Smart Watch Preference Analysis

A Choice-Based Conjoint Analysis Approach to Optimizing Product Strategy

----

Thao Le

Table of Content

I. Problem Description

1.1 Business Objective

WearTech Innovations, a smart watch manufacturer, is struggling in a highly dynamic and competitive market. Their current direct marketing campaigns and product development strategies are not yielding optimal results, largely due to a critical lack of understanding regarding precise customer preferences. Without clear insights into what customers truly value, WearTech risks misallocating resources on features that don't drive sales or overlooking attributes that could be significant differentiators. This can lead to products that are either over-engineered in certain aspects or fail to meet core customer expectations. Current marketing messages often lack the specificity to resonate with distinct customer segments. This results in broad, inefficient campaigns that fail to highlight the most appealing attributes to the right audiences, leading to wasted marketing spend and lower-than-desired conversion rates for their smart watch offerings.

This survey systematically captured customer choices for smart watches, which were described by four key attributes: Color (Black, Lilac, White), Brand (Apple, Samsung, Fitbit), Price (NOK 990, 2390, 3190, 6190), and Battery life (10, 18, 40, 80 hours). Additionally, valuable demographic data, including gender, age, and region, was collected from each participant to enable deeper segmentation. The core problem now lies in effectively transforming this rich, raw survey data into actionable, data-driven insights that can directly inform WearTech's strategic decisions in product design, pricing, and targeted marketing, ultimately improving their market position and profitability.

1.2 Data Mining Objective

This project aims to provide WearTech Innovations with a quantitative understanding of consumer smart watch preferences by exploring the survey data and investigating respondent choices, using appropriate modeling methods like multinomial logistic regression to identify key influencing variables (DV: choice, IVs: attributes & demographics), validate and compare model performance (e.g., linear vs. nominal attribute effects), estimate the best model with full data to report findings on attribute importance (most/least important factors) and ideal product characteristics, and finally offer actionable recommendations for optimized product design, pricing, and targeted marketing campaigns to enhance customer satisfaction, increase market share, and improve financial performance.

II. Research Methodology and Results

2.1 Exploratory Analysis

Our dataset is complete with no missing values, structured across three interconnected dataframes. The alt_data dataset contains the specification of 144 alternative products that were shown to the respondents. Each product has 4 attributes including color, brand, price, and battery life. For example, the color of the product number 1 is black, the brand is Apple, the price is 990 NOK and the battery_life is 10 hours.

The sets_data data frame contains the choice sets that have been shown to respondents so that they can make their decisions. There are 1000 choice sets that have been shown to 500 respondents. For example, the first choice set (setID = 1) has 3 alternatives: the first alternative (productID1) is product number 12, the second alternative (productID2) is product number 126, and the last alternative (productID3) is product number 60.

The response_data data frame contains information about the respondents' choices. Each respondent only has to do 12 tasks. Each task consists of 3 different types of products from which respondents need to choose. For example, respondent number 1031 (respID = 1031) made 12 different choice tasks. In her first task (taskID = 1), she was exposed to the choice set number 165 (setID = 165) which contains product number 112, 64, and 42. Her choice was the third alternative (choice = 3), meaning that she had chosen product number 42.

The gender distribution of our survey participants is fairly even, although there are slightly more male respondents than female respondents. The age range of our dataset is between 18 and 60, with the majority of respondents being around 30 years old.

We note that all columns in response_data and sets_data file are categorical variables. Therefore, we convert those columns into factor type for further analysis. For alt_data, we only convert the `product` column and keep the type of other variables the same for further analysis.

2.2 Variable Identification

To model customer choices effectively, we identify the following variables:

- Dependent Variable (DV):
    - *choice*: This binary variable indicates whether a specific smart watch alternative was chosen by the respondent within a given task. When the data is transformed to a long format, each row will represent an alternative, and choice will be 1 if that alternative was selected, and 0 otherwise.
- Independent Variables (IVs): These are alternative-specific attributes that describe each smart watch alternative and vary across them within a choice set.
    - *color*: A categorical variable with levels 'Black', 'Lilac', and 'White'.
    - *brand*: A categorical variable with levels 'Apple', 'Samsung', and 'Fitbit'.
    - *price*: A numerical variable with four specified levels (NOK 990, 2390, 3190, and 6190).
    - *battery_life*: A numerical variable with four specified levels (10, 18, 40, and 80 hours).

