

CANDIDATE NUMBER (C-NUMBER)	C2096215
MODULE NAME	Data Visualisation
WORD COUNT	3--
SUBMISSION DATE	12/06/2024
DECLARATION I certify that this assessment submission is entirely my work and I have fully referenced and correctly cited the work of others, where required. I also confirm the contents of my submission have not been generated by a third party, or through an Artificial Intelligence generative system*. I have read the Student Discipline Regulations (Student Discipline Regulations) and understand any Assessment Related Offence/ Academic Misconduct may result penalties being applied. By submitting this assessment submission, I am confirming that I am fit to sit according to the Assessment Regulations. I declare that: <ul style="list-style-type: none">• This is my own unaided work. Yes <input checked="" type="checkbox"/> No <input type="checkbox"/>• The word count stated by me is correct. Yes <input checked="" type="checkbox"/> No <input type="checkbox"/>• I'm happy for my work to be retained on the Elite repository and made available to staff and future students** Yes <input checked="" type="checkbox"/> No <input type="checkbox"/> *Please note that all the assignments are submitted to Turnitin. **Please note personal information (such as names) will be deleted.	

Table of Contents

1. Portfolio Activity 1: Responses to Topics	1
1.1. Unit 1 Consolidate	1
1.2. Unit 2 Consolidate	2
1.3. Unit 3 Consolidate	4
1.4. Unit 4 Consolidate	5
1.5. Unit 5 Consolidate	6
1.6. Unit 6 Consolidate	7
2. Portfolio Activity Two: Accessibility of Data Visualisation	9
2.1. Visualization's clarity	9
2.2. Including a Variety of Audiences	9
2.3. Ease of Interpretation	9
2.4. Importance of Accessibility in Global Organizations	9
2.5. Accessibility in Practice	10
2.6. Conclusion	11
3. Portfolio Activity Three: Comparison of Tableau and Power BI Visualisations	12
3.1. Popularity	12
3.2. Number of Functions verity of Type variable Supported in Functions	13
3.3. Data Sources	13
3.4. Data preprocessing	13
3.5. User Interface	14
3.6. Conclusion	16
4. Portfolio Activity Four	17
4.1. Preparing Data for EDA	17
4.2. Statistical Analysis	20
4.3. Univariate Analysis	21
4.4. Multivariate Analysis	26
4.5. Conclusion	35
5. References	36

Figure 1: Unit 1.....	1
Figure 2: Unit 2-1	2
Figure 3:Unit 2-2	3
Figure 4:Unit 3-1	4
Figure 5:Unit 3-2	4
Figure 6:Unit 4.....	5
Figure 7:Unit 5-1	6
Figure 8:Unit 5-2	6
Figure 9:Unit 6-1	7
Figure 10:Unit 6-2	8
Figure 11 Compare Two Plot	10
Figure 12: (5 tips on designing colour-blind-friendly visualizations, no date)	10
Figure 13: Visualization Sample	11
Figure 14: google trend (Google Trends, 2024)	12
Figure 15:Gartner.....	12
Figure 16	13
Figure 17: Tableau Prep- Power Query	13
Figure 18	17
Figure 19:Columns Information	18
Figure 20: Dnusual Data-Part 1.....	18
Figure 21: Unusual Data-Part 2.....	19
Figure 22:Statistical Description.....	20
Figure 23:BoxPlot	21
Figure 24:Heatmap	23
Figure 25:Count Education, Countries, and Marital	24
Figure 26 sales per product.....	25
Figure 27 Product by Marital Status	26
Figure 28:Average Age-Marital State	27
Figure 29: Products Based on Marital State.....	27
Figure 30 Product by Education	28
Figure 31: Average Income Based Education	29
Figure 32: Kids in Home Based On Education	29
Figure 33:Rate of Accept Campaign based Education	30
Figure 34:Rate of Buying Per Media Based Education	31
Figure 35: Sum of Each Product Buying Based on Number of Kids at Home	31
Figure 36: Sum of Each Media Sales Based on Number of Kids at Home	32
Figure 37: Average Response based Incoming Group.....	33
Figure 38: Average Response based Age Group	33
Figure 39:Distribution of Type of Product Based on Income Categories	34
Figure 40	35

1. Portfolio Activity 1: Responses to Topics

1.1. Unit 1 Consolidate



Farzam FERYDOONI

27/03/2024, 13:52



Within the scope of this survey (Mattia, no date), it was not expressly stated that Tableau is the tool that is recommended or favoured over the others. According to the findings of the survey, the following are the most important aspects of Tableau:

1. The authors discovered that the process of producing dashboards in Tableau was relatively comparable to that of Excel, but created graphs more easily.
2. Although trendlines were unable to display a connected moving average, they were able to display a moving average that was broken down by month.
3. It was possible to generate data labels and slicers (for filtering by quarter) with a reasonable amount of ease.

The overall purpose of the study is to demonstrate the advantages and disadvantages of each tool in an exploratory manner, without clearly favouring one tool over the others. It appears that the authors are suggesting that business schools should conduct an analysis and select acceptable data visualisation technologies to incorporate into their curriculums depending on the requirements of the market and the requirements of the students. On the other hand, Tableau is not presented in a way that makes it appear to be the most obvious or dominating option based on their exploratory investigation.

References:

- Mattia, A. (no date) 'Data visualization: an exploratory study into the software tools used by businesses', 18.

[Reply](#)

Figure 1: Unit 1

There is a comprehensive review of the Diamond and Mattia (2017) work, which It takes a close look at the study and points out that it doesn't definitively say that Tableau is better than other tools. The discussion based on the specific data, like how easy it is to make screens and how limited trendlines are. It's important to note that this study wasn't meant to recommend one tool; instead, it was meant to find the pros and cons of different tools. In terms of education, It says that business schools should not use a one-size-fits-all method when analysing tools, but instead should base their decisions on what the market needs. At all times, It keeps pointing out Tableau's benefits without making it sound like the clear winner.

1.2. Unit 2 Consolidate



Farzam FERYDOONI

24/03/2024, 16:35

There is no doubt that Tableau's set of tools for building stories out of data is tempting. Even so, it is important to be aware of the problems that might come up when trying to meet every cognitive benefit. We'll look at the pros and weaknesses in more detail below (Greenfield & Moorjani, 2019):

- Advantage:

Processing what we see is a natural thinking skill that comes naturally to humans (Ware, 2013). Tableau uses this to make it easier to understand patterns and trends by turning complicated data into maps, charts, and graphs.

Visually appealing and carefully written narratives can be used to effectively share insights and keep people's attention, even if they have different levels of data literacy.

-Disadvantage:

The bad things are that too much knowledge or a cognitive load that is too complex can possibly overburden viewers, especially those with learning disabilities or cognitive impairments. Accessibility features like alternative text descriptions and colorblind-friendly colours can be added to fix this problem (W3C WAI, 2022).


Individual differences, like visual, auditory, and physical learning styles, can affect how people understand what they are learning. How much each person connects with Tableau's storytelling tools may be different.

Comments other posts:

based on Mohammed Yusuf Shaikh's post There are a few key features of Tableau that make it easier for people with different ways of thinking to see data and tell stories. To begin, Tableau's focus on visual representation helps understanding by showing data in maps and dashboards that are both visually appealing and mentally stimulating.

Another important factor is interactivity, which lets users look at material at their own pace and takes into account differences in how people think. Users can tailor visualisations to their needs and make them more relevant by using customisation choices.

Figure 2: Unit 2-1



Farzam FERYDOONI
 24/03/2024, 16:35

Another important factor is interactivity, which lets users look at material at their own pace and takes into account differences in how people think. Users can tailor visualisations to their needs and make them more relevant by using customisation choices.

Tableau's narrative settings make it easier to tell interesting stories about data by letting users change the amount of detail and verbosity to fit different audiences.

Lastly, the post praises Tableau's accessibility features, which let users of all abilities use its powerful data visualisation tools. This encourages truly inclusive data stories.

Upgradable Items in Tableau:


1. For people who aren't good with numbers, use simple graphics and highlight the most important ideas
2. -Add dynamic features like drill-down or filtering options so users can look into information on their own
- 3.

Finally, Even though Tableau can be a useful tool for sharing ideas through data visualisation, it is important to note that it can't meet all cognitive needs. By using inclusive design techniques (Glinert, 2018) and tailoring the story, we can make sure that our data stories can be understood and shared by a wider audience.

References:

- Glinert, E. (2018). Inclusive design for dummies. John Wiley & Sons.
- Greenfield, A., & Moorjani, S. (2019). Tableau for dummies. John Wiley & Sons.
- Ware, C. (2013). Visual thinking for information age. Morgan Kaufmann Publishers.
- W3C WAI (World Wide Web Consortium Web Accessibility Initiative). (2023, December 13). Web Content Accessibility Guidelines (WCAG) 2.2. <https://www.w3.org/WAI/> (<https://www.w3.org/WAI/>)

[Reply](#) | [Hide Replies \(1\)](#) • **NEW**



Maulik SUTHAR
 Overall, your analysis addresses the topic of cognitive diversity well by acknowledging both the benefits and limitations of Tableau for users with different learning styles and abilities.

However, please confirm To what extent do you agree that telling a story in Tableau can adapt to all cognitive needs?

[Reply](#) • 31/03/2024, 23:09 • **NEW**

Figure 3:Unit 2-2

The post talks about how important it is to be welcoming and suggests adding tools that are easy for color blind people to use. There are comments on other posts that acknowledge the usefulness and originality of what other users have said. it gives a fair view and stresses how important it is for storytelling methods to be flexible enough to fit the needs of different people. The discussion about cognitive needs in data-driven stories is better because of the critical engagement with earlier comments and the focus on inclusivity.

1.3. Unit 3 Consolidate

[Reply](#)



Farzam FERYDOONI

27/03/2024, 14:37

Chartability, as described by Fizz Studio [2023], is a new way to make data visualisations easier for everyone to use that came from academic study. Let's look at how it's different from old ways of doing things and use it to visualise something from real life.

Similarities and Differences of Chartability

When it comes to traditional data visualisation accessibility standards, they tend to focus on specific technical issues, such as colour contrast or keyboard navigation. Chartability, on the other hand, looks at things in a more complete way. For people with disabilities, it uses a set of rules (testable questions) organised into principles that make sure they can access data (Fizz Studio [2023]).

Here's a table summarizing the key differences:

Feature	Traditional Approach	Chartability
Focus	Technical specifications	User experience and outcomes
Comprehensiveness	Limited scope	Broad range of accessibility factors
Ease of Use	Can be complex	Designed for quick and easy application

Figure 4:Unit 3-1

Chartability Evaluation:

- Relevant Content- The visualisation uses colour and arrows to show how energy flows, but people who are blind or have low vision could benefit from hearing more about it.
- Colour contrast between parts is good, but users who have trouble seeing colours could be given other colour schemes to choose from.
- Practical-The interaction looks like it can be done with a computer, but more testing is needed.
- That makes sense—the visualisation itself is hard to understand. For people with cognitive disabilities, you might want to include a simpler summary or more information.

Critical Findings:

The image is nice to look at and does tell you something, but it could be better to make sure that everyone has a truly good time with the data. Artists can make it easy for people to understand their work by following Chartability's rules.

References

- Fizz Studio. (2023). Chartability - Fizz Studio [Website]. chartability.fizz.studio/

[Reply](#)



Type a reply

Figure 5:Unit 3-2


The answer is creative because it uses the idea of Chartability to look at a real-life visualisation and then gives a critical analysis based on its principles. The answer does more than just restate the

theoretical differences between Chartability and traditional methods, as shown in the table. It also evaluates a specific case of a visualisation.

This real-world use of Chartability shows that you understand it in a way that goes beyond its academic roots. The answer shows how Chartability can be used to make data visualisations better for a wide range of people by showing where the visualisation could be more accessible and improve the user experience.

Lastly, the answer adds something new to the conversation by going beyond theoretical explanations and showing what Chartability means in real life through a study of a real-life case. Taking this method helps us understand the idea and its possible uses better, adding a new point of view to the talk.

1.4. Unit 4 Consolidate

**Farzam FERYDOONI**
24/03/2024, 17:00

Application Programming Interfaces (APIs) have emerged as a powerful tool for data acquisition, offering new avenues for creating compelling visualizations (Manyika et al., 2013). Twitter's recent transition to "X" and the introduction of its Developer Platform exemplify this shift, providing access to targeted data through the X API.

The best thing about APIs is that they make it easy to get detailed, real-time info. APIs give you organised access to specific data points that are useful for visualisations, unlike scraping public data, which can be messy and not always work. Also, a lot of APIs offer real-time changes that let you make dynamic graphs that show the newest trends (Greenfield & Moorjani, 2019). This simplified method can cut down on the time needed to collect data by a large amount (Manyika et al., 2013).

But there are some things that APIs can't do. Cost and lack of data can be problems, since some platforms may charge or require fees to see certain data points (Greenfield & Moorjani, 2019). Also, APIs often have rate limits or limits on how much data can be retrieved in a certain amount of time, which makes large-scale visualisations less possible (Twitter Developer Platform, n.d.). The source of the data can also change its quality and uniformity, so it needs to be carefully cleaned and checked before it can be visualised (Few, 2009).

It's important to think about the goals of the project when looking at APIs for data visualisation. Important things to think about are:

- Assessing whether the required data is accessible through APIs or alternative methods.
- Weighing the potential costs (financial and time-related) of API access against the project's objectives.
- Ensuring data quality by addressing potential inconsistencies within the API data (Few, 2009).

APIs are definitely helpful for data visualisation because they let you access structured data, get changes in real time, and save time (Manyika et al., 2013). But it's important to know what APIs can't do and carefully tailor your method if you want to use them effectively and make interesting, useful visuals.

References:

1. Few, Stephen. *Information Dashboard Design: The Effective Visual Communication of Data*. Sebastopol, CA: O'Reilly, 2009.


Edited by Farzam FERYDOONI on 30 Mar 2024, 10:50:30

Figure 6:Unit 4

The response provides a balanced and critical analysis of using Application Programming Interfaces (APIs) for data acquisition and visualization. It highlights the strengths of APIs, such as real-time information and organized access to data points, but also acknowledges potential drawbacks like cost constraints and data availability restrictions. It also emphasizes the importance of considering project goals and objectives when assessing the viability of API usage. The response highlights the need for a

tailored approach based on specific contexts and the author's ability to apply theoretical concepts to real-world scenarios.

1.5. Unit 5 Consolidate

**Farzam FERYDOONI**
25/03/2024, 14:35

Based on the "COVID-19 Tweets" dataset available on the George Washington University's TweetSets platform and Social Media Data files with 800,000 raw tweets. Here's an analysis of the data characteristics:

Data Types:

- Covid Tweet IDs data: The primary data points seem to be tweet IDs, which are numeric strings (19 Digit integer) identifying individual tweets related to Coronavirus or COVID-19 (Twitter, n.d.).
- Social Media Data has 800,000 raw tweets from Twitter dated April 1st to April 9th 2022 that discuss BTC. In file 4-4 is oriented over crypto currency. It has different columns but only tweet id column is 19 integer digit and tweet content is a string, other columns are empty.

