** 1.** **Importing and Loading Libraries**

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

2. Loading and Understanding the Data

```
df=pd.read_csv(r'/content/Snitch.csv')
df.head()
df.info()
df.dtypes
df.shape
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 15132 entries, 0 to 15131
Data columns (total 23 columns):
 #   Column         Non-Null Count  Dtype 
---  ------         --------------  ----- 
 0   SKU_GROUP      15132 non-null  object
 1   CATEGORY       15129 non-null  object
 2   OCCASSION_NEW  15132 non-null  object
 3   PRINT_DESIGN   13840 non-null  object
 4   COLLAR_NEW     15132 non-null  object
 5   MATERIAL_NEW   15132 non-null  object
 6   SLEEVE_TYPE    10852 non-null  object
 7   FIT            13836 non-null  object
 8   COLOR          15132 non-null  object
 9   CLOSURE        4458 non-null   object
 10  STYLE          2811 non-null   object
 11  MONTH1         15132 non-null  int64 
 12  MONTH2         15132 non-null  int64 
 13  MONTH3         15132 non-null  int64 
 14  MONTH4         15132 non-null  int64 
 15  MONTH5         15132 non-null  int64 
 16  MONTH6         15132 non-null  int64 
 17  MONTH7         15132 non-null  int64 
 18  MONTH8         15132 non-null  int64 
 19  MONTH9         15132 non-null  int64 
 20  MONTH10        15132 non-null  int64 
 21  MONTH11        15132 non-null  int64 
 22  MONTH12        15132 non-null  int64 
dtypes: int64(12), object(11)
memory usage: 2.7+ MB

(15132, 23)
```

3. Assumptions and Data Cleaning: To ensure consistency, column names were standardized. Missing values in feature columns were filled with "Unknown," and rows with missing or

"Unknown" categories were excluded from the analysis. A new column, TOTAL_SALES, was created by summing across all monthly columns, which were assumed to represent monthly sales figures.

```
# Standardizing column names
df.columns = df.columns.str.strip().str.upper().str.replace(" ", "_")


# In print_design column i have 1292 null values and i am removing
rows because except category and color feature all row has
# NaN values and 0 in all months from 1 to 12.
df[df['PRINT_DESIGN'].isna()]
df[df['PRINT_DESIGN'].isna()].index
df.drop(labels=df[df['PRINT_DESIGN'].isna()].index,inplace=True,errors
='ignore')

# In sleeve_type column we have NaN,i have filled with Unknown rather
than deleting
df['SLEEVE_TYPE'].fillna('Unknown',inplace=True)

# in category column i have 1 NaN, i have deleted because it will not
affect any rows because Month 1 to 12 all are 0.
df[df['CATEGORY'].isna()].index
df.drop(labels=df[df['CATEGORY'].isna()].index,inplace=True,errors='ig
nore')

# Checking FIT Column as it has 4 NaN. Deleting 4 rows as many
features has NaN values and also Month 1 to Month 12 all are 0.
df[df['FIT'].isna()].index
df.drop(labels=df[df['FIT'].isna()].index,inplace=True,errors='ignore'
)

# Now Closure and Style column has only NaN, i am filling with Unknown
in place of NaN so that data might not looks messy
df.fillna({'CLOSURE':'Unknown','STYLE':'Unknown'},inplace=True)

# Creating TOTAL_SALES column
month_cols = [col for col in df.columns if col.startswith('MONTH')]
df['TOTAL_SALES'] = df[month_cols].sum(axis=1)

# Creating a feature column
feature_columns = ['FIT', 'MATERIAL_NEW', 'PRINT_DESIGN',
'SLEEVE_TYPE', 'COLLAR_NEW','OCCASSION_NEW','COLOR']
```

```
<ipython-input-7-c68a0f9390bc>:12: FutureWarning: A value is trying to
be set on a copy of a DataFrame or Series through chained assignment
using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never
work because the intermediate object on which we are setting values
always behaves as a copy.
```

```
For example, when doing 'df[col].method(value, inplace=True)', try
using 'df.method({col: value}, inplace=True)' or df[col] =
df[col].method(value) instead, to perform the operation inplace on the
original object.


  df['SLEEVE_TYPE'].fillna('Unknown',inplace=True)
```

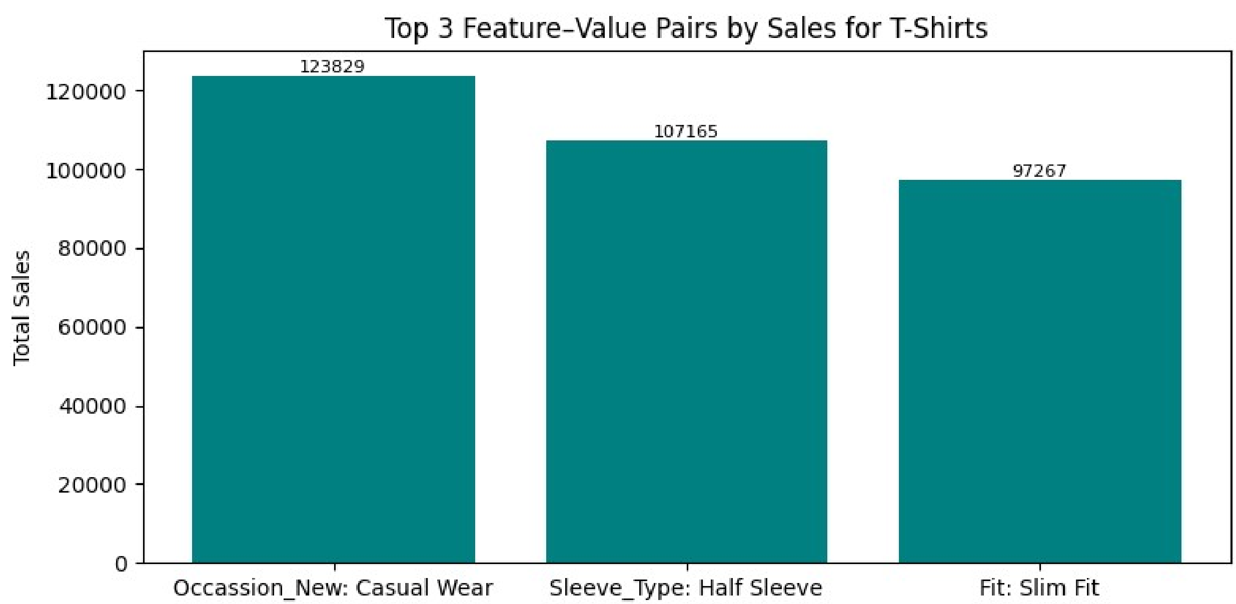
Q1. For each category, show which feature is most influential (highest total sales by feature value)

Approach -> For each product category, I evaluated all specified features to determine which individual feature values contributed the highest total sales. This allowed me to quantify the influence of each feature on overall sales performance. The results were then sorted and visualized using bar charts to highlight the top-performing features within each category.

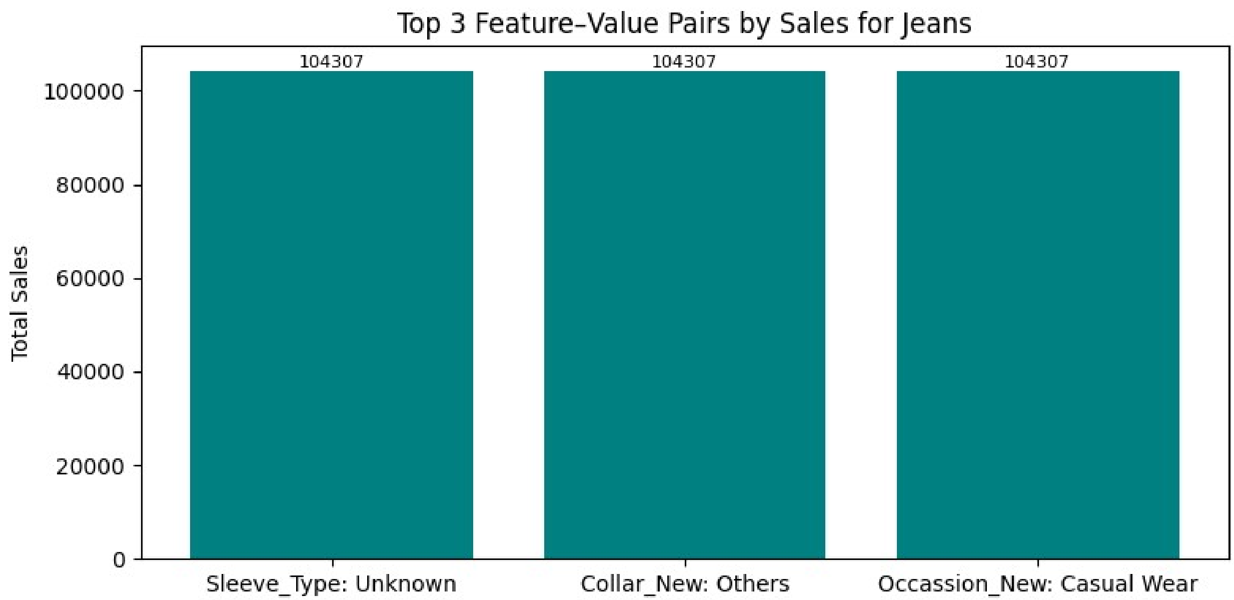
```
# Print available categories
print("Available categories:")
print(df['CATEGORY'].value_counts())

# Plot top 3 (feature, value) pairs per category by total sales
for category in df['CATEGORY'].unique():
    cat_df = df[df['CATEGORY'] == category]
    print(f"\nCategory: {category}, Rows: {cat_df.shape[0]}")

    value_sales = []
    for feature in feature_columns:
        if feature in cat_df.columns:
            sales_by_value = cat_df.groupby(feature)
['TOTAL_SALES'].sum()
            for value, total in sales_by_value.items():
                label = f"{feature.title()}: {value}"
                value_sales.append((label, total))

    # Sort and take top 3 feature-value pairs
    top3 = sorted(value_sales, key=lambda x: x[1], reverse=True)[:3]
    labels = [x[0] for x in top3]
    sales = [x[1] for x in top3]

    # Plot
    plt.figure(figsize=(8, 4))
    bars = plt.bar(labels, sales, color='teal')
    plt.title(f'Top 3 Feature–Value Pairs by Sales for {category}')
    plt.ylabel('Total Sales')

    # Annotate bars
    for bar in bars:
        yval = bar.get_height()
        plt.text(bar.get_x() + bar.get_width()/2.0, yval,
f'{int(yval)}', ha='center', va='bottom', fontsize=8)
```

```
    plt.tight_layout()
    plt.show()
```

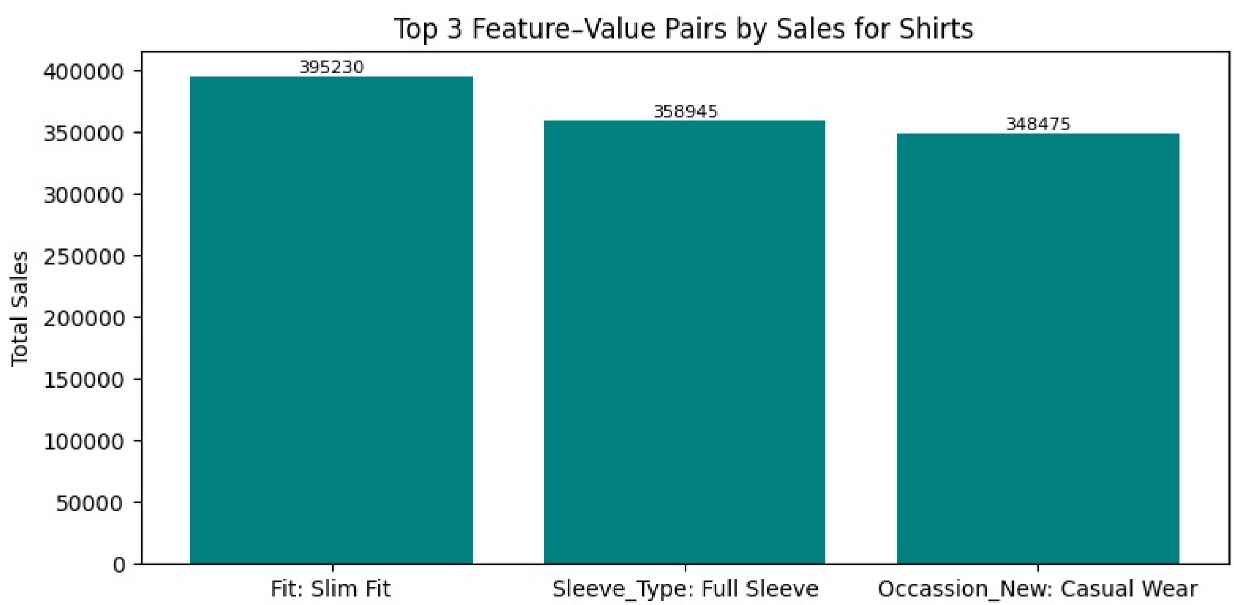
```
Available categories:
CATEGORY
Shirts                     8450
T-Shirts                   1485
Jeans                      1308
Trousers                   1057
Overshirt                   495
Cargo Pants                 429
Jackets                     246
Sweaters                    138
Joggers & Trackpants        107
Sweatshirts                  77
Pyjamas                      43
Name: count, dtype: int64

Category: T-Shirts, Rows: 1485
```



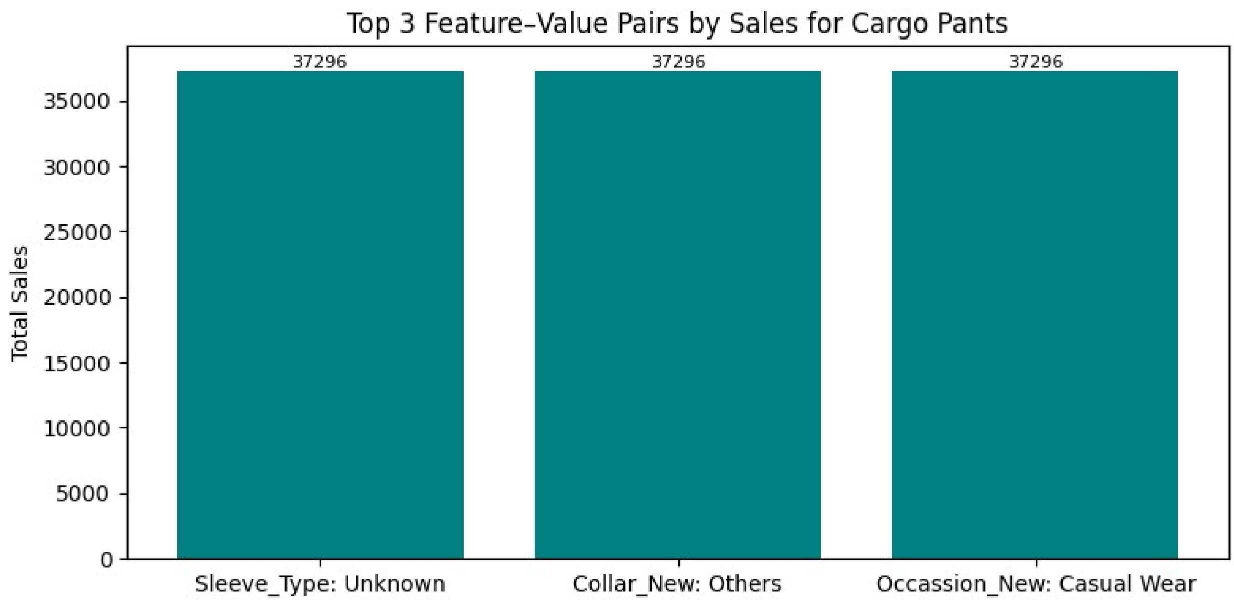

```
Category: Jeans, Rows: 1308
```



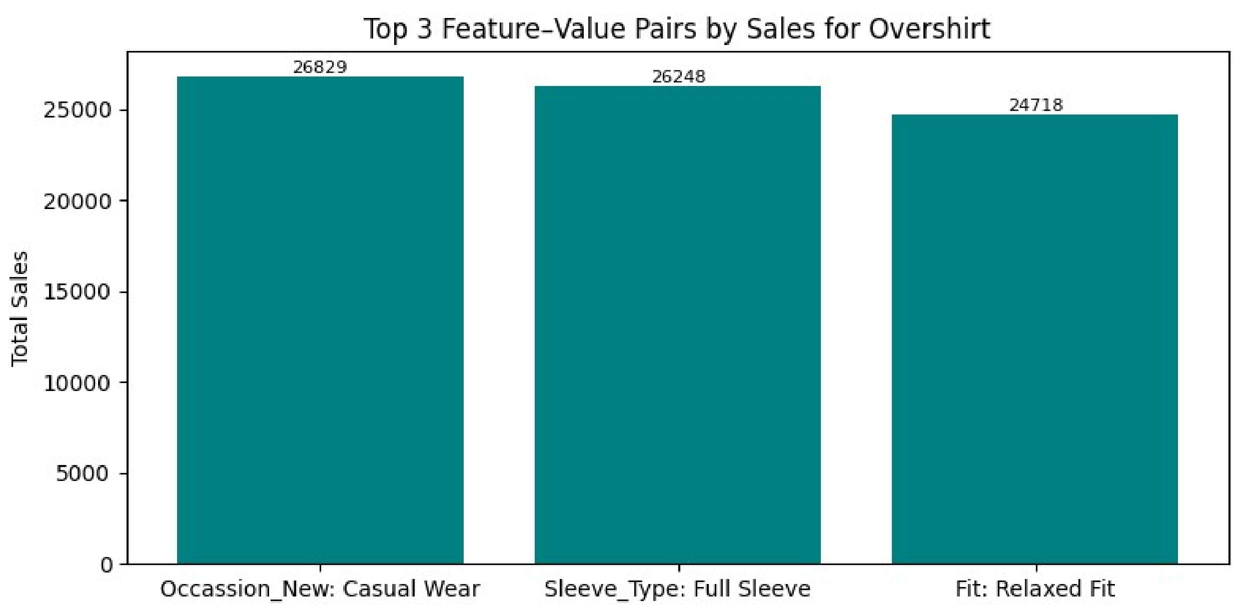

```
Category: Shirts, Rows: 8450
```



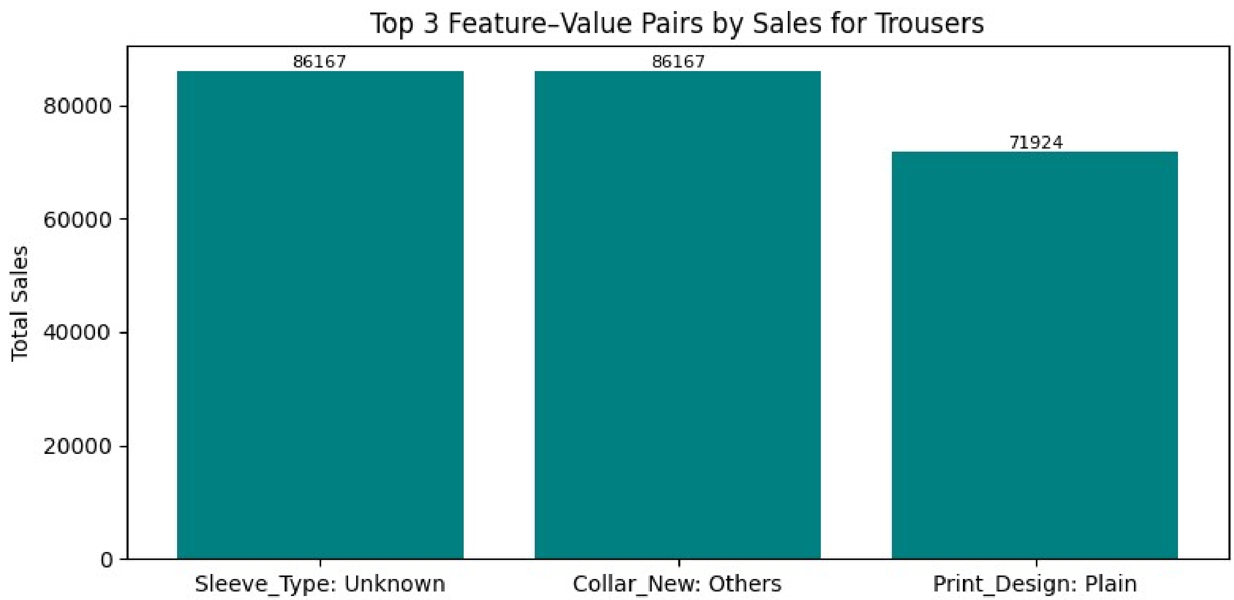

```
Category: Cargo Pants, Rows: 429
```



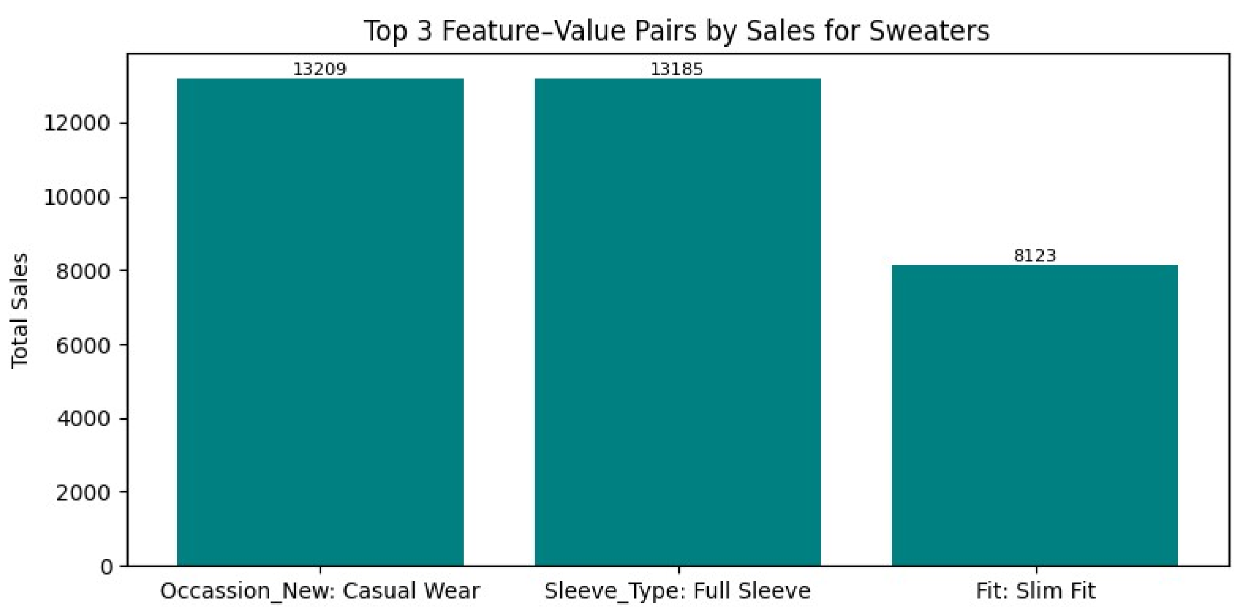

```
Category: Overshirt, Rows: 495
```



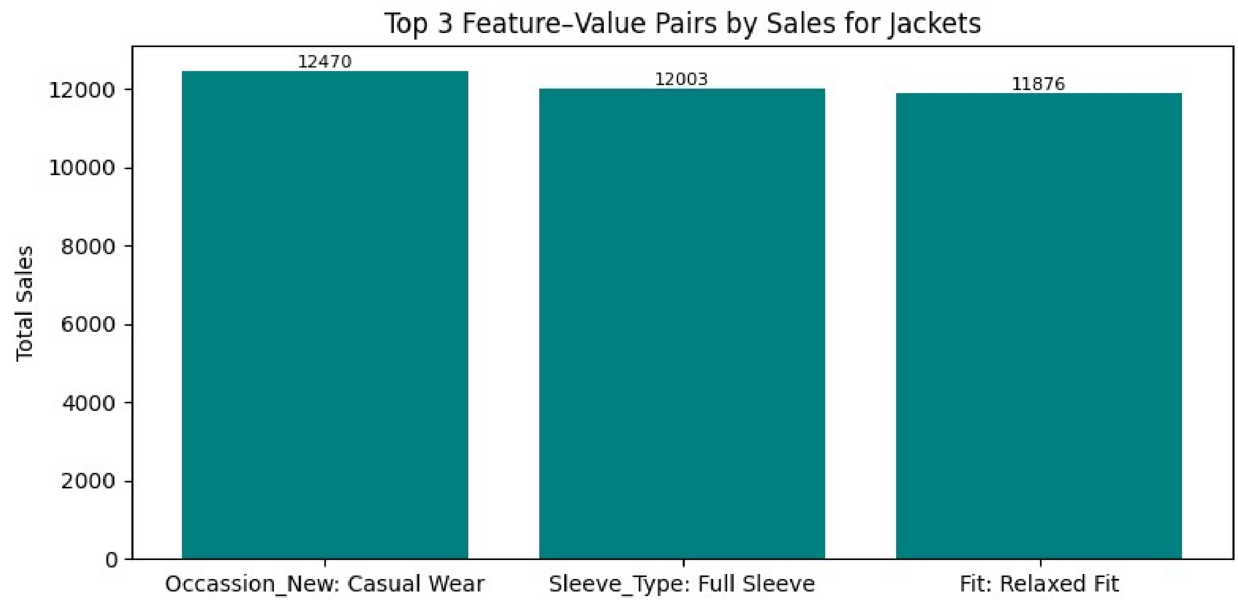

```
Category: Trousers, Rows: 1057
```



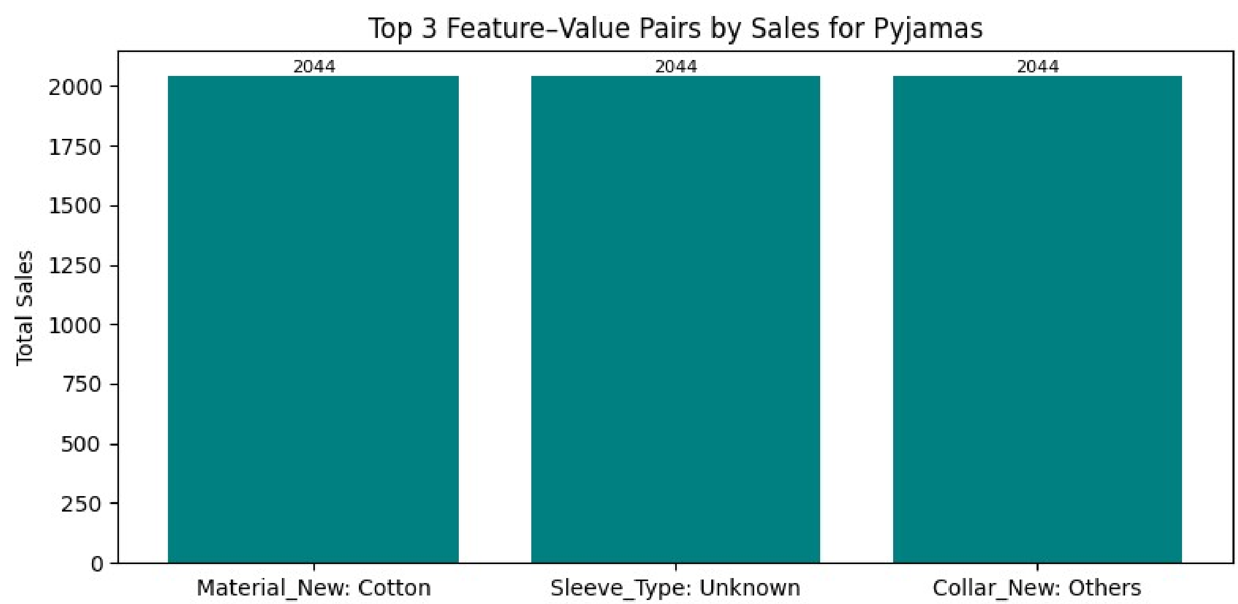

```
Category: Sweaters, Rows: 138
```



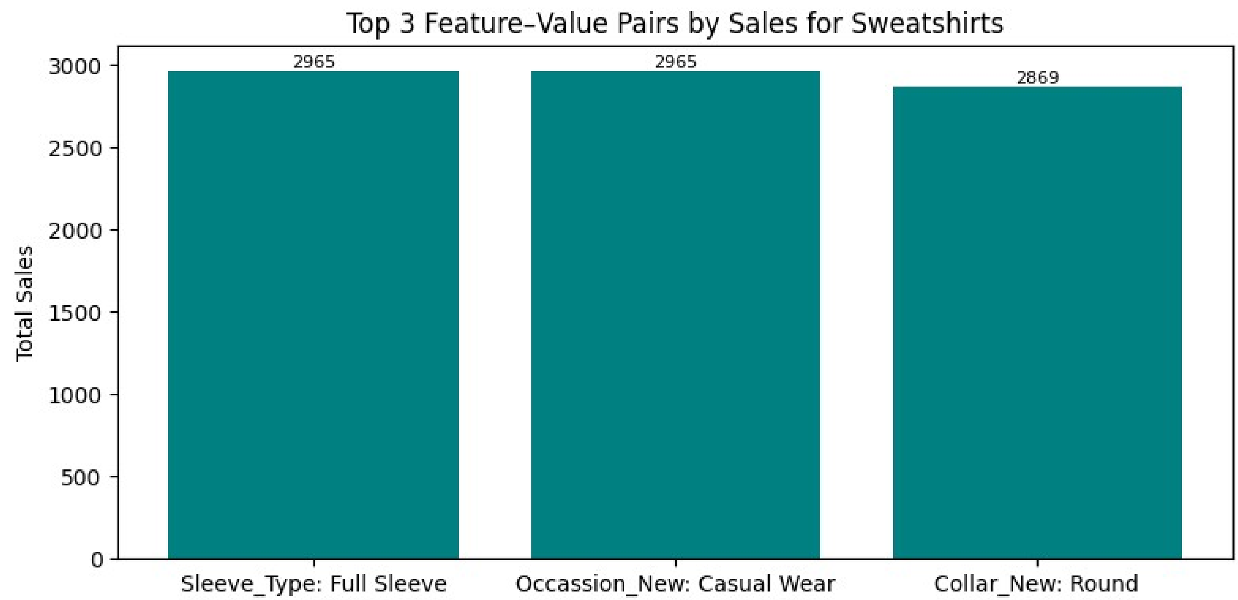

```
Category: Jackets, Rows: 246
```



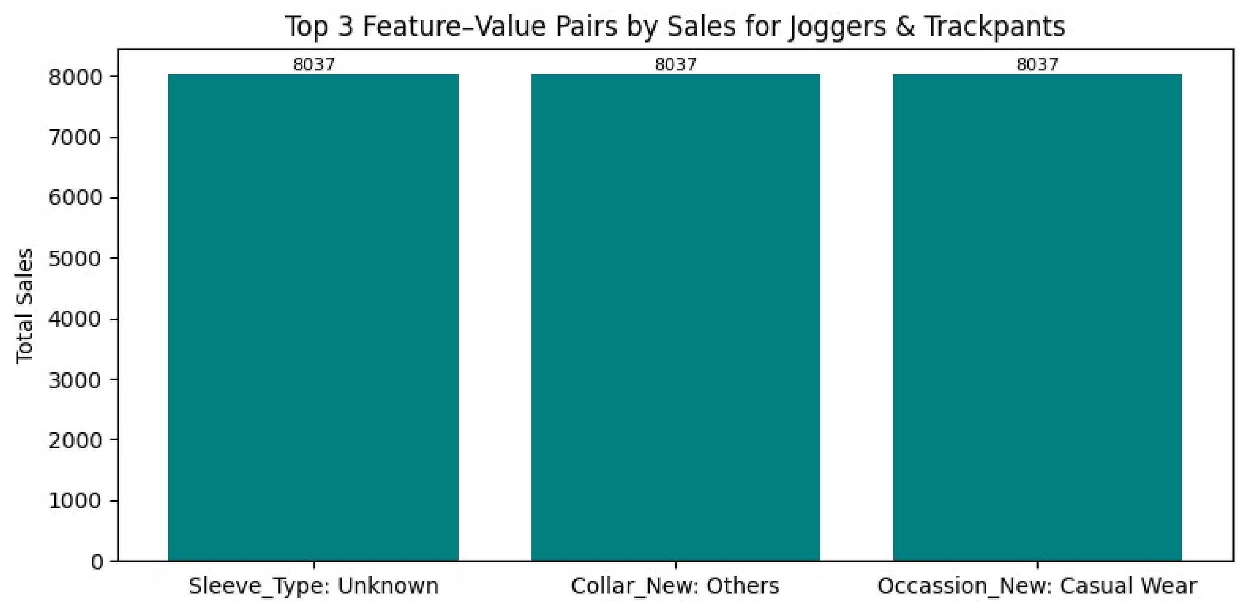

```
Category: Pyjamas, Rows: 43
```




```
Category: Sweatshirts, Rows: 77
```




```
Category: Joggers & Trackpants, Rows: 107
```



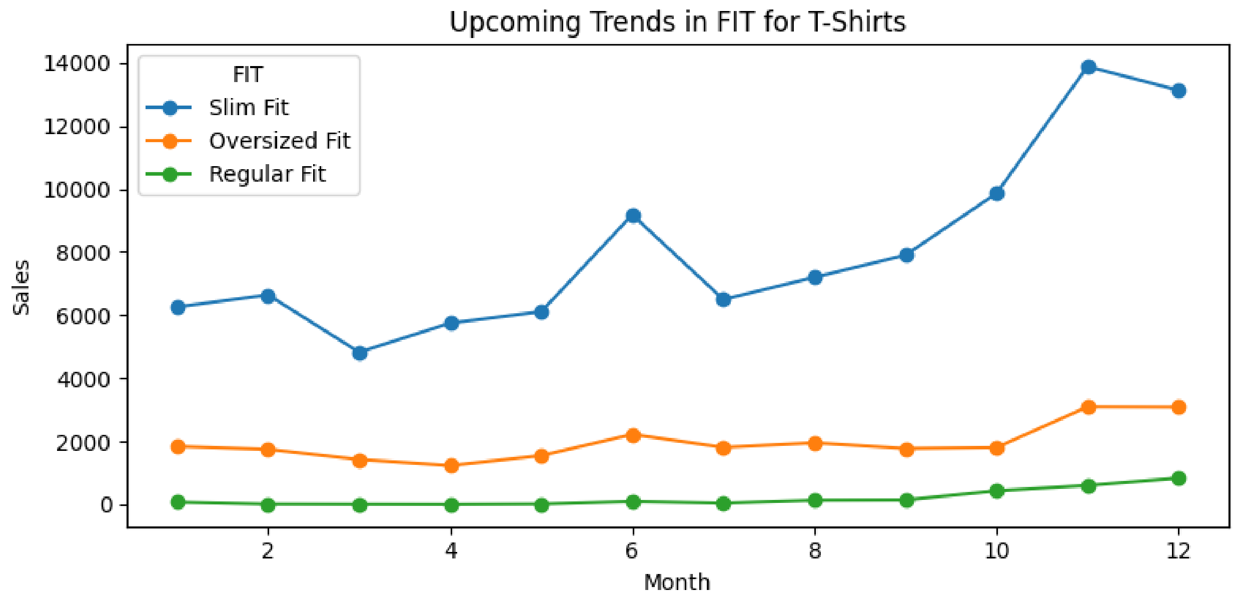
Q2. What are the upcoming trends on the basis of features Category Wise.

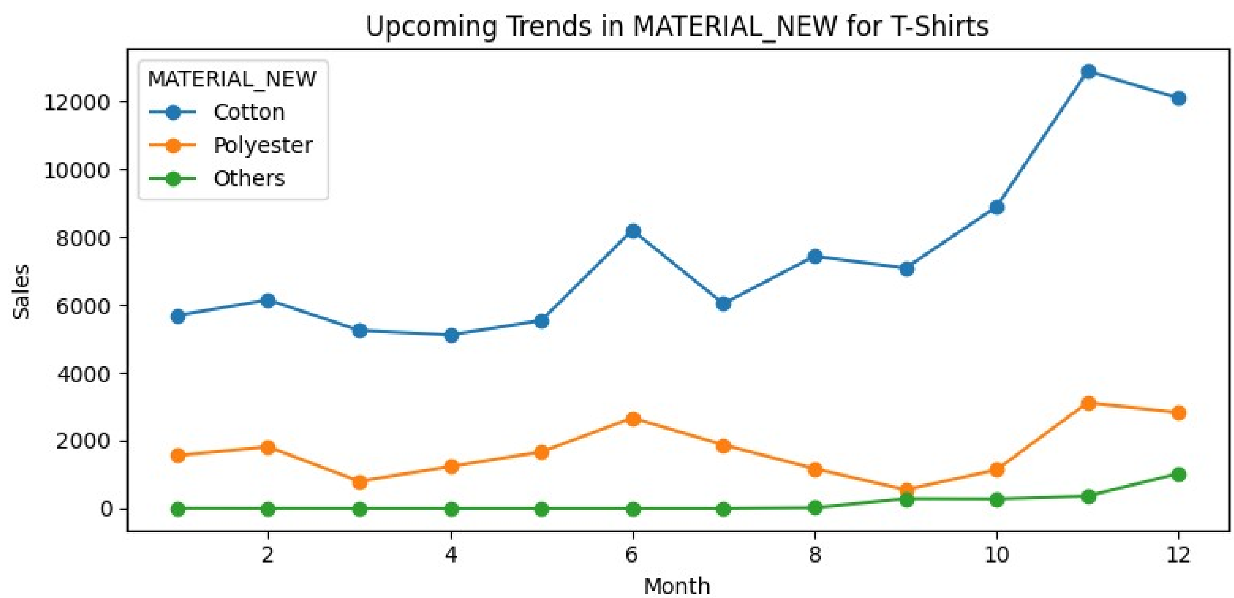
Approach -> The dataset was reshaped into a long format to facilitate time-series analysis of monthly sales. For each combination of category and feature, I computed the average monthly growth rate (slope) in sales. Based on these calculations, the top three feature values exhibiting the highest positive growth were identified as indicators of emerging trends. Their sales trajectories were then visualized using line charts to clearly highlight these evolving patterns.

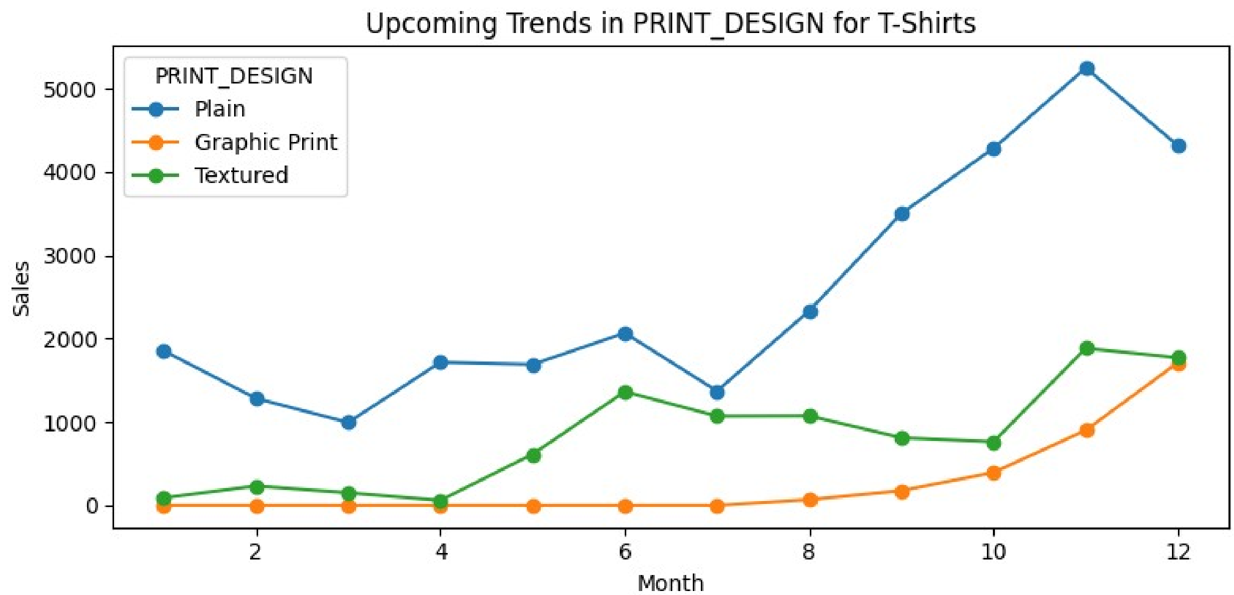
Note: In some cases, only a single graph is displayed because the data for certain category-feature combinations may not contain enough distinct feature values with meaningful monthly growth. The analysis computes growth trends based on the average change in sales over time (slope). If fewer than three feature values show positive or non-zero growth, only those available values are plotted. If all growth values are zero or insufficient data exists, the plot is skipped altogether to maintain the clarity and relevance of the visual output.