2.3 Data Preparation and Evaluation Metrics

For multinomial logit modeling, data preparation begins by merging the alt_data (product specifications) with response_data (respondent choices and demographics). This crucial step links the attributes of each smart watch alternative to the corresponding choice tasks. Subsequently, the merged dataset, initially in a wide format, will be transformed into a long format. This transformation is fundamental, as it structures the data such that each row represents a single alternative within a choice set, essential for estimating the utility (preference score) of each alternative relative to a chosen baseline. Once prepared, multinomial logit models will be developed to quantify these utilities, explaining the probability of a respondent selecting a particular smart watch.

Model validation is crucial to ensure robustness and generalizability. This involves splitting the data into distinct training and testing sets to fit models and evaluate performance, thereby preventing overfitting. We will compare various model specifications, critically assessing whether numerical attributes like price and battery_life are best treated as linear or nominal variables based on their effects on preference. Model performance will be assessed using key metrics such as Hit Rate (Accuracy), AIC (Akaike Information Criterion) for balancing fit and complexity, and Log-likelihood for overall model fit. Through these comparisons, the most robust and parsimonious model, effectively explaining respondent preferences, will be selected.

2.4 Multinomial Logit Models

We developed four distinct multinomial logit models to predict customer smart watch choices and understand the impact of each attribute. In Model 1, all predictors were treated as nominal (factor) variables. Model 2 expanded upon Model 1 by including a squared_price variable, with price, squared_price, and battery_life all considered numeric; other variables remained nominal. Model 3 treated only price as a numeric variable while all others were nominal. Finally, Model 4 was similar to Model 2, but explicitly excluded the squared_price predictor, meaning price and battery_life were numeric, and others nominal. Each model was then used to make predictions on a testing dataset, with performance evaluated using hit rate, AIC, and Log-likelihood to identify the optimal model for fitting the full dataset and subsequently making recommendations for an ideal smartwatch design.

2.5 Model Result Comparison

Our initial evaluation of the four multinomial logit models, based on Hit Rate and AIC, revealed that Model 4 demonstrably outperformed Models 1, 2, and 3. Model 4 achieved a significantly higher hit rate (0.4853) and a lower AIC (9379.7429), indicating superior predictive accuracy and a better balance between model fit and complexity. Furthermore, an examination within Model 2, which included a squared_price term to test for non-linear price effects, showed that this term was non-significant. This suggests that the relationship between price and utility is predominantly linear, justifying a simpler linear treatment of price in subsequent models.

Model	Hit Rate	AIC	Log-Likelihood
Model 1	0.3449	9383.524	-4681.8
Model 2	0.3484	9380.28	-4683.1
Model 3	0.3484	9382.328	-4683.2
Model 4	0.4853	9379.743	-4680.9

To compare the models and select the most appropriate one, we employed the Likelihood Ratio Test, a statistical method used for comparing nested models. Two models are considered nested if one model is a more restricted version of the other, containing all its parameters plus at least one additional parameter. Our analysis identified the following nested relationships: Model 4 is nested within Model 3, and Model 3 is nested within Model 1.

