



Home

Team

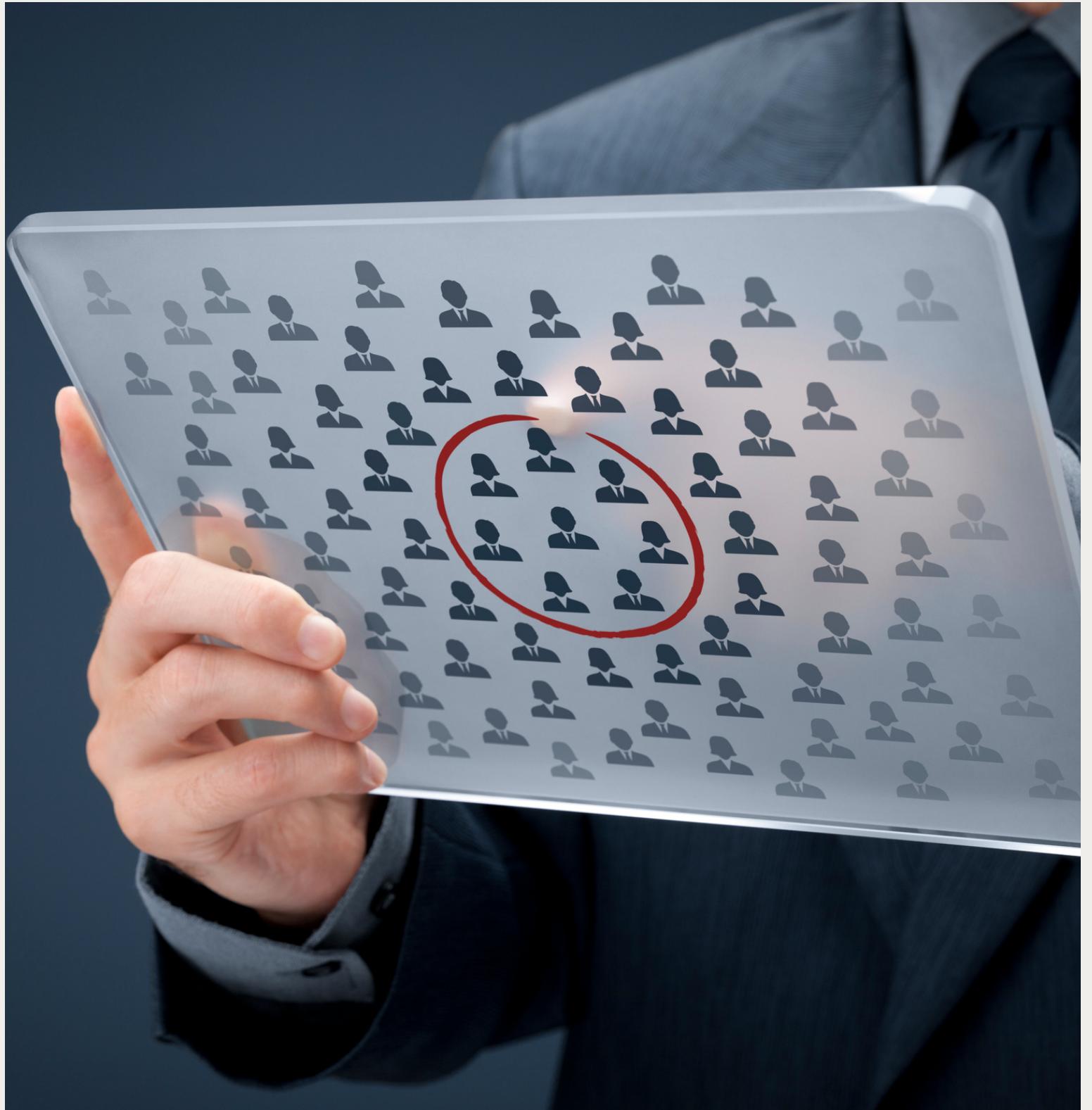
Workflow

Analysis



Customer Segmentation

By Group 3





Home

Team

Workflow

Analysis



Our Team



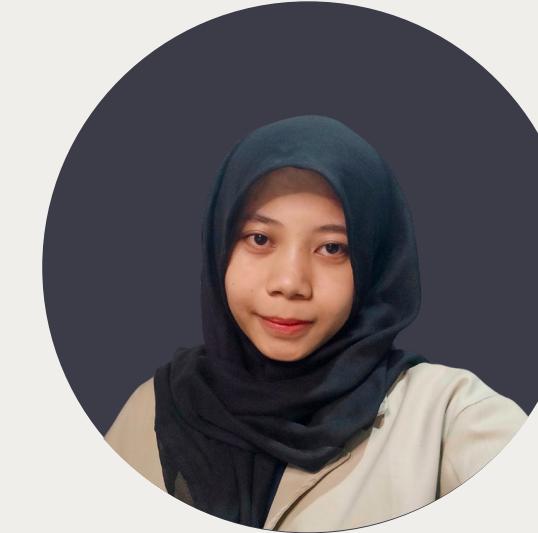
Agung Prayogi

Universitas Pembangunan Nasional
Veteran Yogyakarta



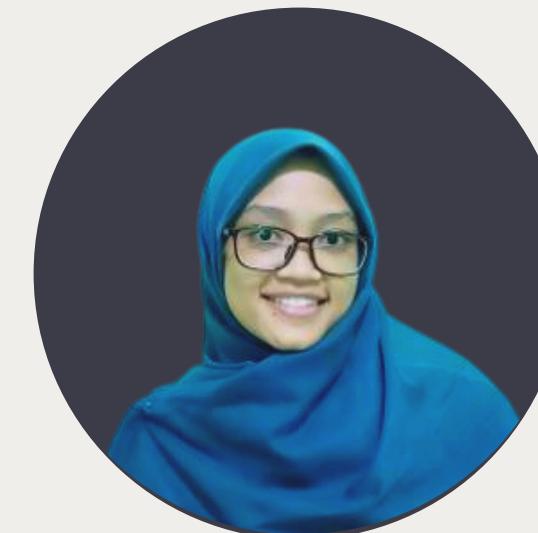
Axel Eldrian H.

Universitas Pendidikan
Indonesia



Dian Maulida

Universitas Diponegoro



Dwi Fitriawati F.

Universitas Pendidikan
Indonesia



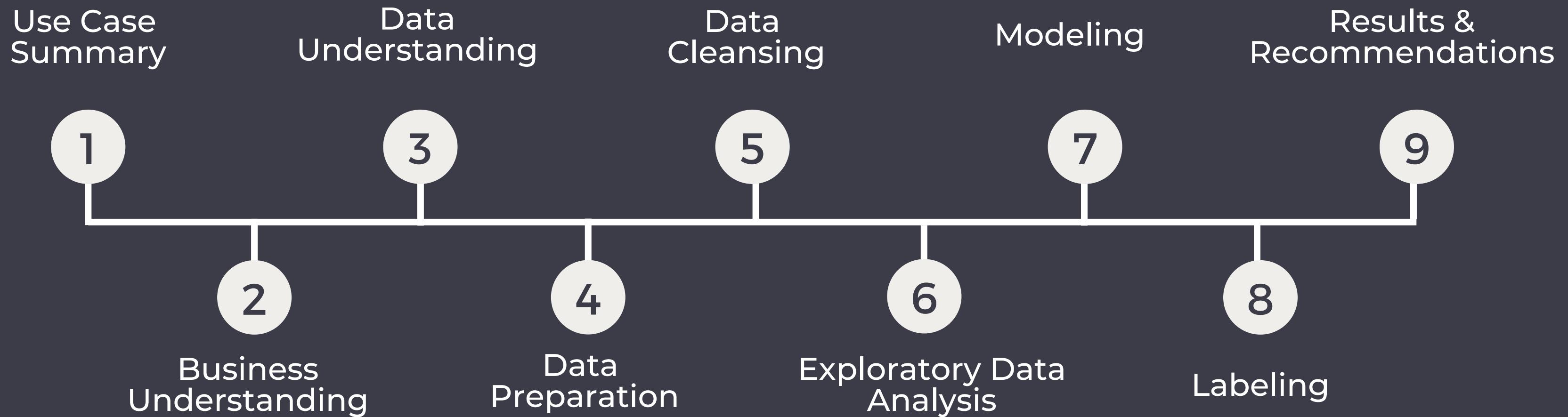
Yesaya Arya D. K.