Data Size:

- The dataset description states it contains 354,903,485 tweet IDs(TweetSets - The George Washington University, n.d.).
- In Social Media Data file 4-4 has 93233 tweets but all files have 800,000 raw tweets.

Range of Values:

- Tweet IDs are likely non-repeating unique identifiers. They may follow a specific format or pattern established by Twitter, but this information isn't readily available from the dataset description or even data.
- In Social Media Data file 4-4 has different columns but only tweet id and tweet contents are available. Tweet Id is 19 integer digit and there is no range for it.

Figure 7:Unit 5-1

Data Condition:

- collected between December 3, 2020 and September 14, 2022 from the Twitter API using Social Feed Manager with this tags "coronavirus, COVID-19, epidemiology, pandemic" (TweetSets - The George Washington University, n.d.).
- Social Media Data dataset description doesn't mention any specific condition but based on observation it use cryptocurrencies tag for example "dogecoin, bitcoin, eth and etc"

This datasets description don't mention any specific data cleaning or validation processes.(it's advisable to perform data cleaning and validation before using the data for analysis or visualization (Few, 2009).)

References:

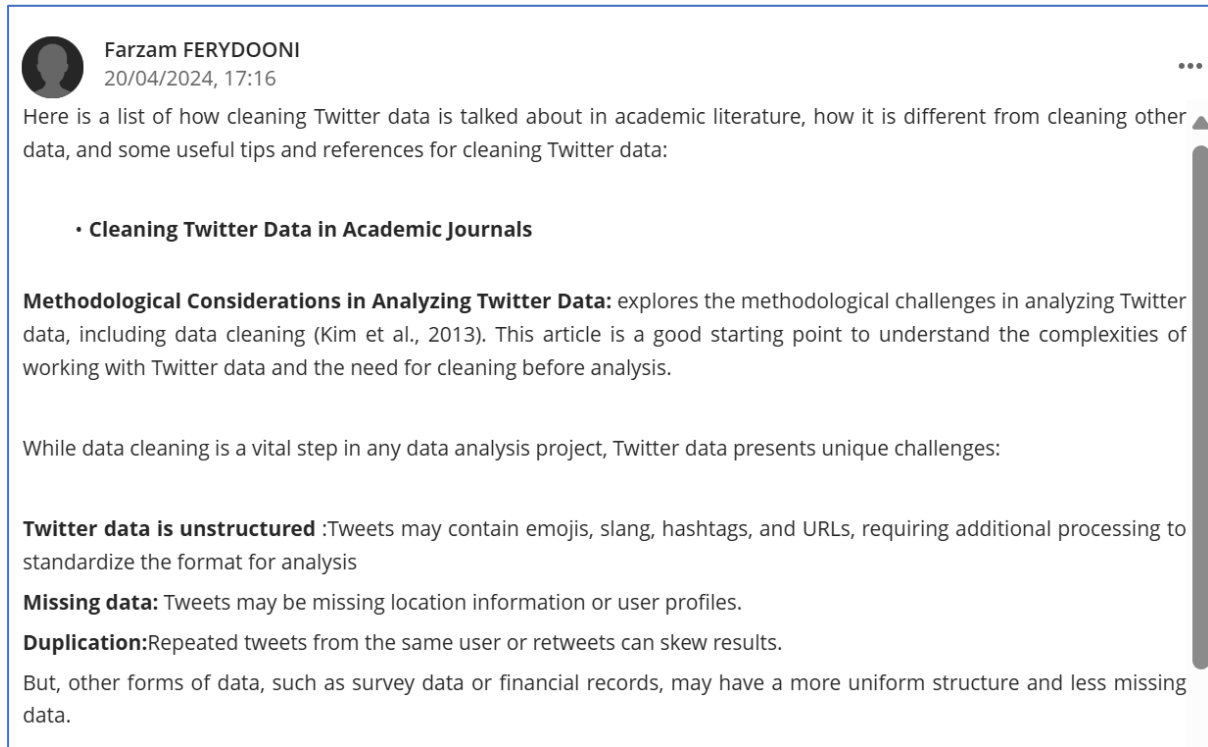
- Few, S. (2009). Information dashboard design: The effective visual communication of data. O'Reilly Media.
- TweetSets - The George Washington University. (n.d.). Select source dataset. [<https://tweetsets.library.gwu.edu/>]
- Twitter. (n.d.). About Tweet IDs. [<https://www.wikihow.com/Find-Your-User-ID-on-Twitter>] (Note: While an official documentation page for Tweet IDs couldn't be located, this source provides context on their function)
- Twitter Developer Platform. (n.d.). Use Cases, Tutorials, & Documentation. [<https://developer.twitter.com/en>]

Figure 8:Unit 5-2

The answer doesn't just give a general outline; instead, it goes into detail about the "COVID-19 Tweets" dataset, giving information that is specific to that data source. The analysis looks at many parts of the information, such as its types, size, range of values, and state, giving a full picture of its structure and what it contains.

Also, It points out the problems with the dataset, like the fact that it doesn't have enough information about how the data is cleaned or checked, and it suggests the right steps to fix these problems. This critical method shows that you know how important good data is and what steps need to be taken to make sure that analysis and visualisation are accurate.

1.6. Unit 6 Consolidate



The image is a screenshot of a tweet from a user named Farzam FERYDOONI, posted on 20/04/2024 at 17:16. The tweet discusses the challenges of cleaning Twitter data and lists several specific issues: unstructured data (emojis, slang, hashtags, URLs), missing data (location, user profiles), and duplication (repeated tweets). It also mentions that other data types like survey data or financial records might have more uniform structures. The tweet is presented in a light blue border with a vertical scrollbar on the right side.

Farzam FERYDOONI
20/04/2024, 17:16

Here is a list of how cleaning Twitter data is talked about in academic literature, how it is different from cleaning other data, and some useful tips and references for cleaning Twitter data:

- **Cleaning Twitter Data in Academic Journals**

Methodological Considerations in Analyzing Twitter Data: explores the methodological challenges in analyzing Twitter data, including data cleaning (Kim et al., 2013). This article is a good starting point to understand the complexities of working with Twitter data and the need for cleaning before analysis.

While data cleaning is a vital step in any data analysis project, Twitter data presents unique challenges:

Twitter data is unstructured :Tweets may contain emojis, slang, hashtags, and URLs, requiring additional processing to standardize the format for analysis

Missing data: Tweets may be missing location information or user profiles.

Duplication:Repeated tweets from the same user or retweets can skew results.

But, other forms of data, such as survey data or financial records, may have a more uniform structure and less missing data.

Figure 9:Unit 6-1

Some Suggestions for Cleaning Twitter Data in social media folder:

Removing irrelevant information: This includes removing URLs, emojis, hashtags, and symbols.

Standardizing text: This can involve converting text to lowercase, removing punctuation, and stemming or lemmatization (reducing words to their base form).

Deduplication: This involves removing duplicate tweets.

Removing empty columns: deleting like geo, media, lang and etc

Dealing with missing data: You can decide to remove tweets with missing data or impute missing values (estimating values based on other data points).

References:

Kim, A. E., Hansen, H. M., Murphy, J., Richards, A. K., Duke, J., Allen, J. A., ... & Fabian, H. (2013). Methodological considerations in analyzing Twitter data. JNCI monographs, 47(1), 140-146.
https://www.researchgate.net/publication/259608350_Methodological_Consideratio

Figure 10:Unit 6-2

This post reply the unique challenges of cleaning Twitter data, citing an article from the Journal of Computer-Mediated Communication, and highlights the methodological considerations and complexities involved in analyzing it. It also distinguishes between cleaning Twitter data and other forms of data, such as survey data or financial records, as Twitter data is often unstructured and may contain elements like emojis, slang, hashtags, and URLs. The response also highlights the potential for missing data and duplication, and provides practical suggestions for cleaning Twitter data, drawing from real-world practices. The response is supported by relevant references, adding credibility to the information presented and facilitating further exploration of the discussed topics. This comprehensive approach deepens the understanding of cleaning Twitter data and highlights the unique considerations and techniques required for effective data cleaning in this domain.

2. Portfolio Activity Two: Accessibility of Data Visualisation

Accessibility in data visualisations is very important for global organisations to make choices that include and reflect the needs of all parties. The tools from Unit 3 show us a lot about how important accessibility is and give us tips on how to make data visualisations that people with eye impairments or other disabilities can use.

2.1. Visualization's clarity

A data visualization's ability to communicate ideas clearly is essential. For people to comprehend the message without any confusion, data points must be clearly labelled and presented clearly (Lonsdale, 2022). The user's capacity to properly evaluate the data in Unit 3 is prevented if the visualisation is unclear or provides information in a confusing way. For Example based on (*Visual Best Practices*, no date)

- **Chart Choice:** Selecting the appropriate chart type in Tableau is essential for accessibility, as it guarantees that the data visualisation is comprehensible and intuitive for all users, including those with disabilities.
- **Layout:** it ensures that the dashboard is organized in a way that can be easily navigated and understood by all users.

2.2. Including a Variety of Audiences

Data visualisation that is inclusive takes people with problems into account. Accessibility requires features including resizable text choices, suitable colour contrast ratios, and alternate text for pictures (Kirk, 2019). The visualisation in Unit 3 has to be assessed to see if it uses these aspects. To reduce colour vision problems, for example, use patterns or textures in addition to colours and make sure that colours contrast sharply (*Visual Best Practices*, no date; *Best Practices for Designing Accessible Views*, no date).

2.3. Ease of Interpretation

Users must be able to easily navigate the visualization and extract meaningful insights. Interactive features such as keyboard navigation and descriptive text accompanying visual elements enhance interpretability, especially for users who rely on assistive technologies like screen readers (Kirk, 2019). If Unit 3's visualization lacks such interactive features, it may not be fully accessible.

2.4. Importance of Accessibility in Global Organizations

Global companies working in many cultural and financial environments need data visualisation to be accessible. Effective data engagement by decision-makers at all levels is ensured by accessible visualisations, which results in well-informed decisions that benefit the company and its stakeholders (Kirk, 2019). For instance, a multinational company that analyses sales data in several locations needs easily understandable graphics so that executives, at all skill levels, can understand important trends and patterns and make strategic decisions. Furthermore, promoting diversity and equality in decision-making processes inside organisations is promoted by an inclusive culture (Lonsdale, 2022).

2.5. Accessibility in Practice

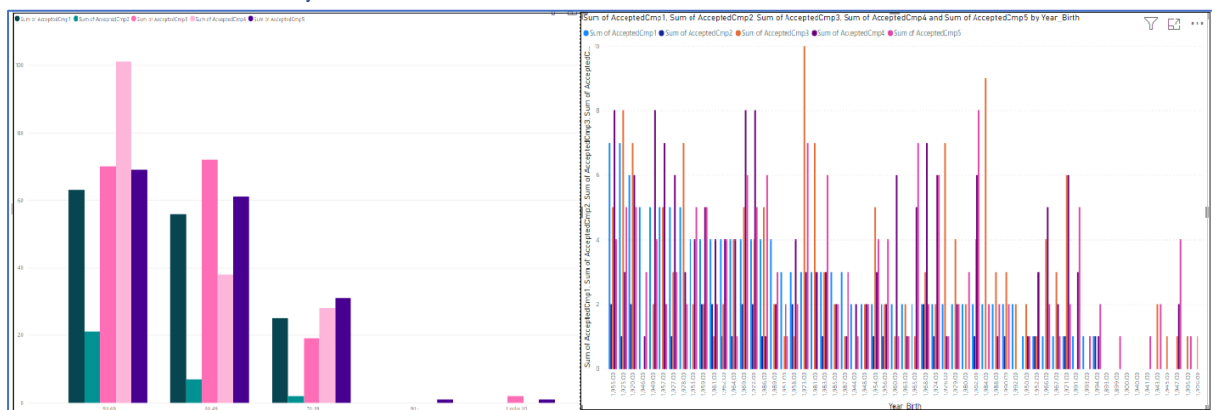


Figure 11 Compare Two Plot

One of the key elements is selecting the kind of chart; in the example above (Figure 11), we selected a bar plot rather than a bunch of lines (Line Plot). Information grouping is one more crucial aspect that may greatly aid in data distinction (Differ From left and right side Plot).

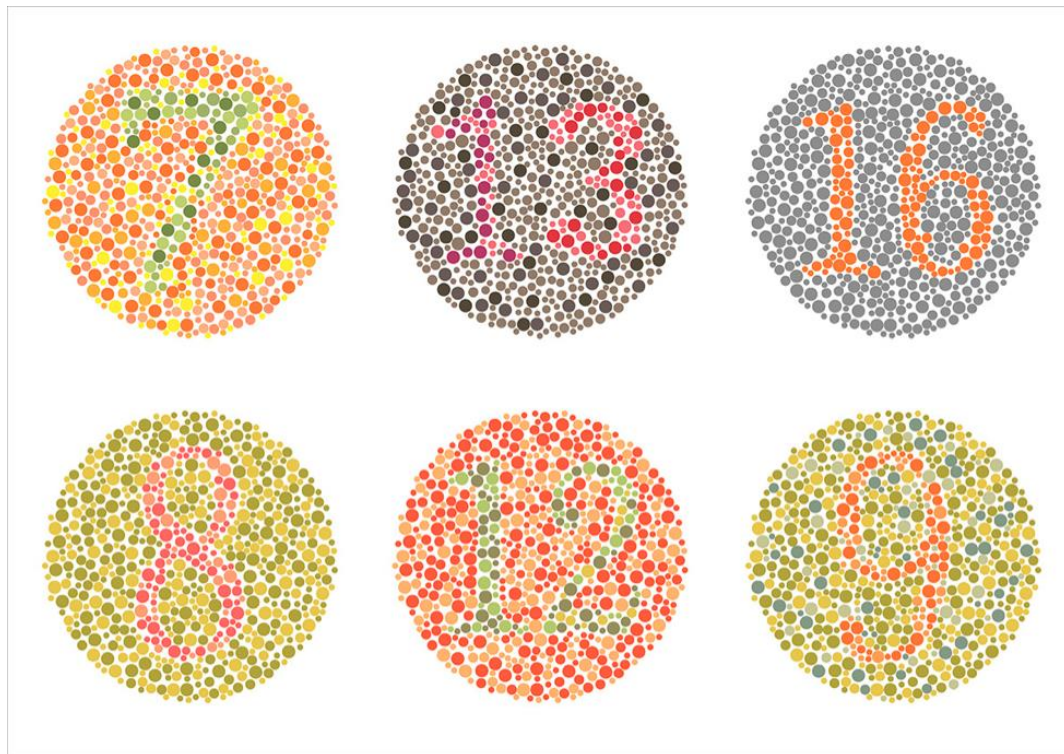


Figure 12: (5 tips on designing colour-blind-friendly visualizations, no date)

Selection of suitable palette of colours is crucial because of the high rate of blind colour ('Colour Blindness in Clinton in Clinton, CT', no date) and elderly visually impaired persons, as stated in Tableau report "5 tips on designing colour-blind-friendly visualisations".

