ST312 Project Cover Sheet

Project Title: Consumer Demographics and Behaviour in the Luxury Shopping Sector

Word count: 4,967

Candidate ID(s): 50537, 44158

Date: 02-05-2025

Permission to use as an example

I do not give permission for my assignment to be used as an example in ST312 (in an electronic form).

1. **Introduction and Literature Review**

The ever-growing luxury product market is discreet and difficult to find relevant datasets on, as the often high-income consumers who often buy luxury items can be less willing to disclose their personal details or purchasing behaviour. Luxury brands aim to maintain an exclusive perception as part of their brand image because rarity is part of how they justify the dearness. Therefore, they keep their sales and client information as private as possible. Given the limited analysis available for this sector and our interest in researching it, we decided to analyse the luxury sector ourselves.

A key driver in influencing people to buy luxuries is the ability to display wealth and social status, known as ‘conspicuous consumption’ (Veblen, 1899) when done by those who can easily afford it, or ‘pecuniary emulation’ when done by those who can’t in an attempt to appear as though they’re wealthier than they are. This is partly why we chose to analyse wearable, more conspicuous luxuries specifically, as they’re likely to display clearer trends in their consumption.

It’s also possible that younger participants want to spend money on more conspicuous luxuries to be perceived as wealthy, whereas older people could be likely to buy luxuries because they enjoy them, or because they can more easily afford them (Eastman et al., 1999). Eastman’s study of six groups of self-report scale surveys validates the idea of age affecting conspicuous consumption but only acknowledges that gender may affect it and the surveys don’t give sufficient evidence to prove this so we will look into it for our research.

Simge Aksu argues that there are differences in behaviour and spending between men and women, such as that women care more about luxuries being flamboyant and men care more about the quality and exclusivity of them (Aksu, 2020). Aksu’s research was very specific, using only 27 low to middle income respondents, aged 17-21, in Turkey, and so in our research we will answer if gender affects luxury purchasing behaviour in more detail with substantially more respondents across a significantly wider range of demographics.

We will analyse this secretive sector so that luxury companies better understand their customers and which aspects of their products or branding appeal to specific demographics such as age, sex and income. Our research will benefit luxury brands as they could strategically market their products to particular demographics, to marketers and advertisers as they will learn about how much different demographics research and consume luxury products, to financial analysts as they will better evaluate growth and risks of this discreet market, and to academics who want to understand luxury consumer behaviours and socio-economic trends.

In this report, we use data from four European countries, with a 53:47 split of men to women and around 36% ‘millennials’ and 29% ‘generation X’, on average, with the rest being made up of ‘gen Z’, and ‘traditionals and baby boomers’. We will look into four of the main categories of luxury items which we believe to be a good representation of wearable luxury items: watches, jewellery, cosmetics and personal care, and fashion and accessories. Our dataset contains multiple research questions across our four product categories, but we found that most of them fall under the three main themes of attitudes and behaviours, expenditure, and market knowledge/ previous research, and so we chose to build our research aims around attitudes and behaviours, and expenditure. The main aims of this study are to discover:

1. What are the main attitudes and opinions shaping consumers’ likelihood to purchase luxury goods and how do they vary across luxury product categories and demographics?
2. How does luxury expenditure vary across luxury product categories and demographics?

By the end of this report, the data we've found will be used to analyse, answer and explain the significance of our research questions.

An interesting outcome of our first research aim relates to how the luxury sector balances exclusivity with marketing their products to as many consumers as possible (Dubois et al., 2005) which we can tell by assessing which demographics are more likely to purchase luxuries if they believe they're scarce. The second part of their study, a quantitative assessment of almost 2,000 management students from twenty countries using finite mixture modelling (Wedel and Kamakura, 1998) on their questionnaire's Likert-scale responses, was conducted in 2005 and so may no longer represent current behaviours, especially as their sample is limited in terms of age and somewhat biased in terms of gender (39.4 % female, mean age 26.5), whereas our data has a wide variety of age groups and a slightly better gender balance.

While researching our second aim, we found that, according to Varga (Varga, 2020), 39% of consumers aren't willing to spend more than €500 on luxury products, around 20% are comfortable spending between €2,001 and €5,000, with 14% are willing to spend over €5,000. Varga's study revealed a positive correlation between income and luxury expenditure, indicating that higher income is correlated with an increase in luxury spending. Our research will expand on how much luxury consumption varies by demographics as part of our second aim because our data is sampled by an online survey rather than convenience sampling via Facebook.

Our analysis' will answer our research questions and provide insights into the significance of relevant factors, and the effects of their interactions.

This report will describe the data we have used, including transformations and interactions, and the methods we've used to transform and analyse the raw data, explain the results and findings of our analyses, disclose any limitations imposed by either our data or methods, and summarise the further implications of our results.

