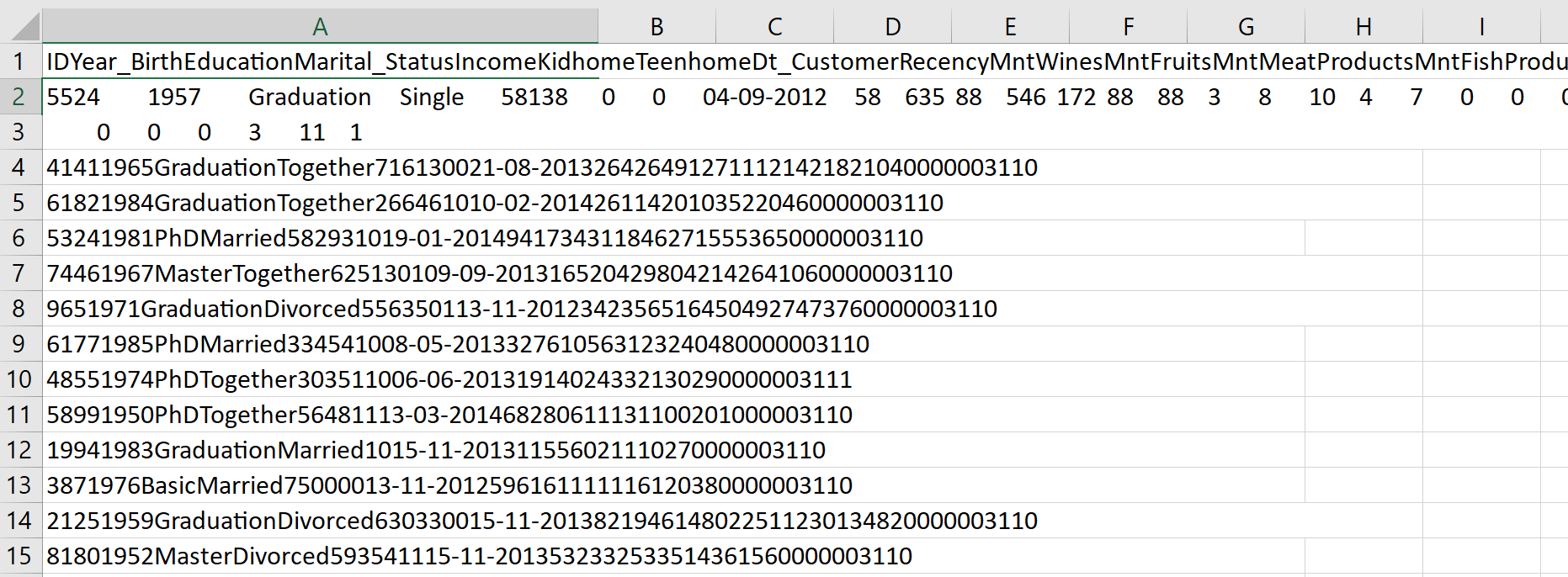
**Analysing Customers Personality for Business Improvement Using Machine Learning**