The result of the likelihood ratio test between model 3 and model 1 shows that Chi-squared value is 2.8046 which is not significant at 5% level (p-value is 0.246), meaning that we cannot reject null hypothesis that the full model is not better than the restricted one. As such, we should use the more parsimonious model, which is the model 3.

```
Likelihood ratio test

Model 1: choice ~ as.factor(color) + as.factor(brand) + price + as.factor(batterylife) |
    0
Model 2: choice ~ as.factor(color) + as.factor(brand) + as.factor(price) +
    as.factor(batterylife) | 0
  #Df  LogLik Df  Chisq Pr(>Chisq)
1   8 -4683.2
2  10 -4681.8  2 2.8046      0.246
```

Next, we compare model 4 and model 3 using the likelihood ratio test. Similarly, we can observe that Chi-squared value is 1.4145 which is not significant at 5% level (p-value is 0.493), meaning that we can not reject the null hypothesis that the full model is not better than the restricted one. As such, we should use model 4 rather than model 3.

```
Likelihood ratio test

Model 1: choice ~ as.factor(color) + as.factor(brand) + price + batterylife |
    0
Model 2: choice ~ as.factor(color) + as.factor(brand) + price + as.factor(batterylife) |
    0
  #Df  LogLik Df  Chisq Pr(>Chisq)
1   6 -4683.9
2   8 -4683.2  2 1.4145      0.493
```

As a result, we will select model 4 to estimate the whole dataset. As a result of these sequential comparisons and the overall performance indicated by hit rate and AIC, Model 4 is selected to estimate the full dataset. In this chosen model, brand and color are treated as nominal (factor) effects on utility, allowing for distinct utility values for each level. Conversely, price and battery_life are treated as linear effects, implying that changes in these attributes have a consistent, additive impact on utility. This final model configuration offers the best balance of predictive accuracy and interpretability for understanding smart watch preferences.

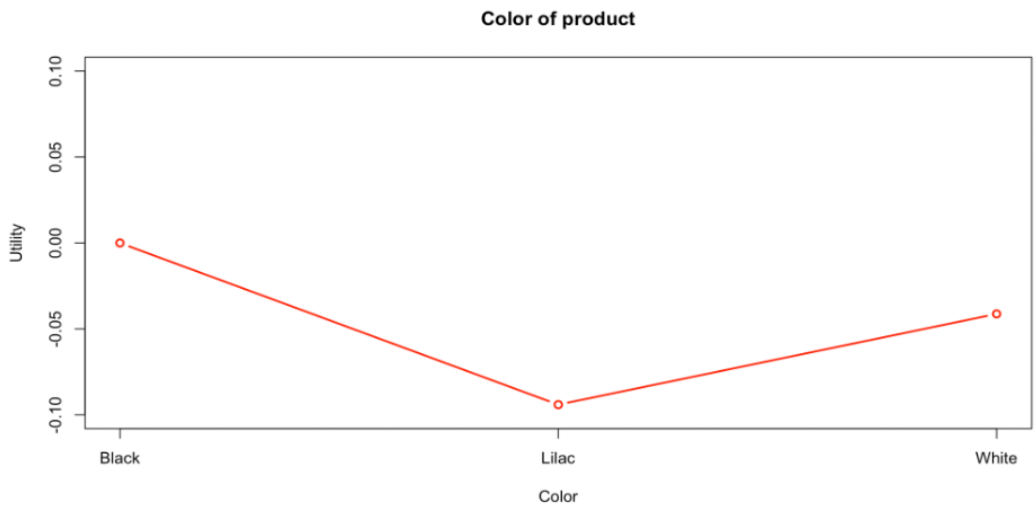
2.6 Model Interpretation

In this step, we fit our selected model 4 to the whole data set. However, before fitting the model, we need to convert the full dataset from wide format to long format. Once we have done this, we can fit the selected model to the entire dataset using the 'mlogit' function, and predict the choice probabilities using the predict function. We then assigned the predicted probabilities to each alternative and computed the predicted choice by selecting the alternative with the highest probability. After fitting the model, we can observe from the fit statistics that the hit rate is slightly higher when we fit the chosen model to the whole data set instead of testing the dataset. However, the AIC performs much worse. According to our model result, the predictors `price` and `(color)Lilac` have statistically significant p-value, meaning that we can use these factors to better predict our customers' choices.

