Project Documentation – Customer Personality Analysis

1. Introduction

Customer Personality Analysis is used to understand a company's ideal customers by studying their demographics, spending behavior, product preferences, and purchase channels.
The dataset contains customer profile details, product-wise spending, promotion responses, and purchase behavior.
This project uses Excel-based analytics—functions, pivot tables, charts, and scenario analysis—to segment customers and identify meaningful business insights.

2. Aim

The primary aim of this project is to:

- Identify customer segments using demographic and behavioural attributes.
- Understand purchasing patterns across product categories.
- Provide insights that can support targeted marketing, product planning, and promotional strategies.

3. Business Problem / Context

Companies often market products to all customers, resulting in wasted resources.
This project answers key business questions such as:

- Which customer segments spend the most?
- Which product categories generate higher revenue?
- How do demographics (age, education, marital status) influence buying behavior?
- Which marketing campaigns are most effective?

- Which channels (web, store, catalog) do customers prefer?

Using segmentation and analysis, businesses can target the most profitable customers effectively.

4. Project Workflow

1. **Data import** into Excel.
2. **Initial data inspection** to understand structure.
3. **Data cleaning** – fixing missing, incorrect, or inconsistent values.
4. **Creating derived metrics** such as Age, Total Spending, Age Category, % Income Spent.
5. **Segmentation & Filtering** based on demographics, income, age groups, etc.
6. **Statistical and formula-based analysis** using Excel formulas.
7. **Visualization with PivotTables & PivotCharts.**
8. **Scenario Analysis** using Goal Seek, What-If Analysis and Scenario Manager.
9. **Power Pivot & Data Modelling** for advanced multi-table analysis.
10. **Insight generation** and reporting.

5. Data Understanding

- **Rows (customers):** typically ~2,000+
- **Columns:** 29 attributes
- **Data categories:**
    - *Demographics:* ID, Birth Year, Marital Status, Education, Income
    - *Product Spending:* Wines, Fruits, Meat, Fish, Sweets, Gold

- *Promotions:* Campaign acceptance statistics
- *Purchasing Behaviour:* Web, Store, Catalog purchases
- *Engagement:* Recency, Complaints

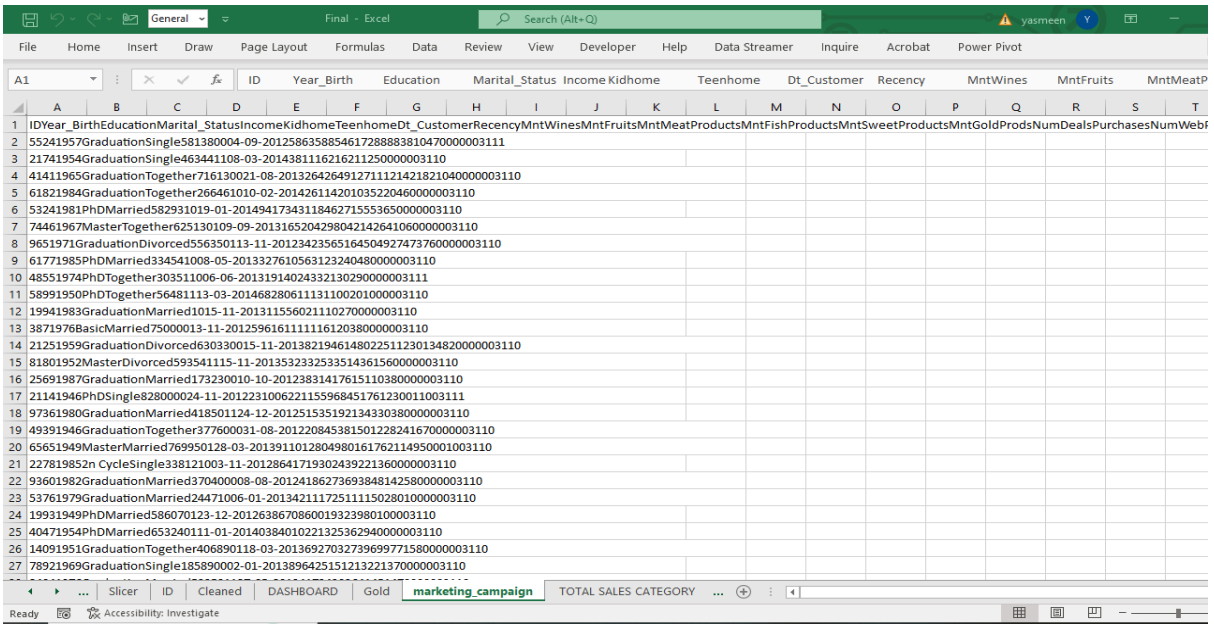
Initial Observations:

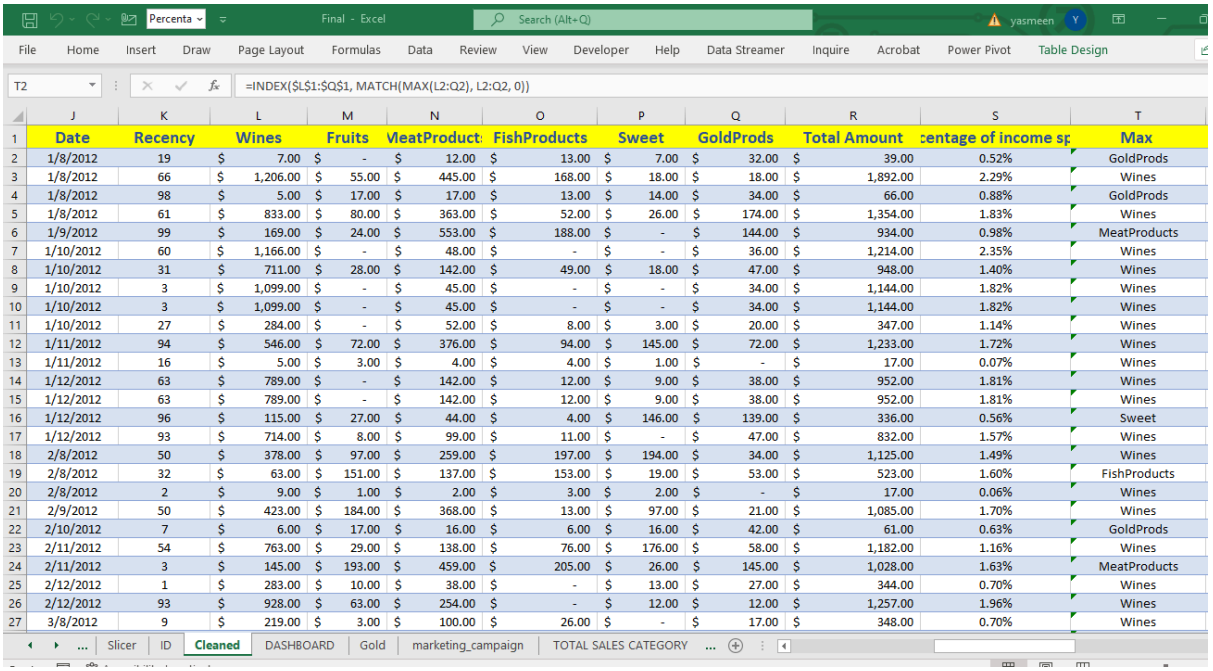
- Wide variation in income levels.
- Uneven spending across products.
- Multiple purchase channels are used but with different frequencies.

6. Data Cleaning

Performed in Excel using:

- Removal of **duplicate entries**.
- Handling **missing values** in income and spending fields.
- Standardizing text formats (Education, Marital Status).
- Converting date columns to proper date format.
- Checking for negative or incorrect values.





Percenta — Final - Excel — Search (Alt+Q) — yasmeen

File Home Insert Draw Page Layout Formulas Data Review View Developer Help Data Streamer Inquire Acrobat Power Pivot Table Design

T2 =INDEX($L$1:$Q$1, MATCH(MAX(L2:Q2), L2:Q2, 0))

