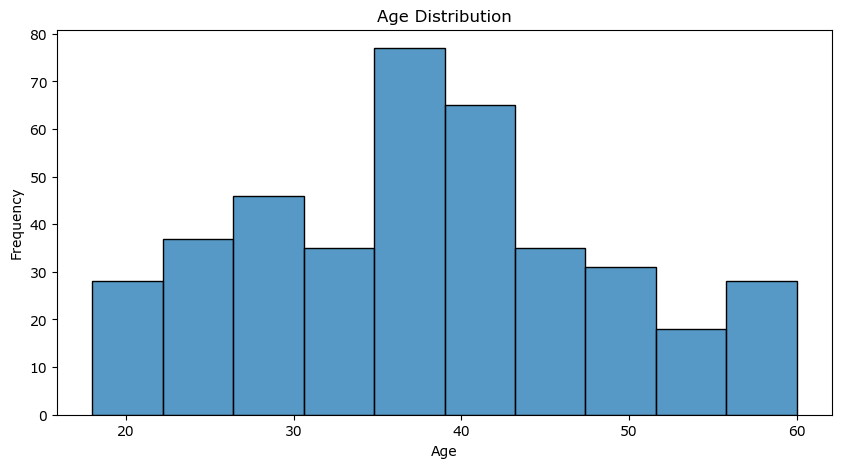
SOCIAL NETWORK DATASET REPORT

The Social Network Dataset gives information about the relationship that exists between the variables and how it affects each other. The variables includes the Age, Estimated Salary and Purchased. This dataset contains 400 rows and 3 columns.

Below is the first 10 rows and columns of the dataset.

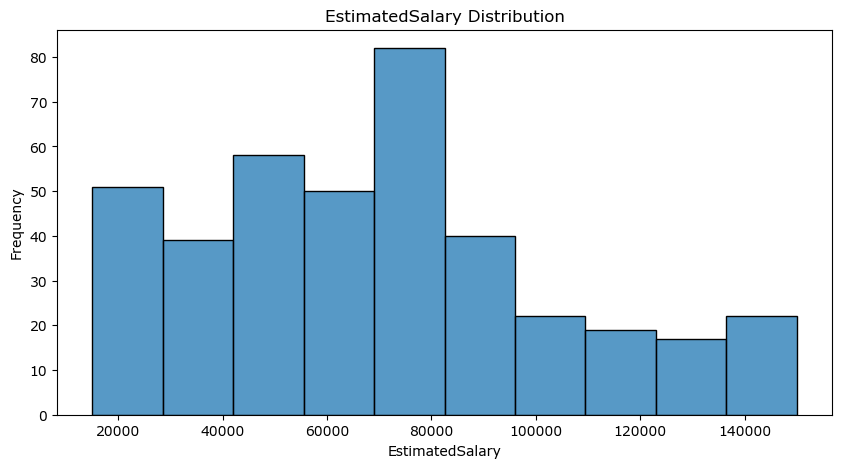
|  | **Age** | **EstimatedSalary** | **Purchased** |
| --- | --- | --- | --- |
| **0** | 19 | 19000 | 0 |
| **1** | 35 | 20000 | 0 |
| **2** | 26 | 43000 | 0 |
| **3** | 27 | 57000 | 0 |
| **4** | 19 | 76000 | 0 |
| **5** | 27 | 58000 | 0 |
| **6** | 27 | 84000 | 0 |
| **7** | 32 | 150000 | 1 |
| **8** | 25 | 33000 | 0 |
| **9** | 35 | 65000 | 0 |

AGE DISTRIBUTION



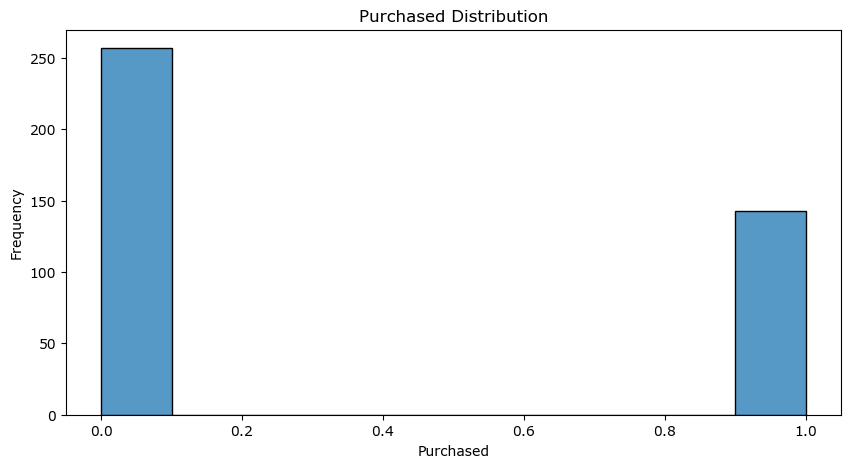
The Age distribution shows that the maximum age is from 35 - 39 and the minimum age range is between 51 - 56

