UNCOVERING CUSTOMER PATTERNS AT COSTCO WHOLESALE: AN EXCEL-DRIVEN ANALYSIS OF DEMOGRAPHICS, INCOME & SPENDING

BY

OYIBO OGHENENYERHOVWO DORCAS

INTRODUCTION

In 2022, Costco Wholesale served a diverse community of shoppers—from young professionals to large families—each with unique spending habits and lifestyles. Beneath every purchase receipt lies a story about income, age, and personal choices that shape how customers engage with the brand.

This analytical report explores those stories through data—examining how factors like gender, age, profession, work experience, and family size influence annual income and spending behaviour. Using Excel as the primary tool, patterns were uncovered that reveal not just who Costco's customers are, but how the company can better connect with them, improve retention, and strengthen its market position.

Image I: A Customer Shopping in Costco Wholesale Store

a. Objective:

The objective of this analysis is to provide actionable insights into customer patterns at Costco Wholesale Corporation for the year 2022. By diving into demographics (age, gender, profession, work-experience, family size) and financial behaviour (annual income, spending score), the goal is to help the organisation understand customer segments, guide marketing & operational decisions, and ultimately support retention and growth.

b. Problem being Addressed

Costco faces the challenge of reaching and retaining its diverse customer base while operating under its "extremely low prices and efficient operations" model. By not fully understanding how customer demographics relate to income and spending behaviour, the company may be missing opportunities in segmentation, targeted campaigns, product-mix optimisation, and personalised service. This analysis addresses the gap: Who are the customers? How do their demographics map to income and spending? What unseen groups are underleveraged?

c. Key datasets and methodologies

The dataset comprises 2,000 customer records (rows) and eight columns (features) representing unique customerID, gender, age, annual income, spending score, profession, work experience and family size (independent & dependent variables). Methodology: All analysis was conducted using Excel (e.g., filtering, formula creation, pivot tables, charts). No programming code (Python/Power BI) was used— showcasing how powerful Excel can be for data-analysis portfolios. We apply an analytical framework to address questions like:

1. What is the percentage of gender by annual income?
2. How can age group be compared to annual income?
3. What is the total annual income earned by each profession?
4. How does family size relate with the spending score?
5. What is the total number of customers by age groups?
6. What is the correlation between work experience & total customers to family size?
7. What is the spending score category by average annual income?

STORY OF DATA

a. Data Source

The data was sourced from **Kaggle—Customer.csv dataset**

b. Data Collection Process

The dataset was collected by the original uploader and publicly shared. For this analysis we assumed it represents customers of Costco in 2022. The process involved downloading the CSV from Kaggle and loading it into Excel for cleaning and analysis.

c. Data Structure

The table has 2,000 rows and eight columns:

- CustomerID (unique)
- Gender (categorical: Male/Female)
- Age (numeric)
- Annual Income (numeric, presumably thousands)
- Spending Score (1–100)
- Profession (categorical)
- Work Experience (numeric, years)
- Family Size (numeric)

d. Important features and their significance

- **Gender**: To compare income/spending by male vs female; useful for targeted marketing.
- **Age & Age Groups**: Age helps map lifecycle & spending behaviour; the derived Age Groups allow segmentation (Gen Z, Millennials, etc).
- **Annual Income**: A key dependent variable showing financial capacity of the customer—important for value-based segmentation.
- **Spending Score (1–100)**: Proxy for how much a customer is likely to spend or engage. Helps to identify high-value vs low-value customers.
- **Profession**: Helps tie income/experience/spending to occupation, enabling vertical segmentation (e.g., "Engineers vs Teachers").
- **Work Experience**: Indicates maturity in profession, possibly correlates with income, family size, and spending.
- **Family Size**: Larger family size may drive higher spending and influence product needs; important for Costco's bulk model.

e. Data Limitations or Biases

- The dataset may not cover all customers of Costco—potential sample bias (only those captured/uploaded).
- Demographics may be skewed toward certain geographies or customer types (no region column).
- Annual Income is likely rounded/approximate and may not reflect household income or part-time work.