2. **Data and Statistical Analysis Method**

i. Data

Data Description

We have four identical survey datasets from the UK, Italy, Germany and France, completed in June 2021 by Statista of adults (18+) who purchased luxury items in the past 2 years. Each of the four datasets contains survey questions about the respondents' attitudes and habits regarding luxury consumption.

The questions in each of our four datasets are divided by luxury products: Watches, Jewellery, Cosmetics and Personal Care, and Fashion and Accessories. The datasets also contain questions on second-hand and counterfeit products, but we won't be analysing those in our research.

The questions for each product are identical across the four countries. For each survey question, we have the total number of respondents who picked each of the response options for that question, and that grand total is broken down by demographics: sex (male or female), age (generation), and income.

Each dataset contains seventy-nine survey questions across the four luxury products. The questions vary from qualitative questions such as which statements the respondents agree with, or which statements apply best to them, to more quantitative questions like "How many watches over $500 do you own?". The breadth of detail and range of questions asked were certainly deciding factors in us choosing to use these datasets for our project, as they allowed for thorough, insightful and in-depth analysis.

The primary challenge we faced in our datasets is that, for each survey question, we only had aggregate data, but no individual data on each respondent's personal demographics. These demographics weren't cross-tabulated, meaning they captured marginal summaries but not the joint distributions of the demographic variables. As a result, it wasn't possible to reconstruct or infer individual-level profiles or conduct multi-dimensional analyses across demographics, preventing the dataset from being analysed as individual-level data.

Given this unusual format, our datasets didn't naturally lend themselves to a standard type of analysis or modelling. Despite the participants being clustered within four countries, multilevel modelling wasn't suitable, as it requires individual-level data to accurately model within- and between-cluster variation. Additionally, pseudo-individual-level datasets would've incorrectly assumed independence and identical characteristics among individuals within each group, failing to capture real variability.

Given the limitations of the datasets, we had to restructure the data into a more standard format for meaningful analysis. We decided to convert it into a series of three-way contingency tables for each question, with each contingency table capturing the joint distribution of the answer choices, one demographic (age group, gender, or income group) and country. This allowed us to summarise the data from the four datasets and provided an easier and more structured format to work with. In particular, this approach also prepares the data for established methods of analysing categorical data.

Data Transformation

A. Transforming the Data into Contingency Tables:

The survey data was extracted from four CSV files. For each product and question, we chose responses to single-pick survey questions and removed multi-pick questions to preserve the assumption of independent observations.

We segmented each question table into three three-way contingency tables: each one contained data on one demographic group (Gender, Age, or Income), Answer Choices for the question, and the Country variable for control purposes.

We finally compiled the tables into a single file for each product, each with all questions for that product.

Full details on how the tables were constructed and cleaned are provided in the appendix.

B. <u>Preparing the Data for Modelling – Long Format</u>:

To prepare the data for model fitting, we reshaped each contingency table into long format, where each row represented a unique combination of demographic group, country, and answer choice, along with the corresponding count.

This structure is common in statistical modelling as it supports flexible inclusion of categorical predictors and interaction terms.

Using a long-format structure is a standard practice in statistical modelling and for fitting Generalised Linear Models (GLMs), as it allows each observation to be represented individually and supports flexible inclusion of categorical predictors and interactions. For all model fitting stages in this project, we first converted to the long-format.

C. <u>Grouping of Survey Questions</u>:

The primary challenge lay in knowing how to carry out the analysis on contingency tables. On one hand, fitting models on all contingency tables and analysing every estimated variable would have been impractical, time-consuming, and unlikely to yield interpretable or meaningful results. On the other hand, averaging coefficients across unrelated questions would have oversimplified the analysis and confounded many variables and effects.

To avoid both extreme cases, we adopted a theme-based approach. We created three distinct groups of related questions, each focusing on a particular theme that relates to one of our two research questions. Any questions that fell outside the scope of our research were excluded.

ii. Modelling

Model Assumptions

Given the categorical nature of the data and the fact that we were working with counts, log-linear models offered a good framework for analysis, as they are specifically designed to handle multi-dimensional contingency tables.

Log-linear models can be viewed as a specific class of Poisson Generalised Linear Models (GLMs) designed for the analysis of categorical count data. In our project, we fit a model on each contingency table, giving us three models per question.

