



Olist Website

Syed Irteza Haseem

Olist is an Ecommerce Platform that allows sellers and businesses to sell their products on the Olist Store website



They connect small businesses from all over Brazil to sell their products through the Olist Store and ship them directly to the customers

Problem statement



Our aim is to provide meaningful and useful insights into the two databases – Market Funnel and Brazilian E-Commerce



We will be performing exploratory data analysis and creating visuals using a BI Tool to answer the questions



Impact: Using the results of this analysis Olist can:

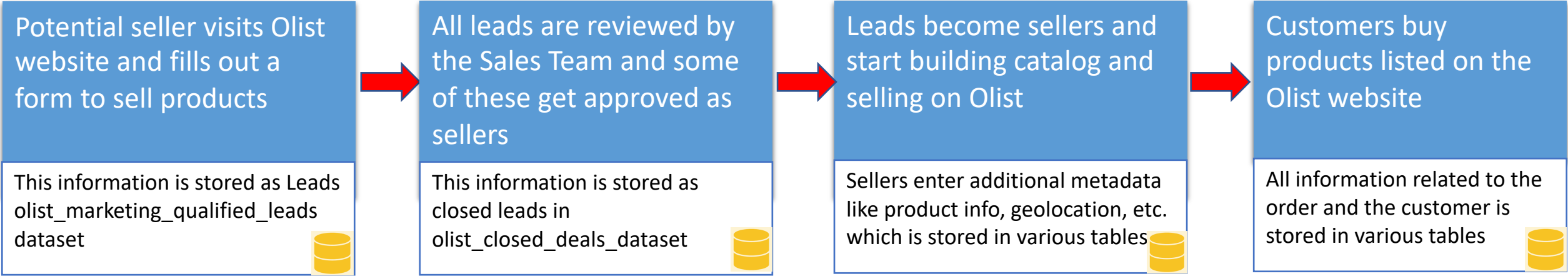
Summarize their performance

Make informed marketing decisions

Predict future revenue and customer behaviour

Optimize customer experience

Olist Business Model



Estamos quase lá!

Preencha abaixo e nossa equipe entrará em contato para explicar como você pode vender no Olist, a Maior Loja de Departamentos dos Marketplaces.

Nome*

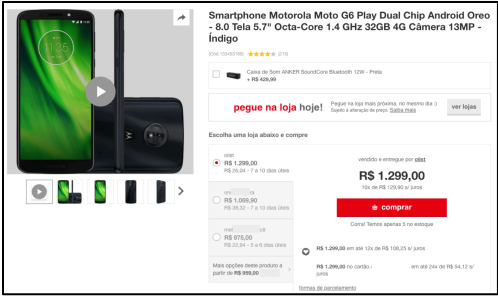
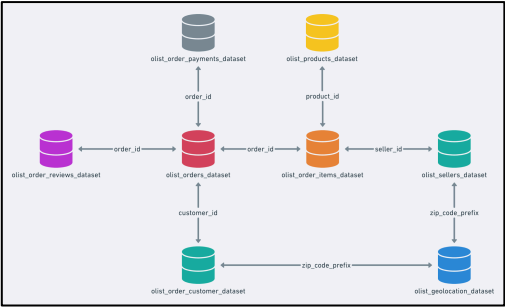
Email*


Telefone*

Seu segmento de negócio:*
Selecione

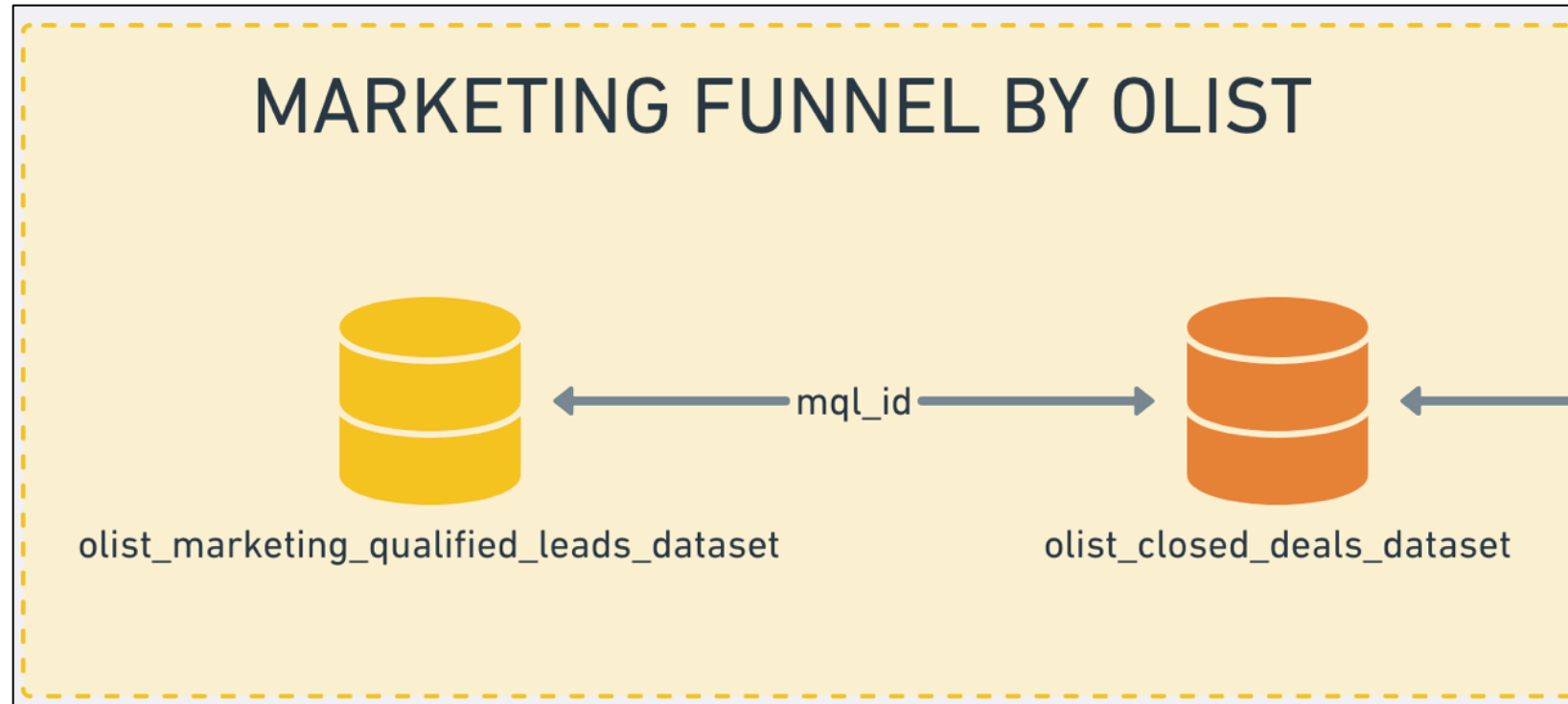
Você pode emitir Nota Fiscal Eletrônica (NFe)?*
Selecione

