Sportswear Group Report

Dataset:

For this project, I'm using full_gen_data which belongs to sportswear group.

Methodology:

For this project, I will be following IBM crisp-dm methodology.

1. Business Understanding:

  a. Problem Definition:
     The provided data represents information from a marketing campaign referred to sportswear group. The team provided information about the product in the campaign that was sent to a specific customer and the result of this campaign. Based on census features, the machine learning task is to predict if the customer will buy or not. before/after sending the offer.

  b. Scope:
     The scope of this sample is:
        i. Create binary machine learning classification model to address whether the customer will buy product or no.
        ii. Create fine marketing strategy to help increase sales.

  c. Metrics:
        i. The performance of the machine learning model will be evaluated on the test dataset provided. Accuracy > 0.8 will be considered acceptable and suitable for deployment.

2. Date Understanding:

  There are total of 10,000 records in the given dataset.
  Target column has 2 classes, either 0 or 1
  Probability of the target variable with' 0,': 86%
  Probability of the target variable with' 1': 14%
 Original dataset has 24 columns.
  Original dataset has 100,000 instances.

3. Data Preparation:
   a. Data Cleaning: The dataset was almost clean. Nothing to mention.
   b. Feature Engineering:
      i. First, I concatenated the RBG layers to new feature called color - numbers to colors.
      ii. Parsed retailweek and extracted feature to know which week per year are that date in.
      iii. Calculated the total sales field from 2 other columns.