Assumptions of Poisson GLM Models:

Table 1 - Poisson GLM Assumptions

Assumption	**Explanation**	**Status/ Action taken**
Independence of Counts	The counts are assumed to be statistically independent (no autocorrelation). For multi-pick questions, the independence of counts assumption may be partially violated, as a single respondent can contribute to multiple counts across different categories, introducing potential dependence between observations. Note that in survey data, true independence between respondents may be difficult to verify due to potential hidden clustering or shared characteristics.	Only single-pick questions were analysed.
Distribution of Counts	Counts follow a Poisson distribution, meaning that the mean and variance of	Overdispersion was tested by calculating the

	each observation are equal (equidispersion). If the data shows overdispersion (variance > mean), we must switch to another model, like Negative Binomial GLM.	Pearson residual deviance / residual df for each model in R. Ratios > 1 were common, so Poisson GLMs were rejected in favour of Negative Binomial GLMs.
Log Link	The natural logarithm of the expected count is modelled as a linear combination of the variables (main effects and interactions).	Inherent to the Poisson GLM model. Satisfied by model choice.
Linearity Holds on the Link (Log) Scale	Linear relationship is assumed between the log of the expected count and the predictors (on the log link scale).	**Already satisfied**: we are only using categorical variables.
No Perfect Multicollinearity among Variables	Variables should not be perfectly linearly dependent, as this prevents unique estimation of coefficients. (high multicollinearity does not violate this assumption but can lead to unreliable coefficient estimates).	Checked by comparing design matrix rank to number of variables in R. No rank deficiency detected; assumption satisfied.

Further Assumptions for Log-linear Models for Contingency Tables:

Table 2 - Log-linear Models for Contingency Table Assumptions

Assumption	**Explanation**	**Status/ Action taken**
Variables are Categorical	All variables are categorical, and the model explains the joint distribution of these variables via cell counts in a contingency table.	**Already Satisfied**: Our variables (Age, Gender and Income) are all categorical.
Each Cell represents a Cell Count	Each cell represents a count corresponding to a combination of categorical variable levels.	Our data contained percentages, which were all removed.
Symmetry	Log-linear models do not distinguish between dependent and independent variables (no single outcome variable); all variables are treated symmetrically as explanatory variables, and we model the joint distributions between the variables, not a response conditioned on predictors. This property makes log-linear models particularly suitable for exploring associations without assuming causal directionality.	**Already Satisfied**: Satisfied by default by the nature of log-linear models, which explain joint distributions.
Variables are Main Effects and Interactions	The model includes main effects and interaction among categorical variables to explain the observed cell counts, rather than treating variables as covariates or distinguishing between predictors and outcomes.	**Already satisfied**: Structure matched this: all models included main effects and two-way interactions between Answer, Group, and Country.

Model Description

For all groups, we fit a Negative Binomial GLM on each contingency table, giving us three models per question.

This Negative Binomial model relaxes the equidispersion assumption of the Poisson distribution by introducing an additional dispersion parameter to capture the overdispersion, allowing the variance to exceed the mean.

As with the Poisson GLM, the expected counts were modelled using a log link function as a linear combination of the variables. Importantly, the Negative Binomial GLM model retains the core assumptions of Poisson GLMs we have mentioned. As our data did not violate any of them, we safely proceeded with this model.

<u>We model each count as</u>:

$$\log\left(E\left[y_{ijk}\right]\right) = \lambda + \lambda_i^{Response} + \lambda_j^{Group} + \lambda_k^{Country} + \lambda_{ij}^{Response \times \text{Group}} + \lambda_{ik}^{Response \times \text{Country}} + \lambda_{jk}^{Group \times \text{Counry}} + \lambda_{ijk}^{Response \times \text{Group} \times \text{Country}}$$

<u>where</u>:

1. y_{ijk}: Observed count for the cell corresponding to level i of Response, level j of Group and level k of Country

2. λ: Intercept (overall mean log effect)

3. $\lambda_i^{Response}, \lambda_j^{Group}, \lambda_k^{Country}$: Main effects for levels i, j, k of Response, Group, Country respectively

4. $\lambda_{ij}^{Response \times \text{Group}}, \lambda_{ik}^{Response \times \text{Country}}, \lambda_{jk}^{Group \times \text{Counry}}$: Two-way Interaction effects

5. $\lambda_{ijk}^{Response \times \text{Group} \times \text{Country}}$: Three-way interaction effect across all variables

One level of each variable is designated as the reference category, and the corresponding parameters for these reference levels are set to zero to ensure model identifiability.

Each count, y_{ijk} , is assumed to follow a Negative Binomial distribution:

$$y_{ijk} \sim \mathrm{NegBin}(\mu_{ijk}, \theta)$$

where $\mu_{ijk} = \mathrm{E}(y_{ijk})$ denotes the expected count for the cell corresponding to level i of variable $\mathrm{Response}$, level j of variable Group, and level k of variable $\mathrm{Country}$, and θ is a dispersion parameter accounting for overdispersion in the data.

The variance of y_{ijk} is modelled as:

$$\mathrm{Var}(y_{ijk}) = E(y_{ijk}) + \frac{E(y_{ijk})^2}{\theta}$$

iii. Analysis

Throughout our analysis, we have only focused on two-way interactions between survey responses and demographic groups (age, gender, income), as these captured how answer patterns varied across subgroups, beyond what would be expected from the main effects only. These interaction terms allowed us to detect over- or under-representation of particular demographics in specific responses, relative to reference levels.