Enviar



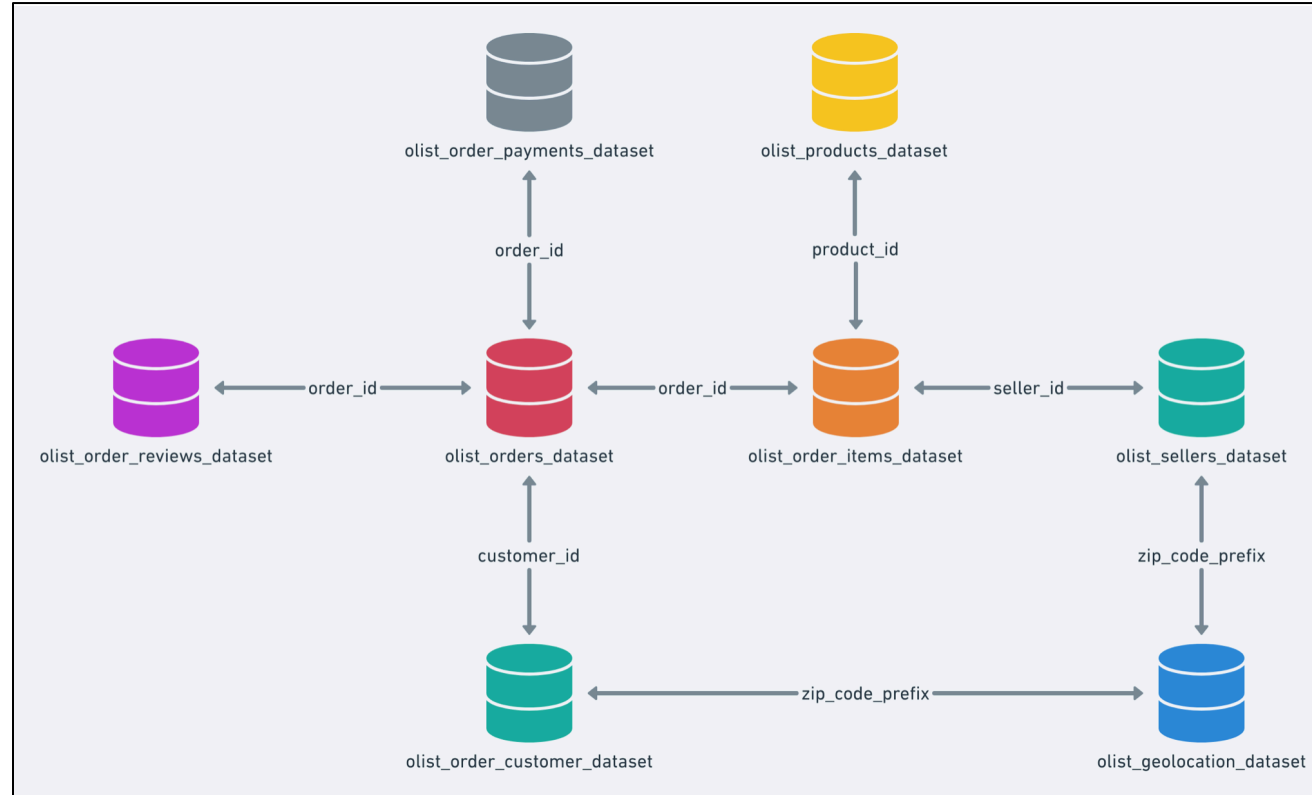
 = Data stored in database

The marketing funnel database consist of two tables which can be connected with a common field



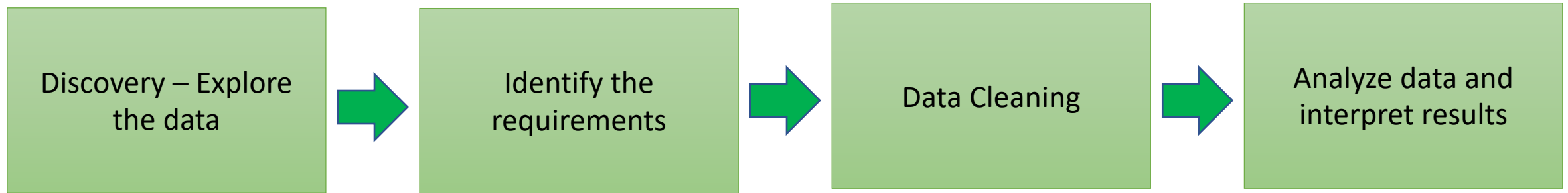
- Table 1: `olist_marketing_qualified_leads_dataset` – This contains information entered by a seller who wishes to sell on Olist
- Table 2: `olist_closed_deals_dataset` – This contains information on all the sellers that have been approved to sell on Olist

The Brazilian E-Commerce Dataset contains data of orders placed at multiple marketplaces in Brazil



- The tables contain information on the sellers, their products and the customers
- The dataset also contains information on the geolocation of the customers and the sellers , reviews left by customers for the products delivered, etc.

The Analysis Process



Discovery – Exploring the Data

Discovery –
Explore the data

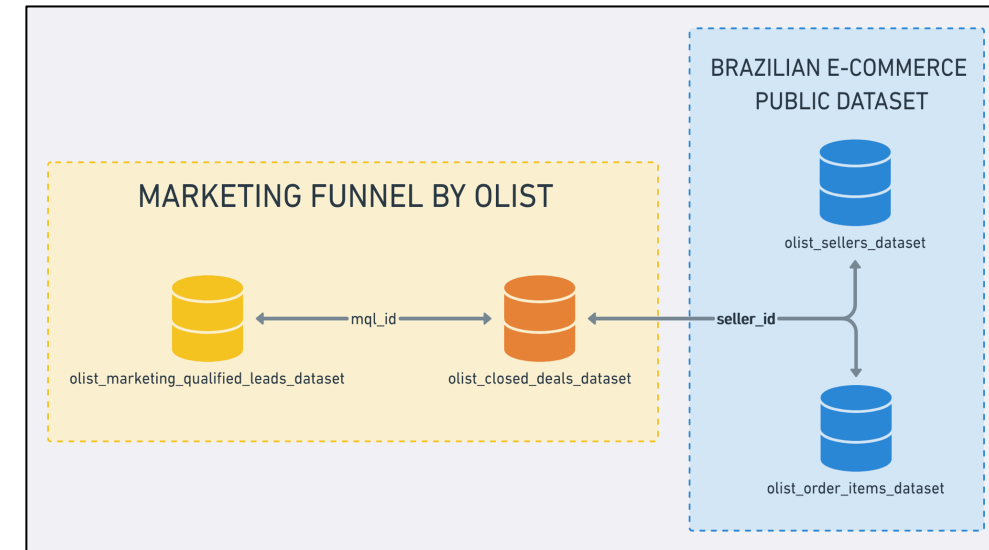
Some Interesting finds:

Marketing Funnel By Olist Dataset

- There seems to be a linear growth in the number of Leads as the company matures in terms of time *(See Appendix)*
- Earliest first contact date by a seller is June 14, 2017 and the first deal closed was on December 5th which indicates an area for improvement in the approval process

Brazilian E-Commerce Dataset

- olist_order_items_dataset contains information on items purchased within each order. This also includes sellers who are not from the Marketing funnel
- A majority of the customers and sellers are from São Paulo
- A majority of the customers have rated their orders as 5 indicating good customer satisfaction



Identifying the requirements - Questions I needed to answer with the help of the dataset

Identify the requirements

Q1. Summarize the seller funnel from marketing sign up to launching products on the platform