4. Analysis and Recommendations:

Weekly sales - Mean and Median
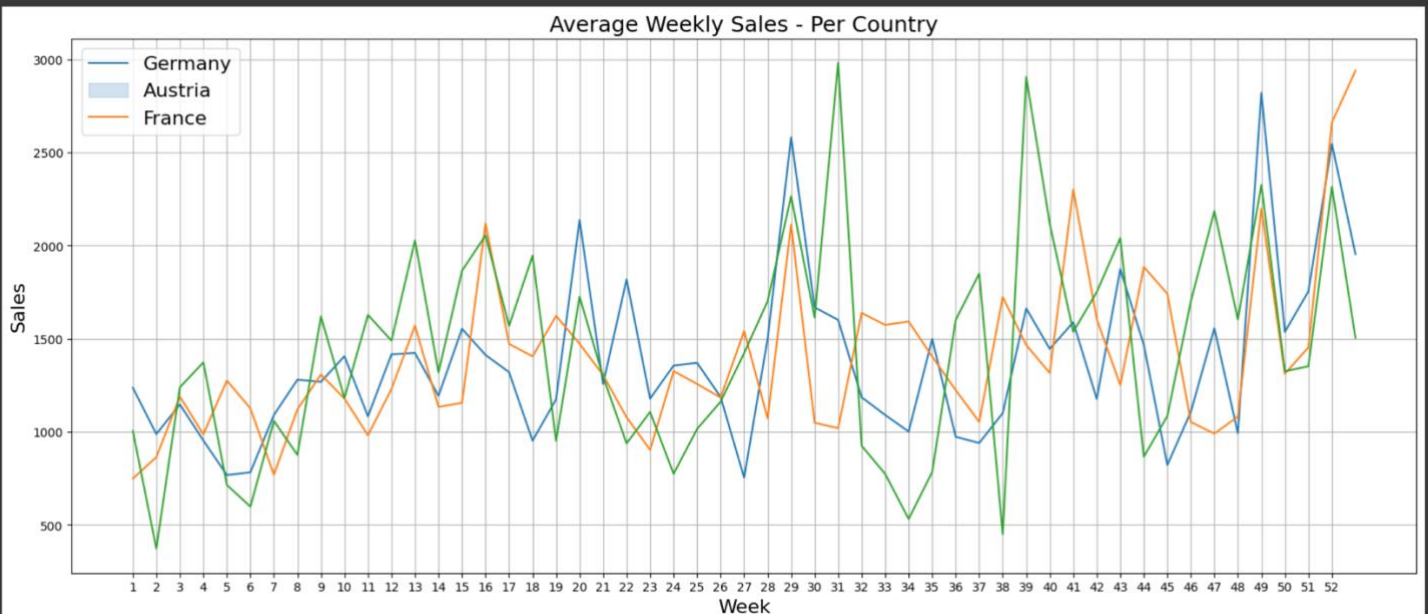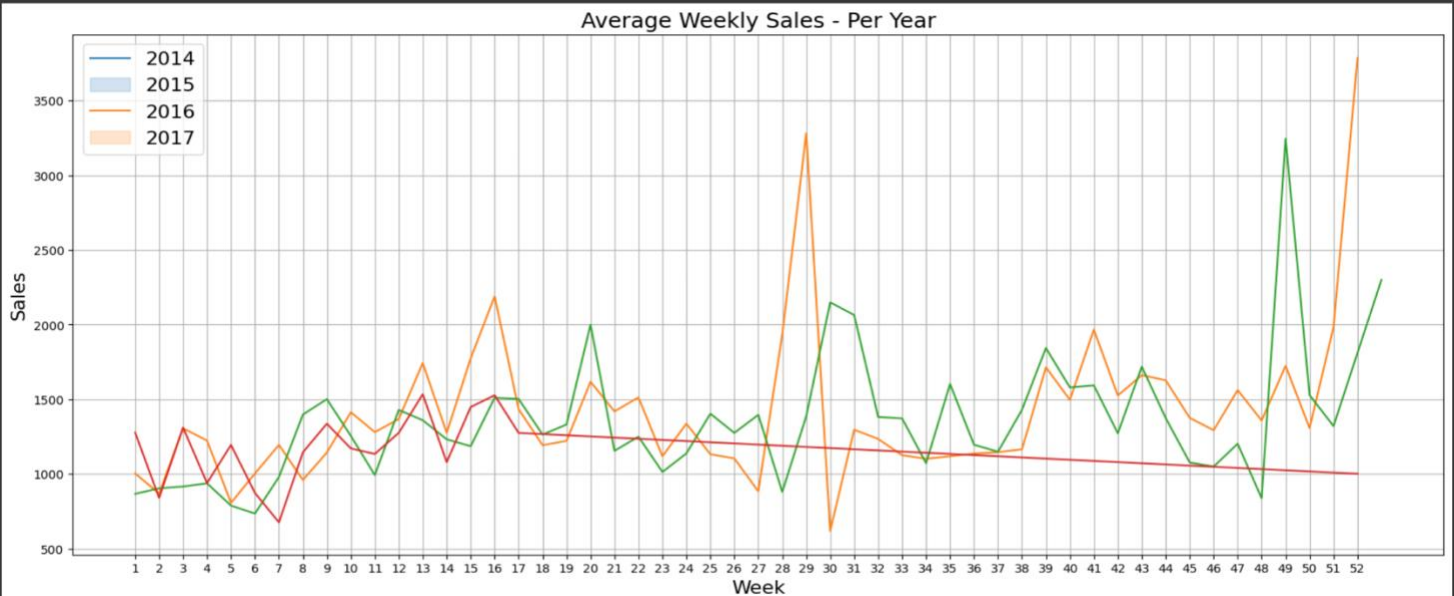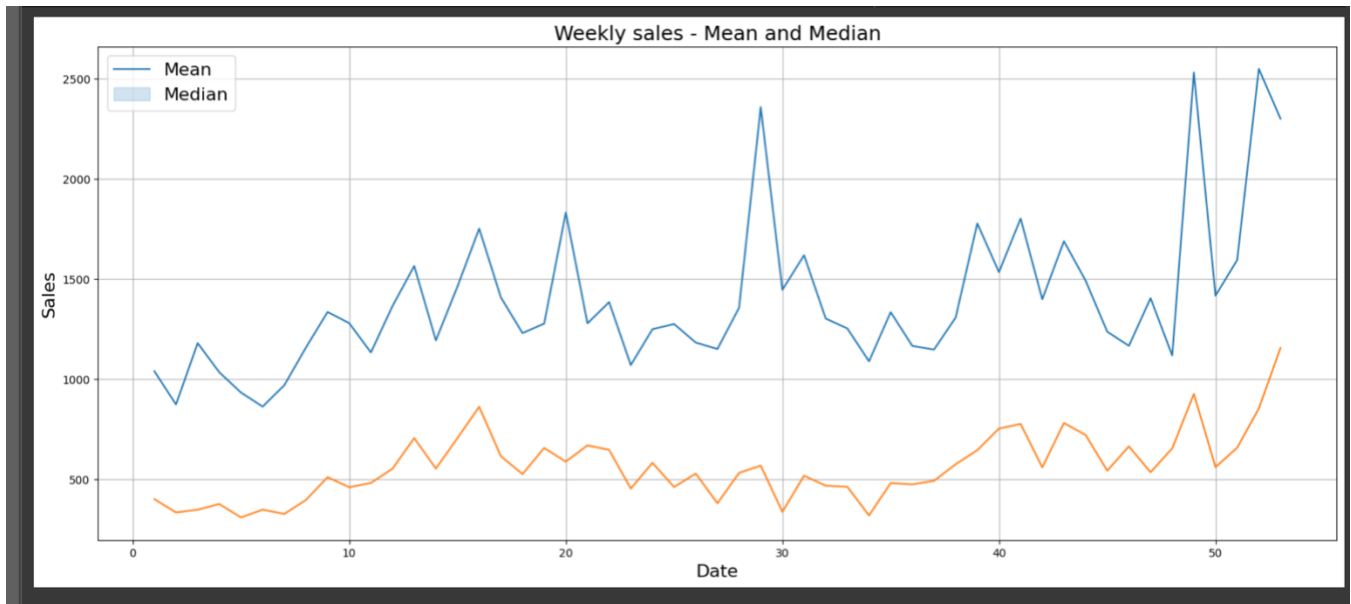
1ˢᵗ quarter:

1. Average sales during 2015 was the highest among others.
2. Average sales during 2017 are the 2ⁿᵈ highest had a good uprising sales curve, but data was insufficient.
3. Germany has the highest sales among other countries.

