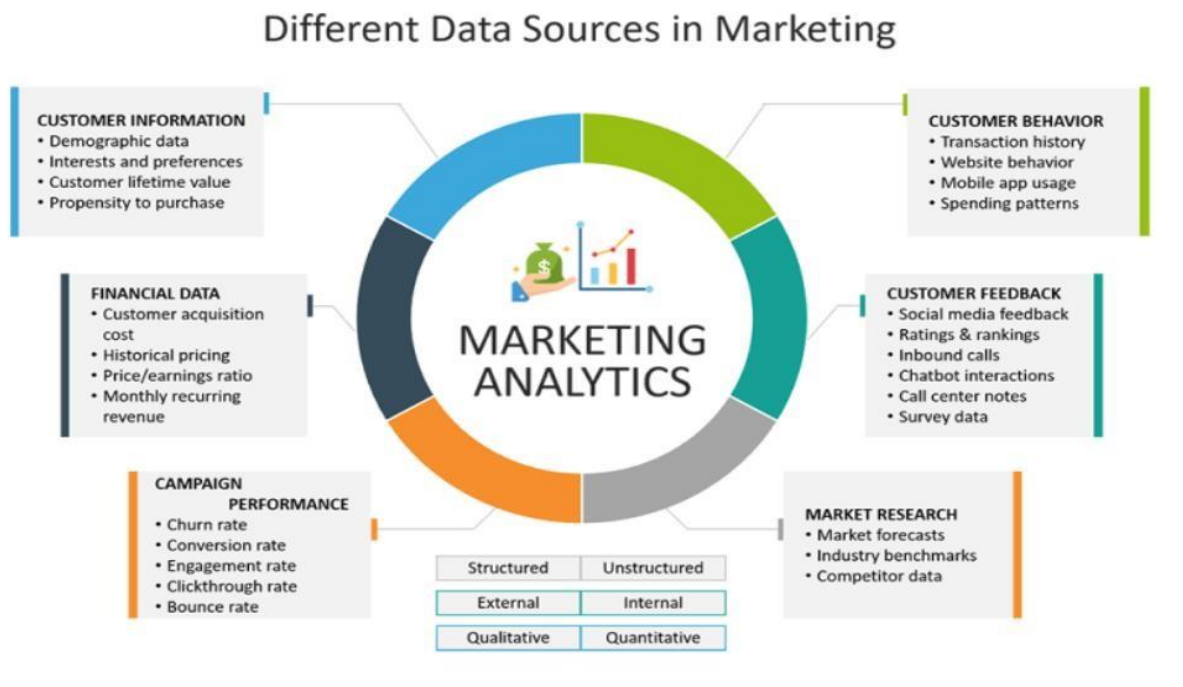


## Marketing Data Technical Report

### Introduction

This project was an exercise in marketing data analytics by applying analytical processes to marketing-related data. Marketing data analytics has the potential to help businesses to better understand what drives consumer action to help refine their marketing strategies.

There are different data sources in marketing.



In this project, I analyse data on:

[www.salesforce.com](https://www.salesforce.com)

1. Customer information (demographic data)
2. Customer behaviour (transaction history, website behaviour, spending patterns)
3. Campaign Performance (engagement rate, conversion are)

### Methodology

1. **Marketing data source:** marketing\_data.csv files and ad\_data.csv files.
2. **Data Selection:** Excel, SQL and Tableau were utilised to help select datasets of interest relevant for my project.
3. **Data Cleaning.** Cleaning of initial datasets in the marketing\_data.csv files and ad\_data.csv files was required to remove invalid data. For example, several customers had 'YOLO' and 'Absurd' listed as marital statuses. Corresponding IDs in the ad\_data.csv file were also removed.
4. **Data analysis and Data Presentation**

Following data selection and cleaning, charts were created in Tableau. Filters were added that allowed for the end user to toggle a single graph to see how measures would change

depending on different demographic factors. Different types of charts and colour schemes were selected to optimal presentation.