Interactions involving countries were included in all models as control terms to account for structural national differences, but were not interpreted, since they were not of substantive research interest.

Likelihood Ratio Tests (LRTs)

For each group, we carried out a LRT to compare two nested models:

A. Main effects only model
B. Model with main effects and Group × Answer interaction

This was done for every Question × Demographic interaction independently in each of our three groups.

Formally, we tested (*A full breakdown is provided in the appendix.):*

H_0: There is no interaction between response and group.

H_1: There is an interaction between response and group.

This allowed us to determine whether including the interaction led to a statistically significant improvement in model fit, so we could see whether response patterns differed systematically across demographic groups. Identifying such differences helped us assess whether demographic characteristics were meaningfully associated with the way individuals responded to particular survey questions, within the scope of our data.

It is important to note that the interactions were assessed across all levels of the categorical variables, not just single dummy pairs. This allowed us to capture overall association structure, rather than focusing on isolated comparisons.

The resulting p-values were summarised in a pivot table (questions × demographics) with conditional formatting applied to highlight statistically significant results ($p < 0.05$). This informed the selection of questions for deeper analysis.

Dummy Variables Interactions Analysis

For questions that showed some evidence of a demographic–response association, we examined individual interaction coefficients at the dummy-variable level between the question and the demographic.

This approach enabled a more detailed, context-specific interpretation of how particular demographic groups responded to individual survey items.

They are interpreted as follows:

Positive interaction coefficient:

This indicates that the observed count for the demographic subgroup selecting that specific answer is greater than expected under an additive main effects model, relative to the reference levels reference response and reference group).

This means there is overrepresentation for that subgroup-answer combination compared to the baseline.

Negative interaction coefficient:

This shows that the observed count for the demographic subgroup selecting that specific answer is less than expected under the additive main effects. This suggests underrepresentation compared to the baseline expectation.

These interpretations were made within the bounds of this dataset, and do not imply population-level generalisations or causality.

3. **Findings and Results**

i. LRT Results – Purchase Frequency and Decision Making

The two following tables show the results for purchase frequency and decision-making groups.

Table 3 - LRT Results for Purchase Frequency

Demographic Question	**Age**	**Gender**	**Income**
How often do you buy cosmetic products from premium/luxury brands for yourself or someone else?	0.9947	0.9984	0.9946
How often have you bought fashion and accessories from a premium/luxury brand in a physical store in the past 2 years? If you are not sure, please estimate.	0.8638	0.9744	0.8533
How often have you bought fashion and accessories from a premium/luxury brand on the internet in the past 2 years? If you are not sure, please estimate.	0.8427	0.9927	0.9150

Table 4 - LRT Results for Decision Making

Demographic Question	**Age**	**Gender**	**Income**
Before buying your last cosmetic product from a premium/luxury brand, how long did you think about it?	0.9984	0.9957	0.9976
Before buying your last piece of jewellery from a premium/luxury brand, how long did you think about it?	0.9653	0.9831	0.9926

Before buying your last watch from a premium/luxury brand, how long did you think about it?	0.9529	0.9904	0.9817
How long did you usually think about/plan before buying fashion and accessories from a premium/luxury brand?	0.9357	0.9994	0.9982

Looking at two tables for our decision making and purchasing frequency groups, we notice that all our p-values are very high, all well above typical thresholds for statistical significance. The smallest value was 0.8533, which is still very high. This suggests that there is insufficient evidence to conclude that, in this survey, any of these nine interactions yielded any significant improvement in the fit of the main effects model.

There is little evidence to suggest a meaningful statistical relationship between the respondents’ demographics and their patterns of answers to questions about luxury purchasing habits and behaviours. This means that, in this survey, respondents across different age, gender, and income groups answered similarly when asked about their luxury shopping habits and decision-making.

Therefore, according to our data, belonging to a particular demographic didn’t appear to influence how spontaneous their luxury purchases were, or how much thought they gave before doing so.

One possible explanation is that attitudes and behaviours around luxury consumption may be becoming more consistent across demographics, though further research would be needed to confirm this. Although differences between demographics may still exist in the broader population, this analysis provides no statistical evidence of such associations in this dataset.

ii. LRT Results – Spending Habits

Table 5 - LRT Results for Spending

Demographic Question	Age	Gender	Income
How much are you willing to pay for a cosmetic product from a premium/luxury brand for yourself or someone else?	0.9653	0.9211	0.3374
How much are you willing to pay for a piece of jewellery from a premium/luxury brand for yourself or someone else?	0.8559	0.5837	0.1722
How much are you willing to pay for a watch from a premium/luxury brand for yourself or someone else?	0.9933	0.7190	0.1265
How much are you willing to pay for fashion and accessories from a premium/luxury brand for yourself or someone else? / Willingness to pay: clothes	0.9999	0.9793	0.2323

Similarly to the decision making and purchase frequency tables, the p-values for the age and gender demographics are insignificant, providing little evidence for any difference in how people from different demographics responded to the willingness-to-pay questions. People across these groups reported similar price points for our products.

