National College of Ireland

Project Submission Sheet

Student Name:	Bolormaa Mendbayar, Akash Narayan Pal, Mitali Vilas Sopte		
Student ID:	x23176725, x22211420, x22198121		
Programme:	MSc in Data Analytics	**Year:**	2023
Module:	Business Intelligence & Business Analytics		
Lecturer:	Vikas Sahni		
Submission Due Date:	06/12/2023		
Project Title:	Implementation Report – Business Analytics on Sephora using Tableau and Power BI		
Word Count:			

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the references section. Students are encouraged to use the Harvard Referencing Standard supplied by the Library. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action. Students may be required to undergo a viva (oral examination) if there is suspicion about the validity of their submitted work.

Signature:	Bolormaa Mendbayar, Akash Narayan Pal, Mitali Vilas Sopte
Date:	06/12/2023

PLEASE READ THE FOLLOWING INSTRUCTIONS:

1. Please attach a completed copy of this sheet to each project (including multiple copies).
2. Projects should be submitted to your Programme Coordinator.
3. **You must ensure that you retain a HARD COPY of ALL projects**, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. Please do not bind projects or place in covers unless specifically requested.
4. You must ensure that all projects are submitted to your Programme Coordinator on or before the required submission date. **Late submissions will incur penalties.**
5. All projects must be submitted and passed in order to successfully complete the year. **Any project/assignment not submitted will be marked as a fail.**

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

AI Acknowledgement Supplement

[Insert Module Name]

[Insert Title of your assignment]

Your Name/Student Number	Course	Date
N/A		

This section is a supplement to the main assignment, to be used if AI was used in any capacity in the creation of your assignment; if you have queries about how to do this, please contact your lecturer. For an example of how to fill these sections out, please click here.

AI Acknowledgment

This section acknowledges the AI tools that were utilized in the process of completing this assignment.

Tool Name	Brief Description	Link to tool
	N/A	

Description of AI Usage

This section provides a more detailed description of how the AI tools were used in the assignment. It includes information about the prompts given to the AI tool, the responses received, and how these responses were utilized or modified in the assignment. **One table should be used for each tool used**.

[Insert Tool Name]	
[Insert Description of use]	
[Insert Sample prompt]	[Insert Sample response]

Evidence of AI Usage

This section includes evidence of significant prompts and responses used or generated through the AI tool. It should provide a clear understanding of the extent to which the AI tool was used in the assignment. Evidence may be attached via screenshots or text.

Additional Evidence:

[Place evidence here]

Additional Evidence:

[Place evidence here]

Implementation report

1)Data analysis

For data analysis we created some visualizations which gives us more information and we can make better decisions and make business more profitable at lower cost

We have created a bar chart with range in Power bi where we can select the year and get our customized visuals and which makes more confident to take better decisions

The below mentioned fig1 comprises of *Sales* for Sephora , *Country_name* , & *the year*

For instances if we the see the visuals in year range in 2020 and 2021 we can see that the European countries named as France , Poland, Sweden and many more countries have better total sales compared to other European countries such as Ireland and Italy

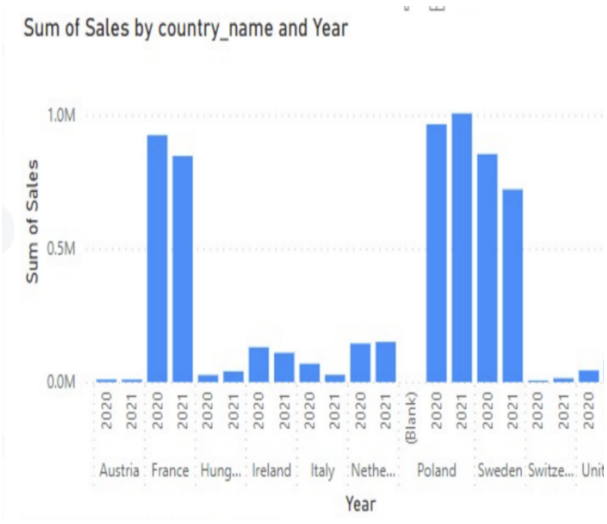

Fig1

The below mentioned fig 2 shows the country which has maximum sales

We have use maximum as aggregate functions on sales and here is the desired output wherein Poland has the maximum sales in the year 2020 to 2021

Fig 2

The next fig 3 shows the visuals to see which country and which brand has more sales and we have used columns as ***Sales , brand_name , country_name . The visuals are for the year 2020-2021-2022-2023. The brand Tom ford has more sales 159400$ in Sweden***
the sales for