ESTIMATEDSALARY DISTRIBUTION



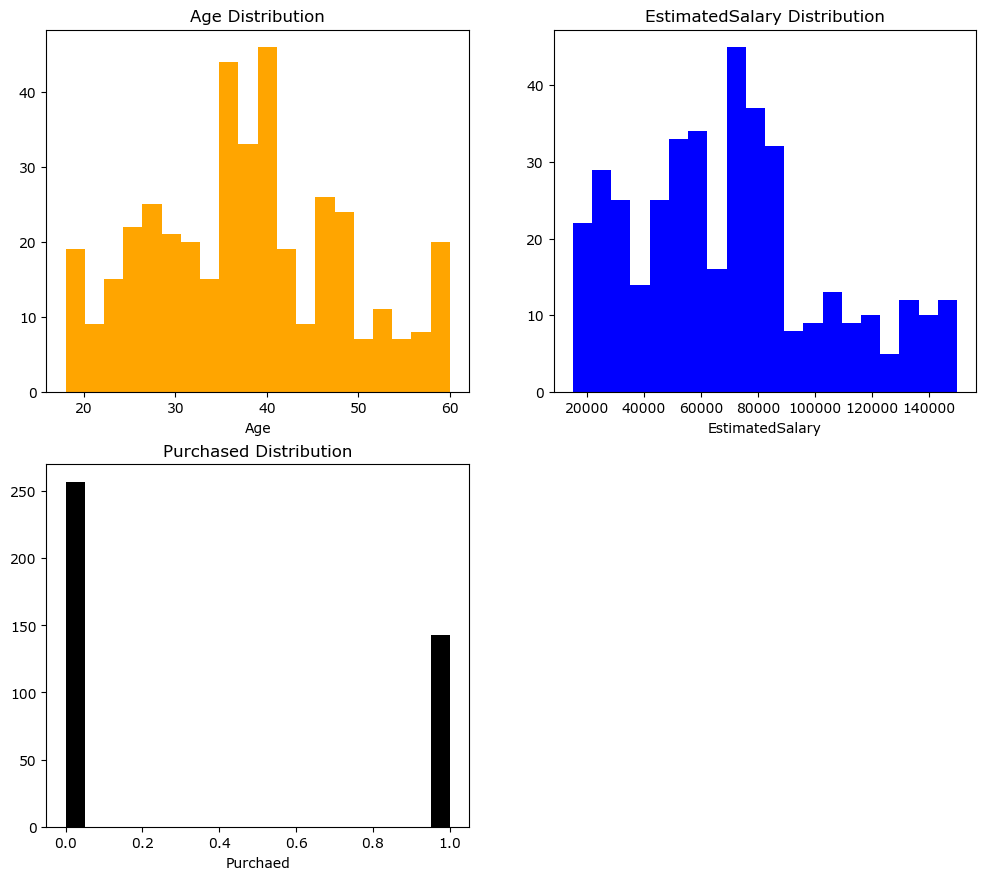
The maximum Estimated salary in this distribution is between 70000 - 81000 and the minimum estimated salary is between 121000 - 139000.

PURCHASED DISTRIBUTION



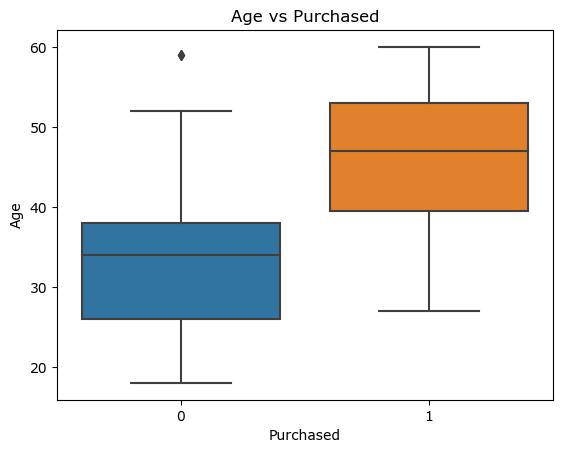
The maximum Purchase in this distribution falls between 0.0 - 0.1 and the minimum falls between 9.0 - 1.0.