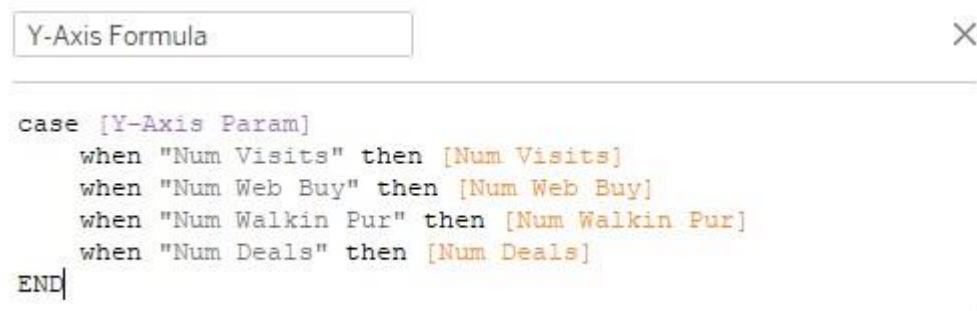
**Ages were grouped into age bins using a calculated field, to allow for the end user to better filter measures based on age ranges.** Here is my code to do this:



The screenshot shows a calculated field editor with a title bar 'age grouping' and a close button 'X'. The formula entered is: `IF [Age (copy)] <= 80 THEN [Age (copy)] ELSE 80 END`

I changed column names, to provide clarity to the end user.

**To allow the end user to toggle the axes in the chart, 'Customer Interaction Trends', I used calculated fields and parameters to facilitate this.** Here is the code I used in my calculated field to change the Y-Axis:



The screenshot shows a calculated field editor with a title bar 'Y-Axis Formula' and a close button 'X'. The formula entered is: `case [Y-Axis Param]  
 when "Num Visits" then [Num Visits]  
 when "Num Web Buy" then [Num Web Buy]  
 when "Num Walkin Pur" then [Num Walkin Pur]  
 when "Num Deals" then [Num Deals]  
END`

In 'Success of Advertising Channels', I also used parameters and calculated fields to swap between sheets.

In 'Trends in Customer Interaction', I included size and detail marks to ID to show where data points are clustered.

I made use of SQL to further query my data to isolate simple demographic data (e.g. the distribution of income), and to also determine how different advertising channels were popular among different demographics.

I created a database, and then an initial table, 'marketing\_data', with corresponding columns to the market\_data.csv file, and imported the latter.

To then display the total count of customers by demographic factor, I wrote my code using this following template:

```
SELECT m."Column1", COUNT("Column1")
FROM public."market_data" m
GROUP BY "Column1"
```

I wanted to **analyse data from the marketing data with advertising data**, to better understand how different advertising channels (y) were popular amongst different demographics (x). To do this, I imported and created a table for the ad\_data.csv file in the same way I did the marketing\_data.csv file. I then used the INNER JOIN clause to create a results table with the desired data.

```
SELECT m."Column(x)", SUM(a."Column(y)") As Total_(y)_By(x)
FROM public."ad_data" a
INNER JOIN public."market_data" on a."ID" = m."ID"
GROUP BY m."Column(x)"
ORDER BY m."Column(x)";|
```