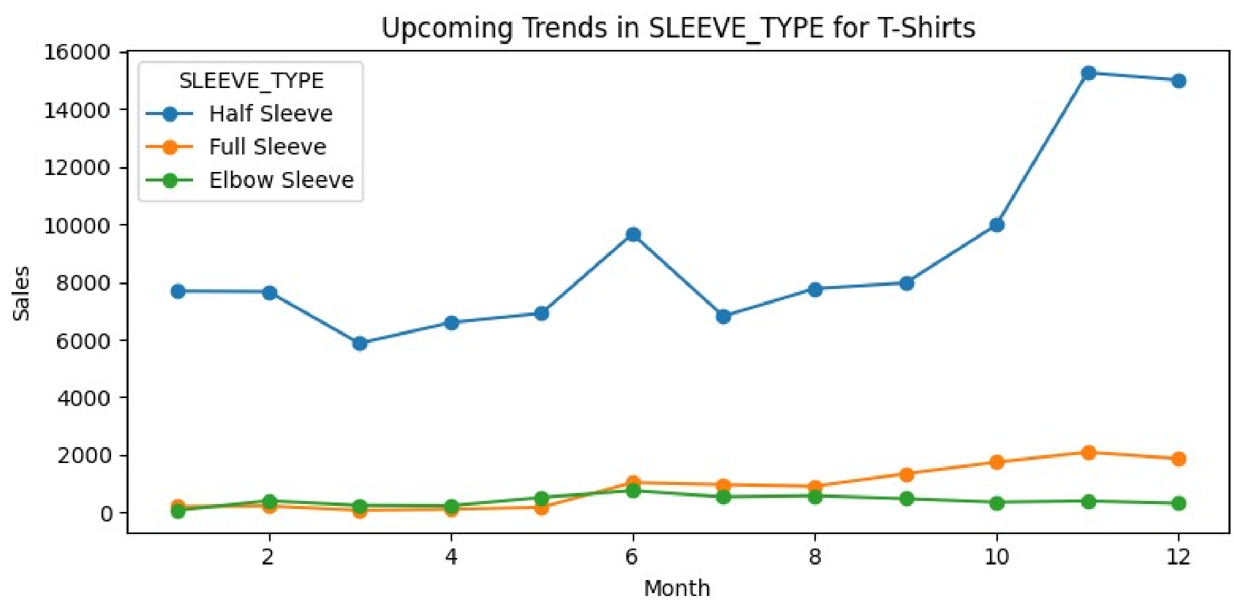
```
# Melt to long format for time series
df_long = df.melt(
    id_vars=['CATEGORY'] + feature_columns,
    value_vars=month_cols,
    var_name='MONTH',
    value_name='SALES'
)
df_long['MONTH_NUM'] = df_long['MONTH'].str.extract('(\
d+)').astype(int)

# For each category and feature, plotting top 3 growing feature values
for category in df_long['CATEGORY'].unique():
    cat_df = df_long[df_long['CATEGORY'] == category]
    for feature in feature_columns:
        # Get total sales per value per month
        pivot = cat_df.groupby([feature, 'MONTH_NUM'])
['SALES'].sum().reset_index()
        # Calculating growth (slope) for each value
        growth = {}
        for val in pivot[feature].unique():
            sales_series = pivot[pivot[feature] ==
val].sort_values('MONTH_NUM')
            if len(sales_series) > 1:
                slope =
pd.Series(sales_series['SALES'].values).diff().mean()
                growth[val] = slope
        # Get top 3 growing values
        top_vals = sorted(growth.items(), key=lambda x: x[1],
reverse=True)[:3]
        if not top_vals or all(v[1] == 0 for v in top_vals):
            continue
        # Plot
        plt.figure(figsize=(8,4))
        for val, _ in top_vals:
            sales = pivot[pivot[feature] ==
val].sort_values('MONTH_NUM')
            plt.plot(sales['MONTH_NUM'], sales['SALES'], marker='o',
label=val)
        plt.title(f"Upcoming Trends in {feature} for {category}")
        plt.xlabel('Month')
        plt.ylabel('Sales')
        plt.legend(title=feature)
```

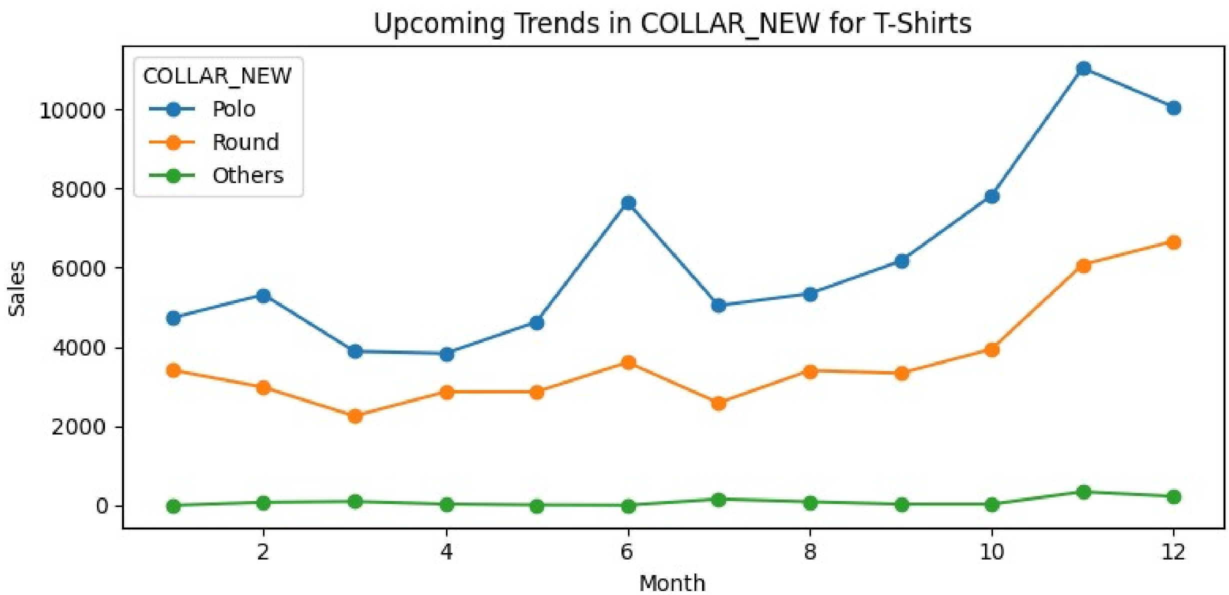
```
plt.tight_layout()
plt.show()
```



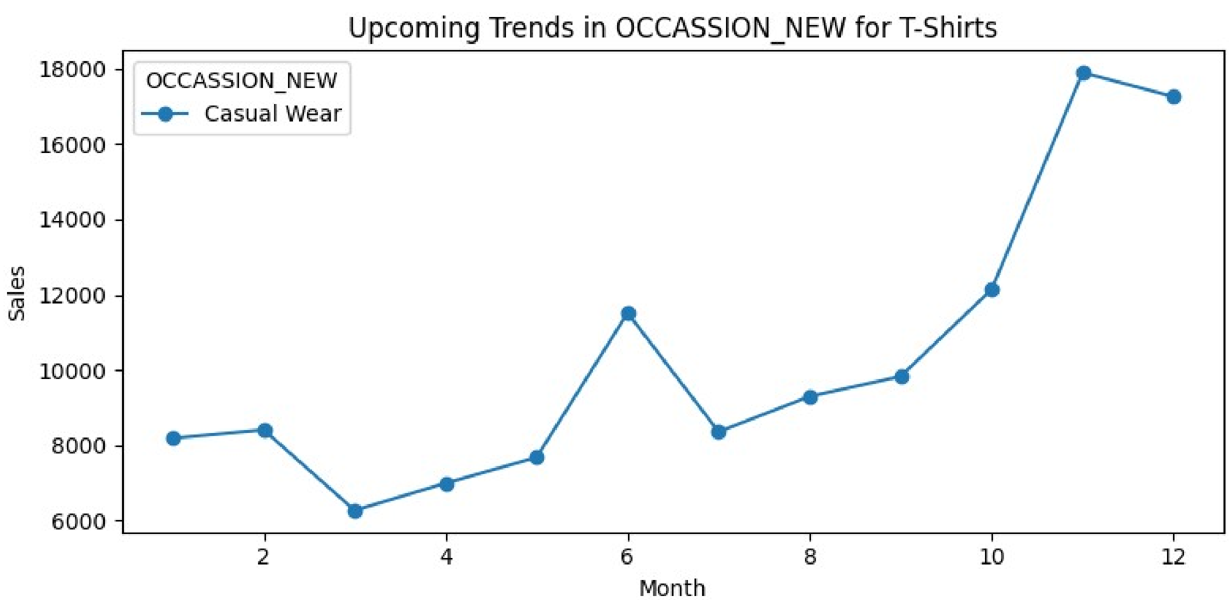



Upcoming Trends in PRINT_DESIGN for T-Shirts
PRINT_DESIGN
Plain
Graphic Print
Textured
Sales
Month
0
1000
2000
3000
4000
5000
2
4
6
8
10
12

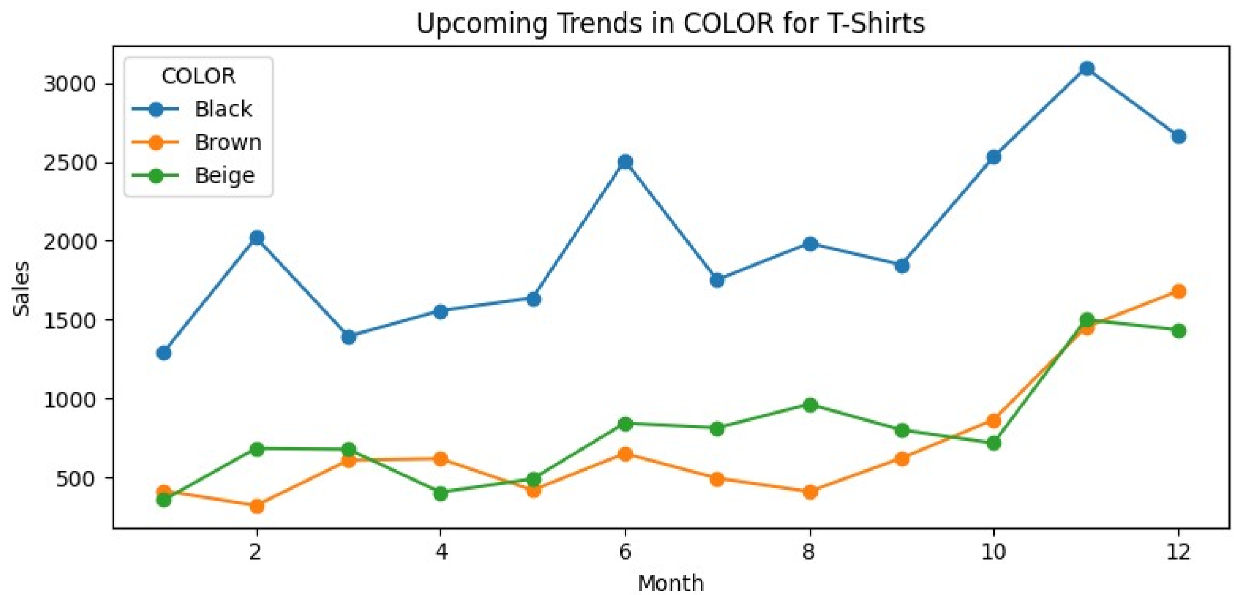

Upcoming Trends in SLEEVE_TYPE for T-Shirts
SLEEVE_TYPE
Half Sleeve
Full Sleeve
Elbow Sleeve
Sales
Month
0
2000
4000
6000
8000
10000
12000
14000
16000
2
4
6
8
10
12

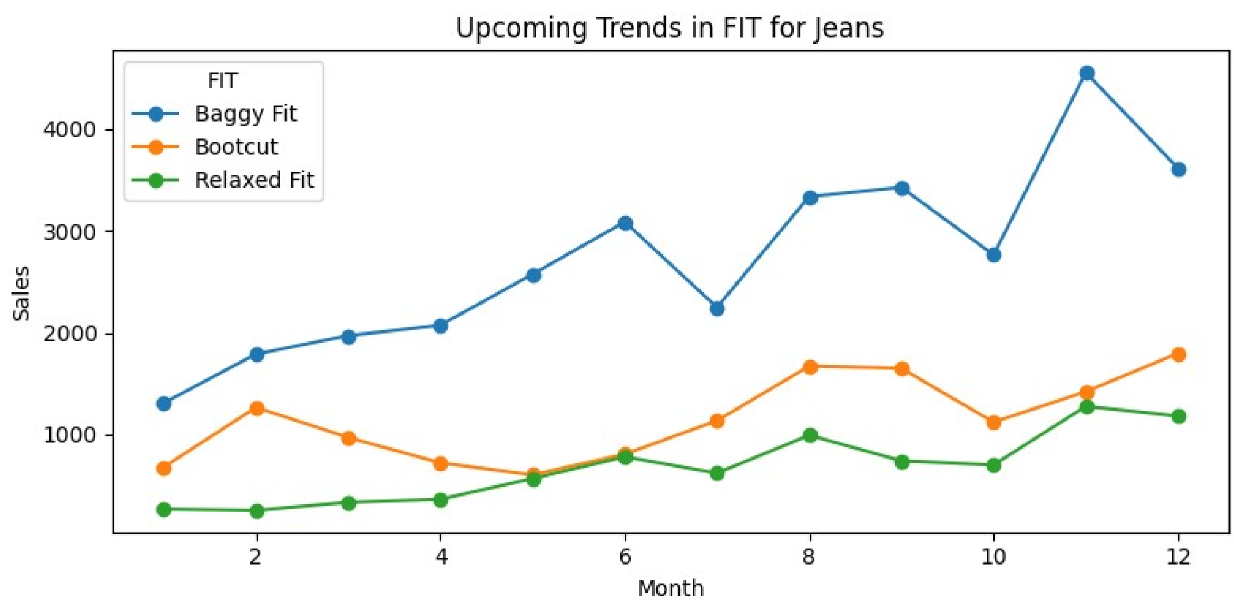


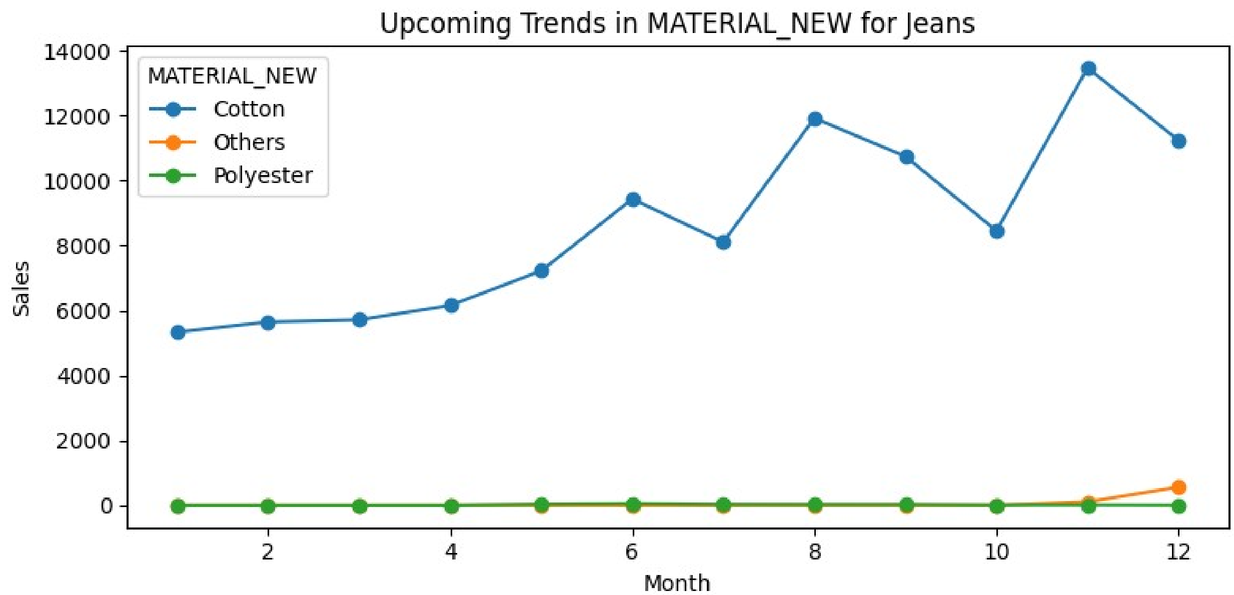
Upcoming Trends in COLLAR_NEW for T-Shirts
COLLAR_NEW
Polo
Round
Others
Sales
Month
0
2000
4000
6000
8000
10000
2
4
6
8
10
12



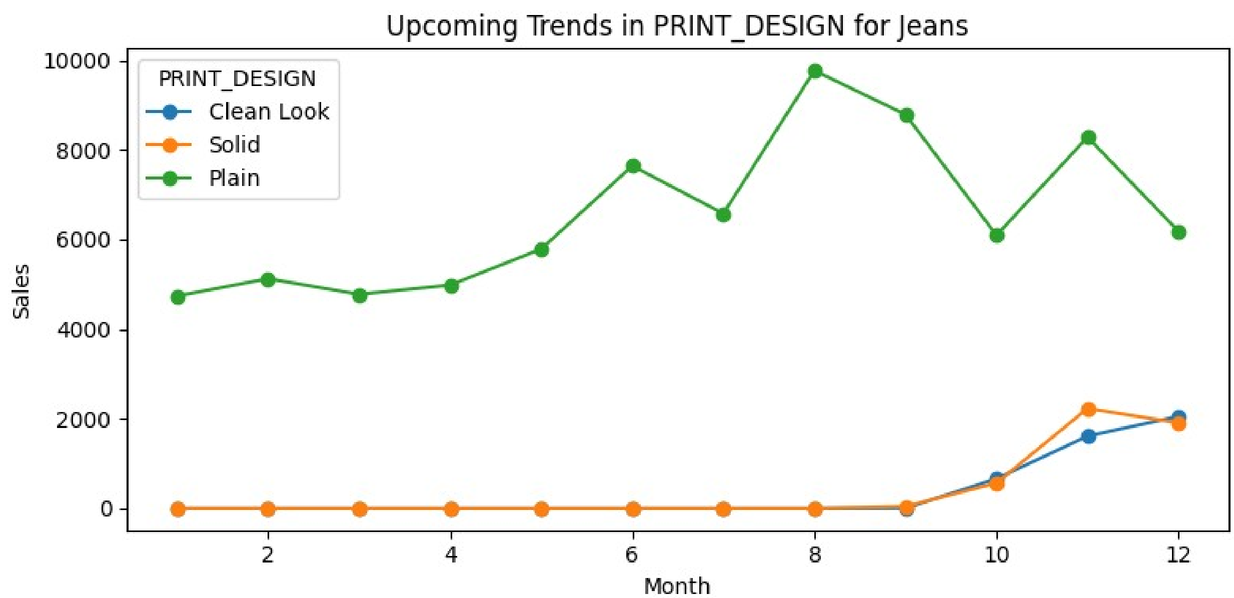
Upcoming Trends in OCCASSION_NEW for T-Shirts
OCCASSION_NEW
Casual Wear
Sales
Month
6000
8000
10000
12000
14000
16000
18000
2
4
6
8
10
12


Upcoming Trends in COLOR for T-Shirts
COLOR
Black
Brown
Beige
Sales
3000
2500
2000
1500
1000
500
Month
2
4
6
8
10
12

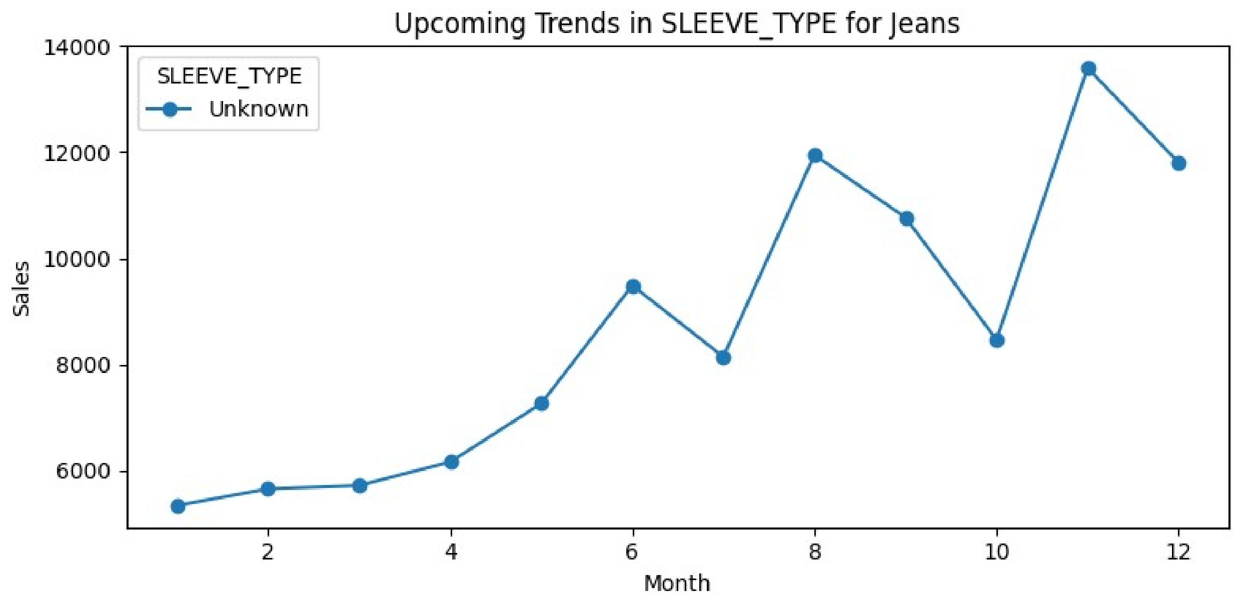

Upcoming Trends in FIT for Jeans
FIT
Baggy Fit
Bootcut
Relaxed Fit
Sales
4000
3000
2000
1000
Month
2
4
6
8
10
12



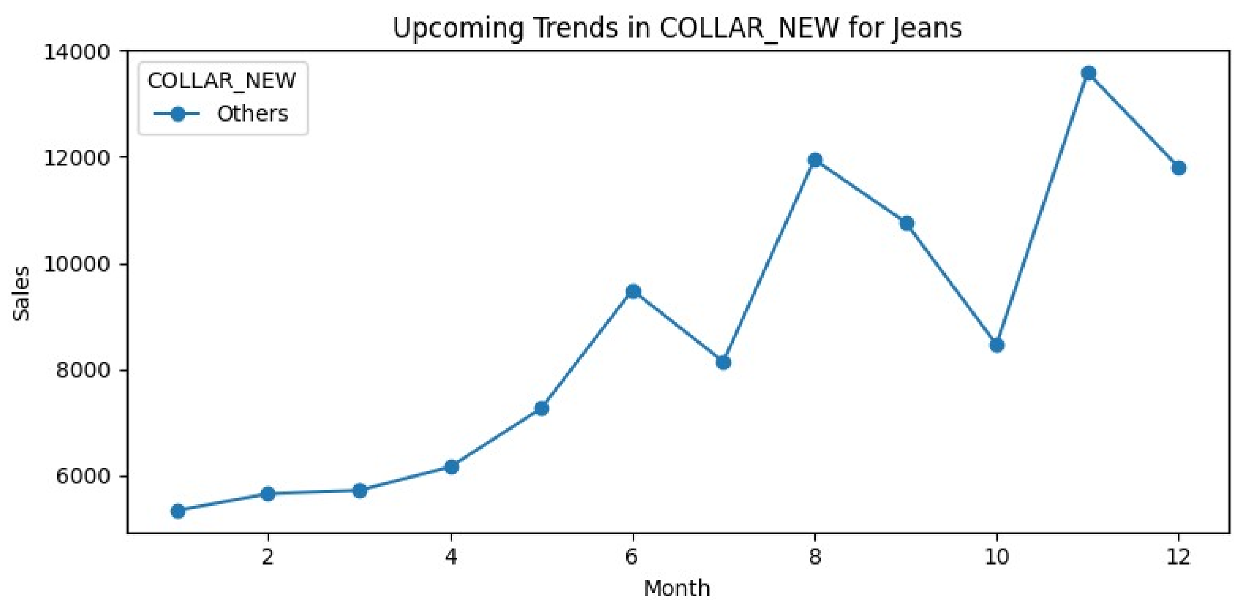
Upcoming Trends in MATERIAL_NEW for Jeans
MATERIAL_NEW
Cotton
Others
Polyester
Sales
Month
14000
12000
10000
8000
6000
4000
2000
0
2
4
6
8
10
12



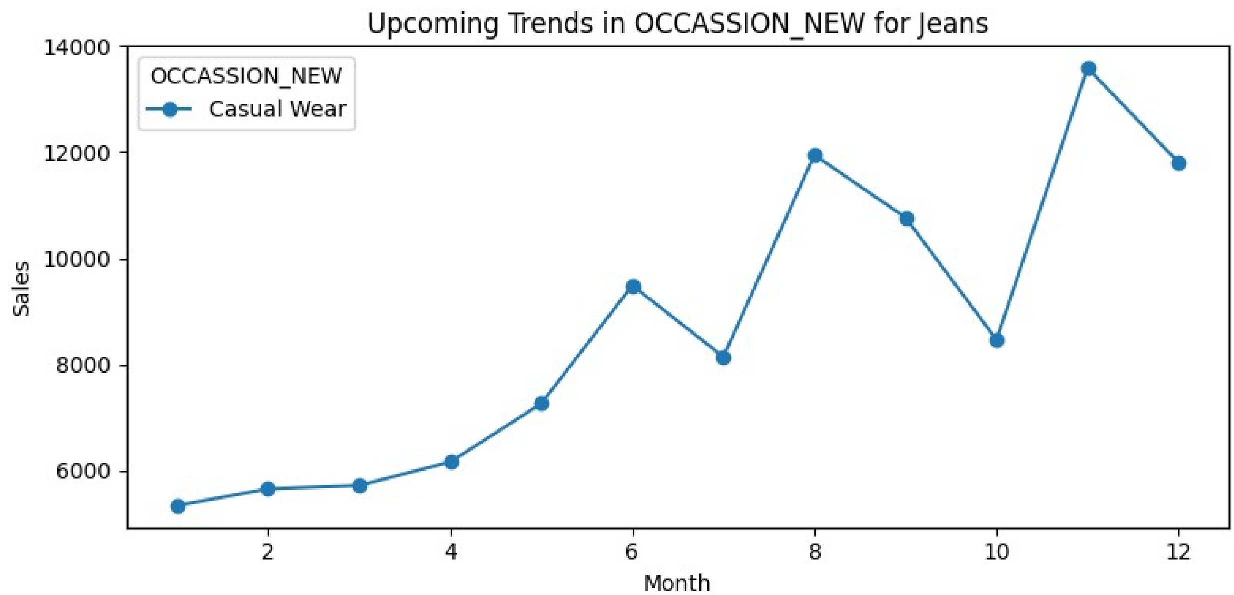
Upcoming Trends in PRINT_DESIGN for Jeans
PRINT_DESIGN
Clean Look
Solid
Plain
Sales
Month
10000
8000
6000
4000
2000
0
2
4
6
8
10
12

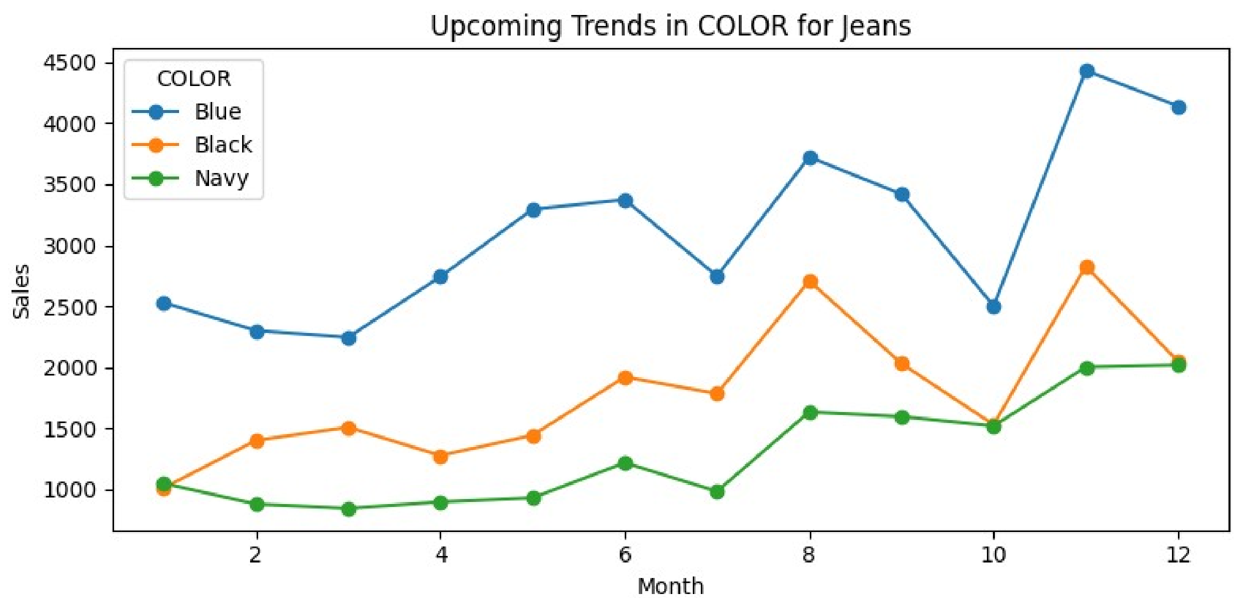


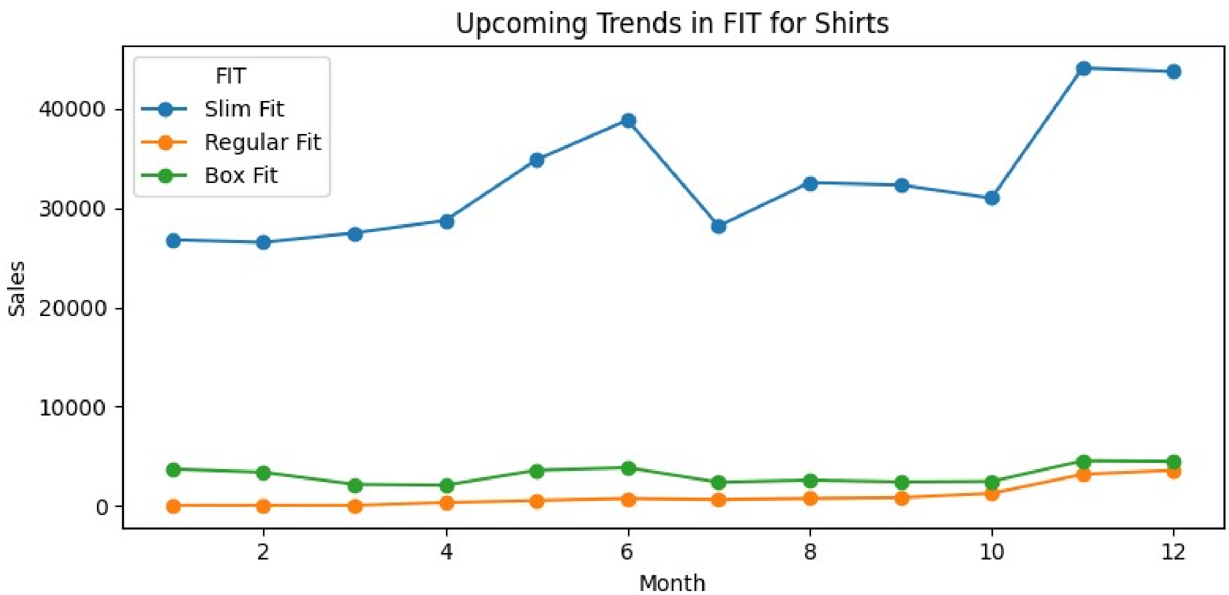
Upcoming Trends in SLEEVE_TYPE for Jeans
SLEEVE_TYPE
Unknown
Sales
Month
14000
12000
10000
8000
6000
2
4
6
8
10
12

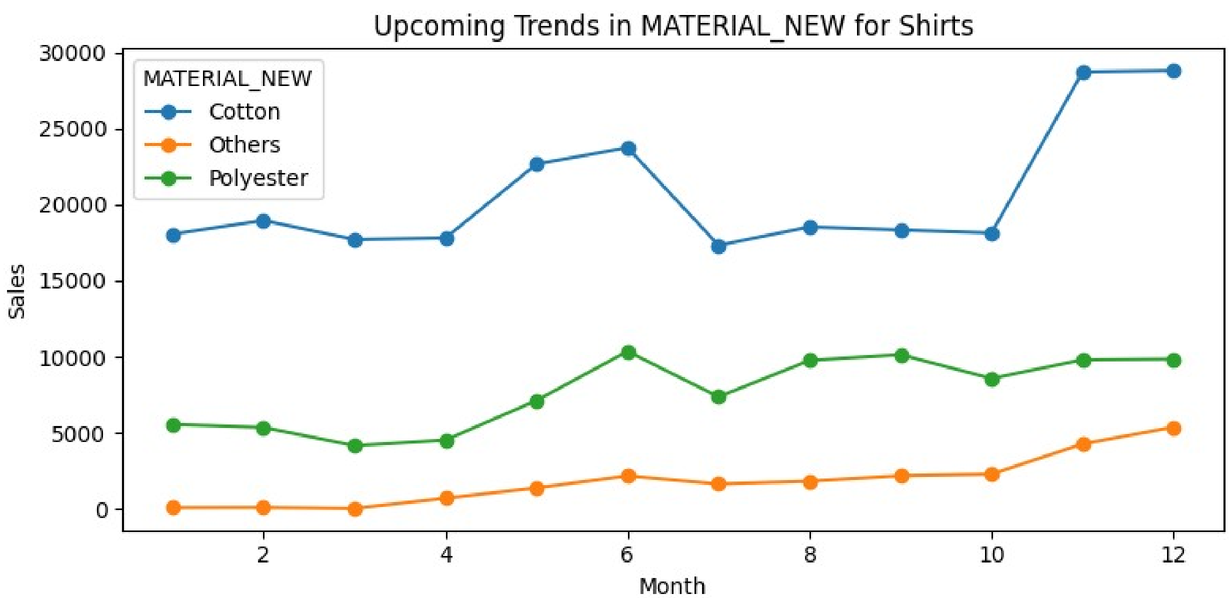


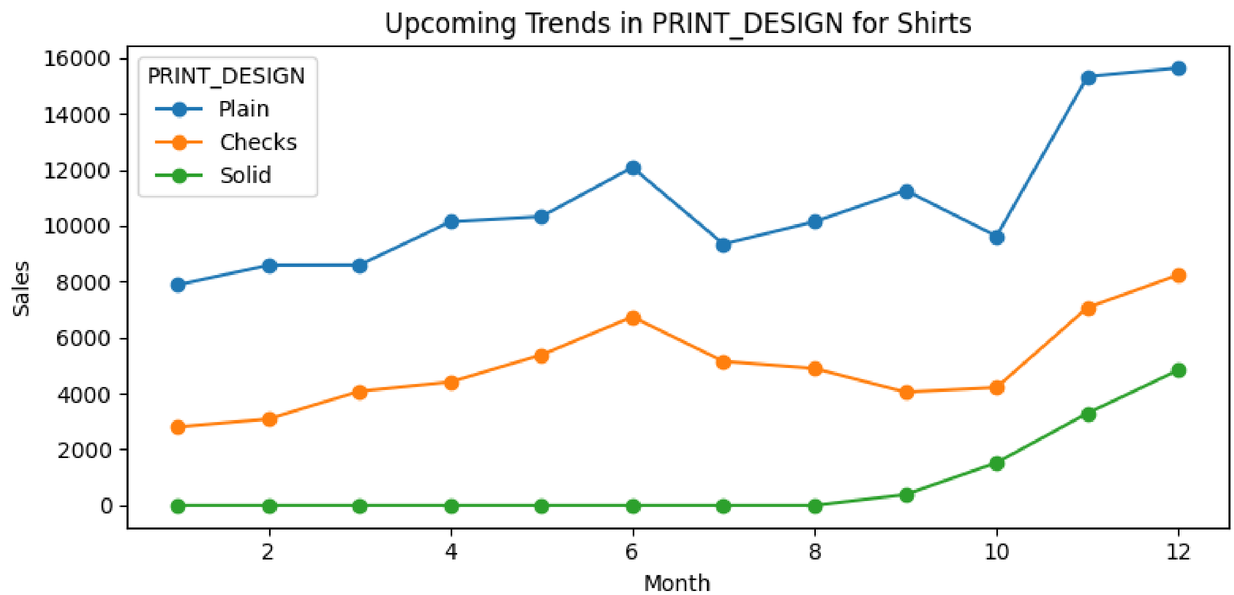
Upcoming Trends in COLLAR_NEW for Jeans
COLLAR_NEW
Others
Sales
Month
14000
12000
10000
8000
6000
2
4
6
8
10
12

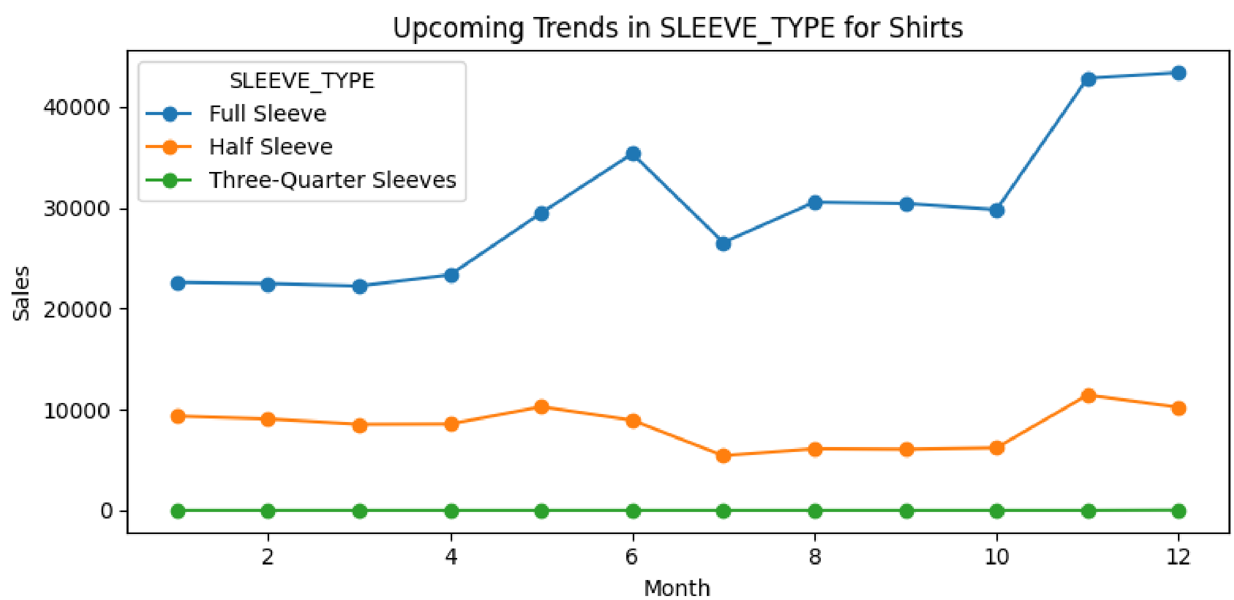

Upcoming Trends in OCCASSION_NEW for Jeans
OCCASSION_NEW
Casual Wear
Sales
Month
14000
12000
10000
8000
6000
2
4
6
8
10
12

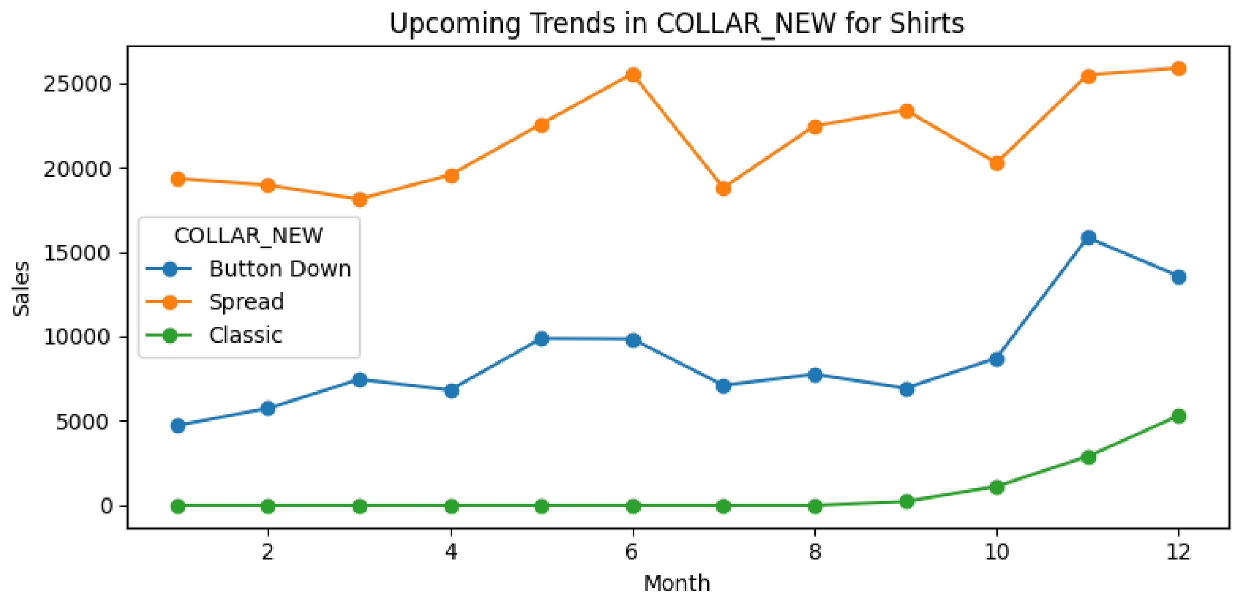

Upcoming Trends in COLOR for Jeans
COLOR
Blue
Black
Navy
Sales
Month
4500
4000
3500
3000
2500
2000
1500
1000
2
4
6
8
10
12

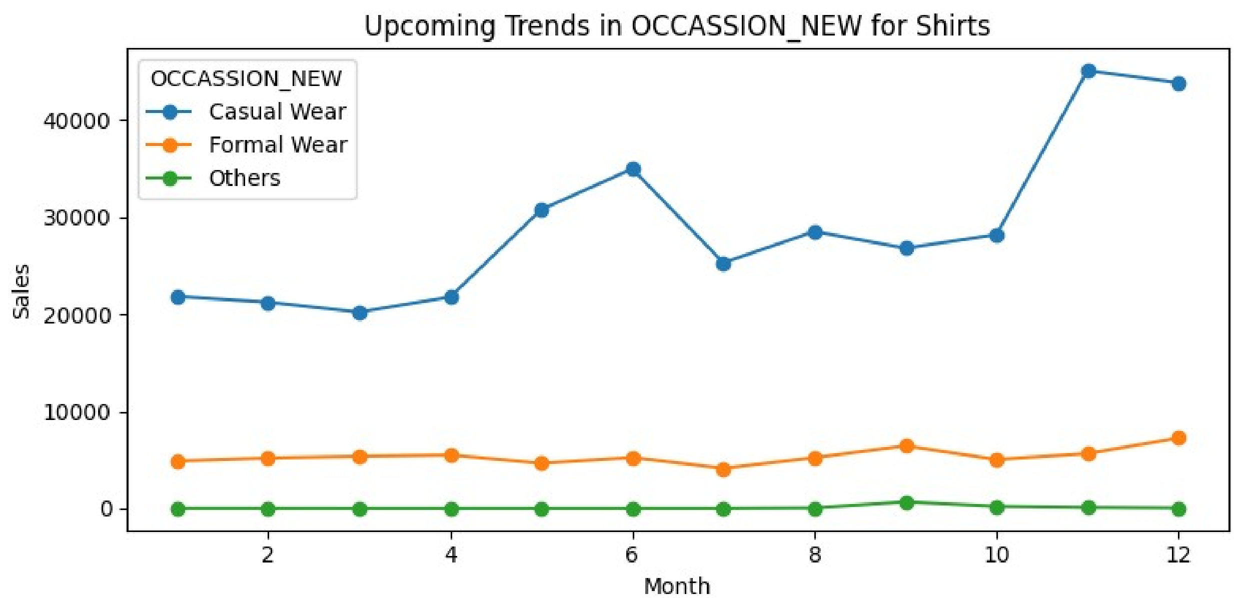

Upcoming Trends in FIT for Shirts
FIT
Slim Fit
Regular Fit
Box Fit
Sales
Month
0
10000
20000
30000
40000
2
4
6
8
10
12

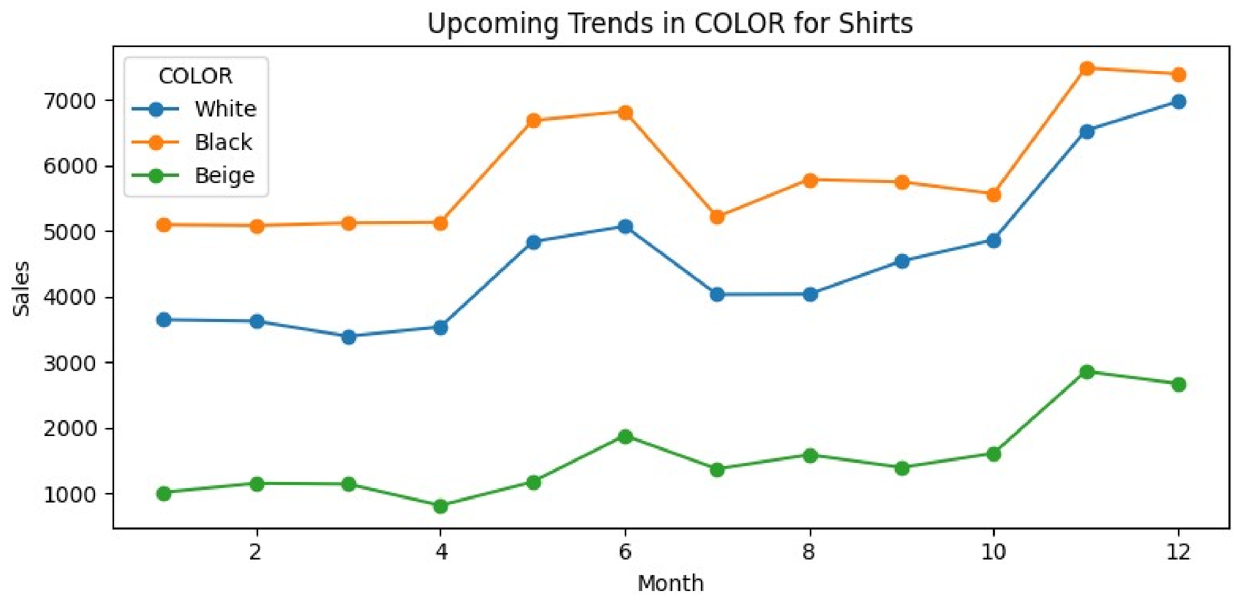

Upcoming Trends in MATERIAL_NEW for Shirts
MATERIAL_NEW
Cotton
Others
Polyester
Sales
Month
0
5000
10000
15000
20000
25000
30000
2
4
6
8
10
12

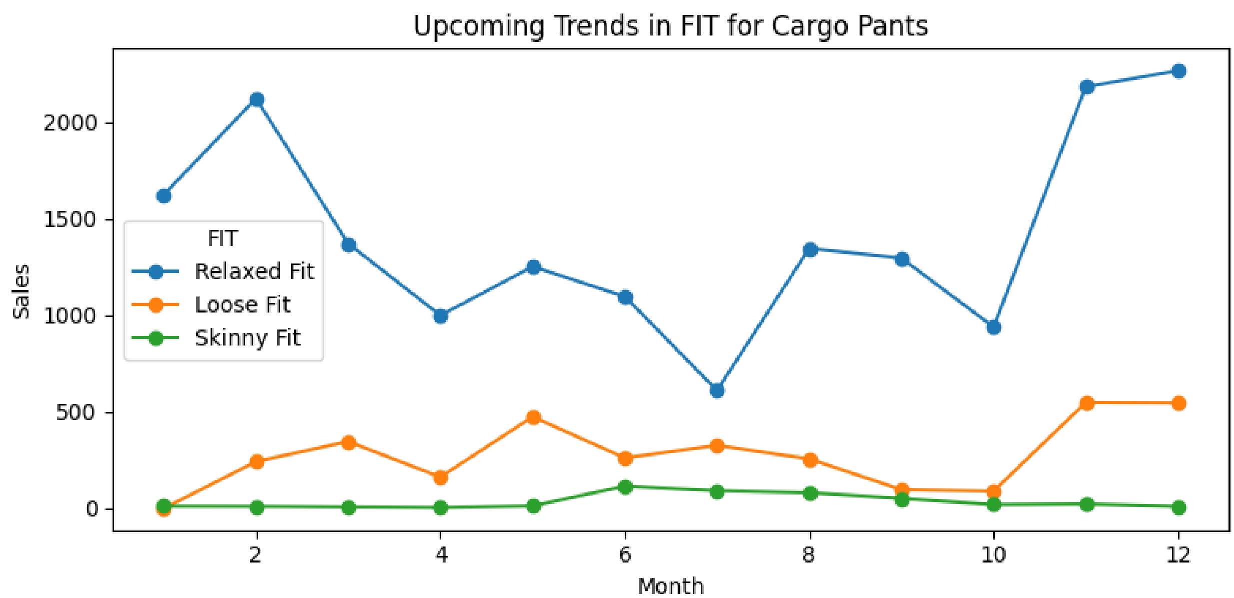

Upcoming Trends in PRINT_DESIGN for Shirts
PRINT_DESIGN
Plain
Checks
Solid
Sales
Month
0
2000
4000
6000
8000
10000
12000
14000
16000
2
4
6
8
10
12

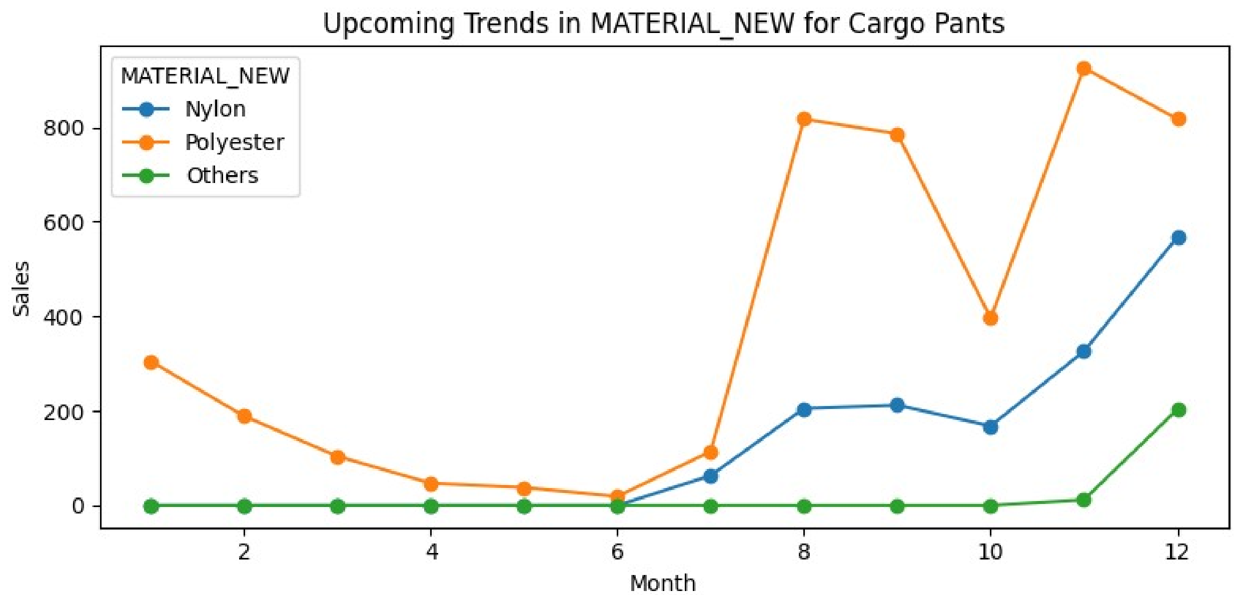

Upcoming Trends in SLEEVE_TYPE for Shirts
SLEEVE_TYPE
Full Sleeve
Half Sleeve
Three-Quarter Sleeves
Sales
Month
0
10000
20000
30000
40000
2
4
6
8
10
12

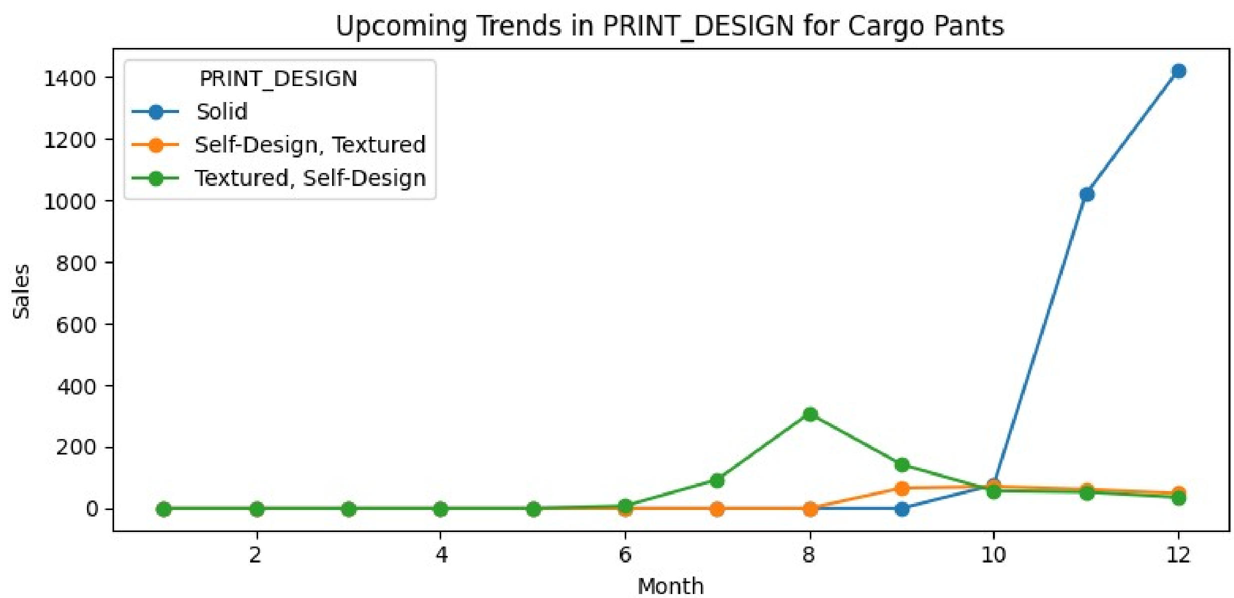

Upcoming Trends in COLLAR_NEW for Shirts
COLLAR_NEW
Button Down
Spread
Classic
Sales
Month
25000
20000
15000
10000
5000
0
2
4
6
8
10
12