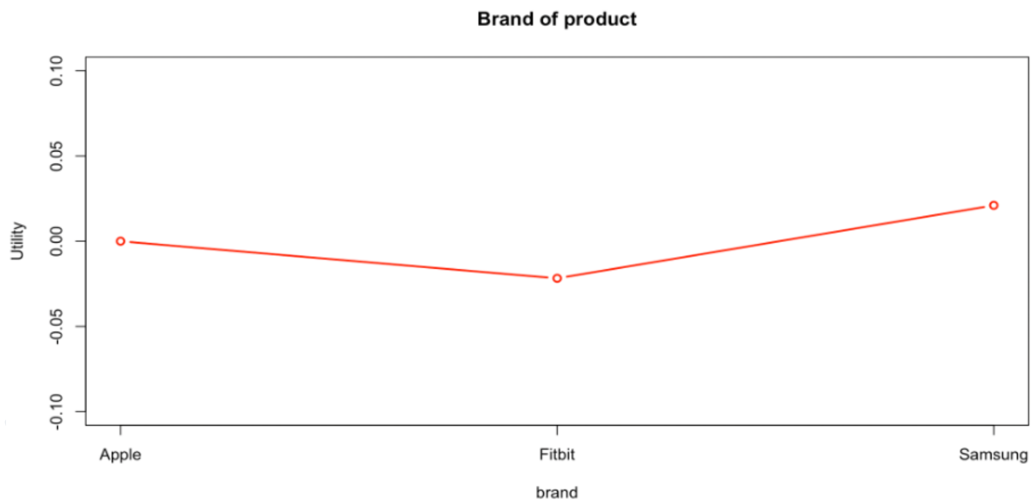
```
Call:
mlogit(formula = choice ~ as.factor(color) + as.factor(brand) +
    price + batterylife | 0, data = response_dat, method = "nr")

Frequencies of alternatives:choice
      1       2       3
0.23467 0.29400 0.47133

nr method
3 iterations, 0h:0m:0s
g'(-H)^-1g = 0.0165
successive function values within tolerance limits

Coefficients :
                           Estimate  Std. Error z-value
as.factor(color)Lilac   -9.2071e-02  3.9287e-02 -2.3436
as.factor(color)White   -4.0575e-02  3.8950e-02 -1.0417
as.factor(brand)Fitbit  -1.9593e-02  3.8337e-02 -0.5111
as.factor(brand)Samsung  2.2998e-02  3.8265e-02  0.6010
price                   -6.3148e-05  8.7261e-06 -7.2367
batterylife             -5.8380e-04  6.2246e-04 -0.9379
                          Pr(>|z|)
as.factor(color)Lilac       0.0191 *
as.factor(color)White       0.2975
as.factor(brand)Fitbit      0.6093
as.factor(brand)Samsung     0.5478
price                    4.599e-13 ***
batterylife                 0.3483
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log-Likelihood: -6560.4
```

In order to interpret the results of our model, we will plot the effect of different levels of attributes on utility and calculate the relative importance of all attributes. Based on these findings, we can give recommendations about the most and least important factor on average as well as what the ideal smartwatch should look like in order to boost sales.

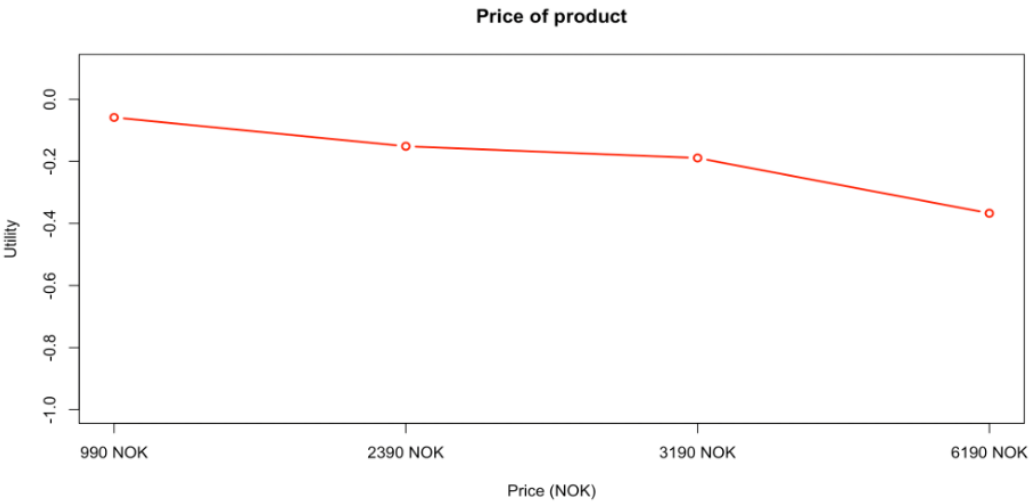
First, we plot the effect of each attribute level on utility. Based on the plot below (Color of product), we can observe that black color has the highest utility among color options, while white color has the lowest utility. This indicates that our products are likely to be sold more if they are black in color.



