Predicting the right location for a new supermarket in Berlin, Germany

Predict the proper location to set up a supermarket is advantageous for stakeholders

- The supermarket business is very lucrative in the United States.
  - In 2017, shoppers made about 1.5 trips to supermarket on average every week.
- With a pandemic that began around March 2020 due to the COVID-19.
  - The statistics showed that the grocery sales boosted.
  - The spending amid pandemics should be conservative.
- Aim: To determine the strategic location, where the supermarket is easily accessible.
- This analysis could be useful for a group of consultants, their business specialists, and the company stakeholders.
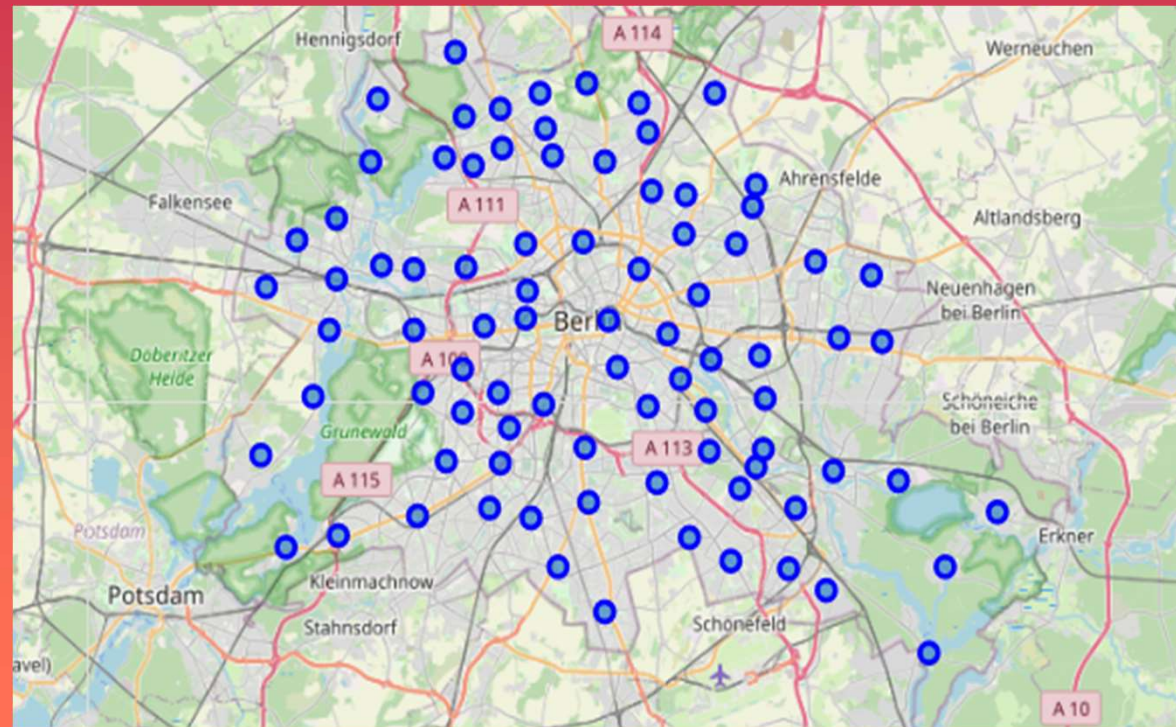
.

Data acquisition

- The data and other relevant information that we need:
    - List of the boroughs and neighborhoods of Berlin
    - Latitudes and longitudes of the boroughs in Berlin through geocoder
    - Top venues of neighborhoods through Foursquare API
- A list of boroughs and neighborhoods will be obtained from Wikipedia.
- All 97 neighborhoods will be used in the experiment.