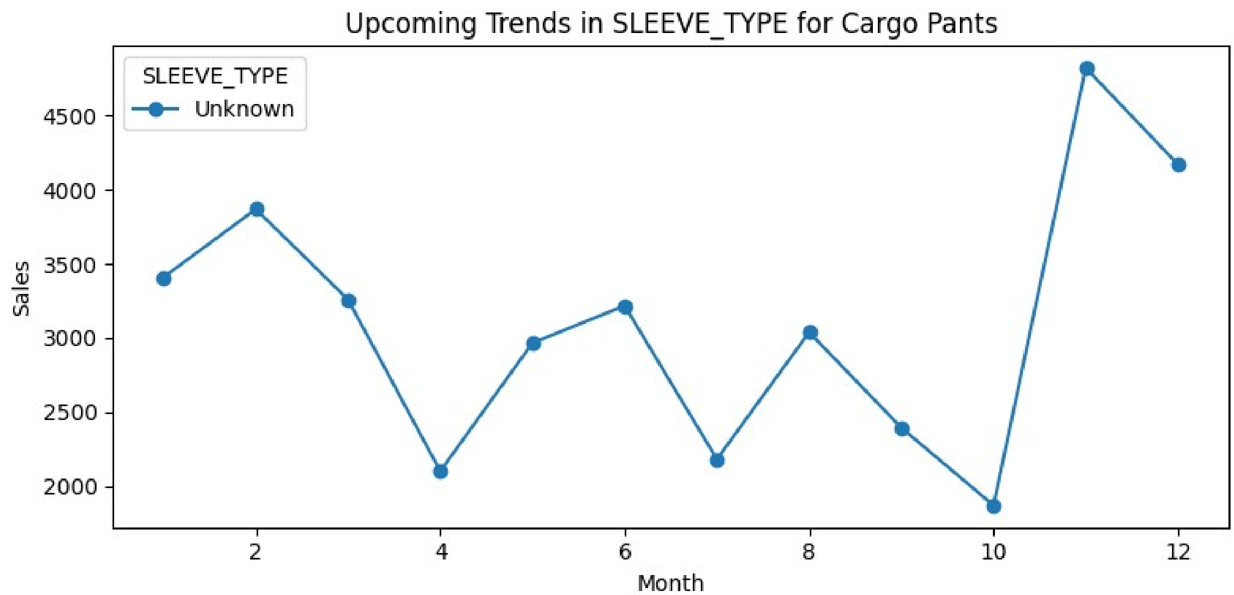

Upcoming Trends in OCCASSION_NEW for Shirts
OCCASSION_NEW
Casual Wear
Formal Wear
Others
Sales
Month
40000
30000
20000
10000
0
2
4
6
8
10
12


Upcoming Trends in COLOR for Shirts
COLOR
White
Black
Beige
Sales
Month
7000
6000
5000
4000
3000
2000
1000
2
4
6
8
10
12

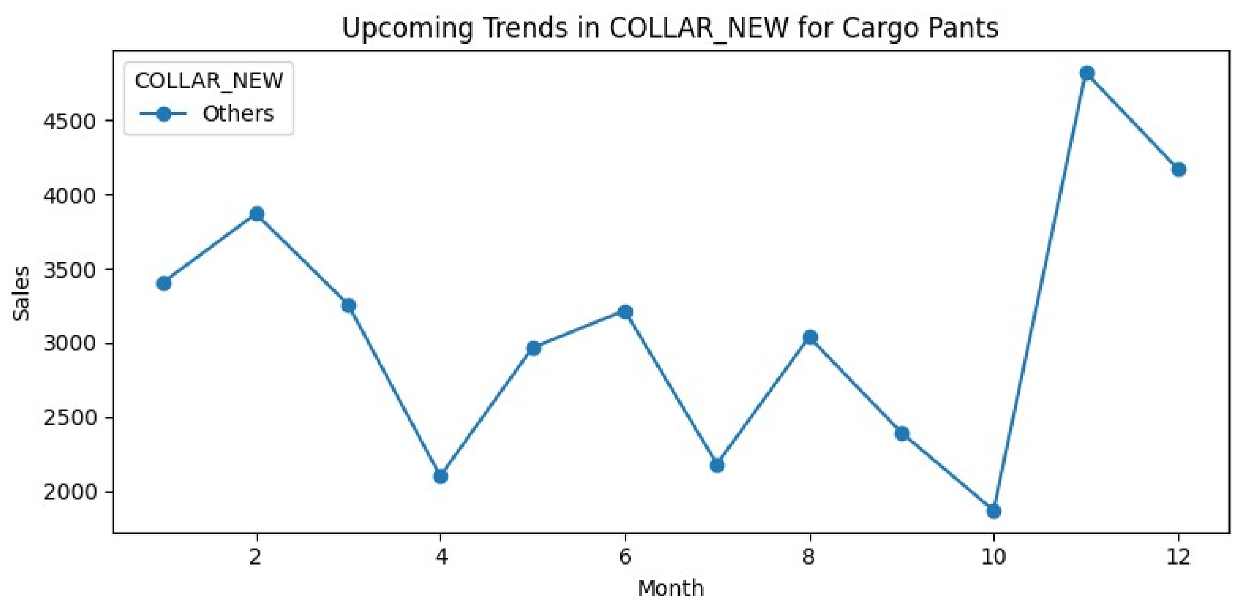

Upcoming Trends in FIT for Cargo Pants
FIT
Relaxed Fit
Loose Fit
Skinny Fit
Sales
Month
2000
1500
1000
500
0
2
4
6
8
10
12


Upcoming Trends in MATERIAL_NEW for Cargo Pants
MATERIAL_NEW
Nylon
Polyester
Others
Sales
Month
0
200
400
600
800
2
4
6
8
10
12

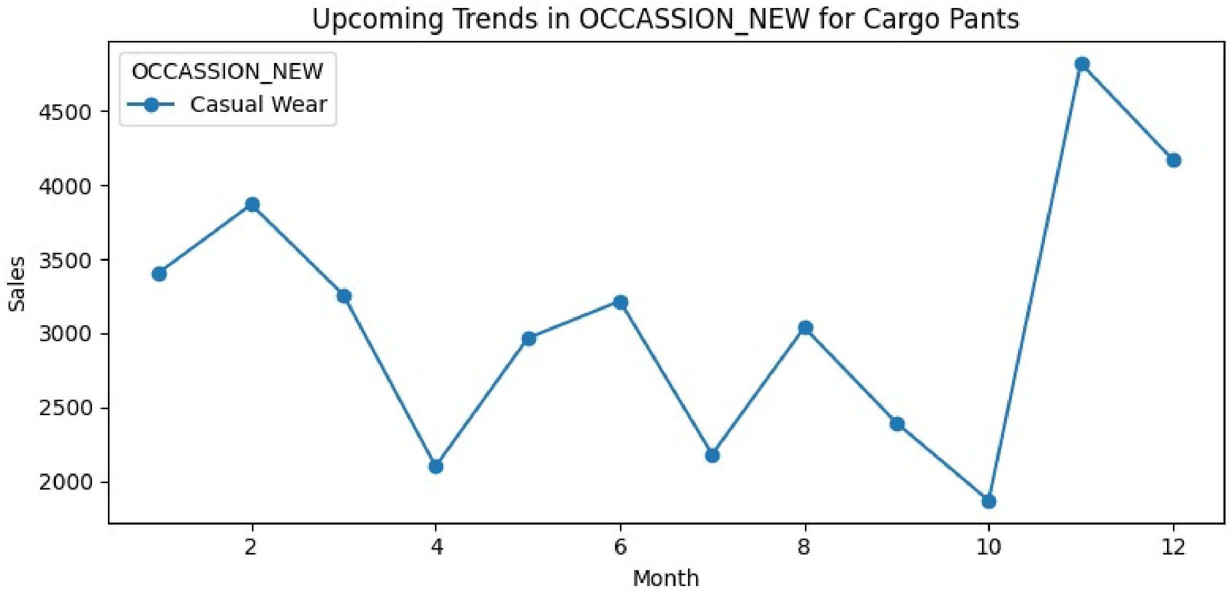

Upcoming Trends in PRINT_DESIGN for Cargo Pants
PRINT_DESIGN
Solid
Self-Design, Textured
Textured, Self-Design
Sales
Month
0
200
400
600
800
1000
1200
1400
2
4
6
8
10
12

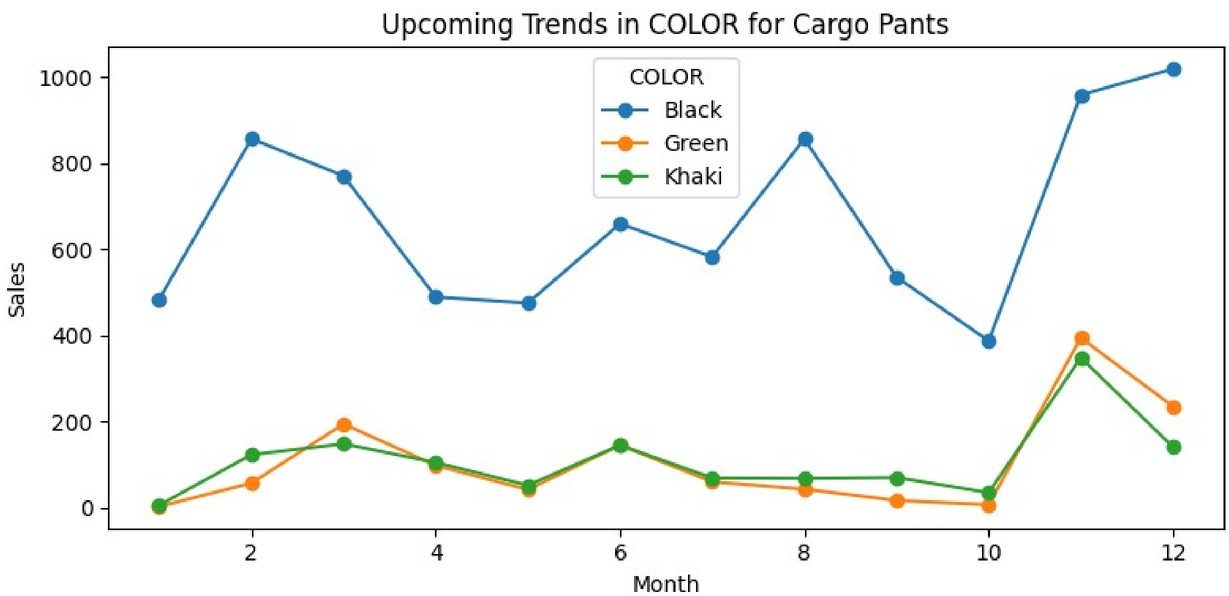


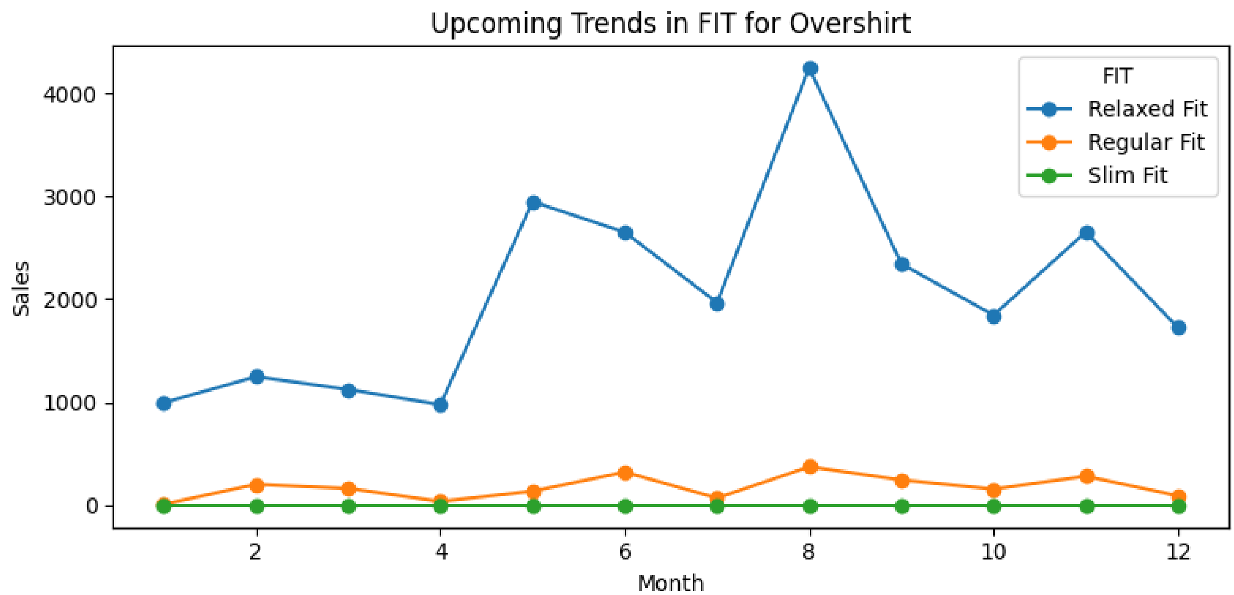
Upcoming Trends in SLEEVE_TYPE for Cargo Pants
SLEEVE_TYPE
Unknown
Sales
4500
4000
3500
3000
2500
2000
2
4
6
8
10
12
Month

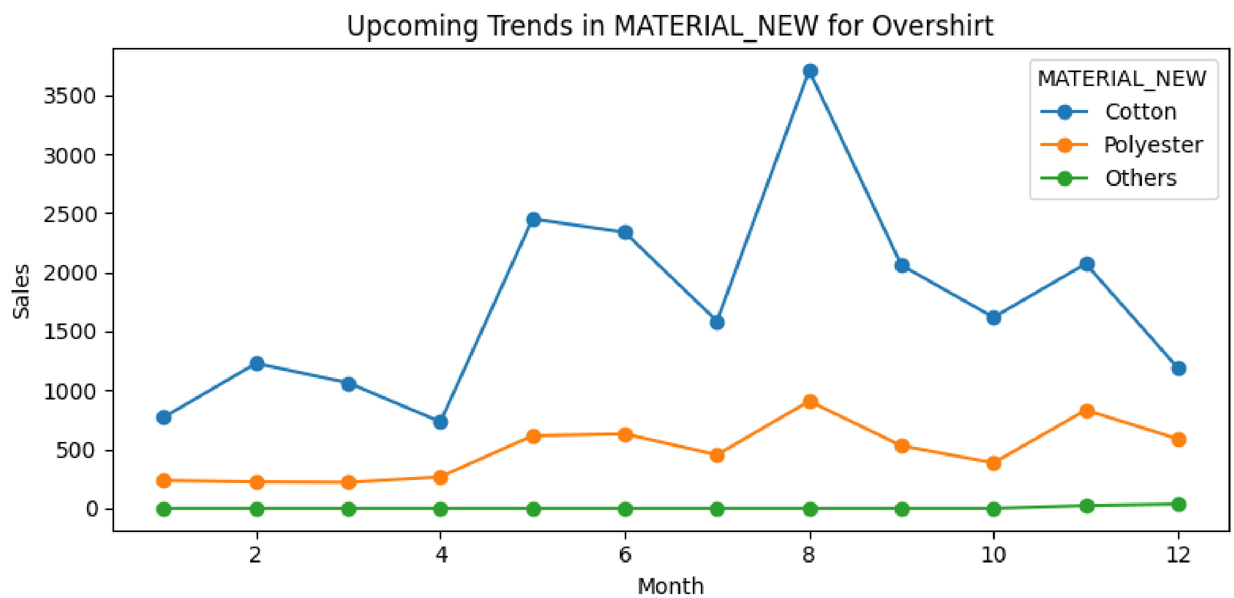


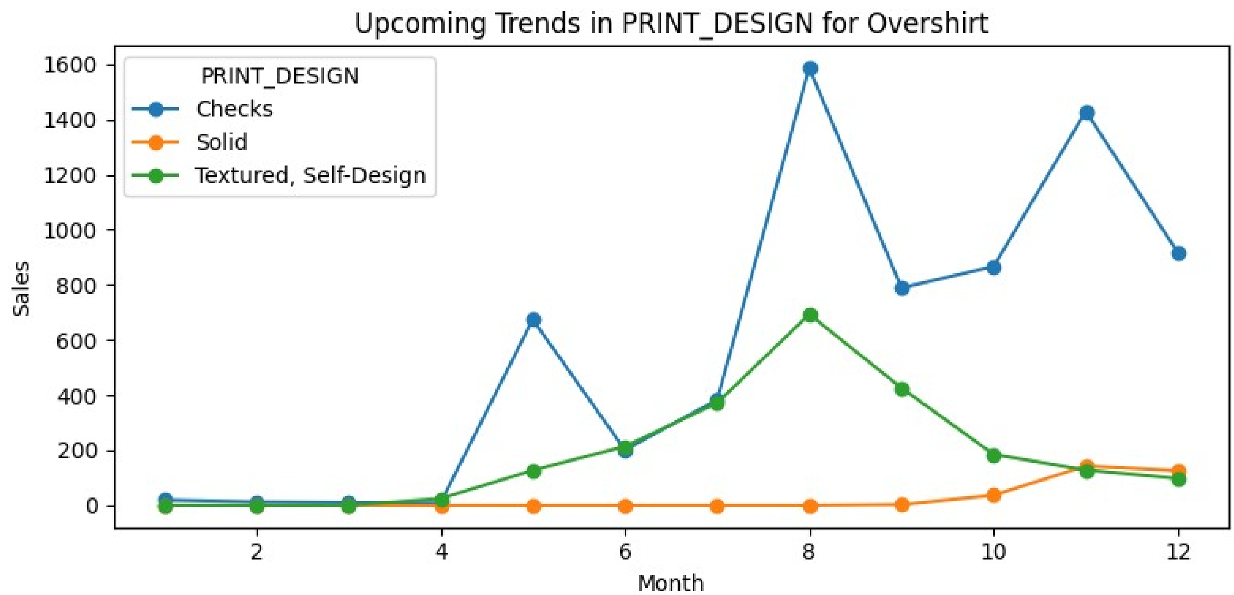
Upcoming Trends in COLLAR_NEW for Cargo Pants
COLLAR_NEW
Others
Sales
4500
4000
3500
3000
2500
2000
2
4
6
8
10
12
Month

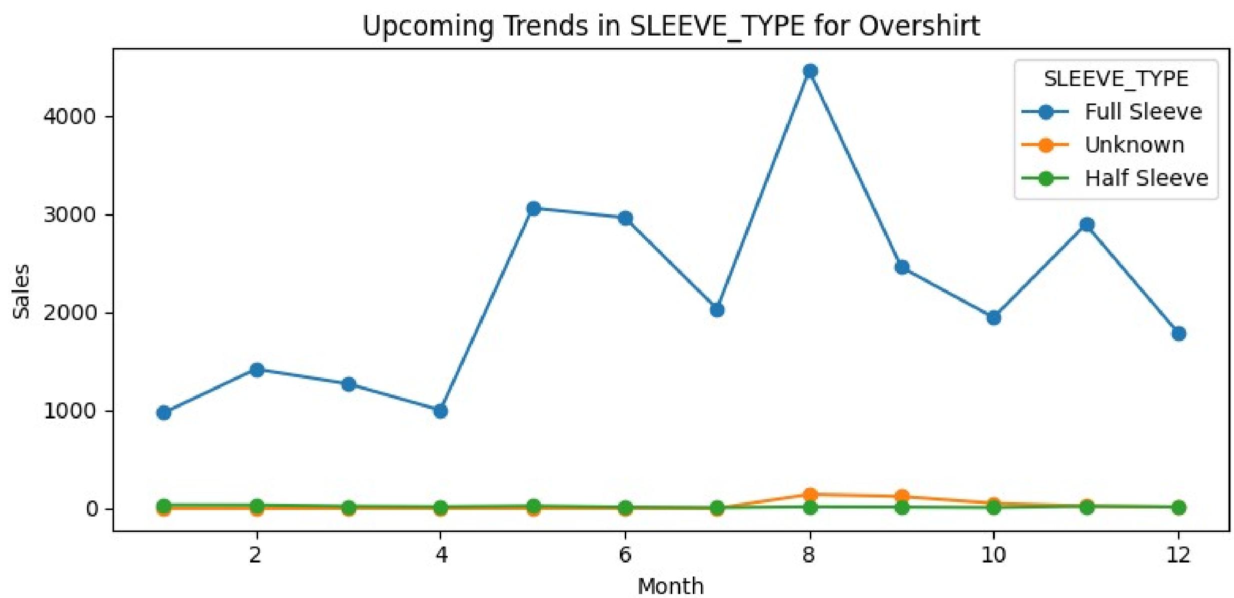

Upcoming Trends in OCCASSION_NEW for Cargo Pants
OCCASSION_NEW
Casual Wear
Sales
Month
4500
4000
3500
3000
2500
2000
2
4
6
8
10
12

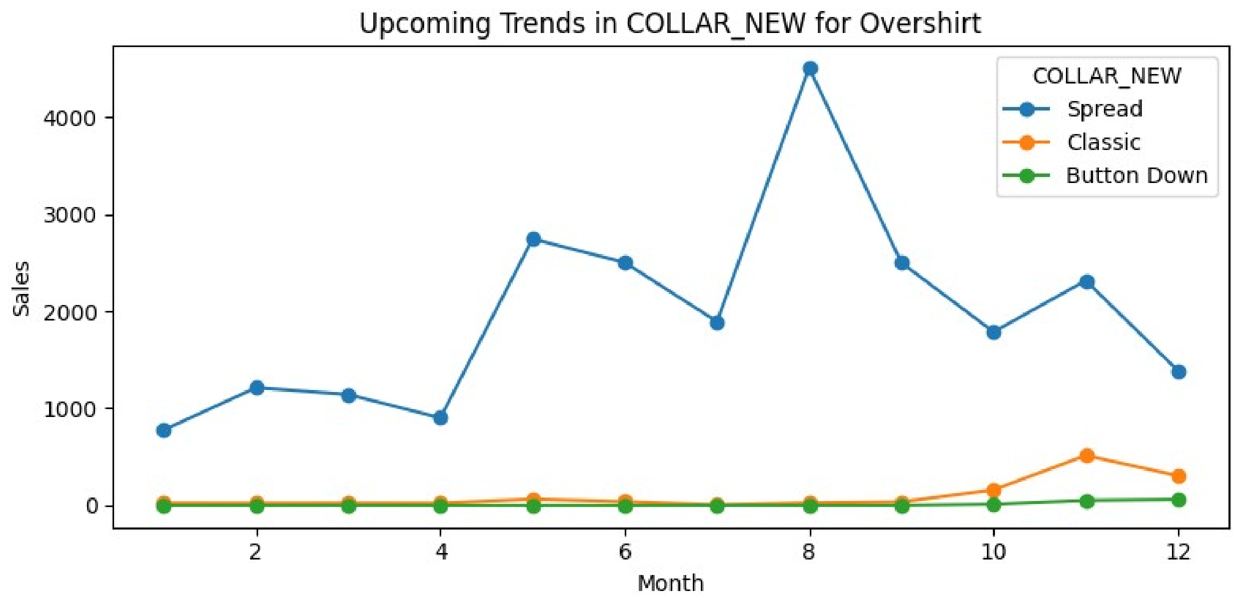

Upcoming Trends in COLOR for Cargo Pants
COLOR
Black
Green
Khaki
Sales
Month
1000
800
600
400
200
0
2
4
6
8
10
12

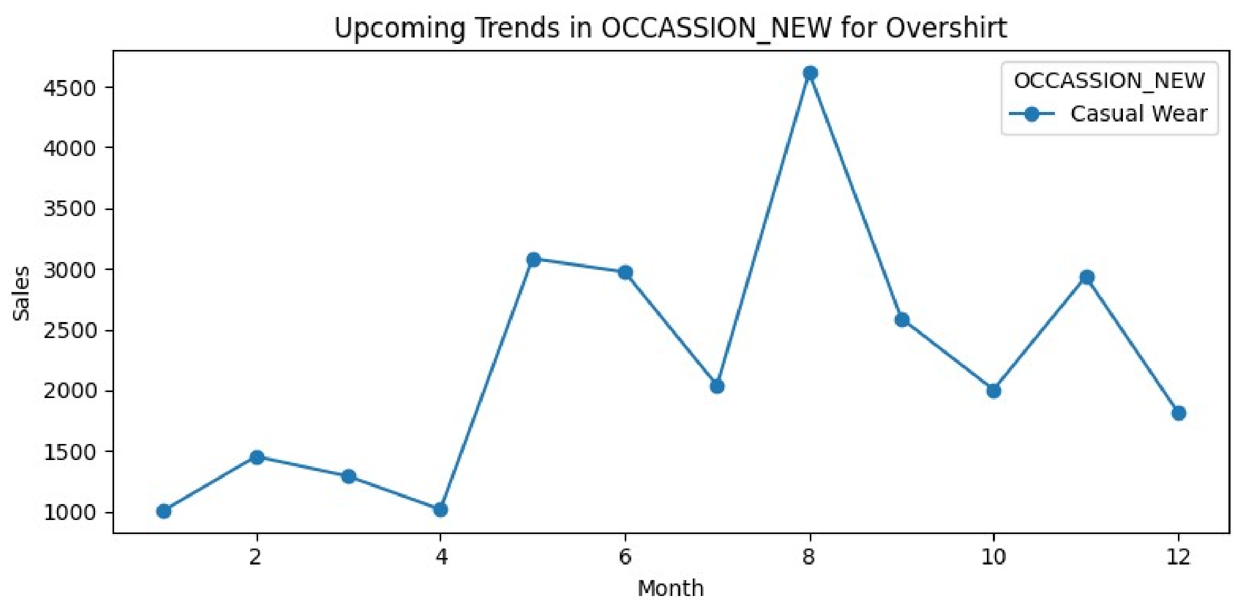

Upcoming Trends in FIT for Overshirt
FIT
Relaxed Fit
Regular Fit
Slim Fit
Sales
Month
0
1000
2000
3000
4000
2
4
6
8
10
12

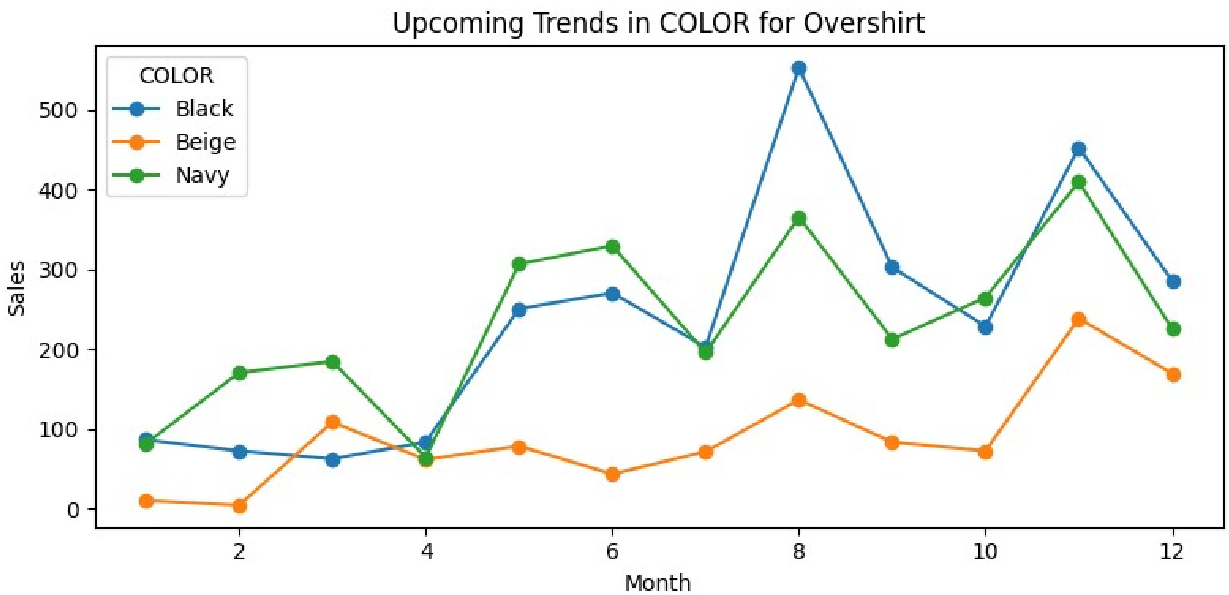

Upcoming Trends in MATERIAL_NEW for Overshirt
MATERIAL_NEW
Cotton
Polyester
Others
Sales
Month
0
500
1000
1500
2000
2500
3000
3500
2
4
6
8
10
12


Upcoming Trends in PRINT_DESIGN for Overshirt
PRINT_DESIGN
Checks
Solid
Textured, Self-Design
Sales
Month
1600
1400
1200
1000
800
600
400
200
0
2
4
6
8
10
12


Upcoming Trends in SLEEVE_TYPE for Overshirt
SLEEVE_TYPE
Full Sleeve
Unknown
Half Sleeve
Sales
Month
4000
3000
2000
1000
0
2
4
6
8
10
12

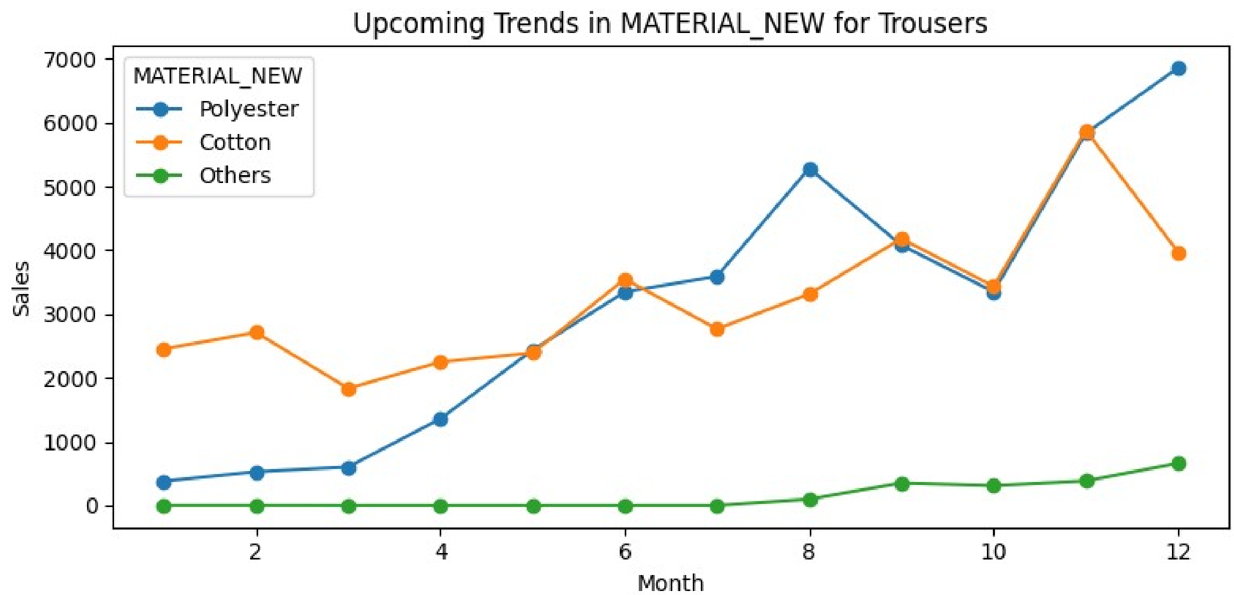

Upcoming Trends in COLLAR_NEW for Overshirt
COLLAR_NEW
Spread
Classic
Button Down
Sales
Month
0
1000
2000
3000
4000
2
4
6
8
10
12

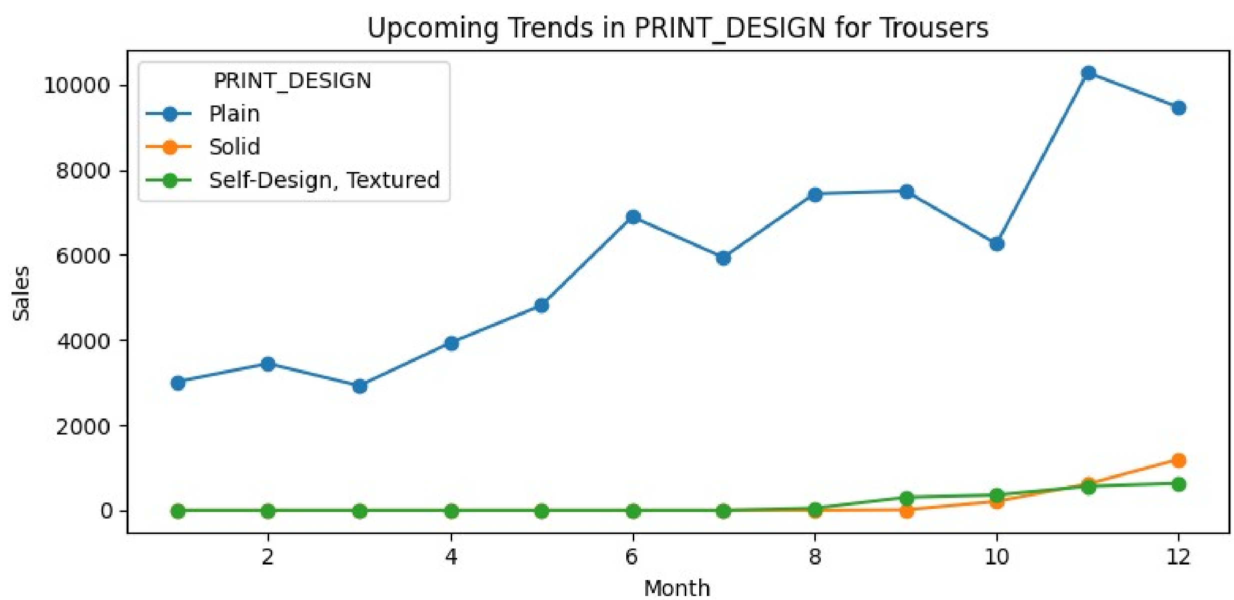

Upcoming Trends in OCCASSION_NEW for Overshirt
OCCASSION_NEW
Casual Wear
Sales
Month
1000
1500
2000
2500
3000
3500
4000
4500
2
4
6
8
10
12

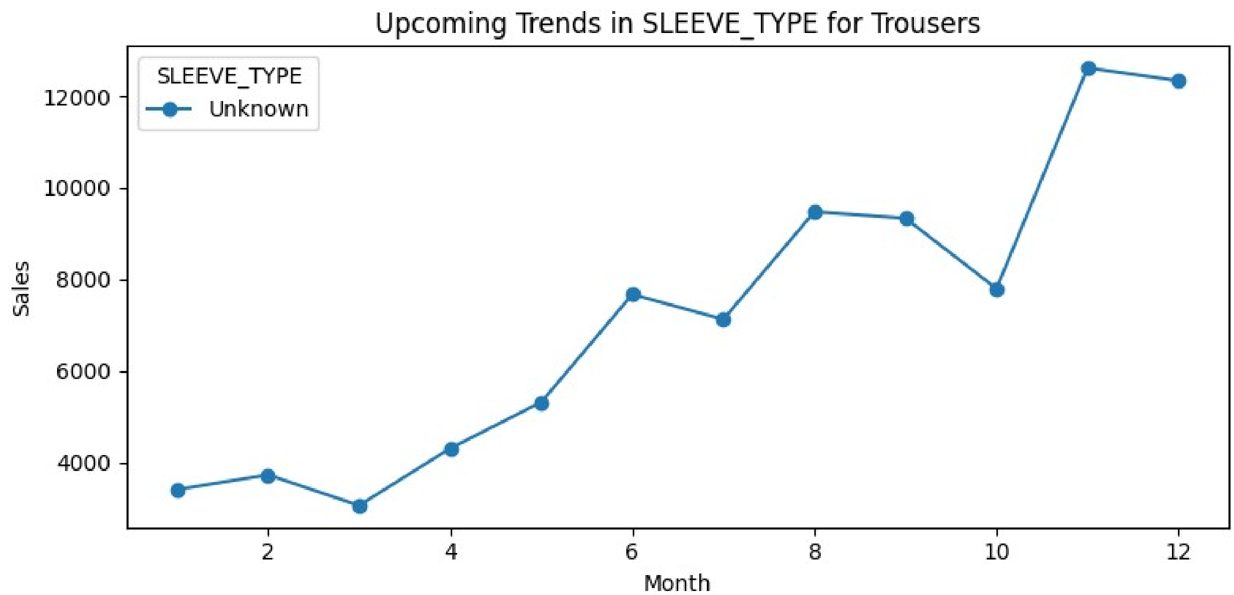


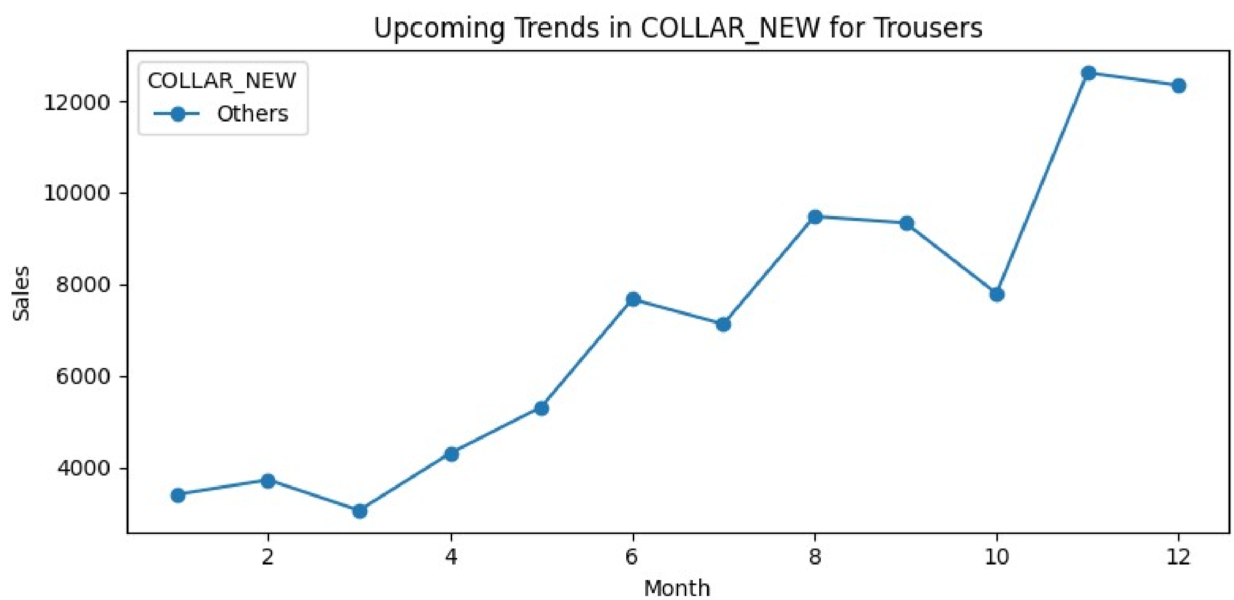
Upcoming Trends in COLOR for Overshirt
COLOR
Black
Beige
Navy
Sales
0
100
200
300
400
500
Month
2
4
6
8
10
12

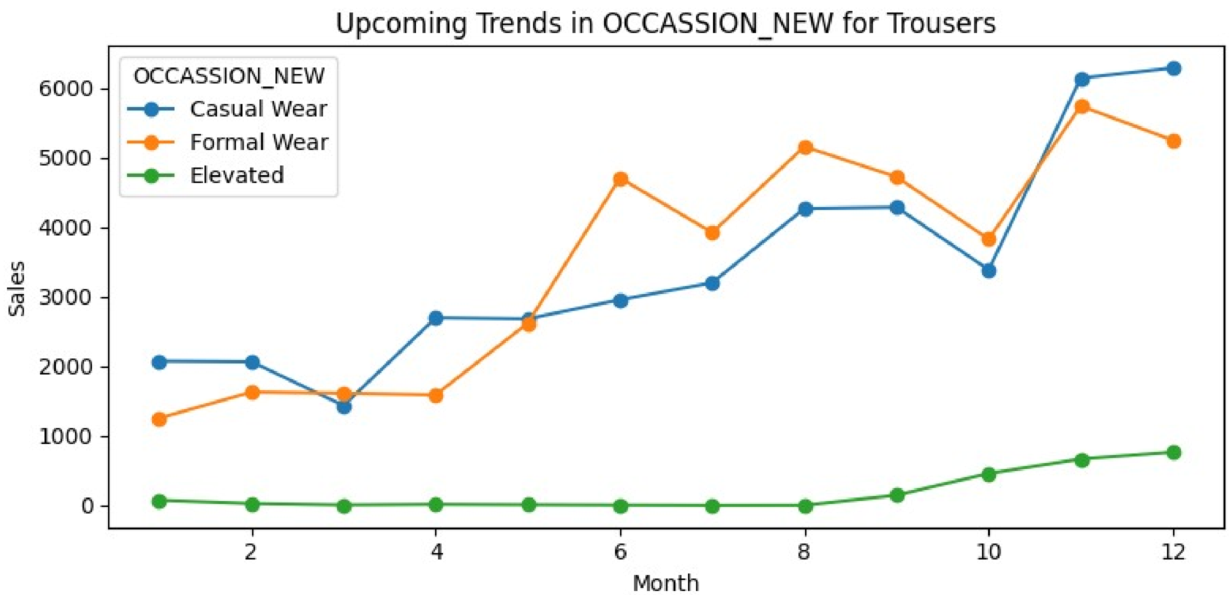
Upcoming Trends in FIT for Trousers
FIT
Slim Fit
Relaxed Fit
Regular Fit
Sales
0
1000
2000
3000
4000
5000
6000
7000
Month
2
4
6
8
10
12

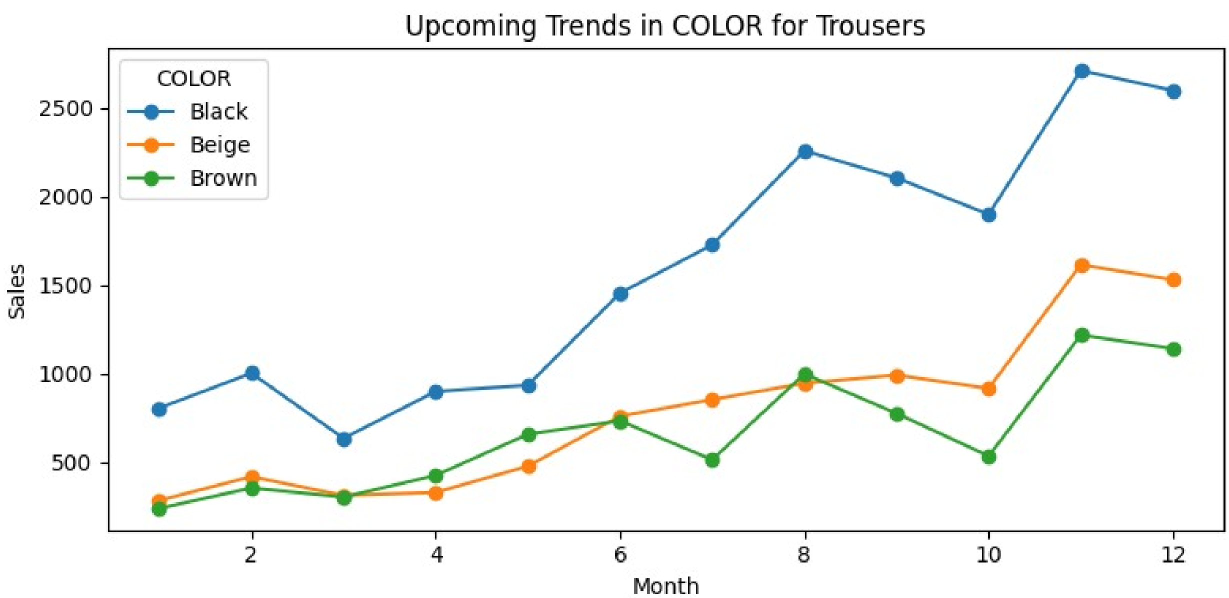

Upcoming Trends in MATERIAL_NEW for Trousers
MATERIAL_NEW
Polyester
Cotton
Others
Sales
Month
0
1000
2000
3000
4000
5000
6000
7000
2
4
6
8
10
12

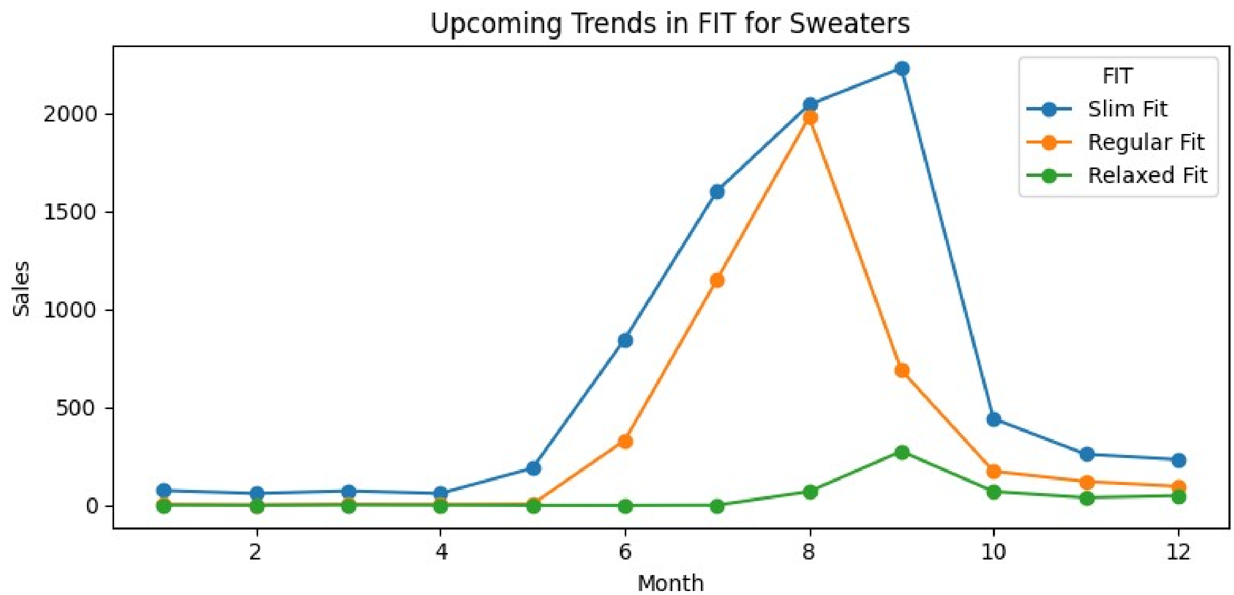

Upcoming Trends in PRINT_DESIGN for Trousers
PRINT_DESIGN
Plain
Solid
Self-Design, Textured
Sales
Month
0
2000
4000
6000
8000
10000
2
4
6
8
10
12