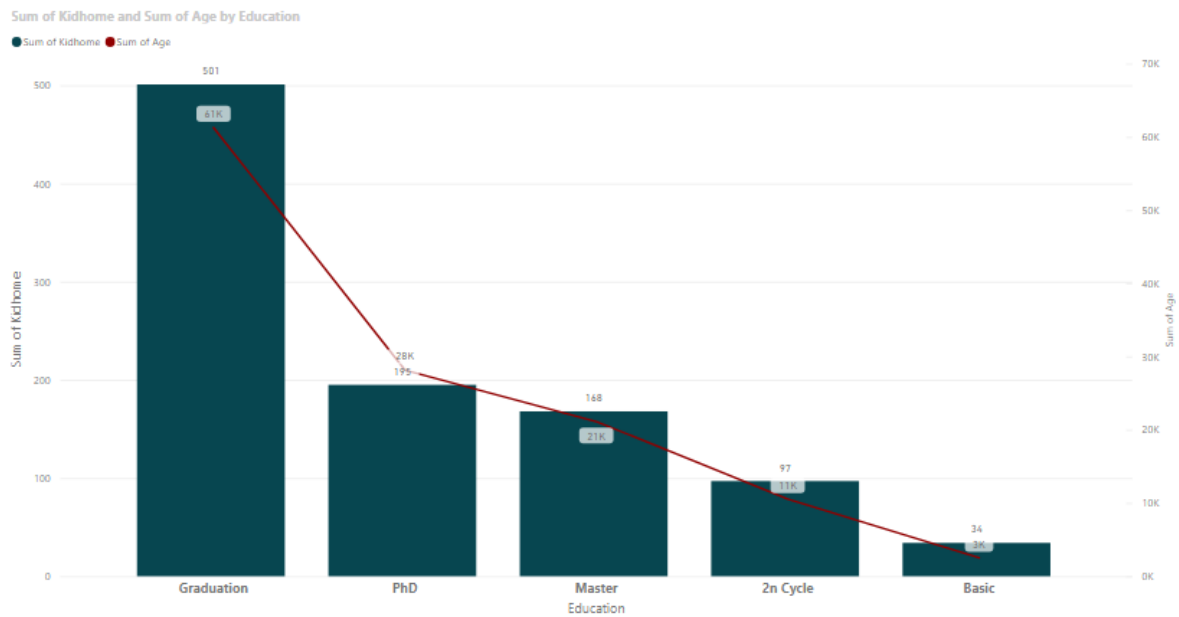


Figure 13: Visualization Sample

Effective visual communication for individuals with disabilities can be achieved through separating diagram elements, using descriptive labels, employing suitable font styles and sizes (particularly bold text), and inserting information directly on lines and bars, as illustrated in the previous figure.

2.6. Conclusion

Effective communication and decision-making depend on the data visualisation in Unit 3 being easily available. All users, including those with impairments, may benefit much from visualisations that are made clear, inclusive, and easy to understand. For multinational companies, easily available data visualisations are a strategic benefit that encourages inclusive corporate culture and educated decision-making in addition to being a compliance need.

3. Portfolio Activity Three: Comparison of Tableau and Power BI Visualisations

Tableau and Power BI are regularly the subject of discussion (Kim, 2024) when it comes to data visualisation tools. Comparing these two software tools, first of all my experience is, Tableau is preferred by beginners because of its user-friendly interface but Power BI is most suitable for enterprise because of its integration with Microsoft products.

3.1. Popularity

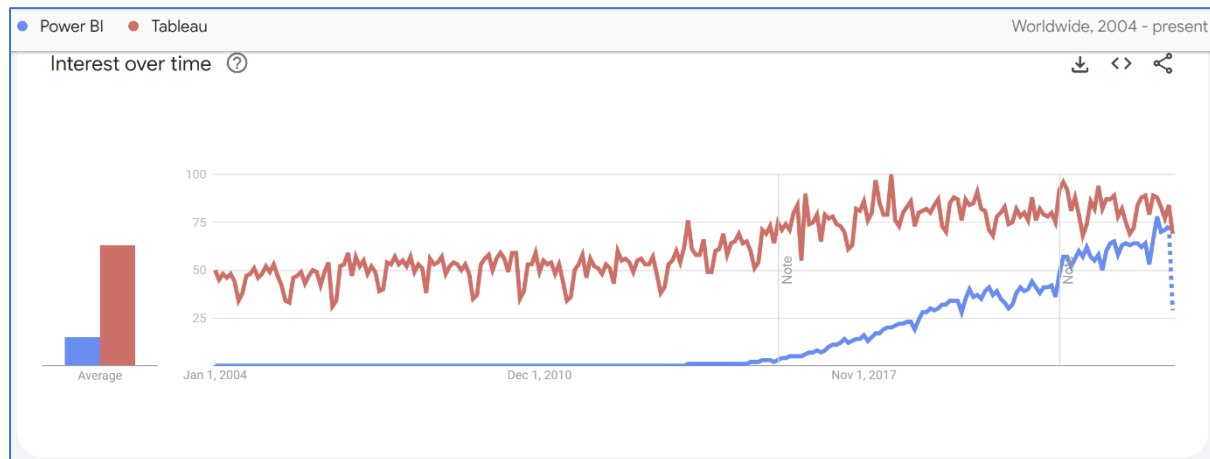


Figure 14: google trend (Google Trends, 2024)

Google Trends indicates that Power BI's popularity has been rising since Jan 2016, matching that of Tableau, based on Figure 14.



Figure 15:Gartner (Microsoft Power BI, Leader in Gartner's Magic Quadrant 2023 | LinkedIn, no date)

In the other side last year Power BI was the leader of Data Visualization software(Figure 15)

3.2. Number of Functions verity of Type variable Supported in Functions

Regarding the number of functions and types of data used in the functions (time, text and date), Power BI has very wide support.

3.3. Data Sources

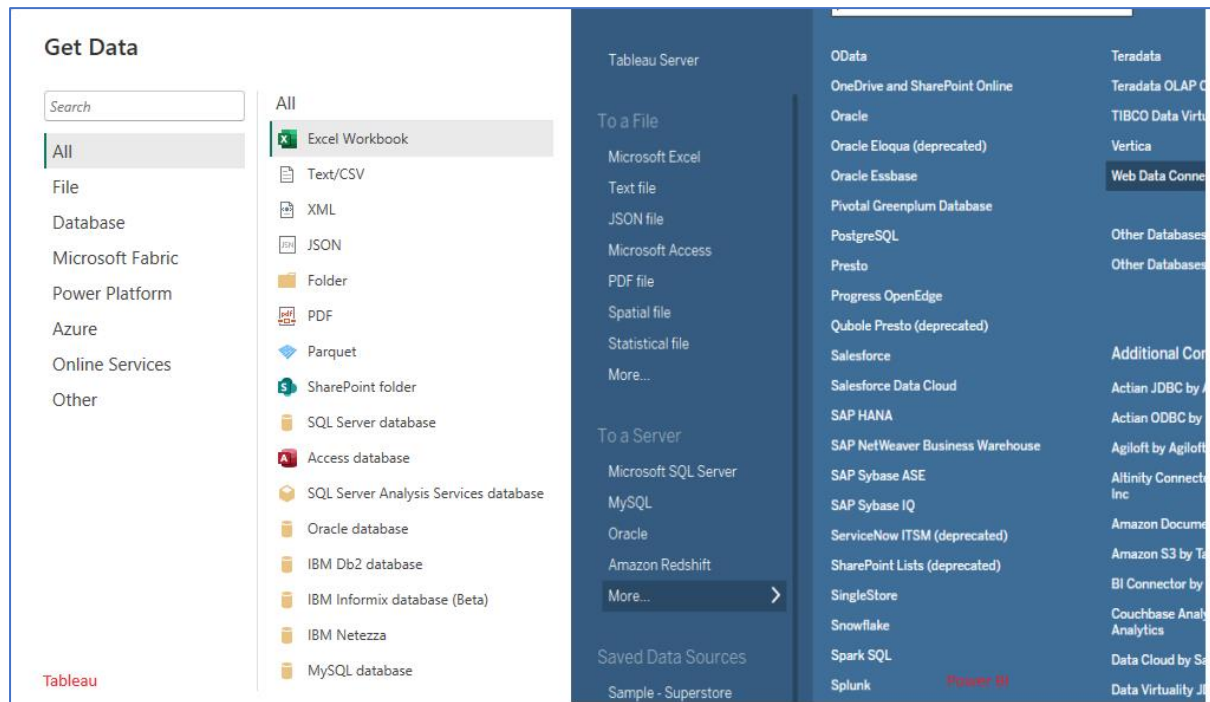


Figure 16

Both software use different sources to enter data, PowerBI has extensive support for Microsoft products, and one of the features of PowerBI is the ability to directly receive information from some websites without the need for API.

Another important feature is that PowerBI can import easily email (exchange and non exchange) attachment to analysis.

3.4. Data preprocessing

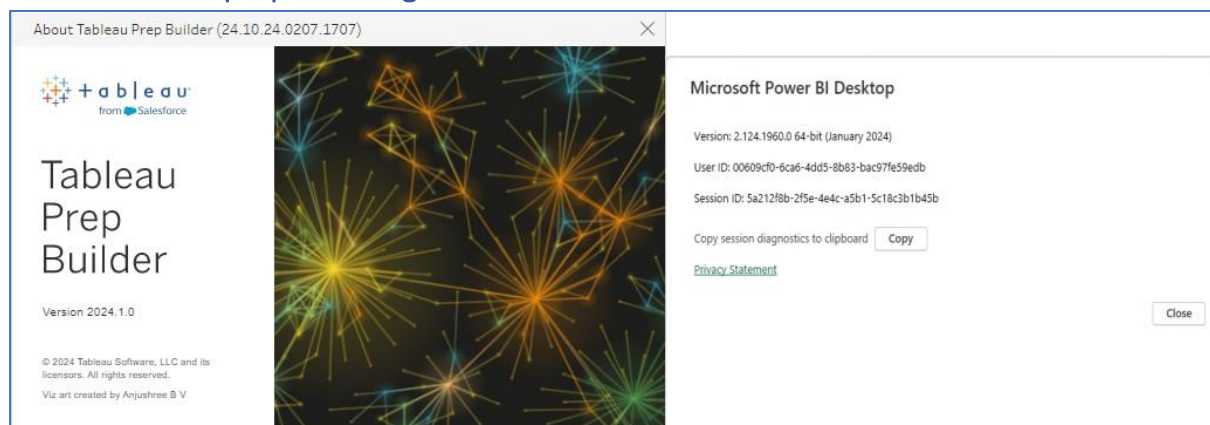


Figure 17: Tableau Prep- Power Query

To perform preprocessing in PowerBI, PowerQuery, which is internally implemented in this software, is used, where there are extensive functions with form of visual and code commend for preprocessing. In the Tableau, a separate software (Tableau Prep) is used for initial processing. It seems that this program is at the beginning of the development process.

3.5. User Interface

Based on the experience gained in LinkedIn courses and some internet references (Edmond and Crabtree, 2023), the tableau user interface is easier to use, especially for first-time users. so that the user's demands can be easily found and used in the environment.

Table 1

Factor	Microsoft Power BI	Tableau
Star Rating	4.5	4.4
Largest Market Segment	Enterprise (44.4% of reviews)	Enterprise (45.4% of reviews)
Entry Level Price	Free	\$70.00
Total Number of Reviews	1065	1992
General Ratings		
Meets Requirements	8.8	8.8
Ease of Use	8.3	8.2
Ease of Setup	8.6	8.3
Ease of Admin	8.3	8.1
Quality of Support	8.3	8.4
Product Direction (% positive)	9.1	8.8
Reports		
Reports Interface	8.8	8.8
Steps to Answer	8.3	8.3
Graphs and Charts	9	9.2
Score Cards	8.8	8.6
Dashboards	9	9.1
Data Preparation		
Connectors	8.8	8.8
Data Governance	8.7	8.6
Data Modeling and Blending		
Data Querying	8.9	8.6
Data Filtering	9	8.9
Data Blending	8.6	8.8
Generative AI		
AI Text Generation	7.8	8.4
AI Text Summarization	7.9	8.4
Self Service		
Calculated Fields	8.5	8.6
Data Column Filtering	8.8	8.8
Data Discovery	8.5	8.5
Search	8.3	8.2
Collaboration / Workflow	8.2	8.2
Automodeling	8.1	7.9
Advanced Analytics		
Predictive Analytics	8.2	7.9
Data Visualization	9.1	9.2
Big Data Services	8.4	8.2
Building Reports		
Data Transformation	8.7	8.4
Data Modeling	8.6	8.1
WYSIWYG Report Design	8.4	8.5
Integration APIs	8.4	8.3

One of the reputable websites that ranks the relevant software using the views of experts in each sector is G2 site. More people clearly choose for tableau, as the table below makes clear. It explains for this software's popularity, and the poll indicates that these two programmes don't vary much that much.

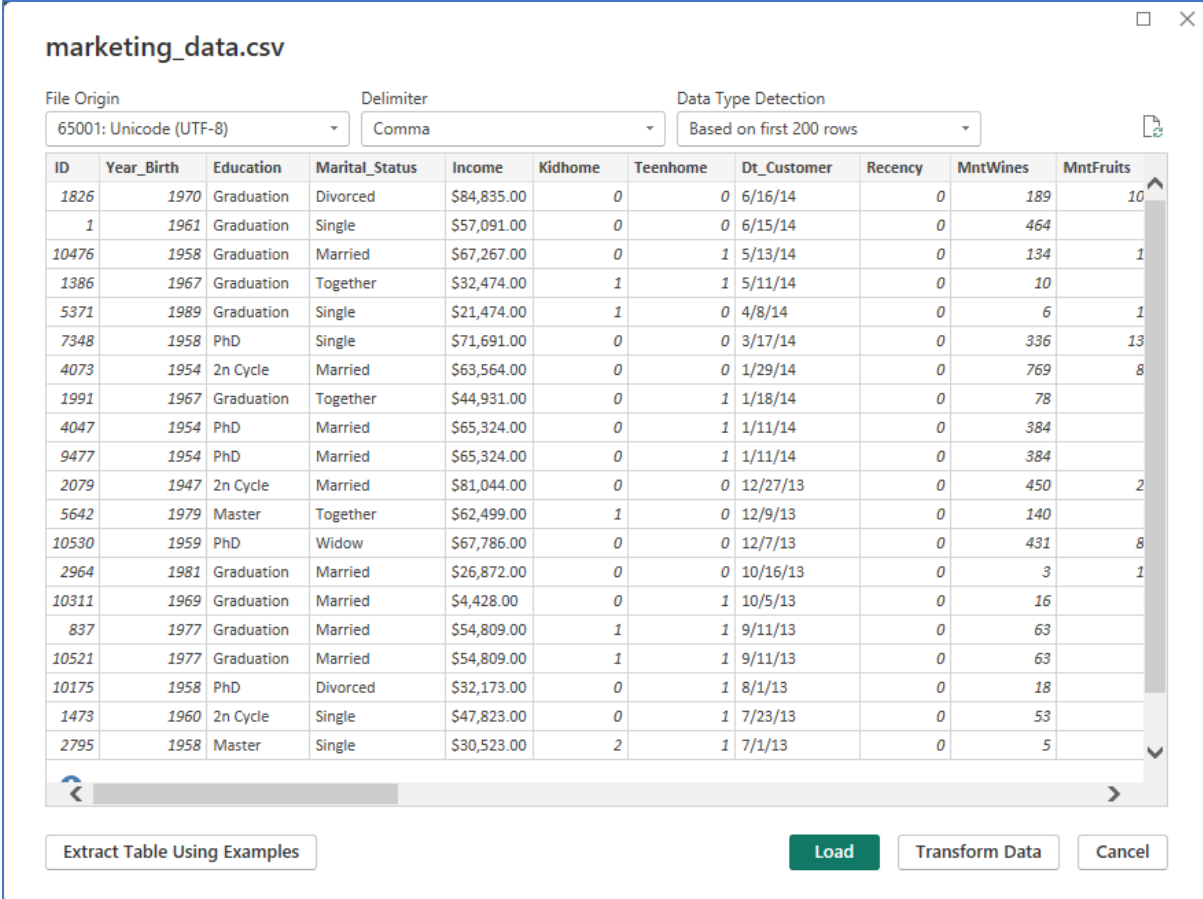
3.6. Conclusion

The study notes Power BI as a more complete and potent tool, especially for enterprise-level data analysis and visualisation jobs, even if Tableau's user-friendly interface is acknowledged, especially for beginner. Power BI is becoming more and more popular. It was ranked as the top solution in Gartner's Magic Quadrant for Data Visualisation in 2023. It also provides more features and support for different data types, easily integrates with Microsoft products, can import data directly from a variety of sources (including email attachments), and has strong data preprocessing capabilities with Power Query integrated into the programme. My report says that Power BI is the better option, matching your taste, despite Tableau's apparent popularity among users as shown by the G2 site rating, given its extensive capabilities and appropriateness for enterprise-level data analysis and visualisation needs.

