

PHÂN TÍCH HIỆU SUẤT HỆ THỐNG CỦA CHUỖI CỬA HÀNG ABC TOYS

Đề bài: TS. Hồ Trọng Long

Thực hiện:

Trần Thị Anh Thư

Trần Công Khanh



NỘI DUNG

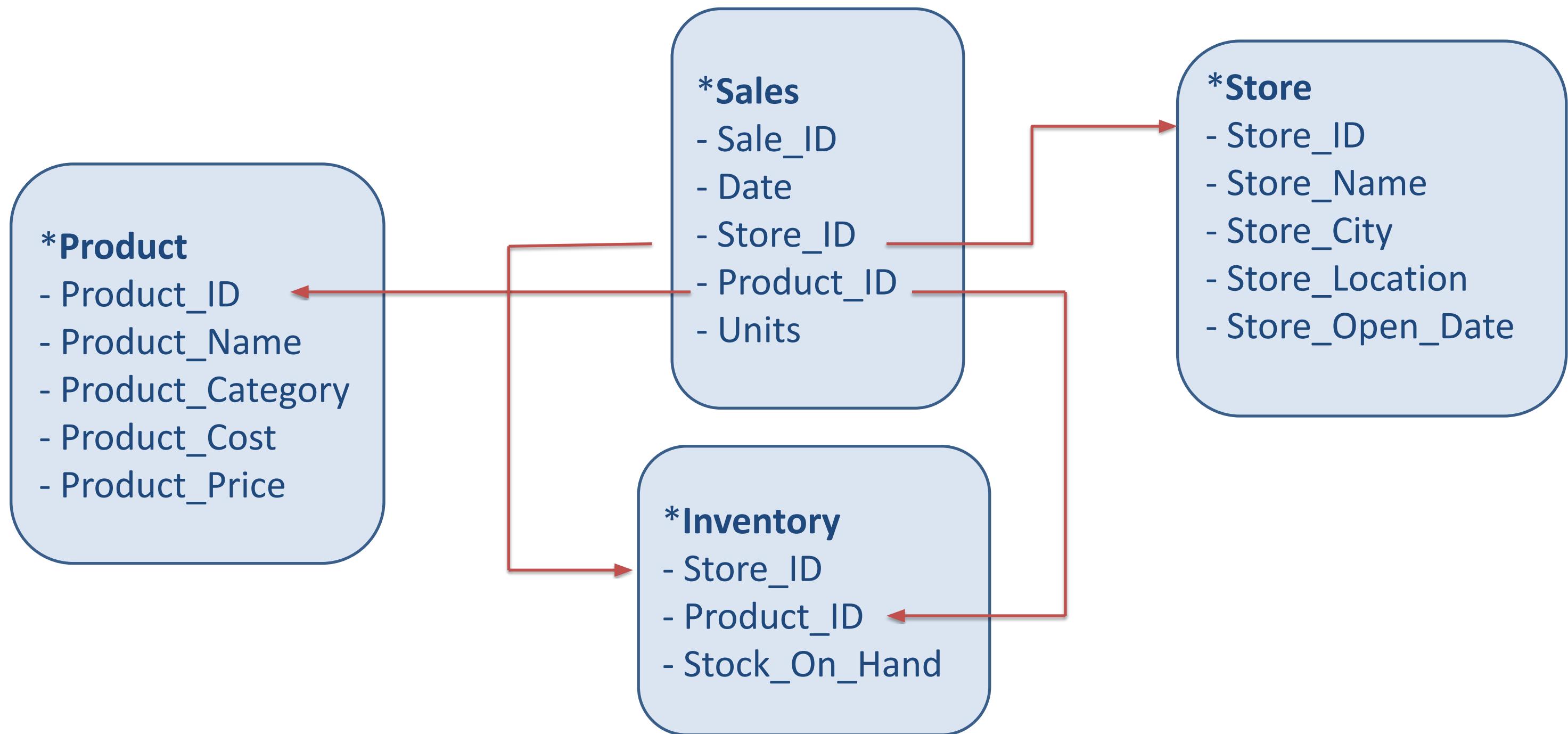
1. Giới thiệu (Sơ đồ ERD)
2. Tổng quan dữ liệu & đặt câu hỏi
3. Xử lý dữ liệu
4. Biểu đồ trực quan hóa (Visualization)
5. Khám phá Insight
6. Tổng kết

GIỚI THIỆU - SƠ ĐỒ ERD

Bộ data có thể là dữ liệu bán hàng của một chuỗi cửa hàng đồ chơi (ABC Toy), với các nhóm thông tin chính:

- Thông tin giao dịch (bảng Sales) : Sale_ID, Date, Units
- Thông tin sản phẩm (bảng Product): Product_ID, Product_Name, Product_Category, Cost, Price
- Thông tin cửa hàng (bảng Stores): Store_ID, Store_City, State, Region, Competitors, Population...
- Thông tin kho (bảng Inventory): Store_ID, Product_ID, Stock_On_Hand

GIỚI THIỆU - SƠ ĐỒ ERD



TỔNG QUAN DỮ LIỆU

Column

- Sale_ID
- Date
- Store_ID
- Product_ID
- Units
- Product_Name
- Product_Category
- Product_Cost
- Product_Price
- Store_Name
- Store_City
- Store_Location
- Store_Open_Date
- Stock_On_Hand
- Mã giao dịch
- Ngày bán
- Mã cửa hàng
- Mã sản phẩm
- Số lượng bán
- Tên sản phẩm
- Danh mục sản phẩm
- Giá vốn
- Giá bán
- Tên cửa hàng
- Thành phố
- Khu vực/Vị trí cửa hàng
- Ngày mở cửa
- Tồn kho hiện có





XỬ LÝ DỮ LIỆU

Table Product:

Bảng Product: có 35 hàng và 5 cột

Dữ liệu của Product_Cost và Product_Price: Đang dạng string -> chuyển về float (Số tiền)

```
products.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 35 entries, 0 to 34
Data columns (total 5 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   Product_ID    35 non-null    int64  
 1   Product_Name   35 non-null    object  
 2   Product_Category 35 non-null    object  
 3   Product_Cost    35 non-null    object  
 4   Product_Price   35 non-null    object  
dtypes: int64(1), object(4)
memory usage: 1.5+ KB
```

```
products.head()
```

	Product_ID	Product_Name	Product_Category	Product_Cost	Product_Price
0	1	Action Figure	Toys	\$9.99	\$15.99
1	2	Animal Figures	Toys	\$9.99	\$12.99
2	3	Barrel O' Slime	Art & Crafts	\$1.99	\$3.99
3	4	Chutes & Ladders	Games	\$9.99	\$12.99
4	5	Classic Dominoes	Games	\$7.99	\$9.99

```
# chuyển cost sang float
```

```
products["Product_Cost"] = products["Product_Cost"].str.replace("$", "").astype(float)
products["Product_Price"] = products["Product_Price"].str.replace("$", "").astype(float)
```

```
products.head()
```

	Product_ID	Product_Name	Product_Category	Product_Cost	Product_Price
0	1	Action Figure	Toys	9.99	15.99
1	2	Animal Figures	Toys	9.99	12.99
2	3	Barrel O' Slime	Art & Crafts	1.99	3.99
3	4	Chutes & Ladders	Games	9.99	12.99
4	5	Classic Dominoes	Games	7.99	9.99



XỬ LÝ DỮ LIỆU

Table Sales & Store:

Dữ liệu của Store_Open_Date (bảng Store) và Date (bảng Sales): Xử lý chuyển về Datatype Datetimes

```
stores.head()  
✓ 0.0s
```

	Store_ID	Store_Name	Store_City	Store_Location	Store_Open_Date
0	1	Maven Toys Guadalajara 1	Guadalajara	Residential	1992-09-18
1	2	Maven Toys Monterrey 1	Monterrey	Residential	1995-04-27
2	3	Maven Toys Guadalajara 2	Guadalajara	Commercial	1999-12-27
3	4	Maven Toys Saltillo 1	Saltillo	Downtown	2000-01-01
4	5	Maven Toys La Paz 1	La Paz	Downtown	2001-05-31

```
sales.head()  
✓ 0.0s
```

	Sale_ID	Date	Store_ID	Product_ID	Units
0	1	2022-01-01	24	4	1
1	2	2022-01-01	28	1	1
2	3	2022-01-01	6	8	1
3	4	2022-01-01	48	7	1
4	5	2022-01-01	44	18	1

```
sales.info()  
✓ 0.1s
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 829262 entries, 0 to 829261  
Data columns (total 5 columns):  
 # Column Non-Null Count Dtype  
---  
 0 Sale_ID 829262 non-null int64  
 1 Date    829262 non-null object  
 2 Store_ID 829262 non-null int64  
 3 Product_ID 829262 non-null int64  
 4 Units   829262 non-null int64  
dtypes: int64(4), object(1)  
memory usage: 31.6+ MB
```

```
stores.info()  
✓ 0.0s
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 50 entries, 0 to 49  
Data columns (total 5 columns):  
 # Column Non-Null Count Dtype  
---  
 0 Store_ID 50 non-null int64  
 1 Store_Name 50 non-null object  
 2 Store_City 50 non-null object  
 3 Store_Location 50 non-null object  
 4 Store_Open_Date 50 non-null object  
dtypes: int64(1), object(4)  
memory usage: 2.1+ KB
```



XỬ LÝ DỮ LIỆU

Table Sales & Store:

Dữ liệu của Store_Open_Date (bảng Store) và Date (bảng Sales): Xử lý chuyển về Datatype Datetimes

```
#Chuyển các cột ngày sang datetime
sales["Date"] = pd.to_datetime(sales["Date"])
stores["Store_Open_Date"] = pd.to_datetime(stores["Store_Open_Date"])
✓ 0.3s
```

```
sales.info()
✓ 0.0s
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 829262 entries, 0 to 829261
Data columns (total 5 columns):
 #   Column      Non-Null Count  Dtype  
---  --          -----          ----- 
 0   Sale_ID     829262 non-null  int64  
 1   Date        829262 non-null  datetime64[ns] 
 2   Store_ID    829262 non-null  int64  
 3   Product_ID  829262 non-null  int64  
 4   Units       829262 non-null  int64  
dtypes: datetime64[ns](1), int64(4)
memory usage: 31.6 MB
```

```
stores.info()
✓ 0.0s
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 50 entries, 0 to 49
Data columns (total 5 columns):
 #   Column      Non-Null Count  Dtype  
---  --          -----          ----- 
 0   Store_ID    50 non-null    int64  
 1   Store_Name  50 non-null    object  
 2   Store_City  50 non-null    object  
 3   Store_Location 50 non-null  object  
 4   Store_Open_Date 50 non-null  datetime64[ns] 
dtypes: datetime64[ns](1), int64(1), object(3)
memory usage: 2.1+ KB
```



XỬ LÝ DỮ LIỆU

Table Inventory:

Dữ liệu của Stock_on_hand:

```
inventory.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1593 entries, 0 to 1592
Data columns (total 3 columns):
 #   Column           Non-Null Count  Dtype  
---  --  
 0   Store_ID         1593 non-null   int64  
 1   Product_ID       1593 non-null   int64  
 2   Stock_On_Hand    1593 non-null   int64  
dtypes: int64(3)
memory usage: 37.5 KB
```



XỬ LÝ DỮ LIỆU

=> Thực hiện Merge các bảng để tiếp tục xử lý

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 829262 entries, 0 to 829261
Data columns (total 14 columns):
 #   Column           Non-Null Count  Dtype  
 --- 
 0   Sale_ID          829262 non-null   int64  
 1   Date             829262 non-null   datetime64[ns]
 2   Store_ID         829262 non-null   int64  
 3   Product_ID       829262 non-null   int64  
 4   Units            829262 non-null   int64  
 5   Product_Name     829262 non-null   object  
 6   Product_Category 829262 non-null   object  
 7   Product_Cost     829262 non-null   float64 
 8   Product_Price    829262 non-null   float64 
 9   Store_Name       829262 non-null   object  
 10  Store_City       829262 non-null   object  
 11  Store_Location   829262 non-null   object  
 12  Store_Open_Date  829262 non-null   datetime64[ns]
 13  Stock_On_Hand   823030 non-null   float64 
dtypes: datetime64[ns](2), float64(3), int64(4), object(5)
memory usage: 88.6+ MB
```

```
# Đếm số giá trị duy nhất check outlier
```

```
numeric_columns = list(df.select_dtypes(['number']).columns)
```

```
print(numeric_columns)
```

```
# đếm số giá trị numeric duy nhất
```

```
for column in numeric_columns:
```

```
    print("*, Column: ", column)
```

```
    print(len(df[column].unique()), "unique values", "\t & \t",
```

```
        df[column].notnull().sum(), "non-null values\t",
```

```
        round(100* df[column].notnull().sum()/len(df[column]),2), "% non-null")
```

```
['Sale_ID', 'Store_ID', 'Product_ID', 'Units', 'Product_Cost', 'Product_Price', 'Stock_On_Hand']
```

```
*, Column: Sale_ID
```

```
829262 unique values & 829262 non-null values 100.0 % non-null
```

```
*, Column: Store_ID
```

```
50 unique values & 829262 non-null values 100.0 % non-null
```

```
*, Column: Product_ID
```

```
35 unique values & 829262 non-null values 100.0 % non-null
```

```
*, Column: Units
```

```
27 unique values & 829262 non-null values 100.0 % non-null
```

```
*, Column: Product_Cost
```

```
16 unique values & 829262 non-null values 100.0 % non-null
```

```
*, Column: Product_Price
```

```
18 unique values & 829262 non-null values 100.0 % non-null
```

```
*, Column: Stock_On_Hand
```

```
94 unique values & 823030 non-null values 99.25 % non-null
```

```
df.isnull().sum()
```

Sale_ID	0
Date	0
Store_ID	0
Product_ID	0
Units	0
Product_Name	0
Product_Category	0
Product_Cost	0
Product_Price	0
Store_Name	0
Store_City	0
Store_Location	0
Store_Open_Date	0
Stock_On_Hand	6232
dtype: int64	

⇒ Cột Stock_On_Hand có 6232 dữ liệu null, chiếm 0,75%

⇒ Giải pháp: Loại bỏ những dữ liệu này



XỬ LÝ DỮ LIỆU

=> Thực hiện: Xoá dữ liệu Null cột Stock_On_Hand

```
# Xoá dữ liệu khoảng >6000 dòng trên tổng hơn 820k dữ liệu  
df= df.dropna(subset = 'Stock_On_Hand')
```

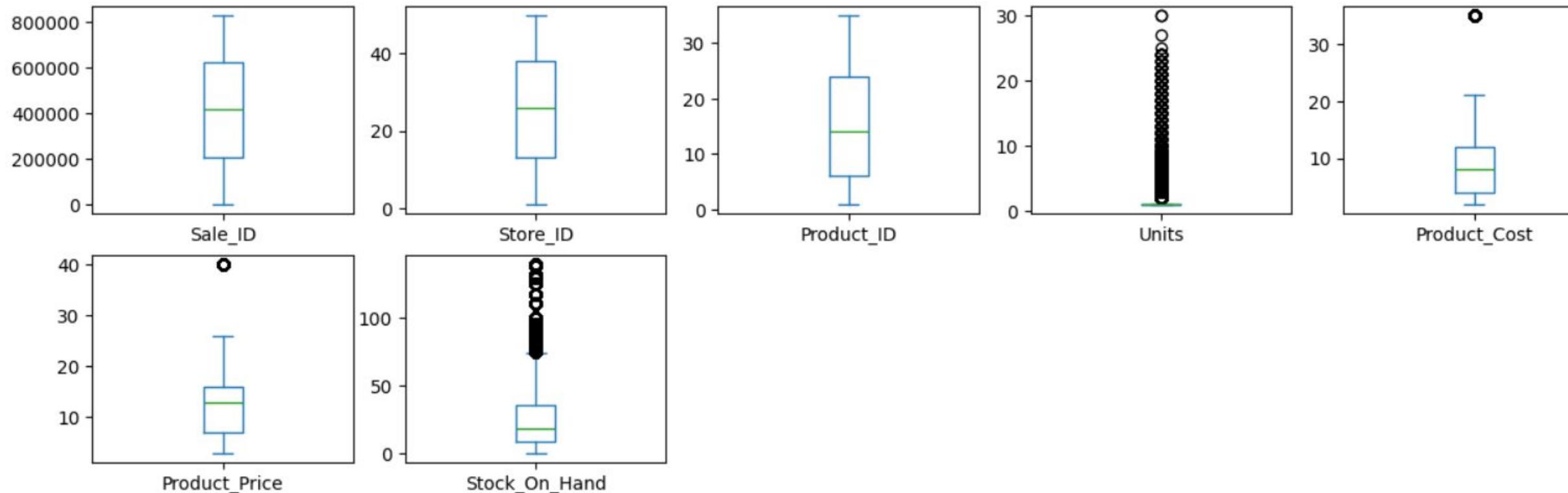
```
df.isnull().sum()
```

```
Sale_ID          0  
Date            0  
Store_ID        0  
Product_ID      0  
Units           0  
Product_Name    0  
Product_Category 0  
Product_Cost    0  
Product_Price   0  
Store_Name      0  
Store_City      0  
Store_Location   0  
Store_Open_Date  0  
Stock_On_Hand   0  
dtype: int64
```



XỬ LÝ DỮ LIỆU

```
plt.show()
```



```
[100]
```

	Sale_ID	Store_ID	Product_ID	Units	Product_Cost	Product_Price	Stock_On_Hand
count	823030.000000	823030.000000	823030.000000	823030.000000	823030.000000	823030.000000	823030.000000
mean	414855.482199	25.267391	15.005944	1.316952	10.026130	13.797757	25.319432
std	239754.243075	14.358599	9.902840	0.833016	7.824972	8.690296	22.955019
min	1.000000	1.000000	1.000000	1.000000	1.990000	2.990000	0.000000
25%	206841.250000	13.000000	6.000000	1.000000	3.990000	6.990000	9.000000
50%	415225.500000	26.000000	14.000000	1.000000	7.990000	12.990000	18.000000
75%	622638.750000	38.000000	24.000000	1.000000	13.990000	15.990000	35.000000
max	829262.000000	50.000000	35.000000	30.000000	34.990000	39.990000	139.000000

- ⇒ Tạo thêm cột doanh thu và Month để check outliers có thể xuất hiện
- ⇒ Price và Cost là giá sản phẩm -> có thể hợp lý, check outlier ở Units



XỬ LÝ DỮ LIỆU

Python

```

df['Revenue'] = df['Units'] * df['Product_Price']
df['Date'] = pd.to_datetime(df['Date'])
df['Month'] = df['Date'].dt.to_period('M').dt.to_timestamp()
df.describe()

```

2.0s

	Sale_ID	Date	Store_ID	Product_ID	Units	Product_Cost	Product_Price	Store_Open_Date	Stock_On_Hand	Revenue	Month
count	823030.000000	823030	823030.000000	823030.000000	823030.000000	823030.000000	823030.000000	823030	823030.000000	823030.000000	823030
mean	414855.482199	2022-12-11 01:14:08.335540736	25.267391	15.005944	1.316952	10.026130	13.797757	2009-08-21 07:57:07.486239744	25.319432	17.466261	2022-11-26 06:22:56.943246336
min	1.000000	2022-01-01 00:00:00	1.000000	1.000000	1.000000	1.990000	2.990000	1992-09-18 00:00:00	0.000000	2.990000	2022-01-01 00:00:00
25%	206841.250000	2022-07-08 00:00:00	13.000000	6.000000	1.000000	3.990000	6.990000	2006-08-30 00:00:00	9.000000	8.990000	2022-07-01 00:00:00
50%	415225.500000	2022-12-28 00:00:00	26.000000	14.000000	1.000000	7.990000	12.990000	2010-09-15 00:00:00	18.000000	14.990000	2022-12-01 00:00:00
75%	622638.750000	2023-05-14 00:00:00	38.000000	24.000000	1.000000	13.990000	15.990000	2014-03-18 00:00:00	35.000000	19.990000	2023-05-01 00:00:00
max	829262.000000	2023-09-30 00:00:00	50.000000	35.000000	30.000000	34.990000	39.990000	2016-05-18 00:00:00	139.000000	879.780000	2023-09-01 00:00:00
std	239754.243075	Nan	14.358599	9.902840	0.833016	7.824972	8.690296	Nan	22.955019	15.566935	Nan