However, the p-values for the income interactions were significantly lower than those for other demographics. However, they are still higher than typical significance thresholds (1%, 5%, 10%), suggesting a possible but inconclusive link between income and willingness-to-pay.

iii. Dummy Variables Interactions – Decision Making

1. Age × Longer Time Periods:

The models revealed that younger respondents showed consistently positive interaction estimates for longer planning times, relative to the reference group (older respondents and the baseline answer choice). For instance, in the context of fashion and accessories, the estimate for "Over a year" among younger individuals was as high as +1.93, and +1.63 for watches. These results indicate that younger respondents were more likely than expected to report longer deliberation periods before making a purchase, contradicting common that younger people are more spontaneous or impulsive. This may reflect greater hesitation, financial caution, or the aspirational nature of luxury consumption among younger consumers.

2. Income × Decisiveness:

Across multiple product categories, high-income respondents showed strongly negative interaction estimates for the response "Don't know", with values such as −1.13 (cosmetics) and −0.76 (watches), relative to the reference levels of answer and income group (low income). This suggests that individuals with higher incomes were less likely than expected to express uncertainty about how long they typically spend deciding on luxury purchases. This may reflect greater decisiveness, or more familiarity with the luxury market among wealthier individuals.

iv. Dummy Variables Interactions – Purchase Frequency

When examining the frequency of luxury cosmetic products, a clear pattern emerged among younger participants. Across all purchase frequency categories, younger individuals showed consistently negative interaction estimates, some as low as −1.21, relative to the reference group (older respondents and the "Several times a month" frequency). This shows that younger respondents were substantially less likely than expected to report regularly buying premium cosmetics. This may reflect differences in affordability, shifting consumer values, or changing cosmetic preferences among younger respondents.

However, this pattern reverses for fashion and accessories, both in physical stores and online, where the same group shows strongly positive interaction estimates, especially at the highest frequency levels. This suggests that younger consumers aren't disengaged from luxury altogether, but rather prefer fashion over beauty and cosmetics. These patterns may reflect shifting priorities, digital-native shopping habits, or a broader redefinition of luxury among the younger demographic.

v. Dummy Variables Interactions – Spending Habits

A. Age × High-End Cosmetic Purchases and Uncertainty:

In the model evaluating how much respondents were willing to pay for luxury cosmetic products, both younger and middle-aged individuals showed high positive interaction estimates for the highest price bracket and for uncertain responses, relative to the reference group (older respondents and the lowest price point). For example, the coefficient for “$200 or more” was +1.35 for younger and +1.24 for middle-aged respondents, while the interaction estimates for “Don’t know” reached +1.60 and +1.27, respectively.

These results suggest that, within the context of this survey, individuals who are younger were more likely than expected to have either a higher willingness-to-pay or a lack of clear pricing anchors, potentially reflecting aspirational behaviour, or inexperience with established luxury price norms.

B. Income × Willingness Across Price Brackets:

Across all product categories, high-income respondents showed substantially positive interaction estimates as price categories rose. For instance, in the watch category, high-income individuals had estimates of +2.82 for the $5,000–$10,000 bracket and +3.21 for $10,000 or more, relative to the reference group and lowest price range. This consistent upward trend highlights a clear economic segmentation, where higher-income consumers are overrepresented in the top spending brackets. Within the bounds of this survey, this indicates a strong relationship between income level and propensity to spend on luxury goods.

vi. Effect of Gender

While most gender-based effects were modest, a few patterns stood out. Male respondents showed moderately higher estimates for willingness to pay in upper price brackets. For example, an interaction estimate of +1.55 for watches priced at $10,000 or more, and +0.89 for luxury cosmetics at $200 or more, relative to the reference group. Additionally, men were slightly more likely to report purchasing cosmetics less frequently and less likely to choose uncertain response categories in decision-making.

Despite these results, gender did not seem to serve as a meaningful factor in determining how respondents approached luxury consumption. Across our categories, the interaction estimates were generally small, directionally mixed, and lacked an interpretable pattern, suggesting that, at least within this survey, gender didn't meaningfully differentiate consumer attitudes or behaviours.

4. **Limitations of Analysis**

While our analyses were effective in uncovering patterns within the survey data, there are several limitations, both statistical and structural, that limit the conclusions we can draw.
Firstly, although we used negative binomial GLMs to account for overdispersion in the count data, these models are descriptive and associative in nature. They detect statistical associations between demographic groups and response patterns but do not imply any causal relationships. For example, finding that younger respondents selected “strongly agree” more often than expected does not mean that age causes agreement — it only shows that, in this dataset, this subgroup over- or under-selected that option relative to the reference levels and marginal totals. A range of unobserved or uncontrolled variables (e.g., lifestyle, education, cultural influences) may be driving these observed associations.