**SCREENS**

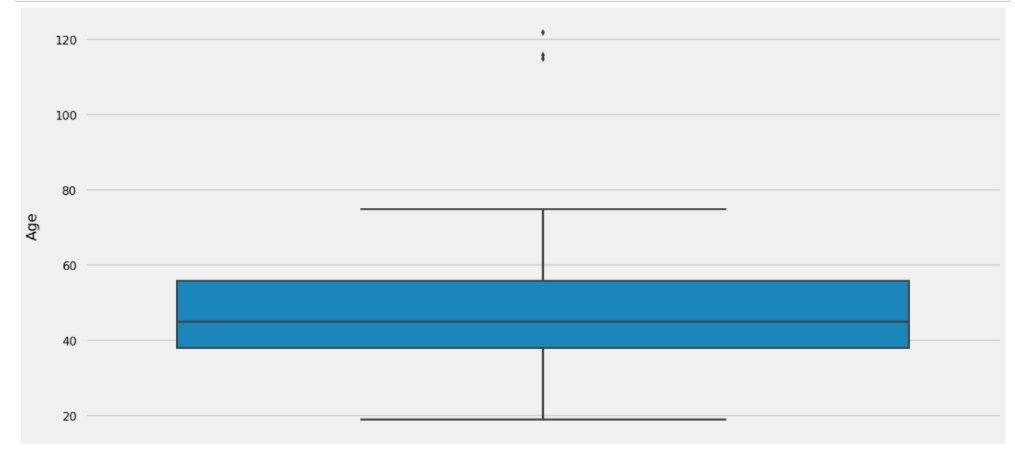
**DATASET**

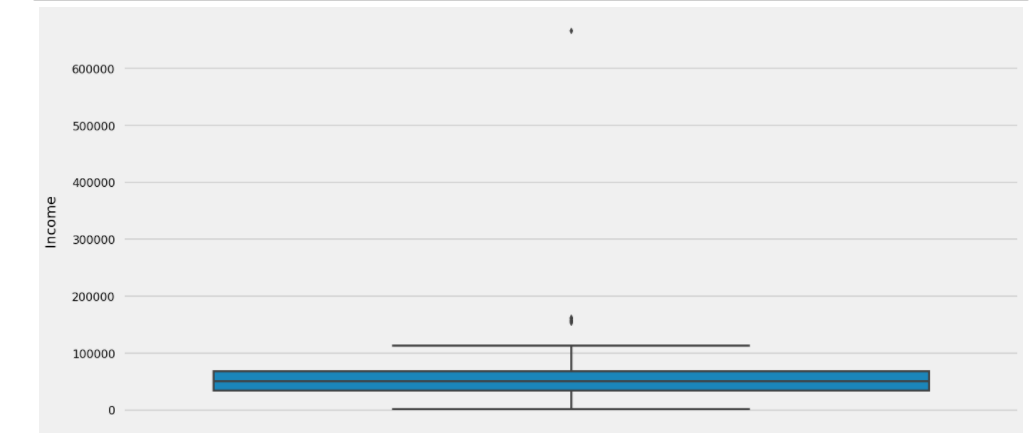
****

**FEATURES:**

ID Year Birth Education Marital Status Income Kid home Teen home Dt\_Customer Recency MN twines Mn Fruits MntMeatProducts MntFishProducts MntSweetProducts MntGoldProds NumDealsPurchases NumWebPurchases NumCatalogPurchases NumStorePurchases NumWebVisitsMonth AcceptedCmp3 AcceptedCmp4 AcceptedCmp5 AcceptedCmp1 AcceptedCmp2 Complain Z\_CostContact Z\_Revenue Response