```

# đếm số dòng Units trên 10
count = (df['Units'] > 10).sum()
print(count)

```

214

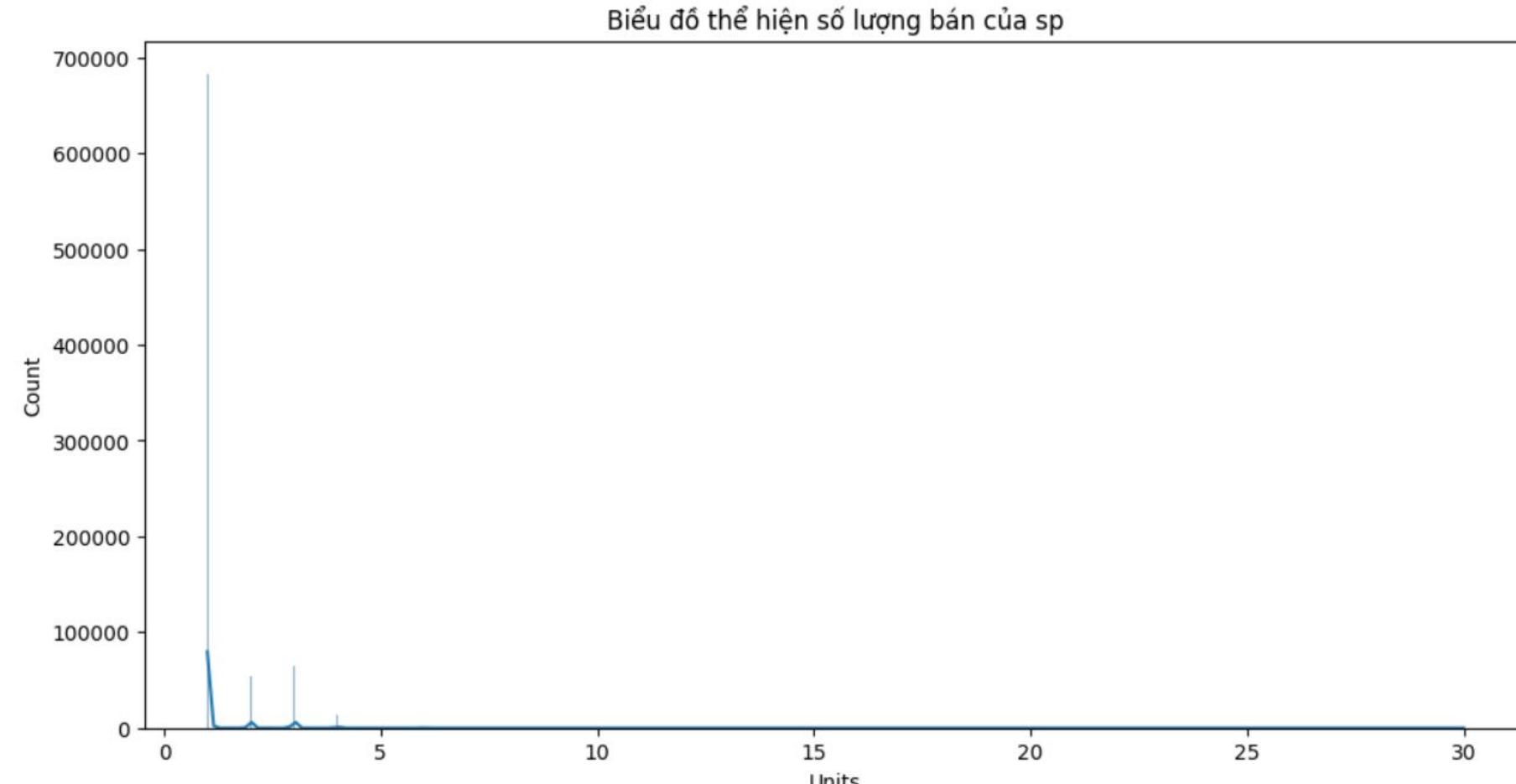
```

# Xử lý dữ liệu (loại thêm outlier trong Units)
# vẽ biểu đồ KDE của Units để xác định ko lại Units theo IQR

plt.figure(figsize=(12, 6))
sns.histplot(data =df , x = 'Units', kde=True )
plt.title('Biểu đồ thể hiện số lượng bán của sp')

# -> Số lượng Units = 1 quá nhiều -> loại giá trị trên 10

```





XỬ LÝ DỮ LIỆU

⇒ Dữ liệu hoàn chỉnh:

```
df.info()
```

✓ 0.1s

```
<class 'pandas.core.frame.DataFrame'>
Index: 822816 entries, 0 to 829261
Data columns (total 16 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Sale_ID          822816 non-null   int64  
 1   Date              822816 non-null   datetime64[ns]
 2   Store_ID          822816 non-null   int64  
 3   Product_ID        822816 non-null   int64  
 4   Units             822816 non-null   int64  
 5   Product_Name      822816 non-null   object 
 6   Product_Category  822816 non-null   object 
 7   Product_Cost      822816 non-null   float64
 8   Product_Price     822816 non-null   float64
 9   Store_Name         822816 non-null   object 
 10  Store_City        822816 non-null   object 
 11  Store_Location    822816 non-null   object 
 12  Store_Open_Date   822816 non-null   datetime64[ns]
 13  Stock_On_Hand    822816 non-null   float64
 14  Revenue           822816 non-null   float64
 15  Month             822816 non-null   datetime64[ns]
dtypes: datetime64[ns](3), float64(4), int64(4), object(5)
memory usage: 106.7+ MB
```