Upcoming Trends in SLEEVE_TYPE for Trousers
SLEEVE_TYPE
Unknown
Sales
12000
10000
8000
6000
4000
2
4
6
8
10
12
Month

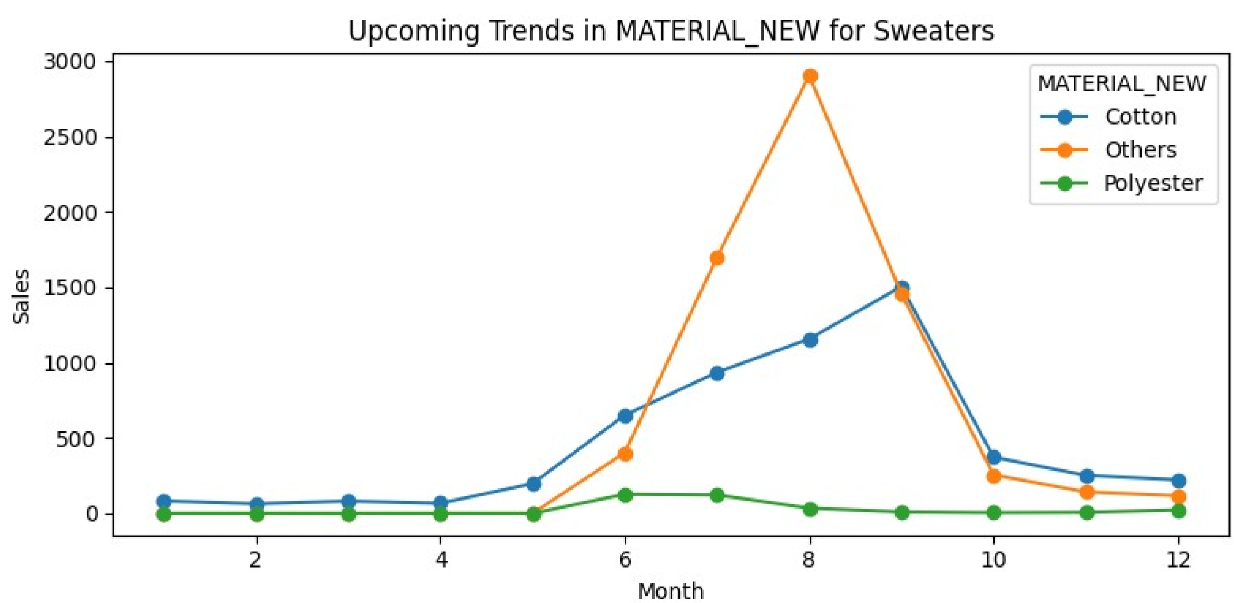

Upcoming Trends in COLLAR_NEW for Trousers
COLLAR_NEW
Others
Sales
12000
10000
8000
6000
4000
2
4
6
8
10
12
Month


Upcoming Trends in OCCASSION_NEW for Trousers
OCCASSION_NEW
Casual Wear
Formal Wear
Elevated
Sales
Month
0
1000
2000
3000
4000
5000
6000
2
4
6
8
10
12

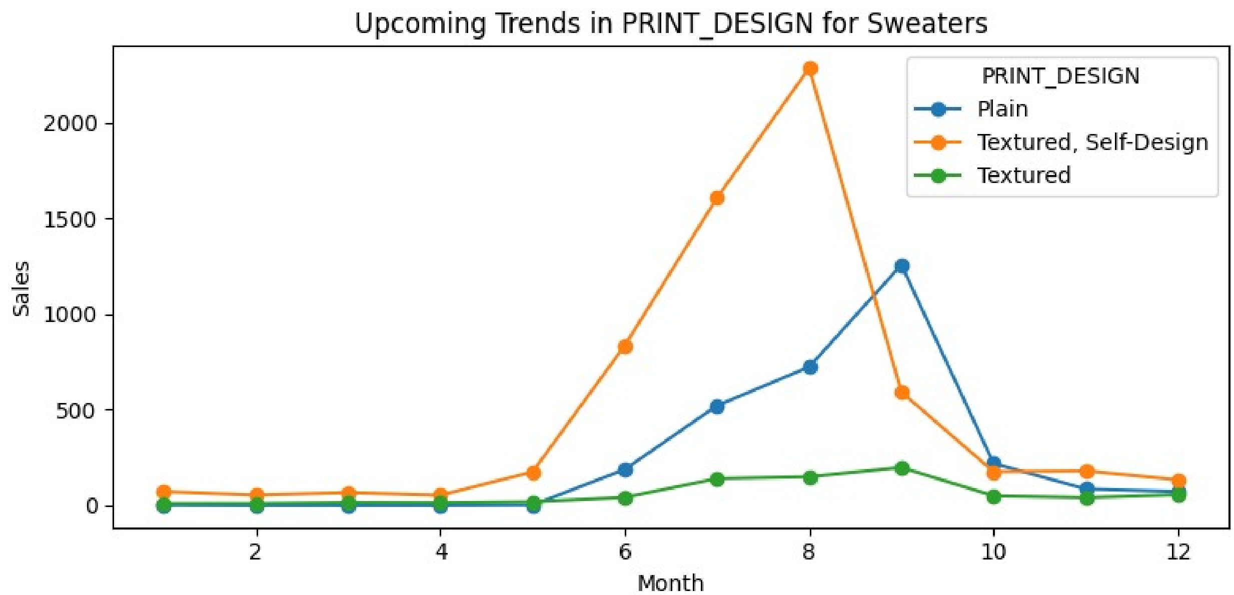

Upcoming Trends in COLOR for Trousers
COLOR
Black
Beige
Brown
Sales
Month
500
1000
1500
2000
2500
2
4
6
8
10
12

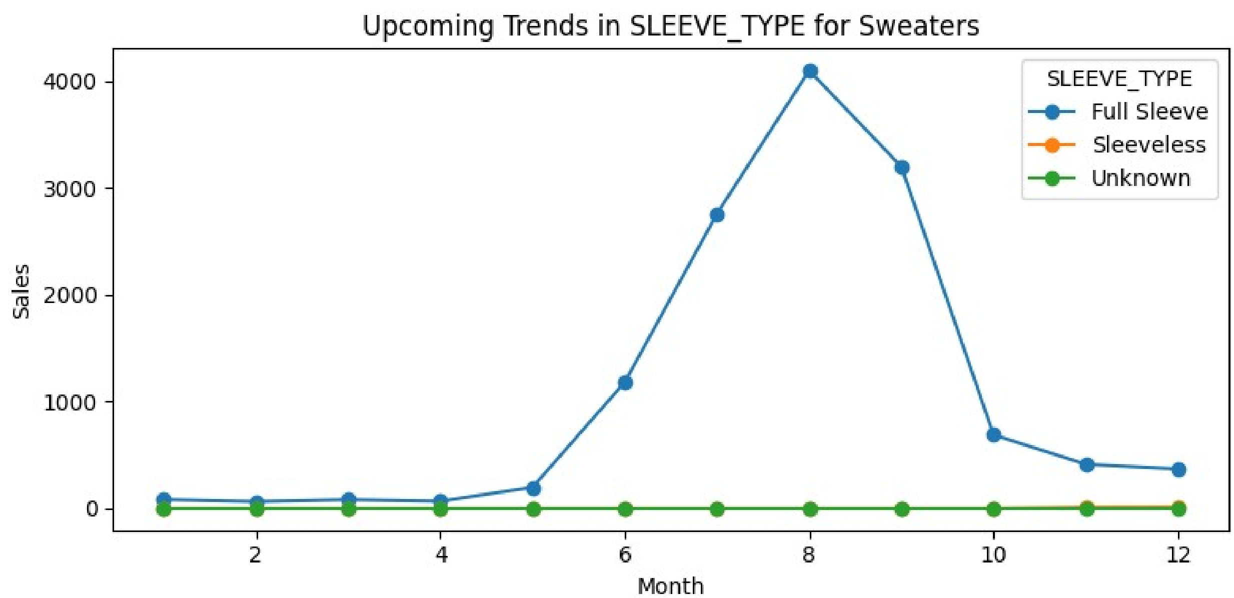


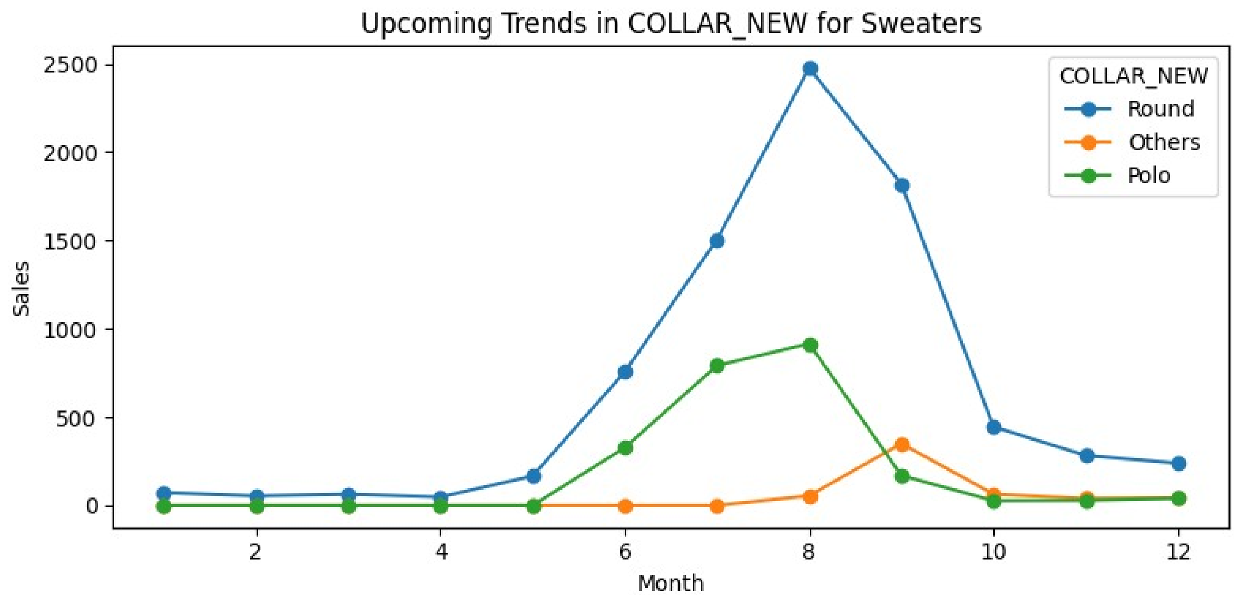
Upcoming Trends in FIT for Sweaters
FIT
Slim Fit
Regular Fit
Relaxed Fit
Sales
Month
0
500
1000
1500
2000
2
4
6
8
10
12

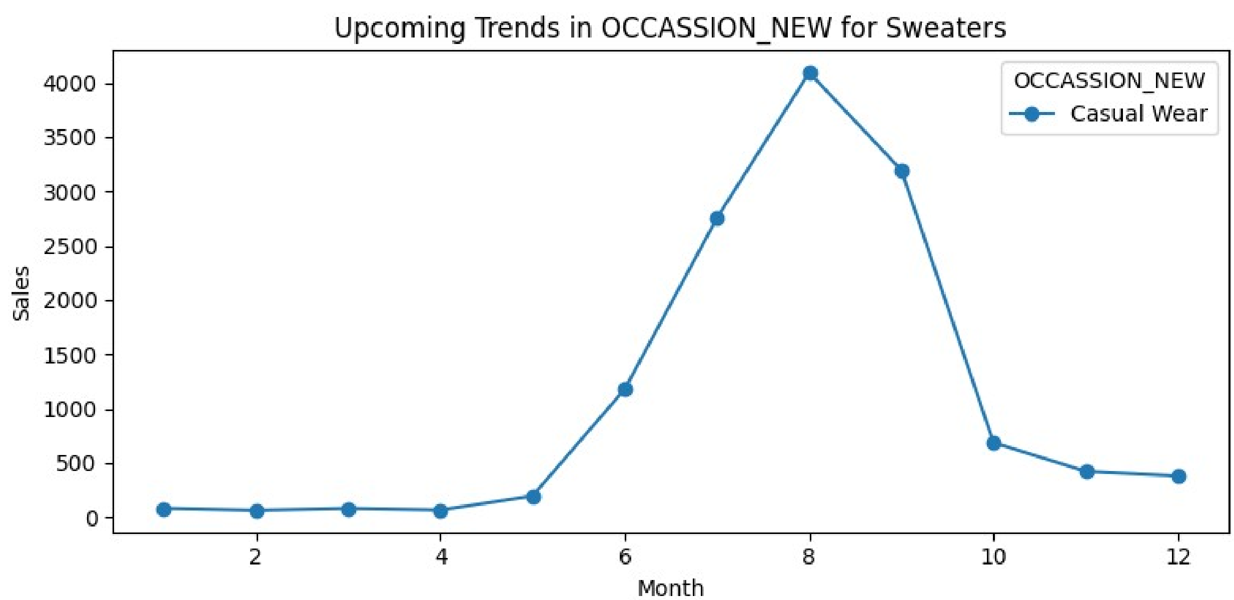


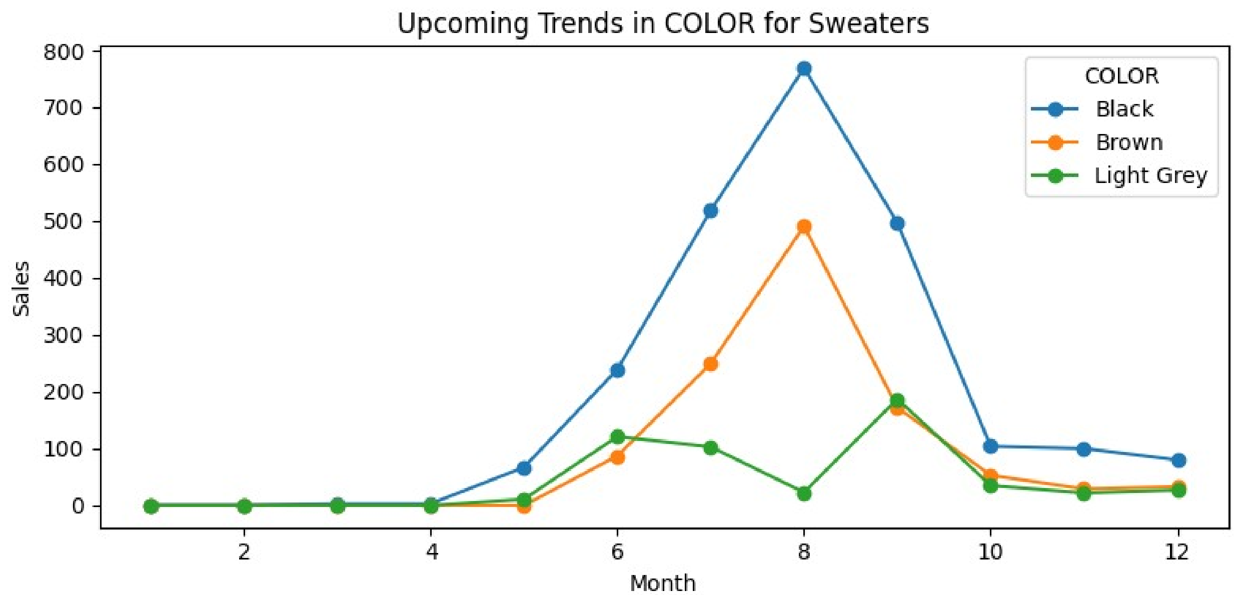
Upcoming Trends in MATERIAL_NEW for Sweaters
MATERIAL_NEW
Cotton
Others
Polyester
Sales
Month
0
500
1000
1500
2000
2500
3000
2
4
6
8
10
12


Upcoming Trends in PRINT_DESIGN for Sweaters
PRINT_DESIGN
Plain
Textured, Self-Design
Textured
Sales
Month
0
500
1000
1500
2000
2
4
6
8
10
12

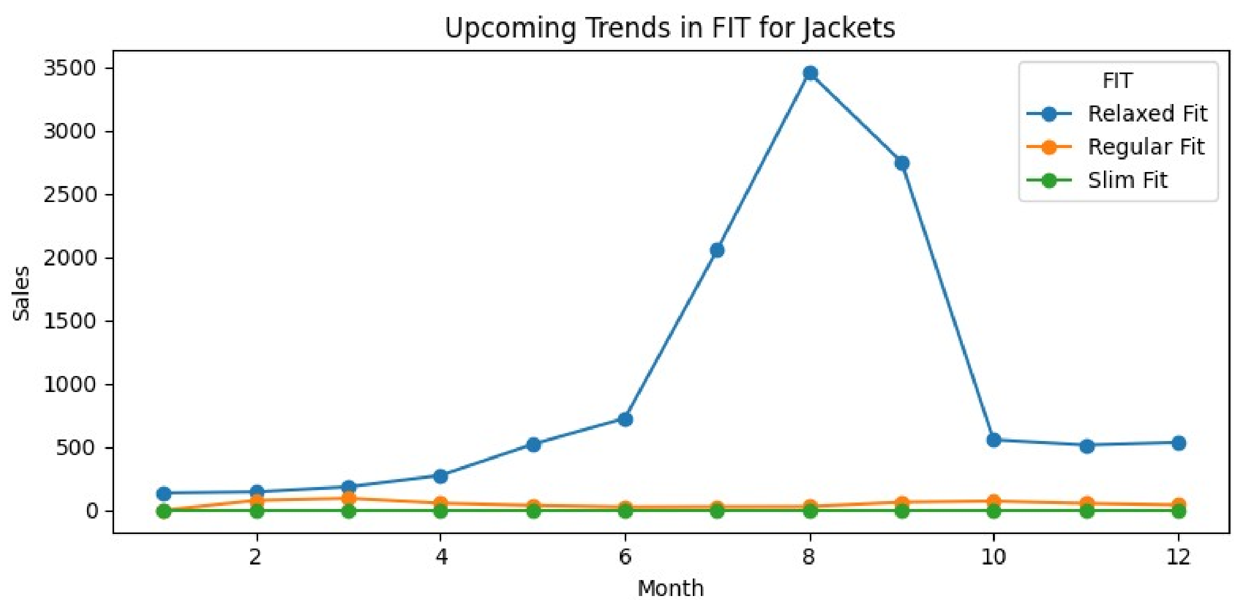

Upcoming Trends in SLEEVE_TYPE for Sweaters
SLEEVE_TYPE
Full Sleeve
Sleeveless
Unknown
Sales
Month
0
1000
2000
3000
4000
2
4
6
8
10
12


Upcoming Trends in COLLAR_NEW for Sweaters
COLLAR_NEW
Round
Others
Polo
Sales
Month
0
500
1000
1500
2000
2500
2
4
6
8
10
12

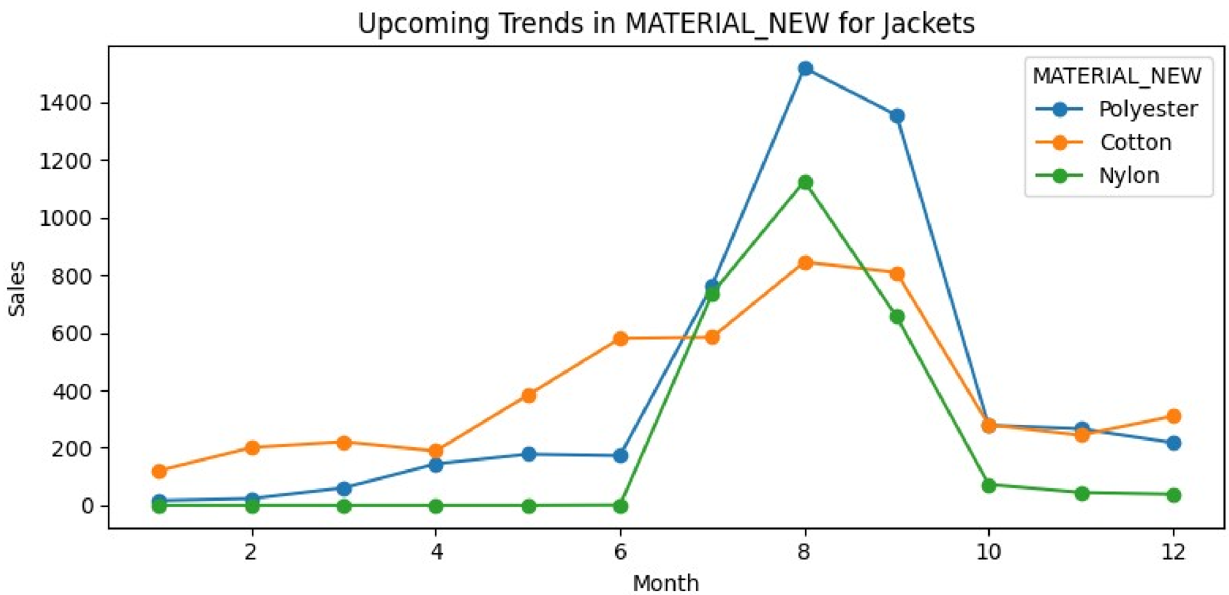

Upcoming Trends in OCCASSION_NEW for Sweaters
OCCASSION_NEW
Casual Wear
Sales
Month
0
500
1000
1500
2000
2500
3000
3500
4000
2
4
6
8
10
12

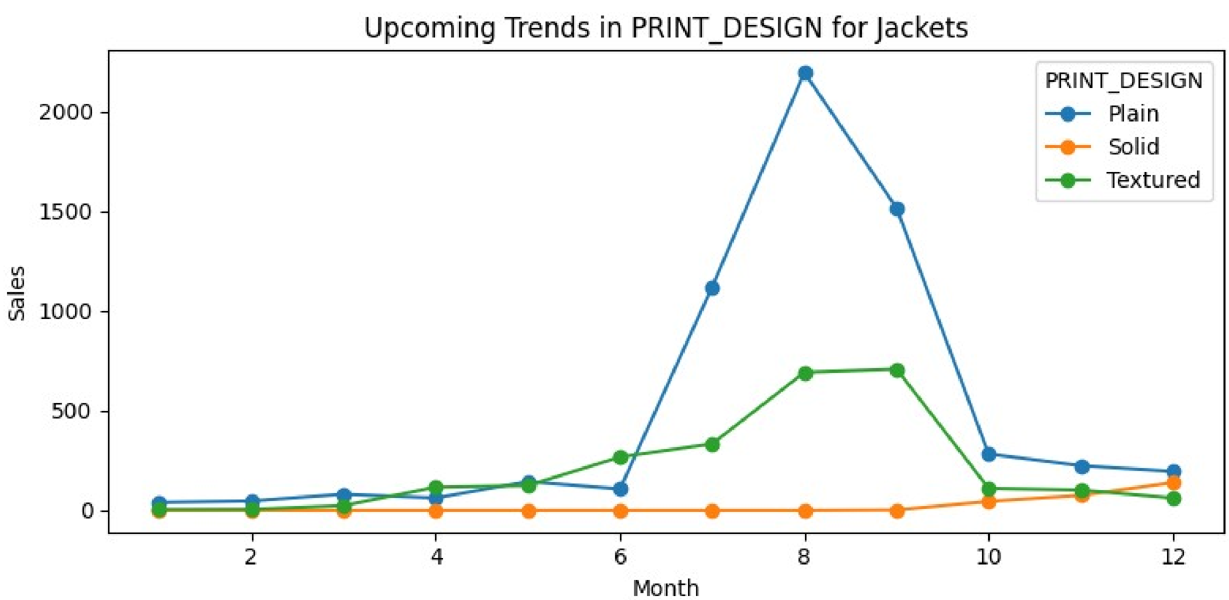


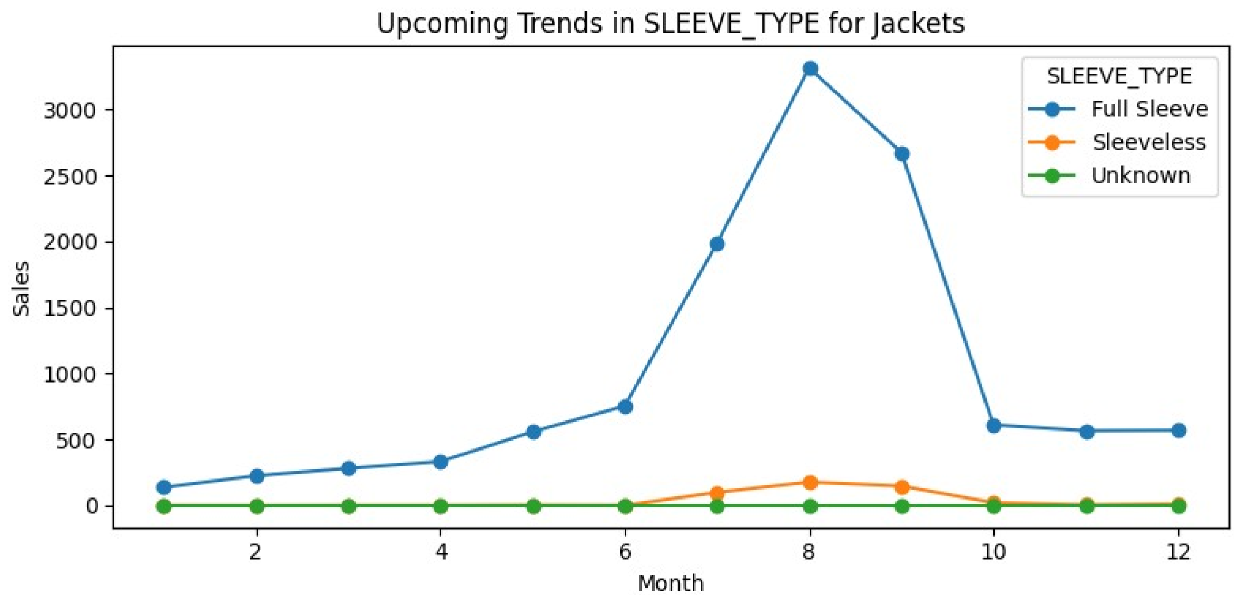
Upcoming Trends in COLOR for Sweaters
COLOR
Black
Brown
Light Grey
Sales
Month
800
700
600
500
400
300
200
100
0
2
4
6
8
10
12



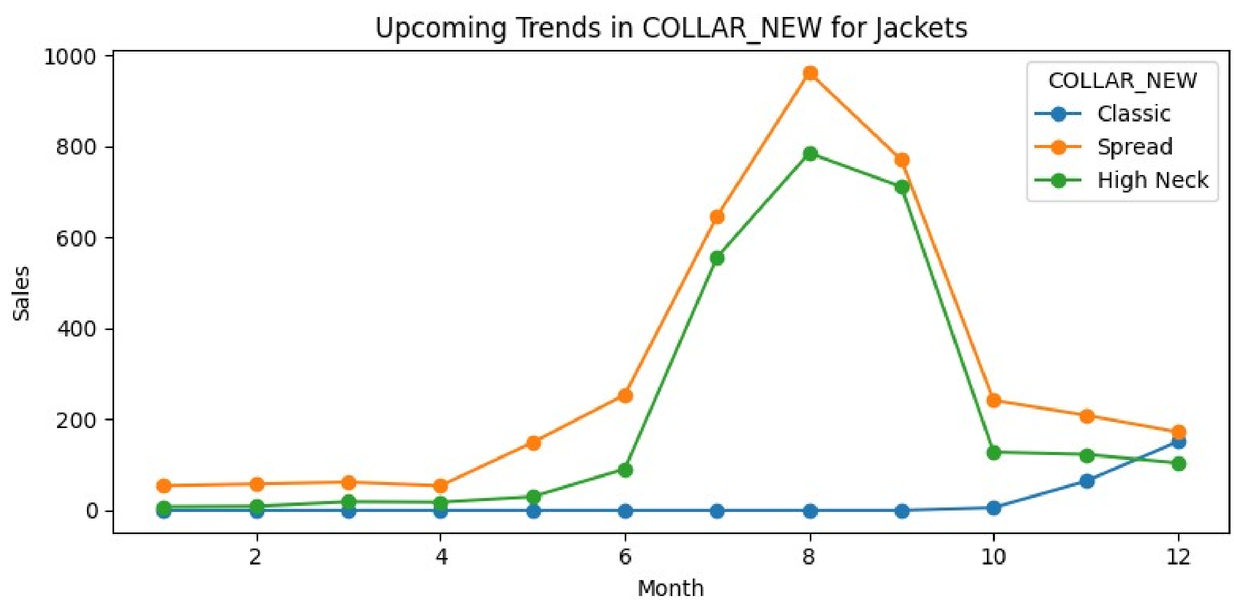
Upcoming Trends in FIT for Jackets
FIT
Relaxed Fit
Regular Fit
Slim Fit
Sales
Month
3500
3000
2500
2000
1500
1000
500
0
2
4
6
8
10
12


Upcoming Trends in MATERIAL_NEW for Jackets
MATERIAL_NEW
Polyester
Cotton
Nylon
Sales
0
200
400
600
800
1000
1200
1400
Month
2
4
6
8
10
12

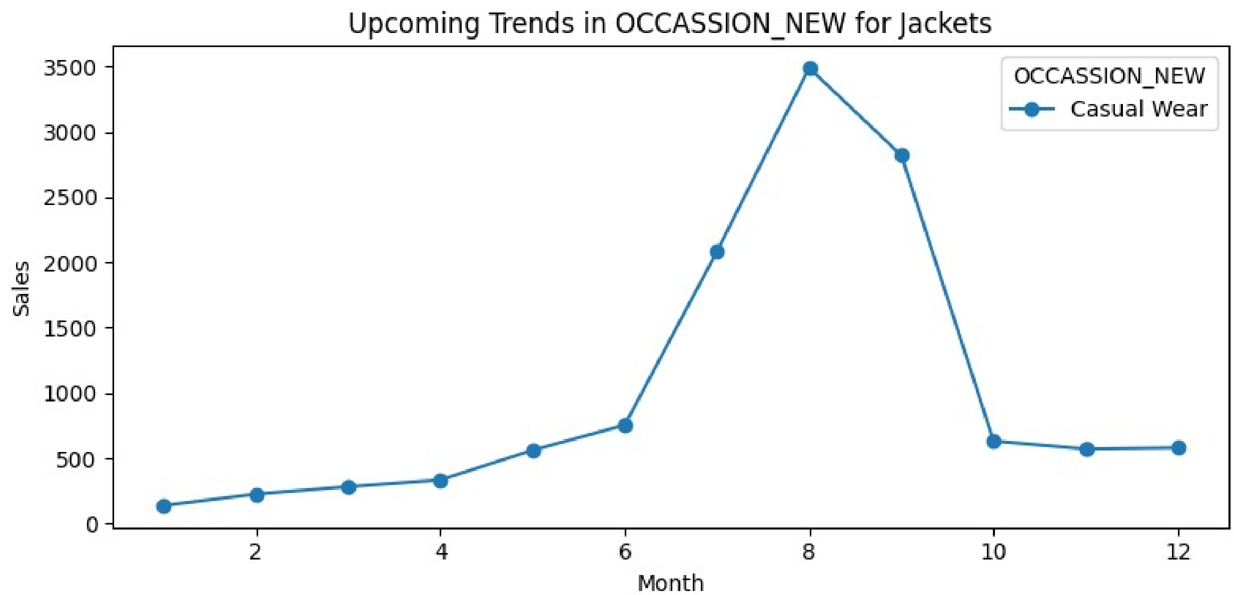

Upcoming Trends in PRINT_DESIGN for Jackets
PRINT_DESIGN
Plain
Solid
Textured
Sales
0
500
1000
1500
2000
Month
2
4
6
8
10
12

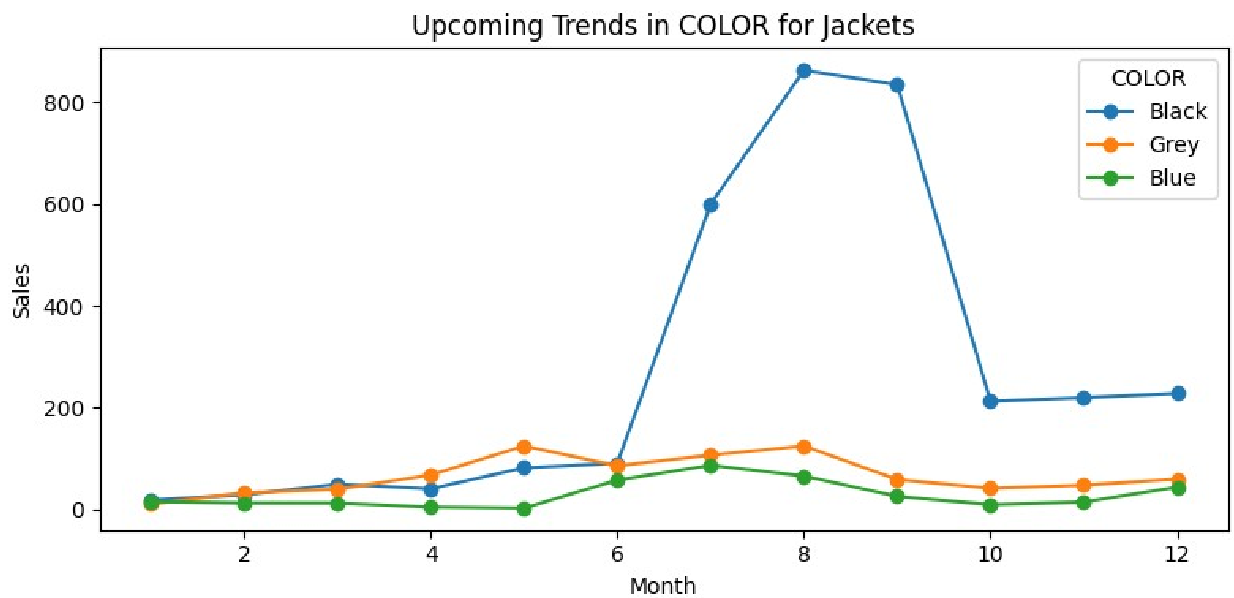


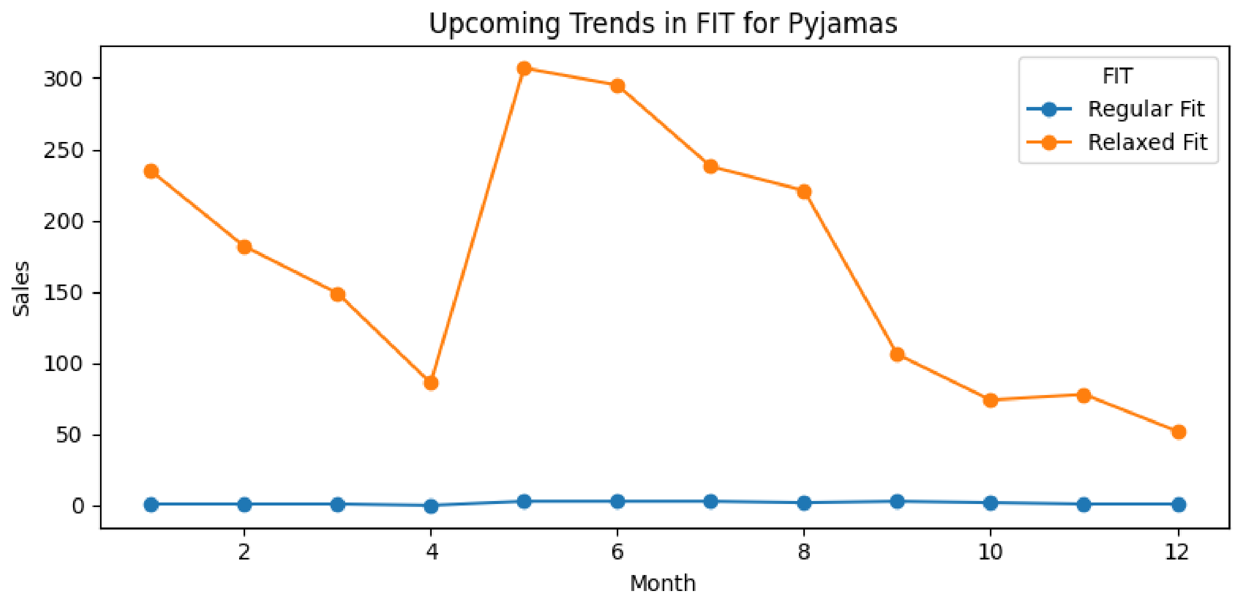
Upcoming Trends in SLEEVE_TYPE for Jackets
SLEEVE_TYPE
Full Sleeve
Sleeveless
Unknown
Sales
Month
0
500
1000
1500
2000
2500
3000
2
4
6
8
10
12

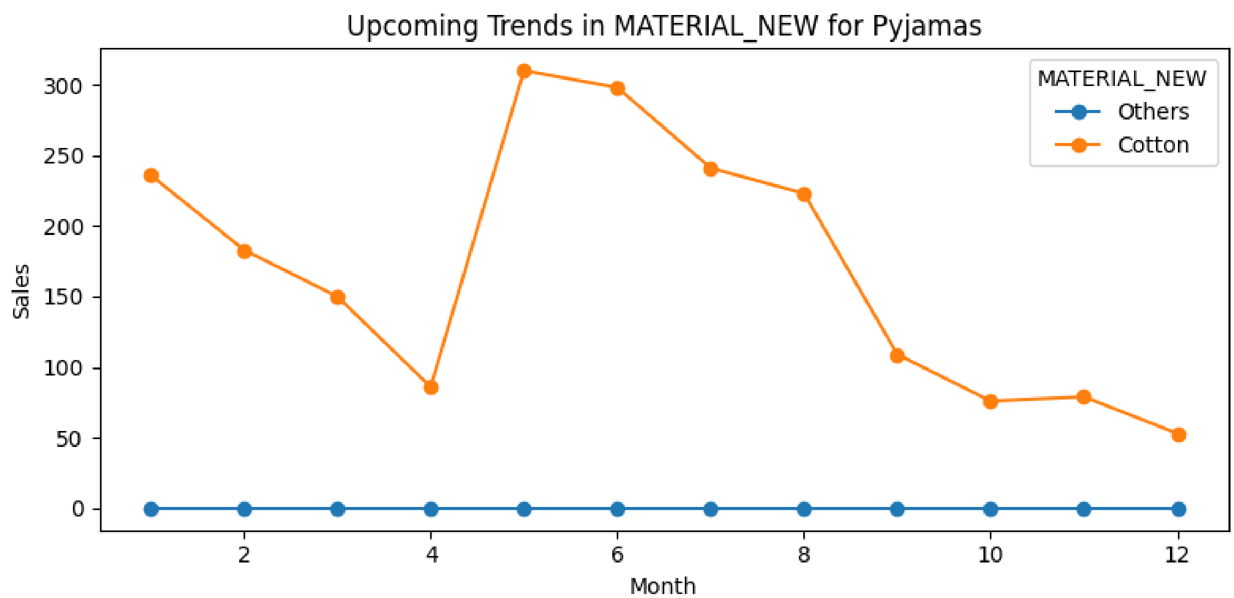


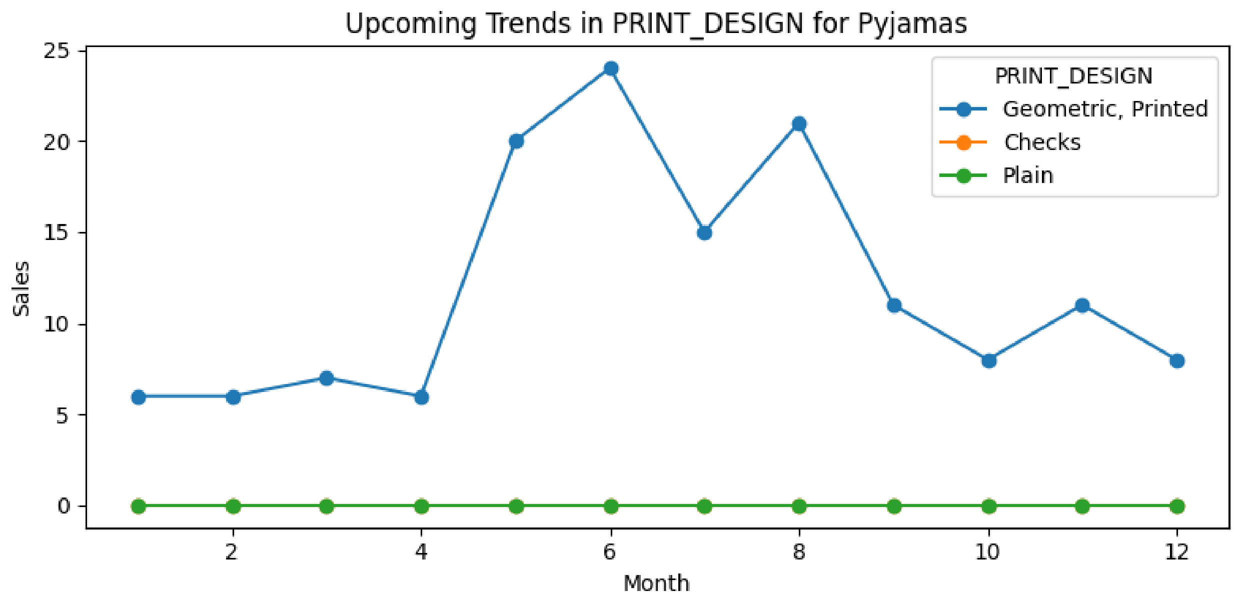
Upcoming Trends in COLLAR_NEW for Jackets
COLLAR_NEW
Classic
Spread
High Neck
Sales
Month
0
200
400
600
800
1000
2
4
6
8
10
12

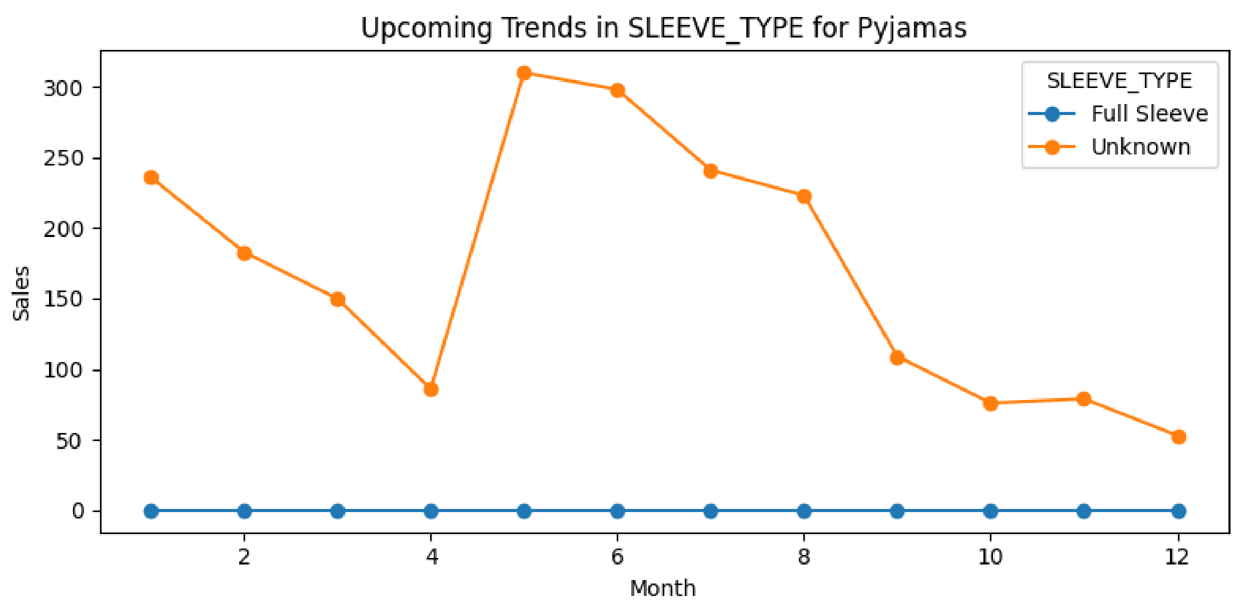

Upcoming Trends in OCCASSION_NEW for Jackets
OCCASSION_NEW
Casual Wear
Sales
Month
0
500
1000
1500
2000
2500
3000
3500
2
4
6
8
10
12

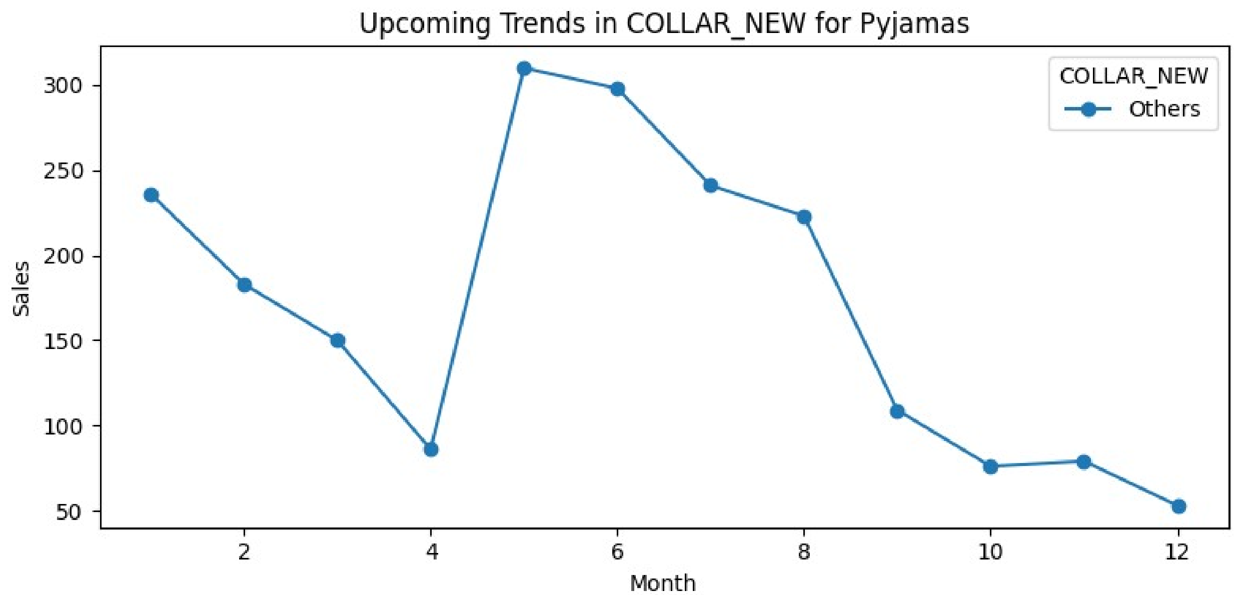

Upcoming Trends in COLOR for Jackets
COLOR
Black
Grey
Blue
Sales
Month
0
200
400
600
800
2
4
6
8
10
12