Methodology

- Using unsupervised machine learning techniques
  - k-mean clustering algorithm to search for clusters of neighborhoods
  - The silhouette score applied for searching the optimal number of clusters
- Latitude and longitude identified for each neighborhood through Geocoder library
- Folium library used to visualize where the neighborhoods are situated in Berlin.
- The venues was collected for each neighborhood via Foursquare API
  - Preparation for cluster labelling

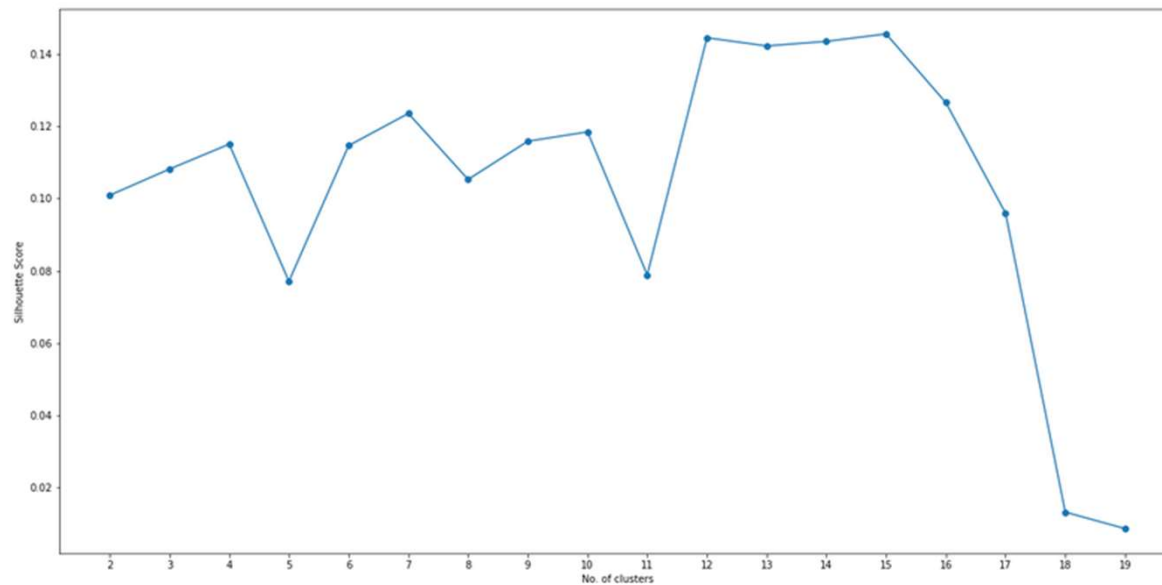# The location of all 97 neighborhoods in Berlin, Germany

First 4 neighborhoods with the top ten most common venues

- **For each neighborhood, 100 venues found using the coordinates via Foursquare API**
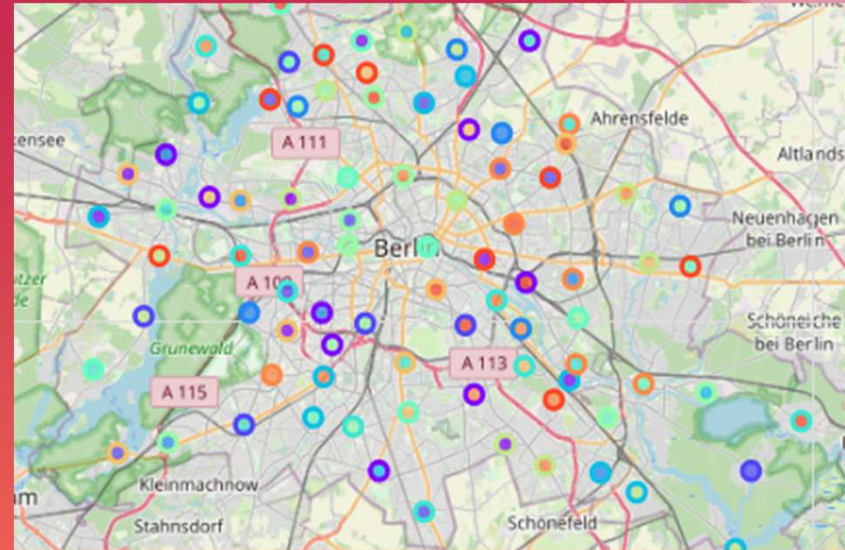- **Function created to sort the top ten most common venues for each neighborhood**

Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
Adlershof	Supermarket	Pizza Place	Italian Restaurant	Steakhouse	Shopping Mall	Greek Restaurant	Trattoria/Osteria	Drugstore	Diner	Dessert Shop
Alt-Hohenschönhausen	Tram Station	Indian Restaurant	Greek Restaurant	Big Box Store	Supermarket	Discount Store	Post Office	Drugstore	Falafel Restaurant	Farmers Market
Alt-Treptow	Italian Restaurant	Bakery	Playground	Tapas Restaurant	Garden Center	Boat or Ferry	Electronics Store	Mexican Restaurant	Juice Bar	Drugstore
Altglienicke	Home Service	Yoga Studio	Ethiopian Restaurant	Food Court	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market	Fish & Chips Shop	Fast Food Restaurant
Baumschulenweg	Supermarket	Ice Cream Shop	Asian Restaurant	Event Space	Fountain	Food Court	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market

The optimal number of clusters to be used is 12 as Input for k-mean clustering

- **When the algorithm finished computing, the cluster labels added to the dataset**

The overall result after appending the cluster labels for each neighborhood

	Postal Code	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	101	Mitte	Mitte	52.517885	13.404060	1.0	German Restaurant	Museum	Hotel	History Museum	Café
1	102	Mitte	Moabit	52.530102	13.342542	1.0	Café	German Restaurant	Burger Joint	Gym / Fitness Center	Hostel
2	103	Mitte	Hansaviertel	52.519123	13.341872	1.0	Café	Art Museum	Pedestrian Plaza	Irish Pub	Metro Station
3	104	Mitte	Tiergarten	50.340922	6.956329	1.0	Hotel	Rental Car Location	Trail	Restaurant	Eastern European Restaurant
4	105	Mitte	Wedding	52.550123	13.341970	7.0	Supermarket	Pharmacy	Ice Cream Shop	Tennis Court	Track
5	106	Mitte	Gesundbrunnen	52.550920	13.384846	1.0	Turkish Restaurant	Drugstore	Supermarket	Clothing Store	Hotel

Understanding the clusters

	Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
14	307	0.0	Asian Restaurant	Turkish Restaurant	Tram Station	Yoga Studio	Ethiopian Restaurant	Food Court	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market
70	1001	0.0	Asian Restaurant	Windmill	German Restaurant	Supermarket	Tram Station	Fish & Chips Shop	Falafel Restaurant	Farmers Market	Fast Food Restaurant	Yoga Studio
83	1101	0.0	Tram Station	Indian Restaurant	Greek Restaurant	Big Box Store	Supermarket	Discount Store	Post Office	Drugstore	Falafel Restaurant	Farmers Market

Cluster 1

	Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	101	1.0	German Restaurant	Museum	Hotel	History Museum	Café	Exhibit	Restaurant	Fountain	Plaza	Drugstore
1	102	1.0	Café	German Restaurant	Burger Joint	Gym / Fitness Center	Hostel	Bar	Doner Restaurant	Hotel	Clothing Store	Drugstore
2	103	1.0	Café	Art Museum	Pedestrian Plaza	Irish Pub	Metro Station	Bus Stop	Mediterranean Restaurant	Farmers Market	Liquor Store	Bistro
3	104	1.0	Hotel	Rental Car Location	Trail	Restaurant	Eastern European Restaurant	Steakhouse	Fish Market	Fish & Chips Shop	Fast Food Restaurant	Electronics Store
5	106	1.0	Turkish Restaurant	Drugstore	Supermarket	Clothing Store	Hotel	Bookstore	Bar	Trail	Garden	Big Box Store
6	201	1.0	Coffee Shop	Bookstore	Bagel Shop	Pub	Doner Restaurant	Café	Middle Eastern Restaurant	Hostel	Vegetarian / Vegan Restaurant	Plaza
7	202	1.0	Café	Italian Restaurant	Bakery	Coffee Shop	Waterfront	German Restaurant	Turkish Restaurant	Bar	Nightclub	Kurdish Restaurant
8	301	1.0	Café	Beer Bar	Cocktail Bar	Park	Organic Grocery	Coffee Shop	Bakery	Falafel Restaurant	Concert Hall	Candy Store
9	302	1.0	German Restaurant	Hotel	Tram Station	Memorial Site	Playground	Vietnamese Restaurant	Beach	Park	Exhibit	Falafel Restaurant
10	303	1.0	Hotel	German Restaurant	Plaza	Supermarket	Park	Café	Yoga Studio	Farmers Market	Exhibit	Falafel Restaurant
15	308	1.0	Café	Miscellaneous Shop	Automotive Shop	Event Space	Fountain	Food Court	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market