```
df.head()
```

✓ 0.0s

Sale_ID	Date	Store_ID	Product_ID	Units	Product_Name	Product_Category	Product_Cost	Product_Price	Store_Name	Store_City	Store_Location	Store_Open_Date	Stock_On_Hand
1	2022-01-01	24	4	1	Chutes & Ladders	Games	9.99	12.99	Maven Toys Aguascalientes 1	Aguascalientes	Downtown	2010-07-31	1.0
2	2022-01-01	28	1	1	Action Figure	Toys	9.99	15.99	Maven Toys Puebla 2	Puebla	Downtown	2011-04-01	6.0
3	2022-01-01	6	8	1	Deck Of Cards	Games	3.99	6.99	Maven Toys Mexicali 1	Mexicali	Commercial	2003-12-13	50.0
4	2022-01-01	48	7	1	Dart Gun	Sports & Outdoors	11.99	15.99	Maven Toys Saltillo 2	Saltillo	Commercial	2016-03-23	28.0
5	2022-01-01	44	18	1	Lego Bricks	Toys	34.99	39.99	Maven Toys Puebla 3	Puebla	Residential	2014-12-27	117.0



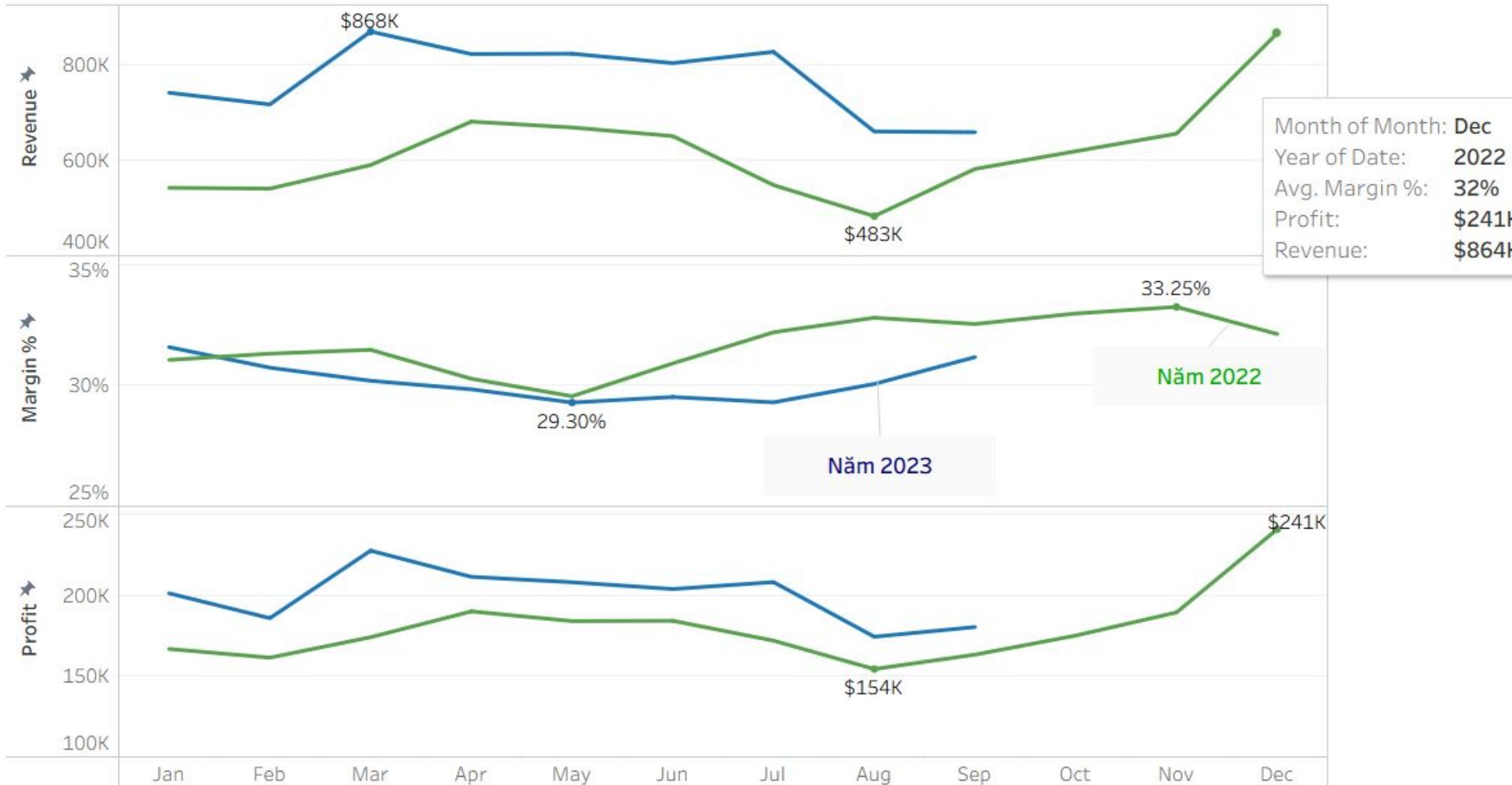
KHÁM PHÁ INSIGHT:

- Tình hình kinh doanh chung:
 - +Doanh thu & lợi nhuận tăng hay giảm
 - +Biên lợi nhuận như thế nào, lời hay lỗ
 - +Ảnh hưởng yếu tố mùa vụ, ngày trong tuần
- Tình hình kinh doanh theo từng cửa hàng
 - +Doanh thu, lợi nhuận có ảnh hưởng từ vị trí của cửa hàng hay không?
(xét khu vực có đông dân hay không)
=> Với những khu vực có doanh thu tốt (như airport) doanh số như thế nào, có nên mở rộng và tăng số lượng cửa hàng hay không?
 - +Top những cửa hàng có tình hình kinh doanh tốt nhất và kém nhất có đặc điểm gì (làm ảnh hưởng đến tình hình chung)? → Giải pháp
- Tình hình kinh doanh theo từng sản phẩm
 - +Danh mục nào bán chạy/bán kém → có đặc điểm gì? tổng quan chung? ⇒ giải pháp cho bán kém
 - +Có những sản phẩm kỳ vọng bán được, nhập giá thành cao → units thấp (Chính sách giá nào có lợi nhuận hiệu quả nhất? (giá cao – biên tốt / giá thấp – bán nhanh)
 - +Tình hình tồn kho: sản phẩm nào đang tồn kho nhiều? sản phẩm nào hết hàng? đề xuất giải pháp

KHÁM PHÁ INSIGHT

Tình hình chung

Year of Date	Revenue	Profit	Store ID	Store City	Product Cate..	Product ID
2022	7,417,428	2,154,018	50	29	5	33
2023	6,912,836	1,799,598	50	29	5	35

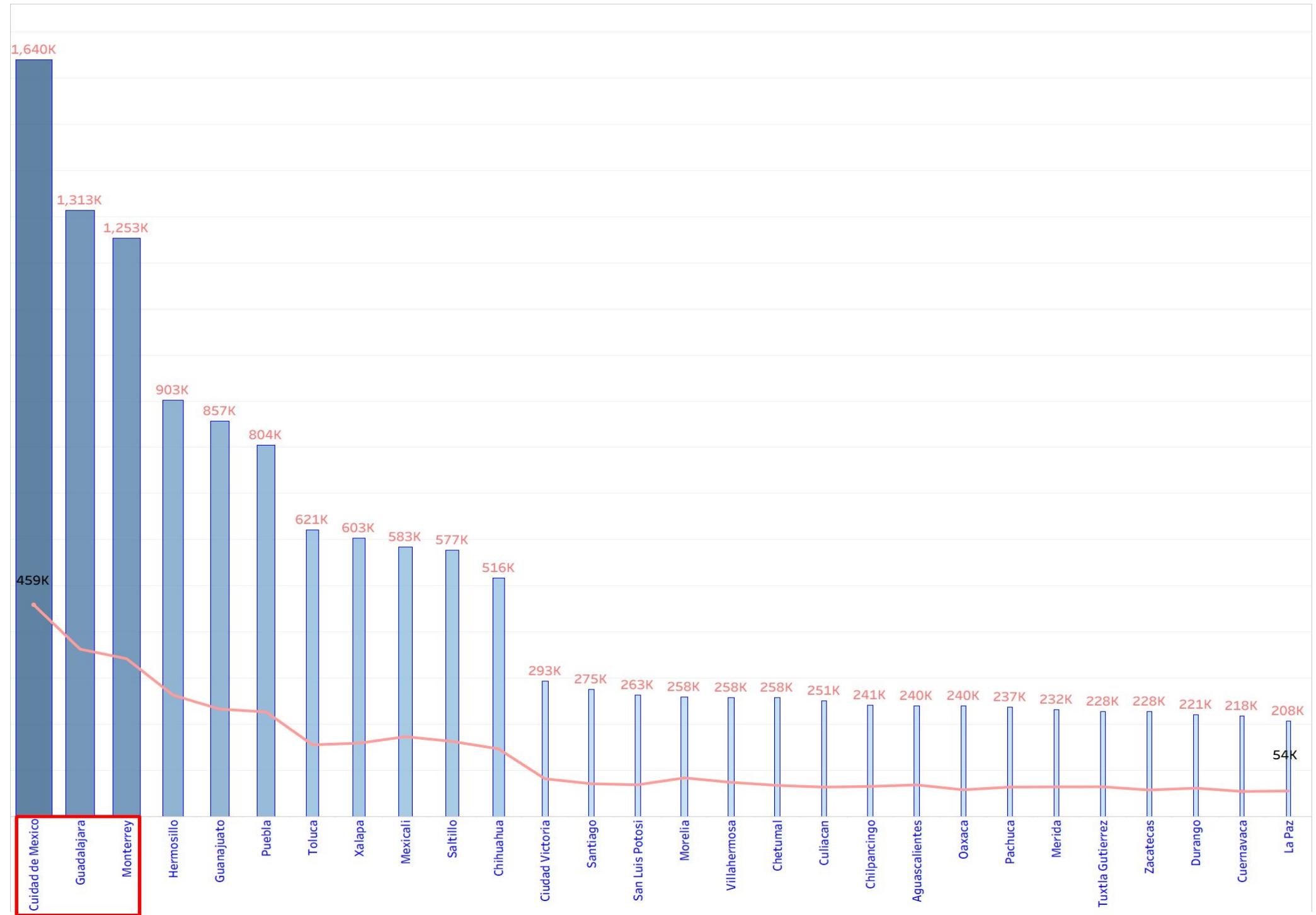


Nhận xét: Năm 2023

- Tất cả các cửa hàng đều kinh doanh có lợi nhuận
- Nhìn chung doanh thu và lợi nhuận tăng so với cùng kỳ năm ngoái
- Nhìn chung biên lợi nhuận % giảm nhẹ so với cùng kỳ năm ngoái
- Số sản phẩm tăng nhẹ (2 sản phẩm)

KHÁM PHÁ INSIGHT

Doanh thu theo thành phố



- 3 thành phố có doanh thu cao nhất: Ciudad de Mexico, Guadalajara, Monterrey → xem xét mở rộng cửa hàng hoặc thêm cửa hàng ở 3 thành phố này