	J	K	L	M	N	O	P	Q	R	S	T
1	Date	Recency	Wines	Fruits	MeatProduct:	FishProducts	Sweet	GoldProds	Total Amount	centage of income sp	Max
2	1/8/2012	19	$ 7.00	$ -	$ 12.00	$ 13.00	$ 7.00	$ 32.00	$ 39.00	0.52%	GoldProds
3	1/8/2012	66	$ 1,206.00	$ 55.00	$ 445.00	$ 168.00	$ 18.00	$ 18.00	$ 1,892.00	2.29%	Wines
4	1/8/2012	98	$ 5.00	$ 17.00	$ 17.00	$ 13.00	$ 14.00	$ 34.00	$ 66.00	0.88%	GoldProds
5	1/8/2012	61	$ 833.00	$ 80.00	$ 363.00	$ 52.00	$ 26.00	$ 174.00	$ 1,354.00	1.83%	Wines
6	1/9/2012	99	$ 169.00	$ 24.00	$ 553.00	$ 188.00	$ -	$ 144.00	$ 934.00	0.98%	MeatProducts
7	1/10/2012	60	$ 1,166.00	$ -	$ 48.00	$ -	$ -	$ 36.00	$ 1,214.00	2.35%	Wines
8	1/10/2012	31	$ 711.00	$ 28.00	$ 142.00	$ 49.00	$ 18.00	$ 47.00	$ 948.00	1.40%	Wines
9	1/10/2012	3	$ 1,099.00	$ -	$ 45.00	$ -	$ -	$ 34.00	$ 1,144.00	1.82%	Wines
10	1/10/2012	3	$ 1,099.00	$ -	$ 45.00	$ -	$ -	$ 34.00	$ 1,144.00	1.82%	Wines
11	1/10/2012	27	$ 284.00	$ -	$ 52.00	$ 8.00	$ 3.00	$ 20.00	$ 347.00	1.14%	Wines
12	1/11/2012	94	$ 546.00	$ 72.00	$ 376.00	$ 94.00	$ 145.00	$ 72.00	$ 1,233.00	1.72%	Wines
13	1/11/2012	16	$ 5.00	$ 3.00	$ 4.00	$ 4.00	$ 1.00	$ -	$ 17.00	0.07%	Wines
14	1/12/2012	63	$ 789.00	$ -	$ 142.00	$ 12.00	$ 9.00	$ 38.00	$ 952.00	1.81%	Wines
15	1/12/2012	63	$ 789.00	$ -	$ 142.00	$ 12.00	$ 9.00	$ 38.00	$ 952.00	1.81%	Wines
16	1/12/2012	96	$ 115.00	$ 27.00	$ 44.00	$ 4.00	$ 146.00	$ 139.00	$ 336.00	0.56%	Sweet
17	1/12/2012	93	$ 714.00	$ 8.00	$ 99.00	$ 11.00	$ -	$ 47.00	$ 832.00	1.57%	Wines
18	2/8/2012	50	$ 378.00	$ 97.00	$ 259.00	$ 197.00	$ 194.00	$ 34.00	$ 1,125.00	1.49%	Wines
19	2/8/2012	32	$ 63.00	$ 151.00	$ 137.00	$ 153.00	$ 19.00	$ 53.00	$ 523.00	1.60%	FishProducts
20	2/8/2012	2	$ 9.00	$ 1.00	$ 2.00	$ 3.00	$ 2.00	$ -	$ 17.00	0.06%	Wines
21	2/9/2012	50	$ 423.00	$ 184.00	$ 368.00	$ 13.00	$ 97.00	$ 21.00	$ 1,085.00	1.70%	Wines
22	2/10/2012	7	$ 6.00	$ 17.00	$ 16.00	$ 6.00	$ 16.00	$ 42.00	$ 61.00	0.63%	GoldProds
23	2/11/2012	54	$ 763.00	$ 29.00	$ 138.00	$ 76.00	$ 176.00	$ 58.00	$ 1,182.00	1.16%	Wines
24	2/11/2012	3	$ 145.00	$ 193.00	$ 459.00	$ 205.00	$ 26.00	$ 145.00	$ 1,028.00	1.63%	MeatProducts
25	2/12/2012	1	$ 283.00	$ 10.00	$ 38.00	$ -	$ 13.00	$ 27.00	$ 344.00	0.70%	Wines
26	2/12/2012	93	$ 928.00	$ 63.00	$ 254.00	$ -	$ 12.00	$ 12.00	$ 1,257.00	1.96%	Wines
27	3/8/2012	9	$ 219.00	$ 3.00	$ 100.00	$ 26.00	$ -	$ 17.00	$ 348.00	0.70%	Wines

Slicer | ID | Cleaned | DASHBOARD | Gold | marketing_campaign | TOTAL SALES CATEGORY ...

7. Derived Metrics

You created several useful new columns:

✔ Age

```
=YEAR(TODAY()) - Year_Birth
```

✔ Age Category

Example categories:

- 18–30
- 31–45
- 46–60
- 60+

✔ Total Spending

```
=SUM(MntWines, MntFruits, MntMeatProducts, MntFishProducts,
MntSweetProducts, MntGoldProds)
```

✔ Spending as % of Income

```
=Total Spending / Income * 100
```

✔ Maximum Product Category

Using `MAX`, `INDEX`, `MATCH` formulas.

These metrics help to generate meaningful segmentation.

8. Filtering / Segmentation

Data was segmented by:

- **Age Groups**
- **Marital Status**
- **Education Level**
- **Income Ranges**
- **Product-wise High Spenders**
- **Top customers by promotion response**

Filters and slicers were applied to create interactive dashboards.