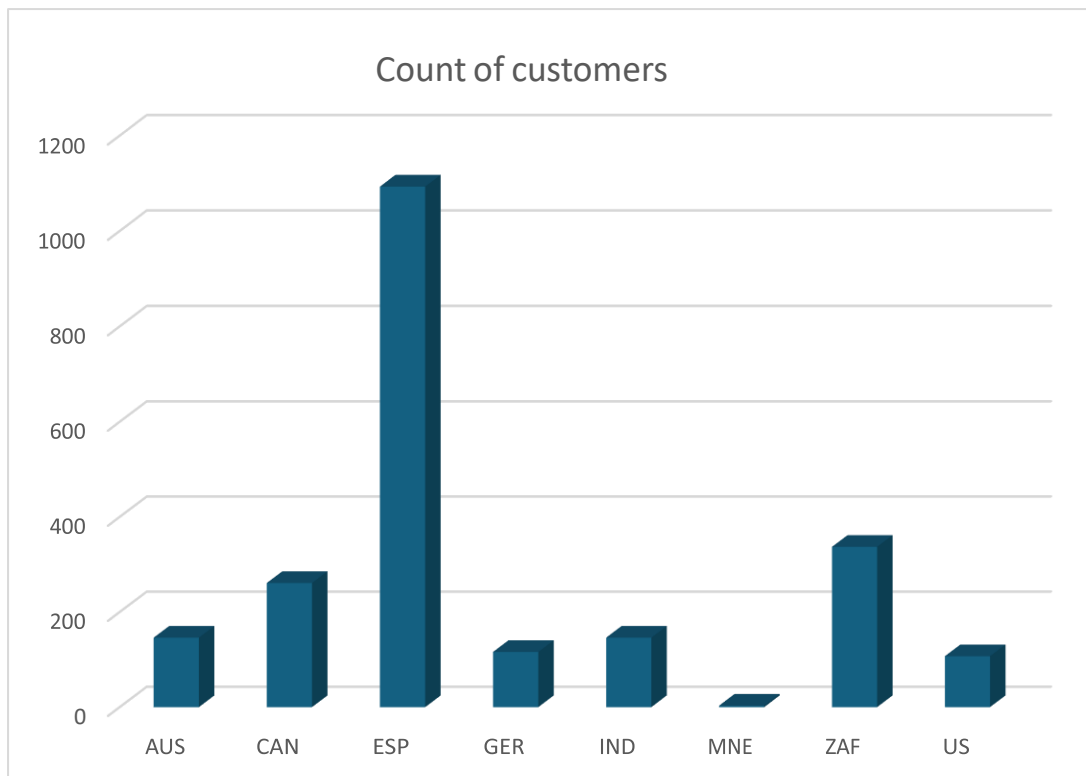
This is how I determined total bulkmail advertisement conversions by country, for example:

```
SELECT m."Country", SUM(a."Bulkmail_ad") as Total_Bulkmail_Ad
FROM public."ad_data" a
INNER JOIN public."market_data" m ON a."ID" = m."ID"
GROUP BY m."Country"
ORDER BY m."Country";
```