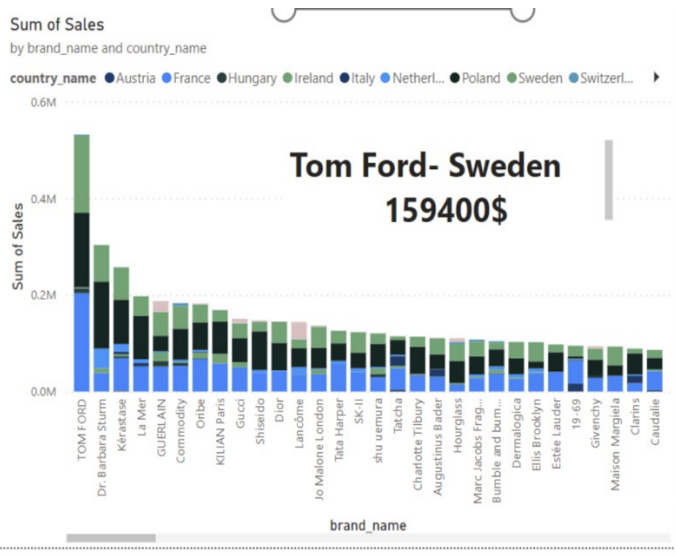

Fig3

The below fig 4 are visuals for all the locations to fetch the sales but on this visuals we have select the country has france and total quantity sold based on location is 67k from 2020 to 2023 and total sales for 3 years is 3 million

Fig 4

The below fig5 visuals shows the distribution for total *sales* by *secondary_category and as per the visuals we see that total sales is for women category for $2million by the year 2020-2023*

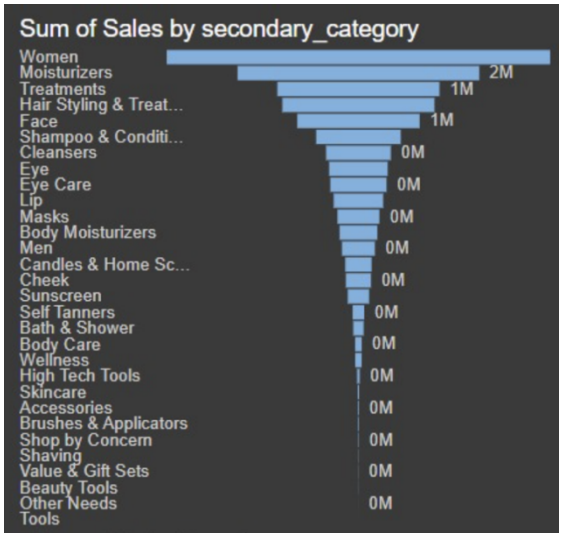

Fig 5

This fig 6 visuals shows that skin care has more sales in primary category and fragrance is the next category which gets sold

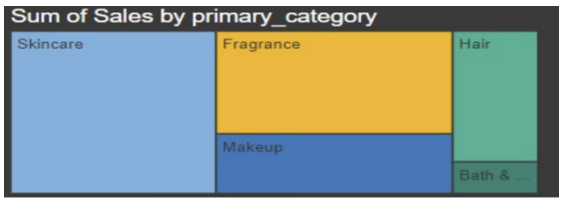

Fig 6

The below mentioned fig 7 visuals is in between *loves_count* and *product_name* ,which states that soft pinch liquid blush is something is most attracted to customers compared to other products **and this is the** only product which is attracted to more than million customers which means that this product is high in demand and we can keep the stock up always in order increase the profit margin