9. Statistical / Formula-Based Analysis

The following Excel functions were used:

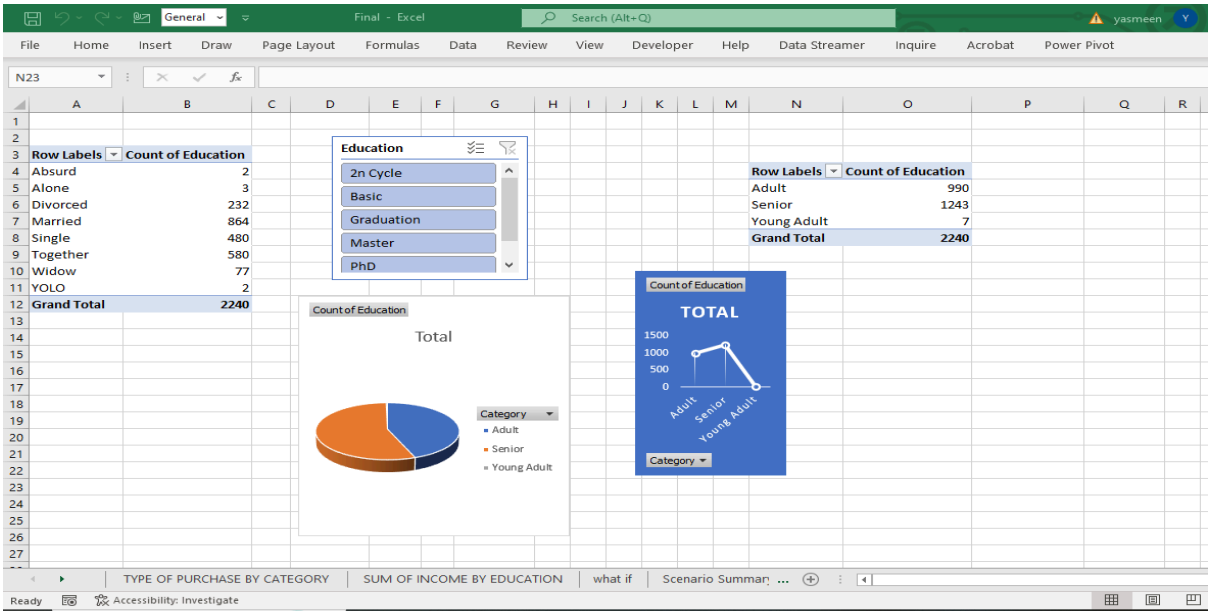
- **Math Functions:** SUM, AVERAGE, MAX, MIN
- **Logical Functions:** IF, AND, OR
- **Lookup Functions:** VLOOKUP, XLOOKUP, INDEX, MATCH
- **Text & Date Functions:** YEAR, TODAY, CONCAT
- **Conditional Formatting:** Highlight high income, high spenders

- **Aggregation:** COUNT, COUNTIF, SUMIF, SUMIFS

These helped compute key customer metrics and spending trends.

10. EDA (PivotTables, PivotCharts, Visualizations)

Extensive Exploratory Data Analysis was performed using PivotTables and PivotCharts:



✔ Product Spend Analysis

- Wine, Meat, Gold, Sweets, Fish spending by **Marital Status**
- Total sales by **Education Level**
- Income vs **Wine Sales Category**