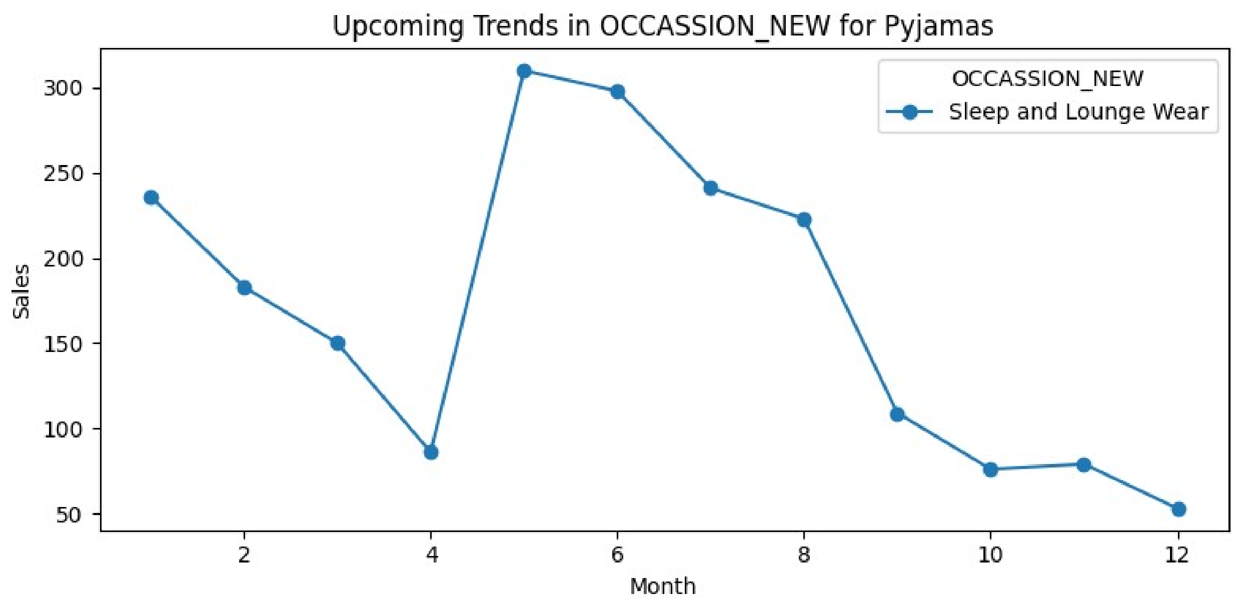

Upcoming Trends in FIT for Pyjamas
FIT
Regular Fit
Relaxed Fit
Sales
Month
0
50
100
150
200
250
300
2
4
6
8
10
12

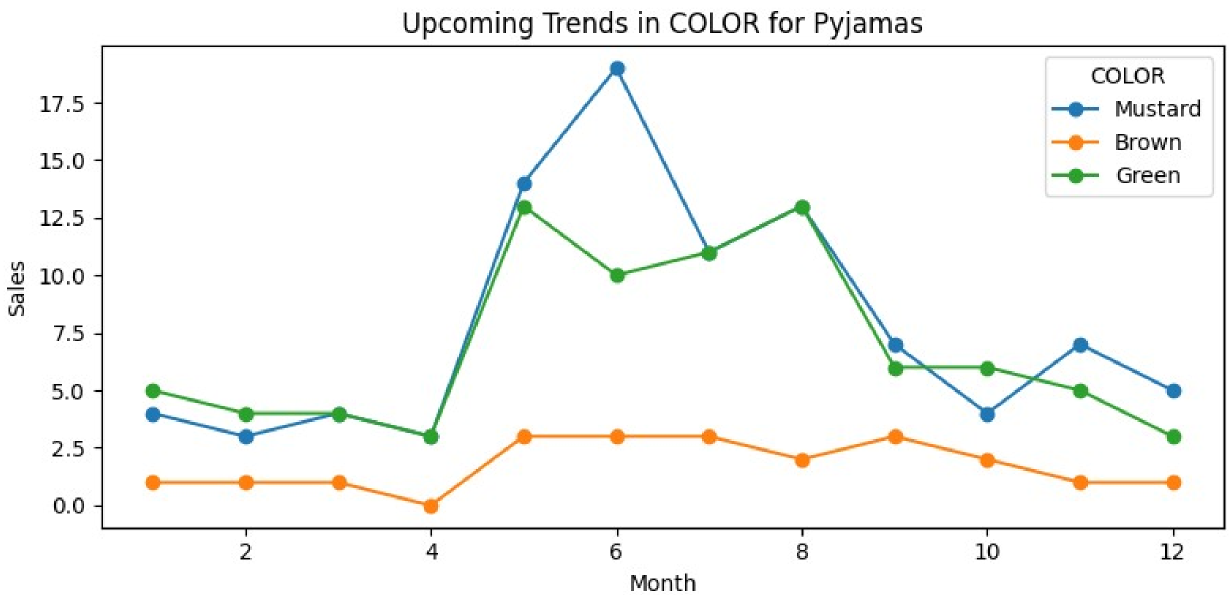

Upcoming Trends in MATERIAL_NEW for Pyjamas
MATERIAL_NEW
Others
Cotton
Sales
Month
0
50
100
150
200
250
300
2
4
6
8
10
12


Upcoming Trends in PRINT_DESIGN for Pyjamas
PRINT_DESIGN
Geometric, Printed
Checks
Plain
Sales
Month
0
5
10
15
20
25
2
4
6
8
10
12

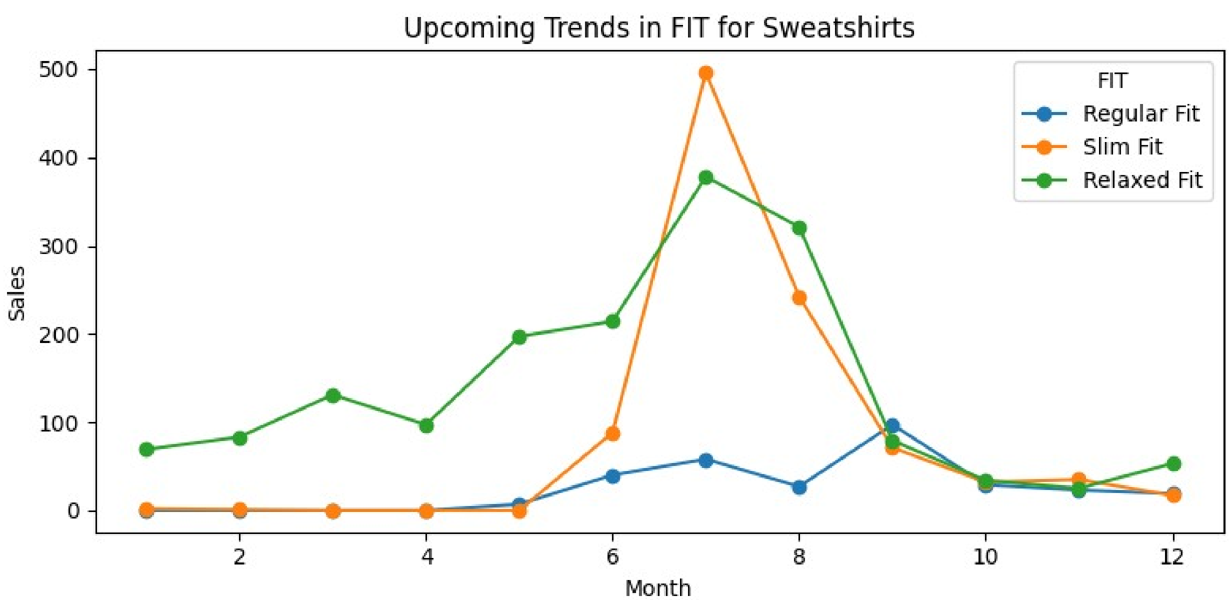

Upcoming Trends in SLEEVE_TYPE for Pyjamas
SLEEVE_TYPE
Full Sleeve
Unknown
Sales
Month
0
50
100
150
200
250
300
2
4
6
8
10
12


Upcoming Trends in COLLAR_NEW for Pyjamas
COLLAR_NEW
Others
Sales
Month
300
250
200
150
100
50
2
4
6
8
10
12

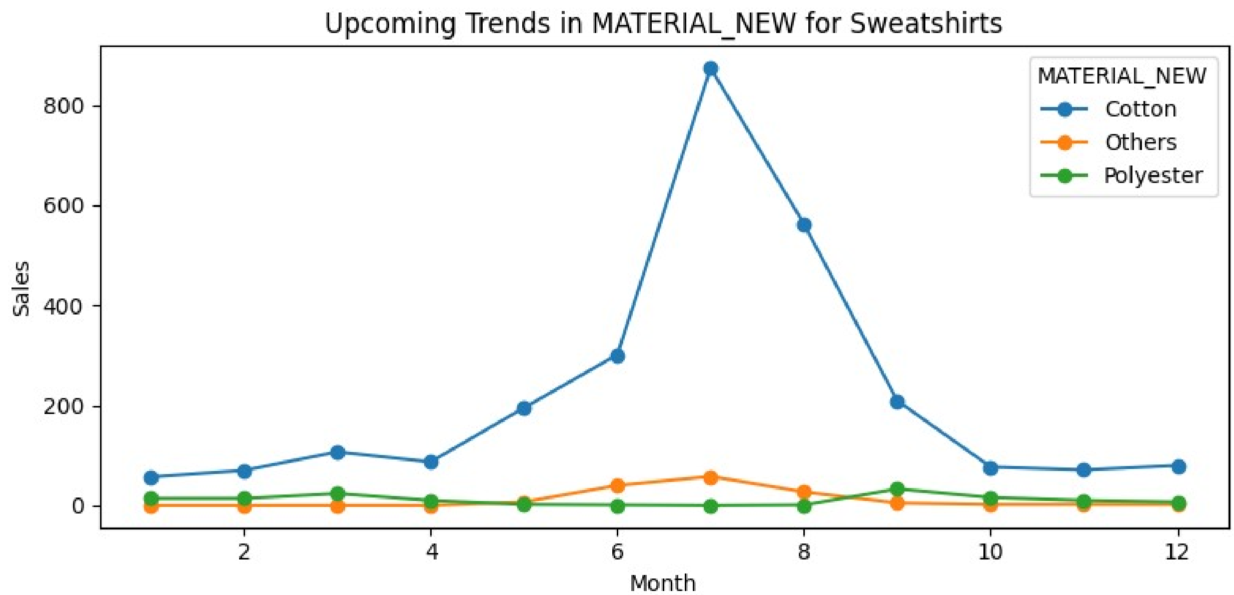

Upcoming Trends in OCCASSION_NEW for Pyjamas
OCCASSION_NEW
Sleep and Lounge Wear
Sales
Month
300
250
200
150
100
50
2
4
6
8
10
12

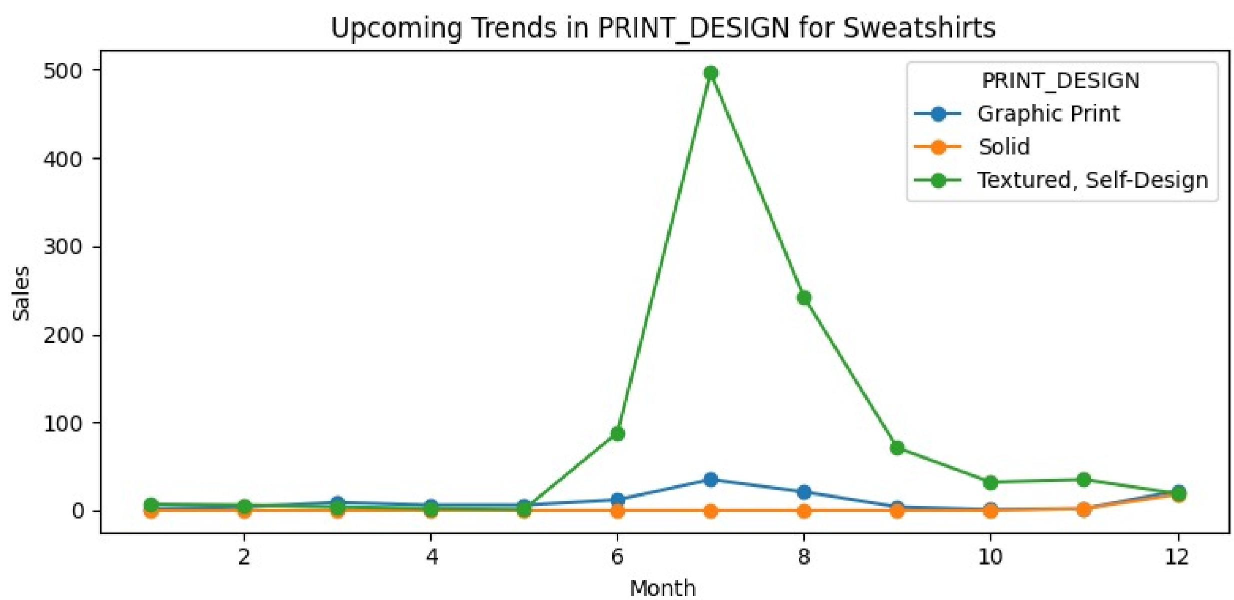


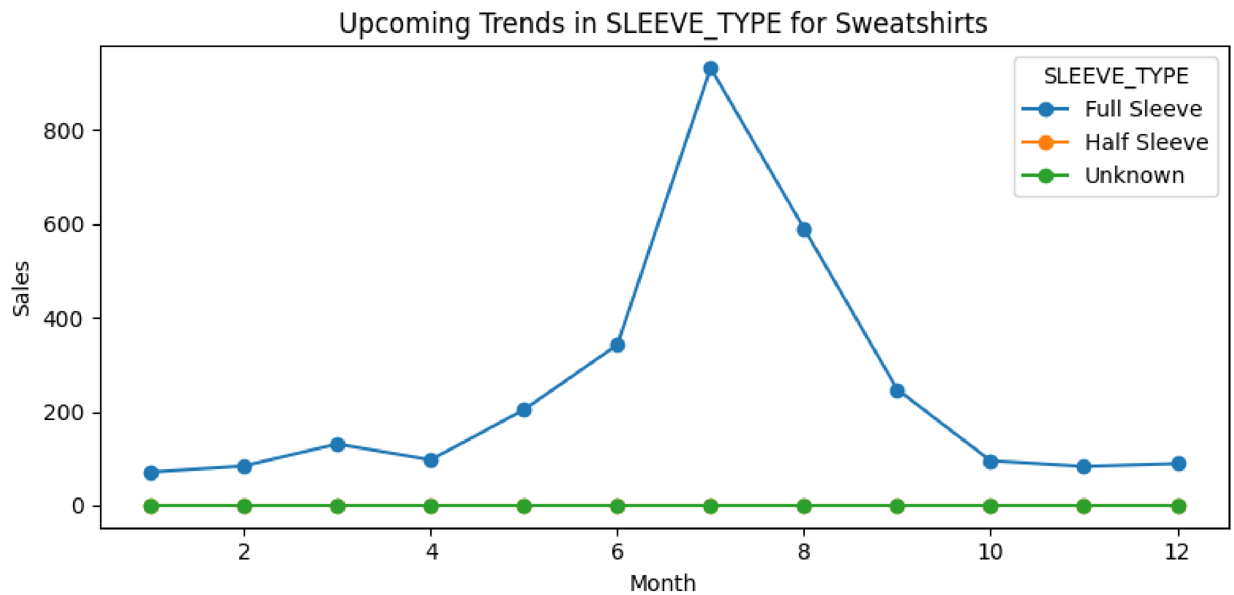
Upcoming Trends in COLOR for Pyjamas
COLOR
Mustard
Brown
Green
Sales
Month
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
2
4
6
8
10
12



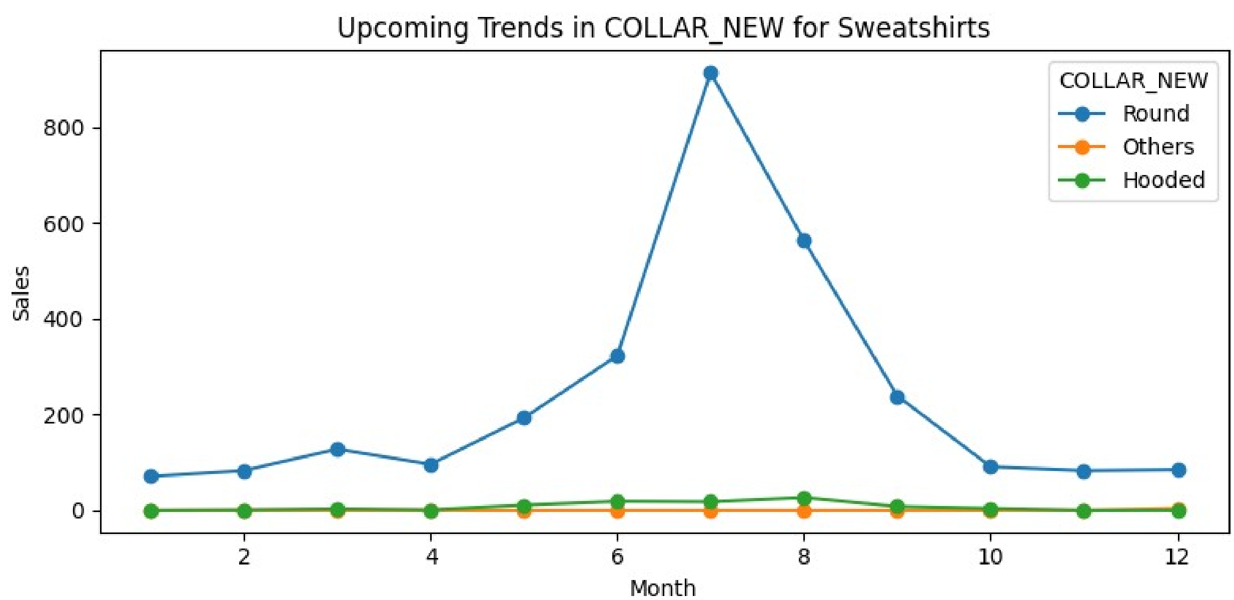
Upcoming Trends in FIT for Sweatshirts
FIT
Regular Fit
Slim Fit
Relaxed Fit
Sales
Month
0
100
200
300
400
500
2
4
6
8
10
12


Upcoming Trends in MATERIAL_NEW for Sweatshirts
MATERIAL_NEW
Cotton
Others
Polyester
Sales
Month
0
200
400
600
800
2
4
6
8
10
12

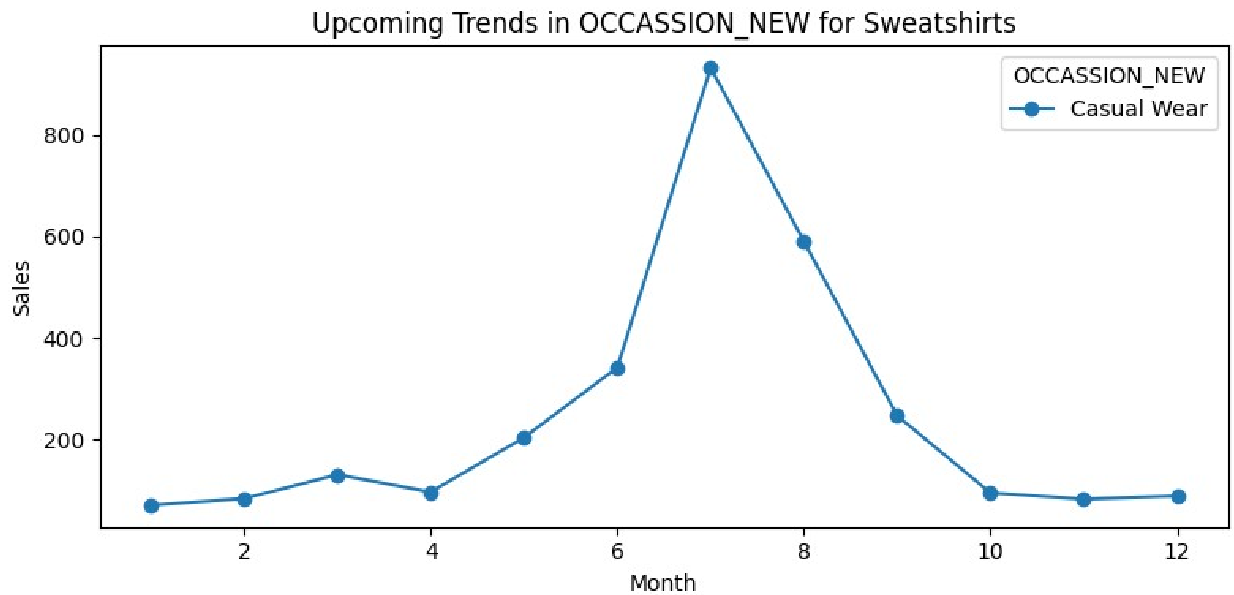

Upcoming Trends in PRINT_DESIGN for Sweatshirts
PRINT_DESIGN
Graphic Print
Solid
Textured, Self-Design
Sales
Month
0
100
200
300
400
500
2
4
6
8
10
12

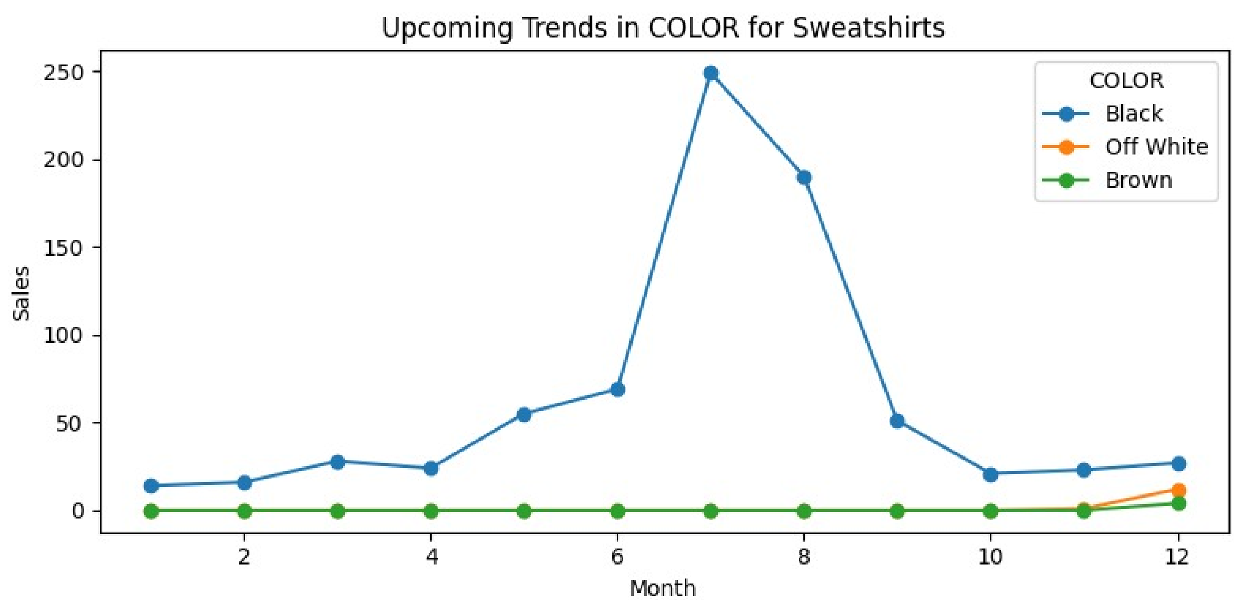


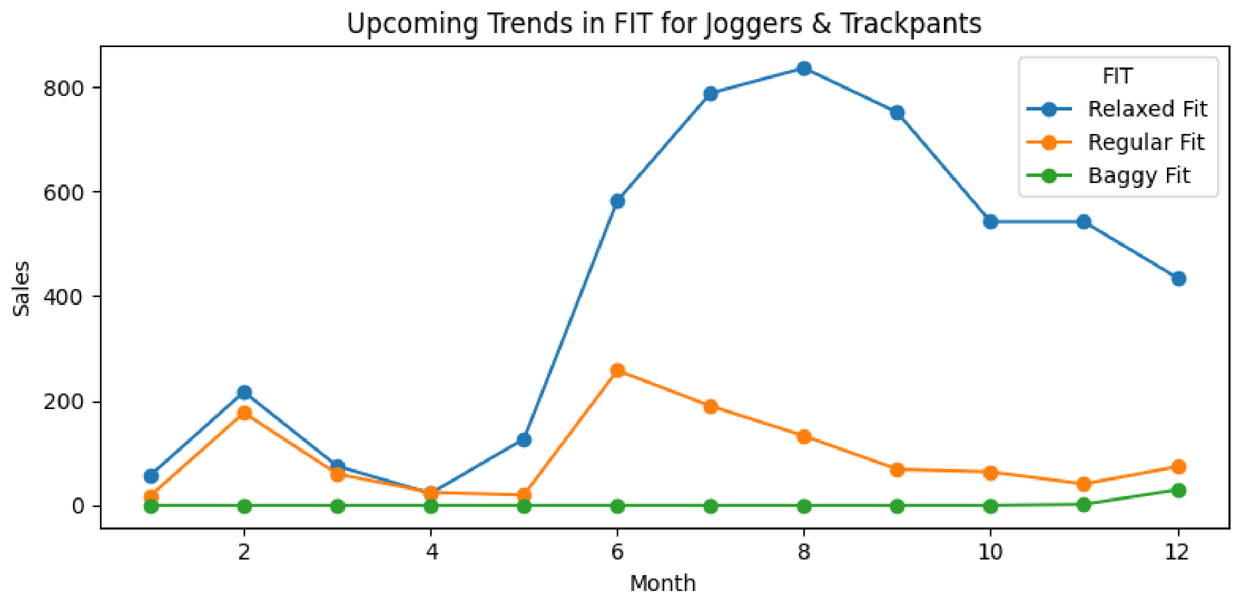
Upcoming Trends in SLEEVE_TYPE for Sweatshirts
SLEEVE_TYPE
Full Sleeve
Half Sleeve
Unknown
Sales
Month
0
200
400
600
800
2
4
6
8
10
12

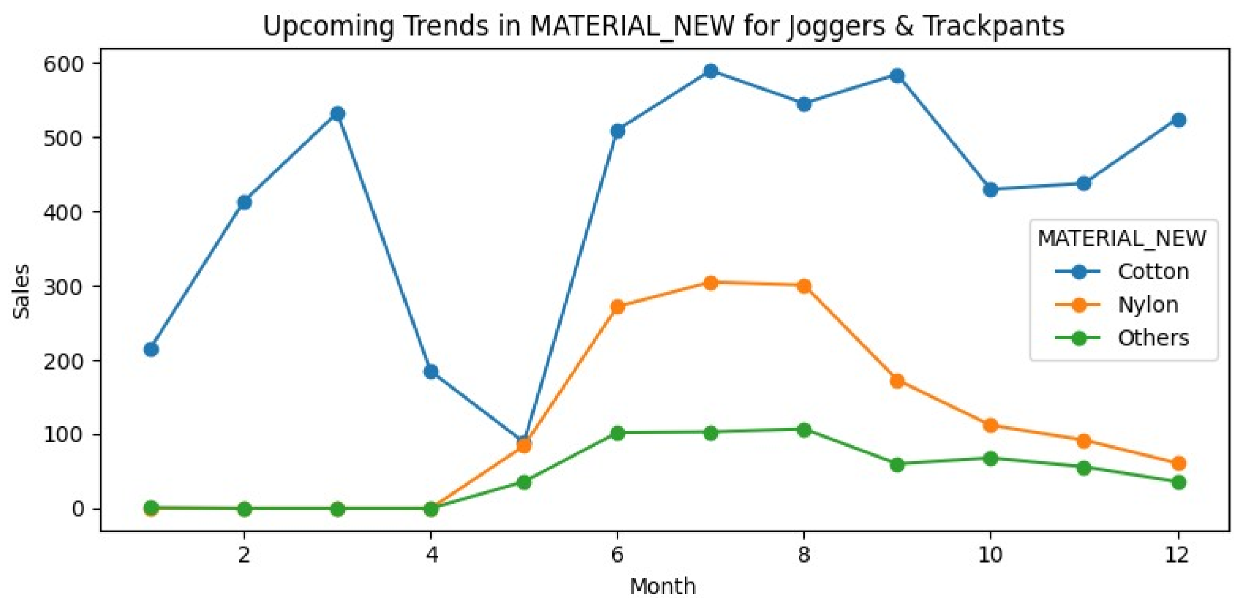


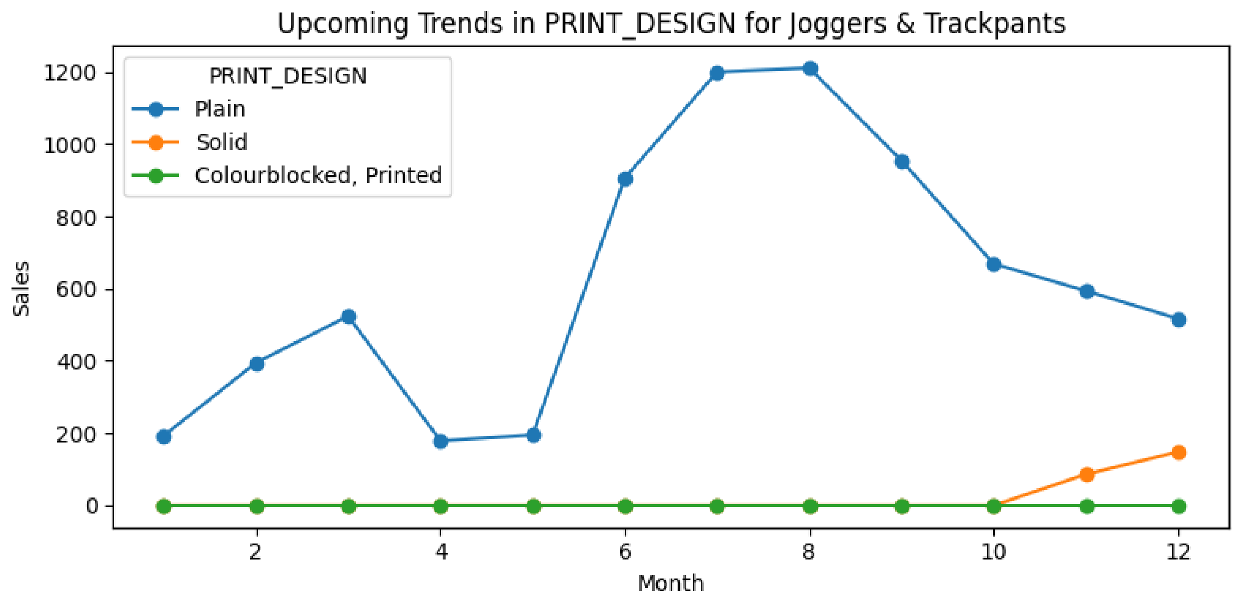
Upcoming Trends in COLLAR_NEW for Sweatshirts
COLLAR_NEW
Round
Others
Hooded
Sales
Month
0
200
400
600
800
2
4
6
8
10
12


Upcoming Trends in OCCASSION_NEW for Sweatshirts
OCCASSION_NEW
Casual Wear
Sales
Month
200
400
600
800
2
4
6
8
10
12

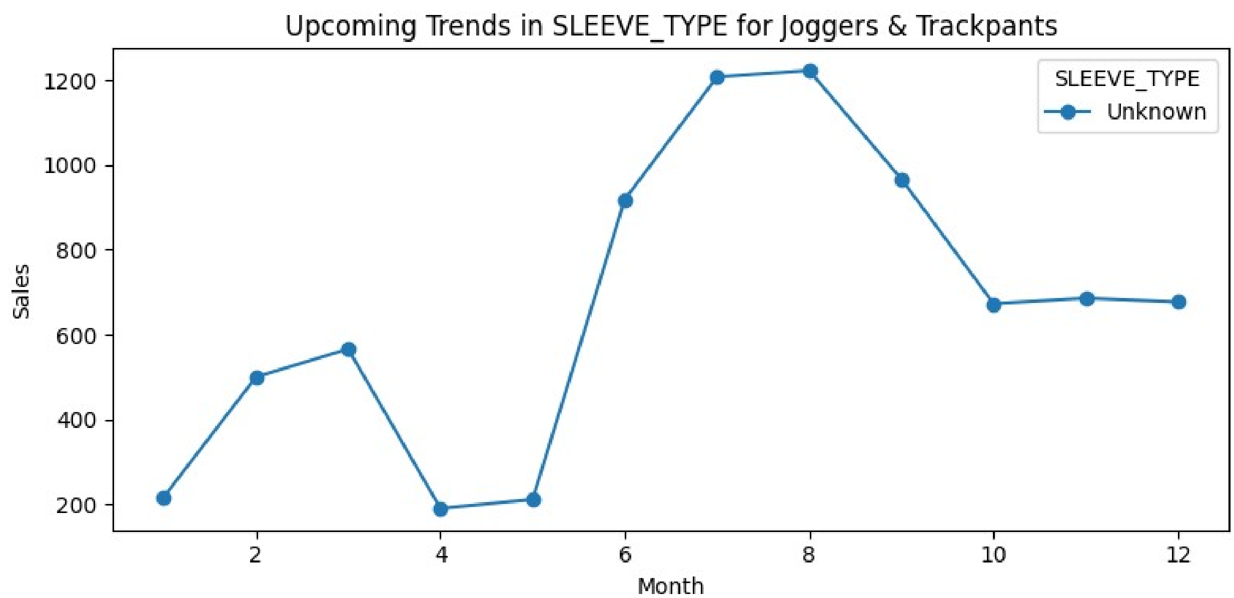

Upcoming Trends in COLOR for Sweatshirts
COLOR
Black
Off White
Brown
Sales
Month
0
50
100
150
200
250
2
4
6
8
10
12


Upcoming Trends in FIT for Joggers & Trackpants
FIT
Relaxed Fit
Regular Fit
Baggy Fit
Sales
Month
0
200
400
600
800
2
4
6
8
10
12

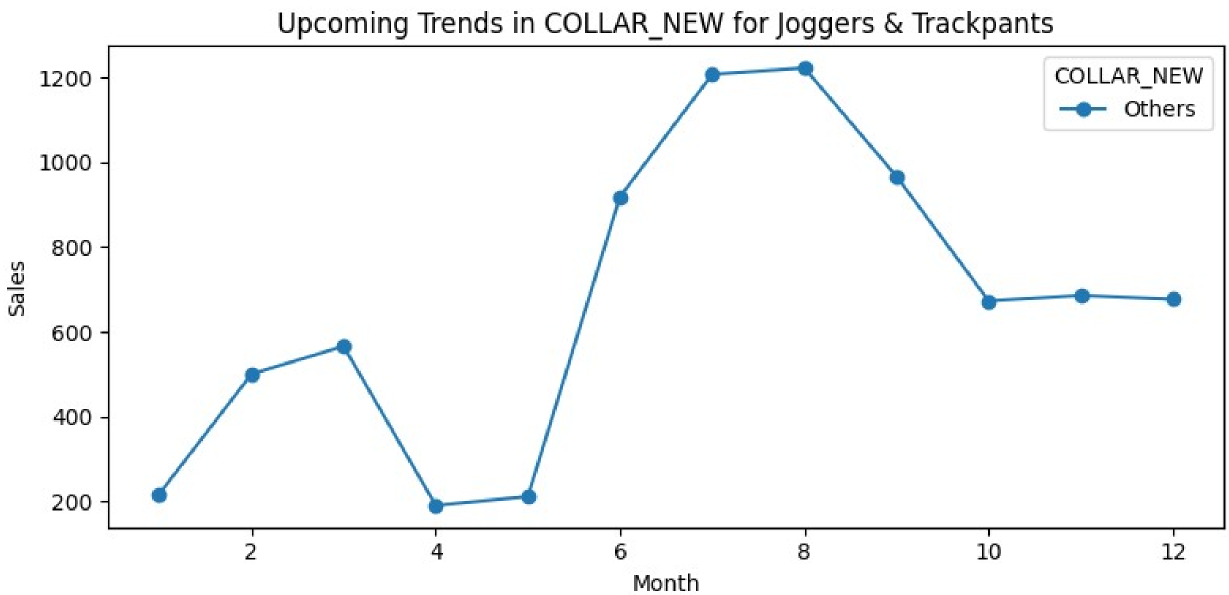

Upcoming Trends in MATERIAL_NEW for Joggers & Trackpants
MATERIAL_NEW
Cotton
Nylon
Others
Sales
Month
0
100
200
300
400
500
600
2
4
6
8
10
12

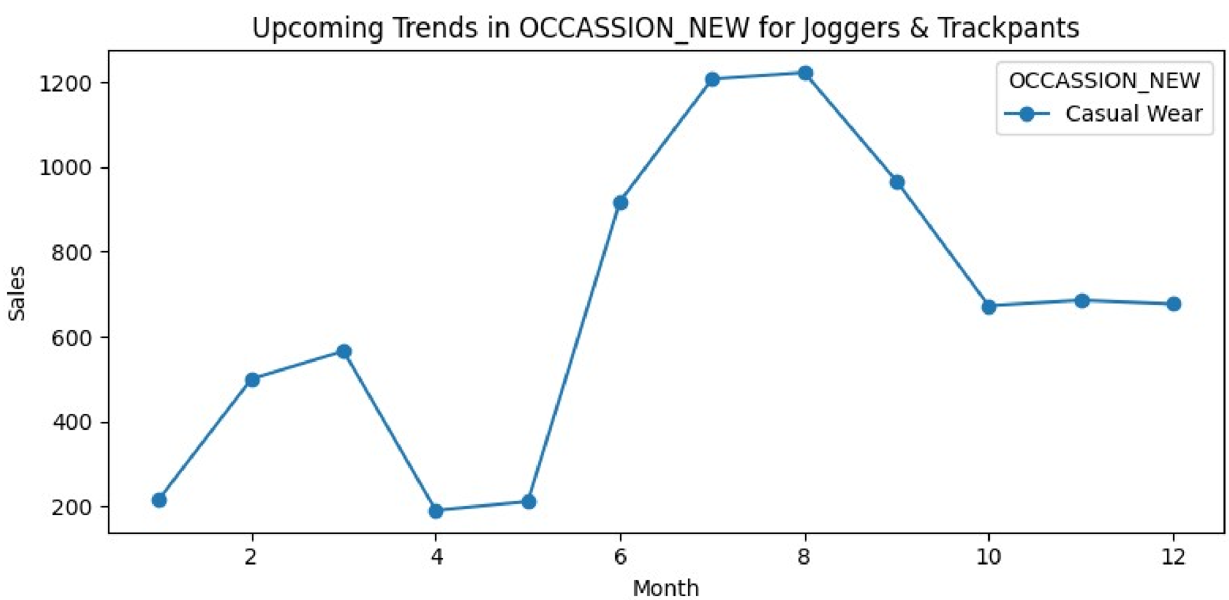


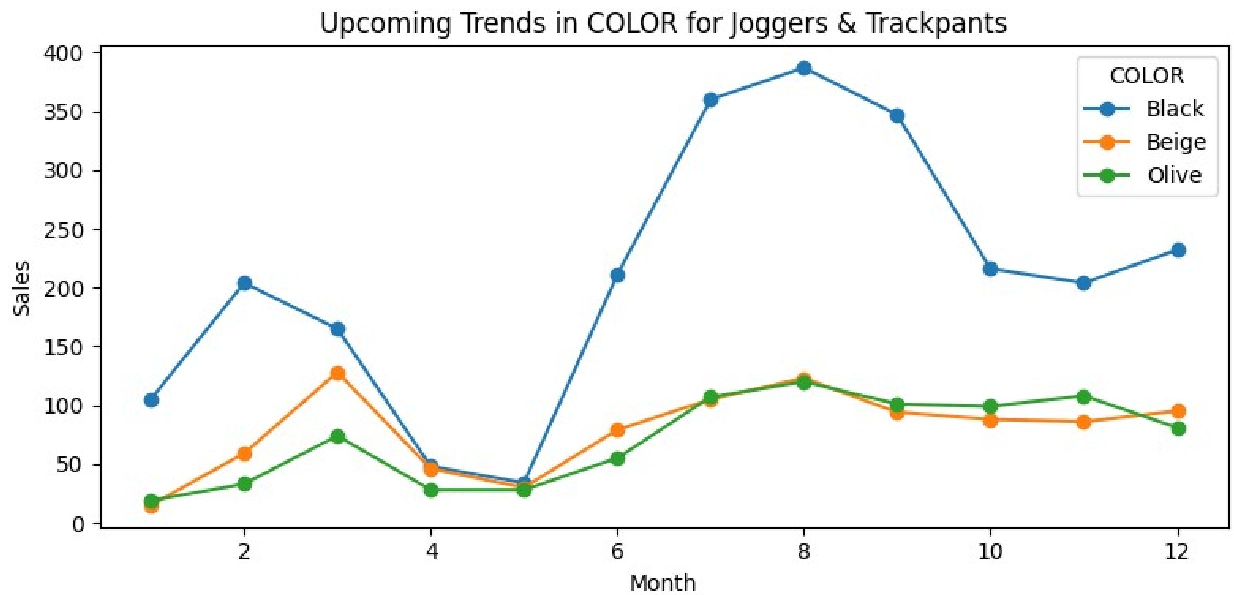
Upcoming Trends in PRINT_DESIGN for Joggers & Trackpants
PRINT_DESIGN
Plain
Solid
Colourblocked, Printed
Sales
Month



Upcoming Trends in SLEEVE_TYPE for Joggers & Trackpants
SLEEVE_TYPE
Unknown
Sales
Month


Upcoming Trends in COLLAR_NEW for Joggers & Trackpants
COLLAR_NEW
Others
Sales
Month
200
400
600
800
1000
1200
2
4
6
8
10
12


Upcoming Trends in OCCASSION_NEW for Joggers & Trackpants
OCCASSION_NEW
Casual Wear
Sales
Month
200
400
600
800
1000
1200
2
4
6
8
10
12



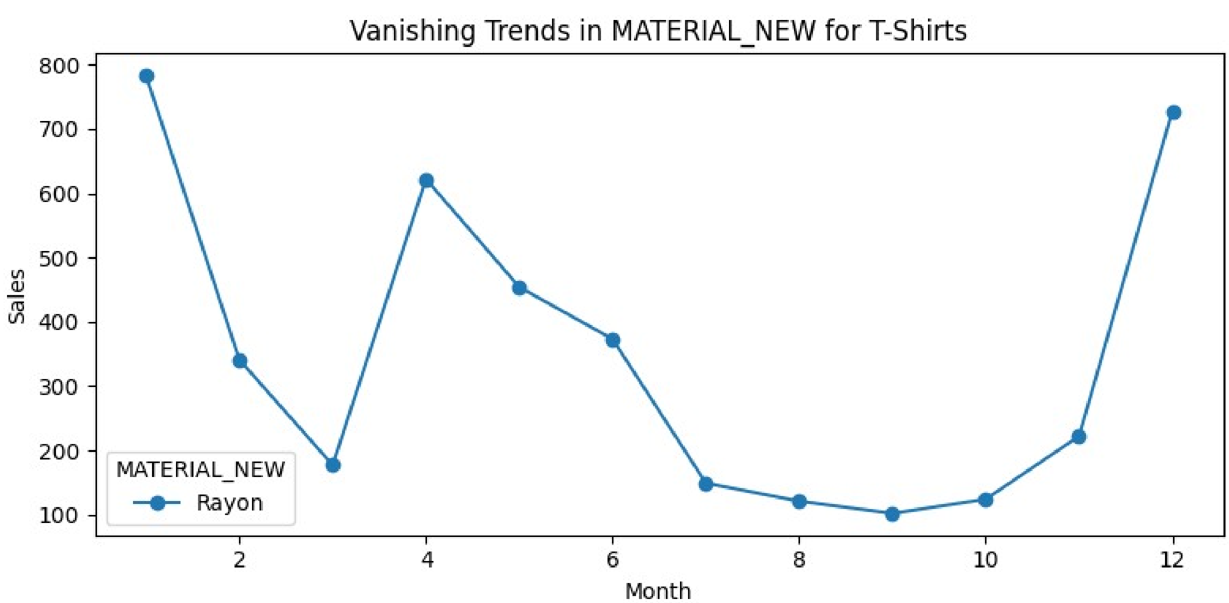
Q3. What are the trends that are vanishing with time.

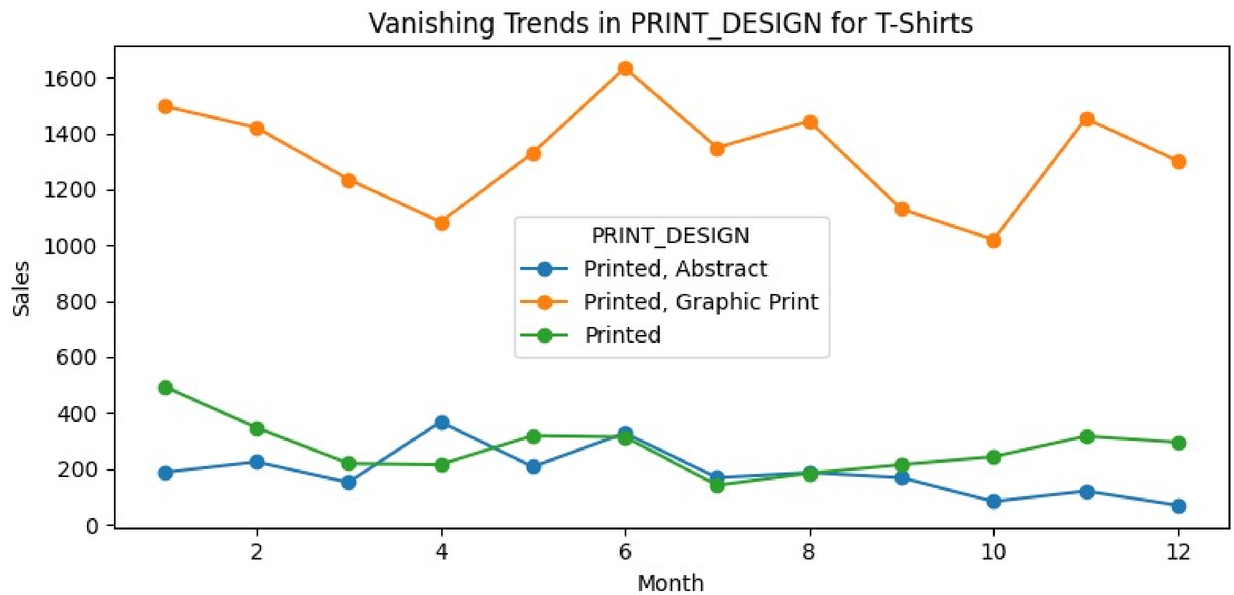
Approach -> The dataset was transformed into a long format to enable time-series analysis of monthly sales. For each combination of category and feature, I analyzed the trend in sales over time by computing the slope of the best-fit line using np.polyfit. Feature values with the steepest negative slopes were identified as those exhibiting the sharpest decline in sales — indicating vanishing trends. The top three such values were selected and their sales trajectories were visualized to clearly illustrate the downward trend.