2ⁿᵈ quarter:

1. Average sales during 2015 was the highest ever.
2. Average sales during 2014 was the lowest among all datasets.
3. Which means the 2ⁿᵈ quarter of each year the sales are low among all genders.
4. All countries have almost the same sales. (People does not buy during 2ⁿᵈ quarter)

3rd quarter:

1. Nothing important to observe.
2. Austria has the highest sales among all countries.

4th quarter:

1. Each year, the 4th quarter average sales are rising, which means clients buy by the end of the year.
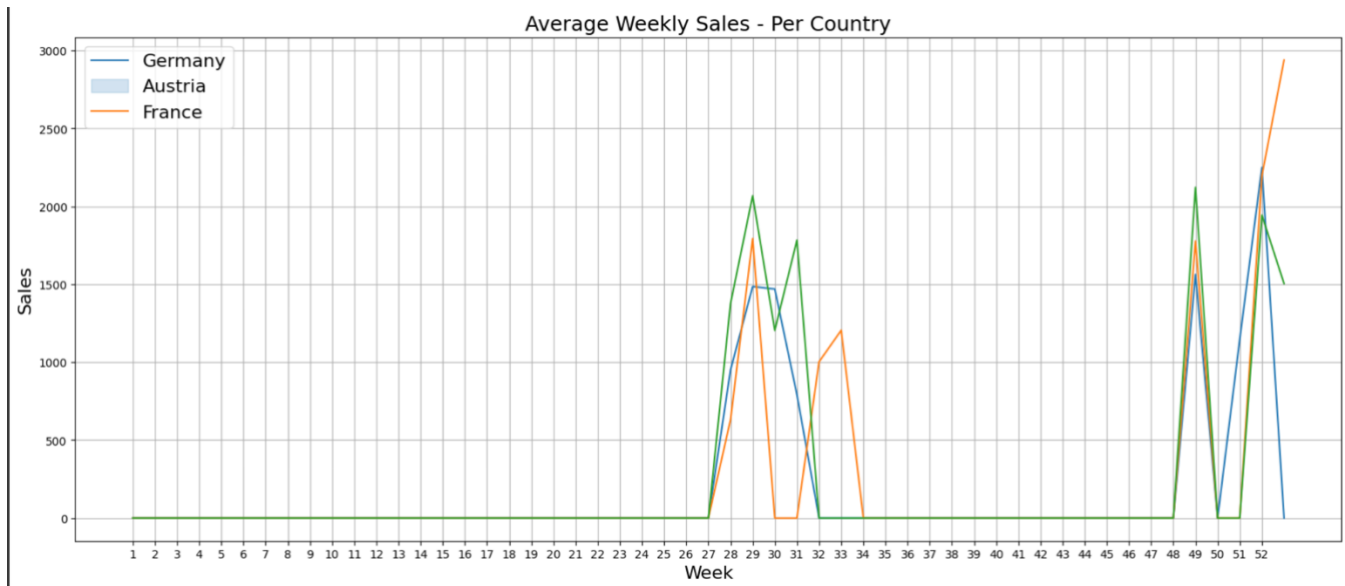2. Countries are almost the same. But 2nd quarter is still the highest.

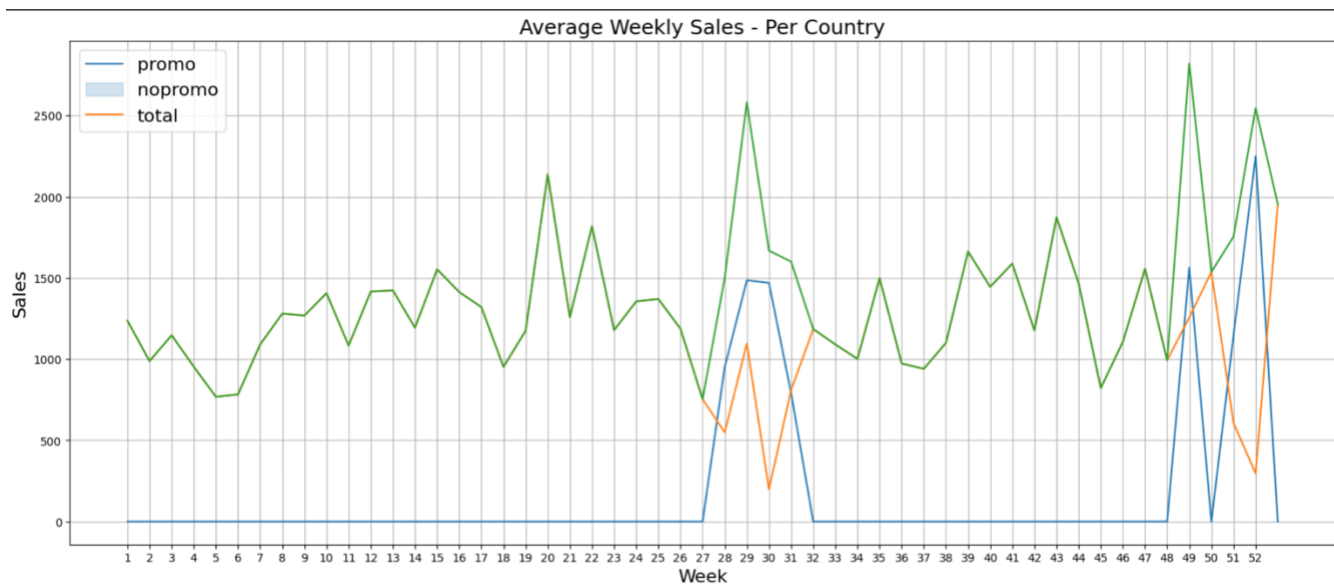*Figure 1 Sales when promotion 1 is activated for countries*