✔ Purchase Behavior

- Type of purchase by **ID**
- Web vs Store vs Catalog purchases

✔ Customer Demographics

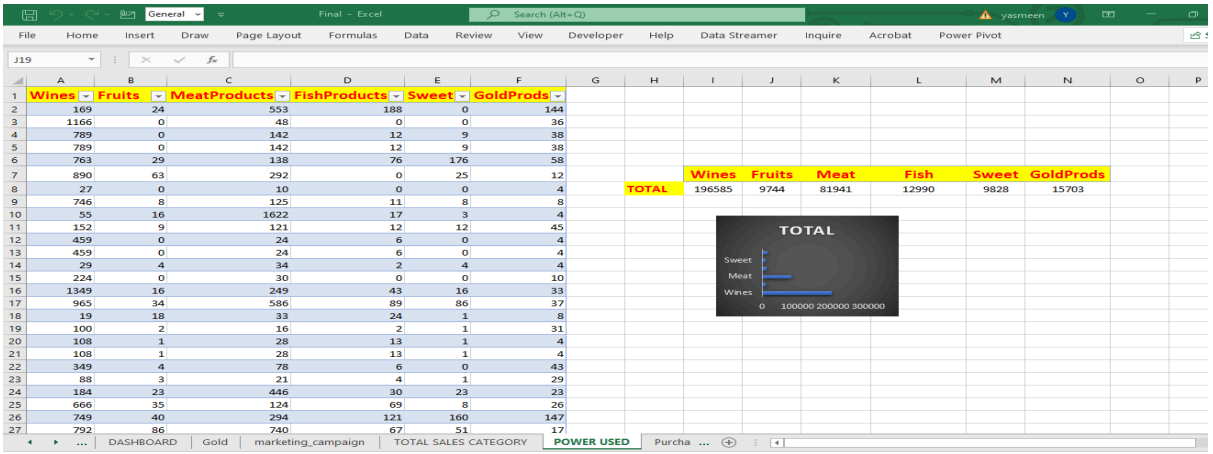
- Income distribution

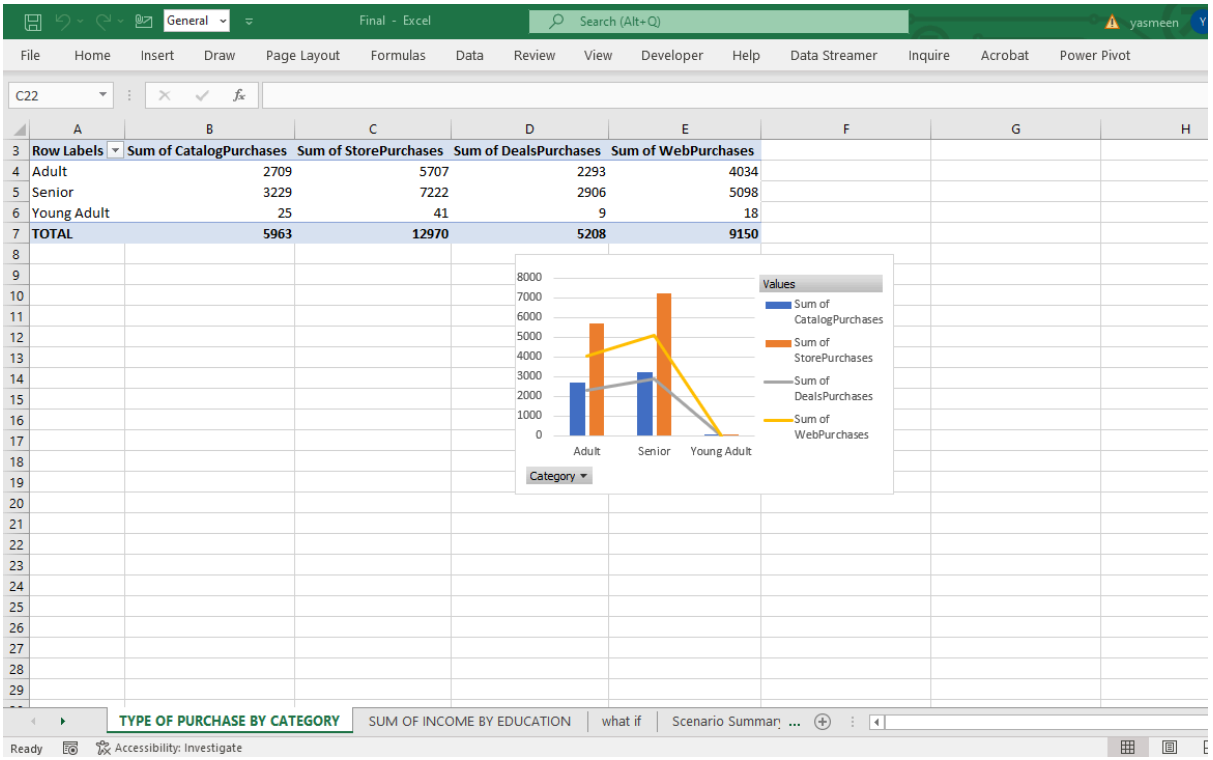
- Age group vs product categories

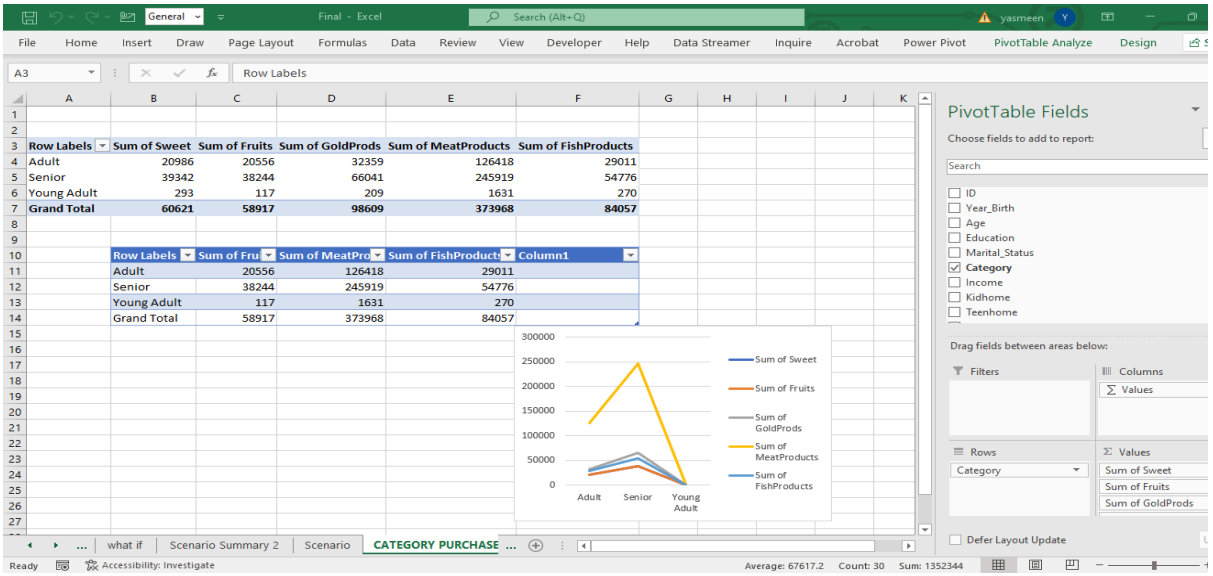
✔ Power Pivot & Data Modelling

- Built relationships between demographic and purchasing tables.