- Also provide the Customer LTV

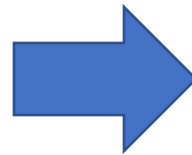
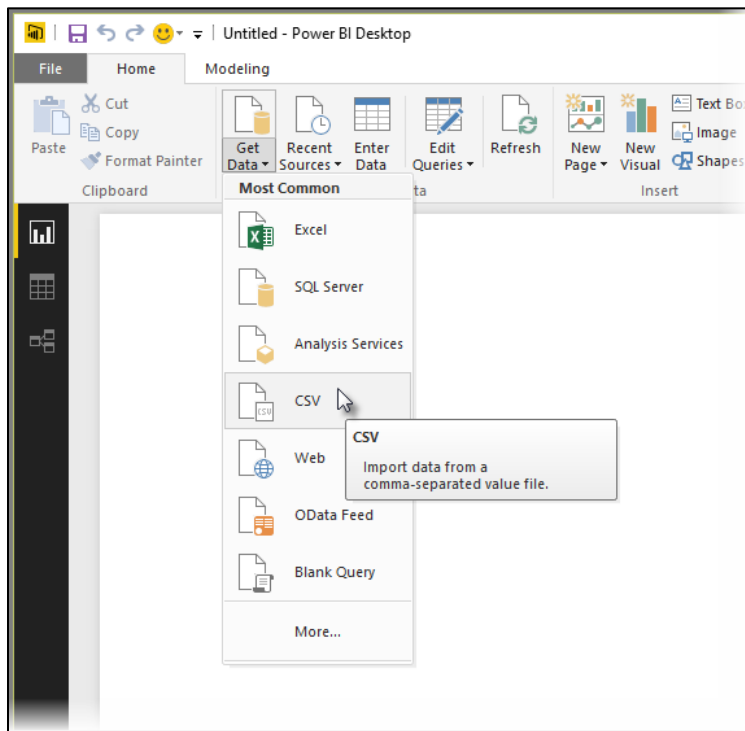
Q2. Summarize the current state of the business

- What has performance been monthly
- What are the best-selling categories

Q3. Predict future revenue for the next 12 months for order volume and revenue

To clean the data and prepare it for analysis and visualization I used Power BI

Data
Cleaning

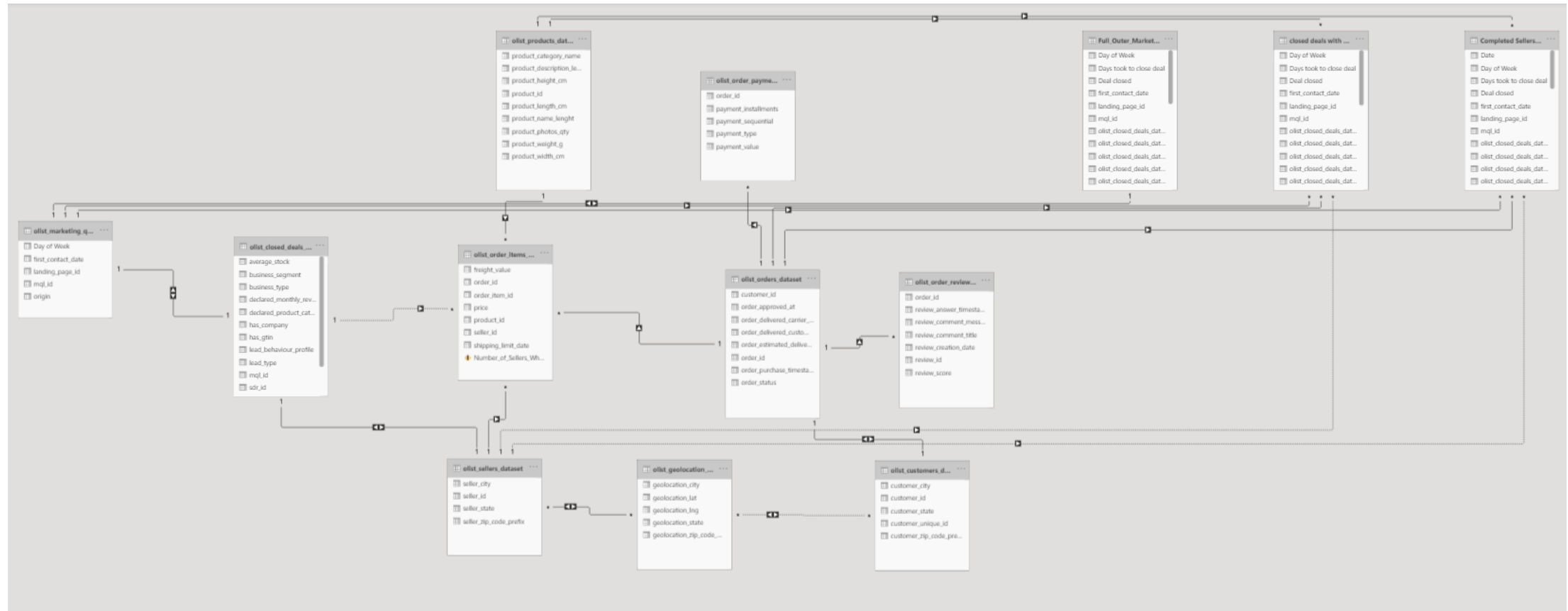


Queries [14]					
	olist_marketing_qualified_le...	olist_closed_deals_dataset	olist_order_items_dataset	olist_sellers_dataset	olist_orders_dataset
	olist_order_reviews_dataset	olist_order_payments_dataset	olist_products_dataset	olist_geolocation_dataset	olist_customers_dataset
1	6dd0ef38856fd5fb21aed5223aea9e66	6/14/2017	8f08d2ef8b9aec2ec5118fbcf7f96f11	display	
2	0b99dab71519032b917dc641cdd7a...	6/14/2017	d41d8cd98f00b204e9800998ecf842...		
3	70bd89abdbdba90e471cd3d0bbdba...	6/16/2017	8f08d2ef8b9aec2ec5118fbcf7f96f11	unknown	
4	98c9832865e4ab224a1649e8e6e9a...	6/20/2017	b6cd08ac77c501fd6857f7916d2c24f3	email	
5	e1e32379b464961f67fba6ddf6d4b3...	7/2/2017	0b37d43dbde6452cdd94c17bf61a88...	direct_traffic	
6	c0a4b0fa8862500a10f65f4d6b1d4490	7/3/2017	63a22f10d1d63d25eaf970d6a9c9df2a	email	
7	0a27b0193970757955db1cae7efb9c...	7/4/2017	c016fd3539eb32d58ae4dca01b4a20...	display	
8	da03a6389035b1470aa3d49b3ba43...	7/4/2017	22c29808c4f815213303f8933030604c	organic_search	
9	7c3a966d88a80726a95c2e16e56c39...	7/4/2017	e492ee5eaf1697716985cc6f33f9cd9b	other_publicities	
10	84403a769c84d34088c8b19584a23...	7/4/2017	d1bad8bf984bd4d7ee08f814ca5898...	paid_search	
11	984a8710e96d3210d30c185060159...	7/4/2017	e92a6f37ebb87a6bfbc569d439cf1635	organic_search	

Using the 'Get Data' feature I was able to connect to the tables

These tables were now available in my workspace to carry out analysis and investigate further

Next, I needed to establish relationships between all the tables in the two databases in order to meet my requirements



This was done by identifying and matching the primary and foreign keys between each table

Before trying to answer these questions, there was some cleaning activities needed to be done

Step 1: Filter data prior to Jan 2017 and after Sep 2018

Add Conditional Column

Add a conditional column that is computed from the other columns or values.