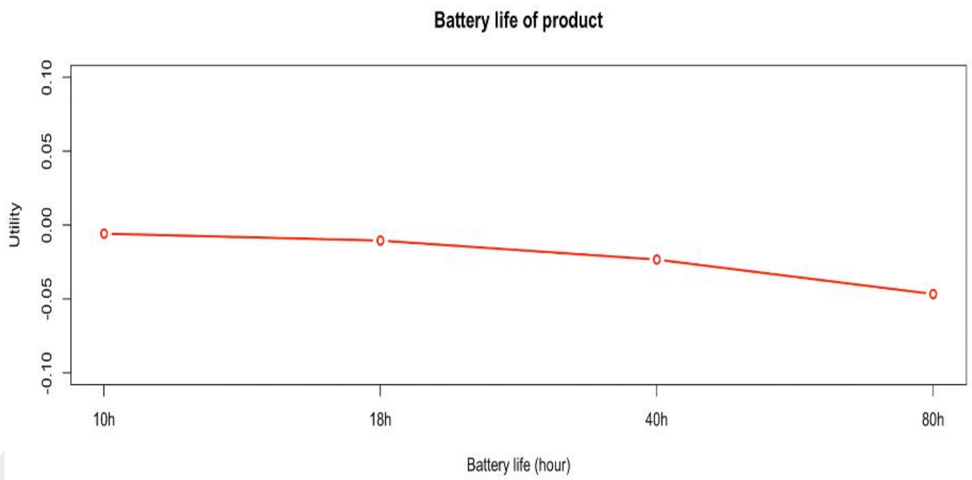
From the plot below (Brand of product), it appears that Samsung is the most preferred brand among options provided, followed by Apple when customers consider buying our products. On the other hand, Fitbit has lowest and negative utility among brands, indicating that it is the least preferred brand.



The plot below (Price of product) suggests that the utility score of lowest price point 990 NOK is the highest. As price increases, the utility of price decreases. This indicates that the customers tend to prefer products with lower prices. The effect of price points at 2390 NOK and 3190 NOK is almost the same.