Universitas Merdeka Malang



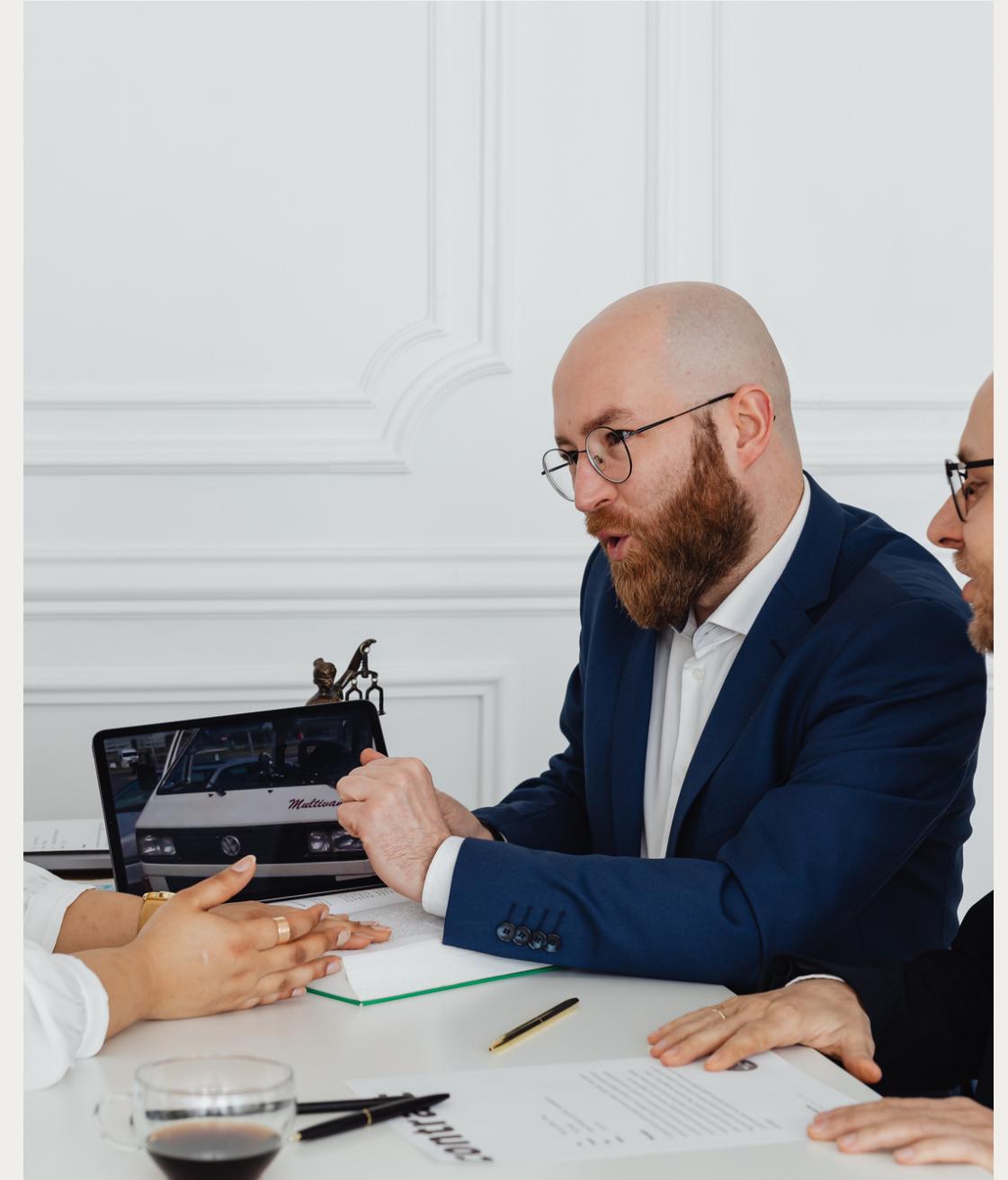
[Home](#)[Team](#)[Workflow](#)[Analysis](#)

Workflow



[Home](#)[Team](#)[Workflow](#)[Analysis](#)

Use Case Summary



[Home](#)[Team](#)[Workflow](#)[Analysis](#)

Objective Statement



- Get insight about how much customer spend every month on every year
- Get insight about how much customer spend every month on 2011
- Get insight about how much customer spend every month on 2012
- Get insight about how much customer spend every month on 2013
- Get insight about how much customer spend every month on 2014
- Get insight about how big they spend on a transaction
- Get insight about the ratio between customer segment
- Get insight about trend on every year
- Get business insight about how, when and whom to promote business on the right time
- Get insight about marketing strategy or planning based on segmentation

[Home](#)[Team](#)[Workflow](#)[Analysis](#)

Challenges

- There is data that has an inappropriate data type



Methodology/Analytic Technique

- Descriptive analysis
- Graph analysis
- RFM (Recency, Frequency, Monetary) Model

Business Benefit

- Help Business to make decision about marketing strategy based on customer segmentation
- Help business to focus what customer need

[Home](#)[Team](#)[Workflow](#)[Analysis](#)

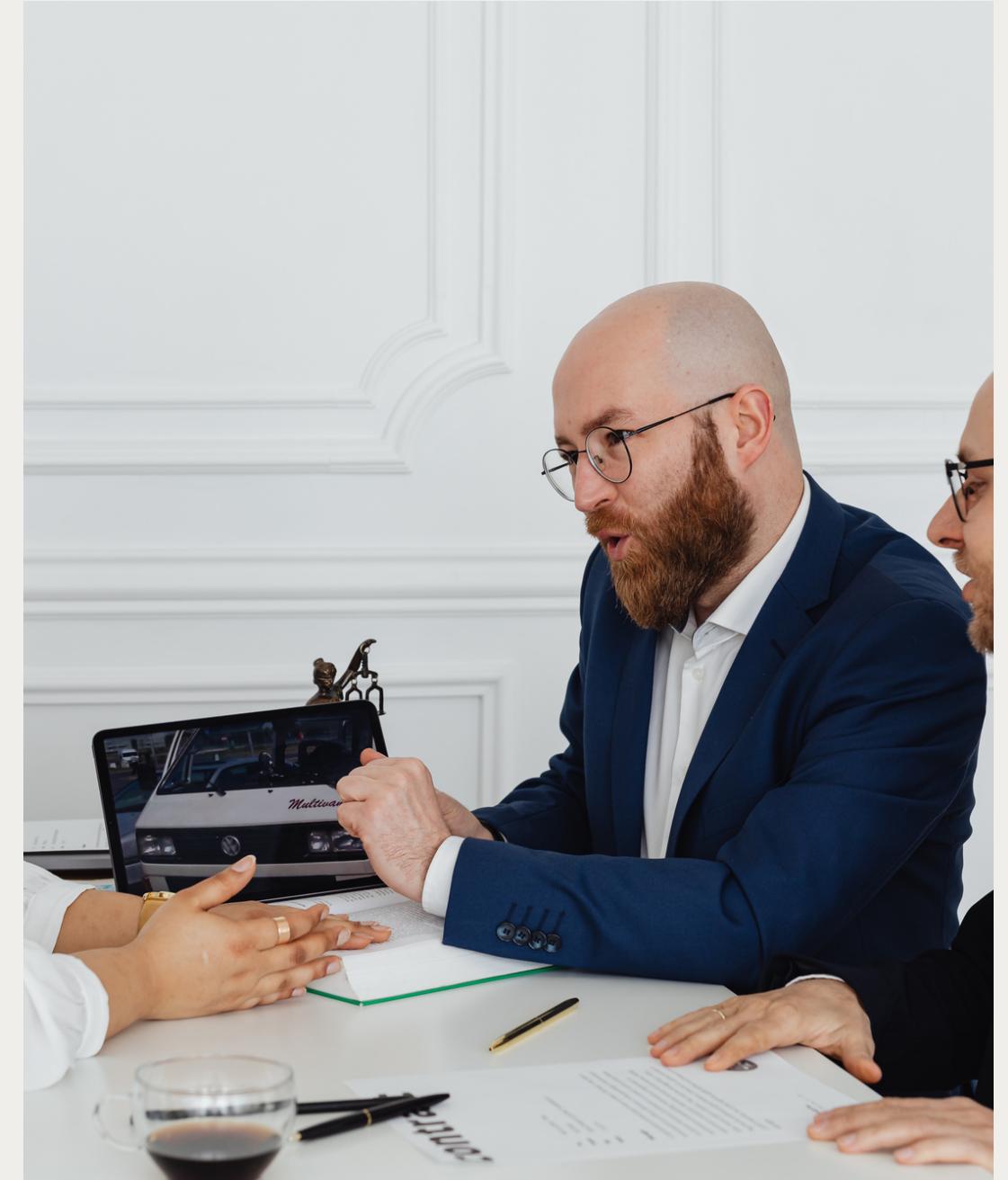
Expected Outcome



- Get to know how many customer spend every month on every year on our business
- Get to know how many customer spend every month on 2011
- Get to know how many customer spend every month on 2012
- Get to know how many customer spend every month on 2013
- Get to know how many customer spend every month on 2014
- Get to know how big they spend on a transaction
- Get to know the ratio between customer segment
- Get to know trend on every year
- Get to know when, how, and for whom the product will be promote
- Recommendation based on customer segmentation

[Home](#)[Team](#)[Workflow](#)[Analysis](#)

Business Understanding



[Home](#)[Team](#)[Workflow](#)[Analysis](#)

