf	200.0	45.0	-155.0
18	Dawlance_MD 10 + DWB 600	23.0	0.0	-23.0
84	Telemall_MM-DR-HB-L	23.0	2.0	-21.0
137	iu_Tickets General Enclosure-Islamabad	20.0	0.0	-20.0
70	RS_Rehmat-e-Shereen Mix Mithai	13.0	0.0	-13.0
141	kansai_NeverWet	10.0	1.0	-9.0
154	sindbad_Sindbad Gift Card-3	7.0	0.0	-7.0
133	emart_00-1	7.0	1.0	-6.0
93	Vouch 365 2016	5.0	0.0	-5.0
4	Am-PTV_ATS-004-M	5.0	0.0	-5.0
129	duma_4561253300294	4.0	0.0	-4.0
155	sockoye_QG in Quarter Grey	4.0	0.0	-4.0
85	The Vitamin Company Kojic Acid Whitening Cream...	4.0	0.0	-4.0
115	aw_Octane Booster-12oz./354ml	3.0	0.0	-3.0
88	Trans2_LW 999	3.0	0.0	-3.0
43	MEGUIAR_G12711	4.0	1.0	-3.0
119	aw_Super Concentrated Fuel Injector (Black)-5....	2.0	0.0	-2.0
156	sstop_Universallensclipkit	2.0	0.0	-2.0
89	Transcend_64GB JetFlash 810	2.0	0.0	-2.0
26	JBS_IFAM-009	2.0	0.0	-2.0

To identify which products in the 'Others' category experienced the most significant decline, here is a list of the 20 products with the highest decrease in sales quantity