Secondly, the use of aggregate contingency tables, rather than individual-level data, limited the granularity of our analysis. Because each table collapsed respondents into demographic-response groupings, we were unable to include covariates or control for other factors simultaneously. If individual-level data was available, we could have used multivariable regression techniques to account for confounders or assess interactions across more dimensions. Even with such data, however, causality would have remained hard to achieve without longitudinal tracking or experimental design.
Additionally, while we used LRTs to assess whether interactions between demographic groups and answers improved model fit, we did not assess the statistical significance of individual coefficients within those interactions. This was a conscious decision, given the risks of multiple testing and overinterpreting isolated dummy-level effects. Nonetheless, more rigorous meta-analyses could have allowed us to gain richer insights.

Our study was subject to several limitations stemming from both the structure of the data and the nature of the sample. The survey was conducted exclusively in Western European countries, with roughly a thousand respondents per country, limiting the geographic scope and representativeness of our findings. As the data was collected via an online survey, it likely excluded high-net-worth individuals, a key segment of the luxury market, who are less inclined to participate in such formats; indeed, only around 1% of respondents reported annual incomes above £123,600. Also, it is possible that rather than demographics, personality types or lifestyles may influence luxury spending behaviours and attitudes, and so we would benefit from finding surveys that better capture personality traits and lifestyle choices.

Moreover, the structure of the data itself had its own limitations. Questions varied in format and clarity across products and countries, and so differences in question format, phrasing, or local brand familiarity across countries may also have introduced additional noise. Also, although we excluded multi-pick questions to preserve count independence, even single-pick formats may have potentially carried implicit dependencies that were not accounted for in the model.

In summary, our findings reflect statistical patterns observed under the specific structure and assumptions of this dataset. They do not generalise beyond it without further evidence. Future work using individual-level, longitudinal, or experimental data would be needed to explore causality and strengthen the external validity of these findings.

5. Conclusion

This project explored how consumer attitudes and expenditure on luxury goods vary across product categories and demographic groups. In relation to our first research question, we found surprisingly little statistical evidence from our LRT tests that age, gender, or income had a strong association with most attitudinal and behavioural outcomes, particularly those related to decision-making and purchasing frequency. Across questions on how often respondents purchased luxury products and how long they considered their purchases, most interactions between demographic groups and responses were statistically insignificant. This suggests that, at least within this dataset, luxury consumption habits and attitudes may be converging across demographic lines, with consumers of different ages and genders behaving more similarly than expected.

However, other specific patterns were more notable. Age was mildly associated with product preference: younger consumers were less likely to purchase premium cosmetics but more inclined to buy fashion and accessories, particularly in higher frequencies. Gender, while mostly neutral, showed some product-specific variation. Fore example, men were somewhat more willing to spend on watches, while women showed relatively higher interest with premium cosmetics. These findings hint at subtle gendered preferences in product category engagement, though not strong enough to generalise broadly. Interestingly, older respondents tended to select more definitive answers to decision-making questions, possibly reflecting more self-awareness or experience in their purchasing behaviour.

Our second research question on how luxury expenditure varies, showed clearer results. While age and gender again played only minor roles, income appeared to be a more substantial factor. The LRTs showed very weakly significant evidence that income level may influence spending. This was further supported by the dummy-level analysis, which suggested that high-income respondents were markedly more willing to pay higher amounts across all product categories. This aligns with existing literature (Veblen, 1899; Varga, 2020) highlighting income as a key driver of conspicuous consumption.

Together, these findings show practical implications for luxury brands. Demographic targeting by age or gender may be less effective than commonly assumed, except in cases where product preferences clearly diverge. Marketing strategies may instead benefit from tailoring to income tiers, especially among consumers with high spending power.

Moreover, brands aiming at younger consumers might consider promoting more accessible luxury items such as fashion accessories, while maintaining aspirational appeal. Our findings are necessarily constrained by the dataset: limited geographic scope (Western Europe), modest sample size (~1,000 per country), and reliance on self-reported online surveys. Future work could enhance generalisability by incorporating a broader international sample and including behavioural or lifestyle data to explore deeper psychological drivers behind luxury consumption. Despite these limitations, our analysis provides a grounded statistical foundation for understanding how demographics intersect with luxury attitudes and spending patterns in today’s market.

Appendix

A. GitHub Repository – containing all of our code and necessary data:

https://github.com/44158-create/ST312_Code.git

B. Transforming the Data into Contingency Tables:

1. **Loading and Identifying Questions:**

   The original survey data for each product and country was loaded from the four CSV files, of which we had four (one per country).

   Single-pick survey questions were identified and selected and multi-pick questions were excluded to maintain the assumption of independent observations (see assumptions below).

2. **Extracting Question Tables:**

   We looped through each question in each dataset (so, through each country).