Business Understanding



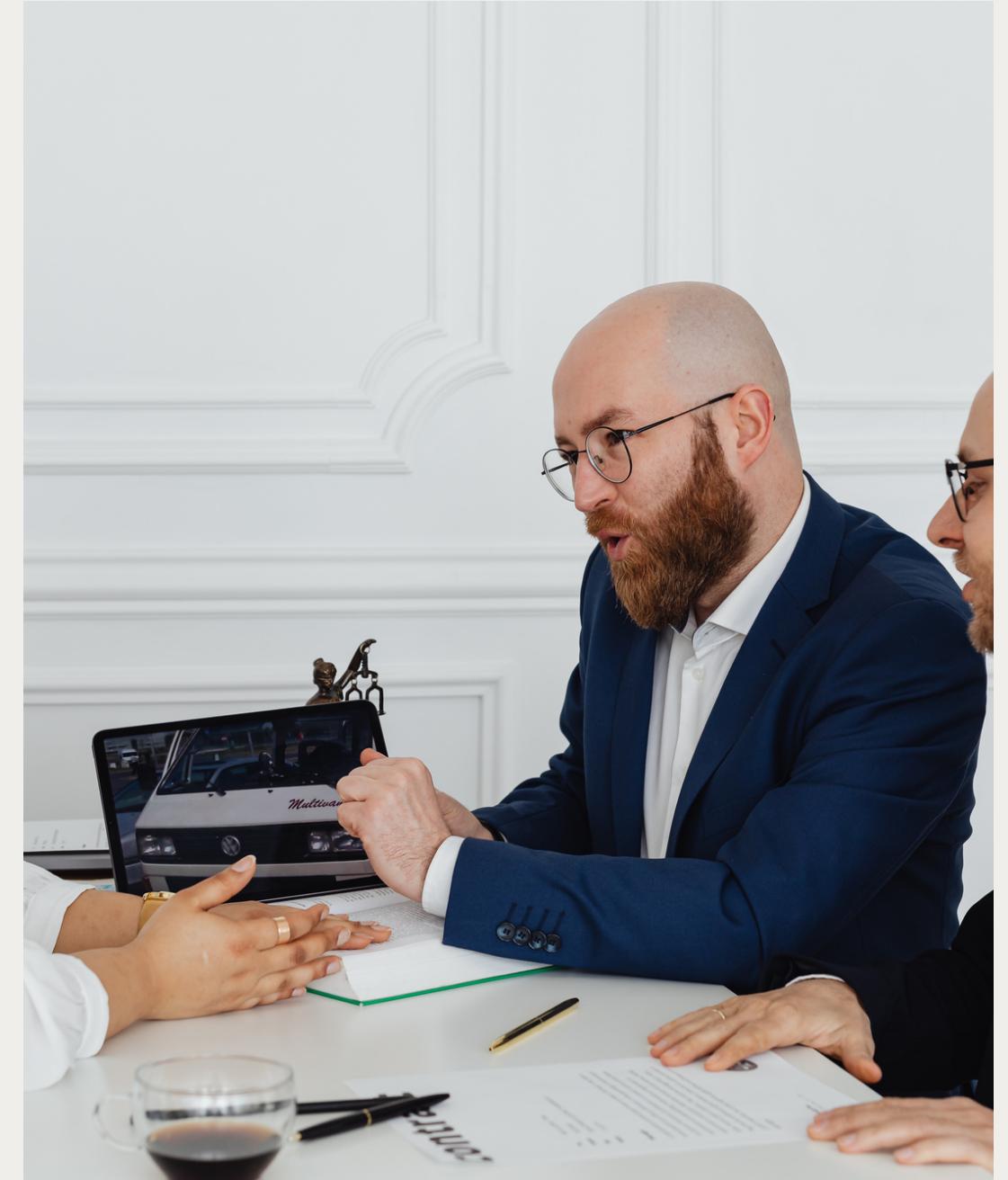
Online retail or as known as online shopping refers to a form of electronic commerce where consumers directly buy goods or services from a seller via the Internet rather. Customers usually visit the company's website, select a product, enter payment information and shipping details, and then order the product. All of these steps are completed entirely online.

This case has some business question using the data :

- How many customer spend every month on every year?
- How big they are spend on a transaction?
- How about Customer segmentation based on customer spend?
- How about recommendation based customer segmentation?

[Home](#)[Team](#)[Workflow](#)[Analysis](#)

Data Understanding



[Home](#)[Team](#)[Workflow](#)[Analysis](#)

Source Data



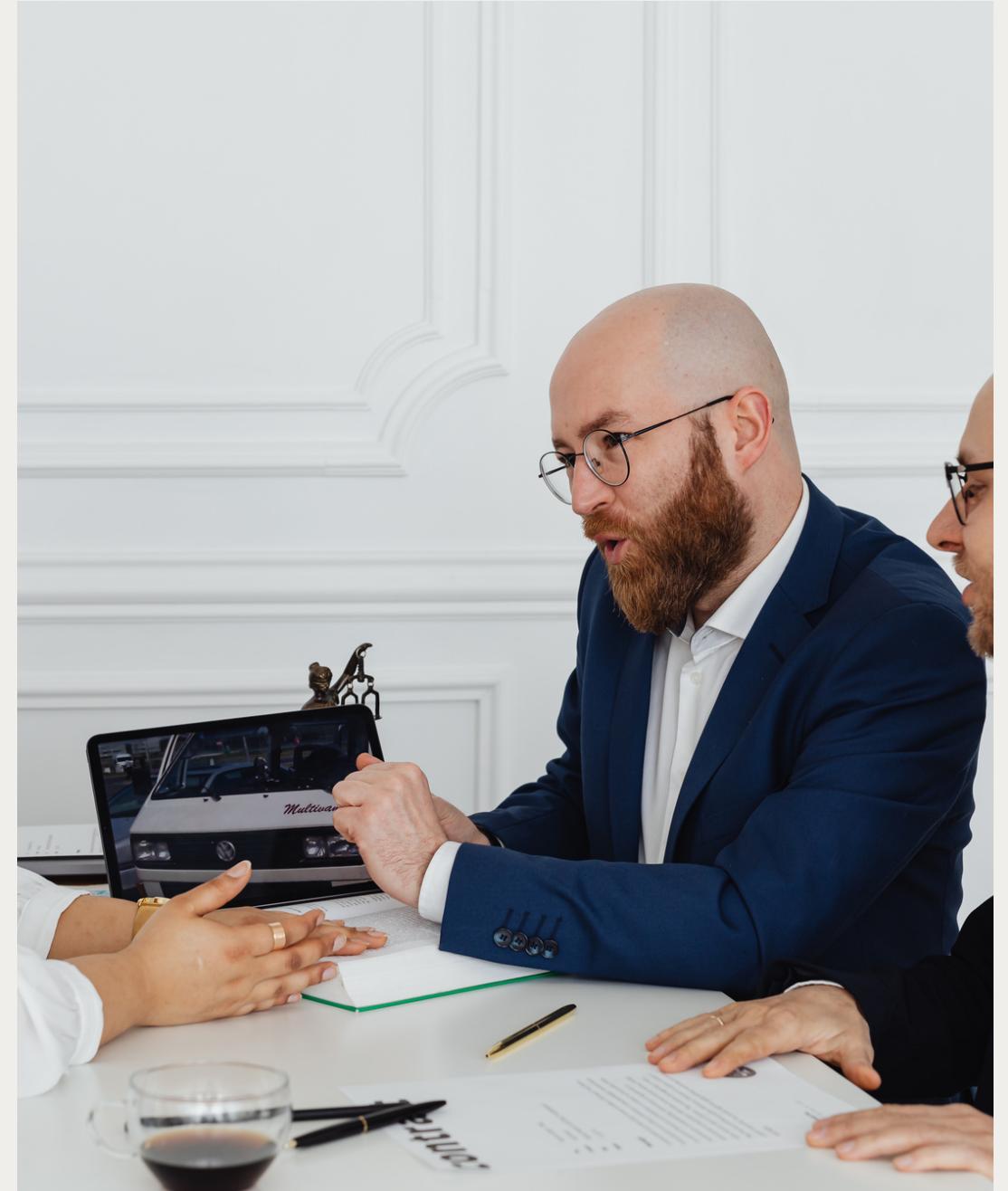
- Order data is a record of customer purchase data from January 4, 2011 to December 31, 2014
- Source Data : <https://www.kaggle.com/datasets/siddinho/sample-orders-dataset-retail>
- The dataset has 4 columns and 5009 rows

Data Dictionary

- `order_date` : Date when customer order
- `order_id` : ID customer order for each transaction
- `customer` : Name of Customer
- `grand_total` : total transaction for each customer

[Home](#)[Team](#)[Workflow](#)[Analysis](#)

Data Preparation



[Home](#)[Team](#)[Workflow](#)[Analysis](#)

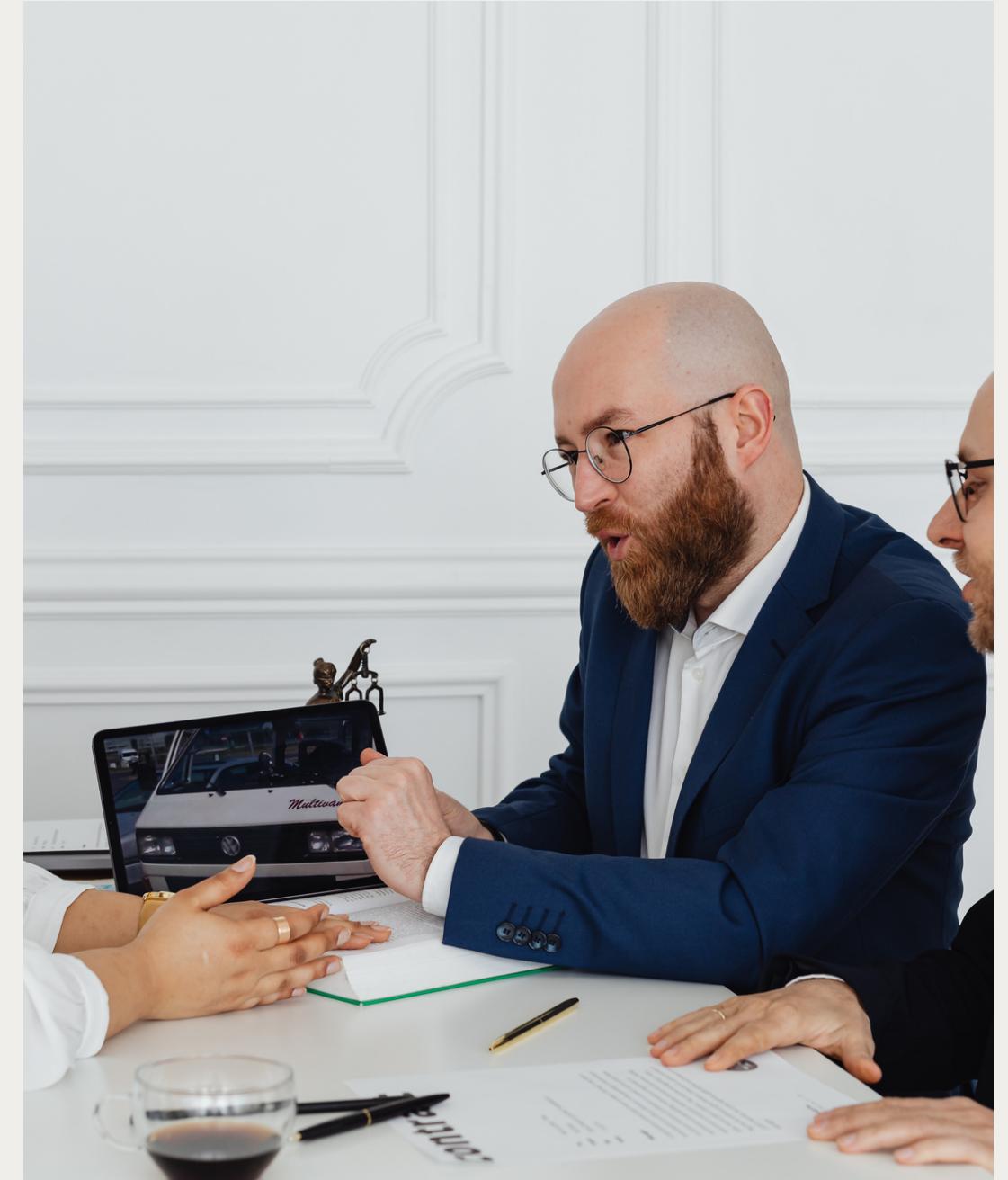
Data Preparation

Code Used:

- Python Version: 3.7.15
- Packages: Pandas, Numpy, Matplotlib, Seaborn, Datetime, Sklearn, Scipy, Plotly and Feature Engine

[Home](#)[Team](#)[Workflow](#)[Analysis](#)

Data Cleansing



[Home](#)[Team](#)[Workflow](#)[Analysis](#)

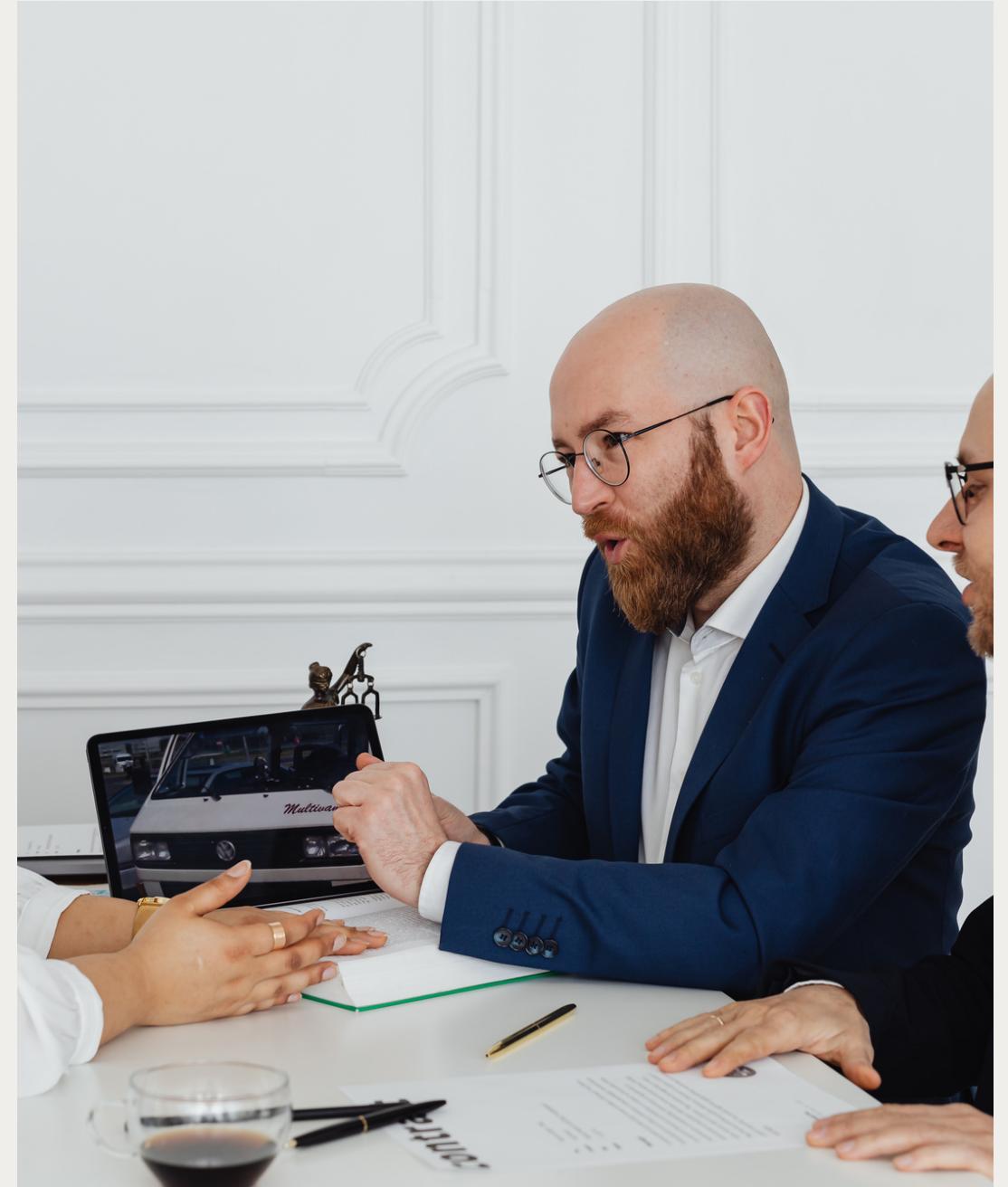
Data Cleansing

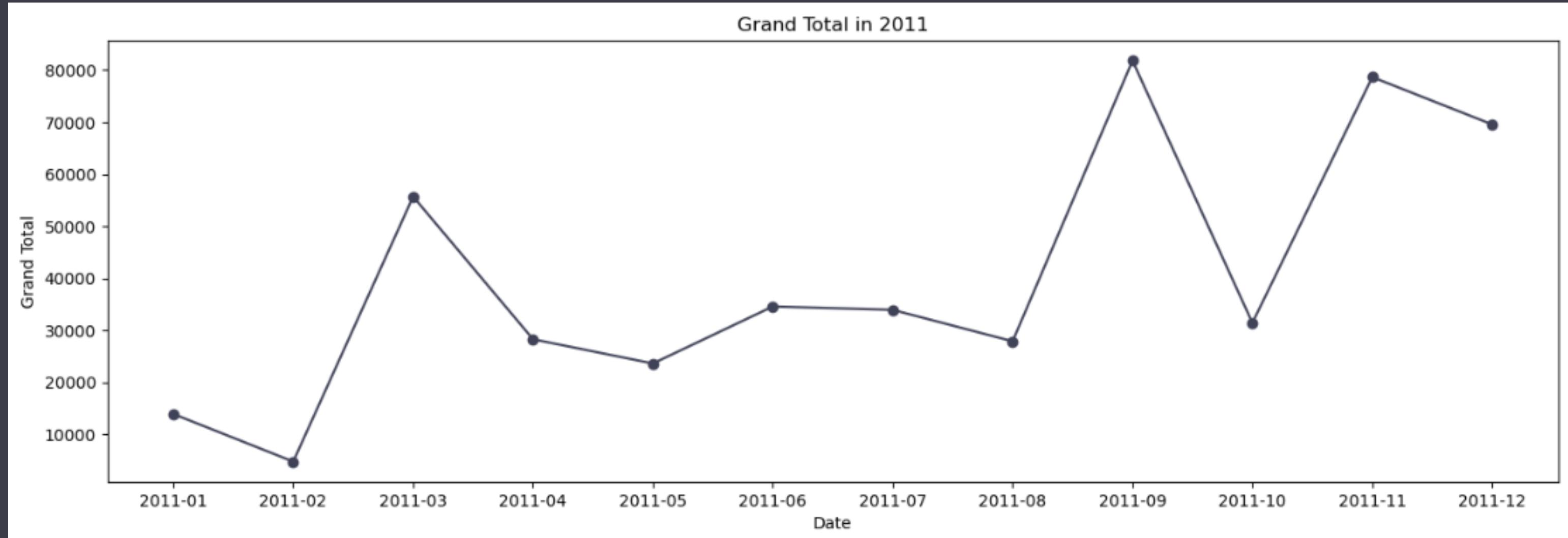
There is 1 column that has an inappropriate data type so it is necessary to adjust the data type.



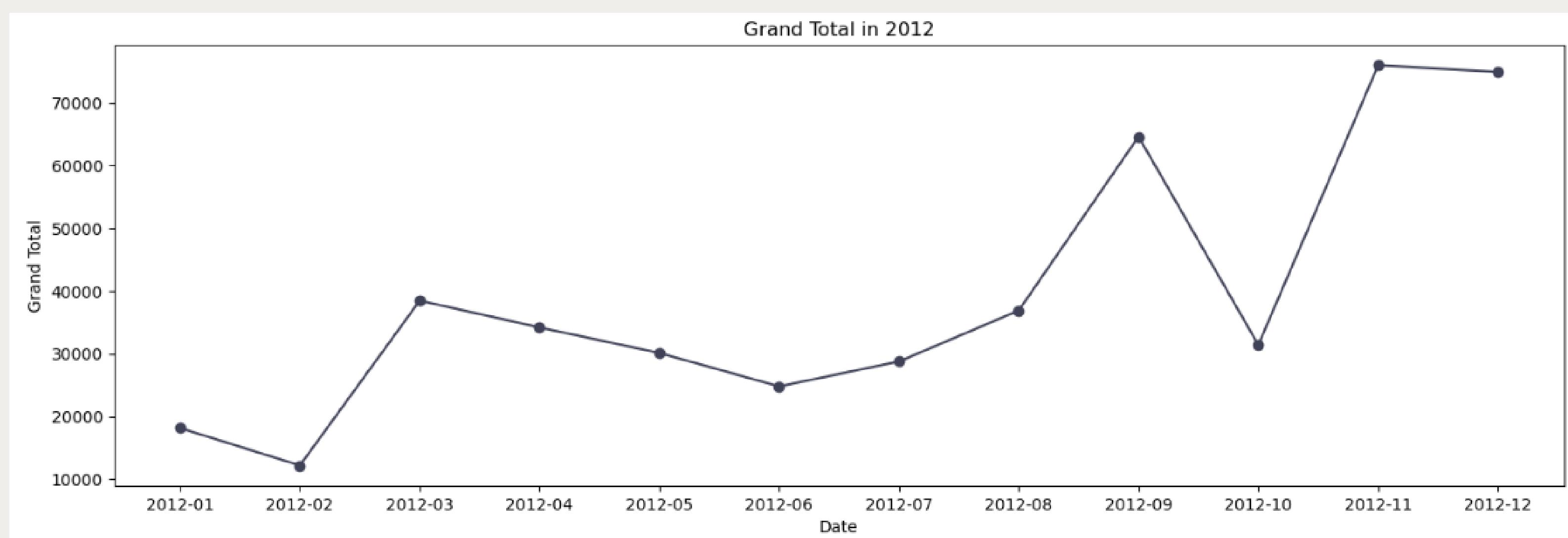
[Home](#)[Team](#)[Workflow](#)[Analysis](#)

