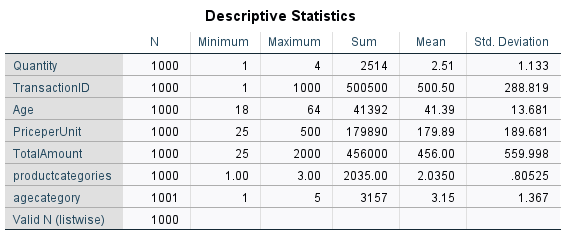
**Descriptive Statistics**

****

N: The data set had a total of 1000 transactions.

Age: Customer ages range from 18 to 64 years.

Price per Unit: Prices per unit range from $25 to $500.

Price per Unit: "The mean price per unit of products sold is $179.89."

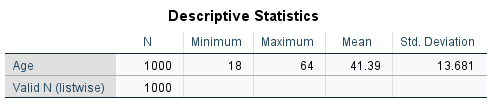
Total Amount: The total amount spent in transactions ranges from $25 to $2,000.

Quantity: A total of 2,514 units were sold in the dataset.

Total Amount: "The total revenue generated across all transactions was $456,000.

**Product Categories:** Products were categorized into 3 distinct categories, with a mean category value of 2.035.

1. **What is the average age of customers in the dataset?**



The average age of the customers in this data set is 41,.

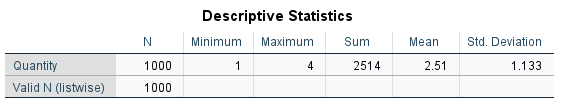
1. **What is the total revenue generated by each product category?**



Electronics lead in revenue generation with a total of 156,905, followed closely by clothing at 155,580. Beauty products rank last, contributing 143,515 to the overall revenue.

1. **What is the average quantity purchased per transaction?**

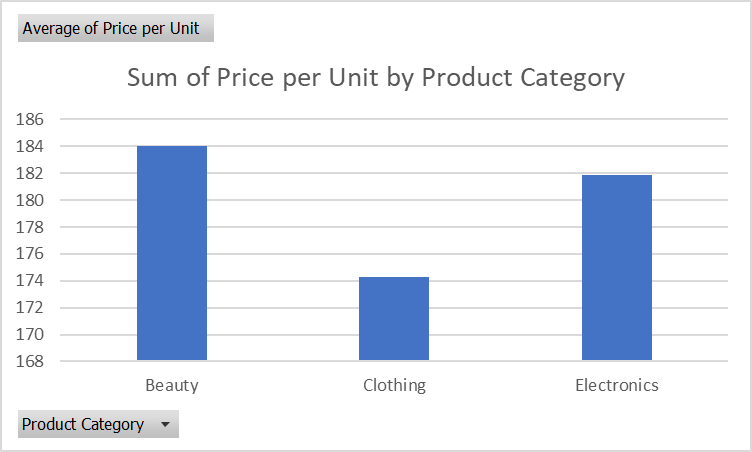






The average quantity of items purchased per transaction is 2.51 which can be rounded up to 3 items.

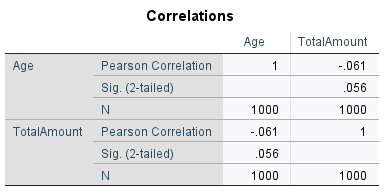
1. **What is the average price per unit across different product categories?**



Beauty products have the highest average price per unit (184.05) followed by electronics (181.90) and Clothing with the lowest average (174.29)

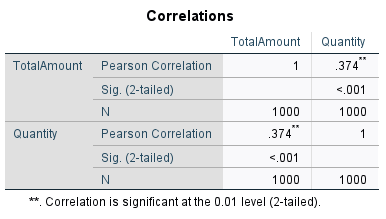
**Correlation Analysis**

1. **Is there a correlation between customer age and total amount spent?**



We obtain a Pearson correlation coefficient of -.061(6%), this implies that there is a weak negative correlation, as age increases, the total amount spent reduces and vice-versa. However, a significant value of 0.056 which is higher than 0.05, means that these results are not statistically significant.