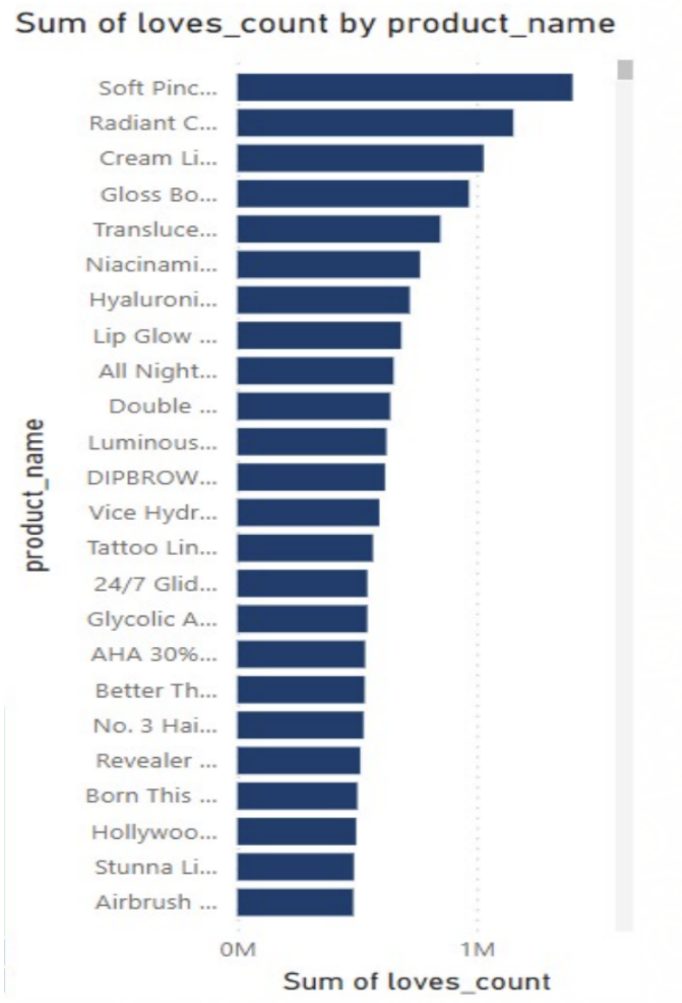

Fig 7

The next fig 8 visuals are based on rating with all the categories

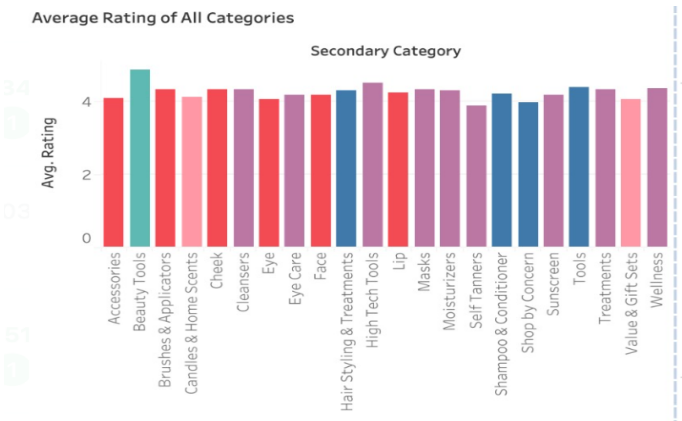

Fig 8

For this bar chart we have used the columns as secondary_category and rating and the

visuals are straight forward that beauty tools are highly rated by the customers which might be in demand and the other one would be high tech tools .
The lowest rating is for self-tanners with an average of 2.7 rating

Below mentioned fig 9 visuals is the comparison between men and women and brand ratings

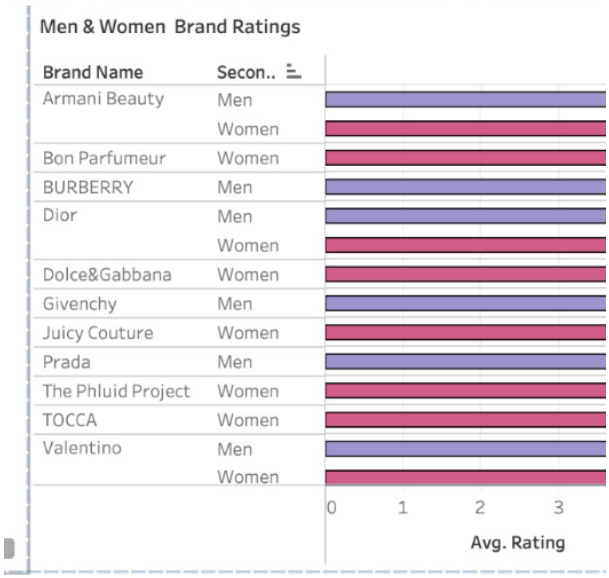

fig 9

On this visuals Burberry , Givenchy, Prada and Valentino are most loved by men's and there is no comparison with women and men's have rated this secondary category beyond 4 which has a positive correlation with men's brands

Men and women both are attracted by *valentino* as they have rated the same .

Visuals for Product sold based on men and women category

Comparison between product sold between men and women
referring to fig10

The perfume quantity is sold the most in the women category and the lowest is Body mist and hair mist .
In Comparison to perfume , only half of rollerballs and travel size were sold

Men's only prefer to buy cologne .

