

END OF BOOTCAMP PROJECT



Sales Analysis & Customer Segmentation

Prepared by Léa Martinez

Project Organization with Trello

The image shows a Trello board with a weekly project organization. The board has five columns: Monday, Tuesday, Wednesday, Thursday, and Friday. Each column contains cards representing tasks or milestones.

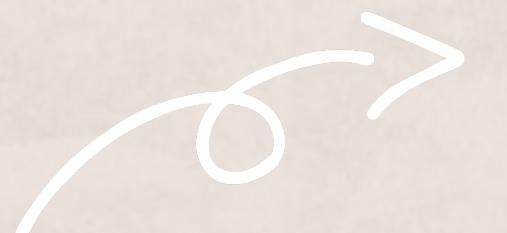
- Monday:**
 - Card: ★ TURN ★ IDEAS INTO REALITY
End of bootcamp Project, let's start !
 - Card: Review the different stages of the project
 - Card: Tableau Dataset Analysis
- Tuesday:**
 - Card: Part 1.
 - Card: Customer Segmentation : K-means & Clusters
 - Card: Clean data
- Wednesday:**
 - Card: Part. 2
 - Card: Comment my code and clean it
 - Card: Clusters Analysis
- Thursday:**
 - Card: Almost finish... Create the presentation now !
 - Card: Make the presentation on Canva
 - Card: Make a presentation trial
- Friday:**
 - Card: PROJECT DONE ! Congrats 🎉

Presentation of my dataset

Well-being E-Commerce Company
Sales data from 2018 to 2020

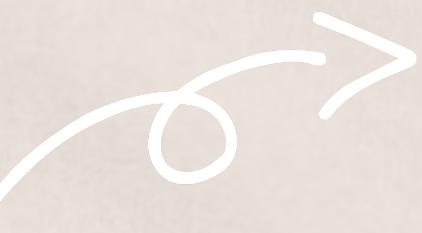
CUSTOMER SPECIFIC

- ID customer
- Date of Birth
- Newsletter opt-in



ORDER SPECIFIC

- Order ID
- Country
- Date of order
- Type of payment
- Order state
- Employee ID



PRODUCT SPECIFIC

- Product ID
- Product quantity
- Product weight

Business Goals

WHAT IS CUSTOMER SEGMENTATION

The practice of dividing a customer base into groups of individuals that are similar in specific ways relevant to marketing

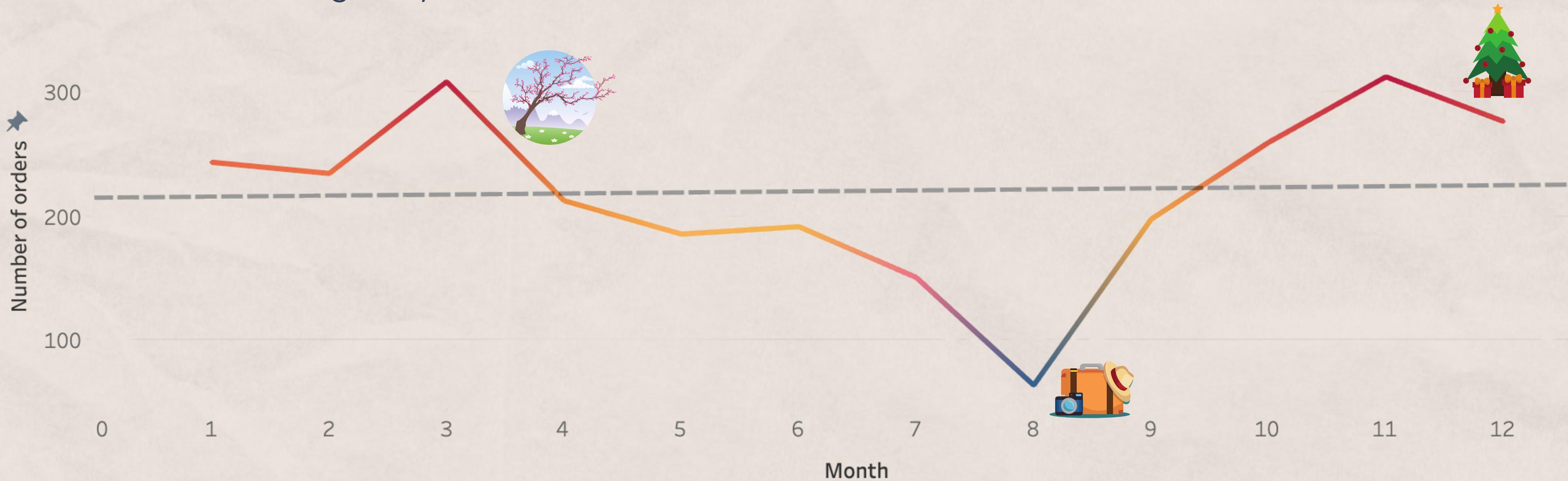
FOR WHICH APPLICATIONS ?