Exploratory Data Analysis

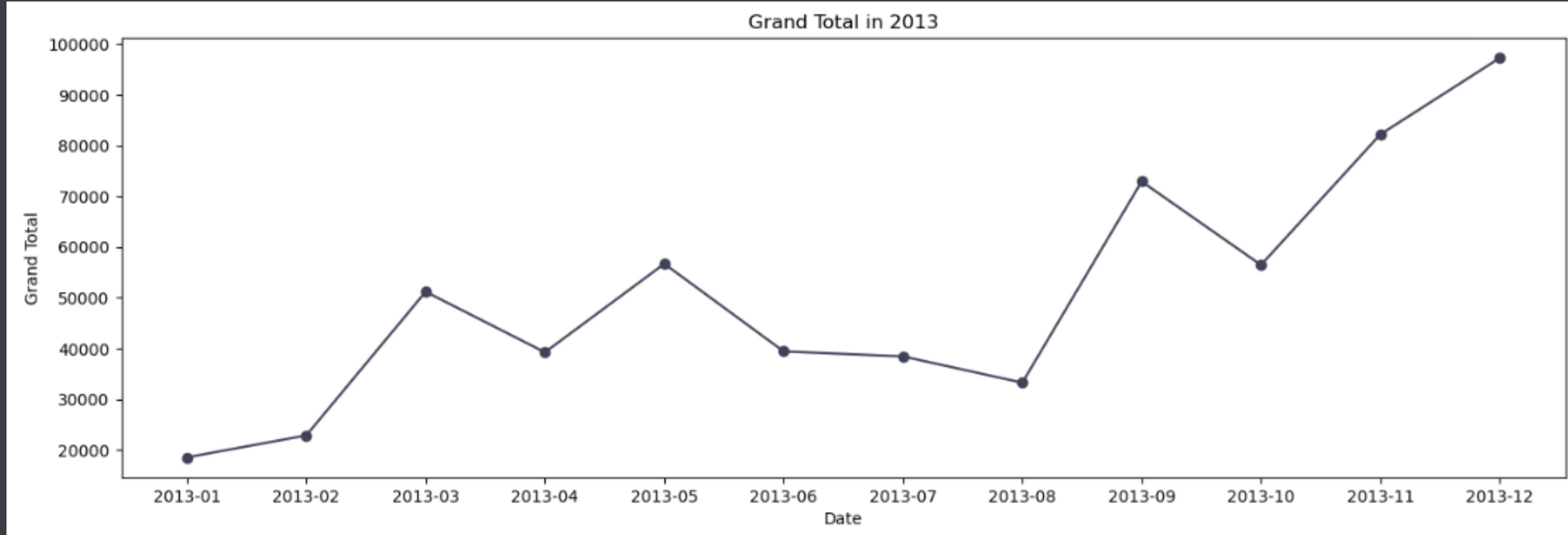


[Home](#)[Team](#)[Workflow](#)[Analysis](#)

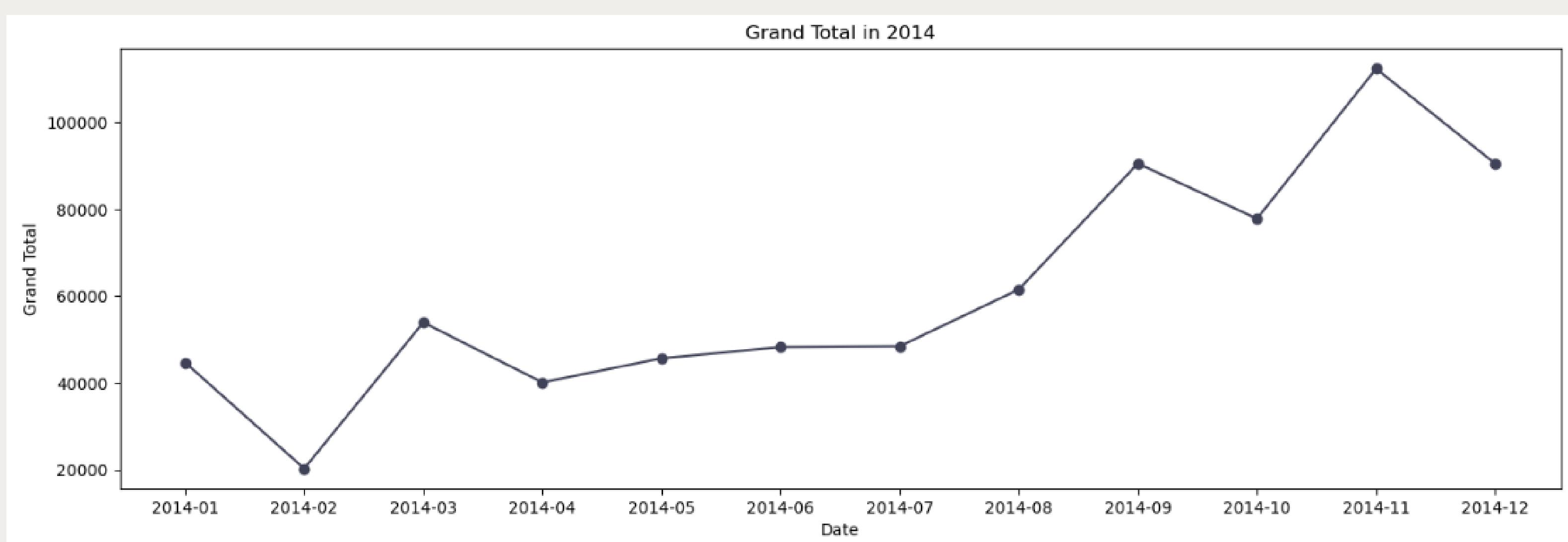
Based on the graph above, it can be seen that in 2011, the trend of total customer purchases increased, although in 2011 there was a significant decrease in the second month which reached the lowest point of 4810, and reached the peak/all time high 2011 in the 9th month of the year which reached 81784. On an annual basis, **Customer spend increased by 398.58%**

[Home](#)[Team](#)[Workflow](#)[Analysis](#)

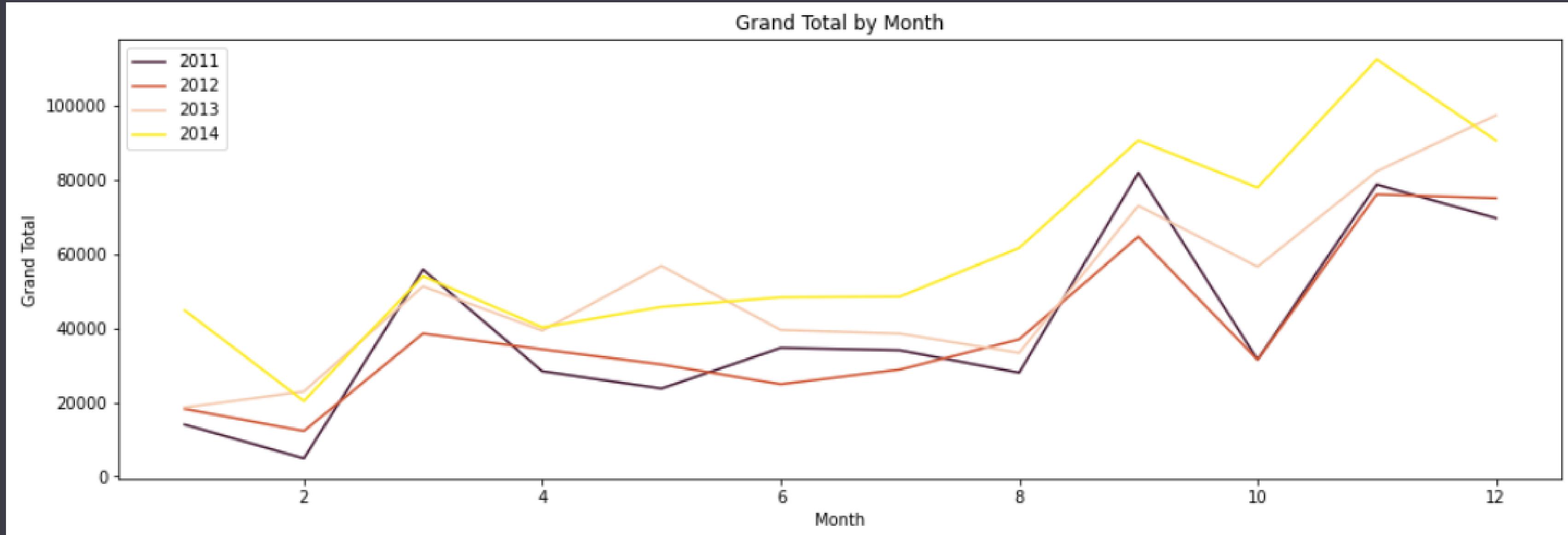
Based on the graph above, it can be seen that in 2012, the trend of total customer purchases increased, although in 2012 there was a significant decrease in the second month which reached the lowest point of 12211, and reached the peak/all time high 2012 in the 11th month of the year which reached 75968. On an annual basis, **Customer spend increased by 312.23%**