4. Portfolio Activity Four

4.1. Preparing Data for EDA

This dataset is provided by Kaggle and is related to a marketing campaign in different countries(Marketing Campaigns_EDA, no date).



The screenshot shows a data loading window titled "marketing_data.csv". It includes settings for File Origin (65001: Unicode (UTF-8)), Delimiter (Comma), and Data Type Detection (Based on first 200 rows). Below these settings is a table preview with 11 columns: ID, Year_Birth, Education, Marital_Status, Income, Kidhome, Teenhome, Dt_Customer, Recency, MntWines, and MntFruits. The table contains 20 rows of data. At the bottom, there are buttons for "Extract Table Using Examples", "Load", "Transform Data", and "Cancel".

ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Dt_Customer	Recency	MntWines	MntFruits
1826	1970	Graduation	Divorced	\$84,835.00	0	0	6/16/14	0	189	10
1	1961	Graduation	Single	\$57,091.00	0	0	6/15/14	0	464	
10476	1958	Graduation	Married	\$67,267.00	0	1	5/13/14	0	134	1
1386	1967	Graduation	Together	\$32,474.00	1	1	5/11/14	0	10	
5371	1989	Graduation	Single	\$21,474.00	1	0	4/8/14	0	6	1
7348	1958	PhD	Single	\$71,691.00	0	0	3/17/14	0	336	13
4073	1954	2n Cycle	Married	\$63,564.00	0	0	1/29/14	0	769	8
1991	1967	Graduation	Together	\$44,931.00	0	1	1/18/14	0	78	
4047	1954	PhD	Married	\$65,324.00	0	1	1/11/14	0	384	
9477	1954	PhD	Married	\$65,324.00	0	1	1/11/14	0	384	
2079	1947	2n Cycle	Married	\$81,044.00	0	0	12/27/13	0	450	2
5642	1979	Master	Together	\$62,499.00	1	0	12/9/13	0	140	
10530	1959	PhD	Widow	\$67,786.00	0	0	12/7/13	0	431	8
2964	1981	Graduation	Married	\$26,872.00	0	0	10/16/13	0	3	1
10311	1969	Graduation	Married	\$4,428.00	0	1	10/5/13	0	16	
837	1977	Graduation	Married	\$54,809.00	1	1	9/11/13	0	63	
10521	1977	Graduation	Married	\$54,809.00	1	1	9/11/13	0	63	
10175	1958	PhD	Divorced	\$32,173.00	0	1	8/1/13	0	18	
1473	1960	2n Cycle	Single	\$47,823.00	0	1	7/23/13	0	53	
2795	1958	Master	Single	\$30,523.00	2	1	7/1/13	0	5	

Figure 18

At first stage, the information is entered in the PowerBI software. For some reports, Python has also been used for data analysis and some Visualization. In the following, the structure of the columns and their types are specified at Figure 19. It has 2240 rows × 28 columns.

Column Name	Description	Data Type
ID	Unique customer identifier	Integer
Year_Birth	Year of customer's birth	Integer
Education	Level of education (e.g., "Graduate Degree", "High School")	Object
Marital_Status	Marital status (e.g., "Married", "Single")	Object
Income	Annual income	Object
Kidhome	Number of children living in the household	Integer
Teenhome	Number of teenagers living in the household	Integer
Dt_Customer	Date the customer became a customer (format may vary)	Object
Recency	Number of days since the customer's last purchase	Integer
MntWines	Total amount spent on wine purchases	Integer
MntFruits	Total amount spent on fruit purchases	Integer
MntMeatProducts	Total amount spent on meat product purchases	Integer
MntFishProducts	Total amount spent on fish product purchases	Integer
MntSweetProducts	Total amount spent on sweet product purchases	Integer
MntGoldProds	Total amount spent on gold product purchases	Integer
NumDealsPurchases	Number of purchases made with a discount	Integer
NumWebPurchases	Number of purchases made through the company website	Integer
NumCatalogPurchases	Number of purchases made using a catalog	Integer
NumStorePurchases	Number of purchases made directly in stores	Integer
NumWebVisitsMonth	Number of visits to the company website in the last month	Integer
AcceptedCmpX	1 if customer accepted the offer in campaign X (X = 1 to 5), 0 otherwise	Integer
Response	1 if customer accepted the offer in the last campaign, 0 otherwise	Integer
Complain	1 if customer complained in the last 2 years, 0 otherwise	Integer
Country	Customer's country of residence	Object

Figure 19: Columns Information

[4]:	df.isnull().sum()
[4]:	<pre> ID 0 Year_Birth 0 Education 0 Marital_Status 0 Income 24 Kidhome 0 Teenhome 0 Dt_Customer 0 Recency 0 MntWines 0 MntFruits 0 MntMeatProducts 0 MntFishProducts 0 MntSweetProducts 0 MntGoldProds 0 NumDealsPurchases 0 NumWebPurchases 0 NumCatalogPurchases 0 NumStorePurchases 0 NumWebVisitsMonth 0 AcceptedCmp3 0 AcceptedCmp4 0 AcceptedCmp5 0 AcceptedCmp1 0 AcceptedCmp2 0 Response 0 Complain 0 Country 0 dtype: int64 </pre>
	<div> <div>+ Code</div> <div>+ Markdown</div> </div>

Figure 20: Dnusual Data-Part 1

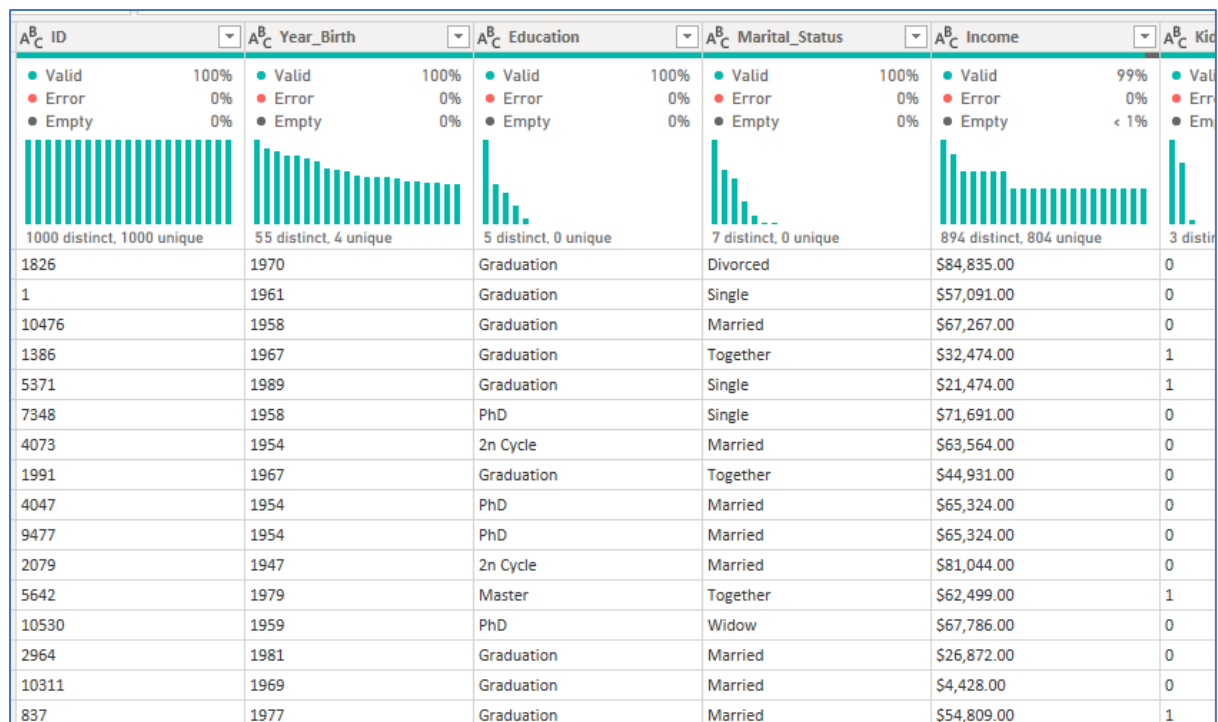


Figure 21: Unusual Data-Part 2

The following procedures are planned for Data Preparation for EDA stage.

1. Delete empty lines.
2. Delete duplicate rows and columns.
3. According to the image number 4 and 5 the 'Income' column is considered as an object, first in the Power Query, the \$ sign should be removed from the beginning of all the figures in this column and then converted into integer.
4. There are 24 empty numbers in the 'income' column, which we change to null.
5. In 'Dt_Customer' has different date format 8/25/12 and 12/18/2012, most converted in same format.
6. Converting 'Year_birth' and 'Dt_Customer' to date format.

Feature Engineering

Based On demand for analysis of this data set we need to create below feature:

1. 'Age' column based on current date and 'Year_birth'.
2. 'Age Group' with classifying 'Age' column to different categories.
3. 'Income Group' with classifying 'Income' column to different categories.

4.2. Statistical Analysis

```
[6]: df.describe().T# describe df horizontally
```

	count	mean	std	min	25%	50%	75%	max
ID	2240.0	5592.159821	3246.662198	0.0	2828.25	5458.5	8427.75	11191.0
Year_Birth	2240.0	1968.805804	11.984069	1893.0	1959.00	1970.0	1977.00	1996.0
Kidhome	2240.0	0.444196	0.538398	0.0	0.00	0.0	1.00	2.0
Teenhome	2240.0	0.506250	0.544538	0.0	0.00	0.0	1.00	2.0
Recency	2240.0	49.109375	28.962453	0.0	24.00	49.0	74.00	99.0
MntWines	2240.0	303.935714	336.597393	0.0	23.75	173.5	504.25	1493.0
MntFruits	2240.0	26.302232	39.773434	0.0	1.00	8.0	33.00	199.0
MntMeatProducts	2240.0	166.950000	225.715373	0.0	16.00	67.0	232.00	1725.0
MntFishProducts	2240.0	37.525446	54.628979	0.0	3.00	12.0	50.00	259.0
MntSweetProducts	2240.0	27.062946	41.280498	0.0	1.00	8.0	33.00	263.0
MntGoldProds	2240.0	44.021875	52.167439	0.0	9.00	24.0	56.00	362.0
NumDealsPurchases	2240.0	2.325000	1.932238	0.0	1.00	2.0	3.00	15.0
NumWebPurchases	2240.0	4.084821	2.778714	0.0	2.00	4.0	6.00	27.0
NumCatalogPurchases	2240.0	2.662054	2.923101	0.0	0.00	2.0	4.00	28.0
NumStorePurchases	2240.0	5.790179	3.250958	0.0	3.00	5.0	8.00	13.0
NumWebVisitsMonth	2240.0	5.316518	2.426645	0.0	3.00	6.0	7.00	20.0
AcceptedCmp3	2240.0	0.072768	0.259813	0.0	0.00	0.0	0.00	1.0
AcceptedCmp4	2240.0	0.074554	0.262728	0.0	0.00	0.0	0.00	1.0
AcceptedCmp5	2240.0	0.072768	0.259813	0.0	0.00	0.0	0.00	1.0
AcceptedCmp1	2240.0	0.064286	0.245316	0.0	0.00	0.0	0.00	1.0
AcceptedCmp2	2240.0	0.013393	0.114976	0.0	0.00	0.0	0.00	1.0
Response	2240.0	0.149107	0.356274	0.0	0.00	0.0	0.00	1.0
Complain	2240.0	0.009375	0.096391	0.0	0.00	0.0	0.00	1.0

Figure 22:Statistical Description

Table 2

Column Name	Finding
Year_Birth	Average birth year: 1968 (suggests middle-aged customer base). Range: 1893-1996 (wide age distribution).
Kidhome	Average: Less than 1 child at home.
Teenhome	Average: About half a teenager at home.
Kidhome & Teenhome (combined)	Many customers may have few or no children living at home.
Recency	Average purchase every 49 days. Range: 0-99 days (varying engagement levels).
MntWines	Highest average spending category (303.9 monetary units). Wine is a significant expenditure.
MntMeatProducts	Second-highest spending category (166.9 monetary units).
MntFruits, MntFishProducts, MntSweetProducts, MntGoldProds	Lower average spending, but still significant. Diverse customer interests across categories.
NumDealsPurchases	Average: 2.325 purchases with discounts. Discounts are a notable factor in purchasing behavior.
NumWebPurchases, NumCatalogPurchases, NumStorePurchases	Store purchases are most common (average: 5.794). Web and catalog purchases are also significant. Highlights a multi-channel sales strategy.
NumWebVisitsMonth	Average: 5.316 visits per month. Online presence and engagement are crucial.
AcceptedCmp1 to AcceptedCmp5	Campaign acceptance rates: 7-8% (relatively low). Room for improvement in marketing campaign effectiveness.
Response	General campaign response rate: 14.9%. Potential for growth in customer engagement.
Complain	Very low complaint rate (0.9%). High customer satisfaction or possibly underreporting of issues.

Summary of Important Finding is:

- Customer Demographics: Predominantly middle-aged with few children or teenagers at home.
- Purchase Behaviour:
 - High spending on wine and meat products.
 - Significant engagement through various purchase channels (store, web, catalogue).
- Marketing and Campaigns:
 - Low but notable acceptance rates for campaigns.
 - Opportunities to enhance campaign effectiveness and customer engagement.
- Customer Satisfaction: Very low complaint rates, indicating generally satisfied customers.