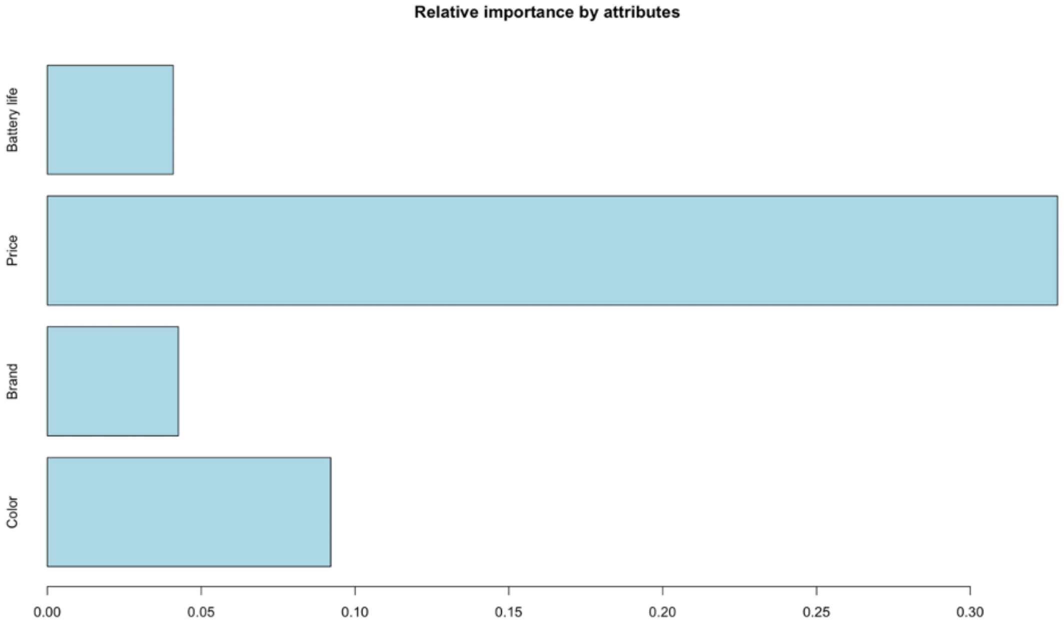
Based on the plot below (Battery life of product), we can see that the utility score of battery life of 10h is the highest, and that the utility score of battery life of 80h is the lowest. This indicates that customers prefer smart watches with battery life of 10h. However, the effects of battery life of 10h and 18h are similar to each other.



Next, we calculate the relative importance of attributes to decide which attribute we should focus the most to increase sales. After calculating the range of each attribute, we compute relative importance of attributes by following the formula: Relative importance = Range/Total ranges of all attributes. Based on the table below, we can plot the relative importance of all attributes.

Attribute <chr>	Range <dbl>	Relative_Importance <dbl>
color	0.09207150	18.271850
brand	0.04259100	8.452305
price	0.32836947	65.165853
battery_life	0.04086609	8.109991
Total	0.50389806	100.000000

5 rows



According to the relative importance plot (Relative importance by attributes), `price` is the most important factor in customers' choice of smart watches, followed by `color`. This outcome is not surprising as `price` is a statistically significant predictor in our model. `Brand` and `battery life` seem to have similar effects on customers' decision-making process.

III. Conclusions and Recommendations

The conjoint analysis of smart watch preferences reveals clear insights into customer valuation of product attributes. Price is overwhelmingly the most critical factor, accounting for over 65% of the overall decision-making importance. This signifies that variations in price have the most substantial impact on a customer's likelihood to choose a particular smart watch. Following price, Color holds moderate importance (approximately 18%), while Brand and Battery Life are considerably less influential, each contributing around 8% to the decision.

Regarding specific attribute levels: As expected, lower prices are strongly preferred, with utility consistently decreasing as price increases, making the 990 NOK price point the most preferred. For colors, customers show a slight preference for Black, with White being moderately less preferred, and Lilac being the least preferred. Among brands, Samsung emerges as slightly most preferred, followed by Apple, with Fitbit marginally less preferred; however, the low overall importance of Brand suggests it's not a primary driver. Finally, while generally less important, longer battery life is slightly preferred, though the utility difference across levels is small, aligning with its low overall importance.

Based on these findings, WearTech Innovations should adopt a price-centric strategy, heavily focusing on developing and marketing smart watches at aggressive price points, particularly around the 990 NOK range, to capture a wider market. For any higher-priced models, price increases must be clearly justified by significant value additions in other highly desired features. Concurrently, WearTech should optimize its color portfolio, ensuring strong availability of black models, which are most preferred, while considering reducing focus on lilac due to its low utility. While Samsung is slightly preferred, and longer battery life shows a minor positive trend, their low overall importance means WearTech should avoid over-investing in brand messaging or costly incremental battery life improvements. Instead, resources should be primarily allocated to competitive pricing and appealing color options (especially black), as these will yield the greatest impact on customer preference and market share.