Cluster 2

- **The cluster 1 is outskirt where retail market is not really represented**
- **Cluster 2 is the biggest cluster with a number of 57 neighborhoods.**
  - **Lots of commercial activities related venues.**

Understanding the clusters

Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
79	1101	2.0	Photography Studio	Grocery Store	Yoga Studio	Food Court	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market	Fish & Chips Shop	Fast Food Restaurant

Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
34	507	3.0	Playground	Yoga Studio	Ethiopian Restaurant	Food Court	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market	Fish & Chips Shop	Fast Food Restaurant

Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
67	913	4.0	Hotel	Yoga Studio	Ethiopian Restaurant	Food Court	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market	Fish & Chips Shop	Fast Food Restaurant

Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
67	913	4.0	Hotel	Yoga Studio	Ethiopian Restaurant	Food Court	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market	Fish & Chips Shop	Fast Food Restaurant

Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
33	506	8.0	Bakery	Flower Shop	Ethiopian Restaurant	Fountain	Food Court	Food & Drink Shop	Fishing Store	Fish Market	Fish & Chips Shop	Fast Food Restaurant

Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
60	906	11.0	Home Service	Yoga Studio	Ethiopian Restaurant	Food Court	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market	Fish & Chips Shop	Fast Food Restaurant

Cluster 3

Cluster 4

Cluster 5

Cluster 6

Cluster 9

Cluster 12

- These cluster contain only one neighborhood each.
- These are locations with minor commercial area which surrounded by neighborhood.

10

Understanding the clusters

	Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
13	306	6.0	Playground	Bus Stop	Restaurant	Yoga Studio	Ethiopian Restaurant	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market	Fish & Chips Shop
31	504	6.0	Bus Stop	Supermarket	Restaurant	Eastern European Restaurant	Yoga Studio	Event Space	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market
56	902	6.0	Bus Stop	Supermarket	Light Rail Station	Yoga Studio	French Restaurant	Food Court	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market
88	1201	6.0	Bus Stop	Supermarket	Italian Restaurant	Restaurant	Yoga Studio	Ethiopian Restaurant	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market

Cluster 7

	Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
75	1101	9.0	Hostel	Historic Site	Yoga Studio	Ethiopian Restaurant	Food Court	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market	Fish & Chips Shop
78	1101	9.0	Hostel	Historic Site	Yoga Studio	Ethiopian Restaurant	Food Court	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market	Fish & Chips Shop
86	1201	9.0	Hostel	Insurance Office	Nature Preserve	Liquor Store	Yoga Studio	Event Space	Food Court	Food & Drink Shop	Flower Shop	Fishing Store

Cluster 10

	Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
49	706	10.0	Soccer Field	Supermarket	Bus Stop	Doner Restaurant	Miscellaneous Shop	Yoga Studio	Exhibit	Food Court	Food & Drink Shop	Flower Shop
81	1101	10.0	Indian Restaurant	Supermarket	Soccer Field	German Restaurant	Fountain	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market	Fish & Chips Shop

Cluster 11

- Cluster 7 is neighborhood where public travel rated top common venue.
  - These are mainly areas with family houses and apartments.
- Cluster 10 with three neighborhoods.
  - Hostel and historic site at the top. It is possibly a tourist attraction region.
- Cluster 11 contains two neighborhoods. Areas with minor commercial area which surrounded by neighborhood.