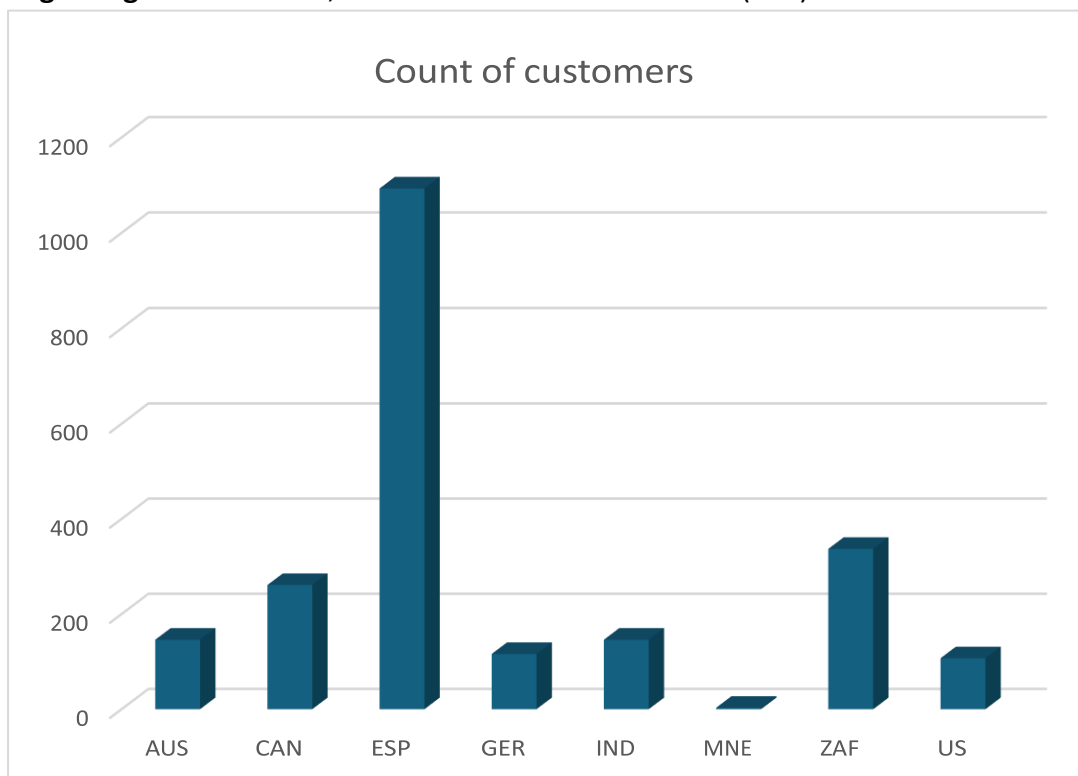
## Results and Findings

### 1. Customer information (demographic data)

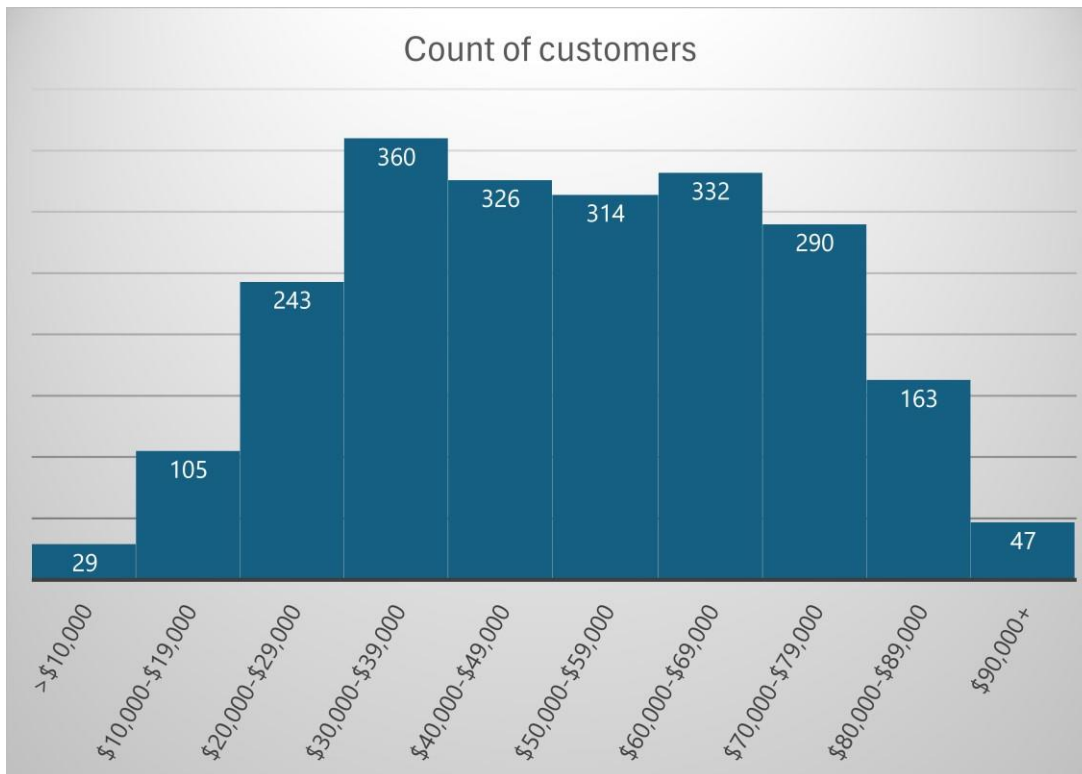
**Regarding customer location**, most customers at the time of collection were in Spain (1093) by an extremely large margin. The country with the lowest count was Montenegro (3).