- 3 thành phố có doanh thu kém nhất: La Paz, Cuernavaca, Durango → cần phân tích nguyên nhân để đề ra giải pháp thích hợp

KHÁM PHÁ INSIGHT

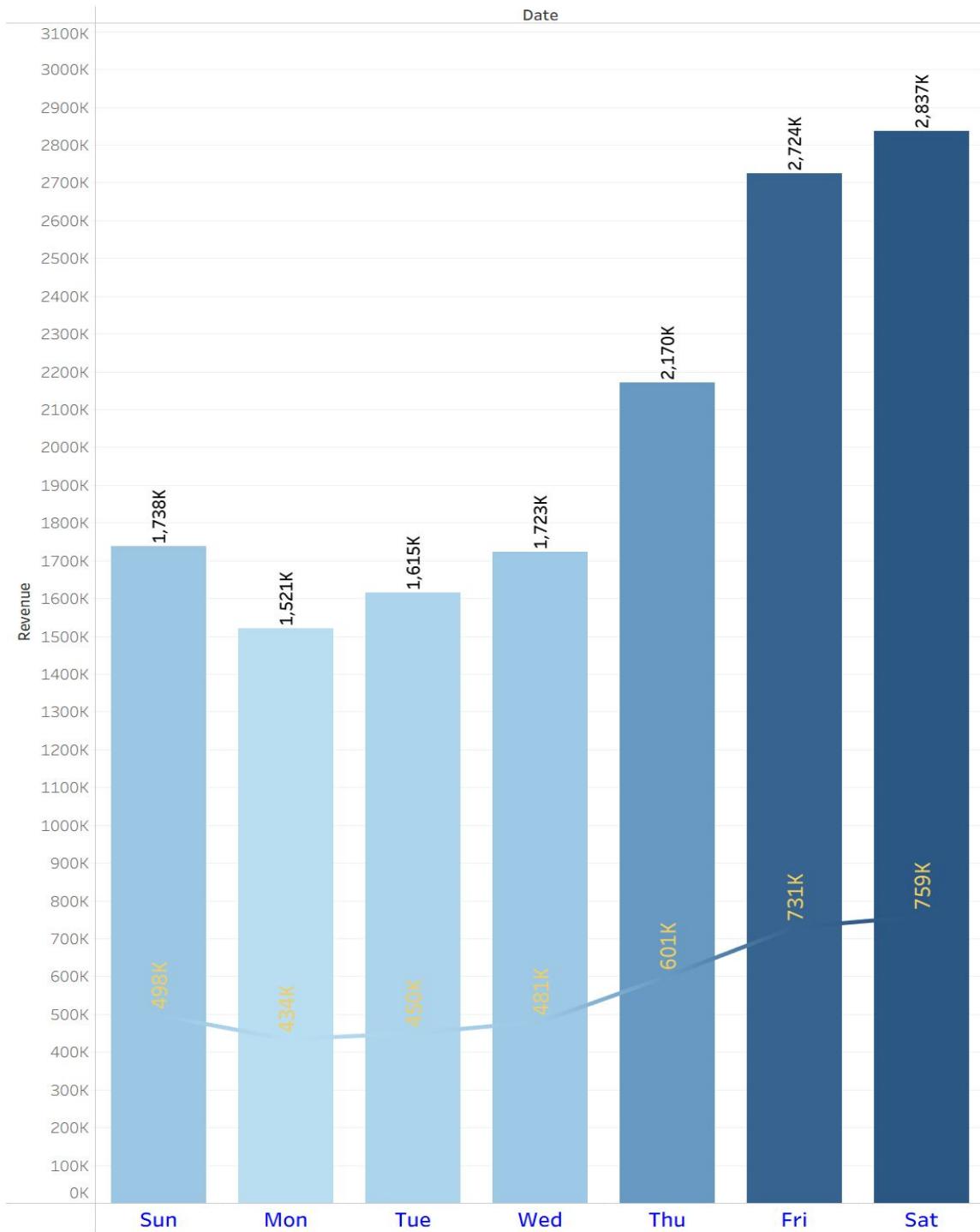
Doanh thu theo khu vực đặc thù

Downtown Ciudad de Mexico 9 \$433K	Downtown Toluca 17 \$404K	Downtown Monterrey 7 \$373K	Downtown Guadalajara 46 \$346K	Downtown Xalapa 39 \$340K	Commercial Hermosillo 42 \$345K	Commercial Guanajuato 22 \$279K	Commercial Xalapa 29 \$264K	Commercial Guadalajara 3 \$259K
Downtown Saltillo 4 \$330K	Downtown Hermosillo 41 \$323K	Downtown Guanajuato 14 \$312K	Downtown Campeche 10 \$305K	Downtown Ciudad Victoria 25 \$293K	Downtown Mexicali 13 \$290K	Commercial Cuidad de Mexico 45 \$322K	Commercial Puebla 19 \$259K	Commercial Toluca 40 \$217K
Downtown Puebla 28 \$279K	Downtown Morelia 36 \$258K	Downtown Chilpancingo 35 \$241K	Downtown Aguascalientes 24 \$240K	Downtown Oaxaca 27 \$240K	Commercial Monterrey 47 \$320K	Commercial Chihuahua 23 \$248K	Commercial Campeche 26 \$206K	Commercial Saltillo 48 \$246K
Downtown Santiago 21 \$275K	Downtown Villahermosa 34 \$258K	Downtown Pachuca 8 \$237K	Downtown Zacatecas 20 \$228K	Downtown Durango 43 \$221K	Residential Cuidad de Mexico 37 \$337K	Residential Puebla 44 \$266K	Residential Guadalajara 1 \$262K	Residential Monterrey 2 \$277K
Downtown Chihuahua 38 \$268K	Downtown Chetumal 12 \$258K	Downtown Merida 18 \$232K	Downtown Cuernavaca 11 \$218K	Downtown La Paz 5 \$208K	Airport Cuidad de Mexico 31 \$548K	Airport Guadalajara 30 \$446K	Airport Monterrey 33 \$284K	
Downtown San Luis Potosi 16 \$263K	Downtown Culiacan 49 \$251K	Downtown Tuxtla Gutierrez 15 \$228K						

- Các cửa hàng có doanh thu cao nhất nằm ở khu vực sân bay (ID 31, 30) → xem xét mở rộng cửa hàng hoặc thêm cửa hàng khu vực này
- Cần kiểm tra, tìm cách cải thiện các cửa hàng có doanh thu kém nhất như ID 5, 26, 40

KHÁM PHÁ INSIGHT

Doanh thu theo ngày trong tuần



- Doanh thu 2 ngày thứ 6, 7 là cao nhất, thứ 2 , 3 là thấp nhất → xem xét bố trí số lượng nhân viên nhiều vào cuối tuần, bố trí ngày nghỉ nhiều vào thứ 2, 3, 4.
- Thực hiện khuyến mãi nhiều vào thứ 2, 3, 4 để cải thiện doanh thu

KHÁM PHÁ INSIGHT

Doanh thu theo thời gian trong năm

	Art & Crafts	Electronics	Games	Sports & Outdoors	Toys
Jan	2,993	4,345	7,026	2,863	5,671
Feb	2,371	4,056	6,252	2,120	9,537
Mar	3,037	4,288	2,184	3,813	7,630
Apr	3,346	2,769	3,644	2,348	10,461
May	4,573	6,229	3,682	3,941	11,930
Jun	2,443	3,518	3,699	3,113	5,432
Jul	1,938	4,140	3,941	2,432	6,127
Aug	2,297	2,220	2,264	2,768	6,614