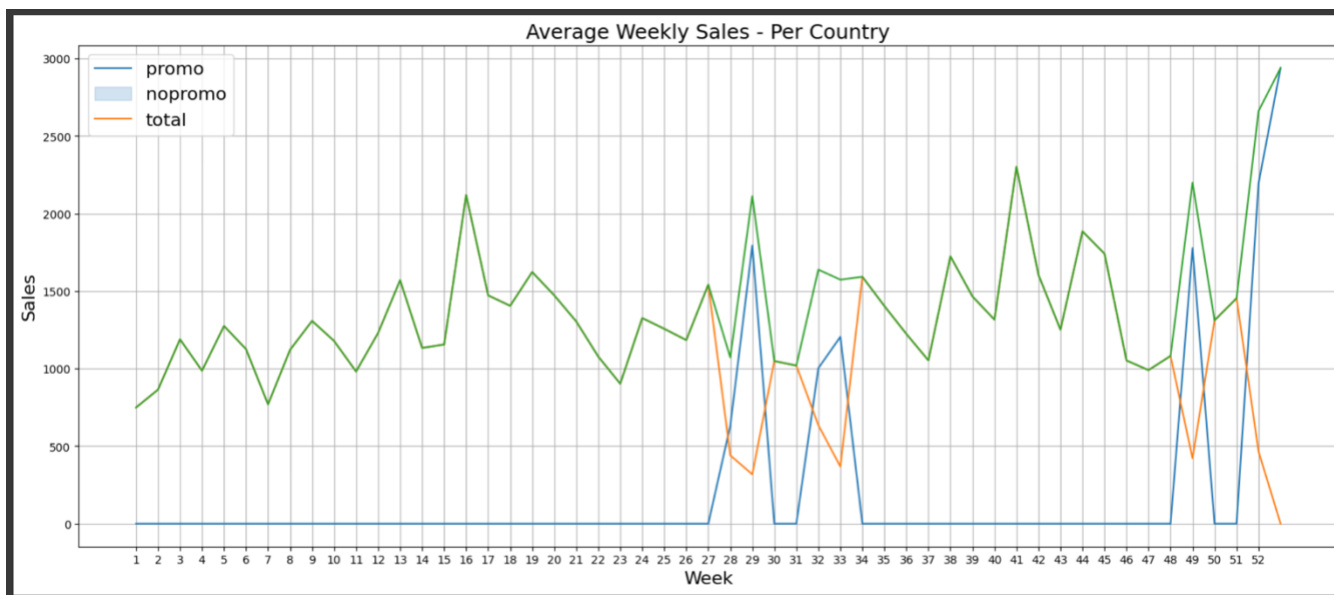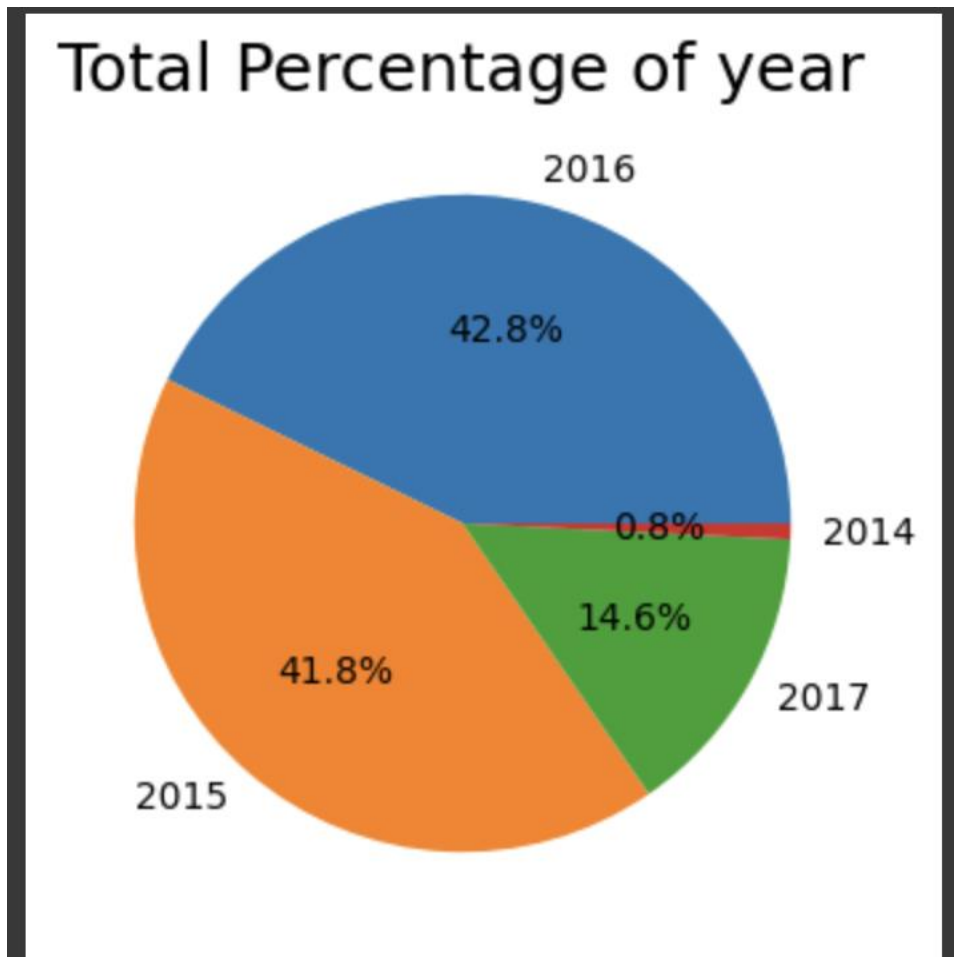*Figure 2 Sales Total