UNIVARIATE DATA EXPLORATION



It visualizes all the univariate variables and show the maximum and minimum age and estimated salary of people that purchased and those that did not purchase.

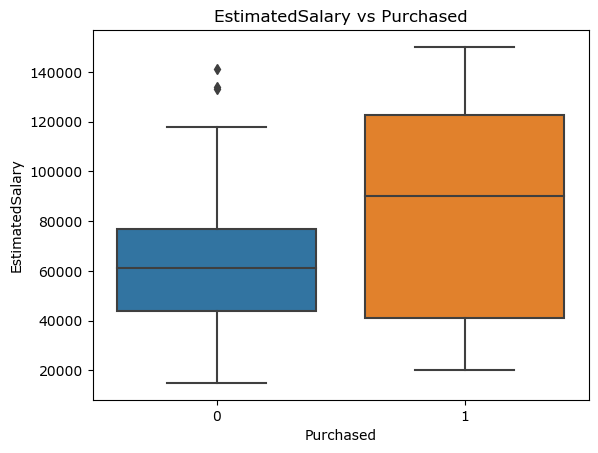
BOX PLOT OF AGE VS PURCHASED



The box plot is used to compare two variables.

Here, it showing the relationship between Age and Purchased. People that purchased are between the age of 40 - 52 and the people that did not purchase falls between the age of 26 - 39. And also we have one outlier.

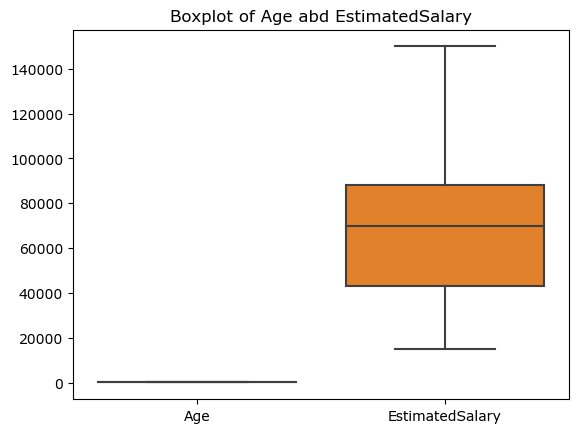
BOX PLOT OF ESTIMATEDSALARY VS PURCHASED



This plot shows the relationship between Estimated salary and Purchased.

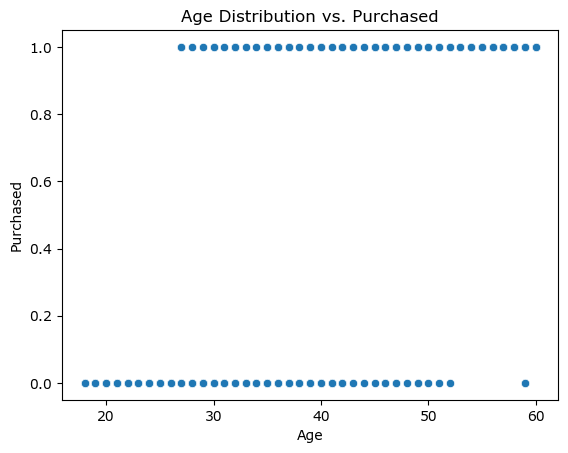
Here, the people that ends 44000 - 78000 salary did not purchased and those that ends 40000 - 124000 purchased . And we have two outliers.

BOX PLOT OF AGE AND ESTIMATEDSALARY



This box plot explains the the relationship between age and estimated salary.

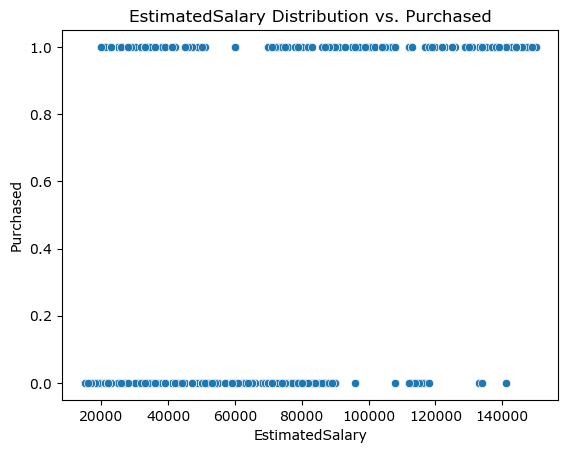
SCATTERED PLOT OF AGE DISTRIBUTION VS PURCHASED



Scattered plot is also used to compare two variables.