4.3. Univariate Analysis

Outlier Detection-Box Plot

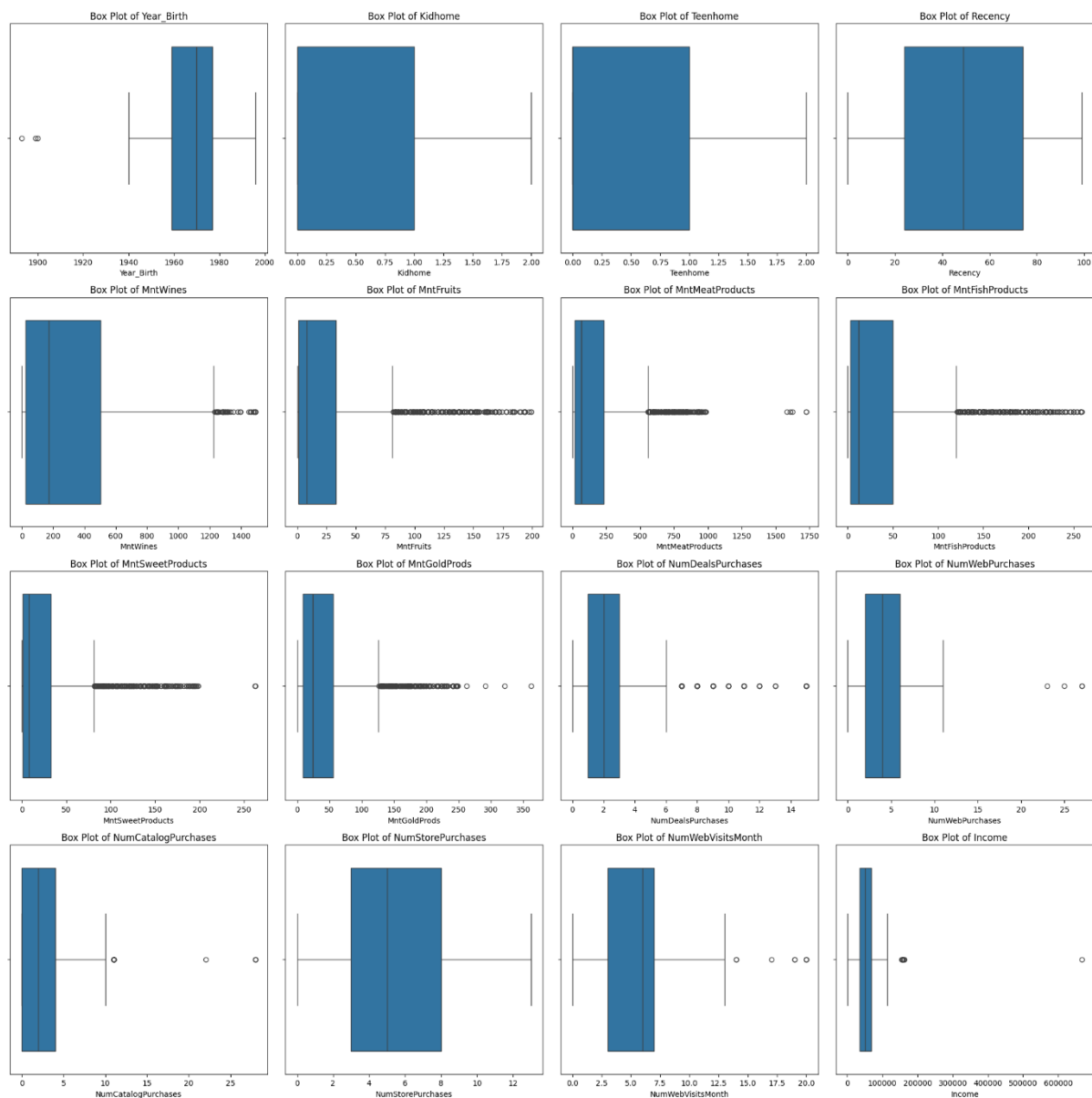


Figure 23:BoxPlot

Table 3

Column Name	Finding (Median & IQR)
Year_Birth	Median: 1970 (IQR not provided). Outliers on both ends (very young & old customers).
Kidhome	Most have 0 or 1 child. IQR: 0-1 (few with 2 children).
Teenhome	Most have 0 or 1 teenager. IQR: 0-1 (very few with 2 teenagers).
Recency	Median: 50 days. IQR: 24-74 days (varied engagement). No significant outliers.
MntWines	Median: 173.5. IQR: 23.75-504.25 (high outliers: heavy spenders).
MntFruits	Median: 6. IQR: 1-33 (several high outliers).
MntMeatProducts	Median: 232. IQR: 16-232 (many high outliers: significant spenders).
MntFishProducts	Median: 12. IQR: 3-50 (several high outliers: higher spending).
MntSweetProducts	Median: 8. IQR: 1-33 (several high outliers: significant spending).
MntGoldProds	Median: 24. IQR: 9-90 (several high outliers: significant spending).
NumDealsPurchases	Median: 2. IQR: 1-3 (outliers up to 15 purchases with deals).
NumWebPurchases	Median: 4. IQR: 2-6 (outliers up to 27 web purchases).
NumCatalogPurchases	Median: 2. IQR: 1-4 (outliers up to 28 catalog purchases).
NumStorePurchases	Median: 6. IQR: 3-8 (outliers up to 13 store purchases).
NumWebVisitsMonth	Median: 5. IQR: 3-7 (outliers up to 20 visits per month).
Income	Median: 51381. IQR: 35380-68487 USD (several high-income outliers).

Summary of Key Insights is:

- **Customer Demographics:**
 - A broad range of ages, with most customers being middle-aged.
 - Few customers have more than one child or teenager at home.
- **Purchasing Behavior:**
 - Significant spending on wine and meat products, with high variability in other product categories.
 - Customers engage in purchases through multiple channels, with notable use of web, catalog, and store purchases.
- **Engagement and Recency:**
 - Variable engagement levels, with a median recency of 50 days.
 - Web visits vary widely, with some customers visiting the website frequently.
- **Income Distribution:**
 - A wide range of incomes, with some customers having significantly higher incomes.

Correlation Analysis-Heat Map

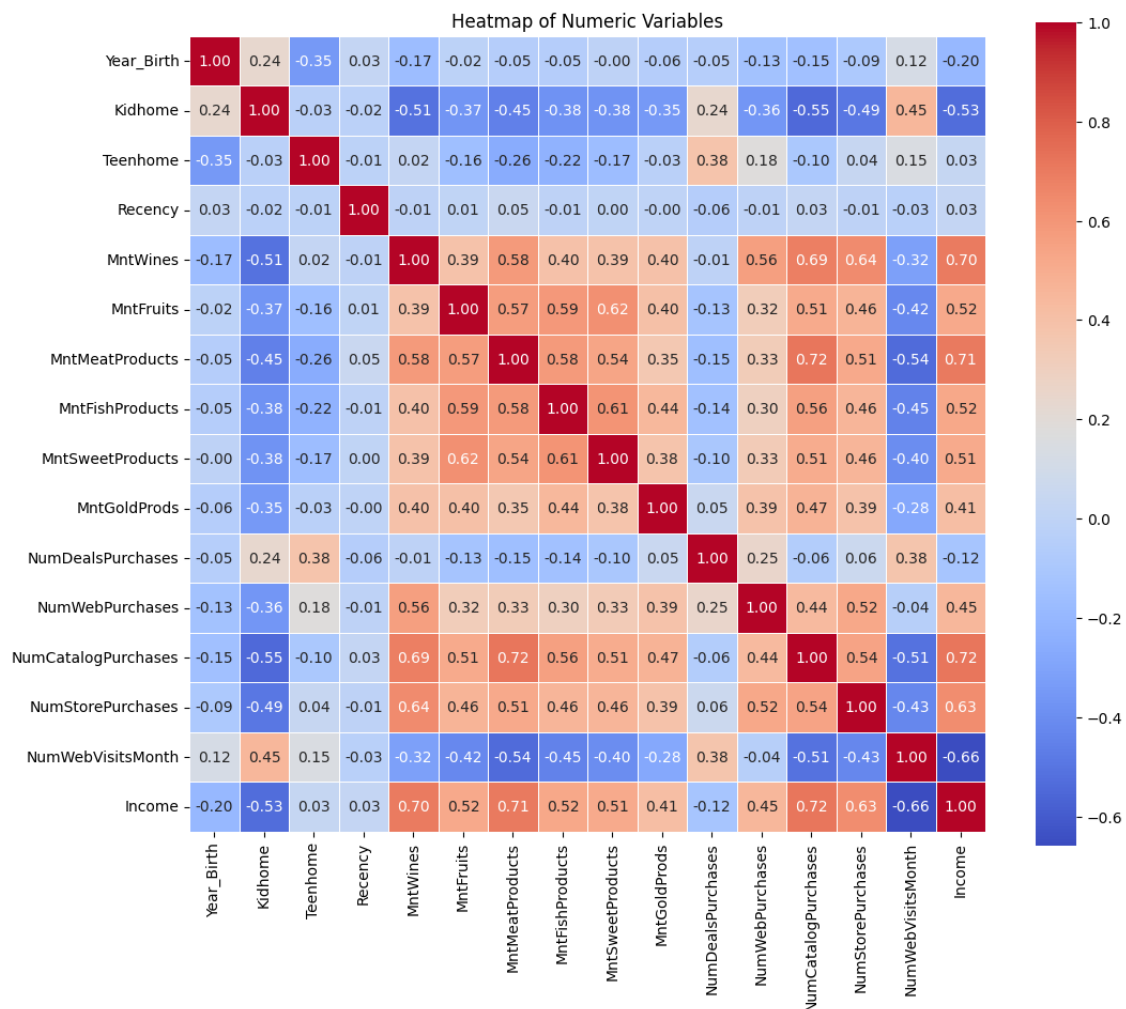


Figure 24:Heatmap

Table 4

Variable	Correlations	Finding
Age (Year_Birth)	Teenhome (-0.35), Kidhome (-0.24), MntWines (-0.17), MntMeatProducts (-0.05)	Older customers tend to have fewer children/teenagers at home and spend slightly less on some products.
Kidhome & Teenhome	Kidhome & Teenhome (0.24), Kidhome & MntWines (-0.51), Kidhome & MntMeatProducts (-0.45)	Households with children tend to have teenagers, but spend less on wine and meat products.
Spending Categories	MntWines & MntMeatProducts (0.58), MntWines & MntFishProducts (0.40), etc.	Customers who spend more on one category tend to spend more on others.
Number of Purchases	NumWebPurchases & NumCatalogPurchases (0.33), NumWebPurchases & NumStorePurchases (0.25)	Customers who use one purchase channel are likely to use others as well.
Website Visits	NumWebVisitsMonth & NumWebPurchases (0.44), NumWebVisitsMonth & Income (-0.66)	More website visits lead to more web purchases. Higher-income customers visit the website less frequently.
Income	Income & all spending categories (positive), Income & NumCatalogPurchases (0.72), Income & NumWebPurchases (0.45), Income & NumWebVisitsMonth (-0.66)	Higher income is associated with more spending across categories and catalog/web purchases, but less frequent website visits.

Summary of Important Insights are:

- **Customer Demographics:**

- Older customers tend to have fewer children and teenagers at home.
- Households with children and teenagers tend to spend less on products like wine and meat.

- **Spending Behavior:**
 - Strong correlations between spending categories suggest that high spenders in one category are likely to be high spenders in others.
 - Higher income correlates with higher spending across all categories.
- **Purchasing Channels:**
 - Positive correlations between different purchase channels indicate multi-channel engagement.
 - Website visits strongly correlate with web purchases but negatively with income, suggesting higher-income customers visit the website less frequently.
- **Income Influence:**
 - Higher income is associated with higher spending and more purchases through catalogs and web channels.
 - Lower frequency of web visits among higher-income customers.

Pie charts of Education, Country, and Marital State

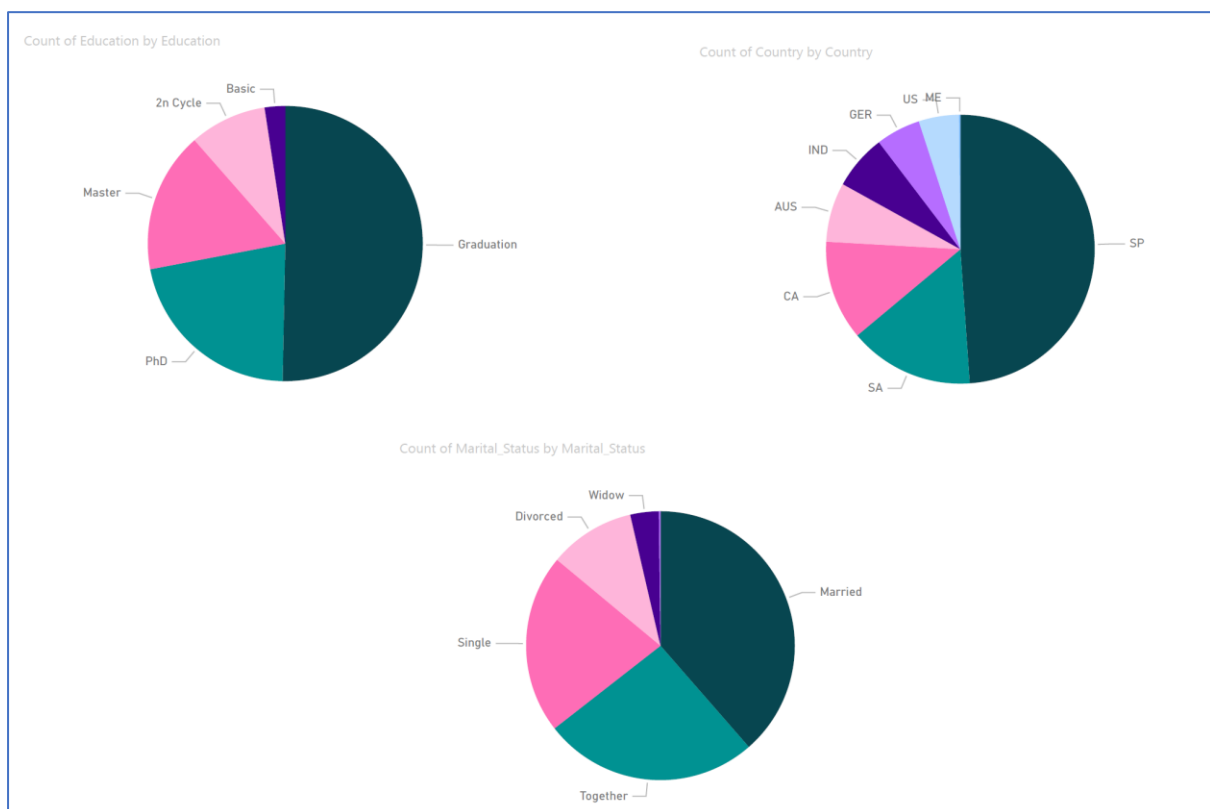


Figure 25:Count Education, Countries, and Marital

The distribution of buyer among different groups is shown in a series of pie charts in the visualisation that is provided. Each pie chart provides details about a certain feature, such as:

- **Levels of Education:**

The pupils' greatest level of education is displayed in this chart. The biggest group, over half (48.88%), having a Specialisation. Additional segments consist of Basic education (11.96%), Graduation (7.14%), and the percentages for Master's, PhD.