**Removing Outliers:**

****

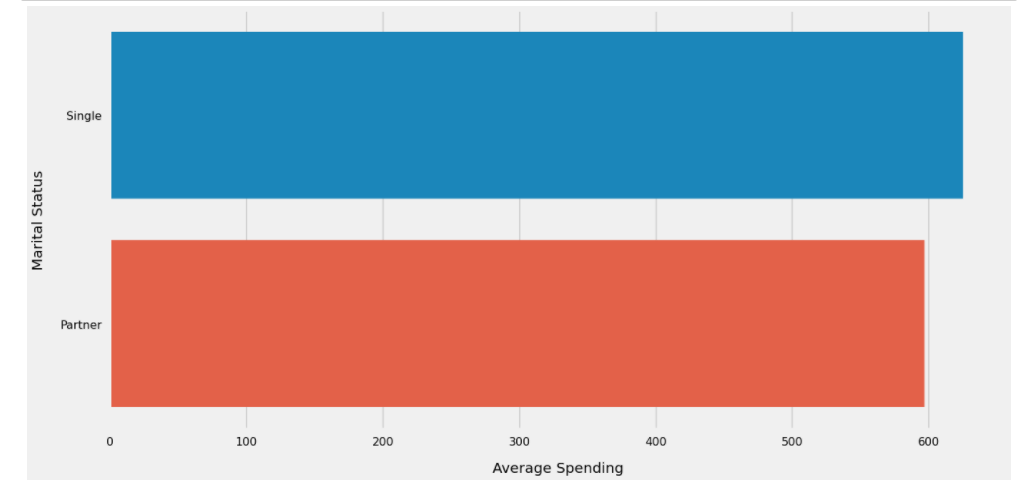
****

1. There are some customers aged above 100. This is unlikely to happen. Let's drop those customers from data

2. There are some customers who are earning more than 120,000 and some of them even more than 600,000. They are clearly the outliers in the data, so we will leave them out

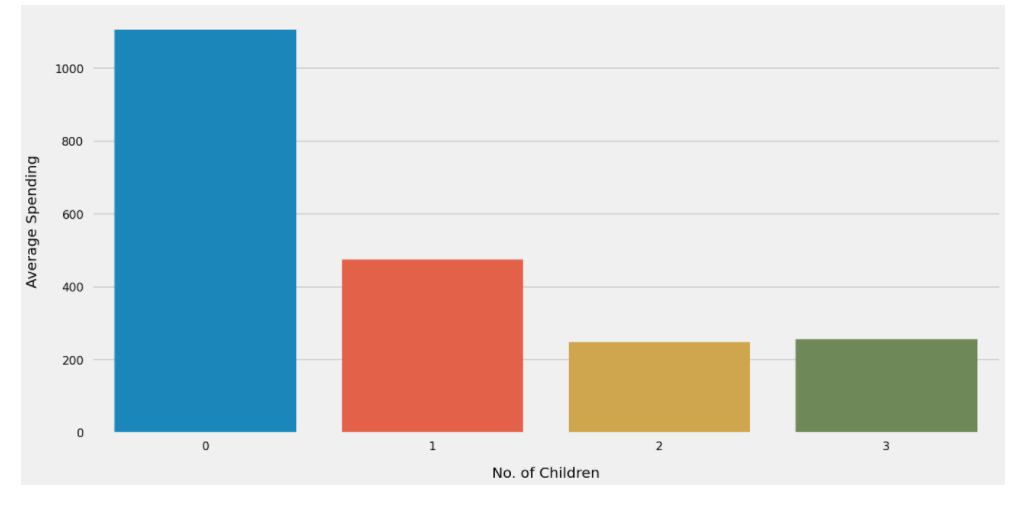
**DATA ANALYSIS:**

**Average spending:**

****

Despite being the minority, the Singles spent more money on the average as compared to the customers having partners