sales, promotion 1 activated, promotion 1 deactivated FOR GERMANY ONLY*
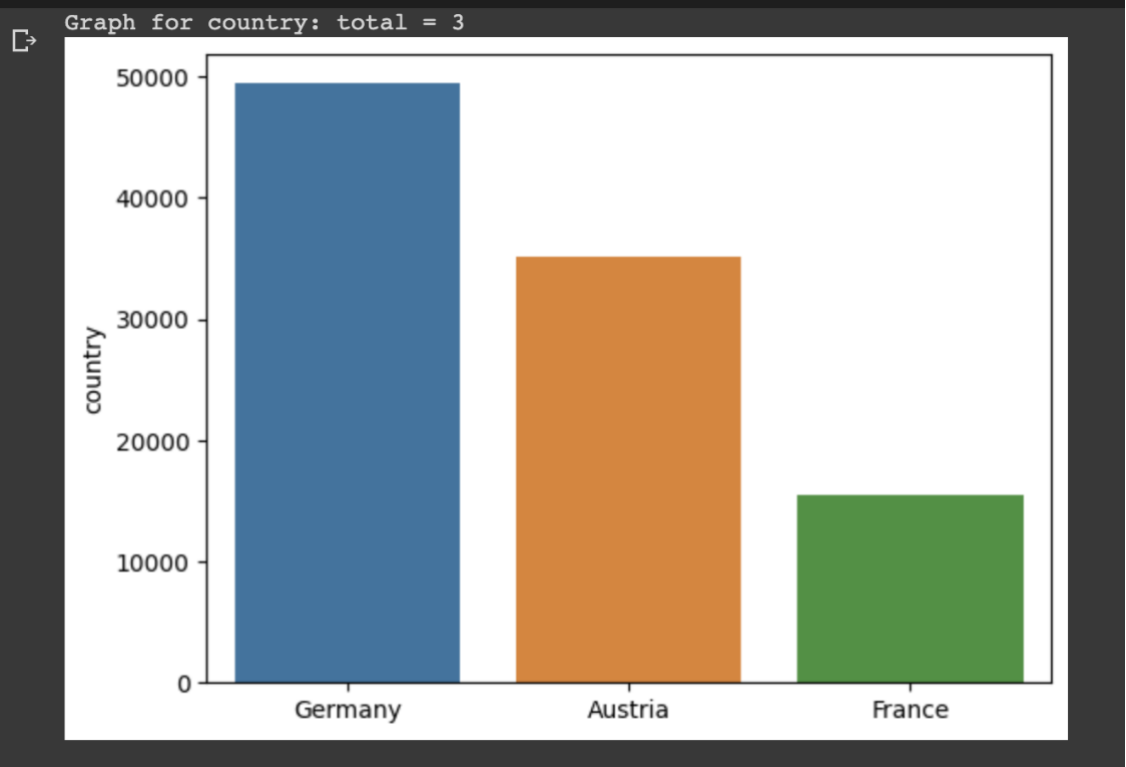