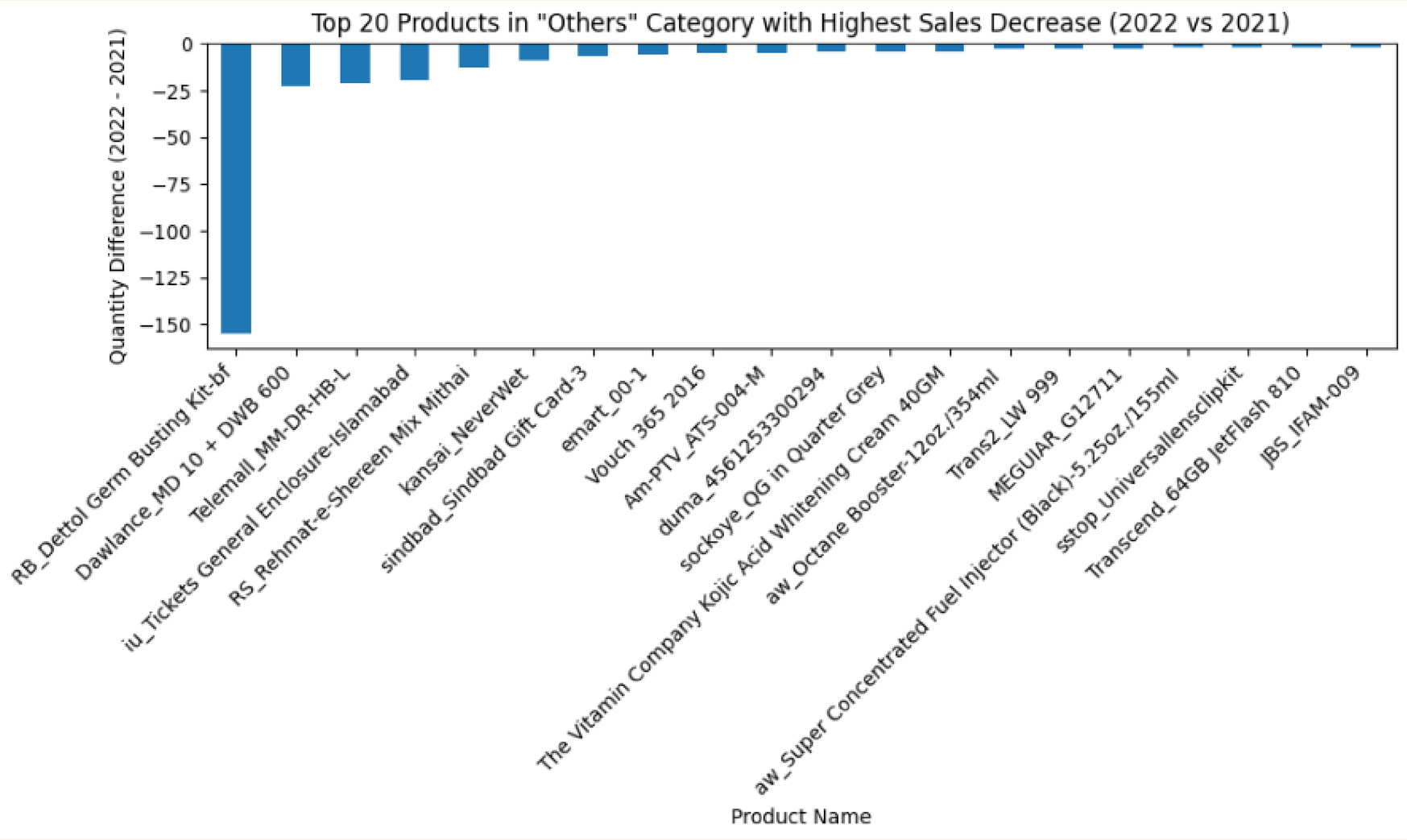
INSIGHT, KEY VALUE, KEY POINT

- **Insight:** The sales decline in the "Others" category is driven by a significant drop in a small number of products. The product "RB_Dettol Germ Busting Kit-bf" shows the steepest decline. Several products even recorded zero sales in 2022 after having sales in 2021.
- **Key Value:** The largest quantity drop occurred in the product "RB_Dettol Germ Busting Kit-bf" with a decrease of 155 units. Additionally, 10 products in the top 20 had no sales in 2022.
- **Key Point:** Identifying these products with the highest decline is crucial for further investigation by the Warehouse and Marketing Teams to understand the causes of the decline and take appropriate actions related to stock management and sales strategy.

C-6

Number 2.2

VISUALISASI



GRAPH INSIGHT

1. The data clearly shows that several specific products in the “Other” category experienced a significant decline in sales quantity from 2021 to 2022.
2. The product **“RB_Dettol Germ Busting Kit-bf”** had the largest drop, followed by **“Dawlance_MD 10 + DWB 600”** and **“Telemall_MM-DR-HB-L.”**
3. Many of the top 20 listed products had no sales in 2022 after recording sales in 2021.

RECOMMENDATION

- **Special Product Investigation:** For each of the top 20 products with the highest decline, a detailed investigation should be conducted to understand the reasons behind the decrease.
  This may include:
- **Marketing and Promotion:** Were these products actively marketed in 2022? Were there any specific campaigns in 2021 that boosted sales but were not replicated in 2022?
- **Pricing and Competitiveness:** Were there any price changes for these products, or did competitors offer similar products at lower prices?
- **Product Life Cycle and Demand:** Have these products reached the end of their product cycle, or has market demand for them decreased?
- **Stock Availability:** Even though high stock levels were mentioned, were there periods of stock-outs for certain products that could have affected sales?
- **Customer Reviews and Feedback:** Were there any negative reviews or customer feedback that could explain the decline?

Question

Number 3

Dear Data Analyst,

In relation to the company’s anniversary in two months, the Digital Marketing Team will be providing promotional information to customers at the end of this month. The customer criteria we need are those who have checked out but have not completed payment (is_gross = 1) during 2022. The data we need are the Customer ID and Registered Date.

We kindly request your assistance in sending this data to the Digital Marketing Team before the end of this month. Thank you very much for your support.