- **Countries Represented:**

The greatest proportion of students, or 50.31 percent, come from the United States. India is home to 15.04% of students worldwide.

- **Marital Status:**

This is the biggest marital status category, with over a third (38.57%) of the students married. Significant segments are additionally unmarried (25.89%) or part of a domestic partnership (10.36% marked as "Together").

Sales Amount Per Product

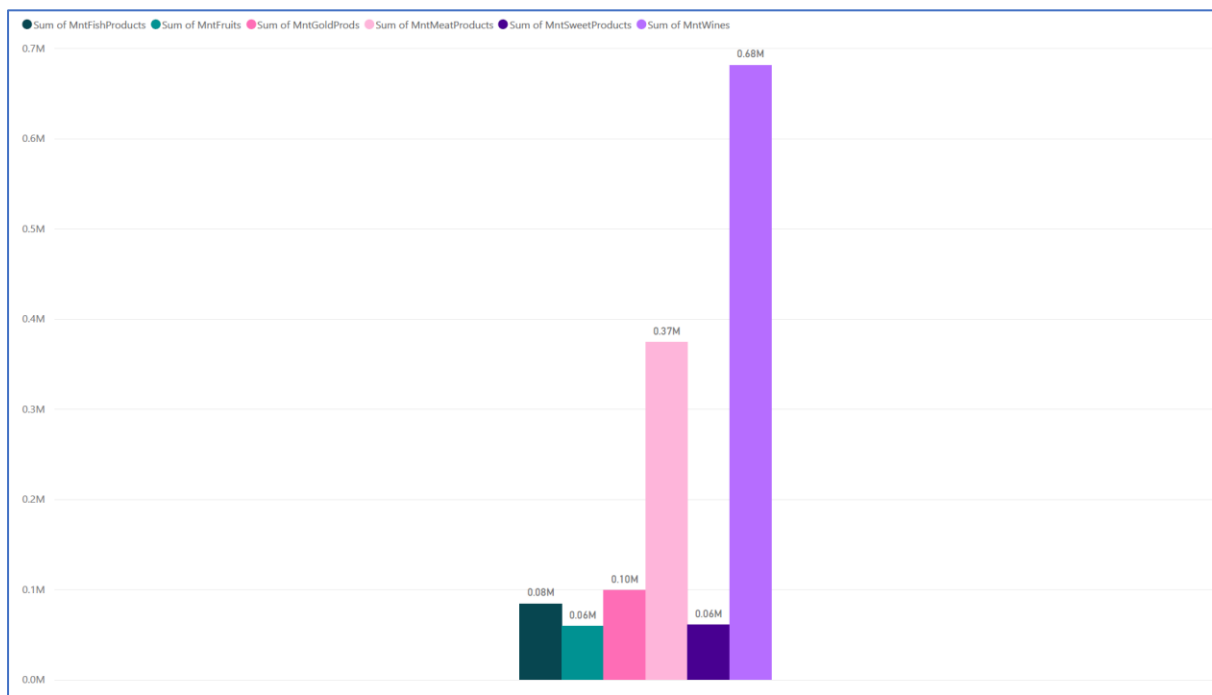


Figure 26 sales per product

Based on the number of sales in all marketing efforts, this chart shows that wine has the most sales, followed by meat goods and gold.

4.4. Multivariate Analysis

Total Sales based of Marital status

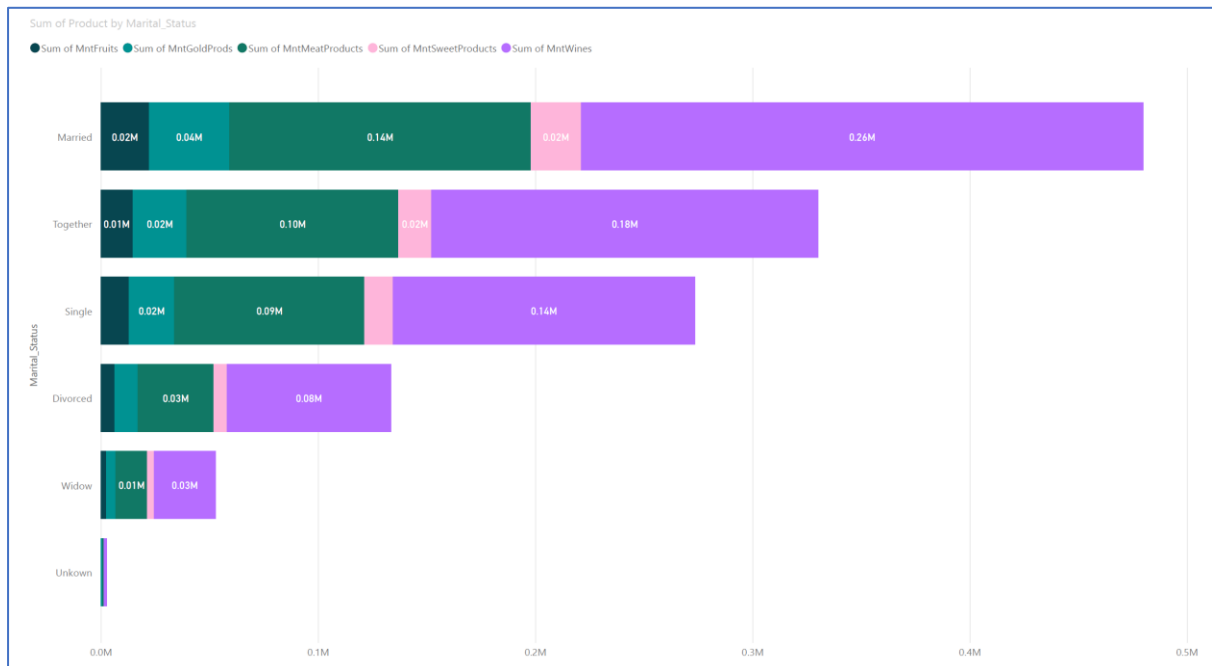


Figure 27 Product by Marital Status

This bar chart displays the sum of buying Fruits, Gold Products, Meat Products, Sweet Products, and Wines grouped by Marital_Status. As can be imagined, married people have more consumption of all goods. But on the other hand, widowed people have less consumption than other people who live alone. This could be due to older age and should be investigated.

- **Hypothesis: Are widowed people who buy less than other single people, because of their old age?**

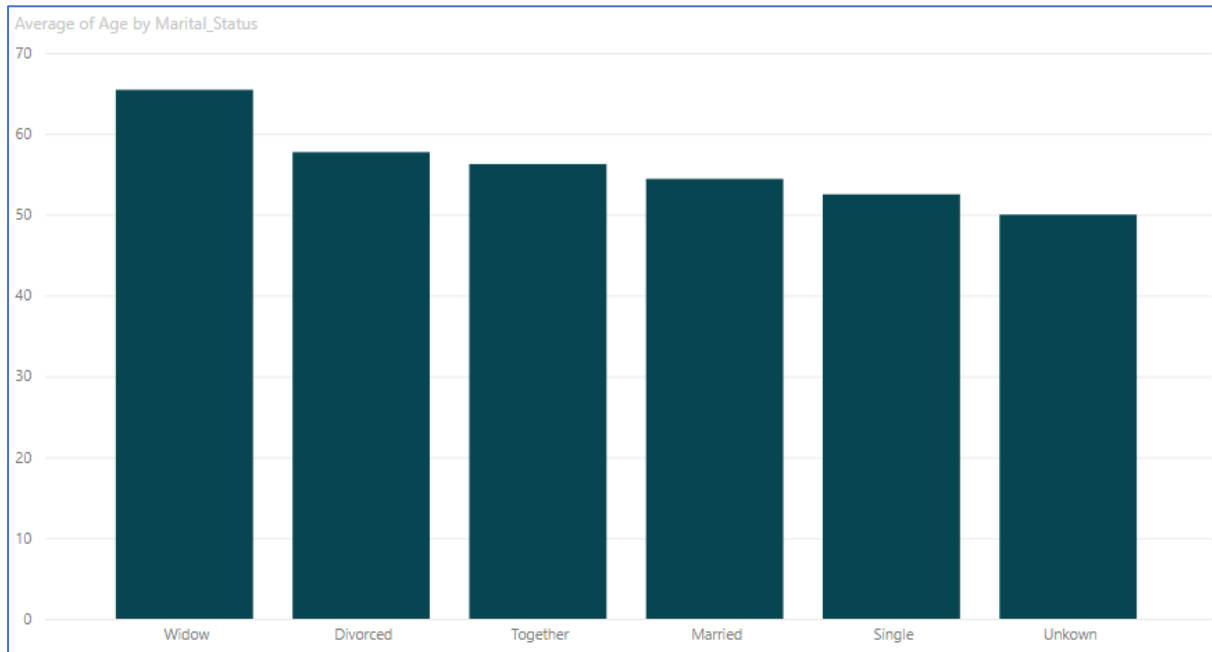


Figure 28: Average Age-Marital State

As can be seen in the figure, widowed people are older on average, so we expect less consumption of goods from them.

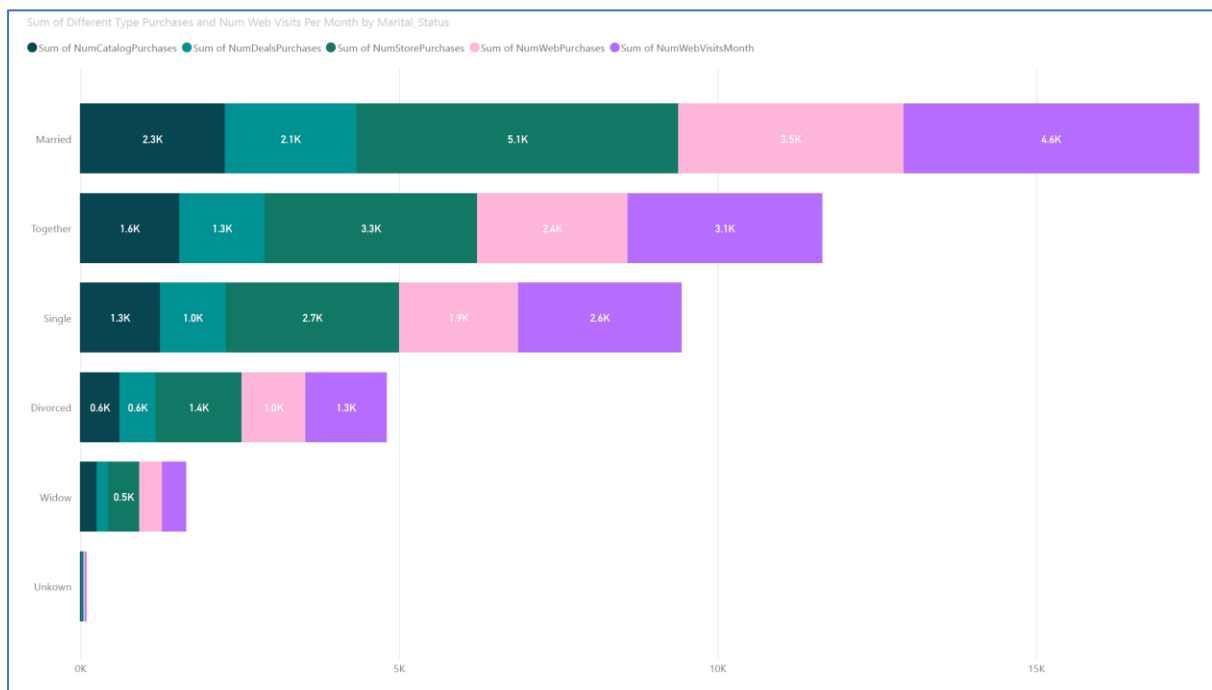


Figure 29: Products Based on Marital State

In chart at Figure 29, the priority of face-to-face sales is higher than online sales, but both of them are much more efficient than sales using catalogs.

Total Sales based of Education Status

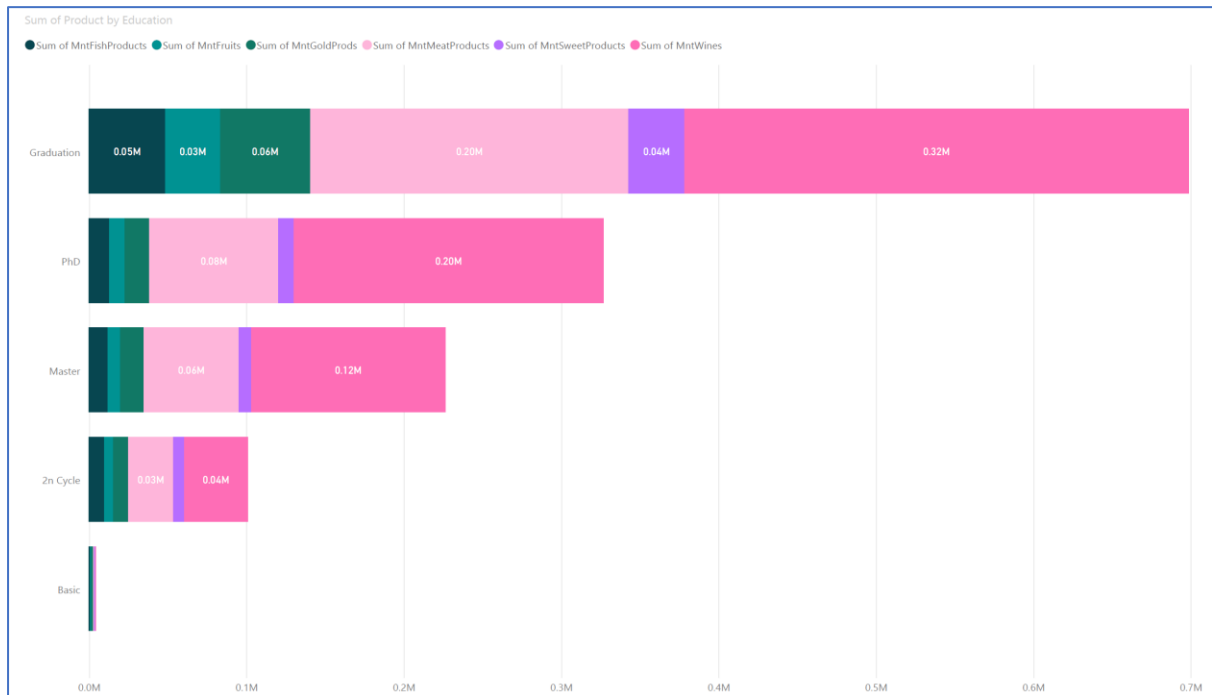


Figure 30 Product by Education

this chart shows the sum of buying Fish, Fruits, Gold, Meat, Sweet Products, and Wines, but grouped by Education level. It is evident that the quantity of spending increases as the level of education increases. Consequently, it is imperative to determine whether the income will rise in tandem with the increase in the number of corrections.