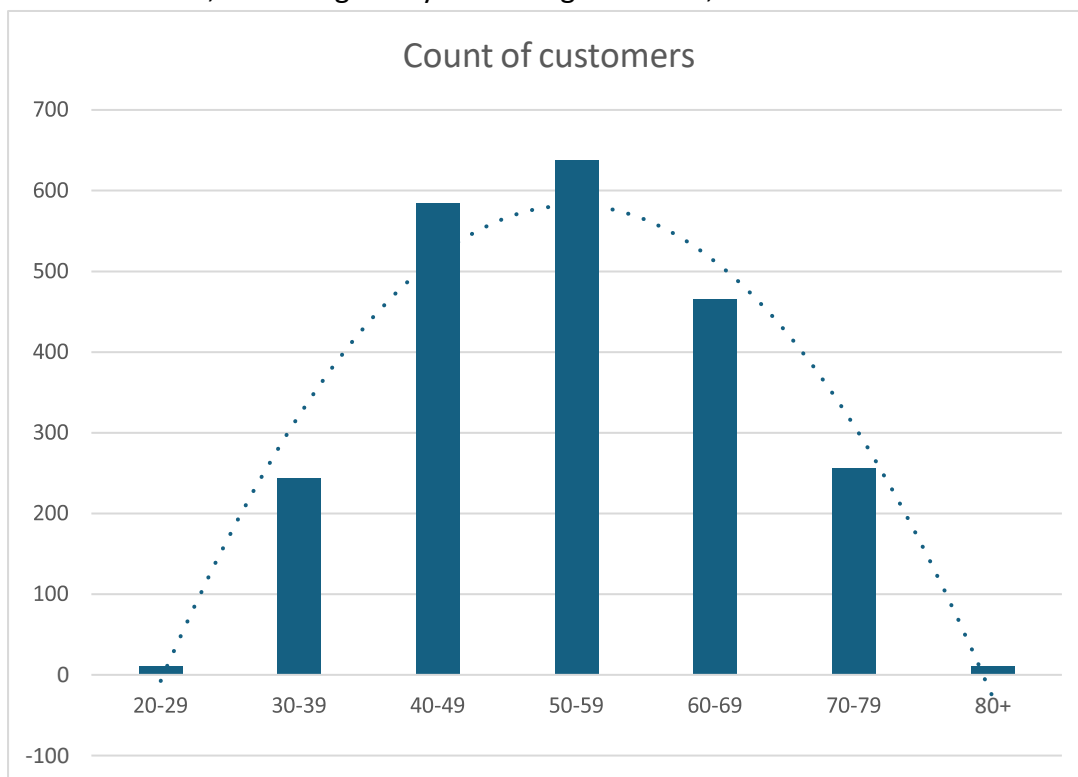
**Regarding marital status,** most customers were married (857)..



**Regarding income,** the highest income bracket was \$30,000-\$39,000 (360), and the lowest was >\$10,000 (29). The skew value was -0.549, indicating a negative skew.