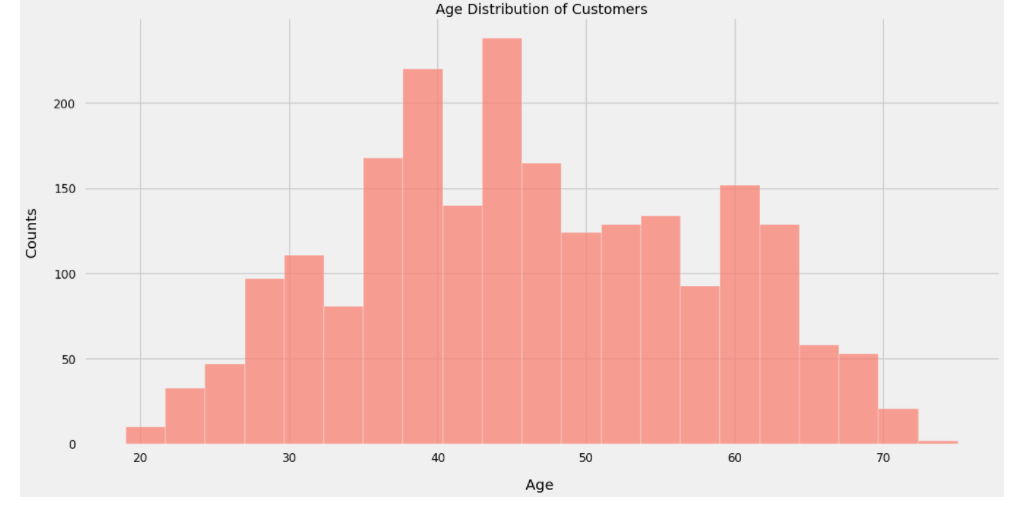
**Average Spendings:**

****

1. Customers who don't have any children at home spent higher than the customers having 1 children

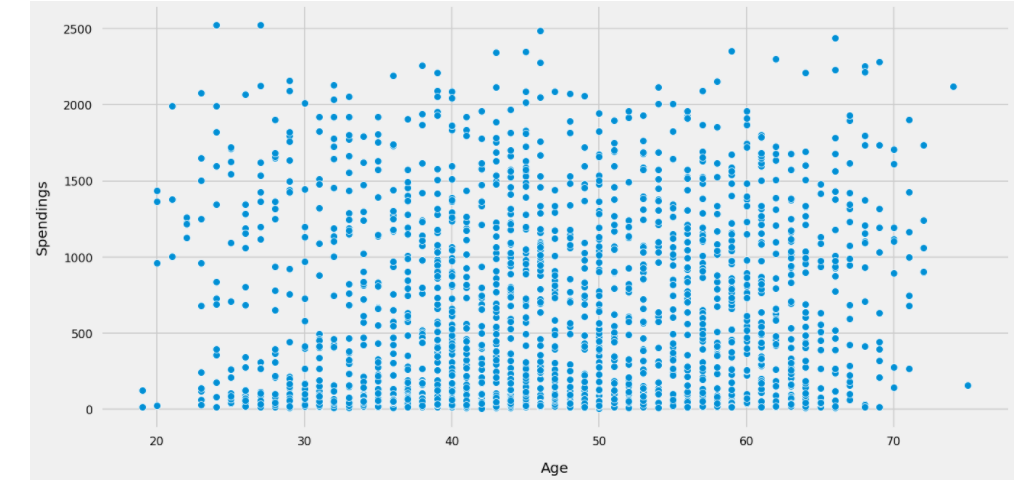
2. The customers having 1 children are spending higher than the customers havin 2 and 3 children

**Age Distribution of Customers:**

****

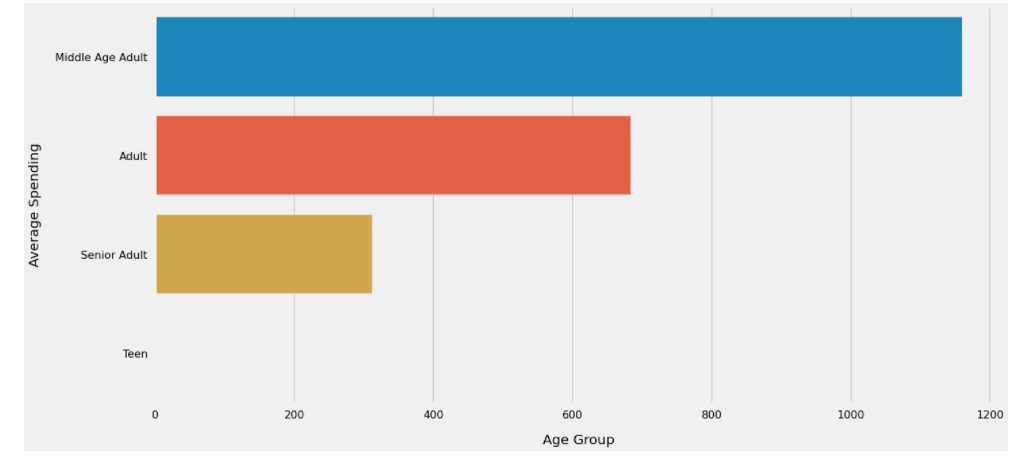
Age of the customers is nearly normally distributed, with most of the customers aged between 40 and 60.