Regards

Tim Digital Marketing

ANSWER

Number 3

SYNTAX

```
# Filter data for transactions where is_gross is 1, is_valid is 0, and is_net is 0
df_gross_unpaid = df[(df['is_gross']==1) & (df['is_valid']==0) & (df['is_net']==0) & (df['order_date'].dt.year==2022)].copy()

# Filter data for the year 2022
df_gross_unpaid_2022 = df_gross_unpaid[df_gross_unpaid['order_date'].dt.year==2022].copy()

# Select the required columns: customer_id and registered_date
customer_data_for_promo = df_gross_unpaid_2022[['customer_id', 'registered_date']]

# Print the resulting dataframe
print(customer_data_for_promo)
```

```
#Jalankan kode ini untuk mendownload file
from google.colab import files

# Simpan dataframe ke CSV
customer_data_for_promo.to_csv('audience_list.csv', encoding = 'utf-8-sig',index=False)
files.download('audience_list.csv')
```

RESULT

Audience_list

RESULT PYTHON

Number 4

	customer_id	registered_date
9	C246762L	2022-05-08
18	C848774L	2021-11-07
19	C693415L	2022-04-12
21	C180595L	2022-04-22
22	C587425L	2022-03-22
...	...	...
5856	C394076L	2021-10-12
5859	C248585L	2022-07-10
5865	C471304L	2022-05-13
5881	C265450L	2022-02-17
5883	C676393L	2021-07-27

[1052 rows x 2 columns]

"For the needs of the Digital Marketing Team regarding the company's anniversary promotion, customer data of those who have checked out but have not completed the payment in 2022 (indicated by is_gross = 1, is_valid = 0, and is_net = 0) has been identified. The requested data includes Customer ID and Registered Date, as shown in the image beside."

INSIGHT, KEY VALUE, KEY POINT

- **Insight:** This data is highly valuable for the Digital Marketing Team to target customers who have shown interest in the products (by checking out) but have not completed their transactions. They can be a great audience for re-engagement campaigns or special offers.
- **Key Value:** The total number of customers that meet this criterion (can be identified from the number of rows in the downloaded CSV file), along with the specific Customer IDs and Registered Dates of each customer.
- **Key Point:** This customer list is ready to be used by the Digital Marketing Team for the anniversary promotion campaign, focusing on converting incomplete transactions from 2022.

Question

Number 4

Dear Data Analyst,

From October to December 2022, we conducted campaigns every Saturday and Sunday. We would like to assess whether these campaigns had a significant impact on sales (before_discount). We kindly request your assistance in providing the following data:

1. The average daily sales on weekends (Saturday and Sunday) vs. the average daily sales on weekdays (Monday–Friday) for each of those months. Indicate whether there was an increase in sales in each month.
2. The average daily sales on weekends (Saturday and Sunday) vs. the average daily sales on weekdays (Monday–Friday) for the overall 3-month period.

We kindly request that you send this data no later than next week. Thank you very much for your assistance.

