**Question 1:**

The data contains three key performance indicators at all retail company outlets. The dataset has three variables.

**Client Satisfaction score** named as “client\_value\_1”.

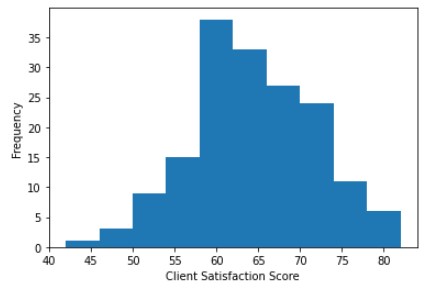
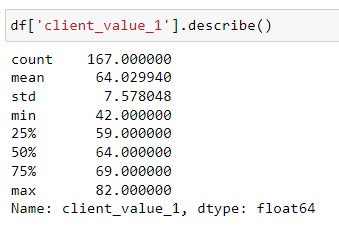
**Location** – customer rating named as “location\_score\_1”.

**Engagement** – summary variable derived from monthly visits to store named “engagement\_index\_1”.

The shape of the dataset is 167 rows and 4 columns.

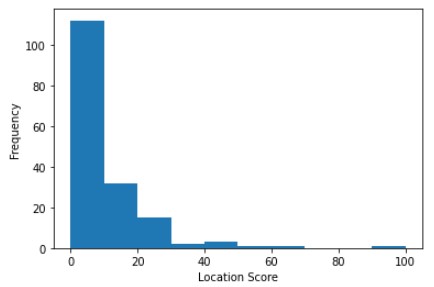
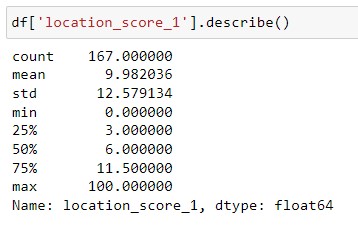
**Client Satisfaction Score**:

* The client satisfaction score measures customers' satisfaction with the retail company's products and services.
* The Client Satisfaction score ranges from 42 to 82, with higher scores indicating greater satisfaction.
* The distribution of scores appears to be normally distributed, with an average score of 64, a standard deviation of 7.5 and a median score of 64.
* This suggests that the majority of customers are somewhat satisfied with the company's offerings, with a smaller number of highly satisfied and dissatisfied customers.



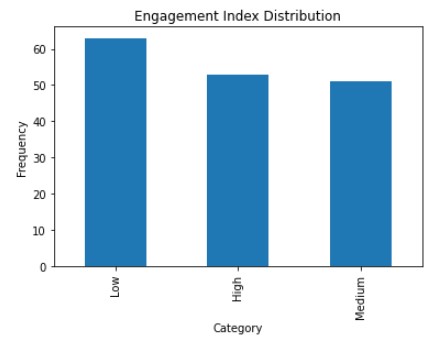
**Location- Customer Rating**:

* The location-customer rating variable measures how customers rate the different outlets of the retail company.
* The Location score ranges from 0 to 100, with higher scores indicating ratings.
* The distribution of the scores is positively skewed, with a mean score of 9.98, a standard deviation of 12.5, and a median score of 6.
* The data suggests that the retail company may want to focus on improving the location scores of their outlets to increase overall revenue.



**Engagement**:

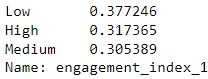
* The Engagement score is divided into three categories: Low, Medium and High.
* The frequencies of Low, Medium and High engagements are 37.7%, 30.5% and 31.7%.
* The data depicts that the majority of the outlets in the dataset have a low engagement index.
* This suggests there is scope to improve the level of engagement between the retail company and its clients.



**Engagement score**

**normalize in each**

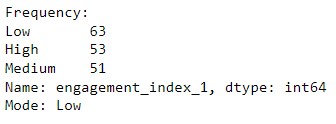
**category**



**Distribution**

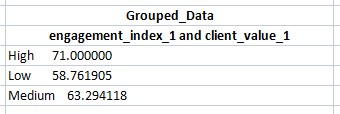
**of Engagement**

**score**

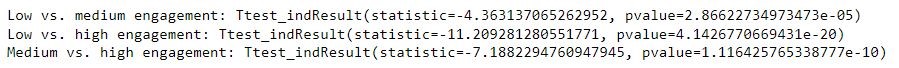


**Question2:**

* ANOVA and t-tests were performed on Client satisfaction scores between the different levels of store engagement to ascertain whether there was a notable difference between the means of groups.
* T-Test: Groups are divided based on their engagement levels, specifically low, medium, and high. The null hypothesis in each test is that there is no significant difference in satisfaction scores between the two groups being compared.
* ANOVA: Determine whether there are significant differences between the means of the three groups.

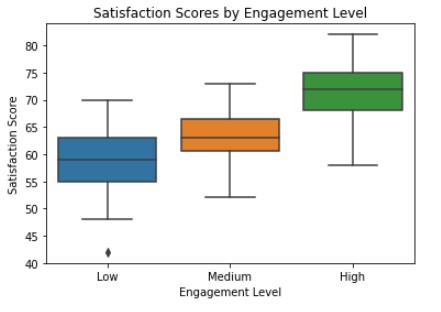


**ANOVA-Test T-Test**



* In both tests, the p-value for the above is less than the significance level of 0.05.
* We conclude by rejecting the null hypothesis that satisfaction levels are significantly different across the different levels of store engagement.
* Higher levels of engagement are associated with higher levels of satisfaction, which is consistent with the idea that engagement positively affects satisfaction.

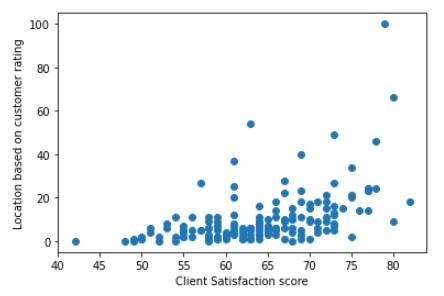
The following visualisation shows the differences in satisfaction scores between the different levels of store engagement.



**Question 3:**



* According to the bar chart, there is a strong relationship between Engagement and Client Satisfaction.
* Highly engaged outlets have a higher proposition of high client satisfaction scores, while lowly engaged outlets have a higher proposition of low client satisfaction scores.



* By scatter plot, a weak positive correlation was observed between Client Satisfaction score and Location based on customer rating.
* The Client satisfaction values are distributed evenly across the range, while the location values are mainly concentrated around 0-30, and very few values are above 30.
* This could suggest that the majority of the stores have a lower customer rating for their location, which could potentially impact their overall KPI scores.