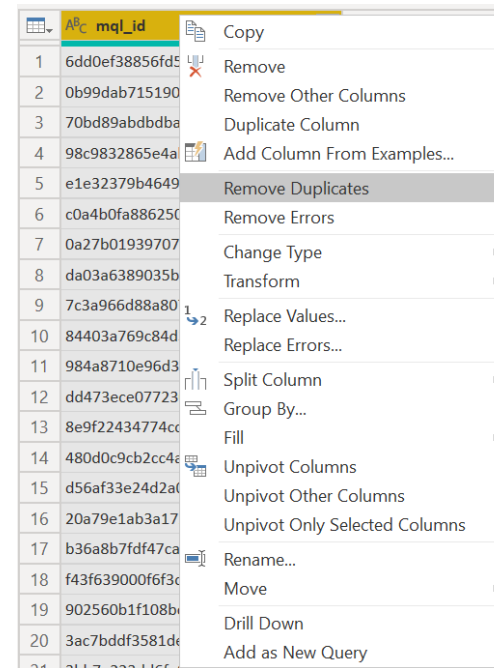
New column name
mql_id1

	Column Name	Operator	Value ⓘ		Output ⓘ
If	first_contact_date	is after	1/1/2017	Then	mql_id
Else If	first_contact_date	is before	10/1/2018	Then	mql_id

- Since the company's inception was January 2017, I ignored all data before January 2017 and since it's currently September 2018 I ignored all data after September 2018
- This was accomplished by creating a Custom Conditional Column for the tables as shown above

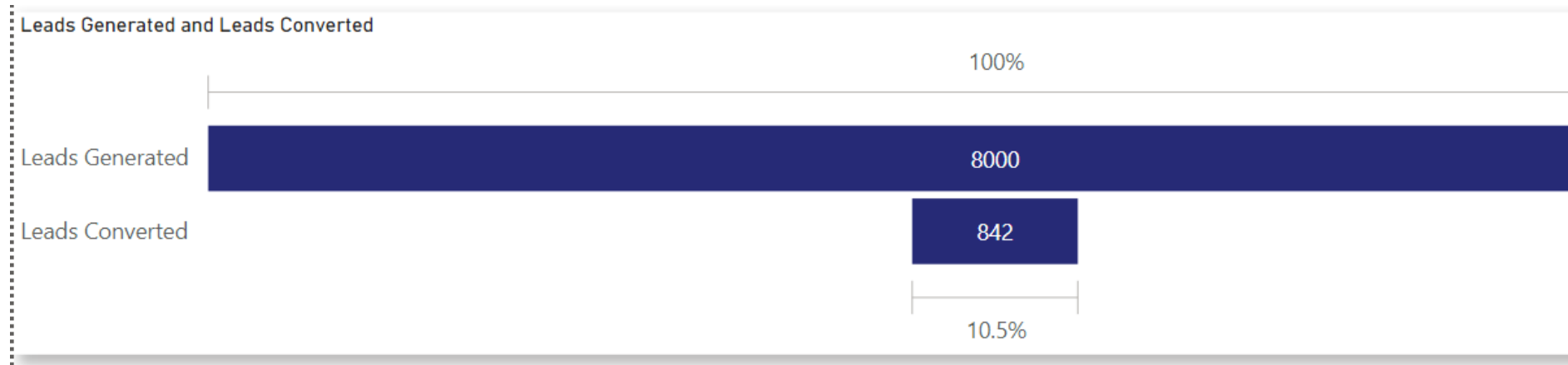
Before trying to answer these questions, there was some cleaning activities needed to be done

Step 2: Remove duplicates from mql_id



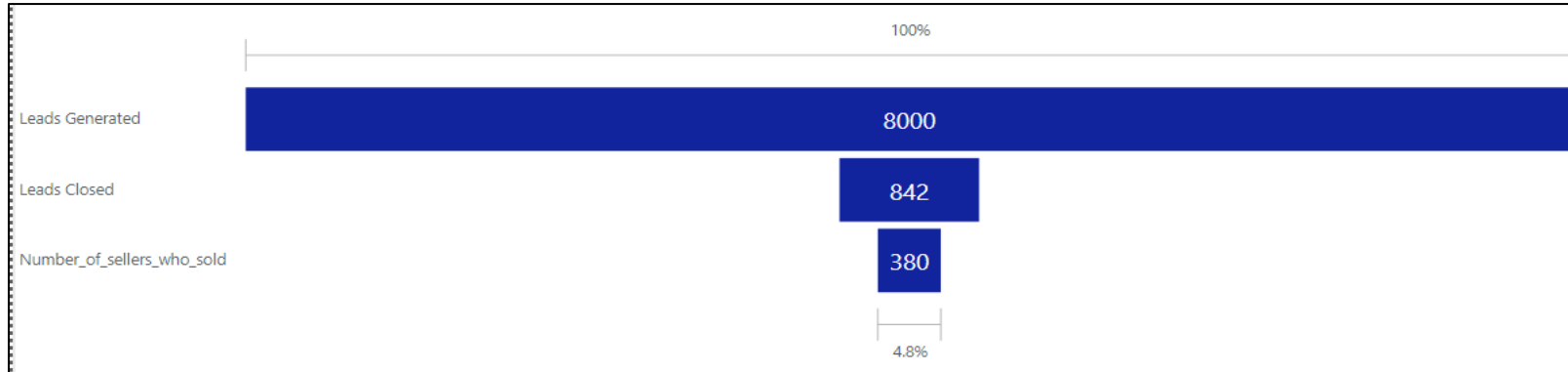
- The remove duplicates functionality within PowerBI allows easy removal of duplicates
- One thing I needed to be cautious about was which data value got deleted – The business logic I implemented was , in case of duplicates, keep the row with the earliest 'First Contact Date'

Q1. Summarize the seller funnel from marketing sign up to launching products on the platform. Also provide the Customer LTV



- The above visual shows the number total number of leads from the Olist website as 8000
- Of these 8000 leads, 842 i.e., 10.5% of the leads were confirmed as sellers

Q1. Summarize the seller funnel from marketing sign up to launching products on the platform. Also provide the Customer LTV