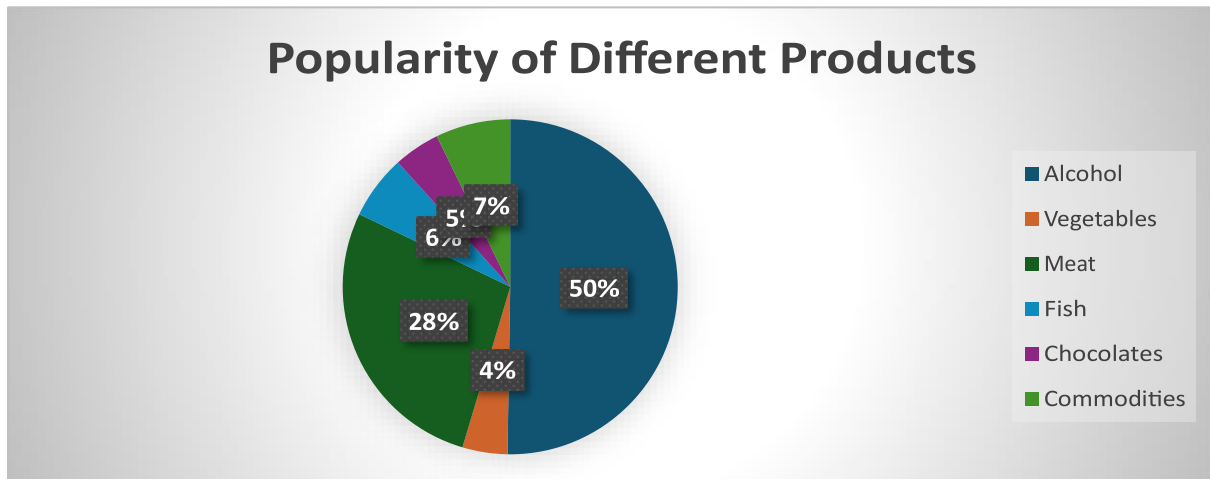


**Regarding age**, most customers were between 50-59 years of age (638), followed by 40-49 (584). Very few people were in the 20-29 age group (10), and the 80+ group (11). The skew value was -0.01, indicating a very minor negative skew, and an almost even distribution.



## 2. Customer behaviour (transaction history, website behaviour, spending patterns)

In **'Popularity of Different Products'**, several insights were unravelled. The general trend across all demographics is that alcohol is by far the most popular product. The second most popular was meat.



However, there were some outliers: among those earning up to \$29,000, meat was by far the most popular, however those in income brackets greater than this stuck to the main trend.

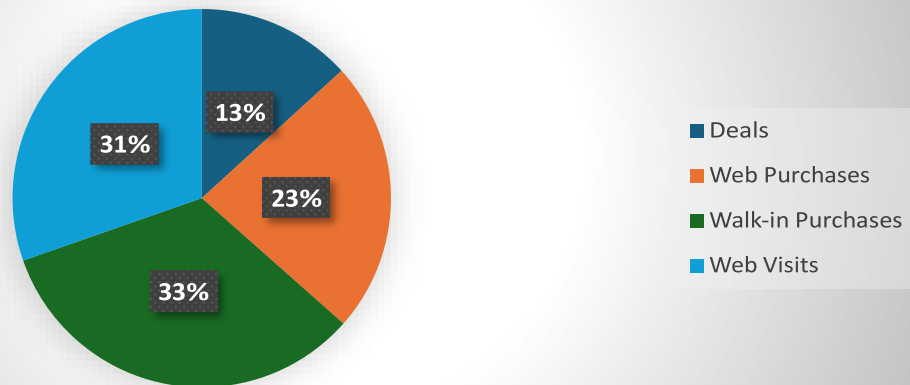
Among all countries, this main trend also continued, apart from Montenegro, where meat was more popular in the 30-39 category.

Among all age ranges, this trend persisted, apart from those in the 20-29 category, where the amounts spent on alcohol and meat was almost equal.

**While the anomaly for Montenegrins and those aged 20-29 is noted, due to a limited sample size, no conclusive findings can be made.**

In **'Popularity of Customer Interactions'**, the most popular interactions were walk-in purchases and website visits across all demographics.

## Popularity of Customer Interaction



### 3. Campaign Performance (engagement rate, conversion are)

I explored how the number of web visits correlated with other customer interaction behaviours, as well in producing lead conversions. Generally, there is a negative correlation aside from the number of deals. There exists a positive correlation between number of web visits and number of deals across all demographics.