- Better targeting of customers
- More personalized communication
- Find the most effective strategy by segment to improve sales

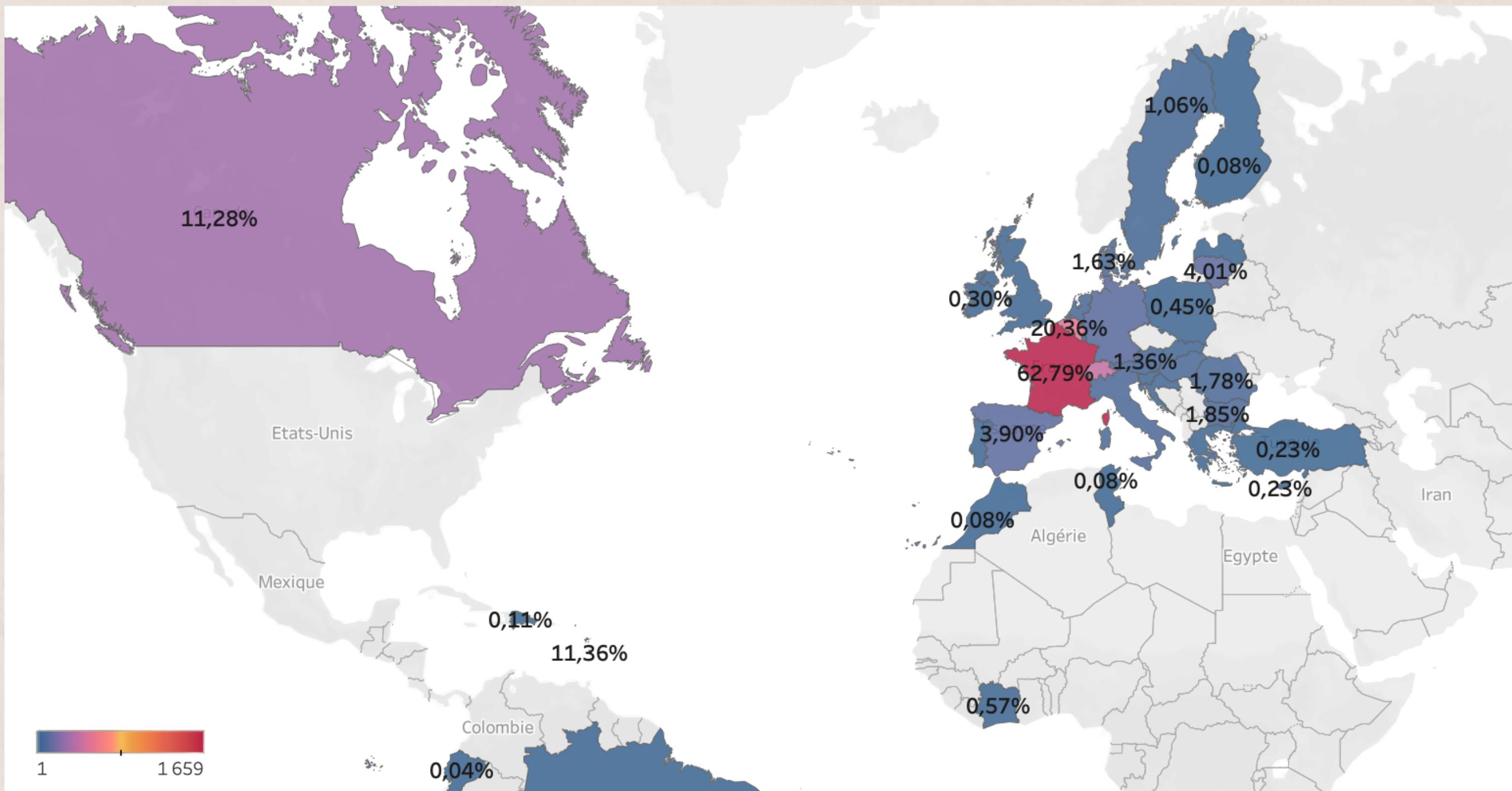


Data Exploration and Visualization

Sales variation during the year

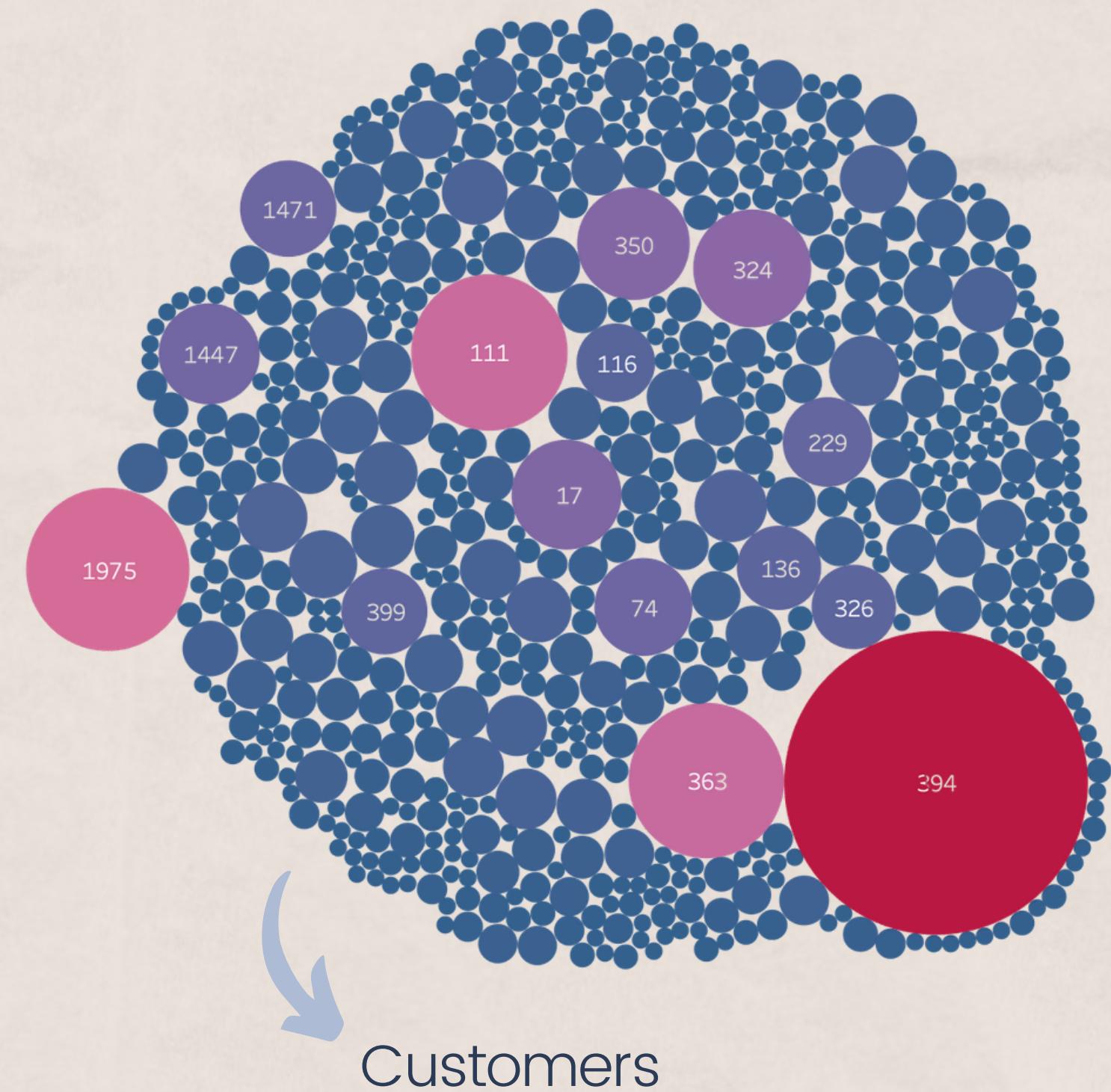


Data Exploration and Visualization