- Created measures for deeper analysis.

All visualizations helped identify high-value customer segments.

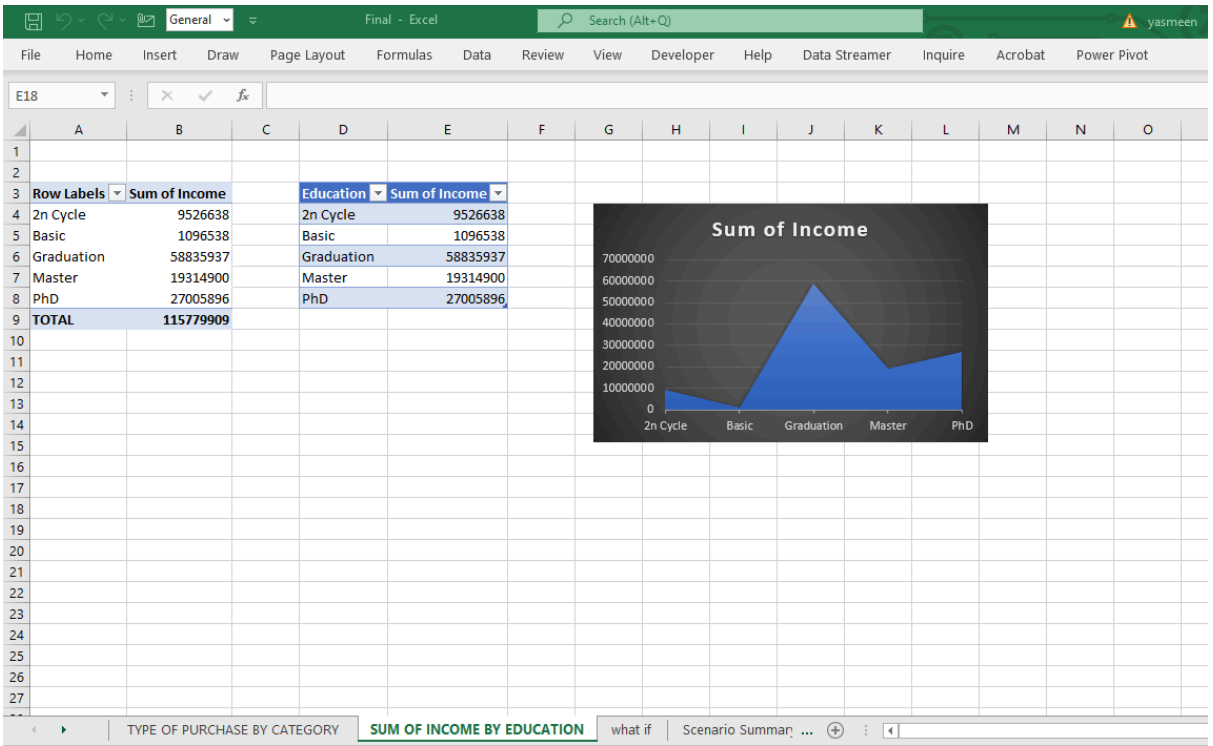






Row Labels	Sum of Sweet	Sum of Fruits	Sum of GoldProds	Sum of MeatProducts	Sum of FishProducts
Adult	20986	20556	32359	126418	29011
Senior	39342	38244	66041	245919	54776
Young Adult	293	117	209	1631	270
Grand Total	**60621**	**58917**	**98609**	**373968**	**84057**