Understanding the clusters

	Postal Code	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
4	105	7.0	Supermarket	Pharmacy	Ice Cream Shop	Tennis Court	Track	Big Box Store	Gas Station	Park	Farmers Market	Exhibit
11	304	7.0	Supermarket	Fried Chicken Joint	Chinese Restaurant	Fish Market	Tram Station	Food Court	Food & Drink Shop	Flower Shop	Fishing Store	Fish & Chips Shop
12	305	7.0	Supermarket	Bus Stop	Skating Rink	Yoga Studio	Event Space	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market	Fish & Chips Shop
17	310	7.0	Supermarket	Garden Center	Drugstore	Organic Grocery	Yoga Studio	Event Space	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market
35	508	7.0	Supermarket	Snack Place	Drugstore	Yoga Studio	Ethiopian Restaurant	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market	Fish & Chips Shop
39	603	7.0	Supermarket	Drugstore	Movie Theater	Fast Food Restaurant	Park	German Restaurant	Bakery	Yoga Studio	Exhibit	Food & Drink Shop
43	607	7.0	Supermarket	Harbor / Marina	Indian Restaurant	Chinese Restaurant	Bank	Bakery	Austrian Restaurant	Liquor Store	Post Office	Farmers Market
47	704	7.0	Supermarket	Ice Cream Shop	Park	Steakhouse	Chinese Restaurant	Greek Restaurant	Diner	Exhibit	Food & Drink Shop	Dance Studio
53	804	7.0	Supermarket	Ice Cream Shop	Italian Restaurant	Drugstore	Bank	Diner	Falafel Restaurant	Fountain	Food Court	Deli / Bodega
57	903	7.0	Supermarket	Ice Cream Shop	Asian Restaurant	Event Space	Fountain	Food Court	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market
61	907	7.0	Supermarket	Pizza Place	Italian Restaurant	Steakhouse	Shopping Mall	Greek Restaurant	Trattoria/Osteria	Drugstore	Diner	Dessert Shop
62	908	7.0	Miscellaneous Shop	Supermarket	Insurance Office	Italian Restaurant	Event Space	Fountain	Food Court	Food & Drink Shop	Flower Shop	Fishing Store
73	1004	7.0	Supermarket	Hotel	Plaza	Park	Light Rail Station	Ethiopian Restaurant	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market
80	1101	7.0	Supermarket	Café	Italian Restaurant	German Restaurant	Seafood Restaurant	Yoga Studio	Event Space	Food & Drink Shop	Flower Shop	Fishing Store
82	1101	7.0	Supermarket	Movie Theater	Gym / Fitness Center	Train Station	Tram Station	Yoga Studio	Falafel Restaurant	Event Space	Exhibit	Fast Food Restaurant
84	1101	7.0	Supermarket	Italian Restaurant	Sushi Restaurant	Fishing Store	Drugstore	Construction & Landscaping	Café	Plaza	Park	Doner Restaurant
92	1201	7.0	Supermarket	Drugstore	Gym / Fitness Center	Italian Restaurant	Light Rail Station	Grocery Store	Fast Food Restaurant	Hardware Store	Shopping Mall	Warehouse Store
96	1201	7.0	Motorcycle Shop	Supermarket	Go Kart Track	Bakery	Yoga Studio	Event Space	Food & Drink Shop	Flower Shop	Fishing Store	Fish Market

Cluster 8

- Cluster 8 is neighborhood
  - a variety of supermarket established in the cluster and plenty of food related venues.
  - There is also lots of commercial activities going around

12

Discussion and limitation

- The potential location for the new supermarket is suggested as below:
  - Clusters 1 and 7 for a family-friendly supermarket that suitable for families to do grocery shopping.
  - Cluster 10 for the supermarket that targets visitors and tourists.
  - Cluster 2 and 8 for a supermarket with luxury facilities and high-class ingredients .

- Four of the neighborhoods were not successfully assigned with cluster labels due to some unknown technical issue with Kmeans function.
- The number of collected venues is limited to 100 per neighborhoods.
- A 1000-meter radius is used to search venues around the center coordinates of the neighborhoods. via Foursquare API.

Conclusion

- The process of finding an answer for the simulated business problem is discussed.
- For clustering 97 neighborhoods in Berlin, a machine learning technique was used.
- The experiment was performed based on the data science tools and the use of Python libraries.
- The outcome of the results provided insight for drafting several recommendations to the mentioned business problem.

Thank you for your attention