**Relation ship age vs spending:**

****

There doesn't seem to be any clear relationship between age of customers and their spending habits

**Average Spendings:**

****

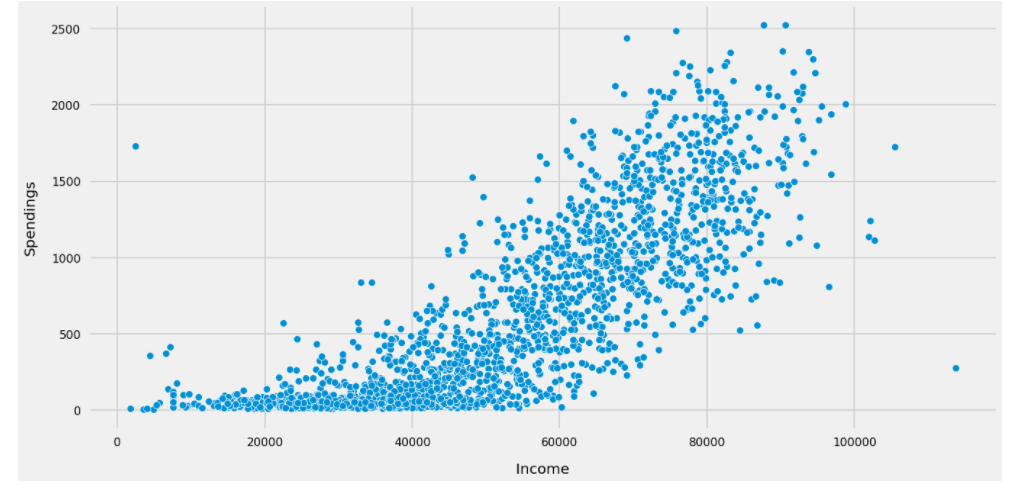
Middle age adults spent much more than the other age groups

**Income Distribution of Customers:**

****

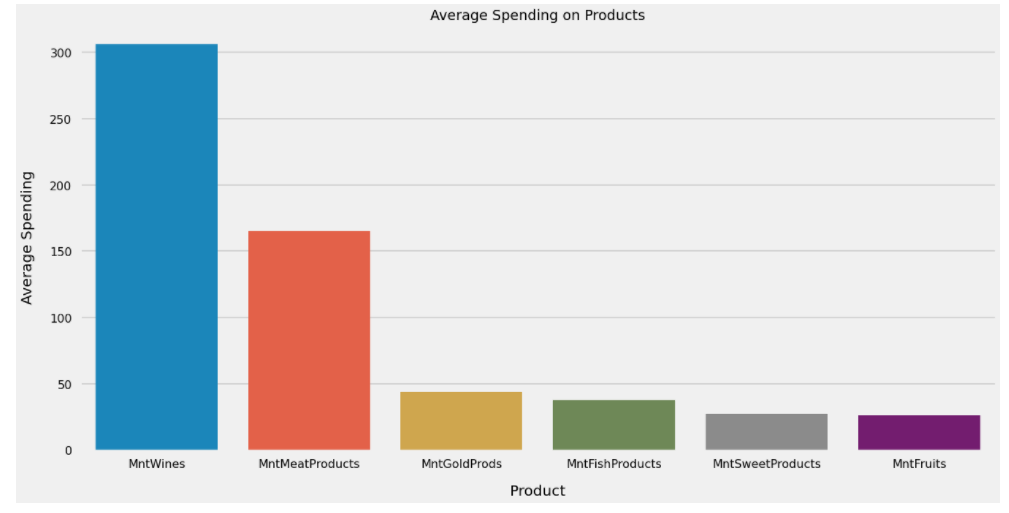
The salaries of the customers have normal distribution with most of the customers earning between 25000 and 85000