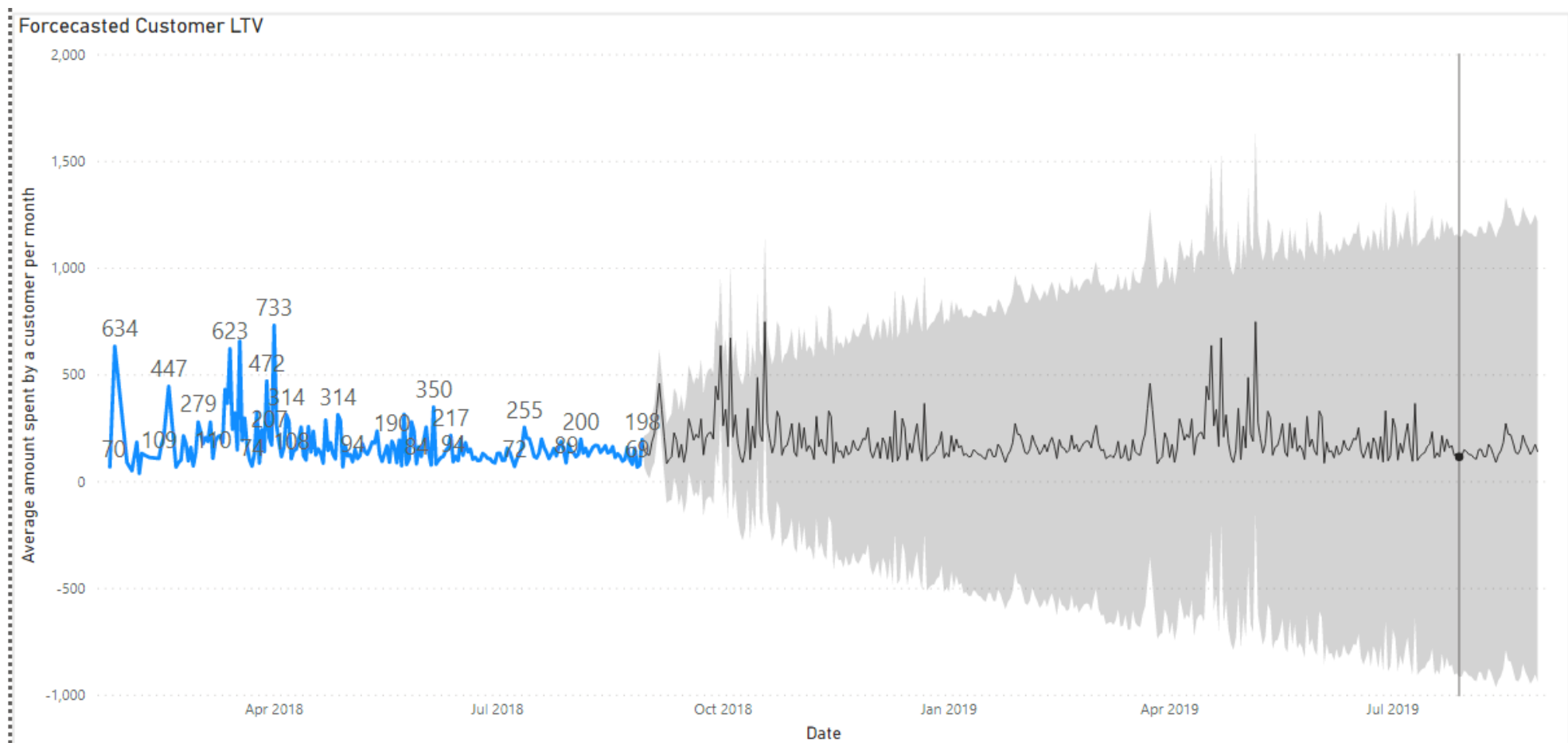


- On further analysis and connecting with the orders table in the Brazilian E-commerce database, I was also able to calculate that of the 842 approved leads, 45% (380) sold products to customers
- The above visual also indicates that only 4.8% of the 8000 leads end up selling products to customers.
- **Conclusion:** This indicates a possible problem area in the marketing funnel and a Business Gap in terms of revenue lost for Olist

Q1. Summarize the seller funnel from marketing sign up to launching products on the platform. Also provide the Customer LTV

Customer Lifetime Value is defined as how much a customer will bring in future revenue

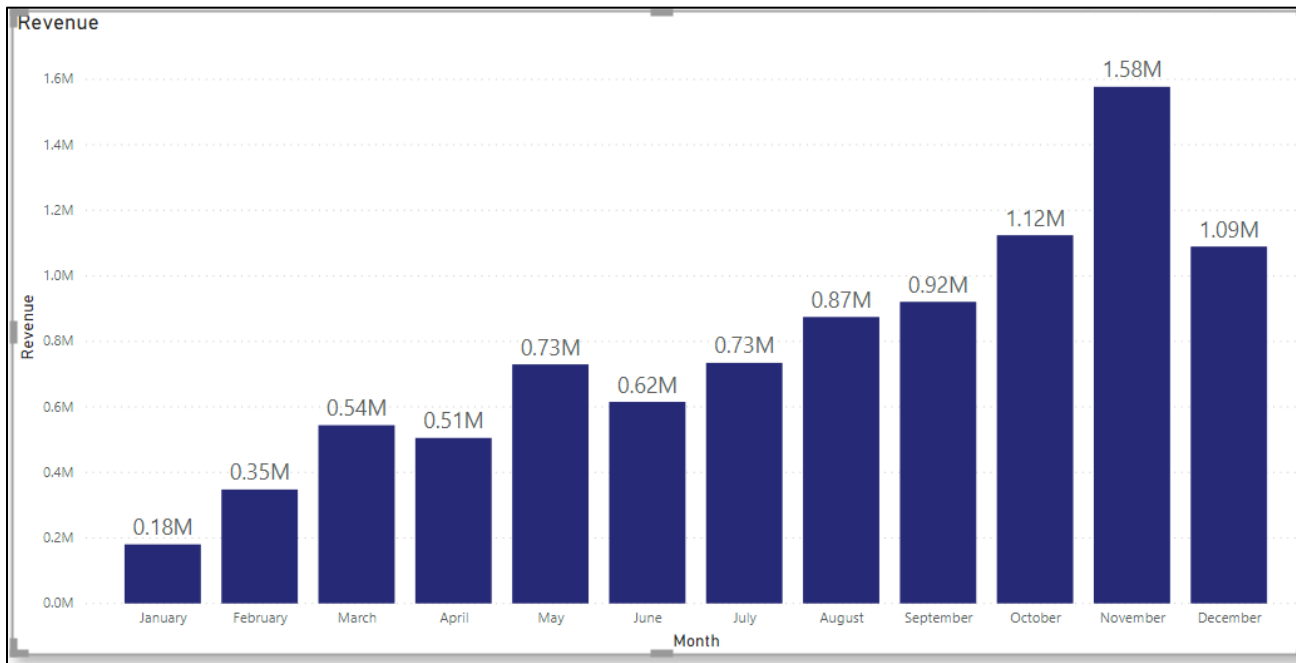
Using the data, I was able to forecast the Customer Lifetime Value for the next 12 months