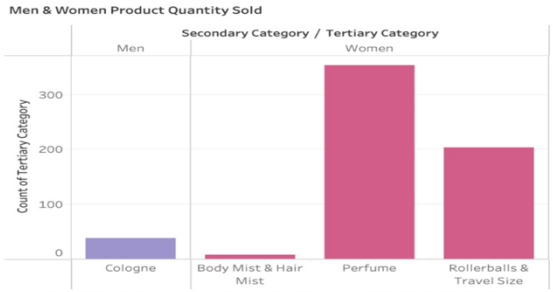

Fig 10

2)Some key KPI' analysis for Sales

***) Sales between weekday and weekend :**

- Sales are the important factors on weekends and weekdays which gives us a general idea what to sell on which specific day .
- It also helps stakeholders to make a better decision in terms of scaling the product and also for manufacturers to deliver the products at the right time .

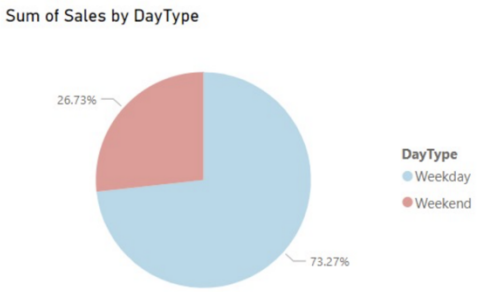

Fig 11

- It is clearly seen from the pie chart that 73.27% of the sales is on weekdays and slightly distributed on weekends .
- With these we can make a better roster for employees and also we can build a great employee relationship in our organization by giving them suitable breaks whenever necessary .

***) Sales based on Variation Type**

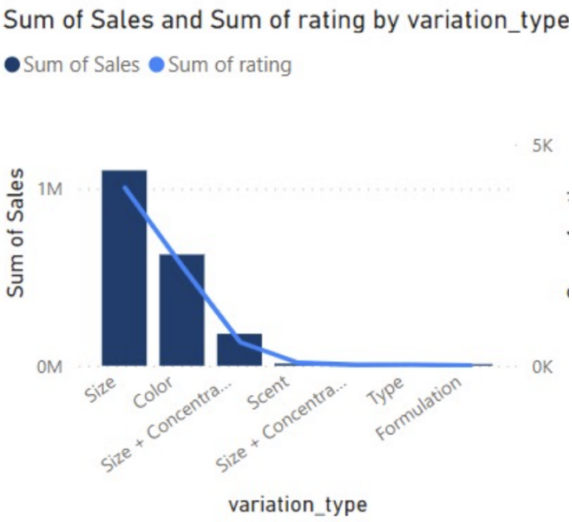

fig 12

The above mentioned fig 12 visuals are basically compared with total sales with respect to total rating and variation type which determines great insights .

- As the scale of rating decreases the sales sometimes decrease at the start of these visuals.
- It is seen that *scent* rating has decreased by the customers and overall sales decreased which shows a direct relationship in terms of rating and sales .

***) <u>Sales based on average price and total loves count by product name</u>**

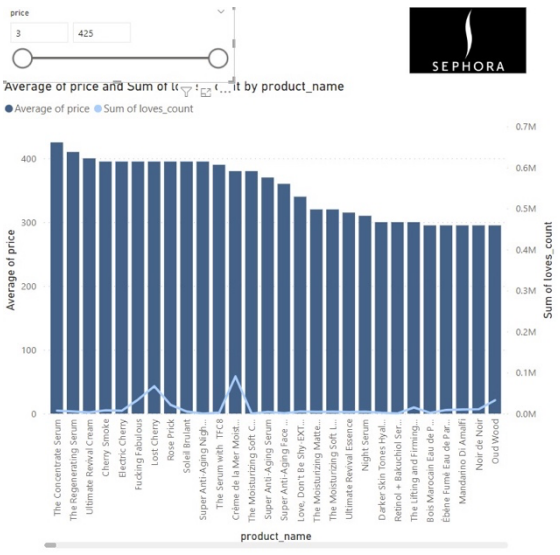

Fig13

The graph is made by average of the price and summary of the loves_count by product_name ,

If we change the price we can see the difference in the visuals

The fig 13 shows that there is high price in the products . and the loves_count is low for which concentration serum price is $425 and the total love_counts is 7514 . which means that customer's don't like expensive products .