Distribution of sales as a percentage of the total for each country

Data Exploration and Visualization

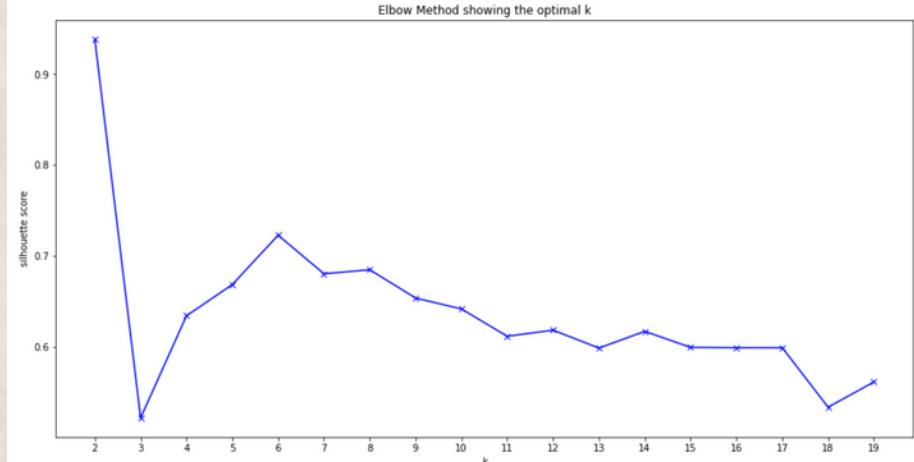


Customer Segmentation via Python

country
id_customer
17 France
22 France
30 France
31 France
34 France
...
2307 Espagne
2310 Croatie

STEP 1.