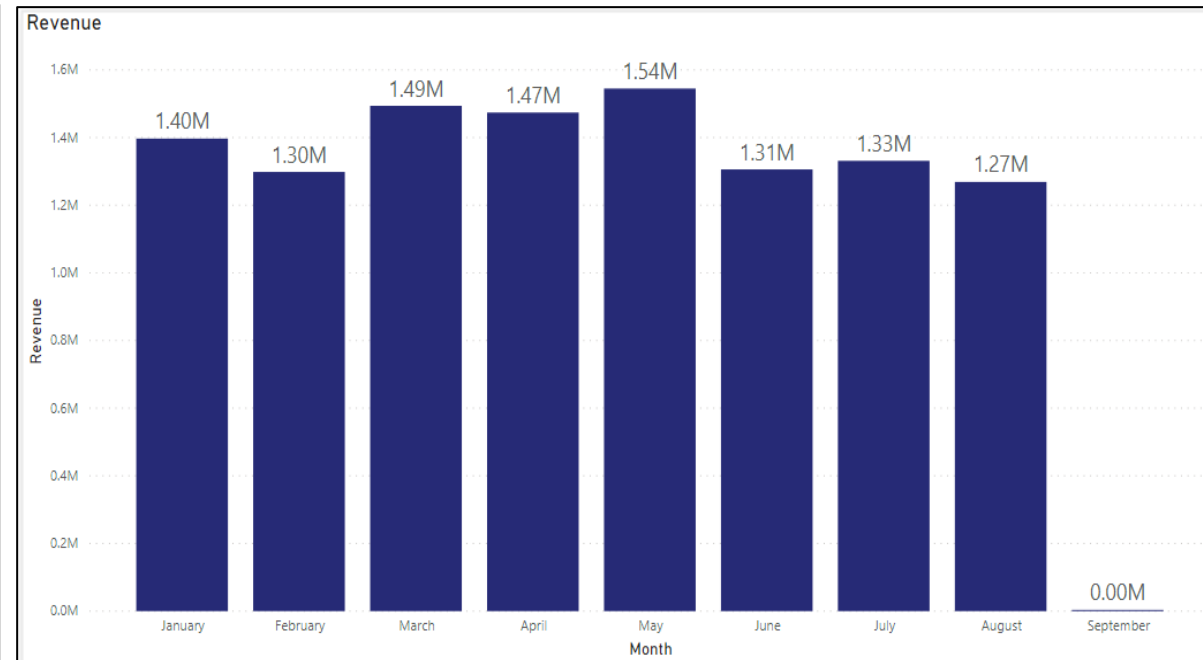
Q2. Summary of the current state of the business:

- What has performance been monthly
- What are the best-selling categories

Revenue for 2017 by month



Revenue for 2018 by month

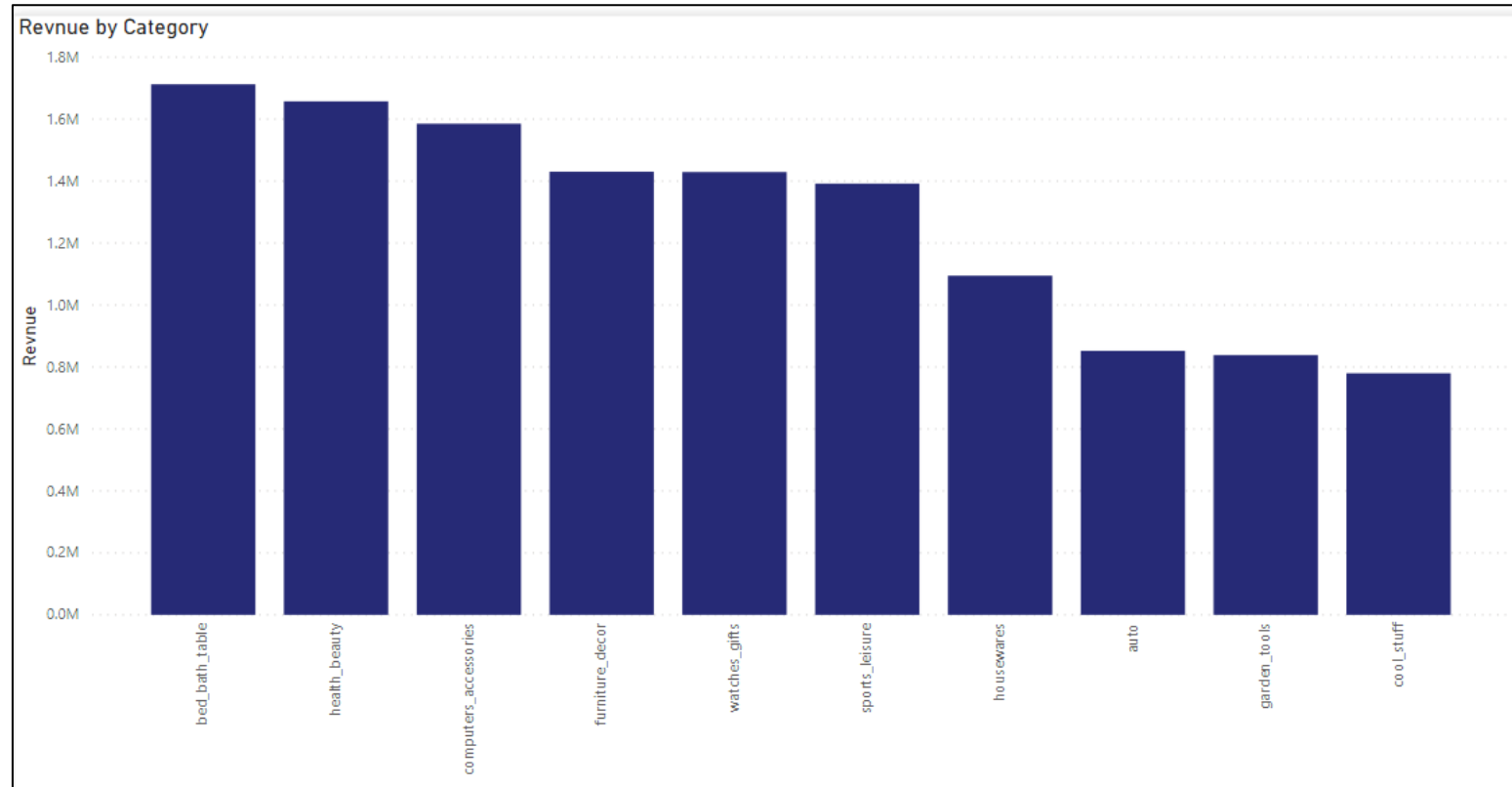


- To create this visual I needed to perform joins between the orders_dataset, payments_dataset, order_items_dataset.
- I then used bar chart visualization in PowerBI and also enabled an interactive drill-down capability to get to the month granularity

Q2. Summary of the current state of the business:

- What has performance been monthly
- What are the best-selling categories

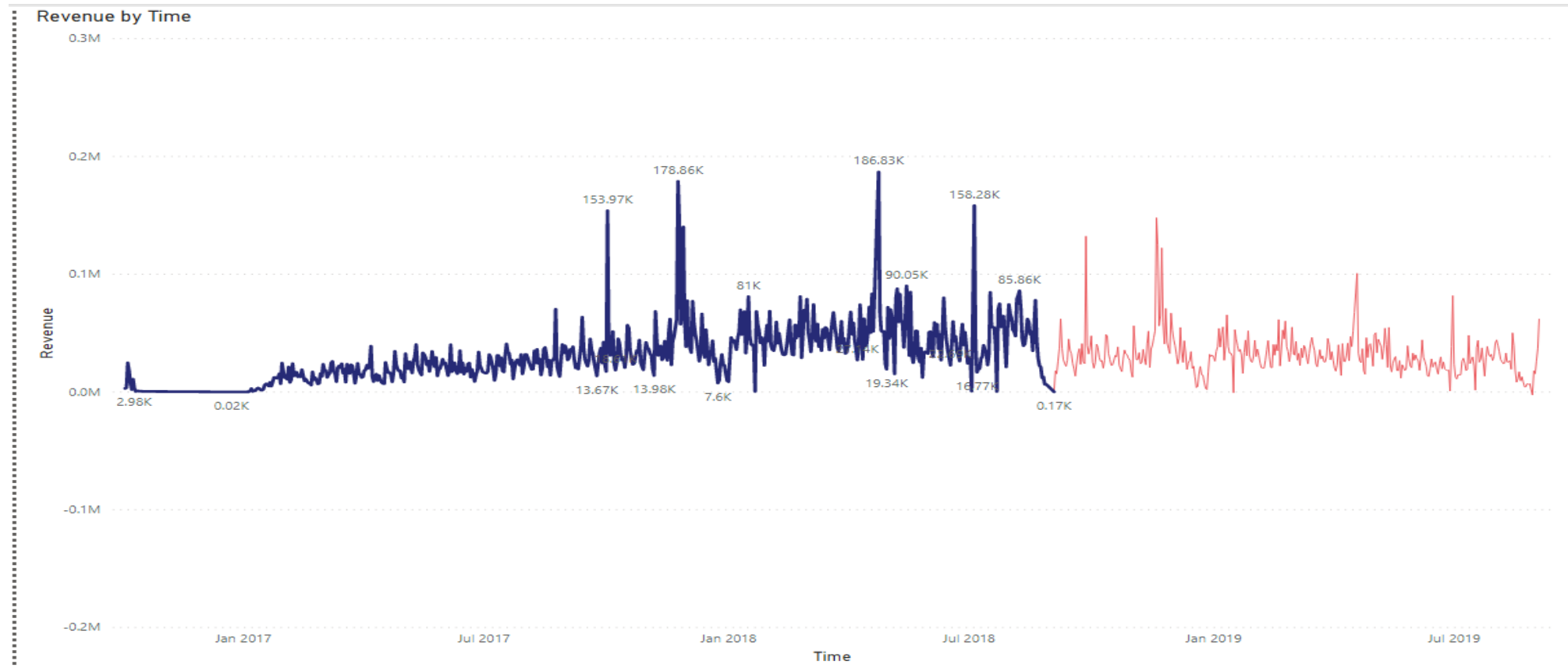
Category	Revenue
bed_bath_table	1,712,553.67
health_beauty	1,657,373.12
computers_accessories	1,585,330.45
furniture_decor	1,430,176.39
watches_gifts	1,429,216.68
sports_leisure	1,392,127.56
housewares	1,094,758.13
auto	852,294.33
garden_tools	838,280.75
cool_stuff	779,698.00



- The table consisted of a total of 71 categories
- To create this visual I needed to perform joins between the orders_dataset, payments_dataset, order_items_dataset.

Q3. Predict future revenue for the next 12 months for order volume and revenue

Analyze data
and interpret
results

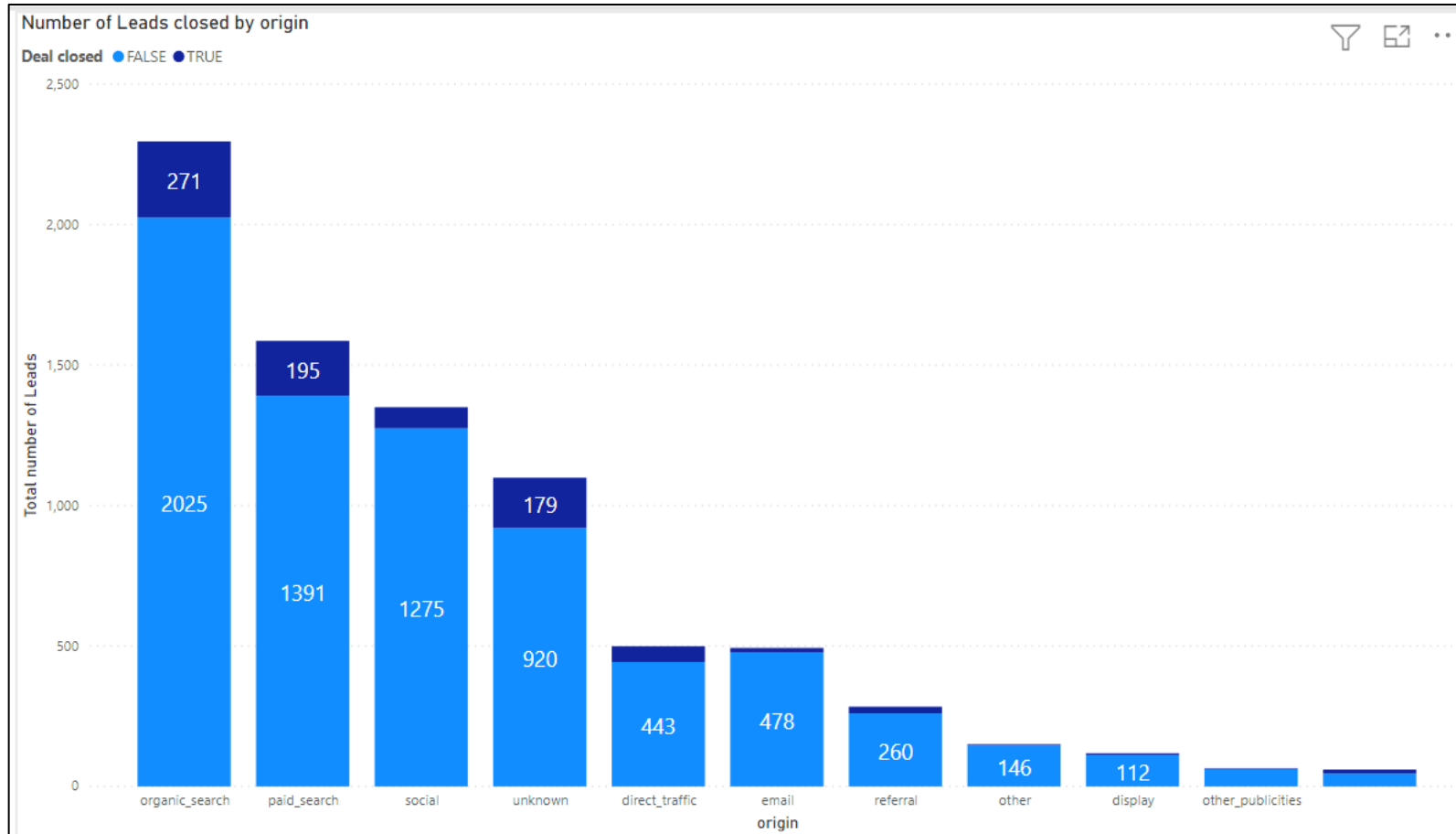


- To forecast this, I used the previous tables that I had joined (orders_dataset, payments_dataset, order_items_dataset)
- The fields used for this was – ‘Purchase Value’, and ‘Order Approved At’ (Date Time Field)