- **Hypothesis: Can the purchase of more educated people be due to their higher income?**

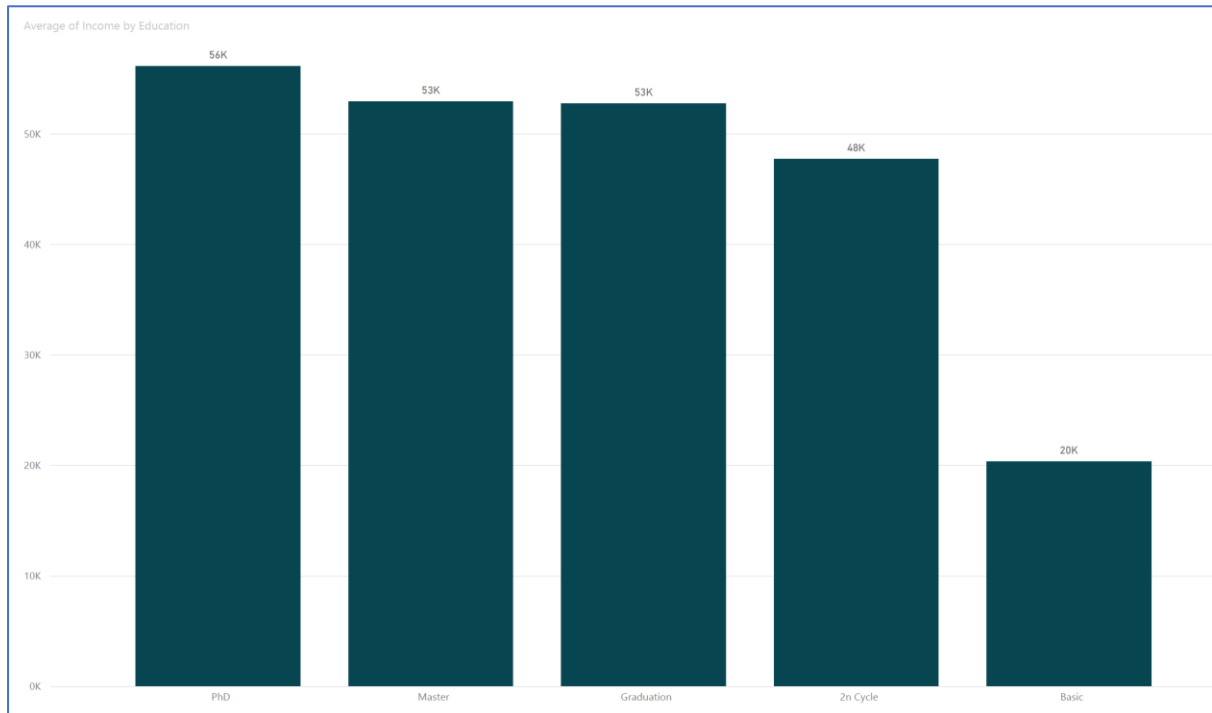


Figure 31: Average Income Based Education

Figure 31 shows that people with more education have a higher income, on the other hand, the next Plot (Figure 32) shows that in our statistical population, people with higher education have children at home earlier, these two factors can determine the reason for more purchases in people with higher education.

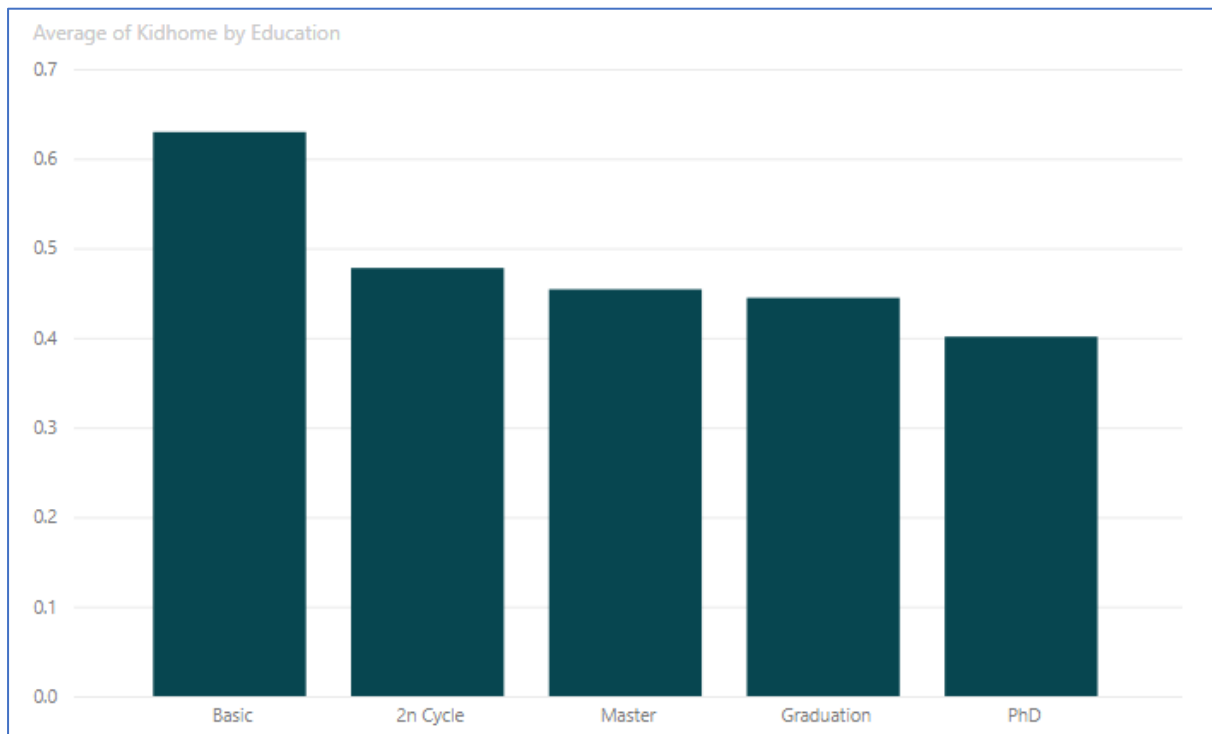


Figure 32: Kids in Home Based On Education

In figure 33 and figure 34 it shows that people with higher education both have more participation in marketing campaigns and due to more purchases, they have used various shopping portals to supply goods to a greater extent.

On the other hand, in all groups, buying directly from the shop was preferred, followed by online shopping. The sale through the website is a small and insignificant part of the business, since it is not very effective.

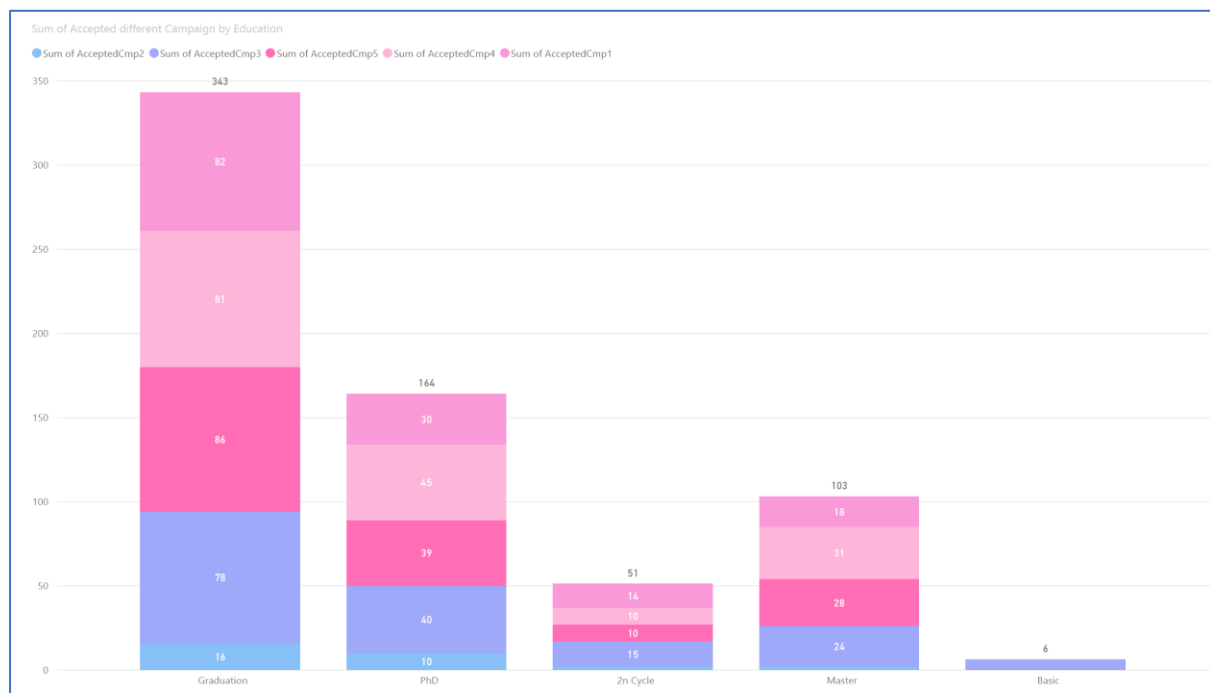


Figure 33:Rate of Accept Campaign based Education



Figure 34:Rate of Buying Per Media Based Education

Impact of Children in Home

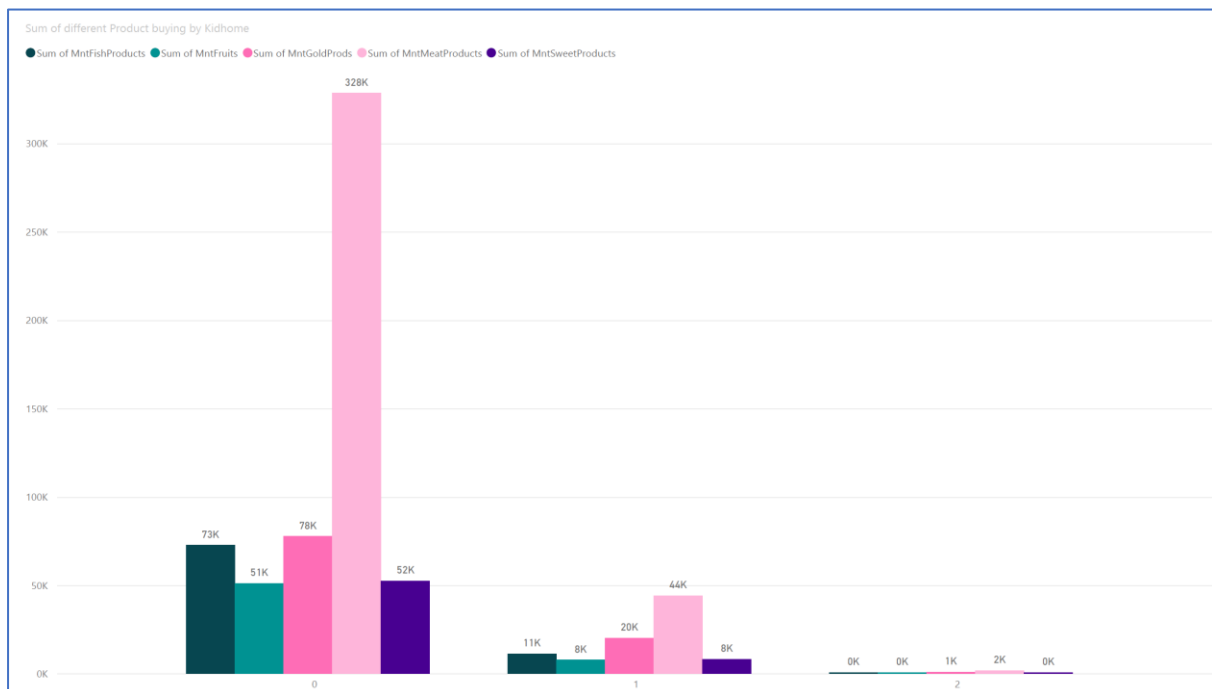


Figure 35: Sum of Each Product Buying Based on Number of Kids at Home

In the Plat above(Figure 35), it is clear that the amount of purchases in campaigns has a reverse relationship with the number of children living at home, which needs to be investigated

Buying meat is much more among people who have less children at home. It can increase sales by focusing on the marketing of meat products in this group.

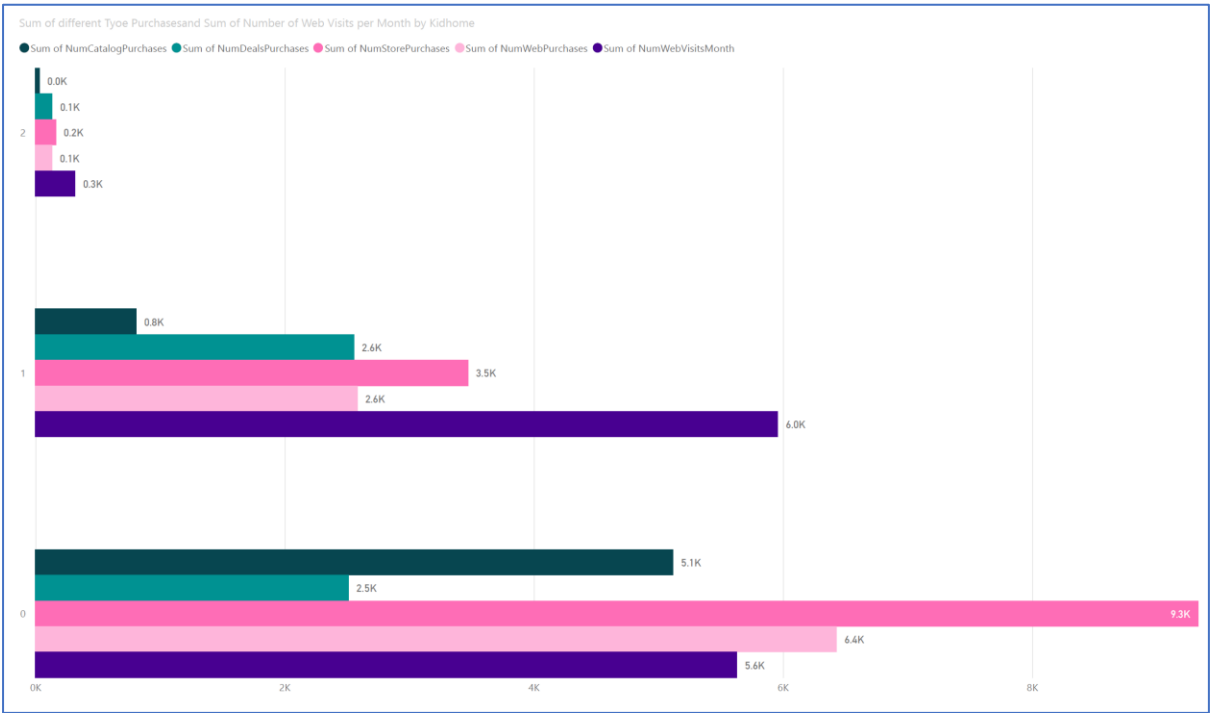


Figure 36: Sum of Each Media Sales Based on Number of Kids at Home

The plot clearly shows that people who have fewer children at home are interested in going out and shopping in person. With the increase in the number of children living at home, the interest in buying promotional offers increases and this can also help in deciding how to market.