**Relationship: Income vs Spendings:**

****

The relationship is linear. Customers having higher salaries are spending more

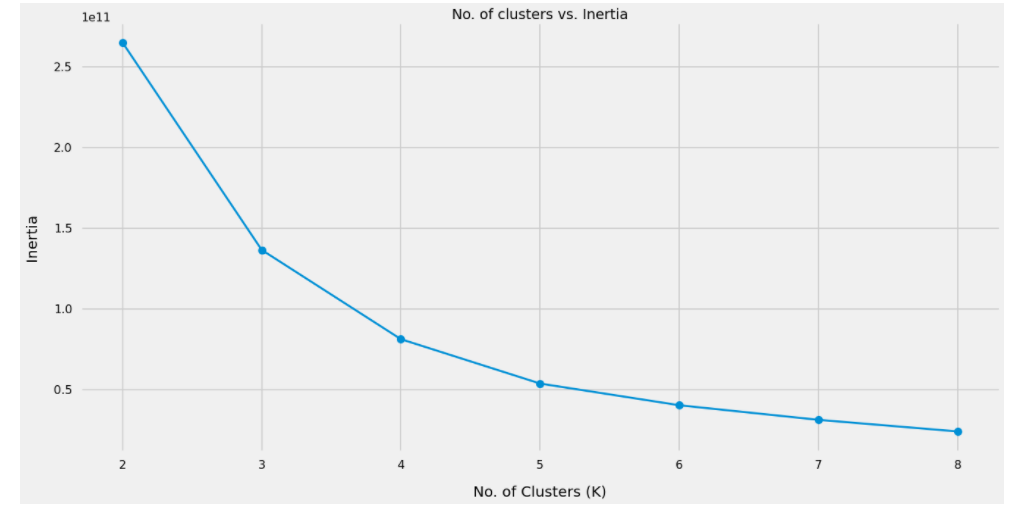
**Most Bought Products:**



1. Wine and Meats products are the most famous products among the customers

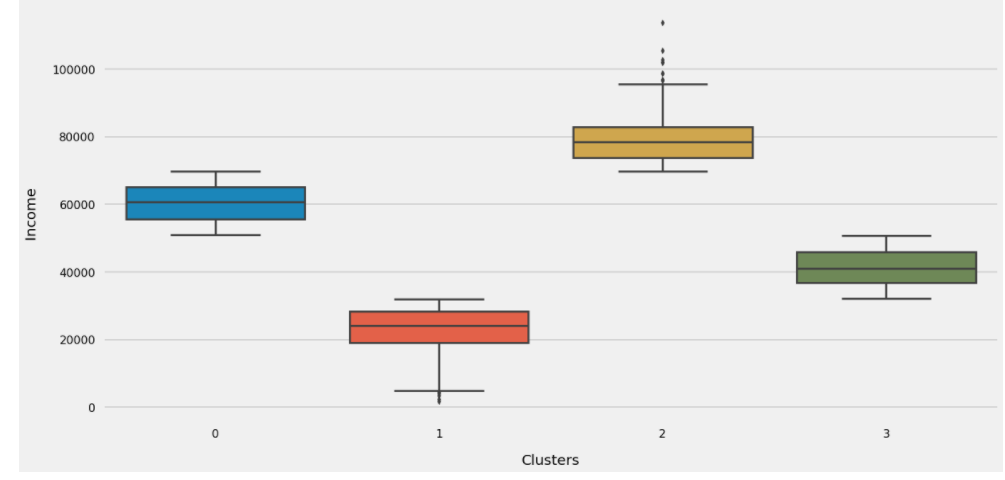
2. Sweets and Fruits are not being purchased often

