Executive Summary

This research aims to analyze whether the more expensive product makes the customers more satisfied. In this research, cosmetics dataset is used which includes cosmetics with price, ranking, brands, and targeted skin type of each cosmetic. If the value of Ranking is low, it means that average satisfaction of customers with the product is high. Both hypothesis testing and linear regression is used to identify the relationship between price and satisfaction level.

In the hypothesis testing, cosmetic datasets were divided into two parts, high-price group and low-price group. Each cosmetic was classified into two groups using the median value as the criteria. Cosmetics were more expensive than the median value of price classified as high-price group. The null hypothesis was defined as "The customer satisfaction is equal for both group". To verify the null hypothesis, T-test was used because the population standard deviation was unknown. The test indicated that the p-value was lower than the significance level, rejecting the null hypothesis. Hypothesis testing shows that the difference of satisfaction between the high-price group and the low-price group is at a significant level.

Next, linear regression was conducted to analyze the confounders and moderators. Initially, only price and ranking was used for linear regression analysis. In other models, brand was used as a confounder variable. Brands were classified into two groups of luxury group and non-luxury group. Brand can cause influence the price because luxury brand aim Veblen effect, higher price causing higher demand, making price higher. Also, brand influence satisfaction by giving customers faith in the product regardless of the quality. Results of two models were compared to identify the effect of the confounder and difference of R-squared. However, adjusted R-squared value showed negative value showing no explanatory power. In the third model, the variable 'Normal' was used. 'Normal' means that the cosmetics is not aimed for specific dry or oily skin. If the cosmetics is not aimed at customer, it is expected for the price to have a larger effect on the ranking since there is more competitors in the market. The third model showed positive R-squared value and the coefficient of the product term showed was negative. This represents that the expected results was confirmed by linear regression results.

To conclude, hypothesis testing results show that high-price group and low-price group has significant difference in satisfaction. Linear regression results showed negative adjusted R-squared value without moderator, but model with moderator showed positive adjusted R-squared value. Also, model with moderator showed that the moderator 'Normal' variable amplified the effect caused to the ranking by price.