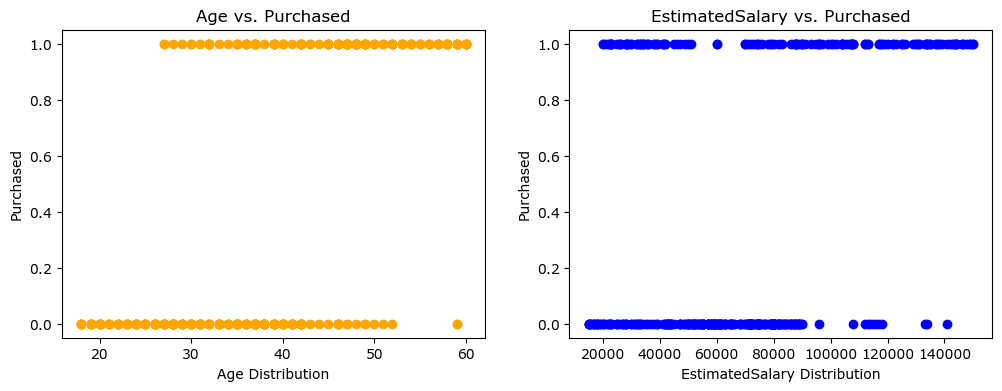
Here it shows the relationship between Age distribution and purchased . The people within the age of 10 - 52 did not purchase, then we had an outlier. The people within the age of 28 - 60 purchased.

SCATTERED PLOT OF ESTIMATEDSALARY VS PURCHASED



User this, we have that the estimated salary of people that did not purchase started from 10000 - 92000 then break and started again from 98000 stops, continues from 109000 - 120000 and then outliers. And the estimated salary for those that purchased started from 20000 - 50000 then break and started again from 60000 then break and continue from 70000 - 158000.

SCATTER PLOT OF AGE VS PURCHASED and ESTIMATEDSALARY VS PURCHASED



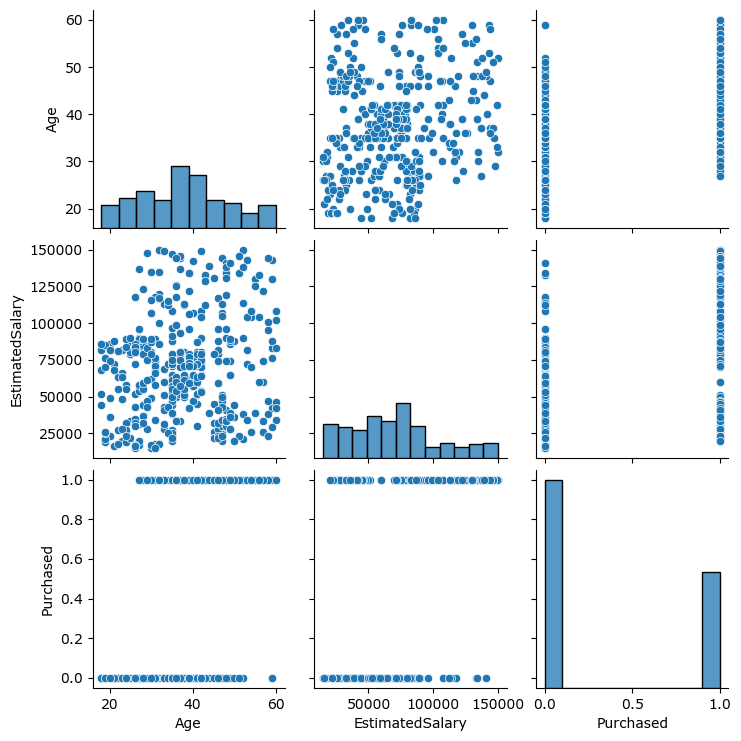
Grouping of bivariate variables

HEATMAP



The heatmap is a representation of a multivariate variables. This shows the relationship between Age, Purchased and Estimated salary. From the map it shows that Age has a higher correlation with 62% compared to Estimated salary and Purchased.

MULTIVARIATE SCATTERED PLOT



This is a multivariate scattered plot showing all the variables im the Social Network dataset.

SUMMARY

Conclusively, the dataset shows that the estimated salary of people that purchased was higher that those that did not purchase. Also in terms of age, the people that purchased were older and more that those that did not purchase.