Sep	5,301	4,339	3,446	4,163	9,542
Oct	704	3,348	3,138	2,010	3,016
Nov	1,735	2,610	2,435	1,479	3,849
Dec	1,884	1,664	2,755	2,261	6,263

- Toys có doanh số cao nhất vào các tháng 4, 5 (trùng bắt đầu nghỉ hè), tháng 9 (trước khai giảng năm học mới)
- Electronics cao tháng 5, hơi cao vào mùa tháng 9 --> có thể do trùng thời gian bắt đầu nghỉ hè và khai giảng năm học mới
- Sport & outdoors và Art & Craft : mạnh vào tháng 5 và tháng 9--> có thể do trùng thời gian bắt đầu nghỉ hè và khai giảng năm học mới
- Game: mạnh nhất tháng 1 và 2, nguyên nhân có thể do nhiều voucher quà tặng dịp giáng sinh và tết dương lịch, kỳ nghỉ mùa đông, các game mới phát hành đầu năm...
- Do đó, có thể chủ động phân bổ số lượng sản phẩm theo mùa vụ tương ứng

KHÁM PHÁ INSIGHT

Tình hình kinh doanh theo từng cửa hàng



=>Nhận xét:

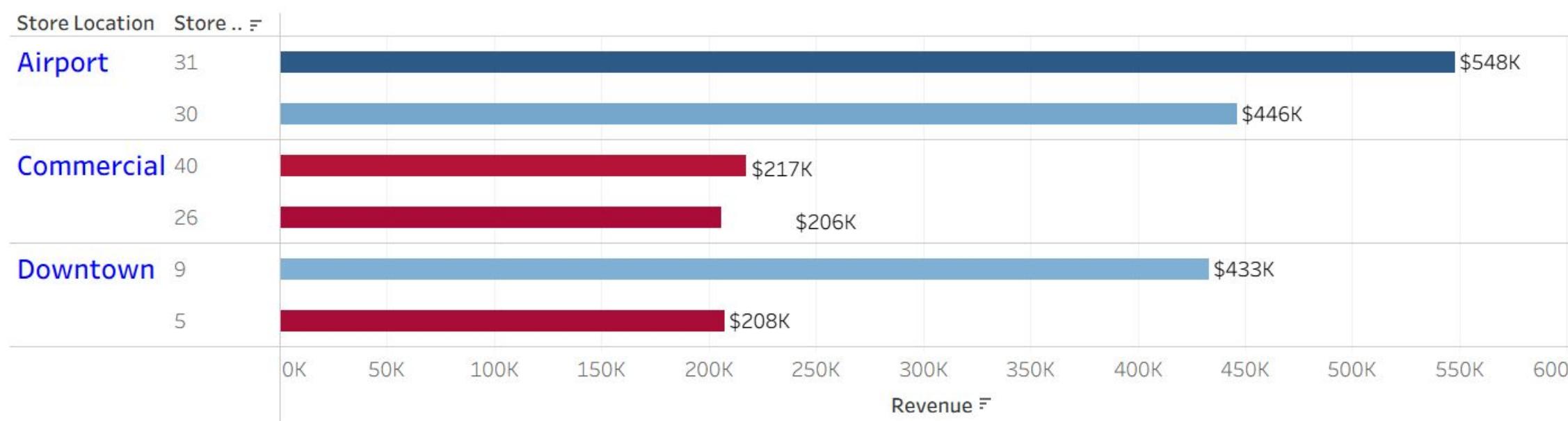
- + 47 cửa hàng có doanh thu tăng
- + 3 cửa hàng (Store 2, 24, 46)
- + Các cửa hàng có doanh thu tăng có lợi nhuận tương đối giống xu hướng với doanh thu
- + Biên lợi nhuận % có xu hướng giảm chung

KHÁM PHÁ INSIGHT

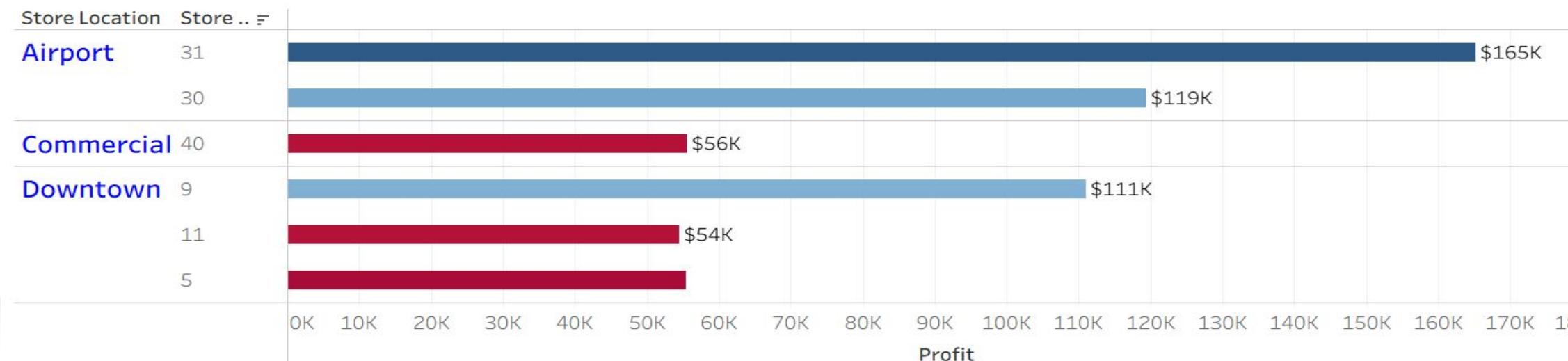
Top cửa hàng có doanh thu và lợi nhuận cao/thấp nhất trong khu vực

Store Location	No.Store	No.Product	Revenue	Avg_Revenue	Profit	Avg_Profit
Airport	3	35	1,278K	426K	370K	123K
Commercial	12	35	3,258K	271K	916K	76K
Downtown	29	35	8,152K	281K	2,211K	76K
Residential	6	35	1,643K	274K	457K	76K

Revenue in Store & Location



Profit in Store & Location



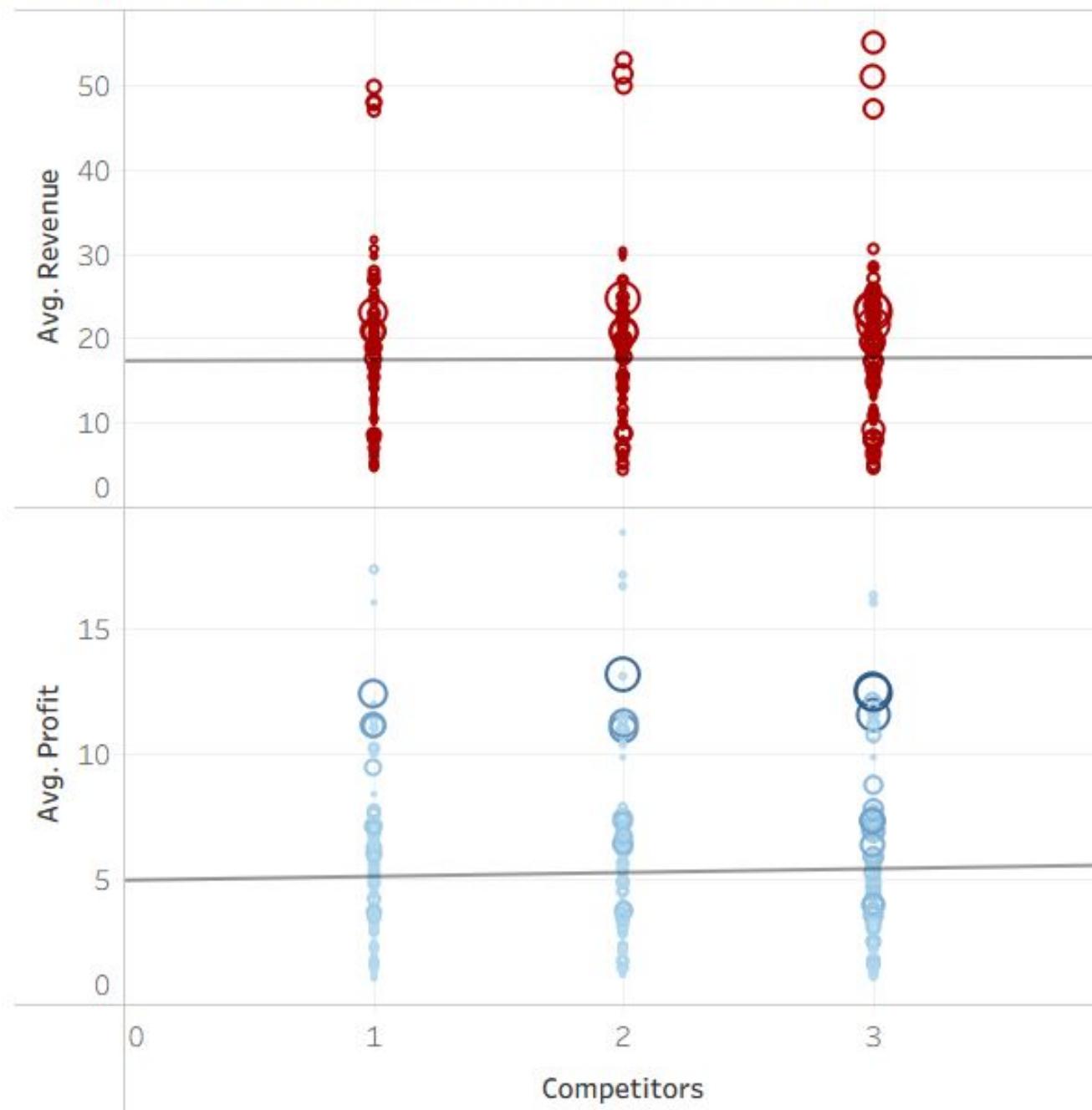
=>Nhận xét:

- + Top 2 cửa hàng có doanh thu cao nhất thuộc khu vực Airport (ID 30, 31)
- + Top 3 cửa hàng có doanh thu cao nhất thuộc khu vực Commercial & Downtown (ID 5, 26, 40)
- + Khu vực Residential có doanh thu trung bình (Top 10)

KHÁM PHÁ INSIGHT

Doanh thu, lợi nhuận có ảnh hưởng do số lượng competitors hay không?

Competitors vs Total Revenue/ Profit



Competitors là số cửa hàng khác (không tính cửa hàng hiện tại) trong cùng 1 thành phố

$$\text{Competitors} = \{ \text{FIXED } [\text{Store}_\text{City}] : \text{COUNTD}([\text{Store}_\text{ID}]) \} - 1$$

=> Nhận xét:

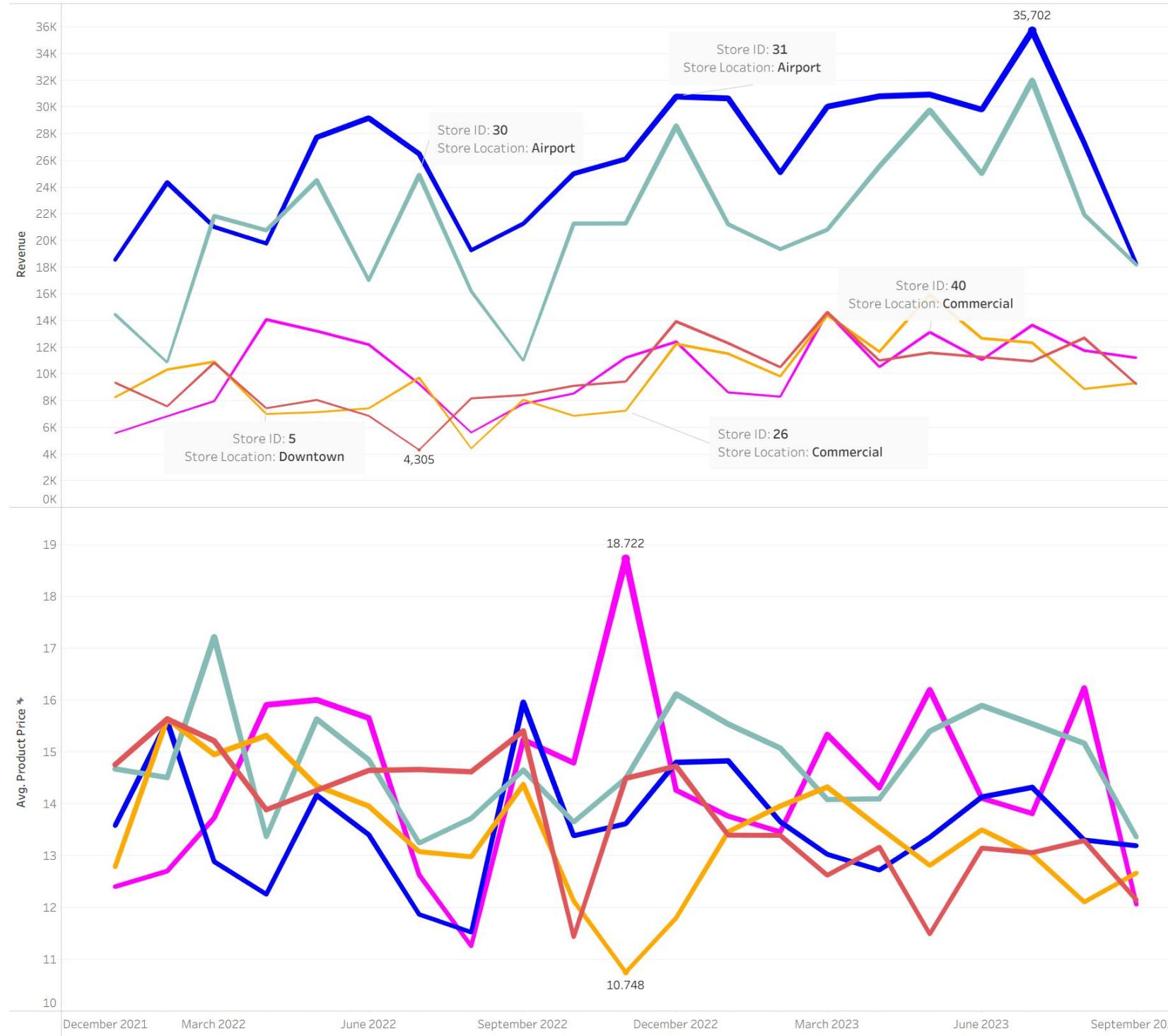
Competitors tăng nhưng trend line Avg doanh thu/lợi nhuận chỉ dốc nhẹ

=> Số lượng Competitors hiện tại **không ảnh hưởng** doanh thu/lợi nhuận , đồng thời **doanh thu cao hơn** khi có nhiều cửa hàng trong thành phố

=> Đề xuất: Mở rộng quy mô cửa hàng hoặc tăng số lượng cửa hàng ở những khu vực có doanh thu cao như sân bay...

KHÁM PHÁ INSIGHT