[Home](#)[Team](#)[Workflow](#)[Analysis](#)

Based on the graph above, it can be seen that in 2013, the trend of total customer purchases increased, although in 2013 there was a significant decrease in the first month which reached the lowest point of 18543, and reached the peak/all time high 2013 in the last month of the year which reached 97244. On an annual basis, **Customer spend increased by 424.42%**

[Home](#)[Team](#)[Workflow](#)[Analysis](#)

Based on the graph above, it can be seen that in 2014, the trend of total customer purchases increased, although in 2014 there was a significant decrease in the second month which reached the lowest point of 44708, and reached the peak/all time high 2014 in the 11th month of the year which reached 112329. On an annual basis, **Customer spend increased by 102.38%**

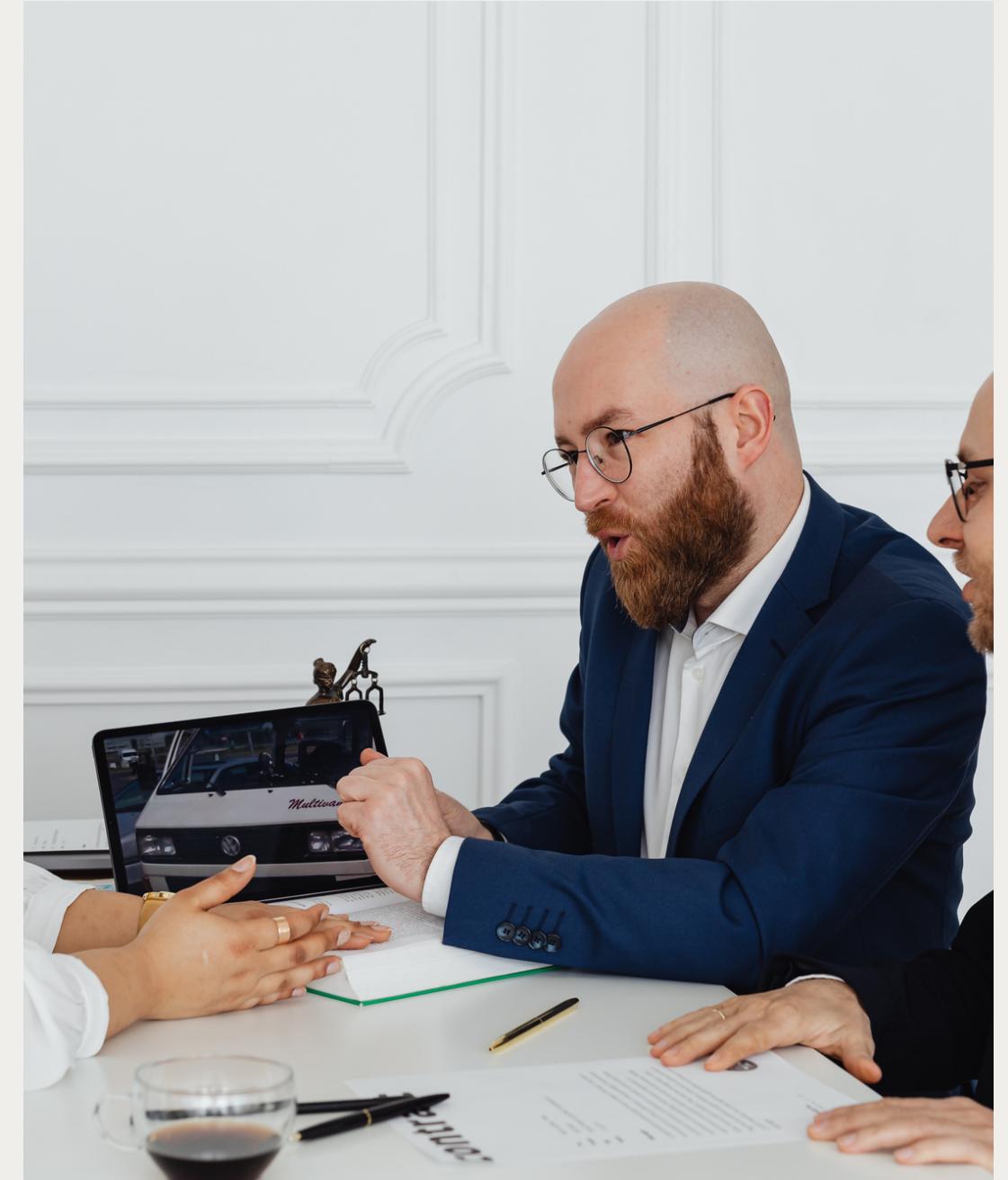
[Home](#)[Team](#)[Workflow](#)[Analysis](#)

[Home](#)[Team](#)[Workflow](#)[Analysis](#)

From the graph before, Customer purchase data 'order data' from January 2011 to 2014 is fluctuated. In 2011 the highest purchase was in September with a total purchase of 81,784 k and the lowest was of 4810 in February, in 2012 the highest purchase was in November with a total purchase of 75,968 k and the lowest was in February with a value of 12,211 k , in 2013 the highest purchase was worth 97,244 k in December and the lowest was in January with a value of 18,543 k, and in 2014 the highest purchase was in November with a total of 112.293 k and the lowest in February with a value of 20,287 k. Overall, the highest purchase in the period 2011 to 2014 was in November 2014 with a value of 112.293 k and the lowest was in February 2011 with a total purchase of 4810. From the data above, there is also a visible pattern where in the three-year period the lowest purchases are generally at the beginning of the year, between January and February, while the highest purchases are at the end of the year between September and December.

[Home](#)[Team](#)[Workflow](#)[Analysis](#)

Modeling



[Home](#)[Team](#)[Workflow](#)[Analysis](#)

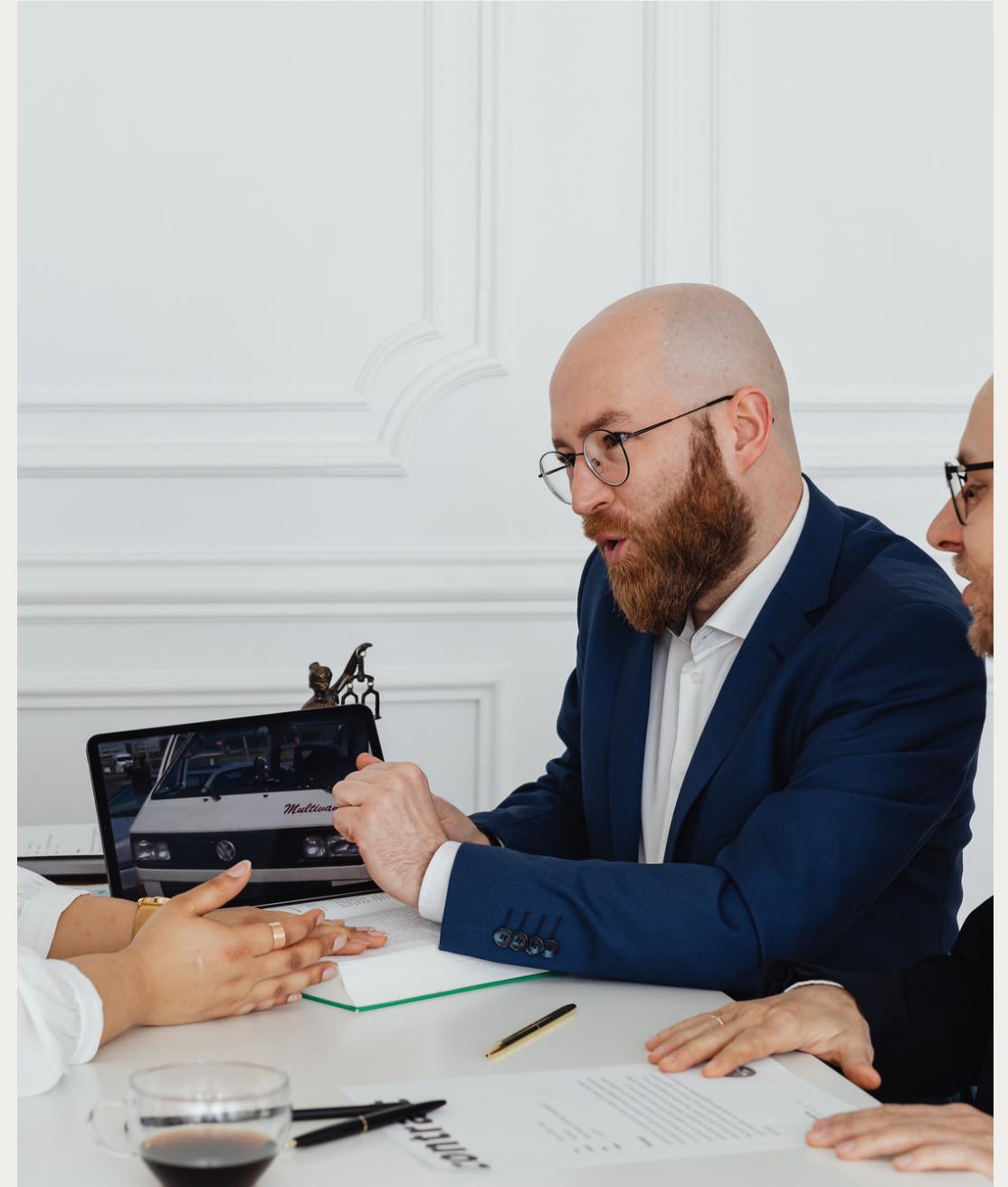
	customer	Recency	Frequency	Monetary	R_quartile	F_quartile	M_quartile	RFM_Segment	RFM_Score
0	Aaron Bergman	415	3	887	4	4	4	444	12
1	Aaron Hawkins	12	7	1744	1	2	3	123	6
2	Aaron Smayling	88	7	3050	3	2	2	322	7
3	Adam Bellavance	54	8	7756	2	2	1	221	5
4	Adam Hart	34	10	3249	2	1	2	212	5