- Spending Score is abstract (1–100) and the methodology for scoring is unclear—may not reflect actual spending currency.
- Profession categories may be broad or inconsistent.
- Derived columns (Age Groups, Spending Score Category) depend on thresholds which may not be universally meaningful.

f. Suggestions
- Collect geographic/region data to enable geo-segmentation.
- Clarify spending score methodology or convert to actual currency spend.
- Expand profession categories for finer granularity.
- Add date/time of transaction or membership tenure for temporal analysis.
- Consider adding loyalty or membership type variables (e.g., regular vs executive).

DATA SPLITTING AND PREPROCESSING

a. Data Cleaning
- Formatted the Annual Income column by adding currency formatting (e.g., US$) to improve readability.
- Verified no duplicate CustomerIDs were present.
- Confirmed no empty cells in the dataset.
- Standardised profession names (e.g., "Teacher " vs "teacher").

b. Handling missing values
- Since no empty cells were found, no imputation was required.

c. Data Transformation:

Four columns were created to enhance interpretability:
- Age Groups
  Formula: =IFS([@Age]<=11,"Gen A",[@Age]<=27,"Gen Z",[@Age]<=43,"Millennials",[@Age]<=59,"Gen X",[@Age]<=78,"Baby Boomers",[@Age]<=96,"Silent",[@Age]>=97,"Greatest") (Chris Beytes, 2024)
- Spending Score Category
  Formula: =IFS([@[Spending Score (1–100)]]<=25,"Low",[@[Spending Score (1–100)]]<=50,"Medium",[@[Spending Score (1–100)]]<=75,"High",[@[Spending Score (1–100)]]<=100,"Very High")
- WorkExperience2

Formula: =CONCATENATE([@[Work Experience]],"years")

- Family Size member

   Formula: =CONCATENATE([@[Work Experience]],"years")

d. Data splitting

- Split the data into training/validation sets is not applicable here (since no predictive modelling). However for segmentation analysis we created separate pivot tables (e.g., by Age Group, by Profession).

e. Industry Context

In the retail wholesale industry (Costco's model), understanding customer segments is vital: membership patterns, bulk purchasing, family size, income levels all influence retention, product-mix and pricing strategy. This analysis fits into the industry context of retail intelligence, customer lifetime value (CLV) modelling, and segmentation.

f. Stakeholders

Key stakeholders include: the Executive & Senior Management team (for strategic direction), the Sales & Marketing team (for targeting and campaigns), and the Customer Service team (for understanding customer profiles and tailoring service).

g. Value to the industry

By leveraging this demographic and behavioural data, Costco can refine its inventory, optimise promotional offers, improve membership segmentation, enhance retention programmes, and ultimately boost profitability via smarter customer-driven decisions.

CustomerID	Gender	Age	Annual Income ($)	Spending Score (1-100)	Profession	Work Experience	Family Size
1	Male	19	15000	39	Healthcare	1	4
2	Male	21	35000	81	Engineer	3	3
3	Female	20	86000	6	Engineer	1	1
4	Female	23	59000	77	Lawyer	0	2
5	Female	31	38000	40	Entertainment	2	6
6	Female	22	58000	76	Artist	0	2
7	Female	35	31000	6	Healthcare	1	3
8	Female	23	84000	94	Healthcare	1	3
9	Male	64	97000	3	Engineer	0	3
10	Female	30	98000	72	Artist	1	4
11	Male	67	7000	14	Engineer	1	3
12	Female	35	93000	99	Healthcare	4	4
13	Female	58	80000	15	Executive	0	5
14	Female	24	91000	77	Lawyer	1	1
15	Male	37	19000	13	Doctor	0	1
16	Male	22	51000	79	Healthcare	1	2
17	Female	35	29000	35	Homemaker	9	5
18	Male	20	89000	66	Healthcare	1	6
19	Male	52	20000	29	Entertainment	1	4
20	Female	35	62000	98	Artist	0	1
21	Male	35	96000	35	Homemaker	12	1
22	Male	25	4000	73	Healthcare	3	4
23	Female	46	42000	5	Artist	13	2
24	Male	31	71000	73	Artist	5	2
25	Female	54	67000	14	Executive	1	3
26	Male	29	52000	82	Artist	1	3
27	Female	45	68000	32	Healthcare	9	8
28	Male	35	78000	61	Artist	1	3