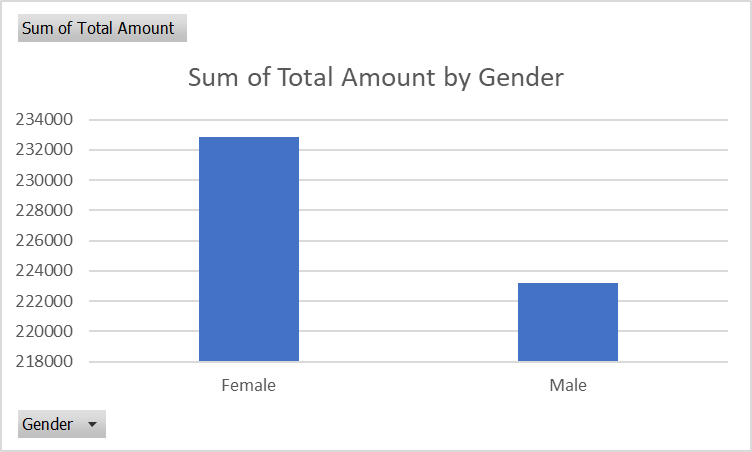
1. **Is there a relationship between the quantity purchased and the total amount spent?**



There is a positive correlation of .374, this implies a positive correlation which means that as quantity increases, the total amount increases. There is a significant value of less than 0.001 which means that these results are statistically significant.

**Gender and Purchase Patterns**

1. **What is the total amount spent by each gender?**

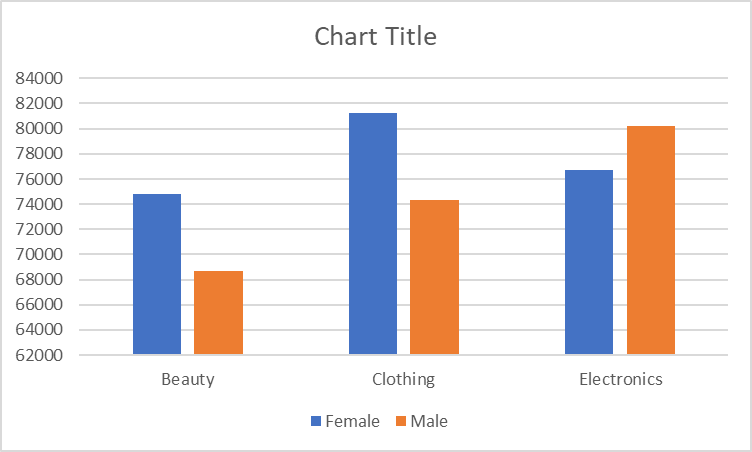


Females spend a total of 232840 while males spend a total of 223160. There is a small margin between the two genders in total amount.

1. **How does the average spending per transaction differ between genders?**

Females have a higher average spending amount of 456.54 while males have an average spending amount of 455.42.

1. **What is the most popular product category among different genders?**





From this, we can tell that the most popular product among females is clothing and that the most popular product among males are electronics.

**Age and Spending Analysis**

1. **How does the average spending per transaction vary across different age groups?**

Younger people have a tendency to spend more according to these results and the amount reduces as age increases.

1. **What is the total revenue generated by different age groups?**

**The age group 46-55 has the highest average of** 97235, followed closely by ages 25-35, the least amount of revenue is generated by ages 18-25.

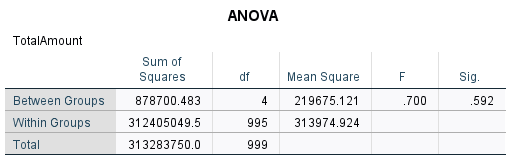
**Product Category Insights**

1. **Which product category generates the most revenue?**

Electronics generate the most revenue. This is closely followed by clothing and beauty generates the least revenue.

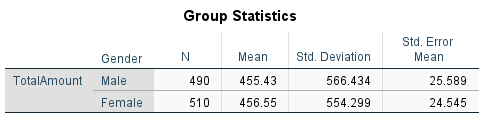
**T-Test or ANOVA**

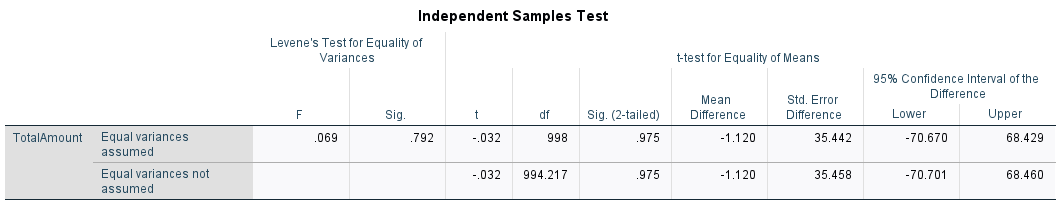
1. **Is there a significant difference in spending between different age groups?** (ANOVA)



Our p-value is 0.592, this is greater than 0.05 which means that the differences between the age groups are statistically insignificant and are likely to be due to random chance.

1. **Is there a significant difference in the average quantity purchased between genders?** (T-Test)





There is a significant value of .975, this means that we can assume equal variances. There is a significant value of .792 which means that the difference is not statistically significant.