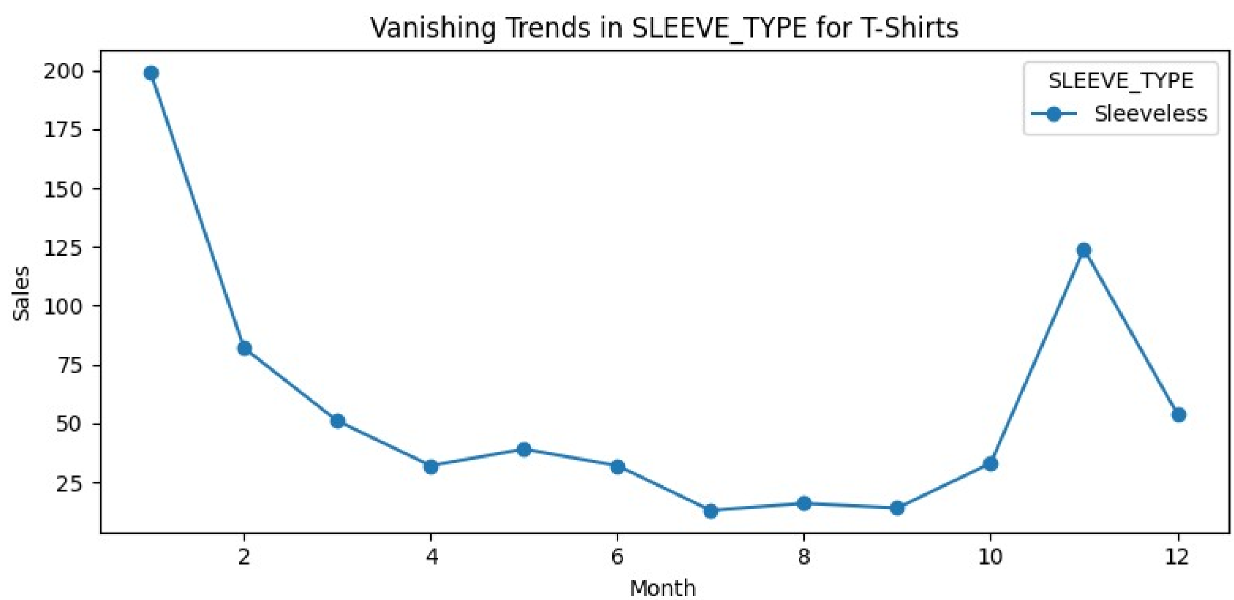
Note: Some graphs display only one plot because the code filters for the top 3 feature values with the most negative sales slopes per category and feature. If fewer than three values have a negative slope (i.e., vanishing trend), or only one meets this condition, the plot will show only those available series, sometimes just one.

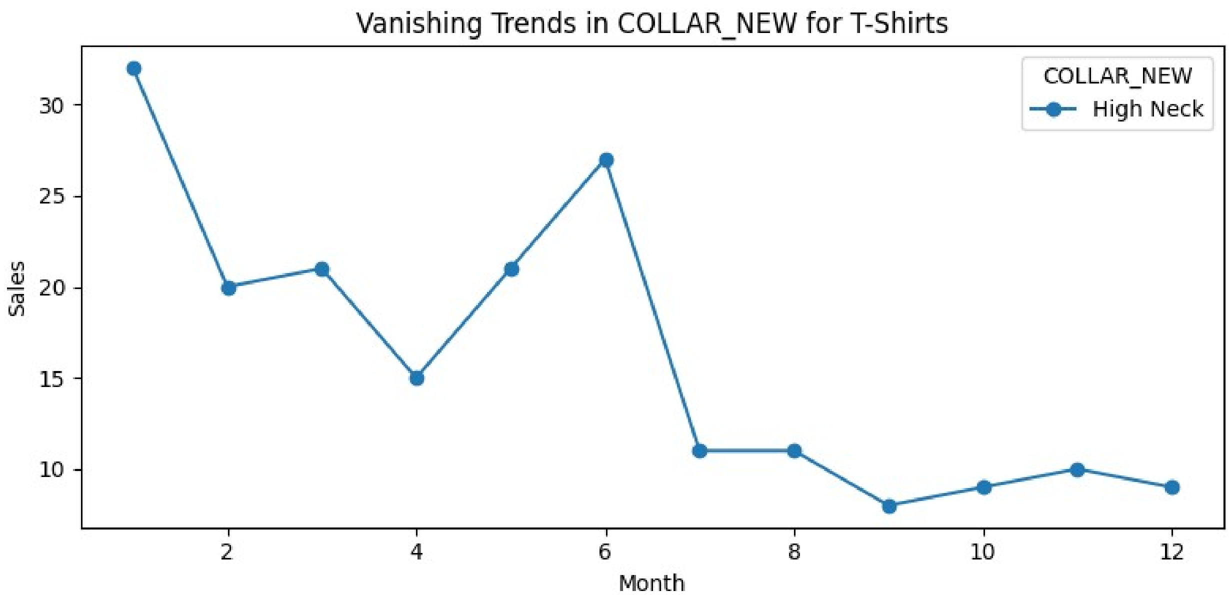
```
# Melt to long format for time series
df_long = df.melt(
    id_vars=['CATEGORY'] + feature_columns,
    value_vars=month_cols,
    var_name='MONTH',
    value_name='SALES'
)
df_long['MONTH_NUM'] = df_long['MONTH'].str.extract('(\
d+)').astype(int)

# For each category and feature, finding and plotting top 3 vanishing
feature values (most negative sales slope)
for category in df_long['CATEGORY'].unique():
    cat_df = df_long[df_long['CATEGORY'] == category]
    for feature in feature_columns:
        pivot = cat_df.groupby([feature, 'MONTH_NUM'])
['SALES'].sum().reset_index()
```

```
        slopes = {}
        for val in pivot[feature].unique():
            sales_series = pivot[pivot[feature] ==
val].sort_values('MONTH_NUM')
            if len(sales_series) > 1:
                x = sales_series['MONTH_NUM'].values
                y = sales_series['SALES'].values
                slope = np.polyfit(x, y, 1)[0]
                slopes[val] = slope
        # Get top 3 most negative slopes (vanishing trends)
        vanishing = sorted(slopes.items(), key=lambda x: x[1])[:3]
        vanishing = [v for v in vanishing if v[1] < 0]
        if vanishing:
            # Visualization only — removed all print statements
            plt.figure(figsize=(8,4))
            for val, _ in vanishing:
                sales = pivot[pivot[feature] ==
val].sort_values('MONTH_NUM')
                plt.plot(sales['MONTH_NUM'], sales['SALES'],
marker='o', label=val)
            plt.title(f"Vanishing Trends in {feature} for {category}")
            plt.xlabel('Month')
            plt.ylabel('Sales')
            plt.legend(title=feature)
            plt.tight_layout()
            plt.show()
```

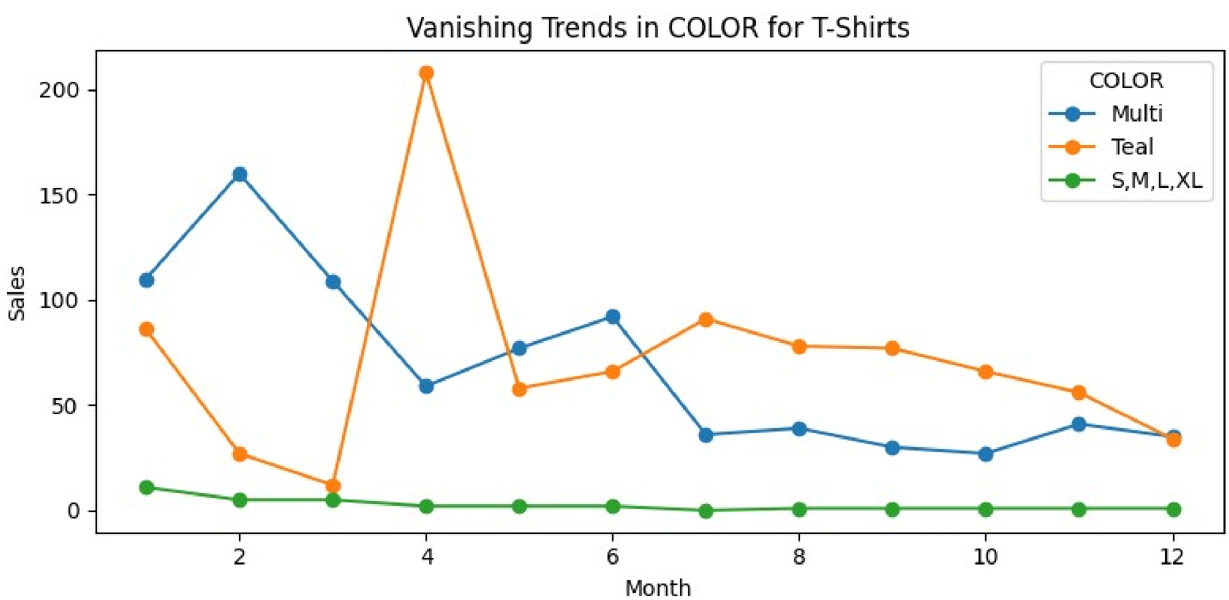



Vanishing Trends in PRINT_DESIGN for T-Shirts
PRINT_DESIGN
Printed, Abstract
Printed, Graphic Print
Printed
Sales
Month
0
200
400
600
800
1000
1200
1400
1600
2
4
6
8
10
12

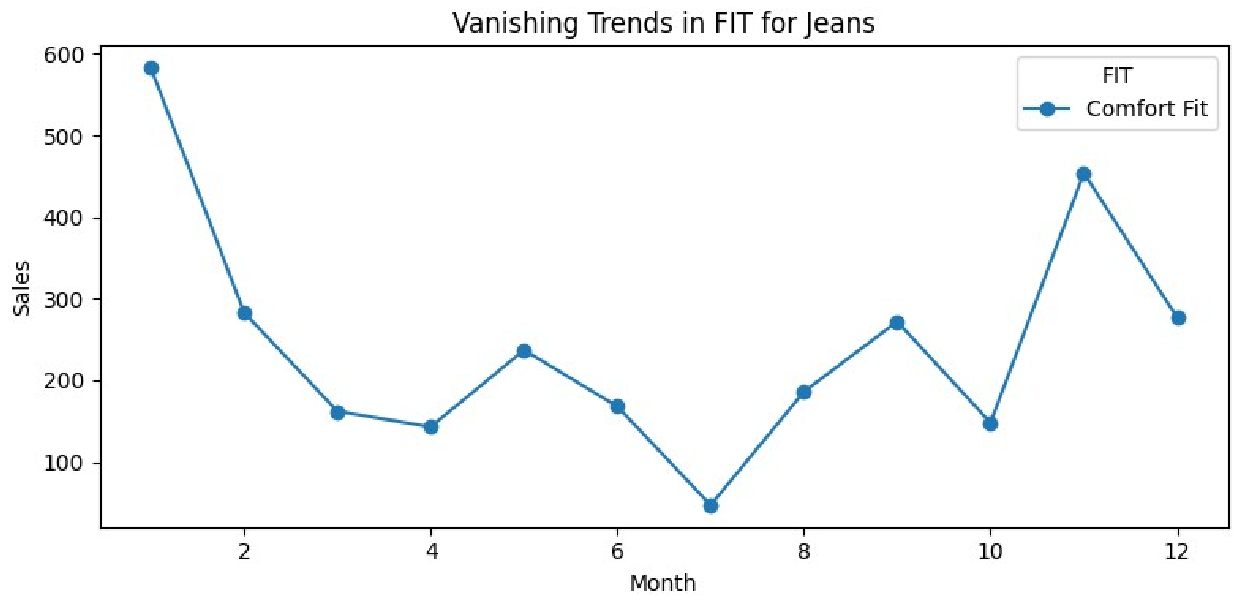

Vanishing Trends in SLEEVE_TYPE for T-Shirts
SLEEVE_TYPE
Sleeveless
Sales
Month
25
50
75
100
125
150
175
200
2
4
6
8
10
12



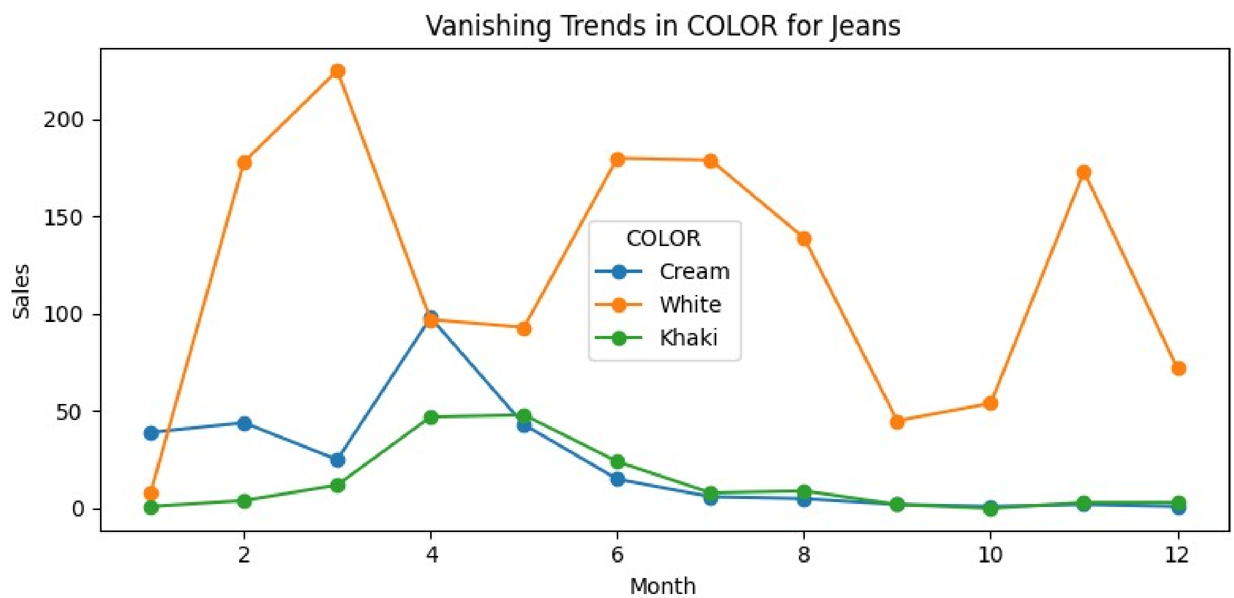
Vanishing Trends in COLLAR_NEW for T-Shirts
COLLAR_NEW
High Neck
Sales
Month
30
25
20
15
10
2
4
6
8
10
12



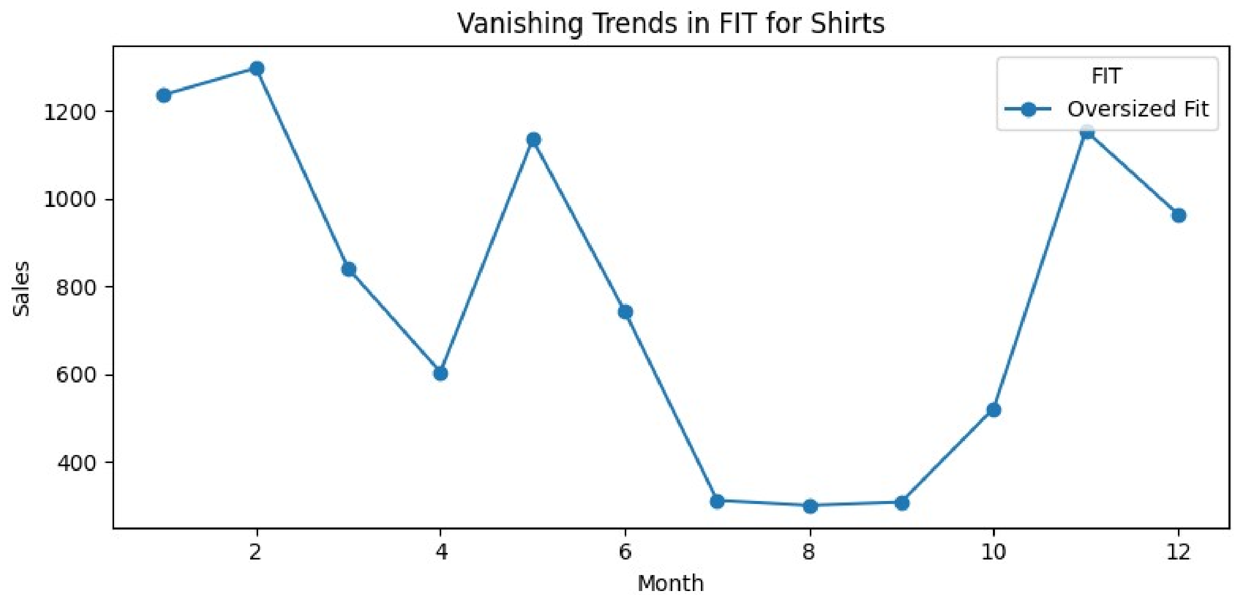
Vanishing Trends in COLOR for T-Shirts
COLOR
Multi
Teal
S,M,L,XL
Sales
Month
200
150
100
50
0
2
4
6
8
10
12



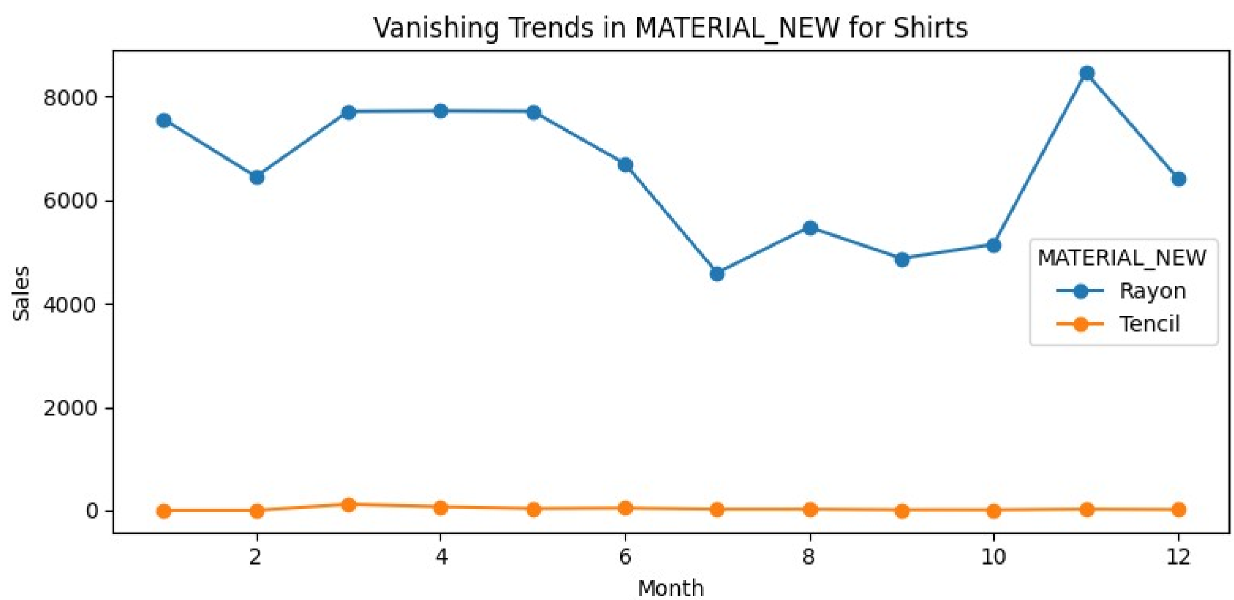
Vanishing Trends in FIT for Jeans
Sales
Month
FIT
Comfort Fit
600
500
400
300
200
100
2
4
6
8
10
12



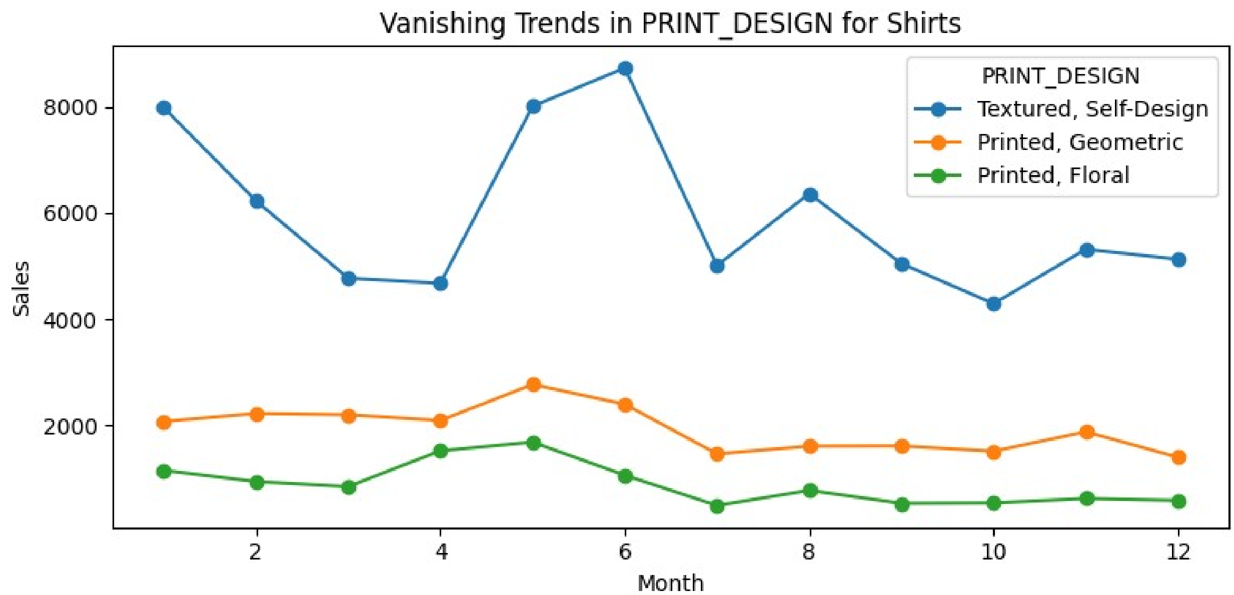
Vanishing Trends in COLOR for Jeans
Sales
Month
COLOR
Cream
White
Khaki
200
150
100
50
0
2
4
6
8
10
12



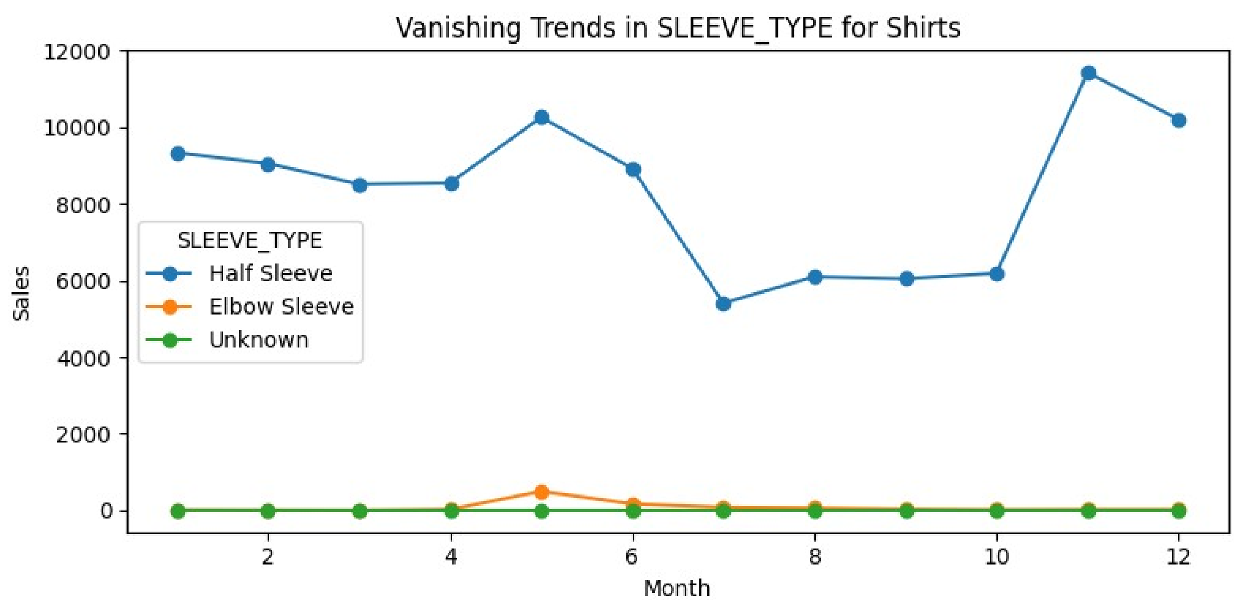
Vanishing Trends in FIT for Shirts
FIT
Oversized Fit
Sales
Month
1200
1000
800
600
400
2
4
6
8
10
12



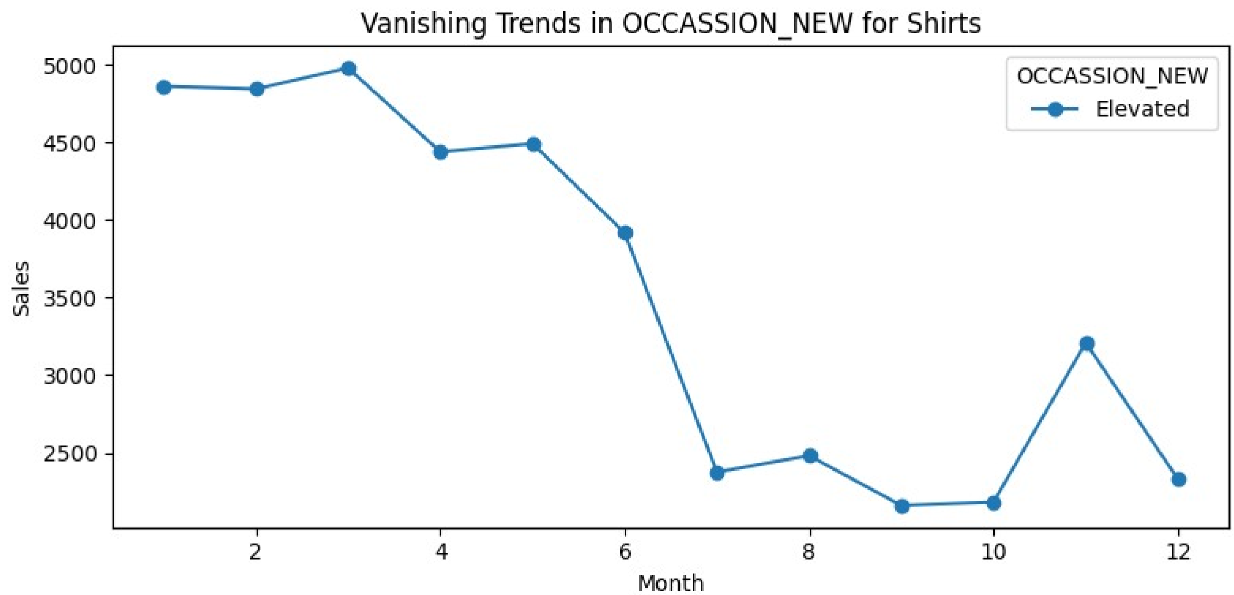
Vanishing Trends in MATERIAL_NEW for Shirts
MATERIAL_NEW
Rayon
Tencil
Sales
Month
8000
6000
4000
2000
0
2
4
6
8
10
12

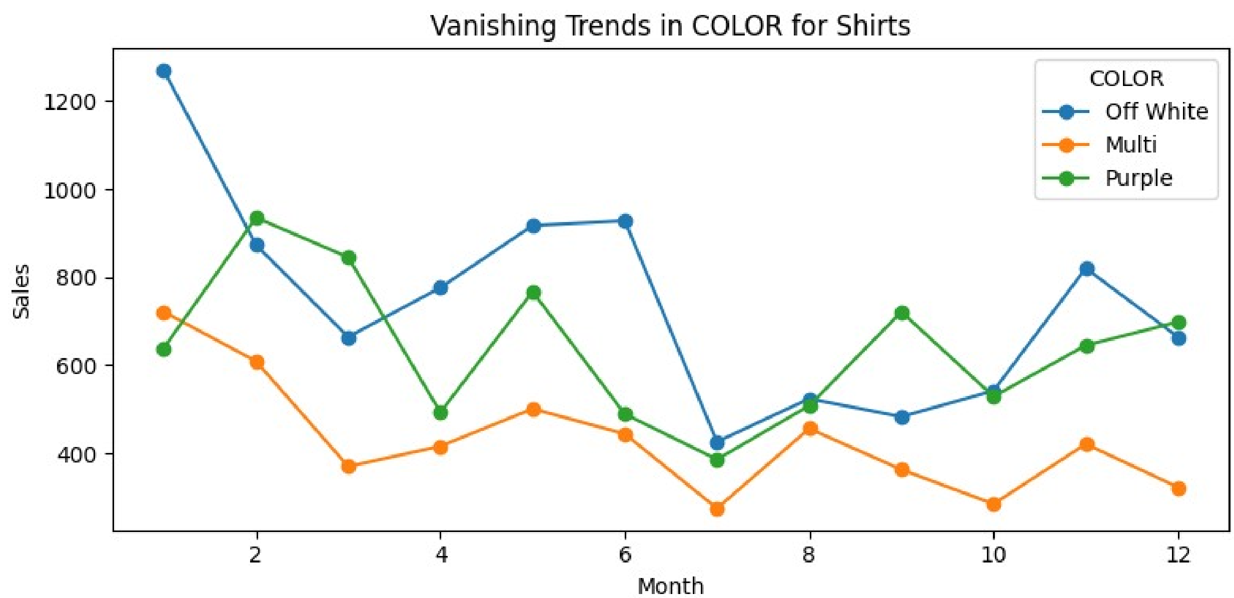


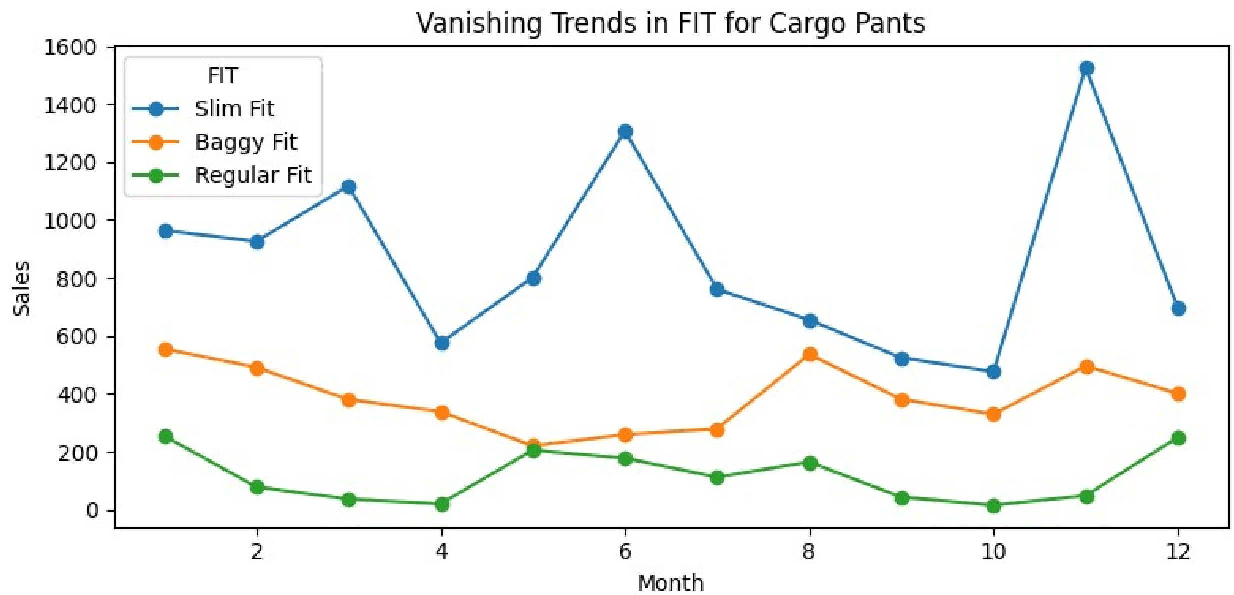
Vanishing Trends in PRINT_DESIGN for Shirts
PRINT_DESIGN
Textured, Self-Design
Printed, Geometric
Printed, Floral
Sales
Month
8000
6000
4000
2000
2
4
6
8
10
12

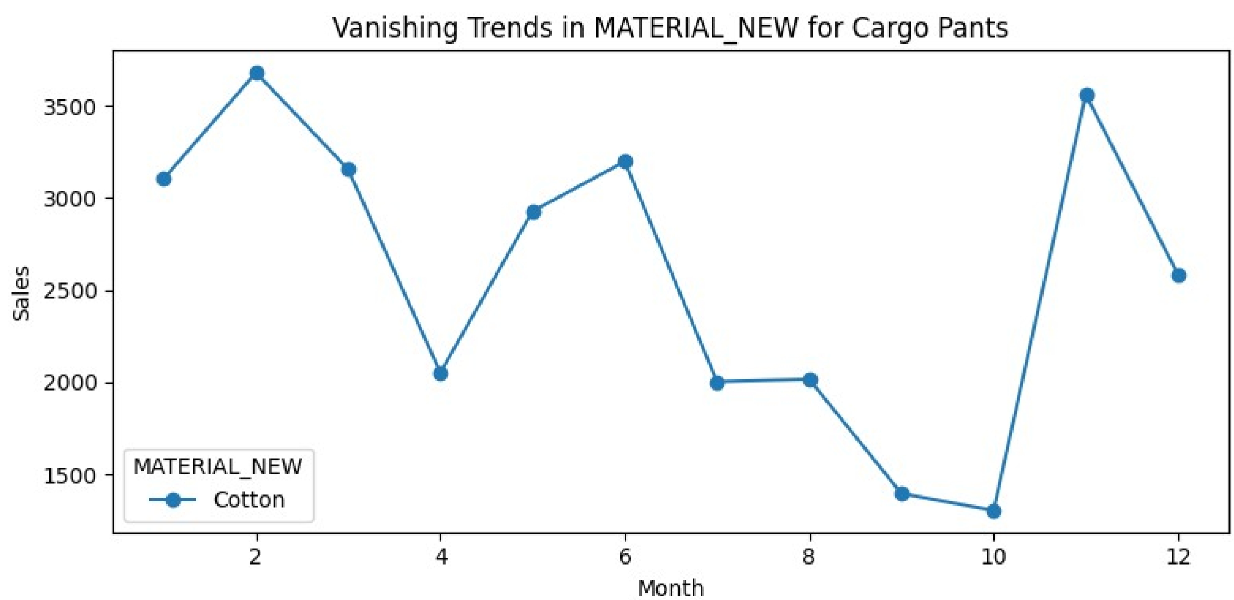


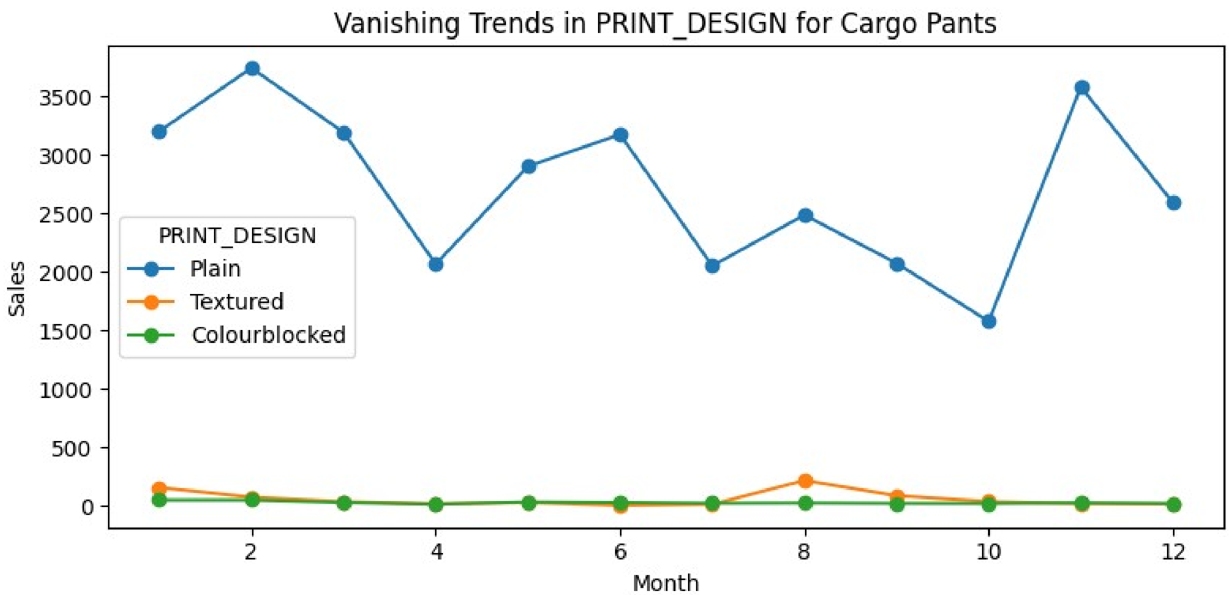
Vanishing Trends in SLEEVE_TYPE for Shirts
SLEEVE_TYPE
Half Sleeve
Elbow Sleeve
Unknown
Sales
Month
12000
10000
8000
6000
4000
2000
0
2
4
6
8
10
12

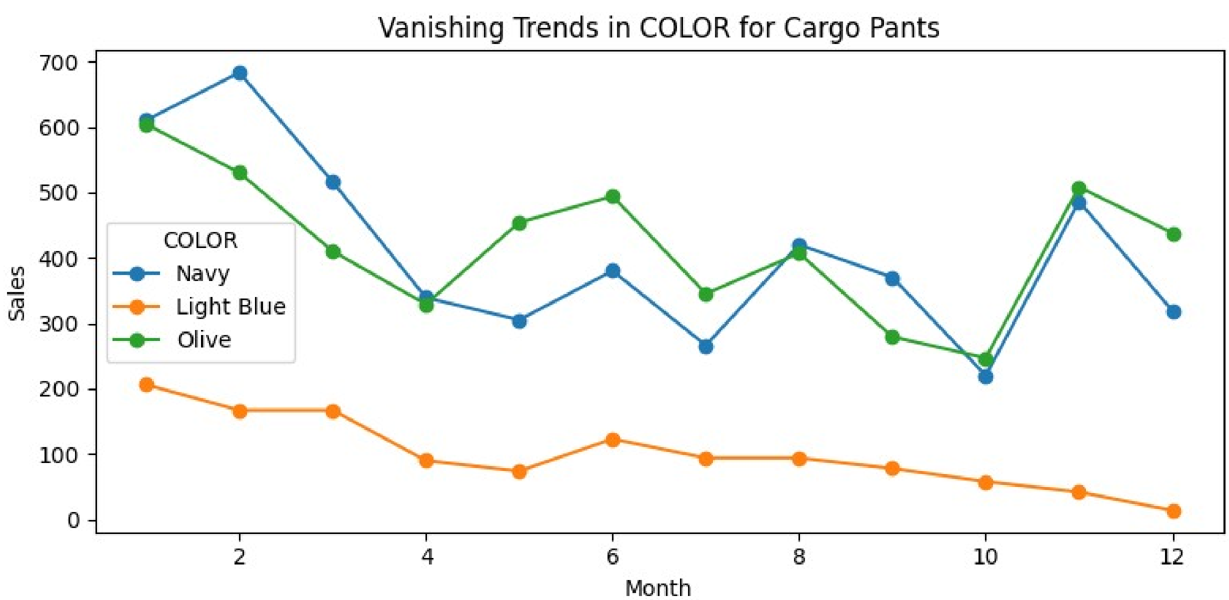

Vanishing Trends in OCCASSION_NEW for Shirts
OCCASSION_NEW
Elevated
Sales
Month
5000
4500
4000
3500
3000
2500
2
4
6
8
10
12

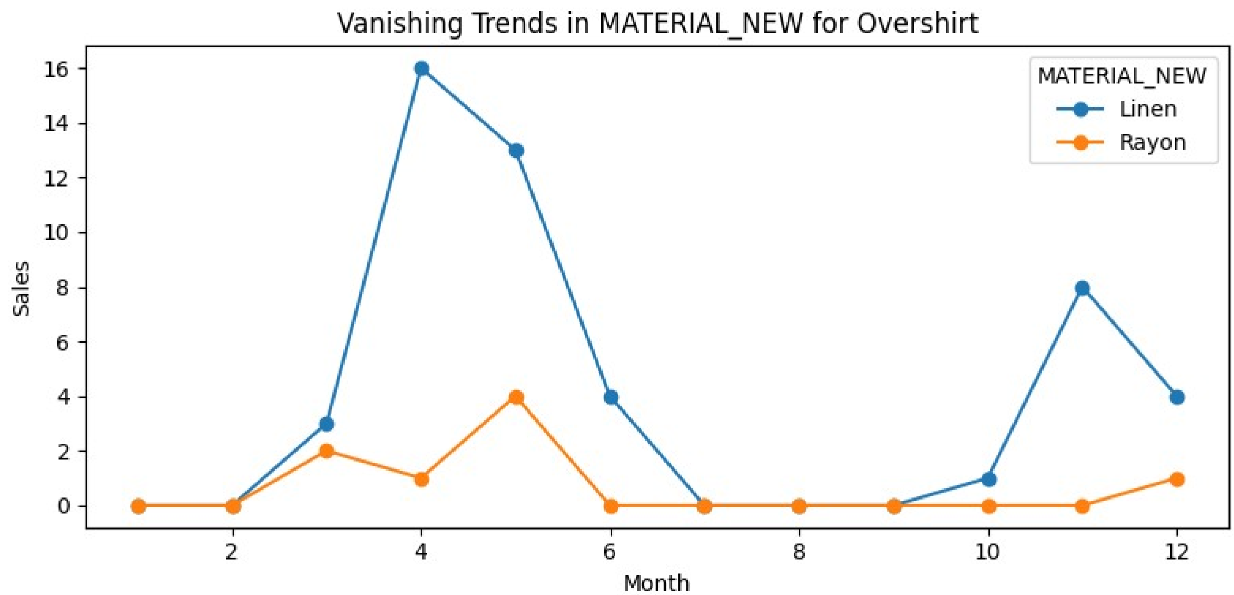

Vanishing Trends in COLOR for Shirts
COLOR
Off White
Multi
Purple
Sales
Month
1200
1000
800
600
400
2
4
6
8
10
12

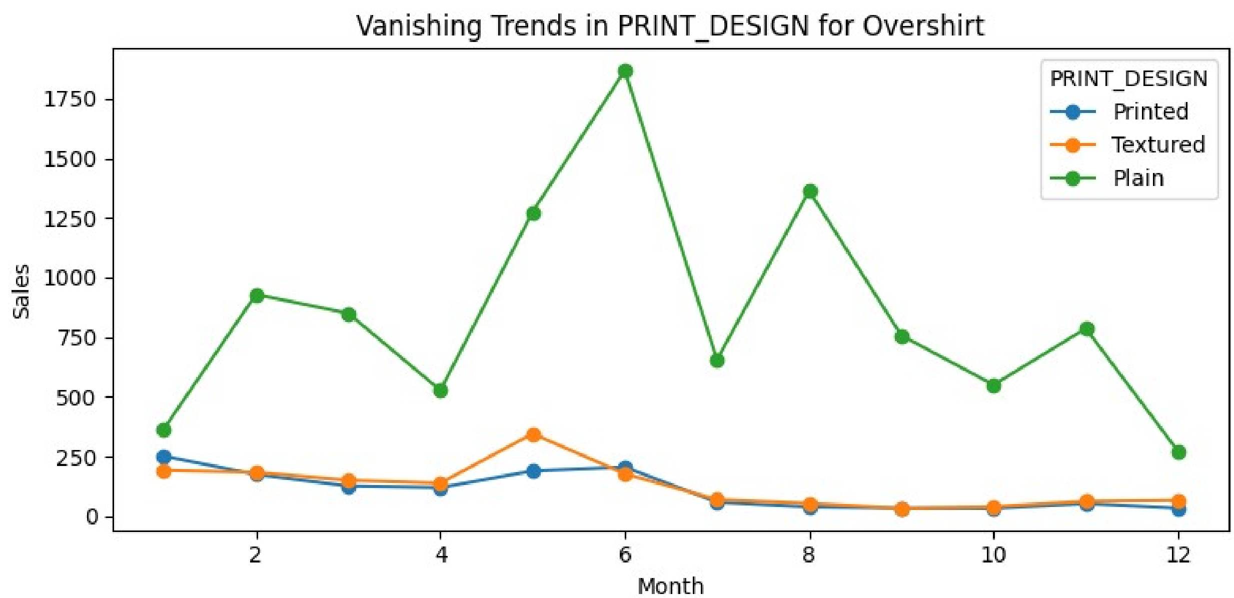

Vanishing Trends in FIT for Cargo Pants
FIT
Slim Fit
Baggy Fit
Regular Fit
Sales
Month
0
200
400
600
800
1000
1200
1400
1600
2
4
6
8
10
12

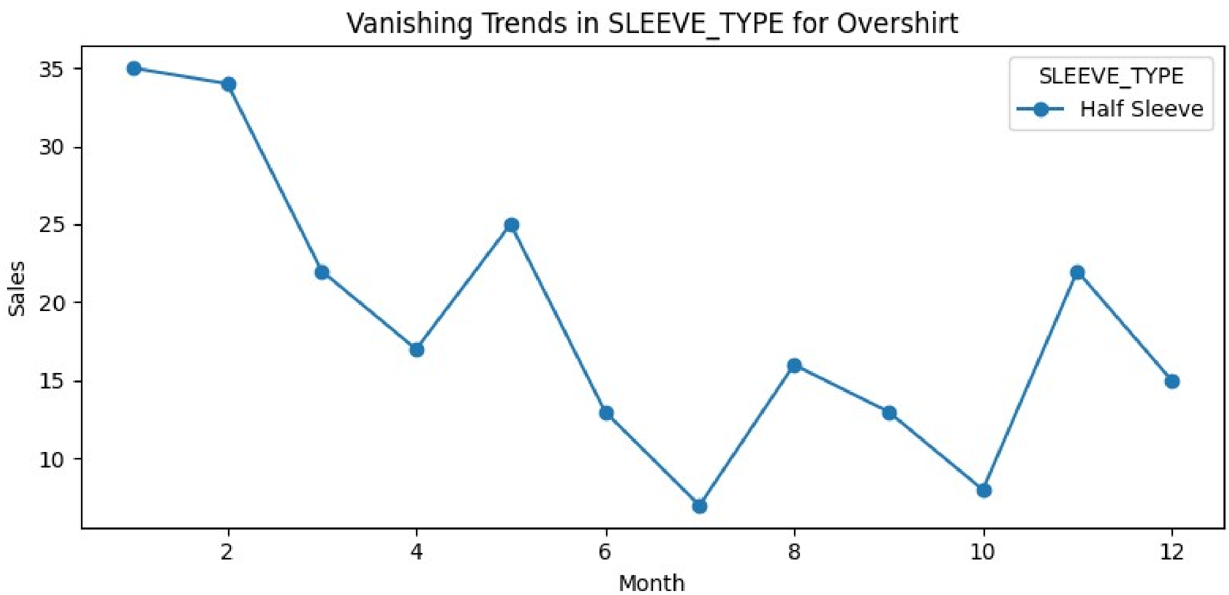

Vanishing Trends in MATERIAL_NEW for Cargo Pants
MATERIAL_NEW
Cotton
Sales
Month
1500
2000
2500
3000
3500
2
4
6
8
10
12