   The block (table) for each question was extracted. This was done based on structural markers within the file (e.g. "Survey Name:", "Base", "Question Type"), which helped us identify the start and end of each question table.

   We segmented each table into three separate tables based on demographics: One for Gender, one for Income and one for Age. We then added a "Country" column to each one.

3. **Removing Non-Count Columns:**

Columns with percentages (labelled "in %") and the grand total for each answer choice were removed, keeping only the raw counts necessary for contingency table construction (see assumptions).

4. **Storing Questions**:

Each survey question was also extracted and stored separately.

5. **Storing the results**:

The questions and the Pandas DataFrames were stored in a nested dictionary. The structure of the dictionary was organised first by product (cosmetics, fashion, watches, jewellery), then by country (United Kingdom, Germany, Italy, France). Within each country, each entry contained a list of questions, with each question linked to a set of the three extracted contingency tables (one each for Gender, Age Group, and Income Group).

6. **Constructing Three-Way Contingency Tables:**

We looped through our nested dictionary. For each question, three separate contingency tables were created:

a. Gender × Answer Choice × Country,
b. Age Group × Answer Choice × Country,
c. Income Group × Answer Choice × Country.

7. **Combining Country-Level Data:**

We combined the contingency tables together into 4 final CSV files, one for each of our four products.

The research question was written to the CSV file, starting with a "QUESTION:" marker for easier identification in later stages. Each question was then followed by three contingency tables, one for each demographic

C. Table of Groups of Questions:

Table 6 - Groups of Questions

Group	**Questions in Group**
Purchase Frequency	How often do you buy cosmetic products from premium/luxury brands for yourself or someone else? How often have you bought fashion and accessories from a premium/luxury brand on the internet in the past 2 years? If you are not sure, please estimate. How often have you bought fashion and accessories from a premium/luxury brand in a physical store in the past 2 years? If you are not sure, please estimate.
Decision Making	How long did you usually think about/plan before buying fashion and accessories from a premium/luxury brand? Before buying your last cosmetic product from a premium/luxury brand, how long did you think about it? Before buying your last piece of jewellery from a premium/luxury brand, how long did you think about it? Before buying your last watch from a premium/luxury brand, how long did you think about it?
Spending	How much are you willing to pay for a cosmetic product from a premium/luxury brand for yourself or someone else? How much are you willing to pay for fashion and accessories from a premium/luxury brand for yourself or someone else? / Willingness to pay: clothes How much are you willing to pay for a piece of jewellery from a premium/luxury brand for yourself or someone else? How much are you willing to pay for a watch from a premium/luxury brand for yourself or someone else?

D. Checking Poisson GLM Assumptions – Steps Taken:

1. Data Extraction:

   For each survey question, three 3-way contingency tables (Gender, Age Group, and Income Group) were extracted and transformed into long-format datasets.

2. Model Specification:

   For each Question × Demographic combination (Gender, Income or Age), a Poisson GLM was fitted with Count as the response variable and Answer, Country, and Group as predictors.

3. Overdispersion Check:

   Overdispersion was assessed by calculating the ratio of the Pearson residual deviance to the residual degrees of freedom for each fitted model.

   Results: The results have revealed substantial overdispersion in the data. In many cases, the ratio of the Pearson residual deviance to the residual degrees of freedom significantly exceeded 1, violating the equidispersion assumption.

   Next Step: Consequently, the Poisson GLM framework was deemed inappropriate for modelling the data, and we decided to use the Negative Binomial GLM instead.

4. Multicollinearity Check:

   It was assessed by checking the rank of the design matrix relative to the number of estimated parameters.

   Result: No perfect multicollinearity was detected in the data.

E. LRT for Main Analysis – Broken Down:

Formally speaking, let:

1. The reduced model be:

$$\log\left(E\left[y_{ijk}\right]\right) = \lambda + \lambda_i^{Response} + \lambda_j^{Group} + \lambda_k^{Country}$$

2. The full model be:

$$\log\left(E\left[y_{ijk}\right]\right) = \lambda + \lambda_i^{Response} + \lambda_j^{Group} + \lambda_k^{Country} + \lambda_{ij}^{Response \times \text{Group}}$$

We test:

H_0: $\lambda_{ij}^{Response \times \text{Group}} = 0$

H_1: $\lambda_{ij}^{Response \times \text{Group}} \neq 0$

In words:

H_0: There is insufficient evidence at the 5% level that there is an interaction between Response and Group

H_1:: There is enough evidence at the 5% significance level to suggest an interaction between Response and Group

We compute the LRT test statistic as:

$\chi^2 = 2(\ell_1 - \ell_0) \sim \chi^2\,(\text{df})$

where:

1. ℓ_0 and ℓ_1 are the log-likelihoods of the full and reduced models, respectively.
2. df: degrees of freedom = number of parameters in full model - number of parameters in reduced model

The resulting p-value was obtained using a chi-squared distribution with degrees of freedom equal to the difference in model parameters.