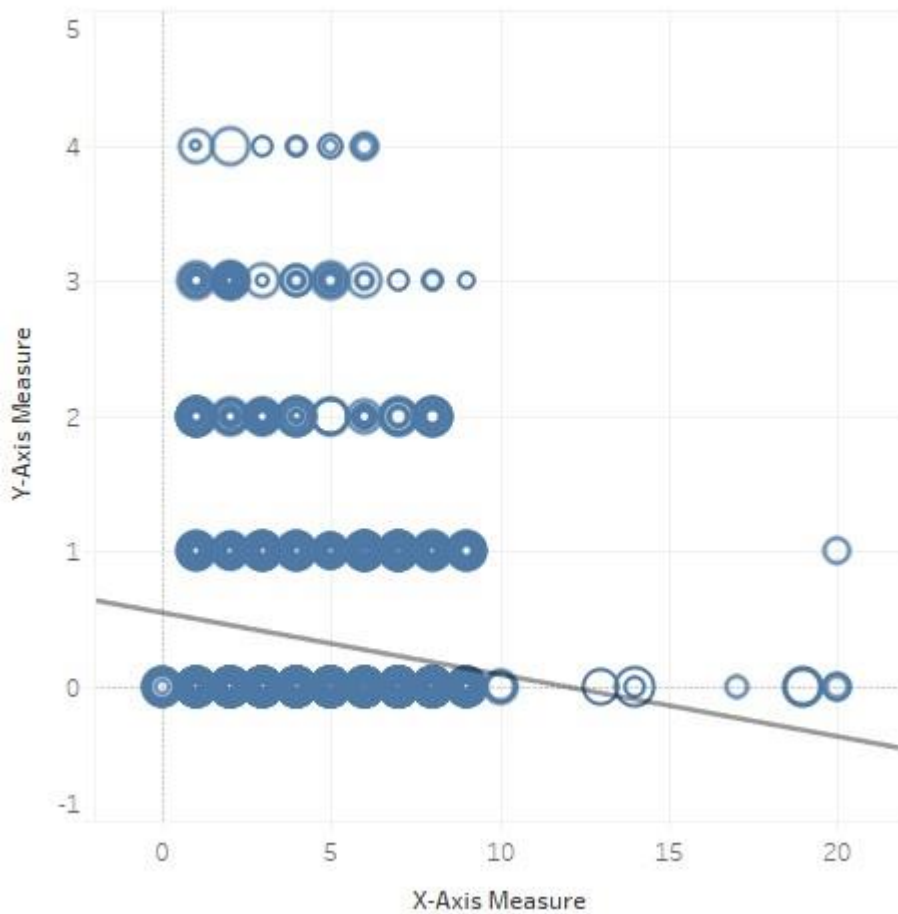
Against walk-in purchases, there was a negative correlation against all demographics, aside from certain income groups, where there was a positive correlation for those in the \$10,000-\$19,000, \$40,000-\$49,000, and \$70,000-\$79,000 income groups

Against web purchases, while there was still a negative correlation against all demographics, the trend was not as clear. There was a positive correlation for those age ranges 60+. There was a negative correlation for those of all marital statuses aside from those divorced and widows. There was a negative correlation for all education levels aside from Basic and Masters. No significant trends could be determined when filtering by income.

Against lead conversions, there was a positive correlation for those with either children or teenagers at home. There was a negative correlation for all age groups aside from 20-29.

**Overall, there was a negative correlation between number of web visits and lead conversion success rate.**

## Trends in Customer Interaction



*A negative correlation between number of visits and lead conversions*

I explored the success of the different advertising channels. Regarding the success of Twitter in producing lead conversions, it was overwhelmingly unsuccessful aside from the 20-29, and 80+ age ranges in Spain. In those cases, the conversion rate was 40% for both groups. However, the sample sizes that fit these specific demographic profiles was too small in this case to draw any meaningful conclusions. The success rate of Facebook, Instagram, Bulkmail and Brochure advertisements was low, with low conversion rates across all demographics.

Advertisement Type	Success rate
Twitter	7.38%
Instagram	7.29%
Facebook	6.43%
Brochure	1.36%
Bulkmail	7.38%



## **Reflections**

In undertaking and completing this assignment, I have learnt fundamentals of Excel, Tableau and SQL. I have begun to think like a data analyst and have learnt to solve datasets in a logical and productive manner.