Row Labels	Sum of Fru	Sum of MeatPro	Sum of FishProducts	Column1
Adult	20556	126418	29011	
Senior	38244	245919	54776	
Young Adult	117	1631	270	
Grand Total	58917	373968	84057	



Row Labels	Sum of Income
2n Cycle	9526638
Basic	1096538
Graduation	58835937
Master	19314900
PhD	27005896
TOTAL	**115779909**

Education	Sum of Income
2n Cycle	9526638
Basic	1096538
Graduation	58835937
Master	19314900
PhD	27005896

11. Scenario Analysis

Using What-If Analysis:

✔ Scenario Manager

- Predicted impact of changing customer income or spending.
- Compared multiple scenarios using summary reports.

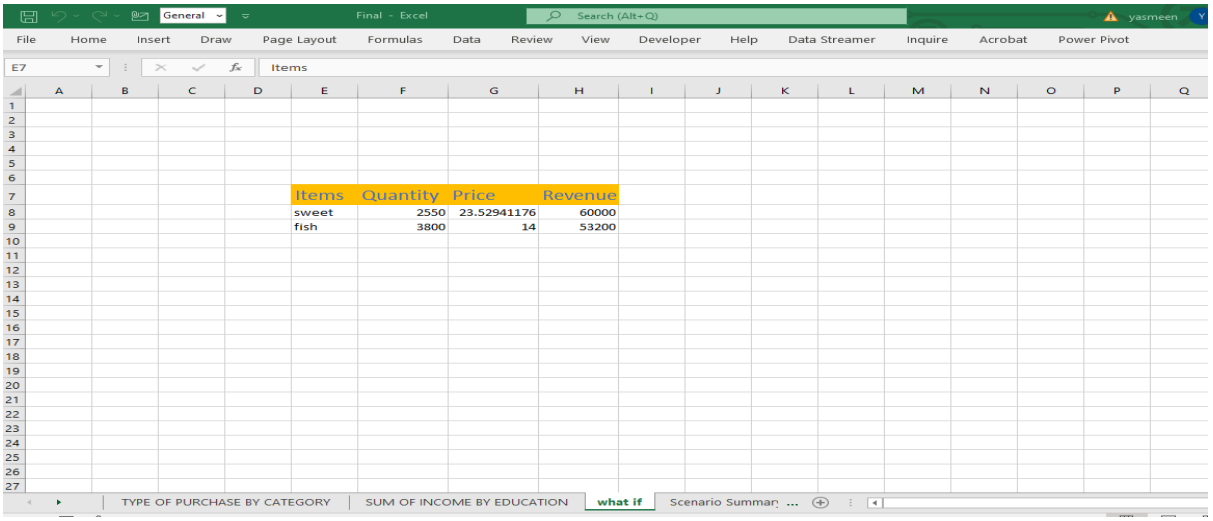
✔ Goal Seek

- Find required spending to reach revenue targets.

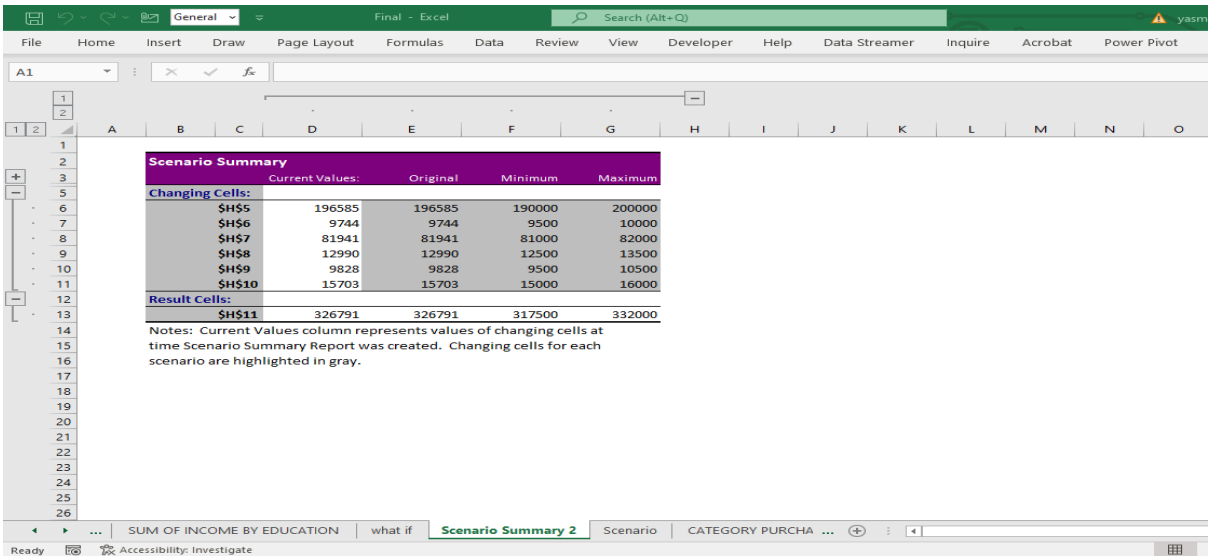
✔ Data Tables

- Tested different income levels and spending assumptions.