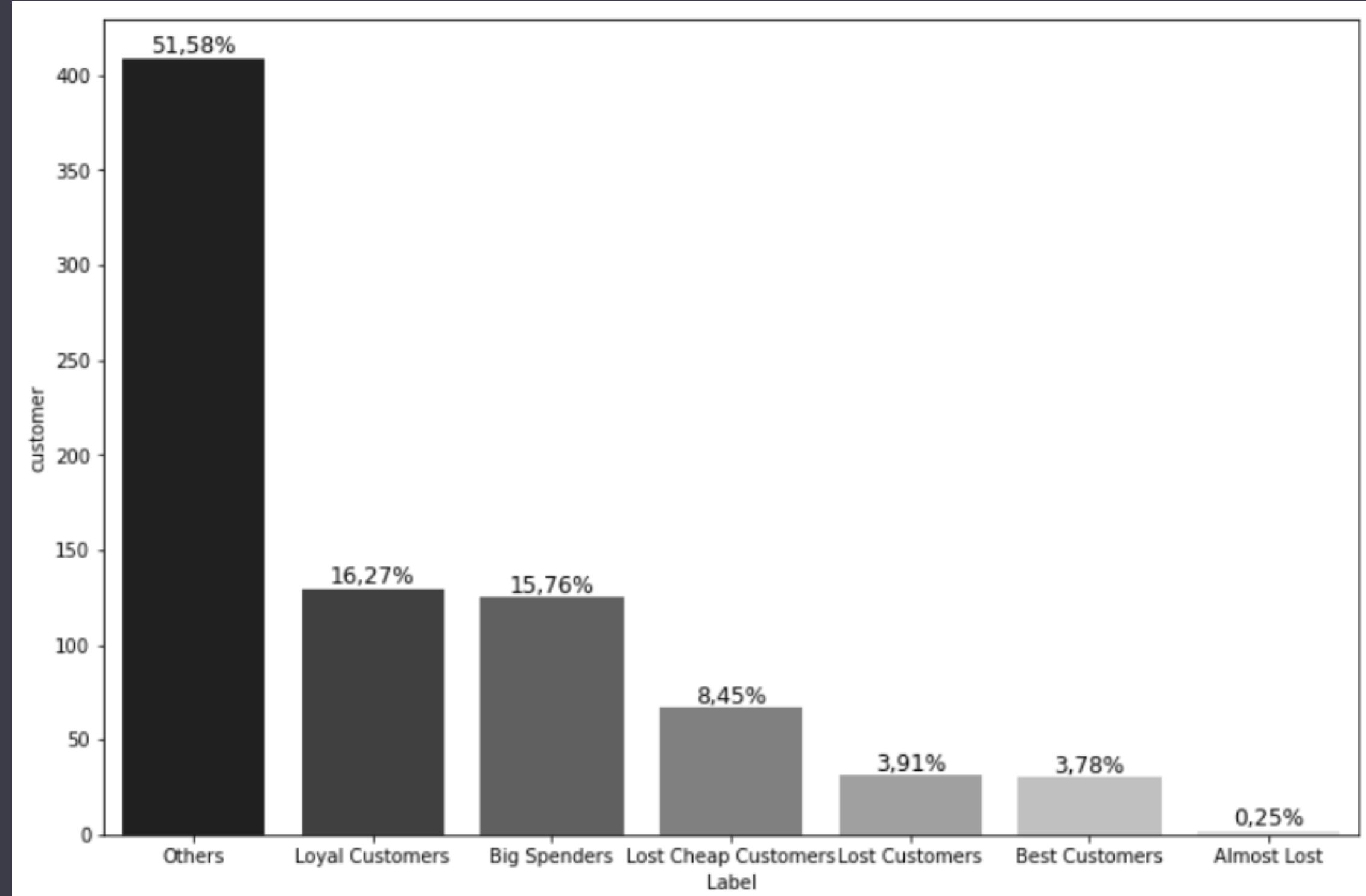


In this RFM Model assigns a score of 1-4 (from best to worst) for customers in each of the three categories. The best RFM score is the one with the smallest values for each variable, we use a 1 to 4 scale for recency, frequency, and monetary, with 1 being the highest, then the perfect RFM score is 111. Customer who get score 1 for R_quartile means that the customer has recently made a purchase from the company. Customer who get score 1 for F_quartile means that customer make a high frequent purchase from the company. Customer who get score 1 for M_quartile means that the customer spend big money for purchase from the company.

[Home](#)[Team](#)[Workflow](#)[Analysis](#)

Labeling



[Home](#)[Team](#)[Workflow](#)[Analysis](#)

[Home](#)[Team](#)[Workflow](#)[Analysis](#)

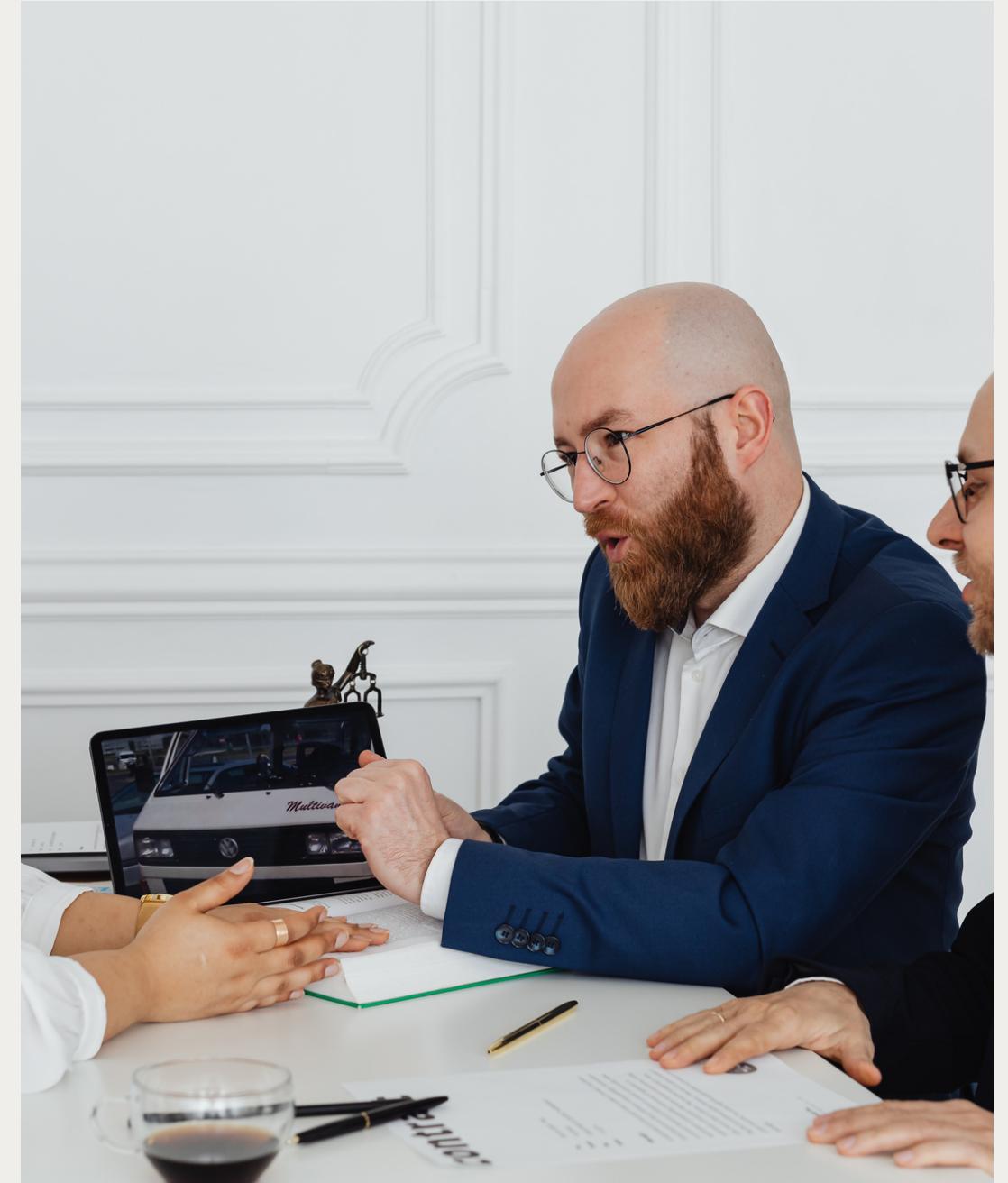
From the graph above, it is known that there are 6 categories of customers. The most customer category is Loyal Customer and the least is Almost lost customer.