Figure 3 Sales for Austria with promotion 1 activated, deactivated and total sales

Promotions are increasing sales.
Promotion 1 increases sales by 240%
Promotion 2 increases sales by 114%

Graph for country: total = 3
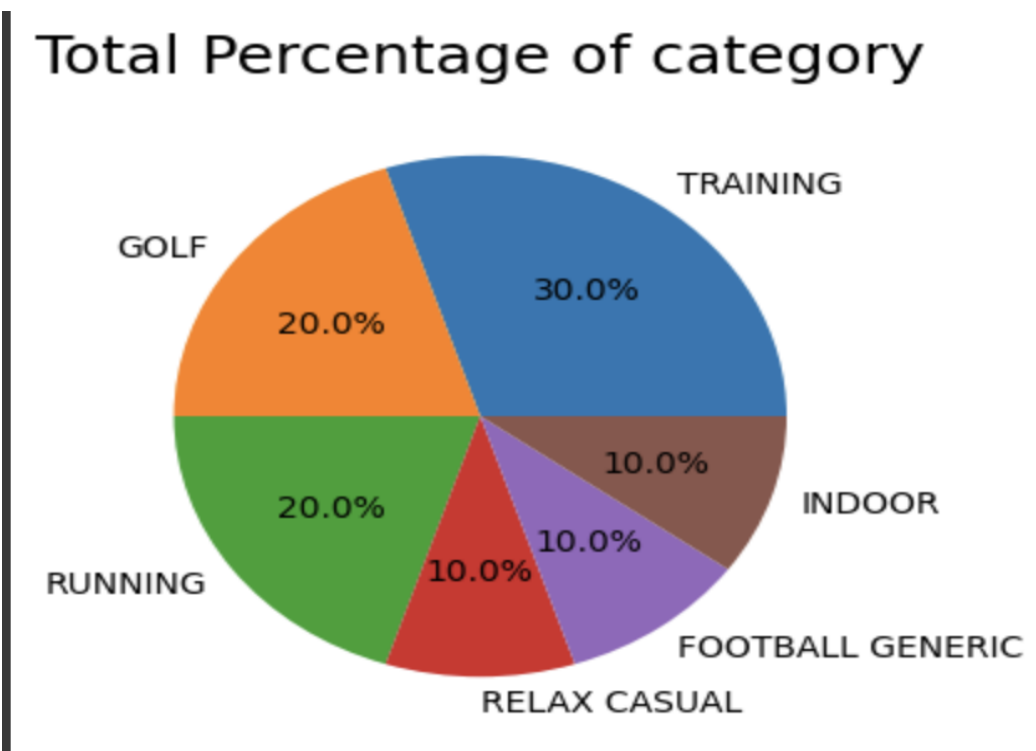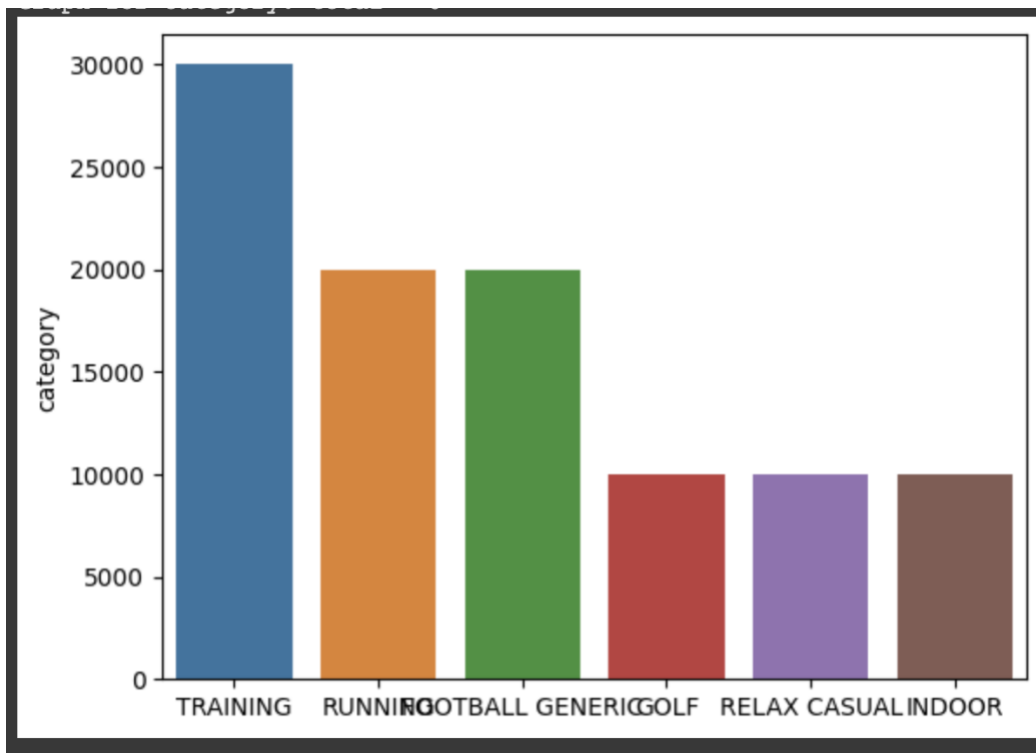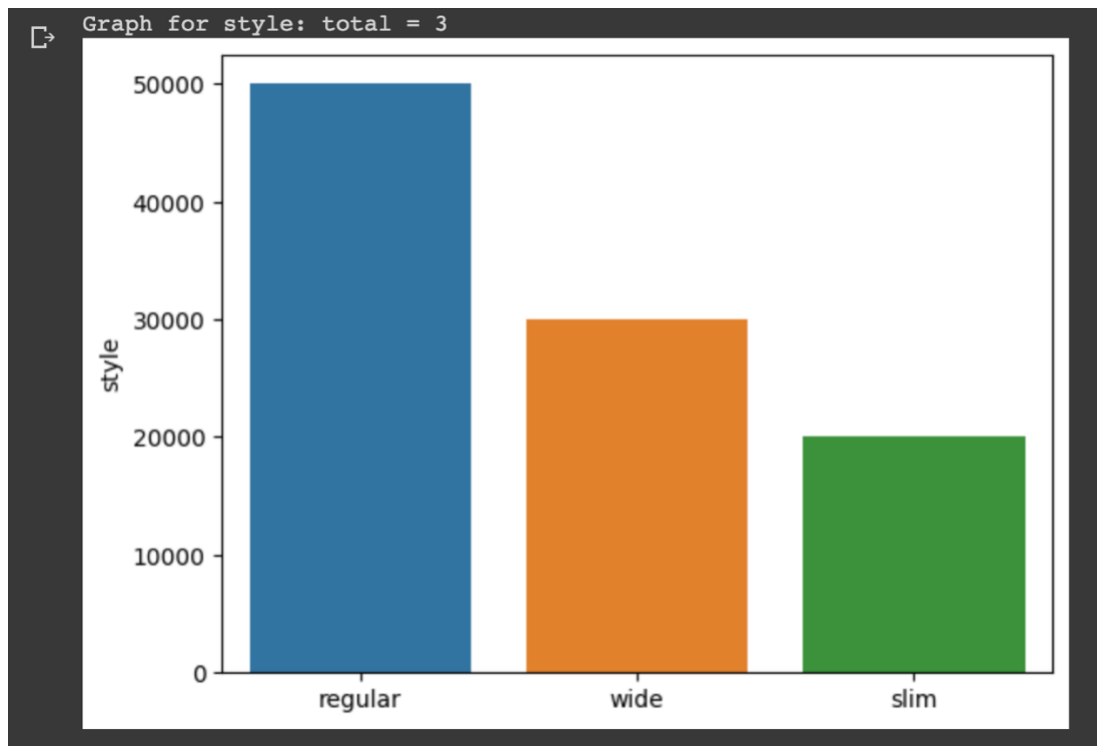