Image II: Dataset Before Data Preprocessing

CustomerID	Gender	Age	Age Groups	Annual Income ($)	Spending Score (1-100)	Spending Category	Profession	Work Experience	Work Experience2	Family Size_memb	Family Size2
1	Male	19	Gen Z	$15,000	39	Medium	Healthcare	1	1years	4members	4
2	Male	21	Gen Z	$35,000	81	Very High	Engineer	3	3years	3members	3
3	Female	20	Gen Z	$86,000	6	Low	Engineer	1	1years	1members	1
4	Female	23	Gen Z	$59,000	77	Very High	Lawyer	0	0years	2members	2
5	Female	31	Millenials	$38,000	40	Medium	Entertainment	2	2years	6members	6
6	Female	22	Gen Z	$58,000	76	Very High	Artist	0	0years	2members	2
7	Female	35	Millenials	$31,000	6	Low	Healthcare	1	1years	3members	3
8	Female	23	Gen Z	$84,000	94	Very High	Healthcare	1	1years	3members	3
9	Male	64	Baby Boomers	$97,000	3	Low	Engineer	0	0years	3members	3
10	Female	30	Millenials	$98,000	72	High	Artist	1	1years	4members	4
11	Male	67	Baby Boomers	$7,000	14	Low	Engineer	1	1years	3members	3
12	Female	35	Millenials	$93,000	99	Very High	Healthcare	4	4years	4members	4
13	Female	58	Gen X	$80,000	15	Low	Executive	0	0years	5members	5
14	Female	24	Gen Z	$91,000	77	Very High	Lawyer	1	1years	1members	1
15	Male	37	Millenials	$19,000	13	Low	Doctor	0	0years	1members	1
16	Male	22	Gen Z	$51,000	79	Very High	Healthcare	1	1years	2members	2
17	Female	35	Millenials	$29,000	35	Medium	Homemaker	9	9years	5members	5
18	Male	20	Gen Z	$89,000	66	High	Healthcare	1	1years	6members	6
19	Male	52	Gen X	$20,000	29	Medium	Entertainment	1	1years	4members	4
20	Female	35	Millenials	$62,000	98	Very High	Artist	0	0years	1members	1
21	Male	35	Millenials	$96,000	35	Medium	Homemaker	12	12years	1members	1
22	Male	25	Gen Z	$4,000	73	High	Healthcare	3	3years	4members	4
23	Female	46	Gen X	$42,000	5	Low	Artist	13	13years	2members	2
24	Male	31	Millenials	$71,000	73	High	Artist	5	5years	2members	2
25	Female	54	Gen X	$67,000	14	Low	Executive	1	1years	3members	3
26	Male	29	Millenials	$52,000	82	Very High	Artist	1	1years	3members	3
27	Female	45	Gen X	$68,000	32	Medium	Healthcare	9	9years	8members	8
28	Male	35	Millenials	$78,000	61	High	Artist	1	1years	3members	3