F. Data Recoding - Collapsing Demographic Categories:

We decided to merge some columns/levels of our income and age variables to reduce their number and give them easier names to interpret.

Table 7 - Income Variable Column Collapse

Original Label	**New Group**	**Action**
up to 22 800	Low Income	Renamed to “Low Income”
22 800 up to 43 200	Medium Income	Renamed to “Medium Income”
43 200 up to 98 400	High Income	Summed into "High Income"
98 400 and more	High Income	Summed into "High Income"
prefer not to say	-	Removed from dataset

Table 8 - Age Variable Column Collapse

Original Label	**New Group**	**Action**
generation x (baby bust) (1965–1979)	Middle Aged	Renamed to “Middle Aged”
traditionals & baby boomer (1922–1964)	Older	Renamed to “Older”
millennials / generation y (1980–1994)	Younger	Summed into "Younger"
igen / gen z (1995–2012)	Younger	Summed into "Younger"

G. Other Data Cleaning carried out:

In addition to collapsing demographic categories into simplified groups, a series of general data cleaning steps were applied to prepare the raw CSV input files for analysis. These steps were necessary to ensure consistent formatting, prevent parsing errors, and support reliable statistical modelling.

The following procedures were performed:

1. **Manually Removing Trailing Commas**:

   We noticed a series of trailing commas in our CSV files, which caused errors in processing, so we manually removed them.

2. **Removal of Leading Spaces**

   Each line in the raw data was stripped of leading whitespace to standardise the structure of question blocks and table headers.

3. **Exclusion of Non-Tabular Lines**

   Lines that did not contain comma-separated values or alphanumeric characters were discarded to eliminate notes, blank lines, and formatting artifacts.

4. **Cleaning Column Headers**

   Column names were:

   - Stripped of whitespace.
   - Converted to lowercase.
   - Normalised by replacing multiple spaces with a single space.
   - Removed non-ASCII and hidden characters.

5. **Removal of Duplicate Columns**

   We noticed columns being duplicated in the initial data retrieval stage. So, any duplicated column names were removed.

6. **Character Normalisation in Data Cells**

   All non-ASCII characters (e.g., accented letters) and embedded quotation marks were removed from string values to ensure compatibility with analysis tools.

These cleaning procedures were applied uniformly for all survey blocks before reshaping the data into long format and fitting the GLMs.

H. Reference Levels:

In all Negative Binomial GLMs, categorical predictors were dummy-coded. For each model, the reference level for the variable of interest (typically the set of answer choices) was consistently set to the first level as it appeared in the dataset. This was done to maintain consistency across models.

Note: Since the specific answer choices varied between questions, the reference category differed across models. However, the selection rule (first level) was applied uniformly.

Table 9 - Reference Levels

Variable	**Reference Level**
Age	
Income	
Gender	
Answer	Varies by model, but always first one in dataset.
Country	UK

Bibliography

(Anderson, 2025)
Anderson, C. J. (2025). Log-linear models for contingency tables [Lecture notes]. University of Illinois at Urbana-Champaign. https://education.illinois.edu/docs/default-source/carolyn-anderson/edpsy589/lectures/6-loglinear/6_loglinear_models_beamer-online.pdf

(Aksu, 2020)
Aksu, S. (2020). Luxury perception of low and middle income Generation Z and their luxury consumption motivations. https://www.researchgate.net/publication/347962407

(Dubois et al., 2005)
Dubois, B., Laurent, G., & Czellar, S. (2005). Consumer rapport to luxury: Analyzing complex and ambivalent attitudes (Working Paper No. 849). HEC Paris. https://www.researchgate.net/publication/4883928_Consumer_Rapport_to_Luxury_Analyzing_Complex_and_Ambivalent_Attitudes

(Eastman et al., 1999)
Eastman, J. K., Goldsmith, R. E., & Flynn, L. R. (1999). Status consumption in consumer behavior: Scale development and validation. Journal of Marketing Theory and Practice, 7(3), 41–52. https://www.researchgate.net/publication/285675797_Status_Consumption_in_Consumer_Behavior_Scale_Development_and_Validation

(Varga, 2020)
Varga, E.-I. (2020). Study on the personal luxury goods market. Acta Marisiensis. Seria Oeconomica, 14(2), 33–42. https://doi.org/10.2478/amso-2020-0011

(Veblen, 1899)
Veblen, T. (1899). The theory of the leisure class: An economic study of institutions. McMillan. https://moglen.law.columbia.edu/LCS/theoryleisureclass.pdf

(Wedel and Kamakura, 1998)
Wedel, M., and Kamakura, W. A. (1998). Market segmentation: Conceptual and methodological foundations. Kluwer Academic Publishers. https://www.researchgate.net/publication/243763942_Market_Segmentation_Conceptual_and_Methodological_Foundations