Analysis Of Response

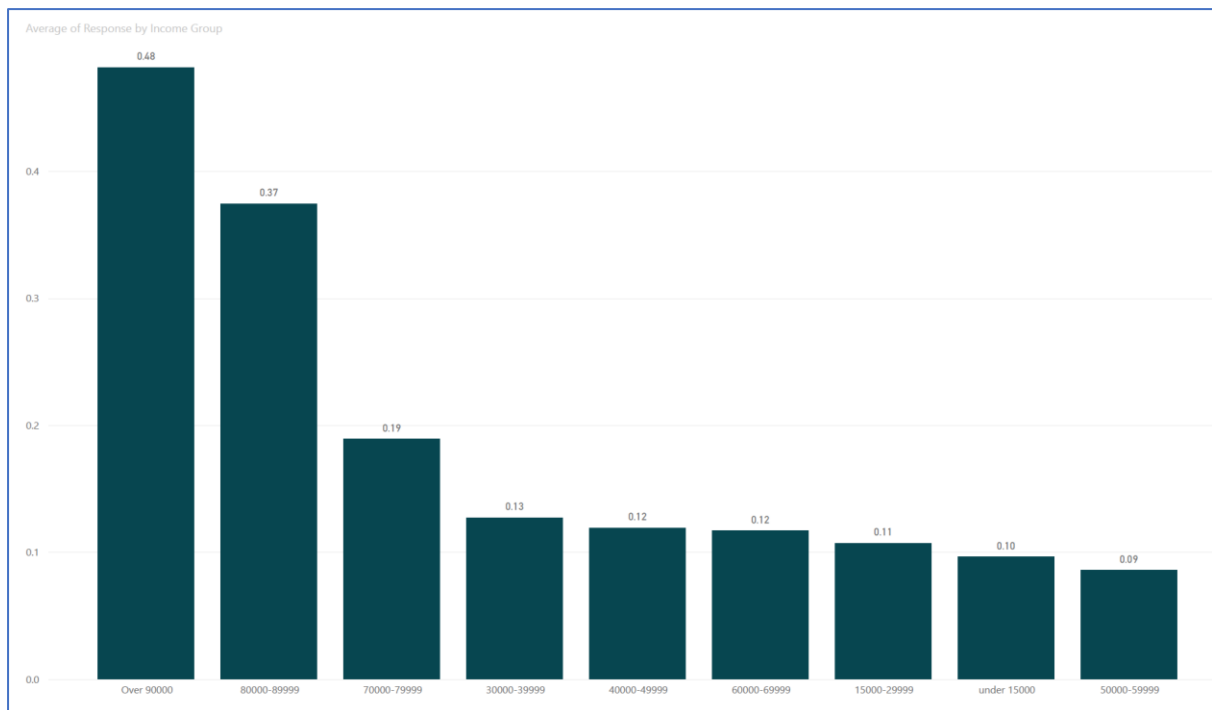


Figure 37: Average Response based Incoming Group

The trend indicates that higher income groups tend to have higher average responses, with the most substantial drop occurring between the 80000-89999 and 70000-79999 income groups.

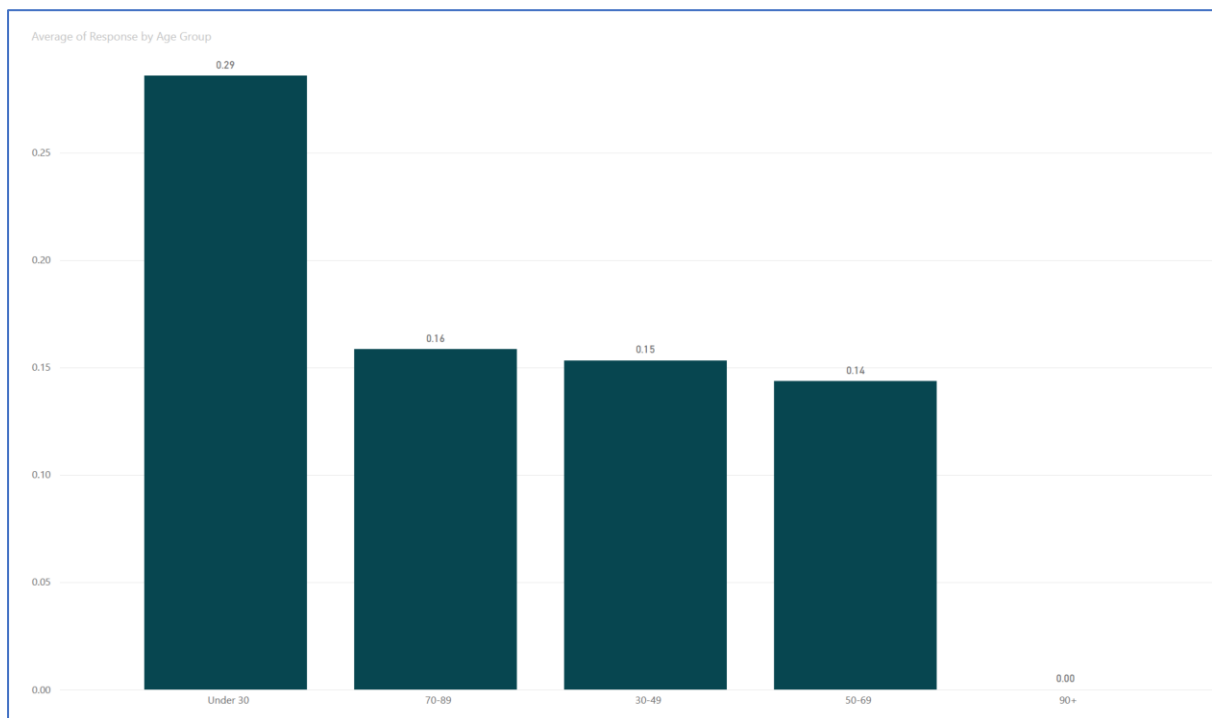


Figure 38: Average Response based Age Group

The trend here shows that younger individuals (under 30) have a significantly higher average response compared to other age groups. The response tends to decrease with age, with a slight increase in the 70-89 age group, but overall, older age groups have lower average responses.

Distribution Of different Products Sale between Different categories of income

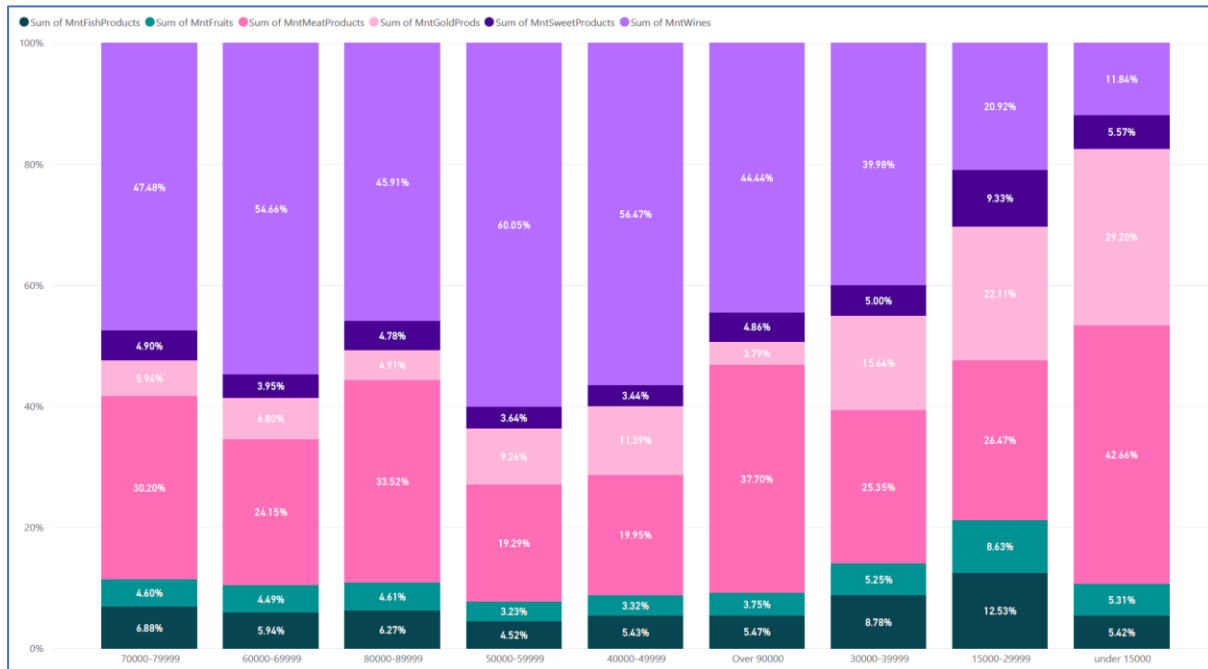


Figure 39: Distribution of Type of Product Based on Income Categories

Key Observations of Plot in Figure 39 is :

- Sweet Products tends to have the highest proportion of expenditures across most income groups, especially in the lower income groups.
 - Meat Products is the second most common expenditure in many groups, notably high in the "under 15000" and "15000-29999" groups.
- Wines has a significant proportion in the "Over 90000" income group.
- Gold Products and Fish Products generally have the smallest shares across all income groups.
 - Higher income groups (Over 90000) show a more balanced distribution with higher expenditure on Wines compared to other groups.
 - Lower income groups show higher expenditure on Meat Products and Sweet Products, indicating these are staple products in their expenditures.

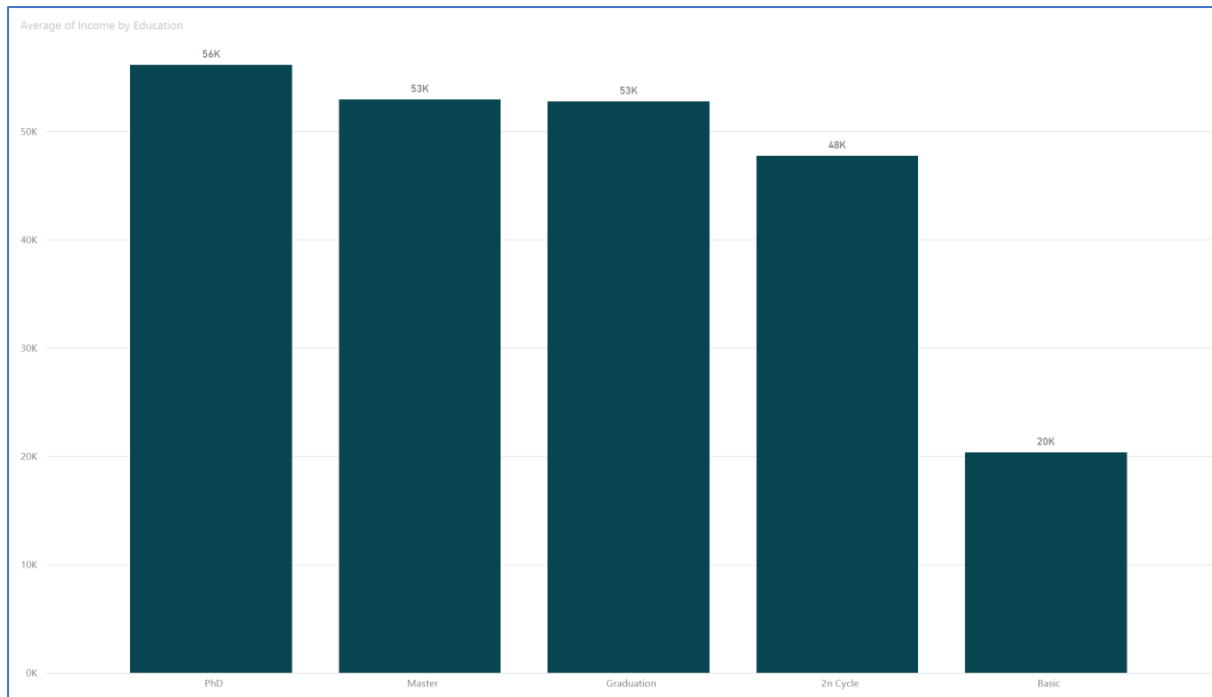


Figure 40

4.5. Conclusion

This report offers a thorough examination of the marketing campaign statistics, underscoring significant information about the buying habits, demographics, and efficacy of the campaign. Several important findings that may be made from the study are:

1. Improving the efficacy of campaigns and raising sales in particular product categories might be achieved by focusing on particular consumer groups according to income, degree of education, and makeup of the home.
2. Meeting the various needs of customers and maximising interaction need the use of several sales channels, especially in-store and online.
3. Better response rates and larger sales might result from concentrating marketing initiatives on younger age groups and higher income people.
4. Given their present lesser impact as compared to other channels, catalog-based sales channels may benefit from investigating ways to improve their efficacy(Or remove).
5. Keeping up a loyal customer base requires ongoing observation of customer satisfaction and resolution of any possible problems.

All things considered, this study offers insightful information that will guide next product offers, marketing plans, and focused initiatives to maximise sales and consumer involvement.

5. References

- (14) Microsoft Power BI, Leader in Gartner's Magic Quadrant 2023 | LinkedIn (no date). Available at: <https://www.linkedin.com/pulse/microsoft-power-bi-leader-gartners-magic-quadrant/> (Accessed: 3 June 2024).
- 'Color Blindness in Clinton in Clinton, CT' (no date) Clinton Eye Associates. Available at: <https://clintoneye.com/services/color-blindness/> (Accessed: 12 June 2024).
- 5 tips on designing colour-blind-friendly visualizations (no date). Available at: <https://www.tableau.com/en-gb/blog/examining-data-viz-rules-dont-use-red-green-together> (Accessed: 12 June 2024).
- Best Practices for Designing Accessible Views (no date). Available at: https://help.tableau.com/current/pro/desktop/en-us/accessibility_best_practice.htm (Accessed: 2 June 2024).
- Compare Microsoft Power BI vs. Tableau (no date) G2. Available at: <https://www.g2.com/compare/microsoft-microsoft-power-bi-vs-tableau> (Accessed: 3 June 2024).
- Edmond, S. and Crabtree, M. (2023) Power BI vs Tableau: Which is The Better Business Intelligence Tool in 2024? Available at: <https://www.datacamp.com/blog/power-bi-vs-tableau-which-one-should-you-choose> (Accessed: 3 June 2024).
- Google Trends (no date) Google Trends. Available at: <https://trends.google.com/trends/explore?date=all&q=Power%20BI,Tableau&hl=en> (Accessed: 2 June 2024).
- Kim, J.H. (2024) 'Tableau vs PowerBI: Here is What You Need to Know', Medium, 4 April. Available at: <https://jefftheanalyst.medium.com/tableau-vs-powerbi-here-is-what-you-need-to-know-79da2f25d744> (Accessed: 11 June 2024).
- Kirk, A. (2019). Data Visualization: A Handbook for Data Driven Design. SAGE Publications.
- Lonsdale, M. S. (2022) Information visualisation: From Theory, To Research, To Practice ad Back. London: SAGE.
- Marketing Campaigns_EDA (no date). Available at: <https://kaggle.com/code/jigsaw13/marketing-campaigns-eda> (Accessed: 6 June 2024).
- Mattia, A. (no date) 'Data visualization: an exploratory study into the software tools used by businesses', 18.
- Visual Best Practices (no date). Available at: https://help.tableau.com/current/blueprint/en-us/bp_visual_best_practices.htm (Accessed: 2 June 2024).
- Ferydooni F(2024), Available at: <https://github.com/fferydooni/Marketing-Campaign-Visualization>