Image III: Dataset After Data Preprocessing

The following Pre-analysis, In-analysis and Post-analysis insights sections are displayed in the Costco wholesale report dashboard below:
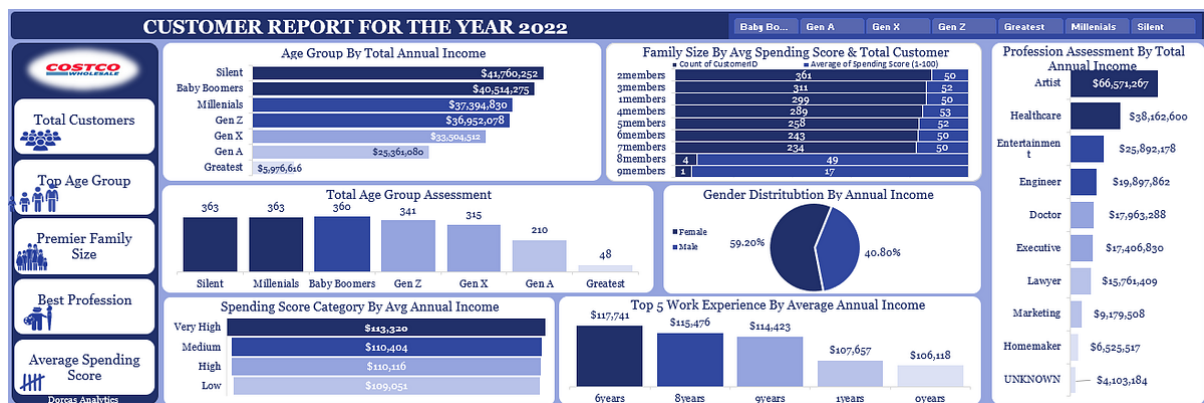

Image IV: Customer Report Dashboard for the year 2022

PRE-ANALYSIS

a. Key Trends

Before deep analysis, quick Excel explorations (sorting, filtering, pivot tables) revealed several notable early trends across the dataset:

- Silent Generation recorded the highest total annual income at $41,760,252, followed by Millennials and Gen Z.

- Most age groups—Baby Boomers, Gen X, Millennials, Greatest— predominantly had 2-member family sizes, while Gen Z leaned toward 3 members and Silent Generation clustered around 4 members.

- Females consistently outnumber males across all generations, ranging between 56% and 65%, suggesting a strong female customer base.

- Across all age groups, the top income-contributing professions remained stable: Artist, Healthcare, and Entertainment.

- Predominant work-experience categories varied by generation but centered around 6–9 years for younger groups and 7–9 years for older groups.

b. Potential Correlations

Although no correlation formula or matrix was used, Excel pivot comparisons suggested some possible relationships:

- Higher work experience groups (7–9 years) generally aligned with higher average annual incomes, especially in Silent, Millennial, and Gen X categories.

- Larger family sizes tended to appear in groups with higher total annual income (e.g., Silent Generation), indicating that income strength may support larger households.
- Profession clusters (Artist, Healthcare, Entertainment) repeatedly showed higher earnings across all generations, hinting at a strong alignment between profession and income potential.

c. Initial Insights

From these early patterns:
- The Silent Generation appears to be Costco's most financially influential segment despite representing an older demographic.
- Millennials and Gen Z provide a strong customer volume base, making them vital for long-term business retention and growth.
- The female-majority distribution indicates a consistent customer segment whose spending behaviour may be central to marketing strategies.
- Several generations share common traits (e.g., 2-member households, same top professions), suggesting predictable patterns Costco could leverage.

IN-ANALYSIS

a. Unconfirmed Insights

As deeper Excel pivot breakdowns continued, some insights appeared promising but required more validation:
- The assumption that larger family sizes always lead to higher income was inconsistent—Gen Z had a larger family size (3) but not the highest income.
- The idea that more work experience always leads to higher income did not hold for every group—some lower-experience brackets (e.g., 1 year in Baby Boomers) still showed strong salary presence.
- The professions with the highest frequency did not always perfectly align with highest spending behaviour, suggesting spending may depend on more than earnings alone.

b. Recommendations (During Analysis Stage)

During this stage, evolving insights began pointing toward strategic ideas such as:

- Focusing on professional clusters (Artist, Healthcare, Entertainment) because they consistently appear as high contributors across all ages.
- Creating age-group-specific value strategies, since spending and income patterns shift significantly between Gen Z, Millennials, and Silent Generation.
- Considering family-size-based product bundles due to the consistency of 2–4 member households across generations.

c. Analysis Techniques Used in Excel

The following Excel tools and methods shaped the analysis:
- Pivot tables to summarise totals, averages, and rank income by profession, age group, work experience, and gender.
- Slicers to dynamically filter profession, gender, and age groups for comparative insights.
- Sorting and filters to identify highest and lowest earning groups.
- IFS formulas for Age Groups and Spending Score Categories.
- Conditional formatting to highlight top-performing groups.
- Simple percentage calculations to evaluate gender contribution and population proportions.