Regards

Tim Campaign

ANSWER & RESULT PYTHON

Number 4.1

```python
# Filter data for valid transactions (is_valid = 1)
df_valid = df[df['is_valid'] == 1].copy()

# Filter data for the months October, November, and December 2022
df_campaign_period = df_valid[
    (df_valid['order_date'].dt.year == 2022) &
    (df_valid['order_date'].dt.month.isin([10, 11, 12]))
].copy()

# Create columns for month and day of the week
df_campaign_period['month'] = df_campaign_period['order_date'].dt.strftime('%B')  # Full month name
df_campaign_period['day_of_week'] = df_campaign_period['order_date'].dt.day_name()

# Separate data into weekends and weekdays
weekends_df = df_campaign_period[df_campaign_period['day_of_week'].isin(['Saturday', 'Sunday'])].copy()
weekdays_df = df_campaign_period[df_campaign_period['day_of_week'].isin(['Monday', 'Tuesday', 'Wednesday', 'Thursday', 'Friday'])].copy()

# Calculate daily sales for weekends and weekdays per month
weekends_sales_per_day_month = weekends_df.groupby([weekends_df['order_date'].dt.date, 'month'])['before_discount'].sum().reset_index()
weekdays_sales_per_day_month = weekdays_df.groupby([weekdays_df['order_date'].dt.date, 'month'])['before_discount'].sum().reset_index()

# Calculate average daily sales for weekends and weekdays per month
avg_weekends_sales_month = weekends_sales_per_day_month.groupby('month')['before_discount'].mean().reset_index()
avg_weekends_sales_month.rename(columns={'before_discount': 'avg_weekend_sales'}, inplace=True)

avg_weekdays_sales_month = weekdays_sales_per_day_month.groupby('month')['before_discount'].mean().reset_index()
avg_weekdays_sales_month.rename(columns={'before_discount': 'avg_weekday_sales'}, inplace=True)

# Merge the average sales data
monthly_avg_sales = pd.merge(avg_weekends_sales_month, avg_weekdays_sales_month, on='month')

# Calculate the difference and percentage change
monthly_avg_sales['sales_difference'] = monthly_avg_sales['avg_weekend_sales'] - monthly_avg_sales['avg_weekday_sales']
monthly_avg_sales['percentage_change'] = (monthly_avg_sales['avg_weekend_sales'] - monthly_avg_sales['avg_weekday_sales']) / monthly_avg_sales['avg_weekday_sales'] * 100

print("Rata-rata Harian Penjualan Weekends vs Weekdays per Bulan (Oct-Dec 2022):")
print(monthly_avg_sales)

# Check for sales increase on weekends for each month
print("\nPeningkatan Penjualan Weekends vs Weekdays per Bulan:")
for index, row in monthly_avg_sales.iterrows():
    month = row['month']
    if row['sales_difference'] > 0:
        print(f"Pada bulan {month}: Ada peningkatan penjualan di weekends sebesar {row['sales_difference']:.2f} ({row['percentage_change']:.2f}%) dibandingkan weekdays.")
    elif row['sales_difference'] < 0:
        print(f"Pada bulan {month}: Terjadi penurunan penjualan di weekends sebesar {abs(row['sales_difference']):.2f} ({row['percentage_change']:.2f}%) dibandingkan weekdays.")
    else:
        print(f"Pada bulan {month}: Tidak ada perbedaan signifikan dalam penjualan antara weekends dan weekdays.")
```

RESULT PYTHON

Number 4.1

Analysis of the average daily sales (before_discount) during the campaign in October, November, and December 2022 shows a comparison between weekdays and weekends.

```
Rata-rata Harian Penjualan Weekends vs Weekdays per Bulan (Oct–Dec 2022):
      month  avg_weekend_sales  avg_weekday_sales  sales_difference  \
0  December       4.105994e+06          8542530.0     -4.436536e+06
1  November       5.774045e+06          6204666.0     -4.306210e+05
2   October       5.708341e+06          7872212.4     -2.163872e+06

   percentage_change
0         -51.934684
1          -6.940277
2         -27.487466

Peningkatan Penjualan Weekends vs Weekdays per Bulan:
Pada bulan December: Terjadi penurunan penjualan di weekends sebesar 4436536.00 (-51.93%) dibandingkan weekdays.
Pada bulan November: Terjadi penurunan penjualan di weekends sebesar 430621.00 (-6.94%) dibandingkan weekdays.
Pada bulan October: Terjadi penurunan penjualan di weekends sebesar 2163871.73 (-27.49%) dibandingkan weekdays.
```

INSIGHT, KEY VALUE, KEY POINT

- **Insight:** The weekend campaigns during the October to December 2022 period appear to have not succeeded in boosting sales compared to weekdays in those months. In fact, the average daily sales on weekends were lower than on weekdays in each month.
- **Key Value:** The drop in average daily sales on weekends compared to weekdays varied each month, with the largest decrease occurring in December (-4,443,653.60).
- **Key Point:** The average daily sales on weekends were lower than the average daily sales on weekdays during the October–December 2022 campaign period.