Total Percentage of year

As state above, Germany has the highest sales among other countries. Austria comes second, France third.
Total sales percentage increased by 1% from 2015, to 2016
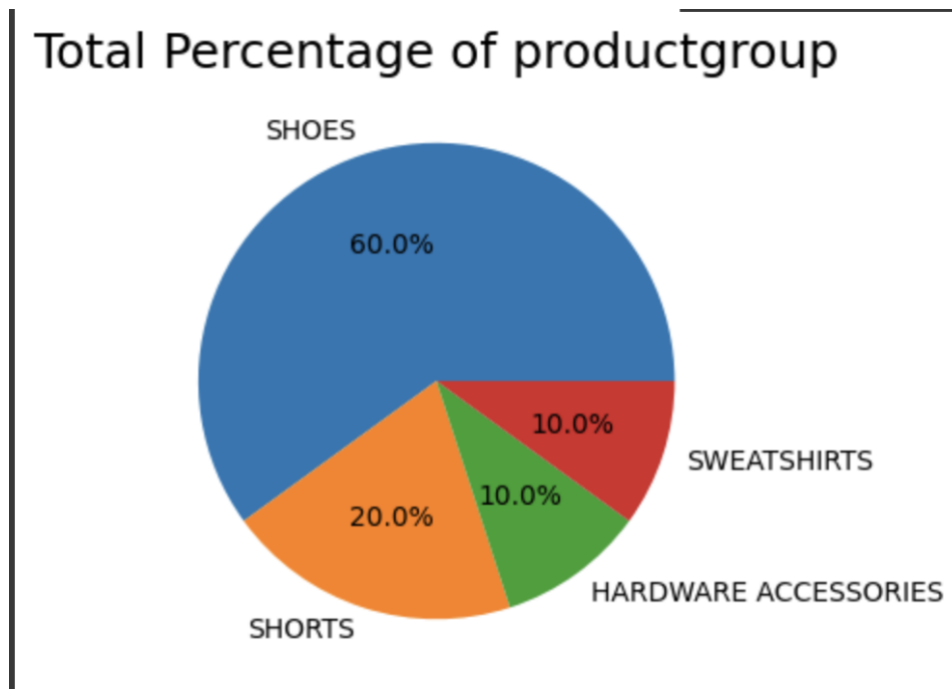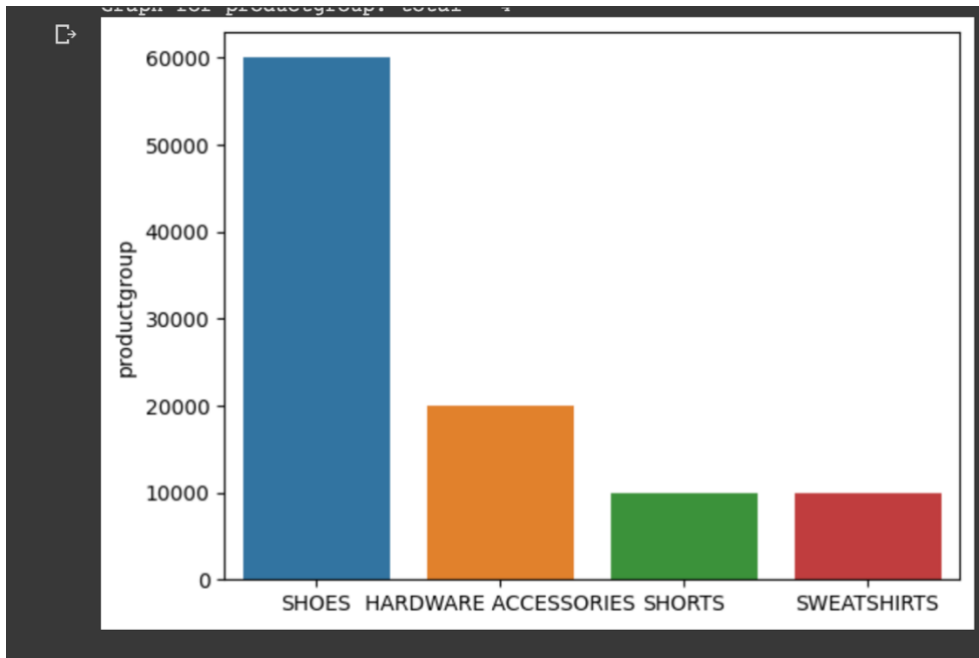
As illustrated, shoes has the highest sales with 60%, shorts comes second with 20%