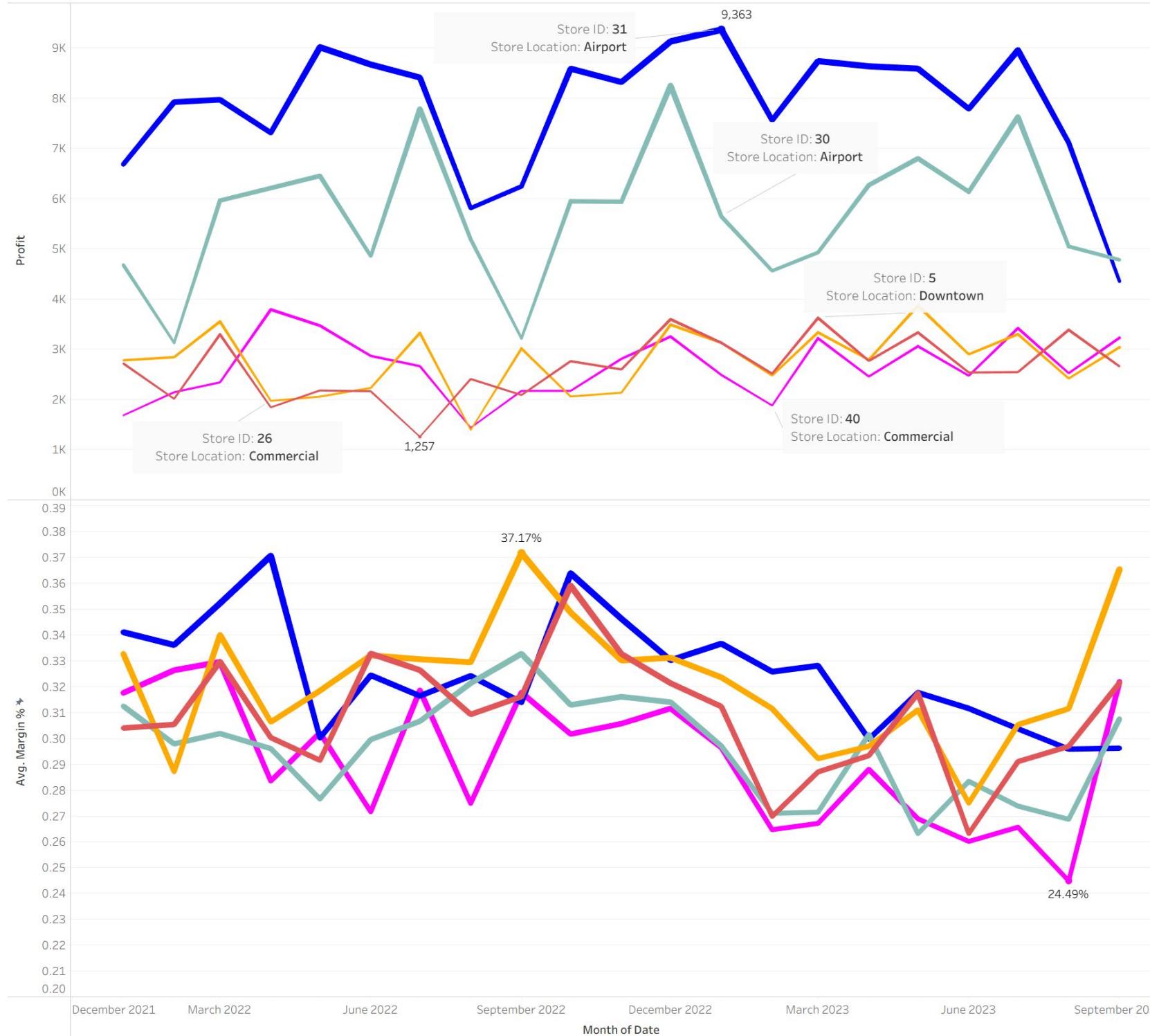
So sánh tình hình kinh doanh của 3 cửa hàng có doanh thu kém nhất với 2 cửa hàng tốt



- Top 2 cửa hàng kinh doanh tốt nhất: 31 và 30
- Top 3 cửa hàng kinh doanh kém nhất: 5, 26, 40
- Doanh số 3 cửa hàng bán kém thấp đều theo thời gian
- Giá bán trung bình 3 cửa hàng kém tương đương 2 cửa hàng tốt
- Vậy yếu tố thời gian, mùa vụ, giá không phải nguyên nhân bán kém

KHÁM PHÁ INSIGHT

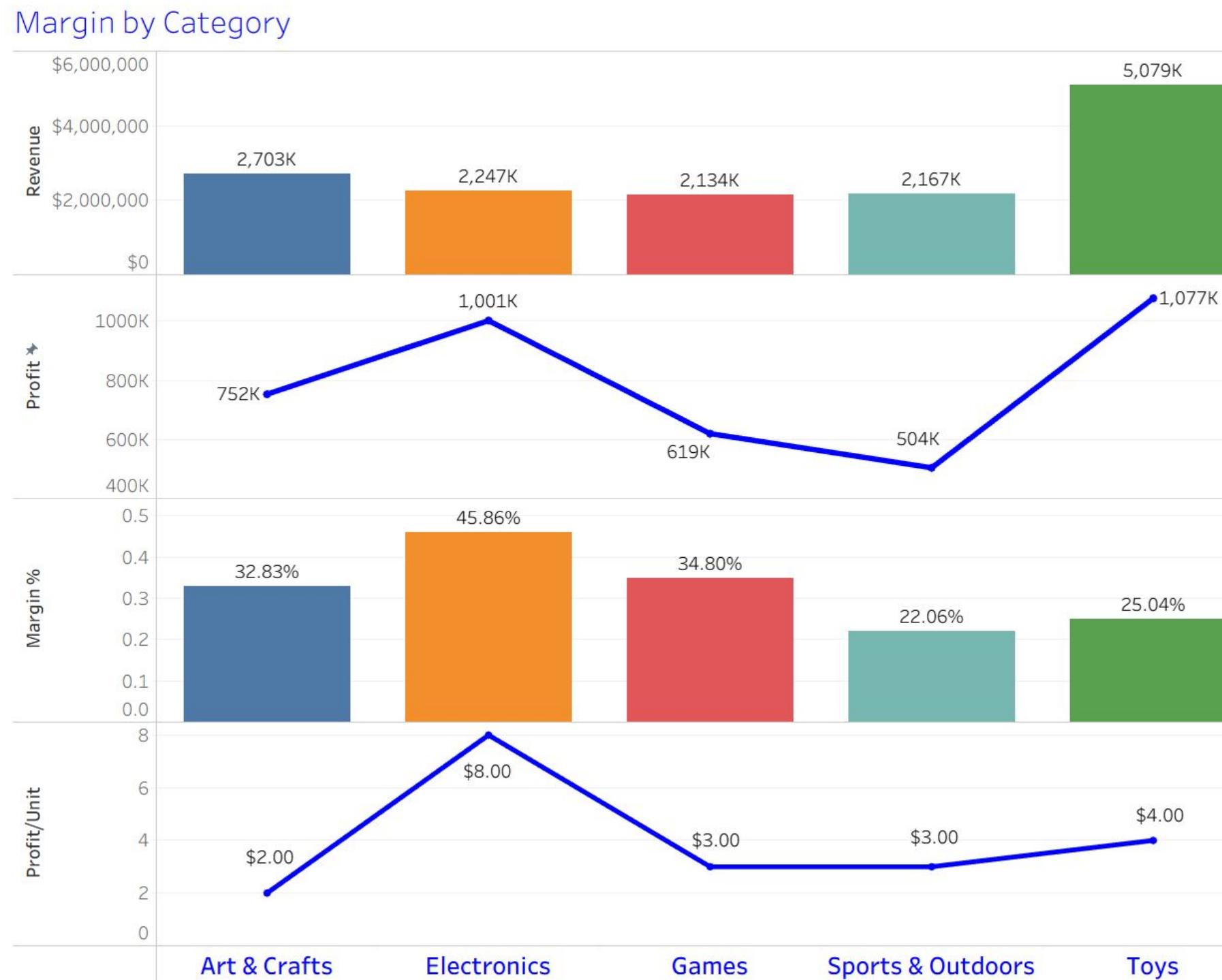
So sánh tình hình kinh doanh của 3 cửa hàng có doanh thu kém nhất với 2 cửa hàng tốt



- Lợi nhuận 3 cửa hàng bán kém thấp tương ứng với xu hướng chung của doanh thu
- Biên lợi nhuận cơ bản bằng giống nhau ở 5 cửa hàng.
- Vậy loại sản phẩm không phải nguyên nhân bán kém
- Kết luận:** 3 cửa hàng bán kém do số lượng bán nhỏ, dẫn đến doanh thu, lợi nhuận kém. Số lượng bán kém có thể kiểm tra và cải thiện những yếu tố sau:
 - Vị trí cửa hàng
 - Hình thức bố trí sản phẩm, trang trí cửa hàng
 - Thái độ phục vụ của nhân viên
 - Tổ chức khuyến mãi, giảm giá
 - Tăng cường marketing

KHÁM PHÁ INSIGHT

Tình hình kinh doanh theo từng loại sản phẩm

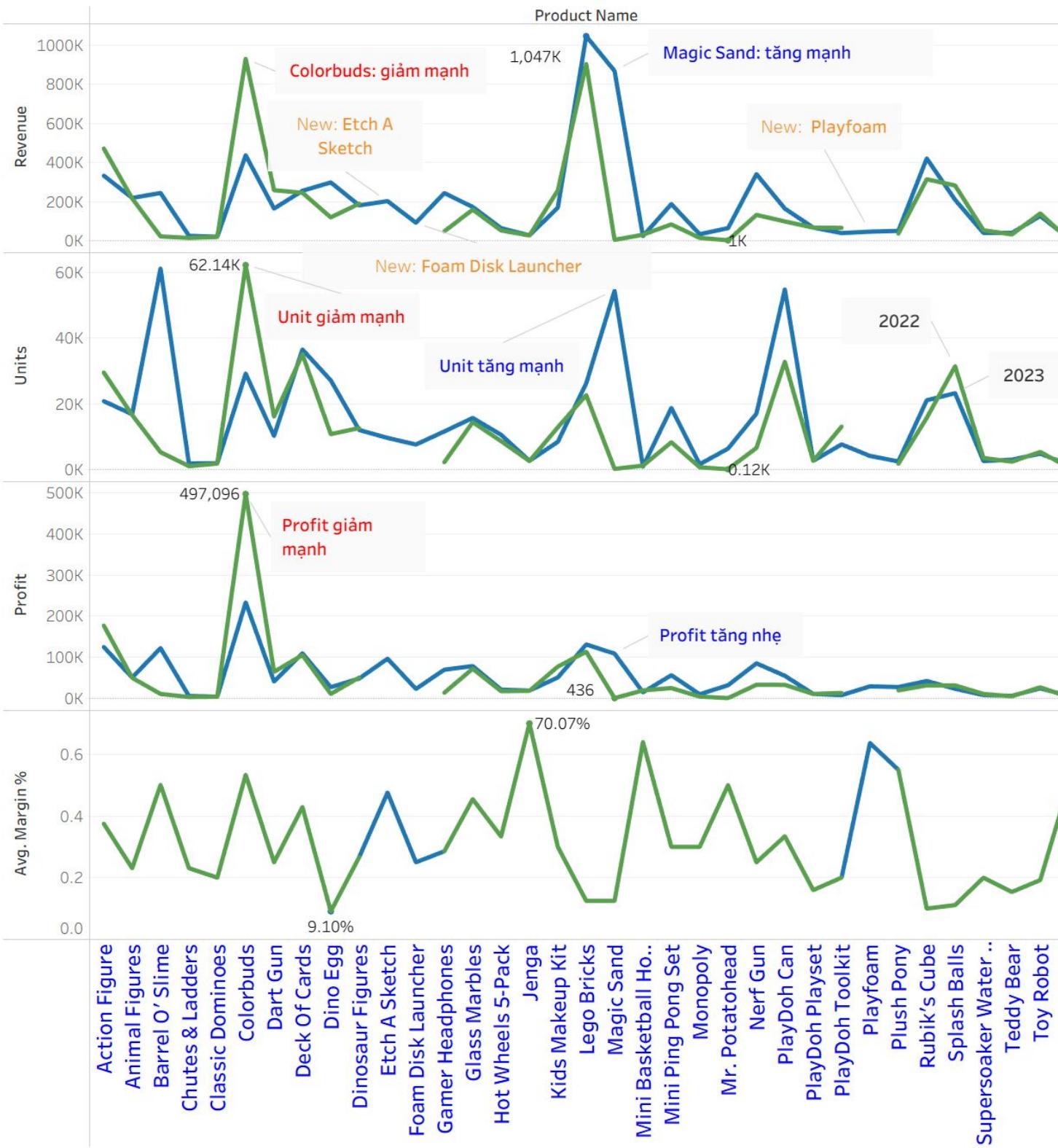


Nhận xét:

- Doanh thu cao: Toys
- Doanh thu thấp: Games, Sport & Outdoors and Electronics
- Toys biên lợi nhuận thấp (25%) nhưng doanh thu cao nên lợi nhuận cao nhất (1077k)
- Electronics doanh thu thấp nhưng biên lợi nhuận cao (45.86%) → đóng góp lợi nhuận cao (1001k)
- Sport & outdoors biên lợi nhuận kém nhất nên lợi nhuận ít nhất.
- Art & Crafts giá thấp, lợi nhuận thấp nhất (2\$), biên lợi nhuận cao (32.83%), bán được nhiều nên lợi nhuận cao
- **Kết luận:** xem xét tăng các mặt hàng Toys và Electronics để tăng lợi nhuận

KHÁM PHÁ INSIGHT

Tình hình kinh doanh theo từng sản phẩm



Nhận xét:

- Thời gian: tháng 1 tới tháng 9 trong 2 năm 2022 và 2023
- Có 3 sản phẩm mới
- Colorbuds giảm mạnh về doanh số (\$497k → \$233k). Ngược lại, Magic sand tăng mạnh về doanh số (\$3k → \$869k)
- Colorbuds có biên lợi nhuận cao (53.37%), chiếm tỉ trọng cao trong lợi nhuận tổng năm 2022. Lợi nhuận Magic Sand nhỏ hơn nhiều so với Colorbuds (chỉ 12.51%). Nên khi Colorbuds giảm mạnh sẽ kéo theo lợi nhuận tổng giảm, biên lợi nhuận giảm tương ứng trong năm 2023.
- Kết luận:** Colorbuds giảm mạnh có thể do Magic sand. Biên lợi nhuận tổng có xu hướng giảm tương ứng

KHÁM PHÁ INSIGHT

Lợi nhuận sản phẩm theo giá và doanh số



Nhận xét:

- Colorbuds thuộc nhóm giá trung bình, biên lợi nhuận tốt, số lượng bán tốt nên lợi nhuận cao nhất