Prepare the dataset and get all numerical columns



STEP 2.

Using K-means, choose the number of K clusters

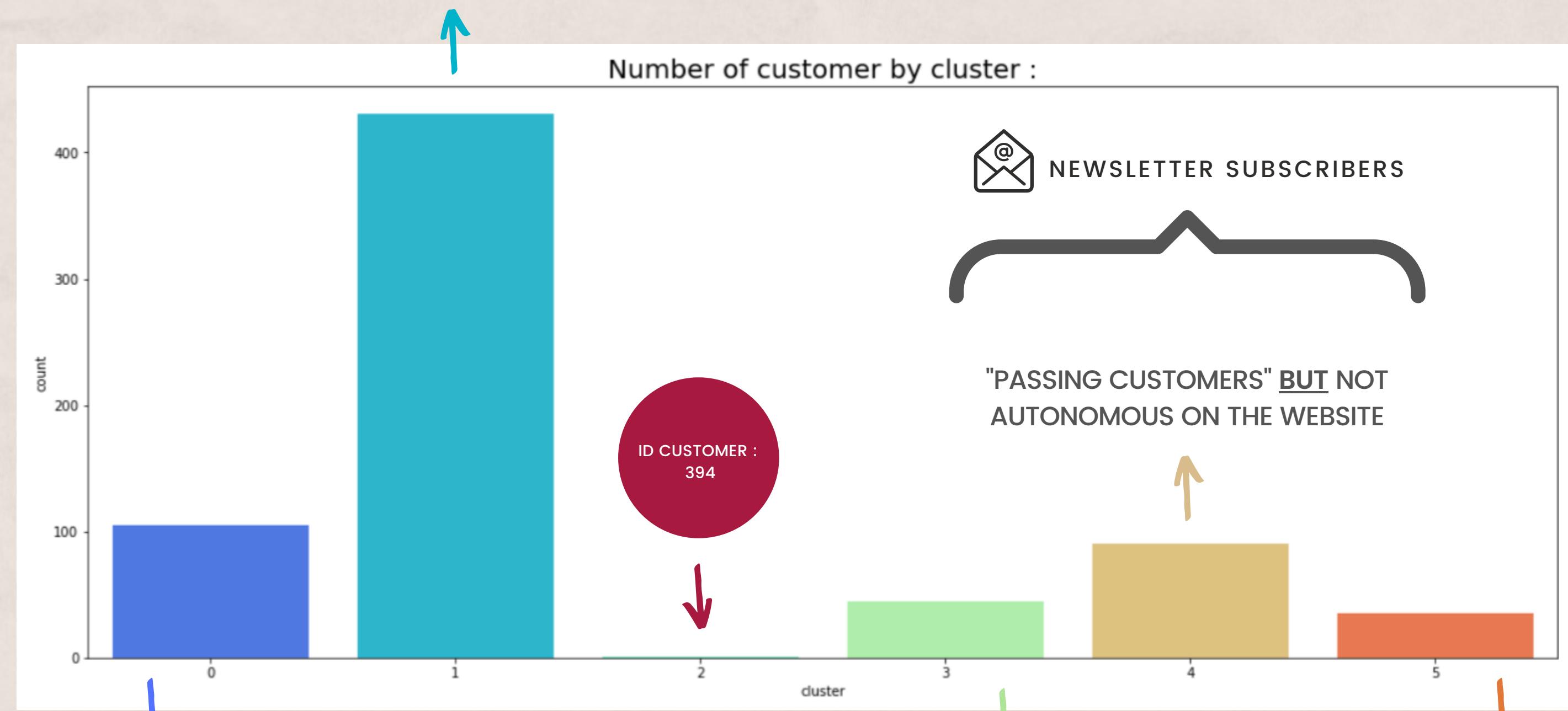
```
1: clusters = kmeans.predict(X_prep_2)
2: pd.Series(clusters).value_counts().sort_index()
0    105
1    431
2     1
3     45
4     90
5     35
dtype: int64
```

STEP 3.