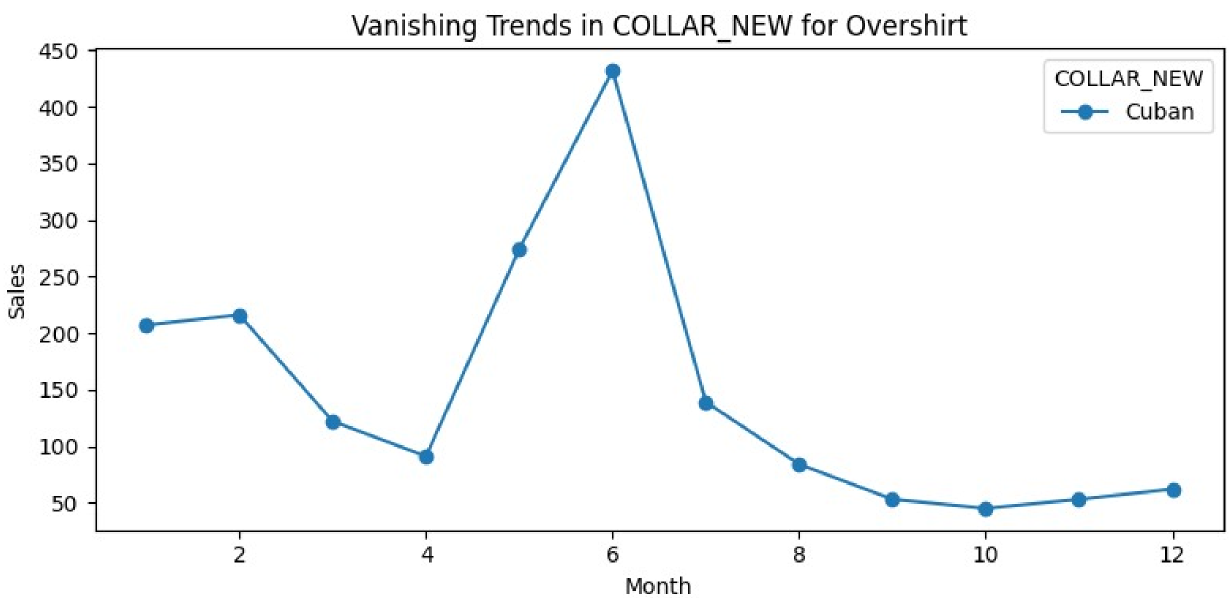

Vanishing Trends in PRINT_DESIGN for Cargo Pants
Sales
Month
PRINT_DESIGN
Plain
Textured
Colourblocked
0
500
1000
1500
2000
2500
3000
3500
2
4
6
8
10
12

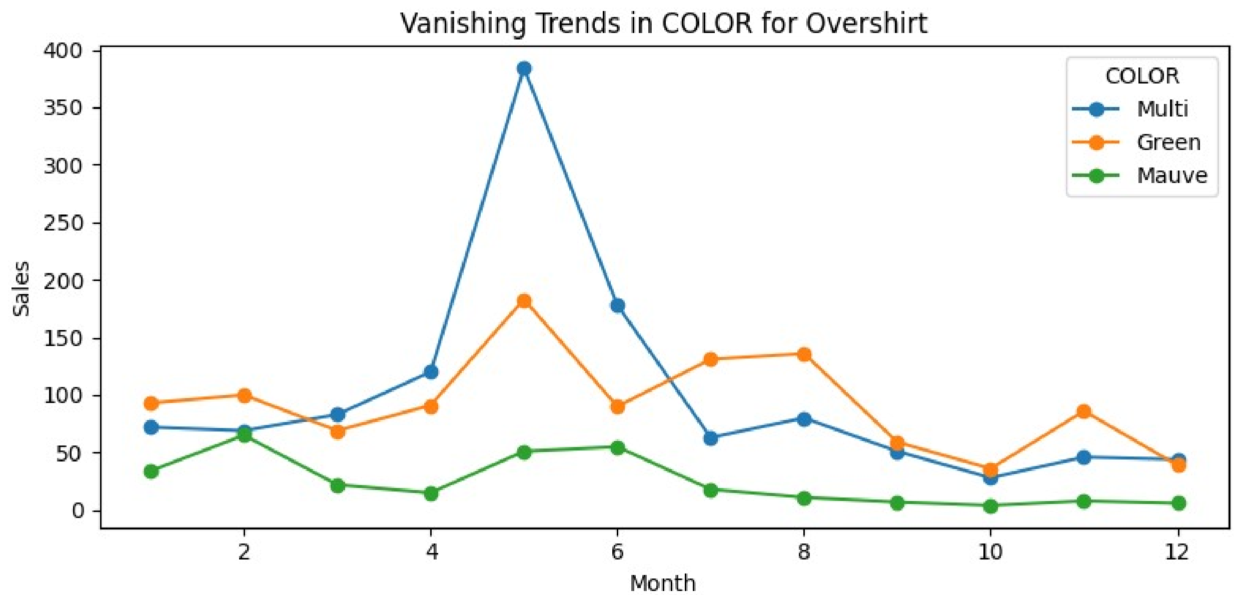

Vanishing Trends in COLOR for Cargo Pants
Sales
Month
COLOR
Navy
Light Blue
Olive
0
100
200
300
400
500
600
700
2
4
6
8
10
12


Vanishing Trends in MATERIAL_NEW for Overshirt
MATERIAL_NEW
Linen
Rayon
Sales
Month
0
2
4
6
8
10
12
14
16

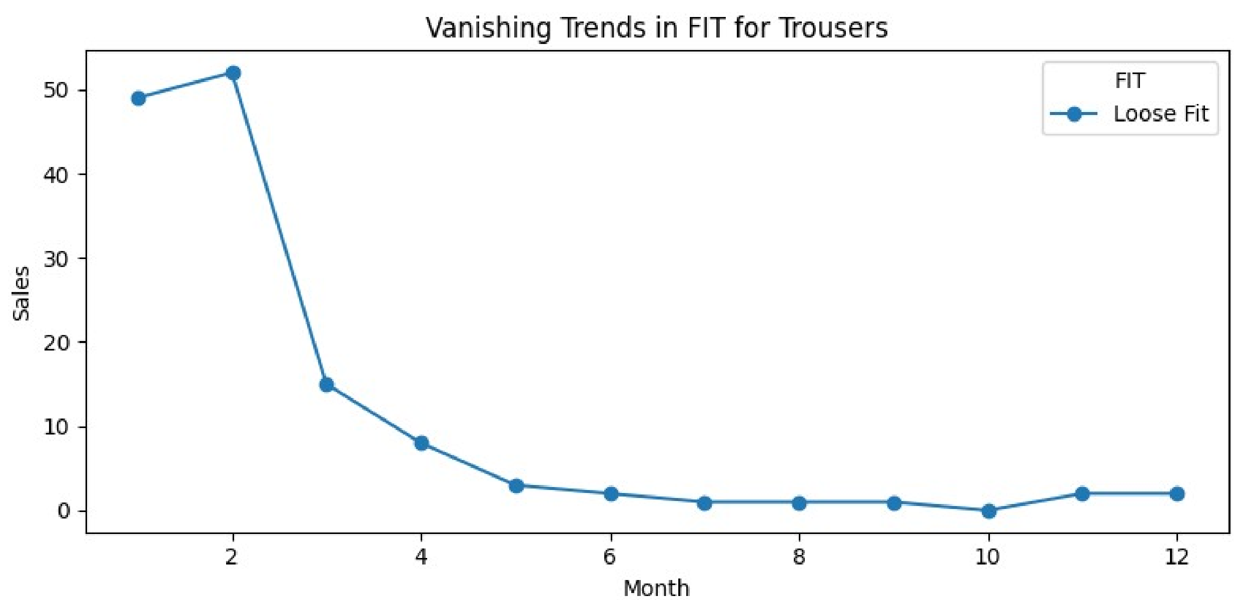

Vanishing Trends in PRINT_DESIGN for Overshirt
PRINT_DESIGN
Printed
Textured
Plain
Sales
Month
0
250
500
750
1000
1250
1500
1750
2
4
6
8
10
12


Vanishing Trends in SLEEVE_TYPE for Overshirt
SLEEVE_TYPE
Half Sleeve
Sales
35
30
25
20
15
10
Month
2
4
6
8
10
12

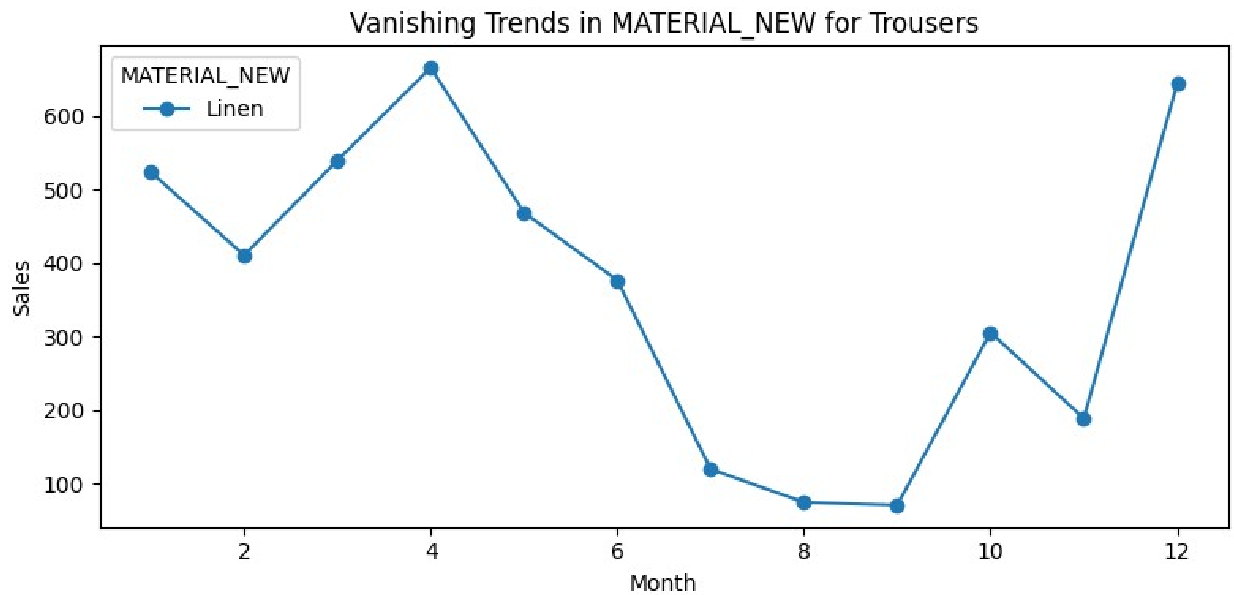

Vanishing Trends in COLLAR_NEW for Overshirt
COLLAR_NEW
Cuban
Sales
450
400
350
300
250
200
150
100
50
Month
2
4
6
8
10
12

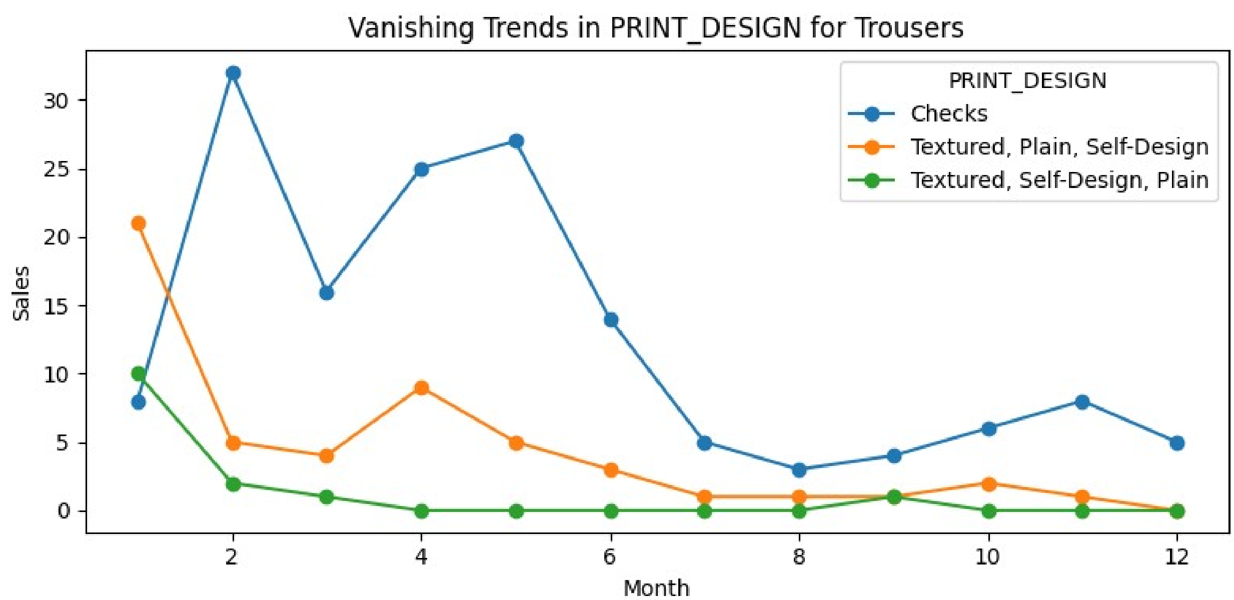


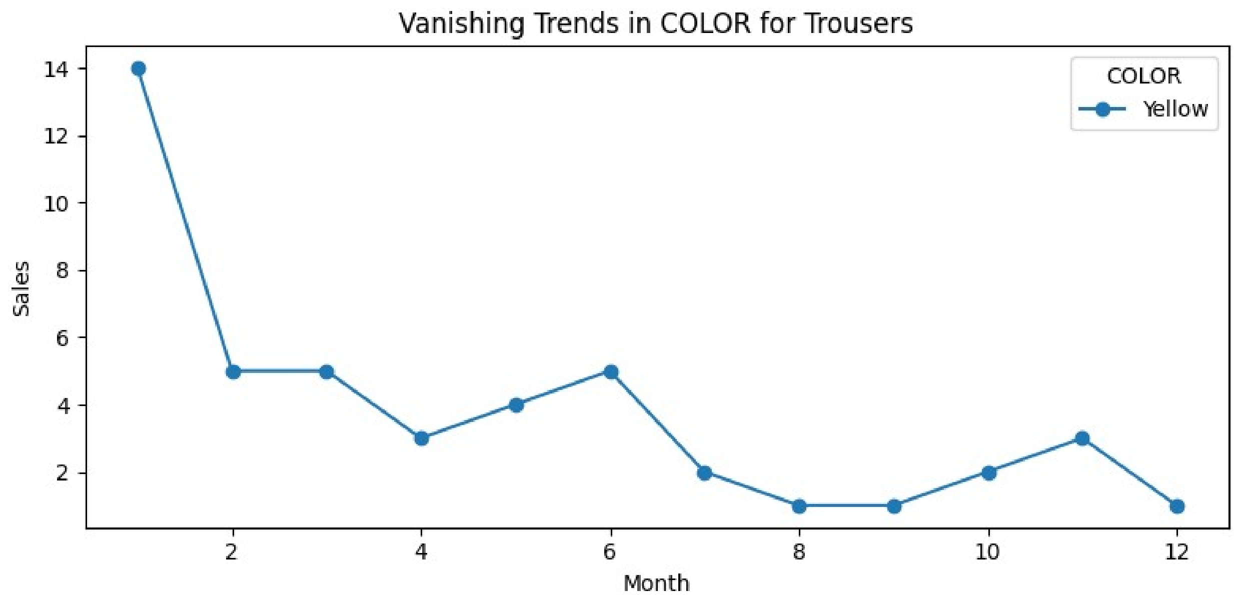
Vanishing Trends in COLOR for Overshirt
COLOR
Multi
Green
Mauve
Sales
Month
0
50
100
150
200
250
300
350
400
2
4
6
8
10
12

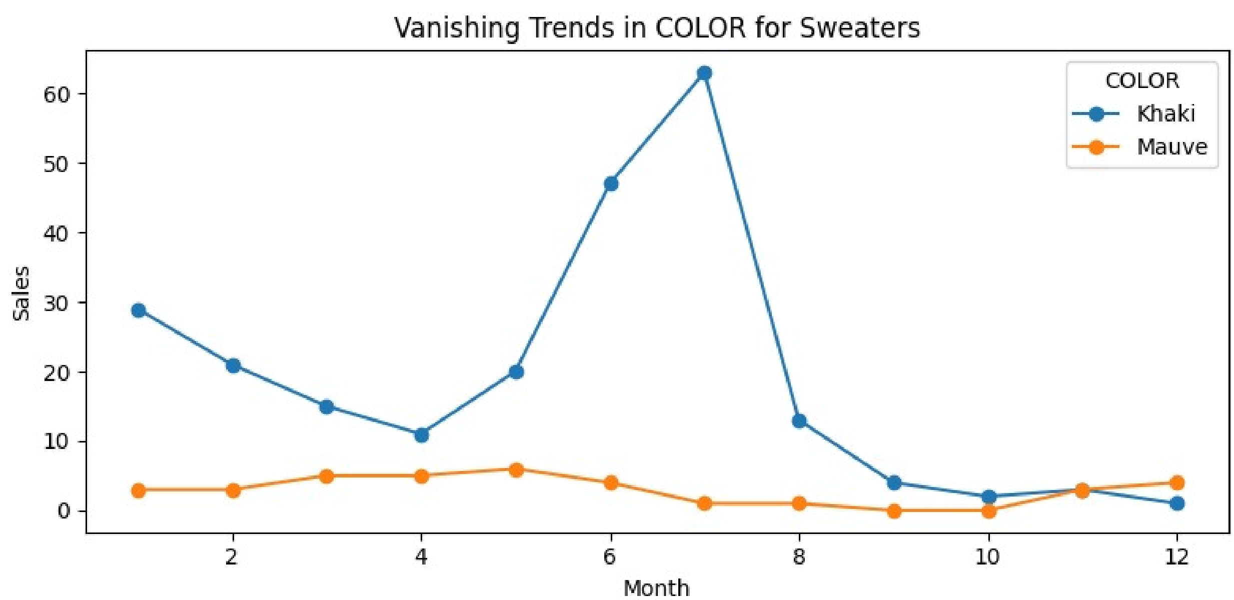


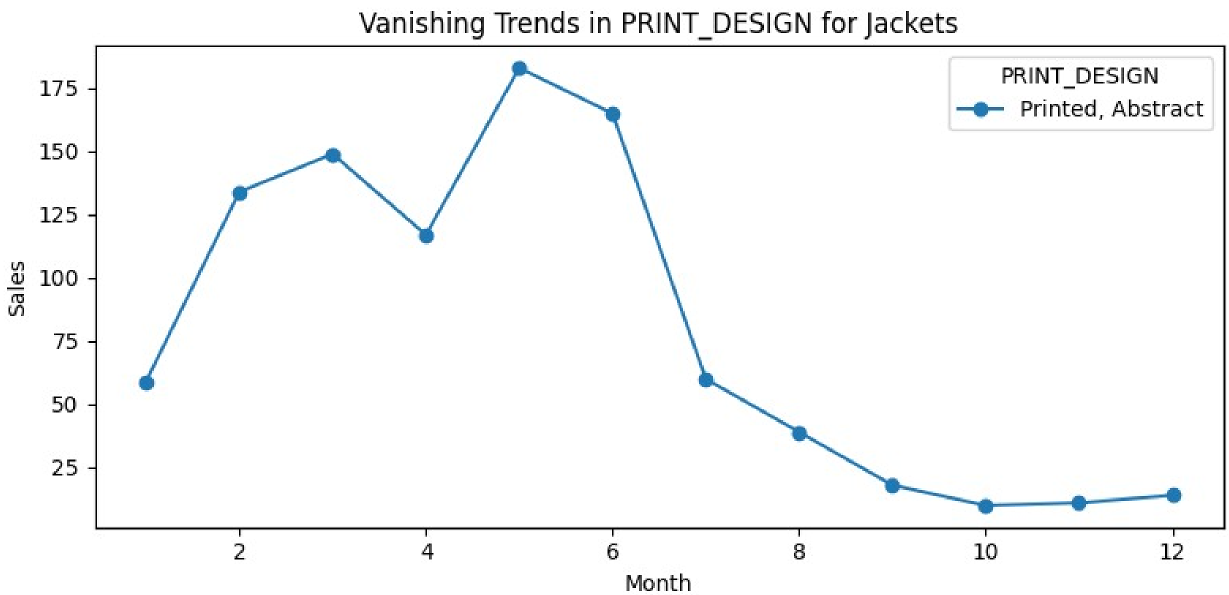
Vanishing Trends in FIT for Trousers
FIT
Loose Fit
Sales
Month
0
10
20
30
40
50
2
4
6
8
10
12

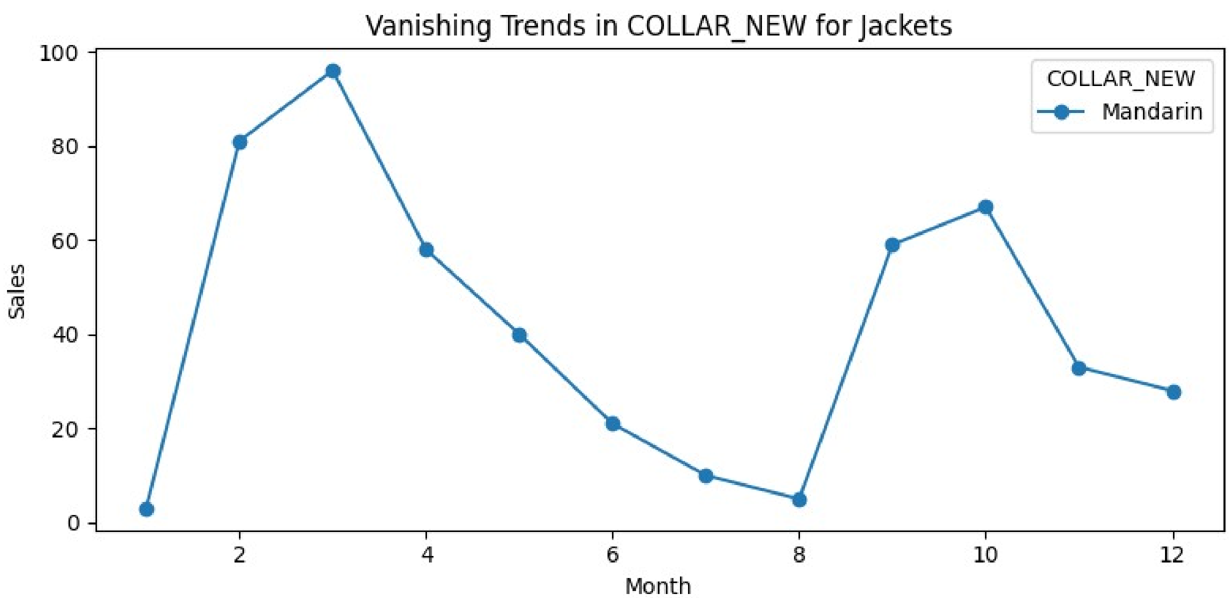

Vanishing Trends in MATERIAL_NEW for Trousers
MATERIAL_NEW
Linen
Sales
600
500
400
300
200
100
2
4
6
8
10
12
Month

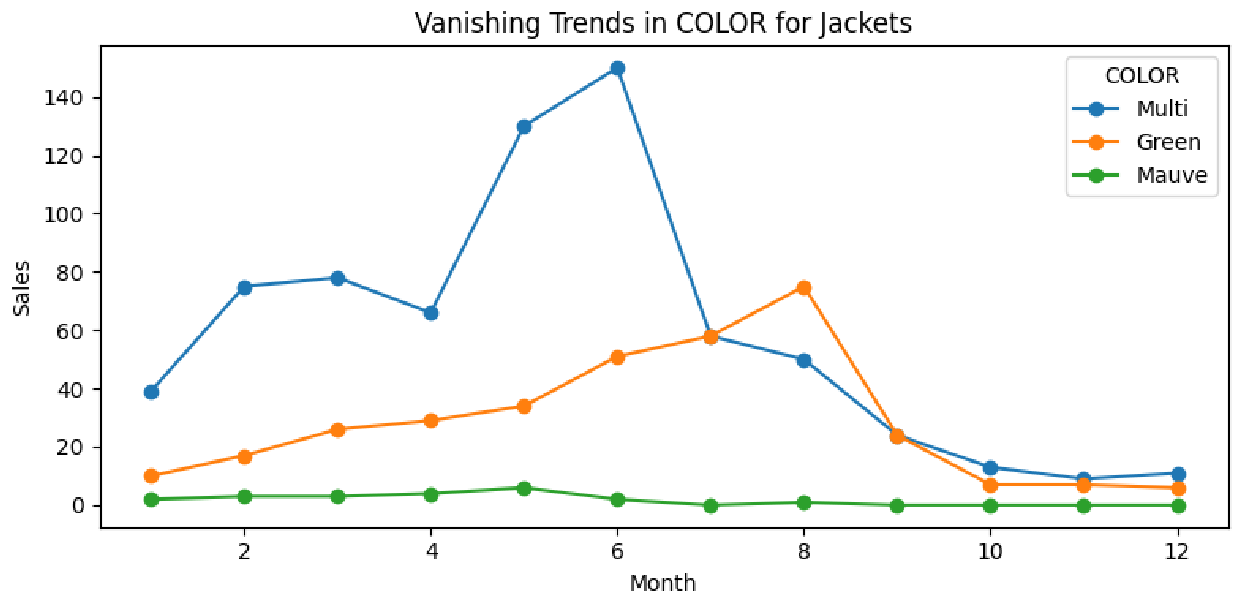

Vanishing Trends in PRINT_DESIGN for Trousers
PRINT_DESIGN
Checks
Textured, Plain, Self-Design
Textured, Self-Design, Plain
Sales
30
25
20
15
10
5
0
2
4
6
8
10
12
Month

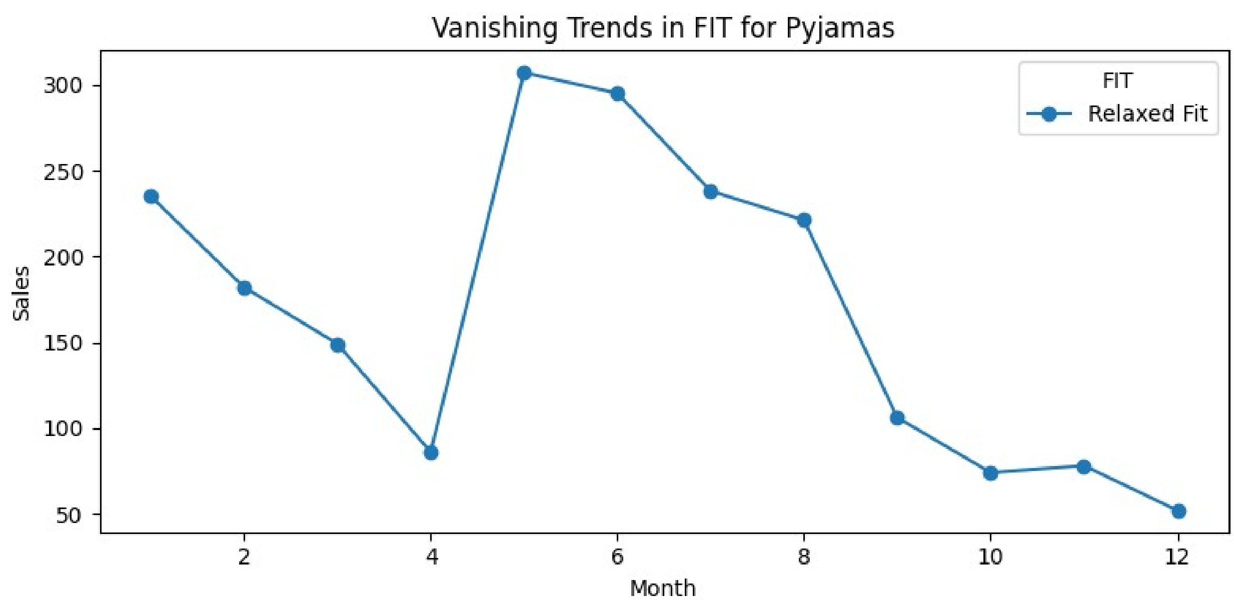

Vanishing Trends in COLOR for Trousers
COLOR
Yellow
Sales
Month
14
12
10
8
6
4
2
2
4
6
8
10
12

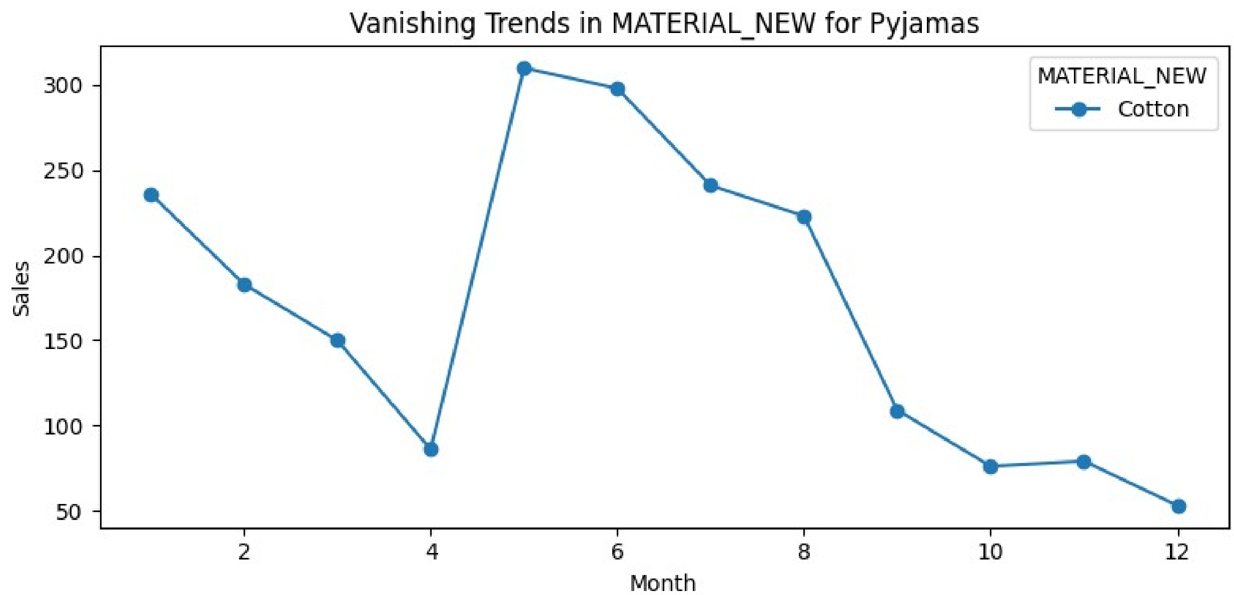

Vanishing Trends in COLOR for Sweaters
COLOR
Khaki
Mauve
Sales
Month
60
50
40
30
20
10
0
2
4
6
8
10
12

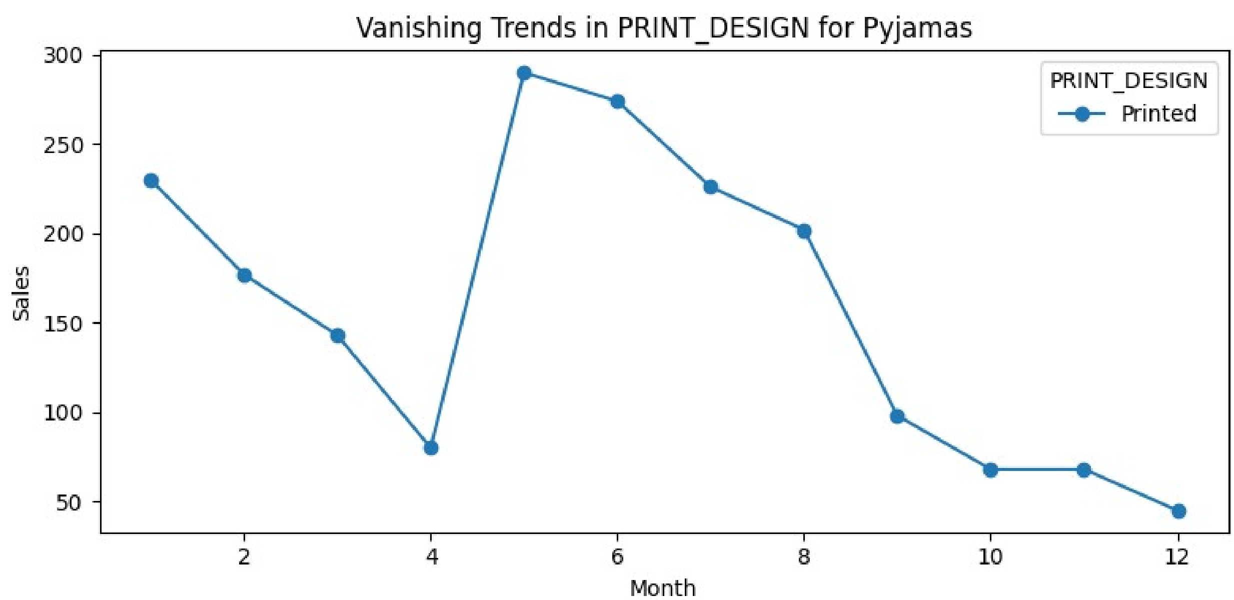

Vanishing Trends in PRINT_DESIGN for Jackets
PRINT_DESIGN
Printed, Abstract
Sales
Month
175
150
125
100
75
50
25
2
4
6
8
10
12

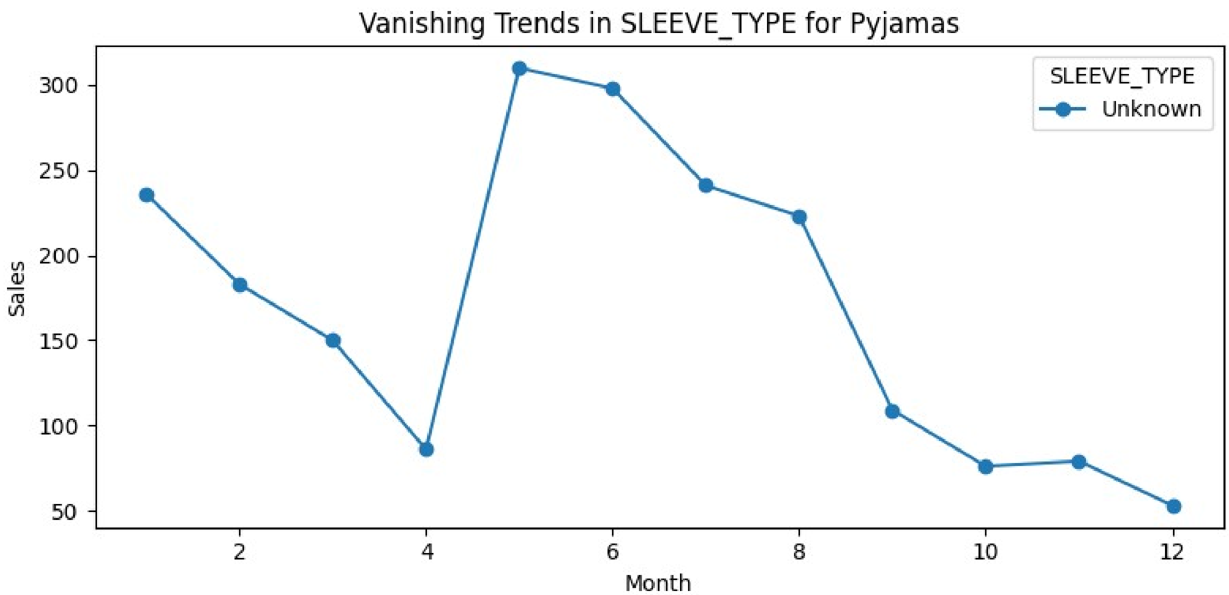

Vanishing Trends in COLLAR_NEW for Jackets
COLLAR_NEW
Mandarin
Sales
Month
100
80
60
40
20
0
2
4
6
8
10
12

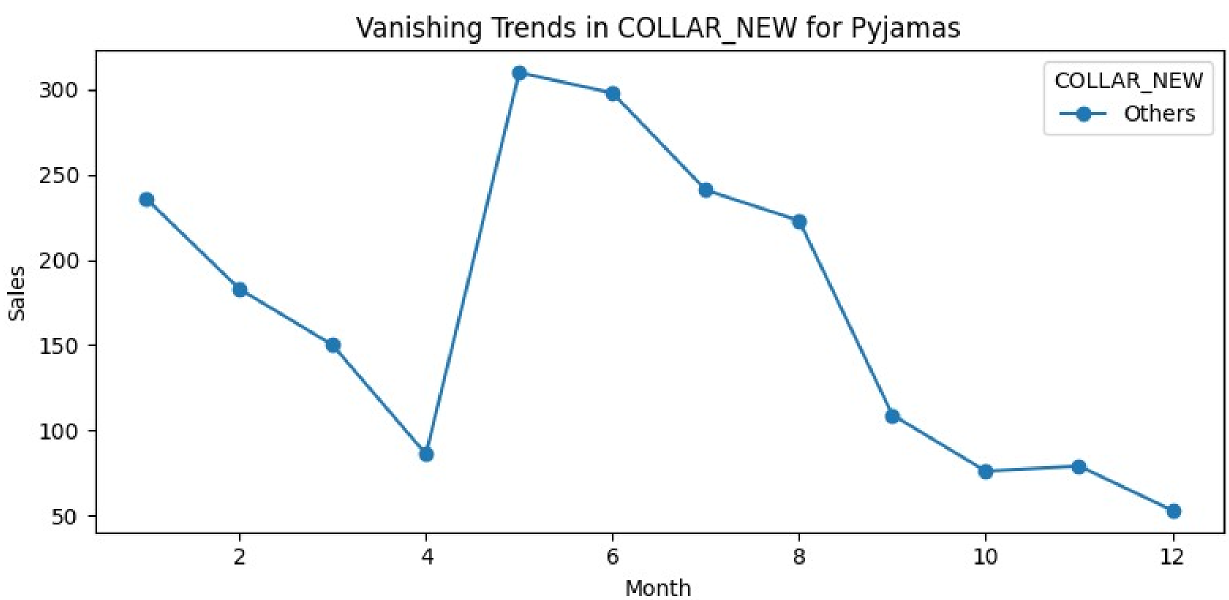

Vanishing Trends in COLOR for Jackets
COLOR
Multi
Green
Mauve
Sales
Month
0
20
40
60
80
100
120
140
2
4
6
8
10
12

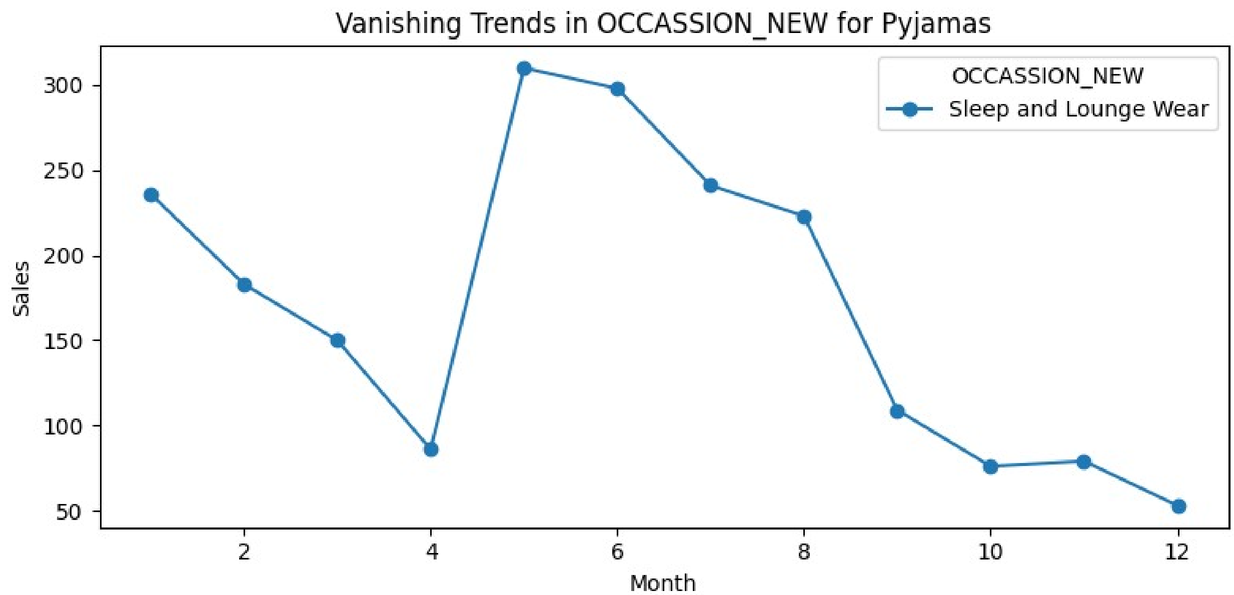

Vanishing Trends in FIT for Pyjamas
FIT
Relaxed Fit
Sales
Month
50
100
150
200
250
300
2
4
6
8
10
12

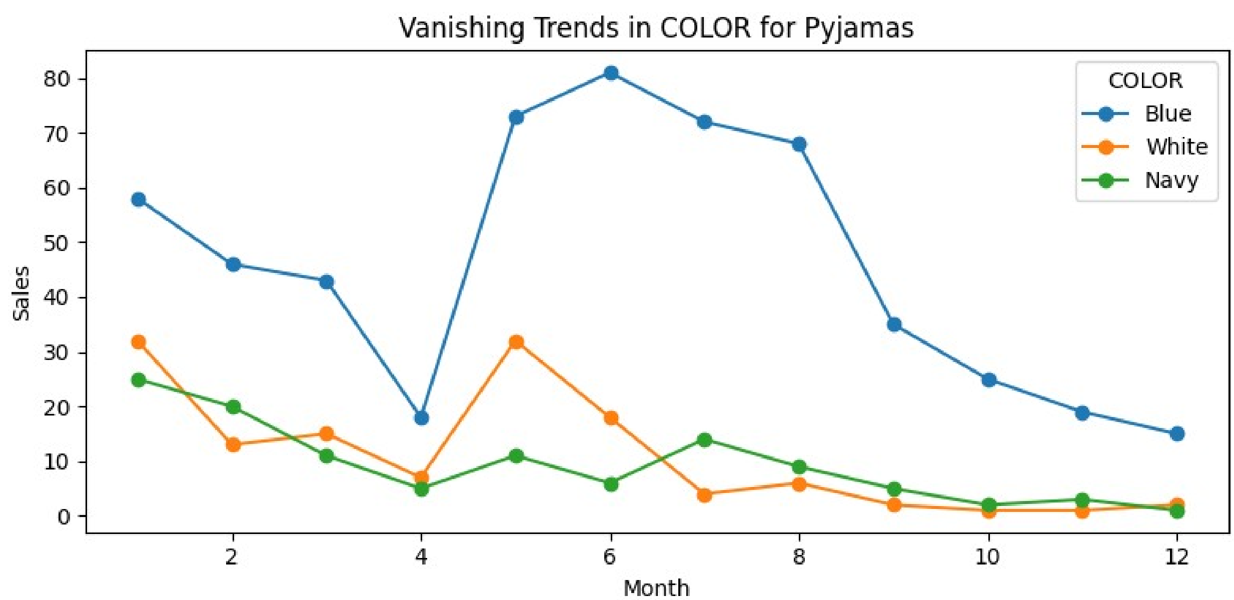

Vanishing Trends in MATERIAL_NEW for Pyjamas
MATERIAL_NEW
Cotton
Sales
Month
300
250
200
150
100
50
2
4
6
8
10
12

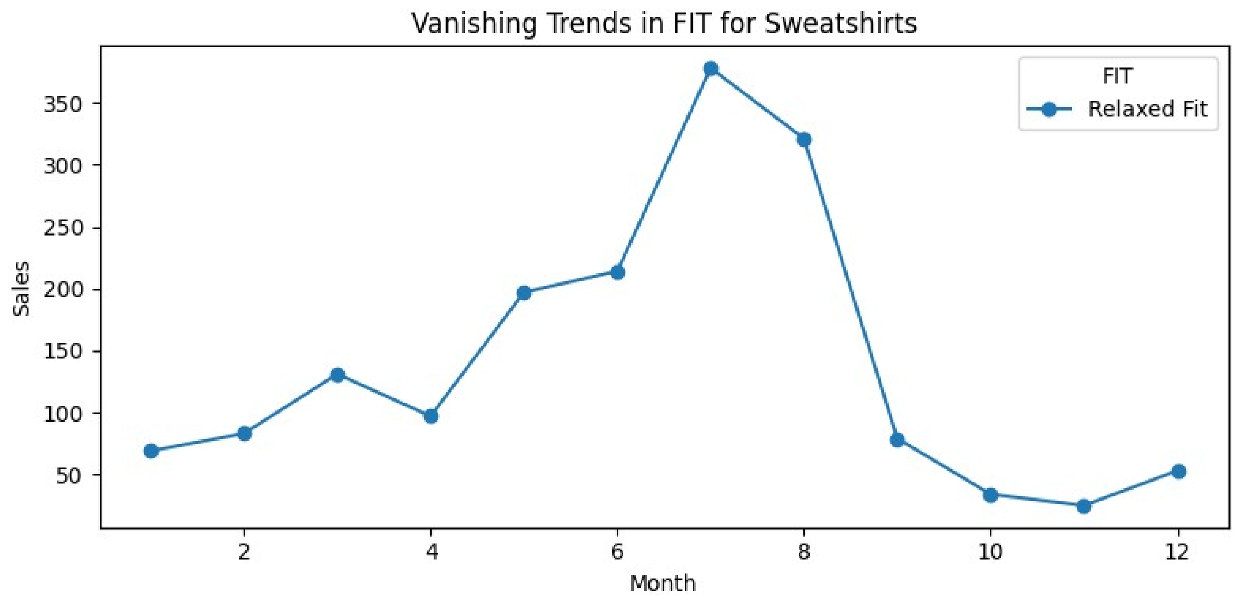

Vanishing Trends in PRINT_DESIGN for Pyjamas
PRINT_DESIGN
Printed
Sales
Month
300
250
200
150
100
50
2
4
6
8
10
12

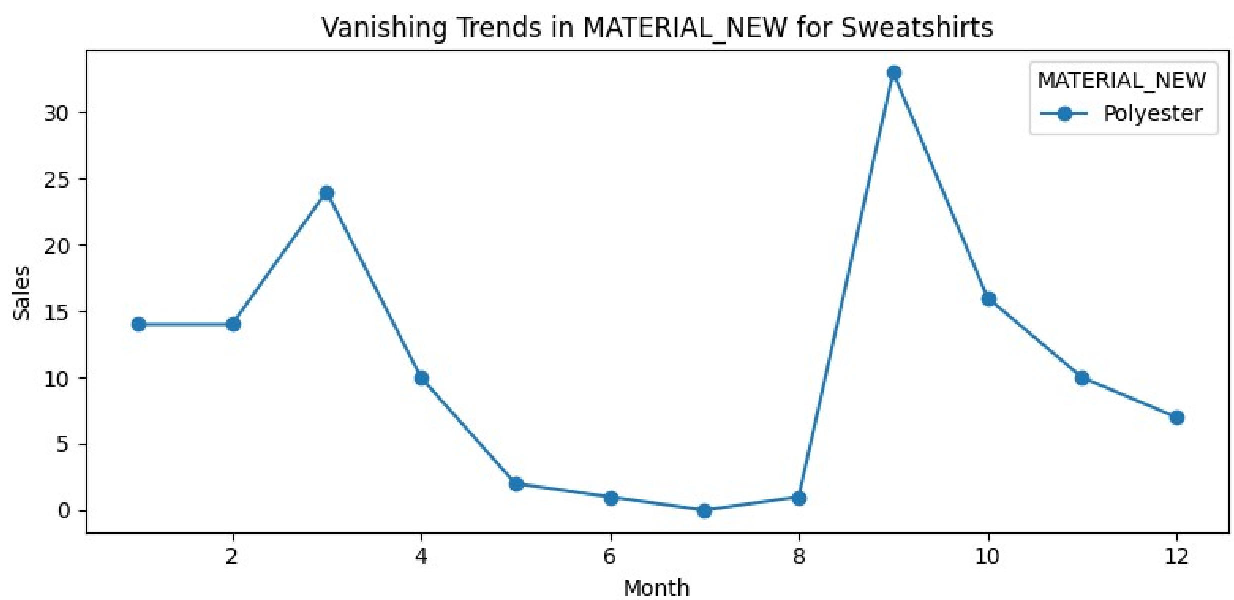

Vanishing Trends in SLEEVE_TYPE for Pyjamas
SLEEVE_TYPE
Unknown
Sales
Month
300
250
200
150
100
50
2
4
6
8
10
12