- Lego Bricks giá cao, số lượng bán tốt nhưng biên lợi nhuận nhỏ nên lợi nhuận thuộc top 3
- Magic sand giá trung bình, số lượng bán tốt, nhưng biên lợi nhuận nhỏ nên lợi nhuận nhỏ
- Nhóm sản lợi nhuận thấp nhưng giá tương đối cao cần xem xét giảm: mini basketball hoop, PlayDo Playset, Monopoly, Plus pony

Kết luận:

- Giá trung bình, biên lợi nhuận cao có thể cho lợi nhuận tốt nhất do số lượng bán nhiều. Cần ưu tiên các sản phẩm như Colorbuds , Action figure, Deck of cards...
- Lego Bricks có doanh thu cao nhất, cần ưu tiên
- Xem xét giảm các sản phẩm như mini basketball hoop, PlayDo Playset, Monopoly, Plus pony

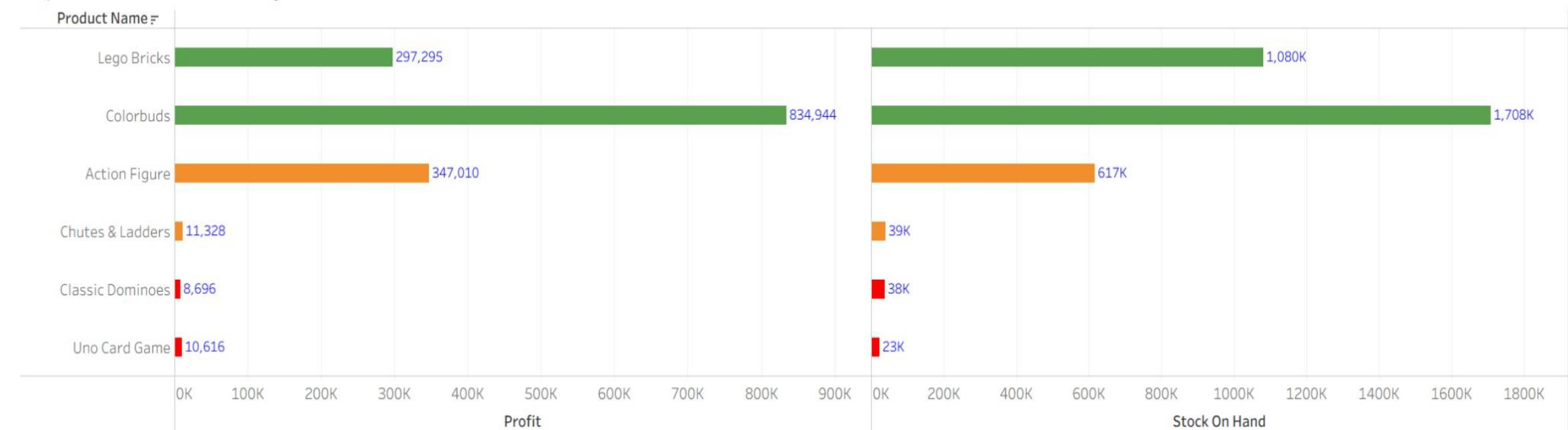
KHÁM PHÁ INSIGHT

Top sản phẩm tốt nhất, kém nhất

Top & Bottom Product by Revenue



Top & Bottom Product by Profit



Nhận xét:

- Colorbuds doanh thu tốt, lợi nhuận cao nhất, tồn kho cao, cần tăng cường bán
- Lego bricks doanh thu tốt và tồn kho trong top 3, xem xét bổ sung kho
- Sản phẩm doanh thu thấp, lợi nhuận kém, xem xét cắt giảm gồm: classic Dominoes, Uno Card Game

KHÁM PHÁ INSIGHT

Stock on hand



Nhận xét

- Sản phẩm tồn kho lớn: Deck of Cards, PlayDoh Can... xem xét lại giá, khuyến mãi
 - Sản phẩm tồn kho ít, margin tốt như: Playfoarm, Jenga... xem xét tăng SOH



KẾT LUẬN

Tổng quan:

- Tất cả cửa hàng đều kinh doanh có lợi nhuận
- Năm 2023, doanh số tăng, lợi nhuận tăng, nhưng biên lợi nhuận giảm so với cùng kì năm 2022 (do số lượng Colorbuds giảm)
- Khu vực kinh doanh tốt (như sân bay...) có thể xem xét mở rộng hoặc thêm cửa hàng mới
- Nhóm cửa hàng có doanh số kém do số lượng bán kém, các giải pháp cải thiện như: tăng cường hình ảnh, trang trí cửa hàng, tăng khuyến mãi, đào tạo nhân viên... (ID 5, 26, 40)
- Sản phẩm doanh thu thấp, lợi nhuận kém, xem xét cắt giảm gồm: classic Dominoes, Uno Card Game
- Nhóm Toys, electronics chiếm lợi nhuận cao nhất và doanh số cao nhất, cần tăng cường bán
- Đẩy mạnh nhóm hàng giá trung bình và biên lợi nhuận cao để tối ưu lợi nhuận (như Colorbuds , Action figure, Deck of cards)
- Sản phẩm tồn kho lớn: Deck of Cards, PlayDoh Can... xem xét lại giá, khuyến mãi
- Sản phẩm tồn kho ít, margin tốt như: Playfoarm, Jenga... xem xét tăng SOH
- Còn 3 tháng cuối năm tranh thủ bán Toys và thêm combo Art+ Carft để giảm lượng tồn kho => tối ưu hóa dòng tiền





LINK THỰC HIỆN CHI TIẾT:

- Demo Visualization trên Tableau:
- File Code thực hiện trên GitHub:
- Data sau khi xử lý trên Drive:
- Mô tả dự án chi tiết trên Drive:



THANK YOU