Total Percentage of category

Graph for style: total = 3

Training categories, regular fit (which include shoes, shorts, etc..) are the highest.

Graph for gender: total = 4
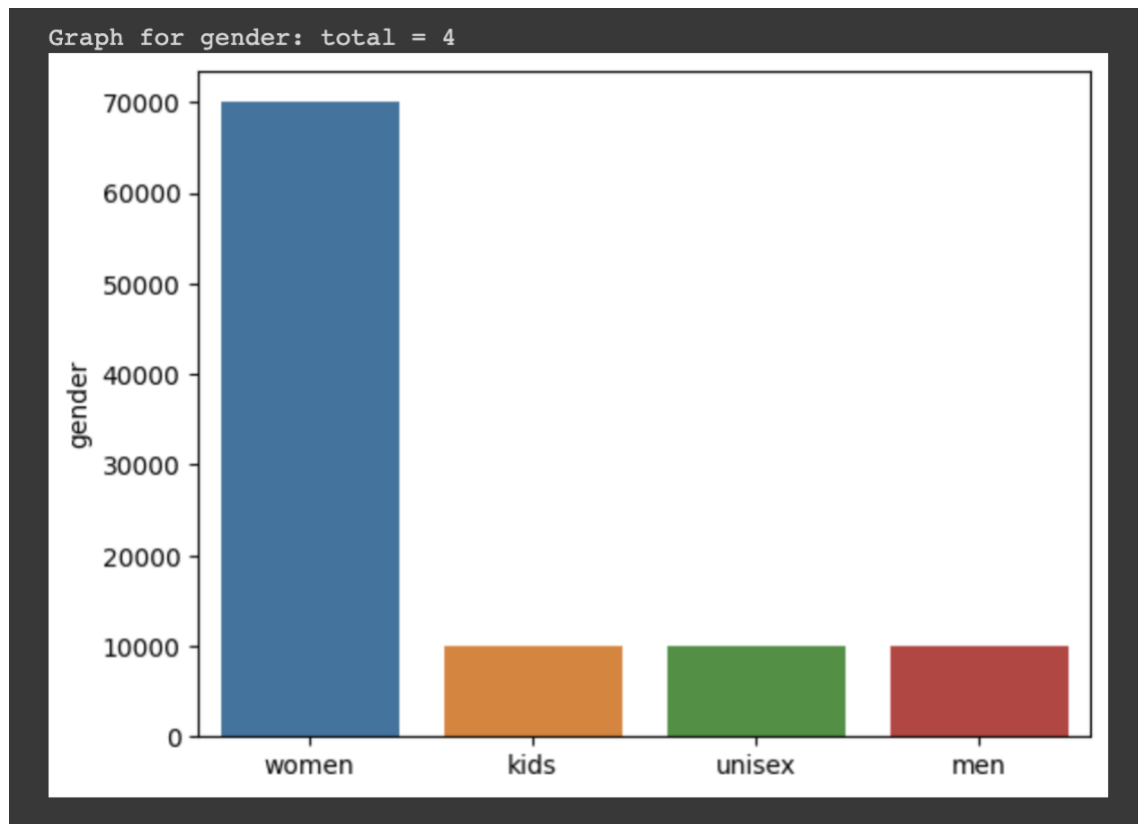
Figure 4 we can state that that women are the main consumers for our sportwear group.

We will discuss more after the pivot tables below.

```
[185]                     total_sales

        country

        Germany        66563324.0

        Austria        48152895.0

        France         21998102.0
```

```
1 pd.pivot_table(pivot_df, index=
```

country	gender	total_sales
Austria	kids	4815289.5
	men	4815289.5
	unisex	4815289.5
	women	33707026.5
France	kids	2199810.2
	men	2199810.2
	unisex	2199810.2
	women	15398671.4
Germany	kids	6656332.4
	men	6656332.4
	unisex	6656332.4
	women	46594326.8

Recommendations:

There are two strategies:
Let's state at first some facts:

1. Since Germany was the highest country in sales among others
2. Since women are the main consumers from the above charts.
3. Since promotion 1 did increase sales by 240% which is much better than promotion 2.
4. Since Shoes are 60% of the sales

   First strategy: To increase promotions among Germany and customize promotions for the main consumers (Women).So that this strategy might make Germany rise more and grab more women to the store.

   Second strategy: To increase promotions at Austria and France. Also customize promotions to be for men as they are not the main consumers. So, the promotions might tempt more customers to the store.

More strategies to be discussed during the meeting.