Appendix

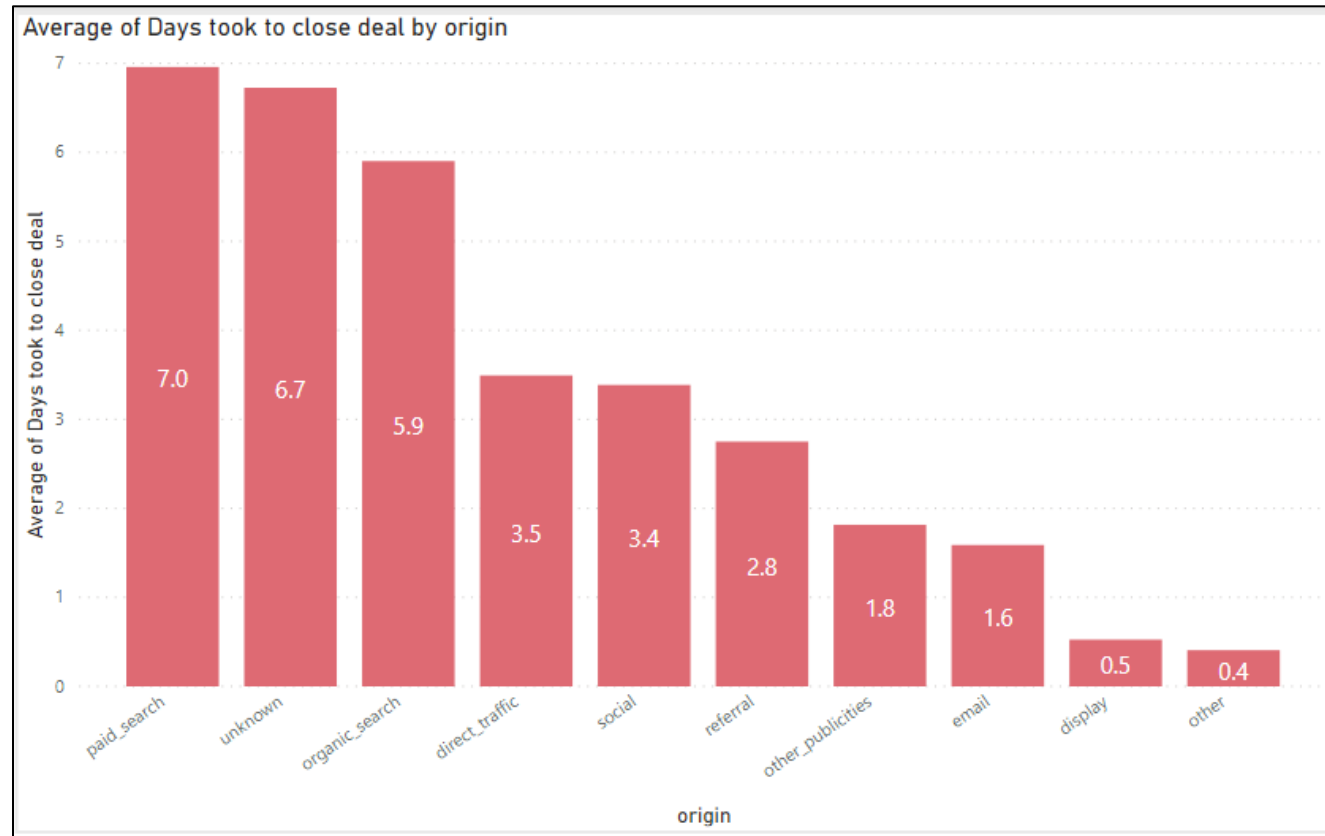
Some additional observations

Total number of Leads closed by origin



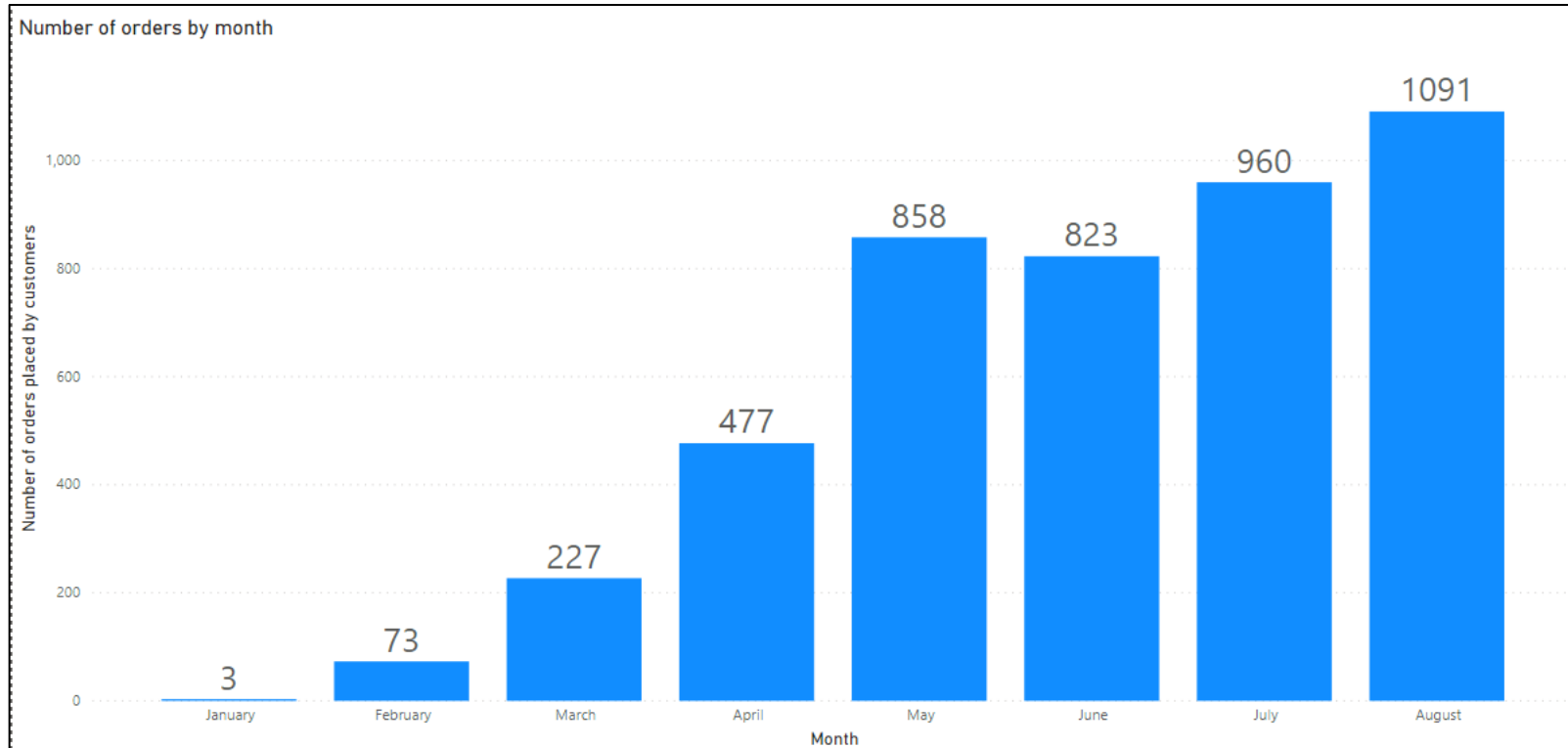
- This chart shows that paid_search is the most number of leads closed by the origin.
- This indicates that the online paid advertising based on search keywords is working well for Olist and they should focus more in this area.

Average days to close deal by Origin



- This chart shows that the number of days to close a deal is the highest for paid_search
- In conjunction with the previous slide, where paid_search has the highest rate of closing by origin, Olist may need to invest more time in bringing these 7 days down

Total number of orders placed by month



- This data shows a linear increase in the number of orders per month for the Olist website.
- This is a great trend which indicates a booming business! 😊

Number of leads by origin and business type