**CLUSTURING METHODS:**

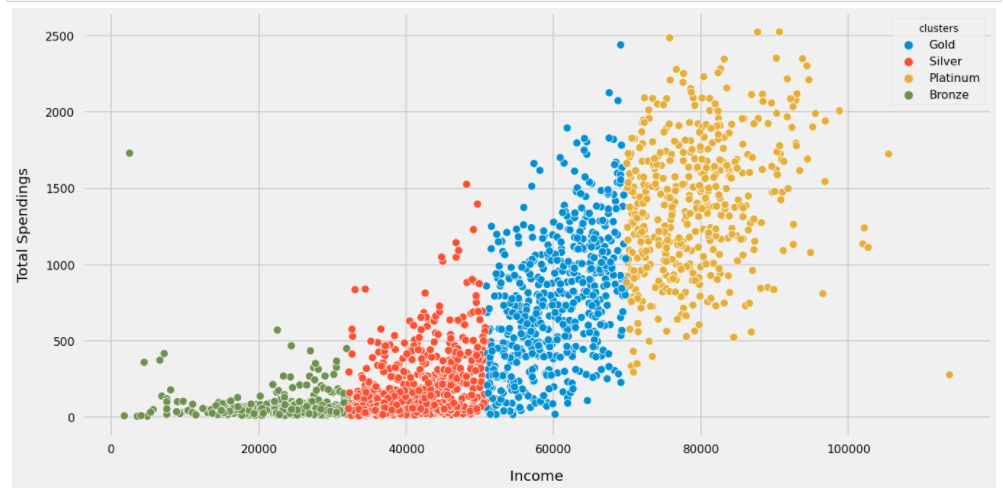


Based on the above plot we will segment the customers into 4 clusters, as the inertia value donot decrase much after 4 clusters

**Clusters Identification:**

****

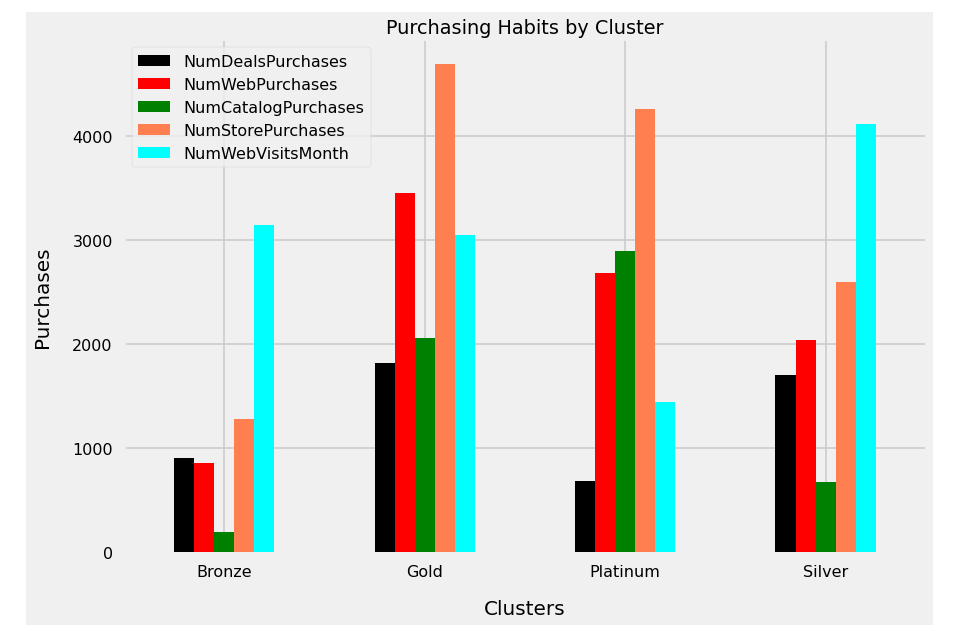
**INCOMMING VS SPENDING CLUSTES:**

****

1. The 4 clusters can easily be identified from the above plot

2. Those earning more are also spending more

**PURCHASE HABBITS BY CLUSTERS:**



1. Platinum and Gold Customers mostly likely to do store purchasing

2. Most of the web and catalog purchases are also done by the customers from Platinum and Gold segments

3. Silver and Gold categoriesnalso like to buy from the stores

4. Deal purchases are common among the Gold and Silver customers

5. Silver category customers made the most number of web visits while customers from Platinum segment have least web visits

**CONCLUSION:**