Assign all the points to the closest cluster

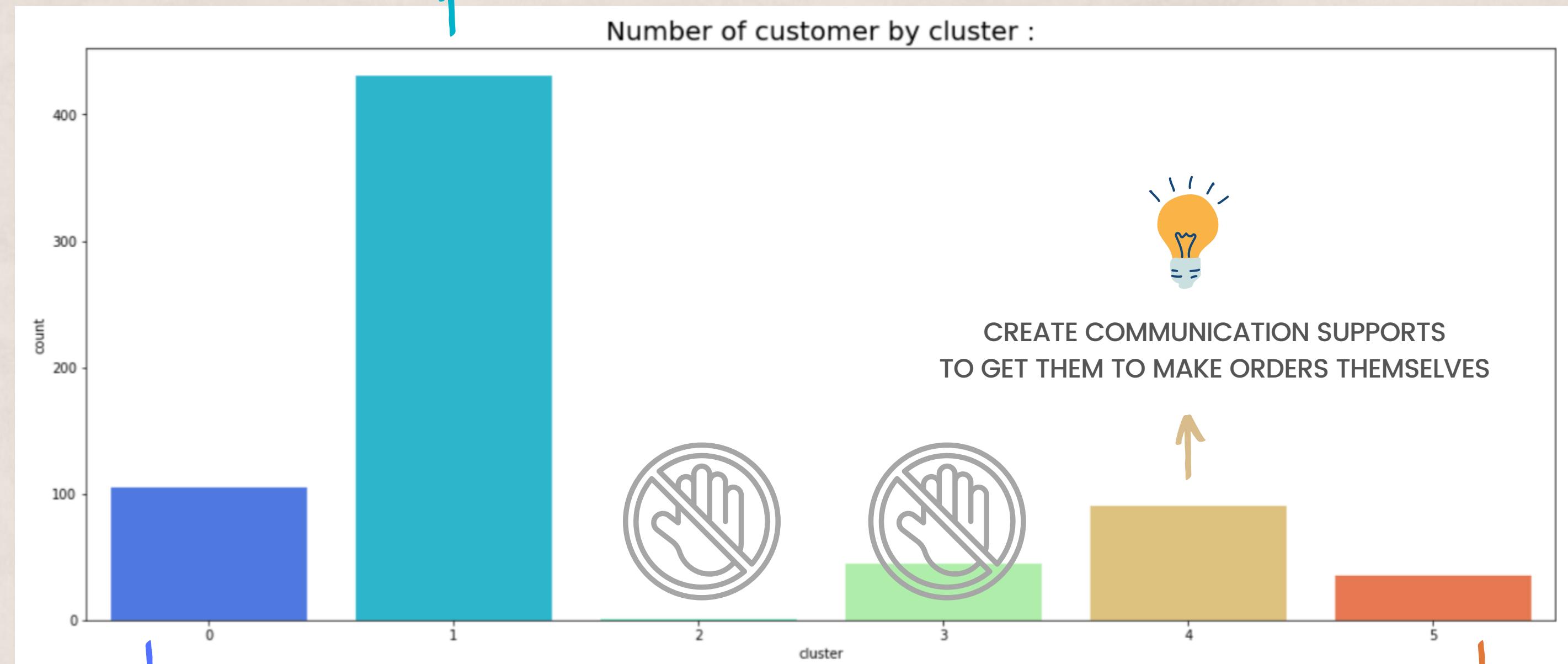
Clusters Review

GOOD CLIENTS BUT NOT
AUTONOMOUS ON THE WEBSITE



Clusters Application in Marketing

OBTAİN FEEDBACK ABOUT WEBSITE
NAVIGATION AND KNOW HOW IT COULD BE
IMPROVED



RETAIN (MAKE A CALL, HAVE MORE INFORMATION ABOUT THEM,...)

REWARD (SALES, PROMOTIONS,...)



Thank you !