This helped analyze future customer behavior patterns.



Items	Quantity	Price	Revenue
sweet	2550	23.52941176	60000
fish	3800	14	53200



Scenario Summary

	Current Values:	Original	Minimum	Maximum
Changing Cells:				
$H$5	196585	196585	190000	200000
$H$6	9744	9744	9500	10000
$H$7	81941	81941	81000	82000
$H$8	12990	12990	12500	13500
$H$9	9828	9828	9500	10500
$H$10	15703	15703	15000	16000
Result Cells:				
$H$11	326791	326791	317500	332000

Notes: Current Values column represents values of changing cells at time Scenario Summary Report was created. Changing cells for each scenario are highlighted in gray.

12. Insights

Demographics & Income

- Higher-income groups spend more on Wines and Gold.
- Older customers (>45 years) show consistent spending across categories.

Product Preferences

- Wine and Meat products are the top revenue-generating categories.
- Sweets and Fruits have the lowest spending.

Channel Behavior

- Store purchases are the most preferred.
- Web purchases are growing but still lower compared to stores.

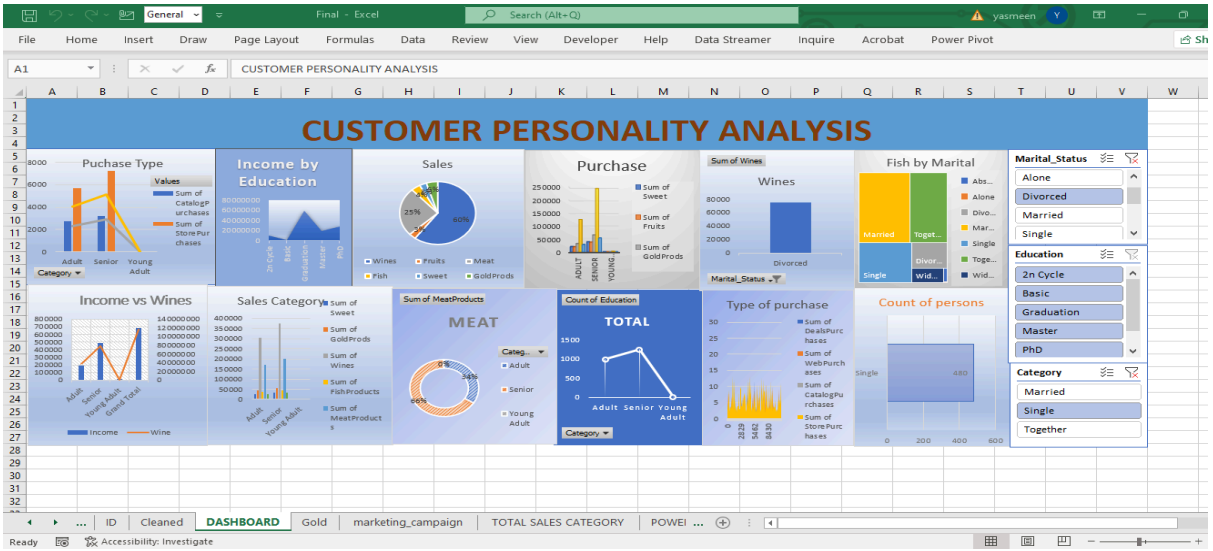
Campaign Effectiveness

- Campaign acceptance is generally low.
- Customers with higher recency (recent purchases) respond better.

Customer Segments Identified

Examples:

- **High-income wine & gold buyers**
- **Middle-age family buyers focusing on meat products**
- **Young low-income customers with low engagement**



13. Conclusion & Recommendations

Conclusion

Customer Personality Analysis revealed clear segmentation based on income, age, and product preferences.
Understanding these groups helps the company focus marketing efforts efficiently.

Recommendations

- Target **high-income wine and gold buyers** for premium campaigns.
- Promote **family-oriented product bundles** to middle-aged customers.
- Improve **web-store experience** to increase online purchases.
- Use personalized email campaigns to improve campaign response rates.
- Offer discounts to low-engagement customers to reactivate them.