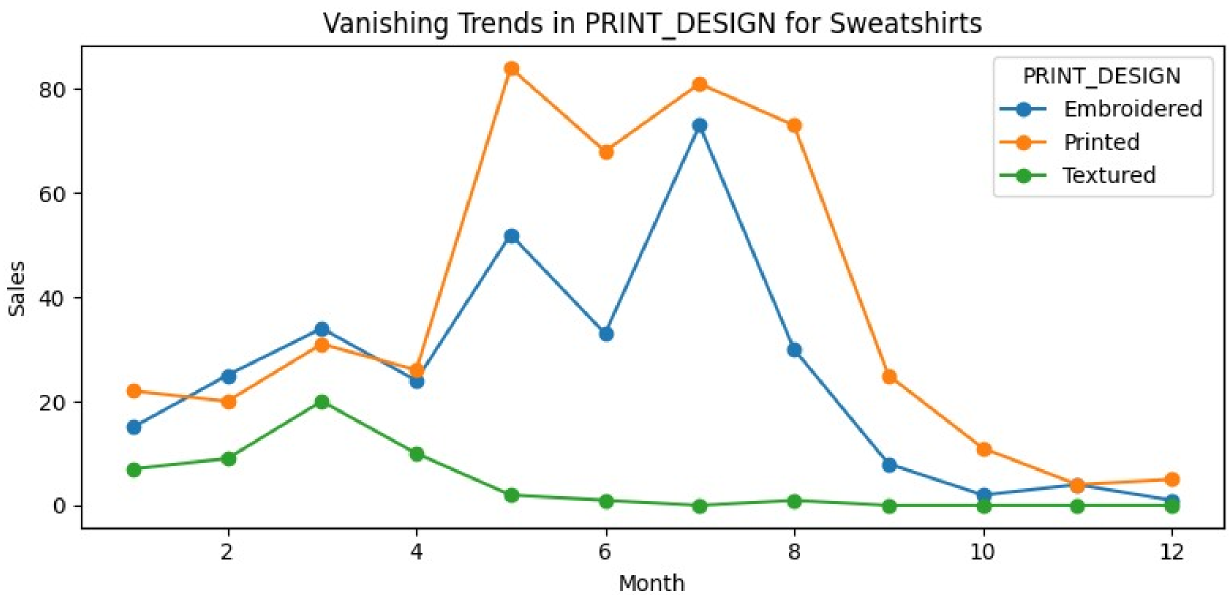

Vanishing Trends in COLLAR_NEW for Pyjamas
COLLAR_NEW
Others
Sales
Month
300
250
200
150
100
50
2
4
6
8
10
12

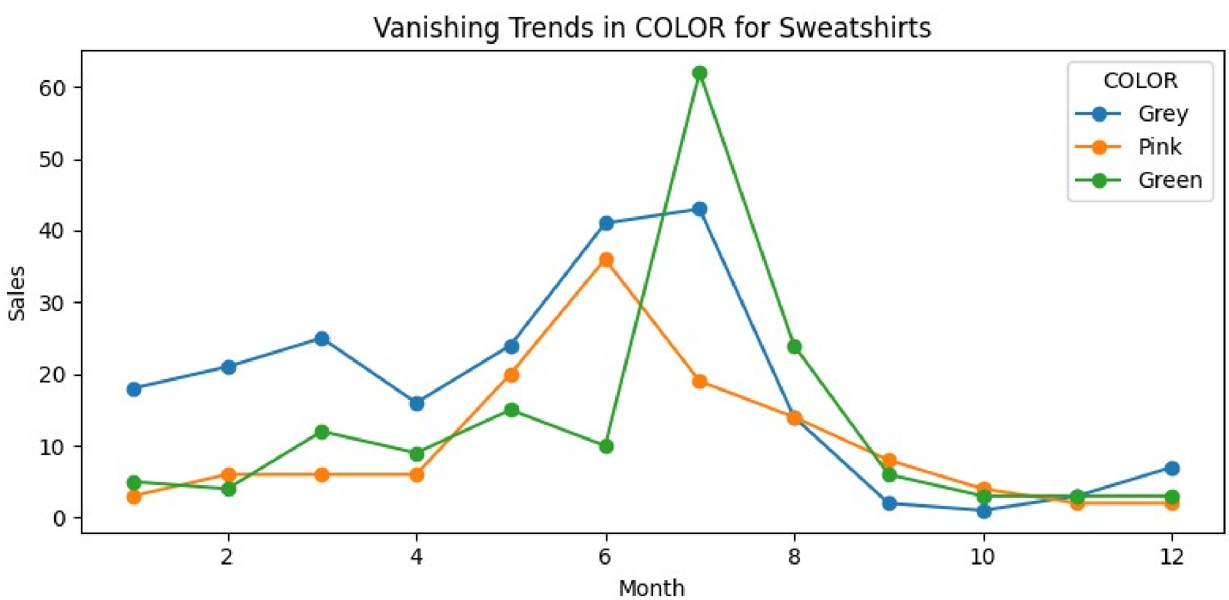

Vanishing Trends in OCCASSION_NEW for Pyjamas
OCCASSION_NEW
Sleep and Lounge Wear
Sales
300
250
200
150
100
50
2
4
6
8
10
12
Month

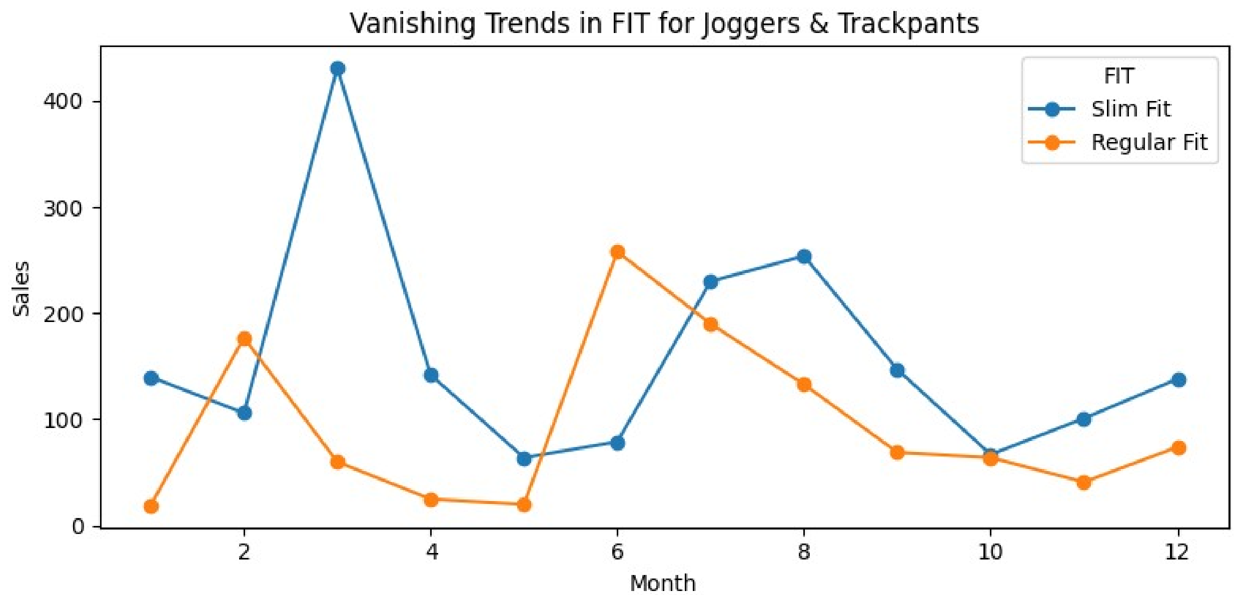

Vanishing Trends in COLOR for Pyjamas
COLOR
Blue
White
Navy
Sales
80
70
60
50
40
30
20
10
0
2
4
6
8
10
12
Month

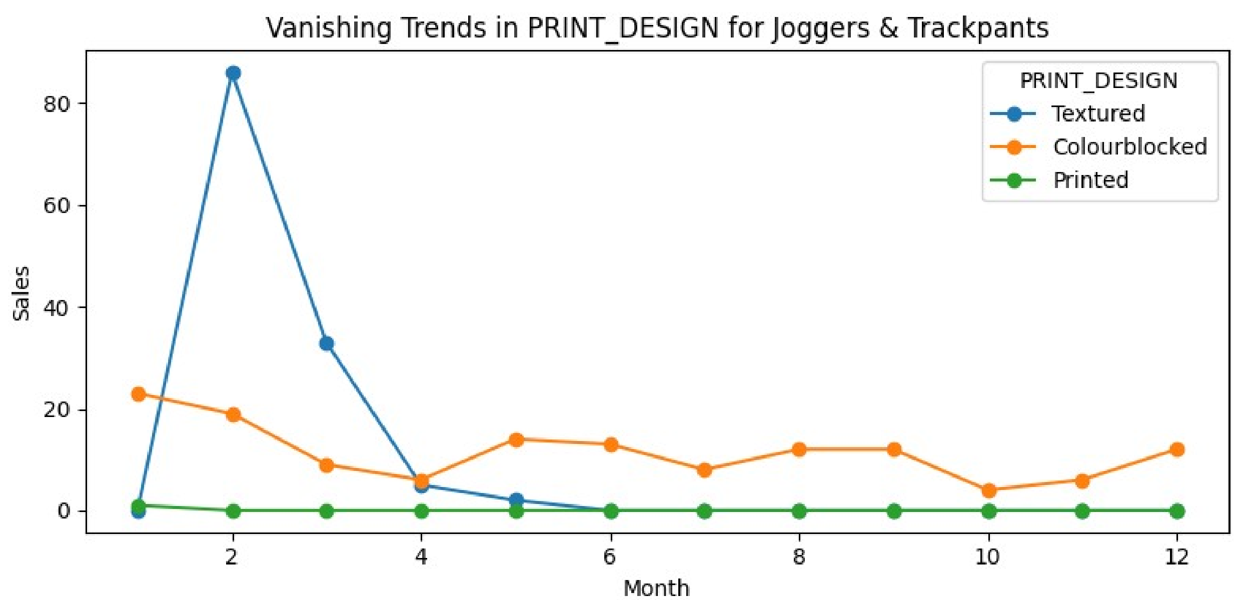

Vanishing Trends in FIT for Sweatshirts
FIT
Relaxed Fit
Sales
Month
350
300
250
200
150
100
50
2
4
6
8
10
12


Vanishing Trends in MATERIAL_NEW for Sweatshirts
MATERIAL_NEW
Polyester
Sales
Month
30
25
20
15
10
5
0
2
4
6
8
10
12


Vanishing Trends in PRINT_DESIGN for Sweatshirts
PRINT_DESIGN
Embroidered
Printed
Textured
Sales
Month
0
20
40
60
80
2
4
6
8
10
12


Vanishing Trends in COLOR for Sweatshirts
COLOR
Grey
Pink
Green
Sales
Month
0
10
20
30
40
50
60
2
4
6
8
10
12


Vanishing Trends in FIT for Joggers & Trackpants
FIT
Slim Fit
Regular Fit
Sales
Month
0
100
200
300
400
2
4
6
8
10
12


Vanishing Trends in PRINT_DESIGN for Joggers & Trackpants
PRINT_DESIGN
Textured
Colourblocked
Printed
Sales
Month
0
20
40
60
80
2
4
6
8
10
12

Q4.Provide composition on category and feature level (important features determined in the first question) for example if Fit is identified as a crucial feature for Shirts then what will be composition of different materials that should be produced next for example Oversized 70% and Regular 30% for Shirts.

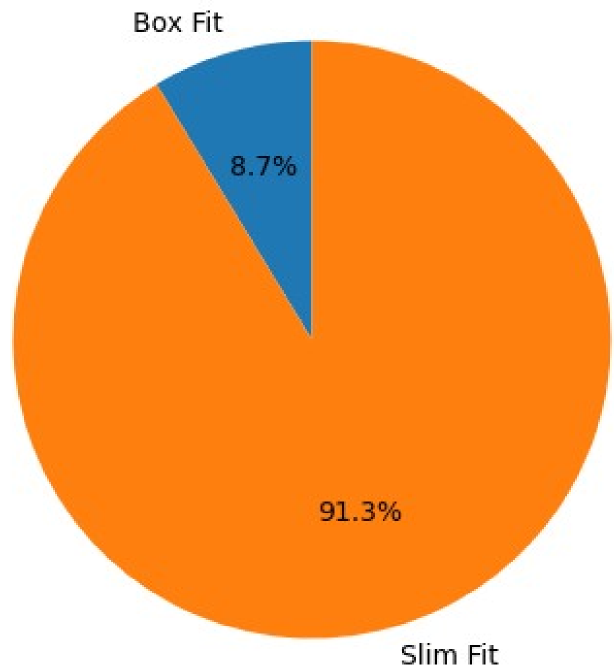
Approach -> For each category, the top 3 significant features (previously identified via importance metrics) were selected to analyze sales composition. Feature values contributing less than 5% of total sales were excluded to eliminate noise and emphasize major contributors. The relative percentage share of each qualifying feature value was computed and visualized through pie charts to inform production prioritization based on dominant sales contributors.

```
# Get top 3 features per category from previous analysis
# (Replacing this with our actual feature importance results)
category_top_features = {
    'Shirts': ['FIT', 'SLEEVE_TYPE', 'OCCASSION_NEW'],
    'T-Shirts': ['OCCASSION_NEW', 'SLEEVE_TYPE', 'FIT'],
    'Jeans': ['FIT', 'MATERIAL_NEW', 'PRINT_DESIGN'],
    'Trousers': ['FIT', 'MATERIAL_NEW', 'COLLAR_NEW'],
    'Jackets': ['OCCASSION_NEW', 'SLEEVE_TYPE','FIT'],
    'Sweaters': ['OCCASSION_NEW', 'SLEEVE_TYPE','FIT'],
    'Overshirt': ['OCCASSION_NEW', 'SLEEVE_TYPE','FIT']
}

def calculate_composition(category, features):
    cat_df = df[df['CATEGORY'] == category]
    compositions = {}

    for feature in features:
        # Calculate sales distribution
        feature_dist = cat_df.groupby(feature)
['TOTAL_SALES'].sum().reset_index()
```

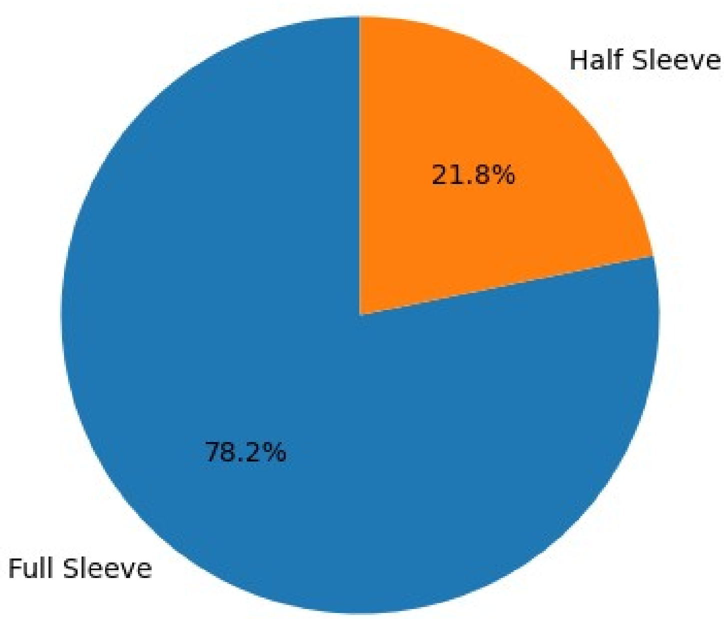
```
        total_sales = feature_dist['TOTAL_SALES'].sum()

        # Calculate percentages and filter out insignificant values
(<5%)
        feature_dist['PERCENTAGE'] = (feature_dist['TOTAL_SALES'] /
total_sales * 100).round(1)
        feature_dist = feature_dist[feature_dist['PERCENTAGE'] >= 5]

        # Normalize to 100%
        total_percentage = feature_dist['PERCENTAGE'].sum()
        feature_dist['PERCENTAGE'] = (feature_dist['PERCENTAGE'] /
total_percentage * 100).round(1)

        compositions[feature] = feature_dist.set_index(feature)
['PERCENTAGE'].to_dict()

    return compositions

# Generate and display compositions for all categories
for category, features in category_top_features.items():
    print(f"\n Production Composition for {category}:")
    composition = calculate_composition(category, features)

    for feature, values in composition.items():


        # Pie chart visualization
        plt.figure(figsize=(5,5))
        plt.pie(values.values(), labels=values.keys(), autopct='%1.1f%
%', startangle=90)
        plt.title(f"{category} - {feature} Composition")
        plt.show()
```

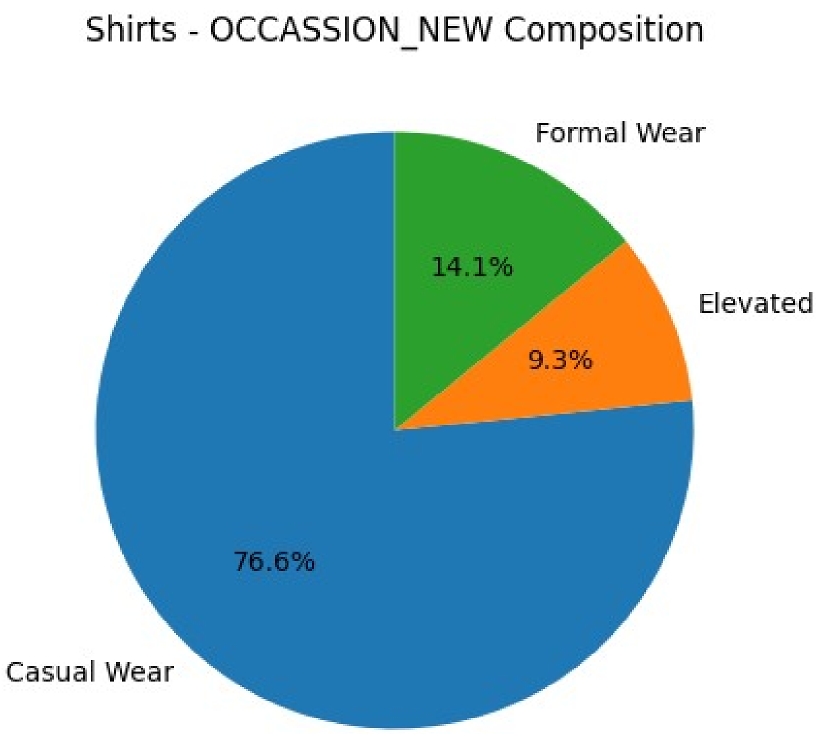
```
 Production Composition for Shirts:
```

Shirts - FIT Composition



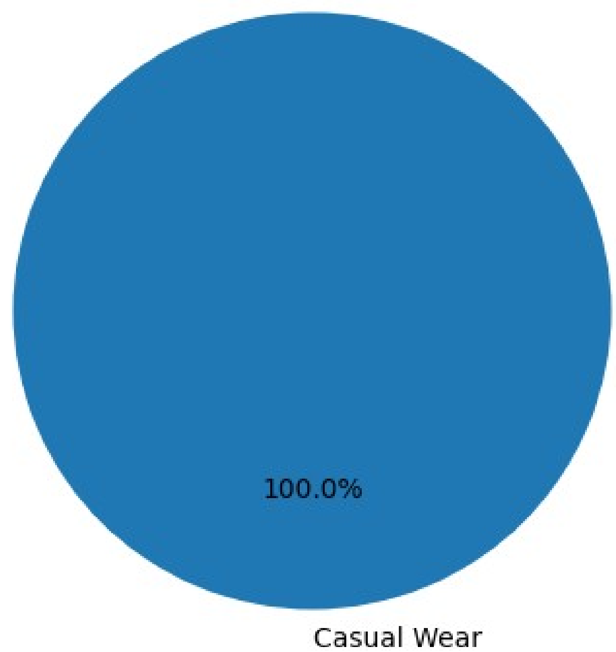
Shirts - SLEEVE_TYPE Composition



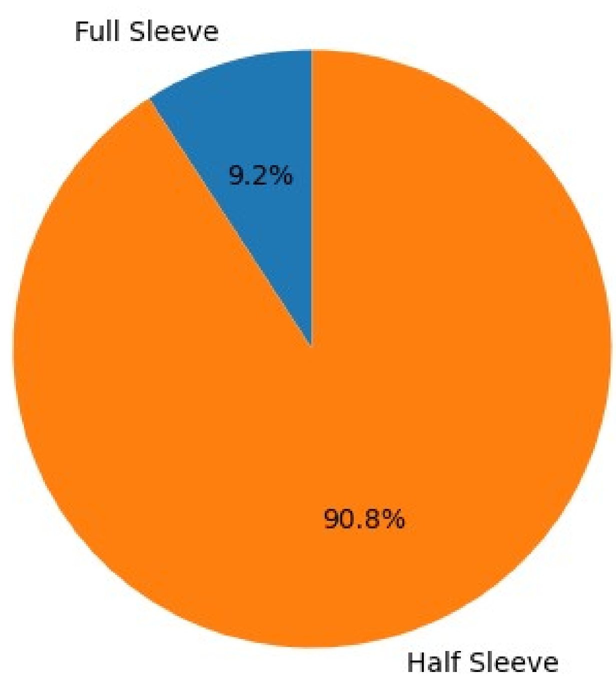


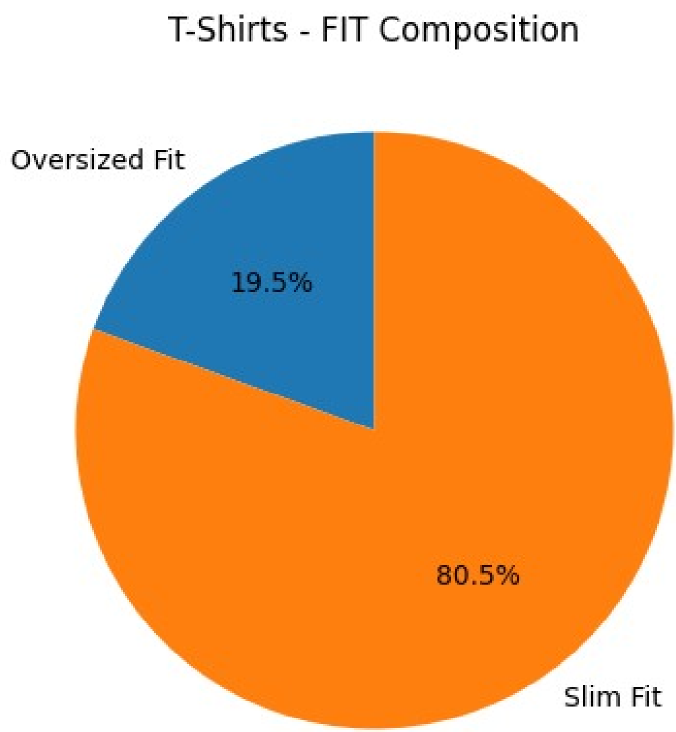

```
Production Composition for T-Shirts:
```

T-Shirts - OCCASSION_NEW Composition



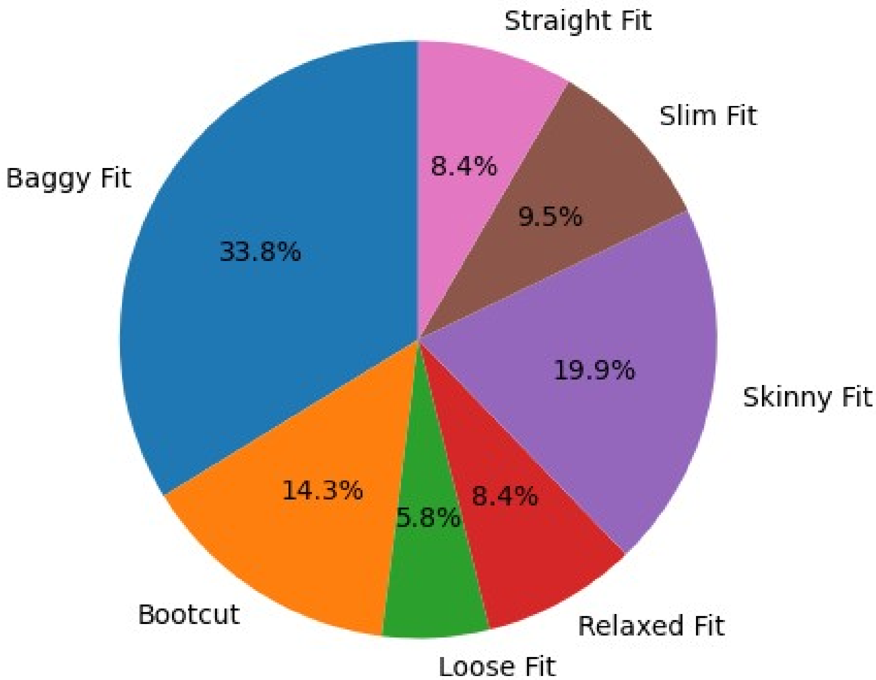
T-Shirts - SLEEVE_TYPE Composition



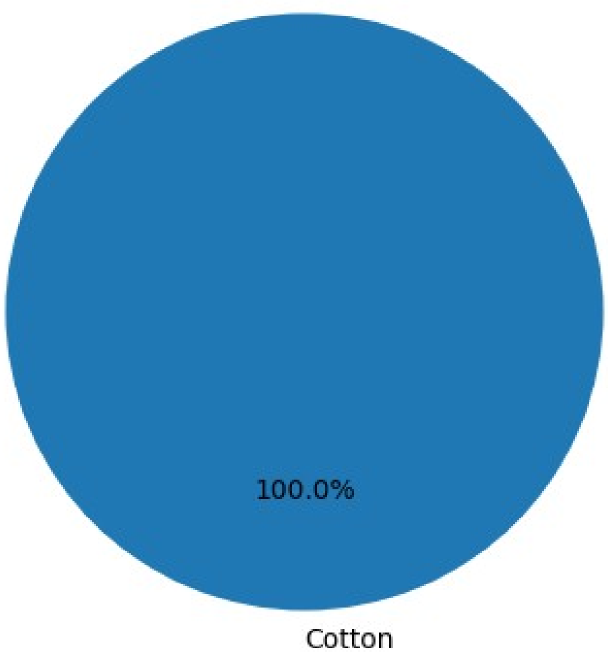


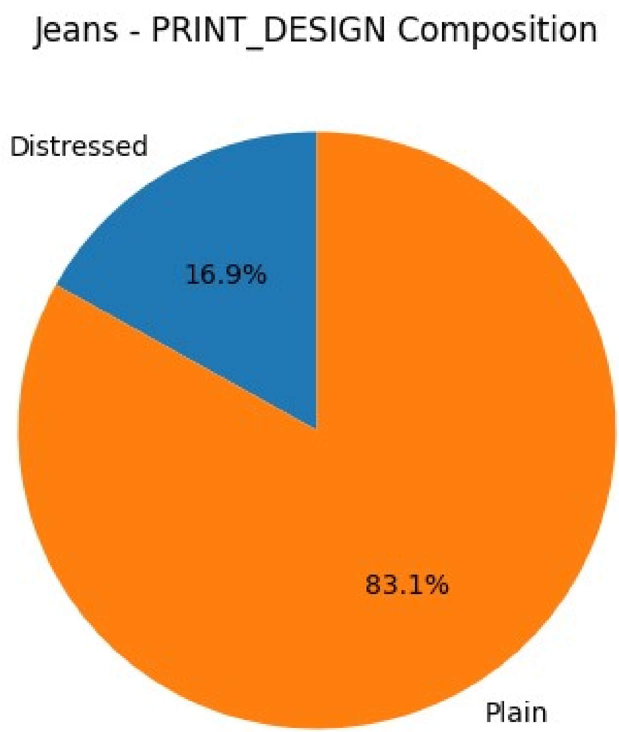

```
Production Composition for Jeans:
```

Jeans - FIT Composition



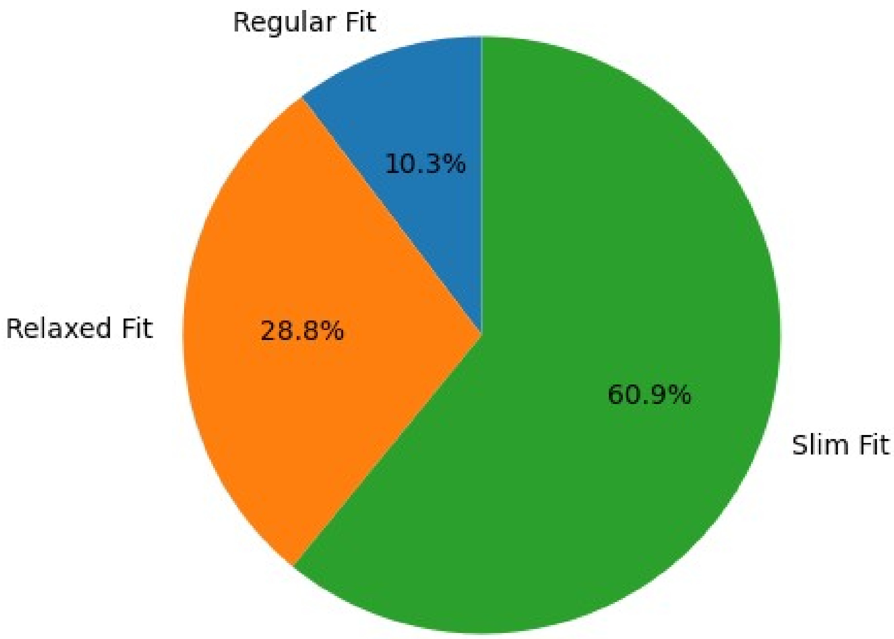
Jeans - MATERIAL_NEW Composition



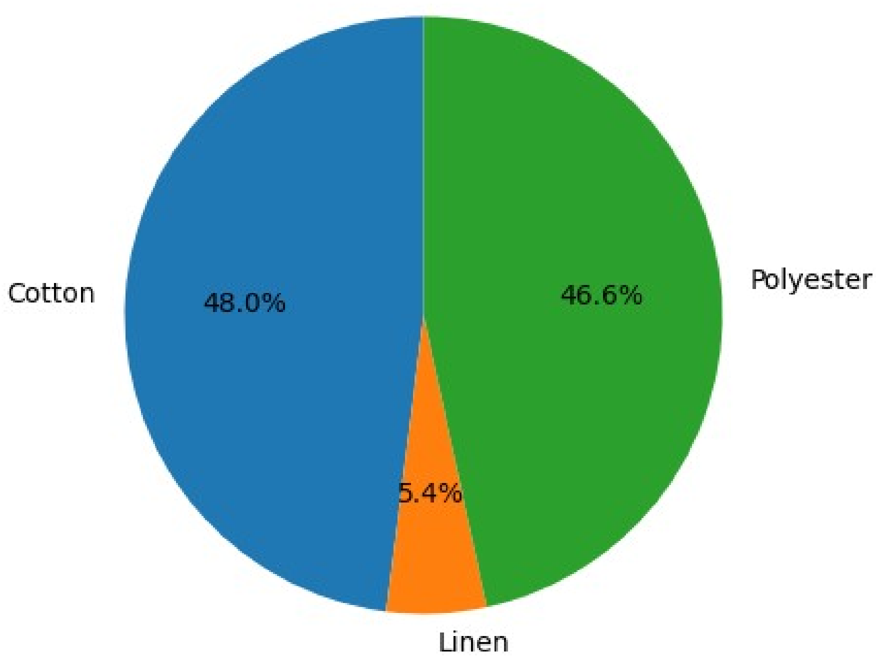


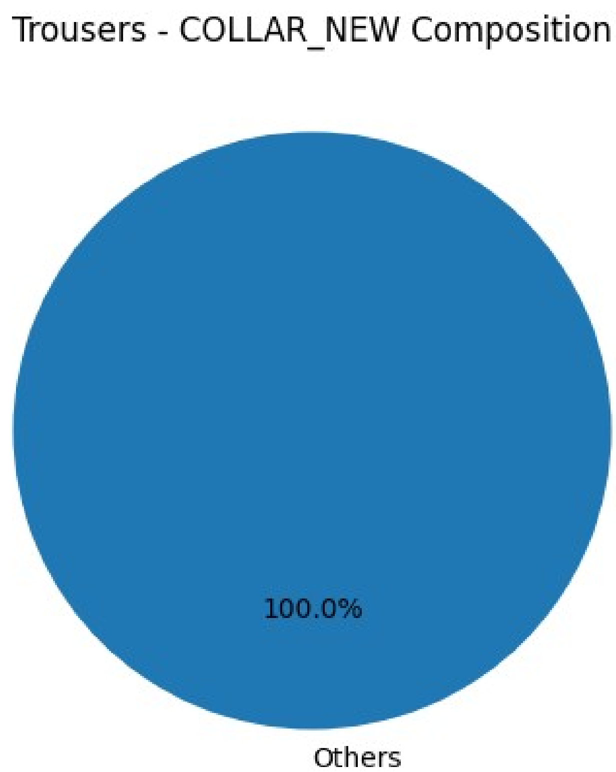

```
Production Composition for Trousers:
```

Trousers - FIT Composition



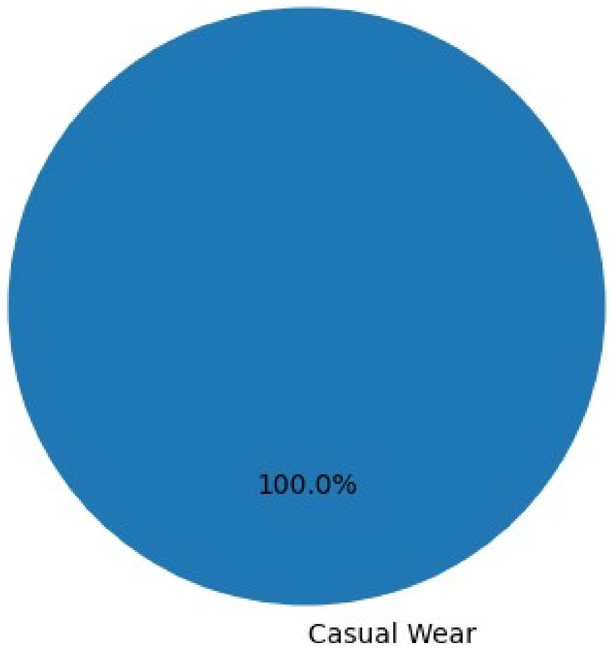
Trousers - MATERIAL_NEW Composition



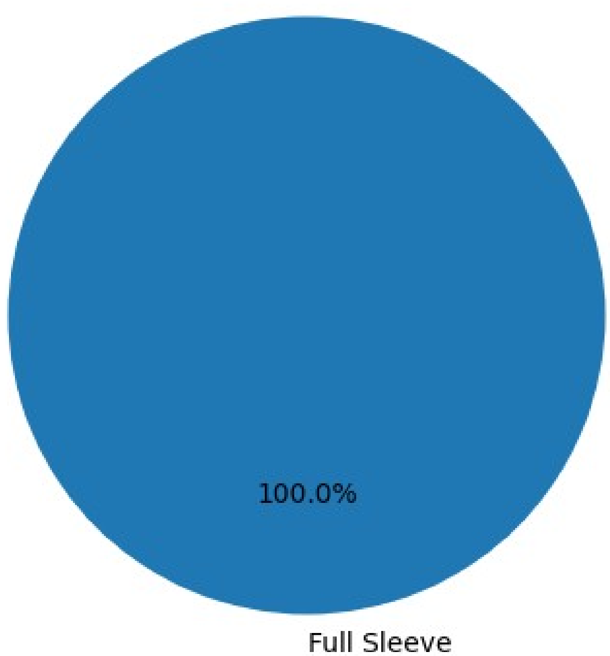


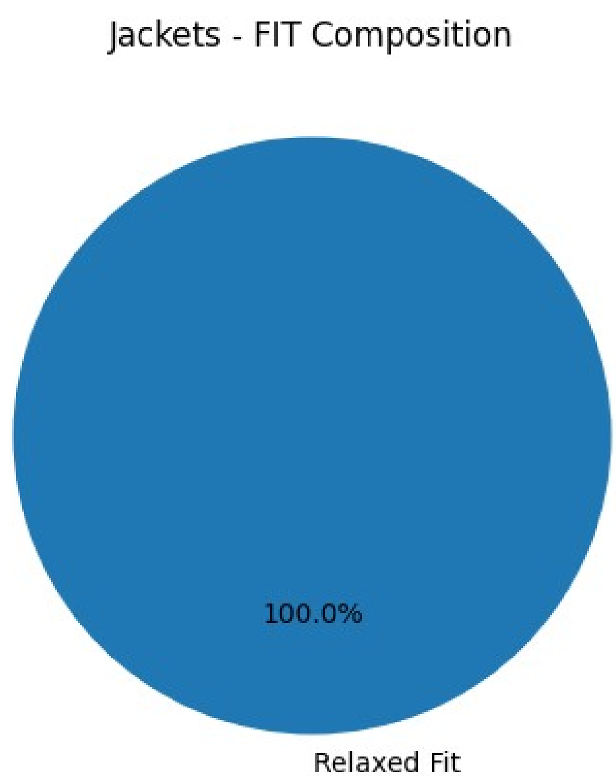

```
Production Composition for Jackets:
```

Jackets - OCCASSION_NEW Composition



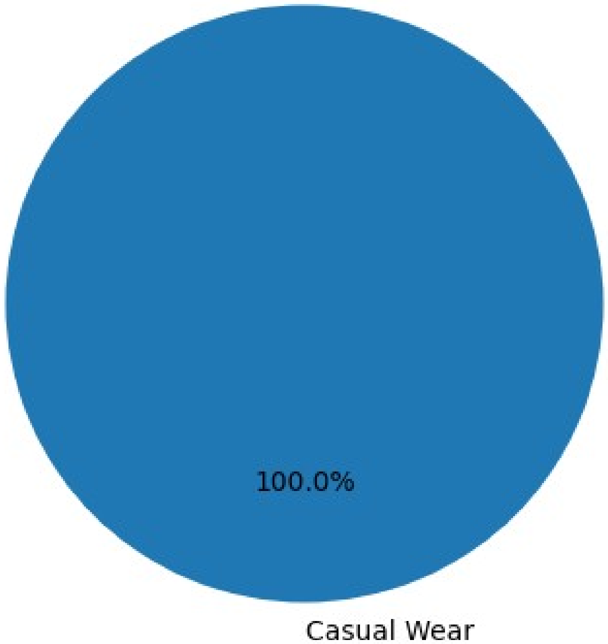
Jackets - SLEEVE_TYPE Composition



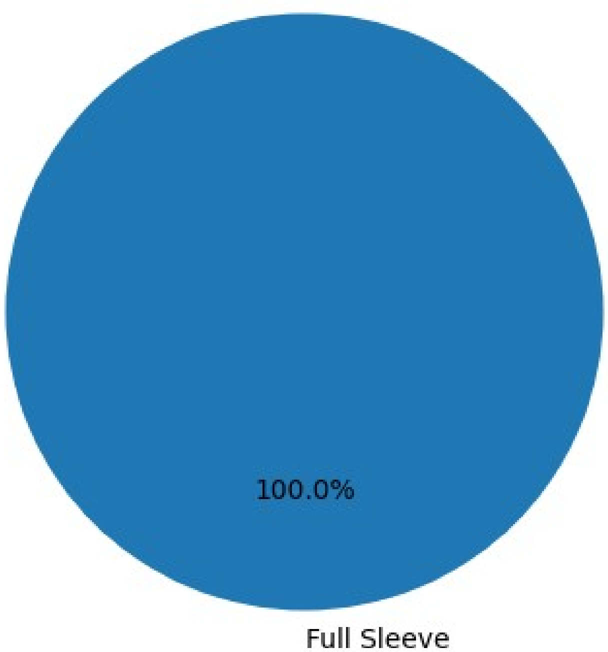



```
Production Composition for Sweaters:
```

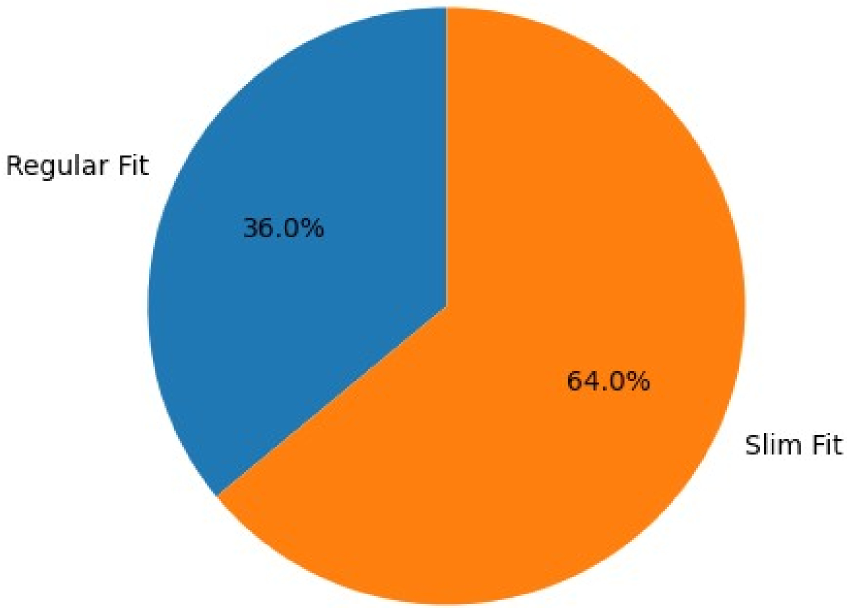
Sweaters - OCCASSION_NEW Composition



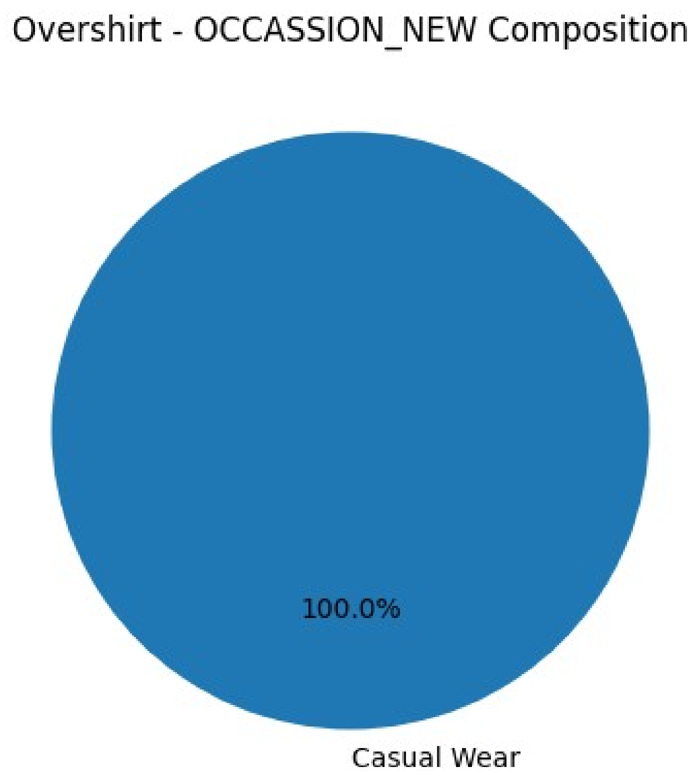
Sweaters - SLEEVE_TYPE Composition

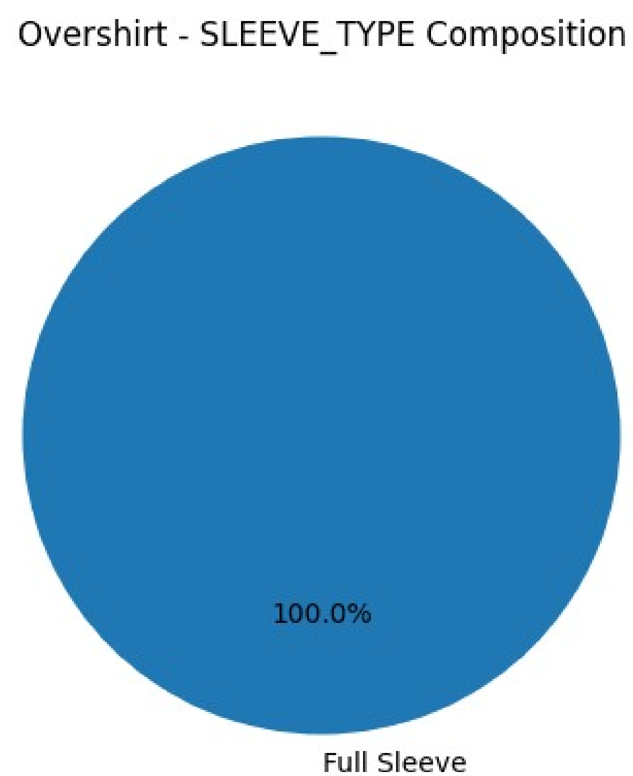


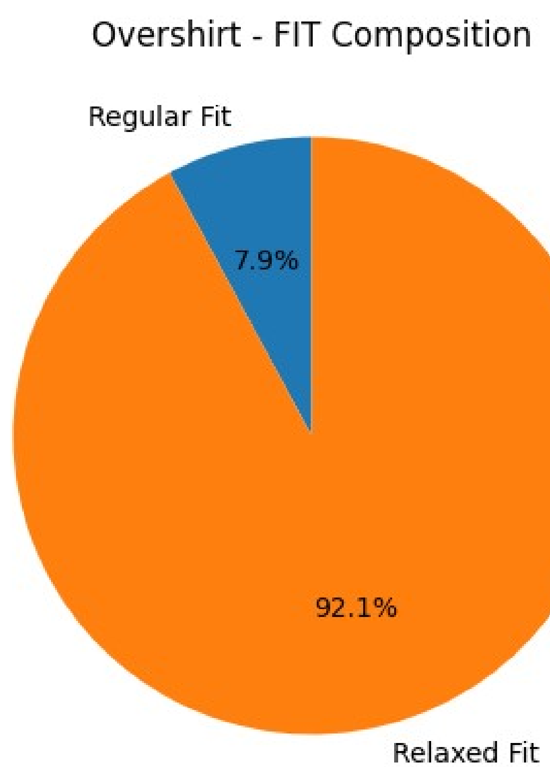
Sweaters - FIT Composition




```
Production Composition for Overshirt:
```


Overshirt - OCCASSION_NEW Composition
100.0%
Casual Wear


Overshirt - SLEEVE_TYPE Composition
100.0%
Full Sleeve


Overshirt - FIT Composition
Regular Fit
7.9%
92.1%
Relaxed Fit