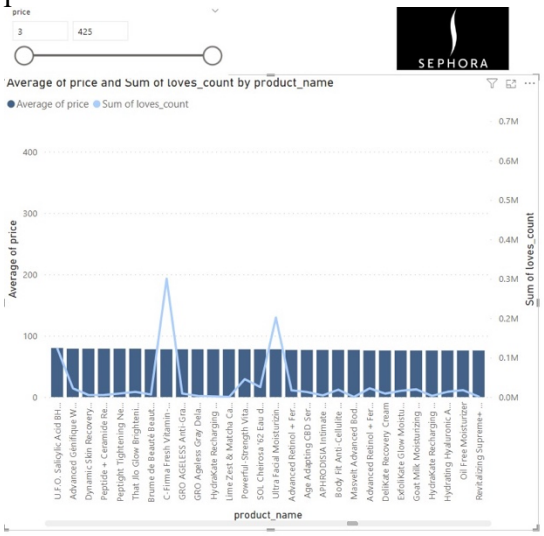

Fig 14

In fig 14 average price of products is around $200 and highest loves_count is 0.3million It means that when the price is cheap customers love the products and might buy for future .

Price	Loves_count
$425	7514
$122	262434
$69	622995

<u>2) Database Migration</u>

- Migrating the raw data into better data we have used R programming language to clean the data
- For cleaning we have to check null values .
- We have not clear the binary values
- This is a very crucial step in loading the data to any tool such as power bi

or Tableau . Once the data is cleaned it is ready for analysis and we can derive insights **.**

3) Key Kpi's

- The columns that we have used for analysis are from product tables and thecolumns are *product_id ,first_name,phone_number,date,time,product_id,product_name ,brand_id,brand_name,loves_count ,rating,reviews,size,variation_type, variation_value,variation_desc,ingridents ,price,limited_edition,new,online_only,out_of_stock,sephora_exclusive ,highlights,primary_category,secondary_category,tertiary_category,quantity,sales*

But the key columns are loves_count, Sales , Brand_name ,Country , secondary_category,Date ,Product_name .

1) WorkFlow :

Below figure provides a graphic overview of the business process . The workflow shows us the step by step how our business intelligence system is developed from start to end

fig 14

Data Collection

- First we gathered our data from Kaggle which is a real data.
- Next we had many tables but we prefer to take the product data .
- We consider the product data as it was aligned with our problem statement and thereby we proceed further .

Data Processing :
The most important factor after collecting the data is to clean the data which affects the results to make desired decision to give a the organization a better results

- For cleaning we used R programming language .
- We looked for null values
- When null values got generated we discard that values .
- Then according to our problem statement we create new columns using mockaroo such as country , quantity , etc to achieve our goals.

Analysis and Reporting :

After getting the clean processed data we moved forward for our analysis using Power bi and Tableau .

Power Bi : It is provided by Microsoft itself and it's a tool where we can create visualization's , dashboards and this tool is effective in creating business intelligence systems at lower cost .

Tableau :
This tool also helps us to create business intelligence systems based on data that we have and can build dashboards which can accept live data as well .

We have create dashboards and visualization on both these tools .

Implementation :

As we have the problem statement with respect to sales we have prepared three dashboard's and each dashboard is developed by every individual using power bi and tableau .

1) Sales Dashboard
2) Product Dashboard
3) Country wise Dashboard

All these dashboards provides great insights with respect to the year ,category , product etc. We have also done the comparison in between men and women .

Team work

Yes the project could not be accomplished with these participants

Team member workload distribution

Team 27:

	X23176725 Bolormaa Mendbayar	**X22211420 Akash Narayan Pal**	**X22198121 Mitali Vilas Sopte**
Data Collection	Decided to choose Sephora dataset from Kaggle	Decided to choose Sephora dataset from Kaggle	Decided to choose Sephora dataset from Kaggle
Data Cleaning	If the columns are unnecessary for next step, removed in the excel, added mock data using Mockaroo	Combined products and review dataset and if the data is big, reduced the rows	Removed NA and NULL values using R
Flow process	Discussed between the group, made by me	Discussed between the group	Discussed between the group
Balanced Scorecard	Discussed between the group, made by me	Discussed between the group	Discussed between the group
ER diagram	Discussed between the group	Discussed between the group, made by him	Discussed between the group
Visualization	Implemented World Wide Sales Dashboard using PowerBI	Implemented Sales Dashboard using PowerBI	Implemented Category Sales Dashboard using Tableau
Specification report			-Background information and organization and scope - System design -Entity relationship diagram
Implementation report	-Management dashboard and my analysis -Solution	-Analysis, including business rules -Team work	
Distribution in percentage	33%	33%	33%