RECOMMENDATION

- **Evaluate and replicate October's strategy for other months:** Review the promotional strategies, featured products, and marketing activities implemented in October 2022, then adapt these approaches for the upcoming months.
- **Strengthen weekend promotions:** Since weekend sales consistently declined from October to December 2022, consider offering weekend promotions (such as weekend flash sales or shopping vouchers) and enhancing digital marketing activities leading up to the weekend.
- **Leverage December momentum:** Since weekday sales increased in December 2022, ensure that year-end moments (Christmas and New Year) are maximized with more aggressive seasonal campaigns to drive sales across all days.
- **Time-based segmentation in sales strategy:** Consider different approaches for weekdays and weekends, such as:
  - **Weekdays:** Focus on routine needs, office-worker bundles, lunchtime discounts, etc.
  - **Weekends:** Focus on family, leisure, and entertainment-oriented or relaxation products.

Number 4.2

ANSWER PYTHON

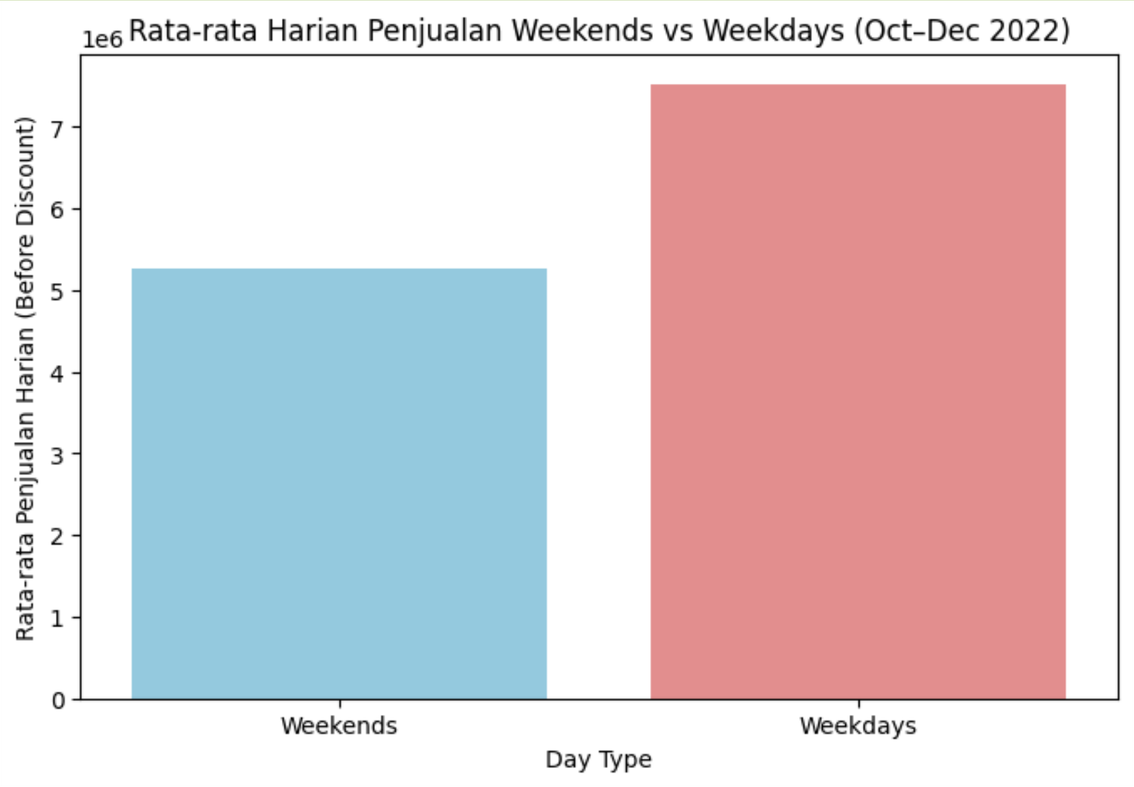
```python
# Tulis kode Anda di bawah ini. Dapat menggunakan lebih dari 1 blok kode
# Calculate the number of unique weekend days and weekday days in the campaign period
num_weekend_days = weekends_df['order_date'].dt.date.nunique()
num_weekday_days = weekdays_df['order_date'].dt.date.nunique()

# Calculate total sales for weekends and weekdays across the 3 months
total_weekend_sales = weekends_df['before_discount'].sum()
total_weekday_sales = weekdays_df['before_discount'].sum()

# Calculate overall average daily sales for weekends and weekdays
overall_avg_weekend_sales = total_weekend_sales / num_weekend_days if num_weekend_days > 0 else 0
overall_avg_weekday_sales = total_weekday_sales / num_weekday_days if num_weekday_days > 0 else 0

print("\nRata-rata Harian Penjualan Weekends vs Weekdays Keseluruhan 3 Bulan (Oct-Dec 2022):")
print(f"Rata-rata Harian Penjualan Weekends: {overall_avg_weekend_sales:.2f}")
print(f"Rata-rata Harian Penjualan Weekdays: {overall_avg_weekday_sales:.2f}")

# Compare the overall averages
if overall_avg_weekend_sales > overall_avg_weekday_sales:
    print(f"Secara keseluruhan selama 3 bulan, rata-rata penjualan harian di weekends lebih tinggi dari weekdays.")
elif overall_avg_weekend_sales < overall_avg_weekday_sales:
    print(f"Secara keseluruhan selama 3 bulan, rata-rata penjualan harian di weekends lebih rendah dari weekdays.")
else:
    print(f"Secara keseluruhan selama 3 bulan, rata-rata penjualan harian di weekends sama dengan weekdays.")

# Optional: Plotting for visualization (similar to 4.1 but for overall)
# Create a DataFrame for plotting the overall averages
overall_avg_data = pd.DataFrame({
    'Day Type': ['Weekends', 'Weekdays'],
    'Average Daily Sales': [overall_avg_weekend_sales, overall_avg_weekday_sales]
})

plt.figure(figsize=(8, 5))
sns.barplot(x='Day Type', y='Average Daily Sales', data=overall_avg_data, palette=['skyblue', 'lightcoral'])
plt.title('Rata-rata Harian Penjualan Weekends vs Weekdays (Oct-Dec 2022)')
plt.ylabel('Rata-rata Penjualan Harian (Before Discount)')
plt.show()
```

RESULT PYTHON

Number 4.2

Overall, during the three-month campaign period (October–December 2022), the comparison of average daily sales between weekends and weekdays is as follows:

```
Rata-rata Harian Penjualan Weekends vs Weekdays Keseluruhan 3 Bulan (Oct-Dec 2022):
Rata-rata Harian Penjualan Weekends: 5269300.00
Rata-rata Harian Penjualan Weekdays: 7520249.29
Secara keseluruhan selama 3 bulan, rata-rata penjualan harian di weekends lebih rendah dari weekdays.
```



INSIGHT, KEY VALUE, KEY KEY POINT

- **Insight:** Consistent with the monthly analysis, the overall average daily sales during the campaign period show that weekdays have higher average sales than weekends. Weekend campaigns do not appear to have had a significant positive impact on increasing overall average daily sales.
- **Key Value:** The average daily weekend sales were **5,269,300.00**, while the average daily weekday sales were **7,520,249.29**.
- **Key Point:** The weekend campaigns during **October–December 2022** failed to raise the average weekend daily sales to match or surpass the average weekday daily sales. The campaign team needs to evaluate the effectiveness of these weekend campaigns.

C-6

OVERALL RESULT

Google Colab Link

THANKYOU