- **Loyal customers** are customers who have frequent purchases ($F=1$) The percentage of loyal customers is 16.27% of all customer.
- **Big Spenders** are customers who make purchases with a large amount of price ($M = 1$), the percentage of Big Spenders customers is 15.76% of all customer.
- **Lost Cheap Customers** are customers who have made a purchase for a long time ($R = 4$), rarely make purchases (low frequency/ $F= 4$) and the total price of the purchase is small ($M= 4$), the number of incoming customers in the Lost Cheap Customer category is 8.45% of all customers.
- **Best Customer** is a customer who has recently made a purchase ($R = 1$), frequently makes purchases (high frequency/ $F= 1$) and the total price of the purchase is large ($M= 1$), the number of customers who fall into the category This is 3.78% of the total customers.
- **Lost Customers** are customers who haven't made a purchase long enough ($R = 3$), rarely make purchases (low frequency/ $F= 4$) and the total price of the purchase is small ($M=4$), customers who enter in this category amounted to 3.91% of all customers.
- **Almost Lost** is a customer who has recently made a purchase ($R = 1$), rarely makes a purchase (quite low frequency/ $F= 3$) and the total price of the purchase is small ($M=4$), a customer who is included in the this category is 0.25% of all customers.



[Home](#)[Team](#)[Workflow](#)[Analysis](#)

Result



[Home](#)[Team](#)[Workflow](#)[Analysis](#)

- Based on the data above, we can see that the average customer purchases products with a total price of 458.626672, the customer with the lowest purchase is at 1.000000 and the highest purchase is at 23661.000000
- Based on the graph above, it can be seen that in 2011, the trend of total customer purchases increased, although in 2011 there was a significant decrease in the second month which reached the lowest point of 4810, and reached the peak/all time high 2011 in the 9th month of the year which reached 81784. On an annual basis, Customer spend increased by 398.58%
- Based on the graph above, it can be seen that in 2012, the trend of total customer purchases increased, although in 2012 there was a significant decrease in the second month which reached the lowest point of 12211, and reached the peak/all time high 2012 in the 11th month of the year which reached 75968. On an annual basis, Customer spend increased by 312.23%
- Based on the graph above, it can be seen that in 2013, the trend of total customer purchases increased, although in 2013 there was a significant decrease in the first month which reached the lowest point of 18543, and reached the peak/all time high 2013 in the last month of the year which reached 97244. On an annual basis, Customer spend increased by 424.42%



[Home](#)[Team](#)[Workflow](#)[Analysis](#)

- Based on the graph above, it can be seen that in 2014, the trend of total customer purchases increased, although in 2014 there was a significant decrease in the second month which reached the lowest point of 44708, and reached the peak/all time high 2014 in the 11th month of the year which reached 112329. On an annual basis, Customer spend increased by 102.38%
- Customer purchase data 'order data' from January 2011 to 2014 is fluctuated. In 2011 the highest purchase was in September with a total purchase of 81,784 k and the lowest was of 4810 in February, in 2012 the highest purchase was in November with a total purchase of 75,968 k and the lowest was in February with a value of 12,211 k , in 2013 the highest purchase was worth 97,244 k in December and the lowest was in January with a value of 18,543 k, and in 2014 the highest purchase was in November with a total of 112.293 k and the lowest in February with a value of 20,287 k. Overall, the highest purchase in the period 2011 to 2014 was in November 2014 with a value of 112.293 k and the lowest was in February 2011 with a total purchase of 4810. From the data above, there is also a visible pattern where in the three-year period the lowest purchases are generally at the beginning of the year, between January and February, while the highest purchases are at the end of the year between September and December.



[Home](#)[Team](#)[Workflow](#)[Analysis](#)

From the graph above, it is known that there are 6 categories of customers.

- The most customer category is Loyal Customer and the least is Almost lost customer.
- Loyal customers are customers who have frequent purchases ($F=1$) The percentage of loyal customers is 16.27% of all customer.
- Big Spenders are customers who make purchases with a large amount of price ($M = 1$), the percentage of Big Spenders customers is 15.76% of all customer.
- Lost Cheap Customers are customers who have made a purchase for a long time ($R = 4$), rarely make purchases (low frequency/ $F= 4$) and the total price of the purchase is small ($M= 4$), the number of incoming customers in the Lost Cheap Customer category is 8.45% of all customers.
- Best Customer is a customer who has recently made a purchase ($R = 1$), frequently makes purchases (high frequency/ $F= 1$) and the total price of the purchase is large ($M= 1$), the number of customers who fall into the category This is 3.78% of the total customers.



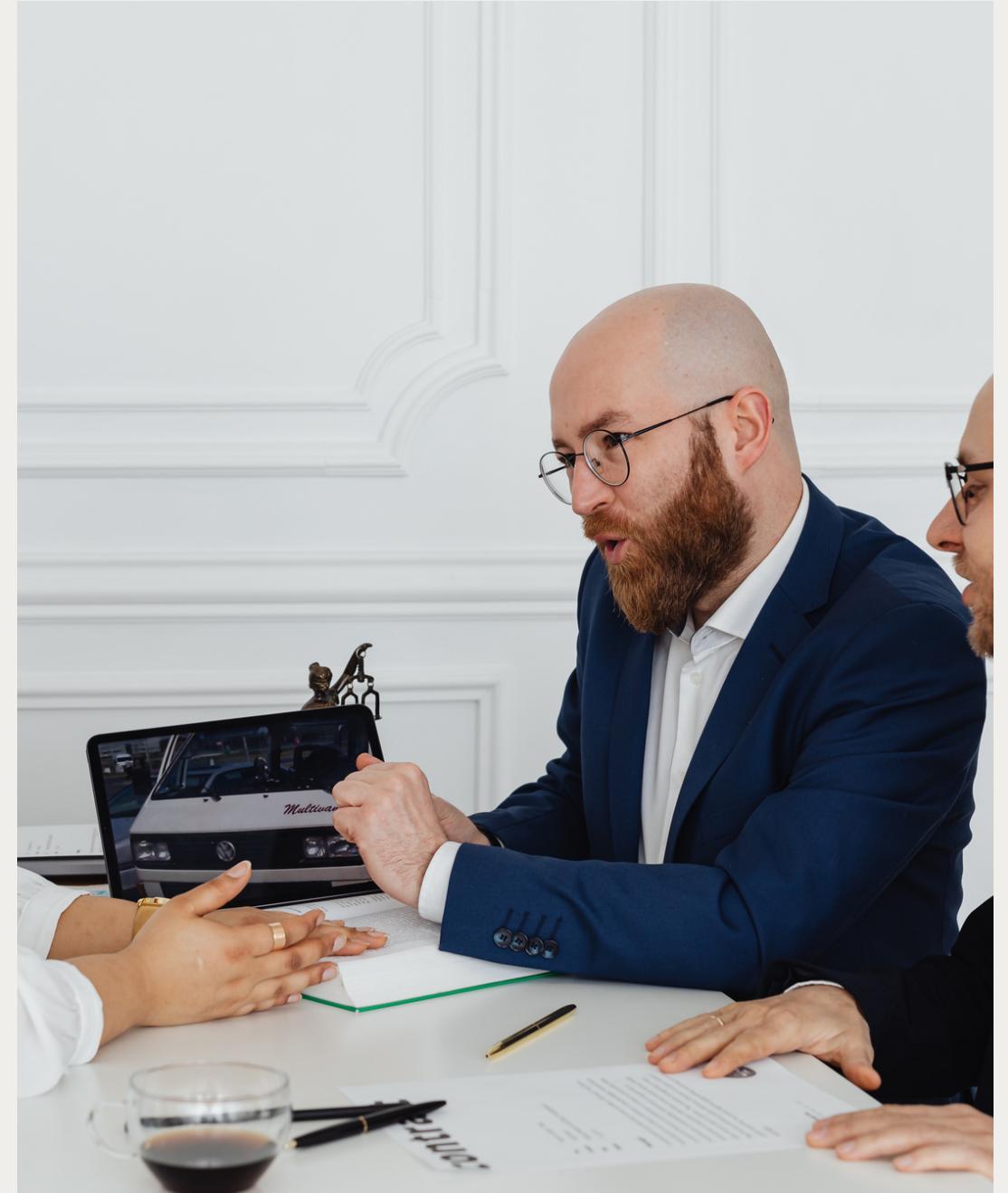
[Home](#)[Team](#)[Workflow](#)[Analysis](#)

- Lost Customers, are customers who haven't made a purchase long enough ($R = 3$), rarely make purchases (low frequency/ $F= 4$) and the total price of the purchase is small ($M=4$), customers who enter in this category amounted to 3.91% of all customers.
- Almost Lost is a customer who has recently made a purchase ($R = 1$), rarely makes a purchase (quite low frequency/ $F= 3$) and the total price of the purchase is small ($M=4$), a customer who is included in the this category is 0.25% of all customers.



[Home](#)[Team](#)[Workflow](#)[Analysis](#)

Recommendations



[Home](#)[Team](#)[Workflow](#)[Analysis](#)

- **Recommendation for "Loyal Customers" segment:** In order to keep this customer group's loyalty and raise their value, the business team must optimize the budget and timing of both campaigns.
- **Recommendation for "Big Spenders" segment:** In order to keep and boosting frequency of order, focus on give loyalty and advantages for every order
- **Recommendation for "Best Customers" segment:** Focus on boosting client purchases; as a result, develop an upselling/cross-selling strategy or make bundling strategy.
- **Recommendation for "Almost Lost" segment:** Since this customer group is more susceptible to churn, concentrate on reactivating consumers and encouraging repeat purchases by developing a reactivation strategy and retention strategy.
- **Recommendation for "Lost Cheap Customers" and "Lost Customer" segment:** The goal of the campaign is to reactivate the customer by creating a Reactivation strategy because this consumer category has churned.



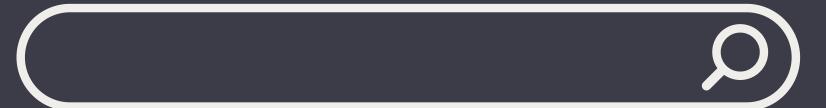


Home

Team

Workflow

Analysis



“**THANK YOU**”