POST-ANALYSIS AND INSIGHTS

a. Key Findings

After completing full pivot-table analysis:
- The Silent Generation emerged as the highest-earning group, supported by strong representation in high-earning professions and long work experience (8–9 years).
- Millennials and Gen Z provide strong customer numbers and stable income contributions, making them strategic long-term segments.
- Profession patterns remained consistent—Artist, Healthcare, and Entertainment dominated annually across all age groups.
- Females contribute the majority share of total income because they form the majority across every generation.
- Most age groups exhibit 2-member families, except Gen Z (3 members) and Silent (4 members), suggesting varied lifestyle and consumption structures.

b. Comparison with Initial Findings

The initial pre-analysis observations were validated, with slight refinements:

- Initial expectations about family size impacting income were nuanced—income varies more by work experience and profession than household size.
- The dominant profession clusters remained the same, reinforcing the importance of these income-driving occupations.
- The Silent Generation's financial strength, hinted early, was strongly confirmed through full analysis.

DATA VISUALIZATIONS & CHARTS

Charts and graphs used:

- Bar chart: Age Group by Total Annual Income and Profession Assessment by Total Annual Income
- Pie chart: Gender Distriubtion by Annual Income
- Column chart: Total Age Group Assessment and Top 5 Work Experience by Average Annual Income
- Funnel chart: Spending Score Category by Avg Annual Income
- 100% Stacked bar: Family Size by Avg Spending Score & Total Customer

These visualisations were embedded in the report to enable both technical and non-technical readers to quickly grasp key patterns.

RECOMMENDATIONS AND OBSERVATIONS

(a) Actionable Insights

- Develop profession-based loyalty packages for Artists, Healthcare workers, and Entertainment professionals who consistently appear as top earners.
- Design family-size-specific bundles (2–4 members), especially since these household patterns dominate across multiple generations.
- Strengthen engagement programs for Gen Z and Millennials, the largest customer groups, focusing on affordable yet value-driven offerings.
- Introduce premium offers for high-income groups like the Silent Generation and Millennials.

- Leverage female-oriented campaigns, since females form 56%–65% of customers in all categories.

b. Business Optimisations or Decisions

- Personalise marketing using age-group segmentation to match lifestyle and spending tendencies.
- Align inventory and product mix with top professions and common household sizes.
- Create multi-tier membership incentives tailored to income levels and generational behaviour.
- Build engagement strategies for groups with lower income representation (e.g., Greatest Generation).

c. Unexpected Outcomes

- Some low-experience brackets (e.g., 1 year in Baby Boomers) showed surprisingly high annual salary presence, contradicting expectations.
- Gen Z, despite larger family sizes, did not show proportionally increased total income compared to older age groups.
- The consistency of Artist appearing as a top profession across *all* generations was unexpected and may suggest dataset-specific skew or a strong customer segment with high engagement.

CONCLUSION

In this Excel-driven analytical report for Costco's customer base in 2022, we uncovered several important patterns: a dominant young-customer base (Gen Z/Millennials), clear differences in income by profession and age group, and nuanced relationships between income, spending score, and family size. The findings highlight that while higher income segments are valuable, moderate-income yet high-engagement segments (such as Teachers) are equally important.

For Costco to optimise its value proposition, tailor membership tiers, personalise marketing and adapt inventory/promotions by segment. This analysis demonstrates how Excel can deliver powerful insights and serves as a strong portfolio piece for recruiters looking for analysts who can bridge business and data.

REFERENCES AND APPENDICES

1. Chris Beytes, 2024. *The Generations.* Available at: https://www.growertalks.com/Article/?articleid=27014 (Accessed at: 15th November 2024)