Mihir Srivastava

[mihir.12107590@lpu.in](mailto:mihir.12107590@lpu.in)

[mihir.srivastava001@gmail.com](mailto:mihir.srivastava001@gmail.com)

**MINI PROJECT: PERFORMING EDA ON A DATASET**

**Q1.**

This dataset is detailed analysis about ideal customers.

By going through the given dataset, following things can be concluded:

* It helps in understanding customers in better way,
* It also helps in easily modification of products according to needs and behaviours,
* It can also help in upgradation of business of the company.

**Q2.**

Income column has highest number of missing values i.e., 24.

Percentage of missing value = 1.071429%

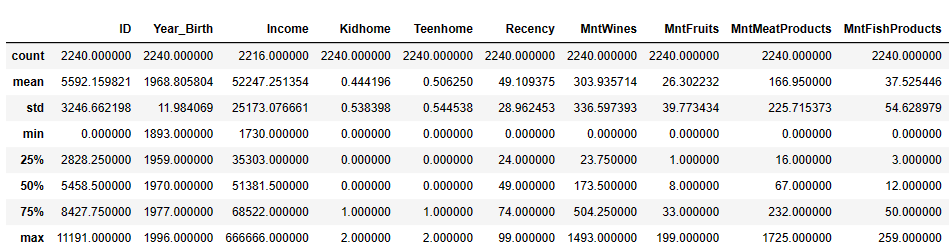
**Q3.**

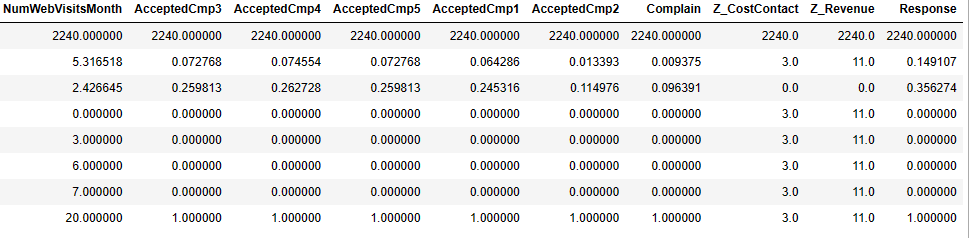
No. of columns not having more than 1 value are 2, i.e.,

* Z\_CostContact
* Z\_Revenue

**Q4.**

Statistical data of numerical values in dataset are as follows:

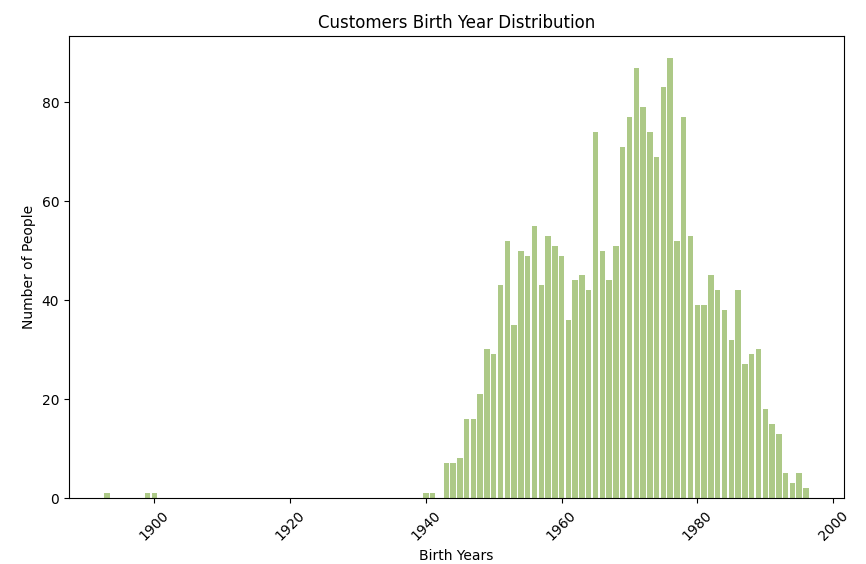


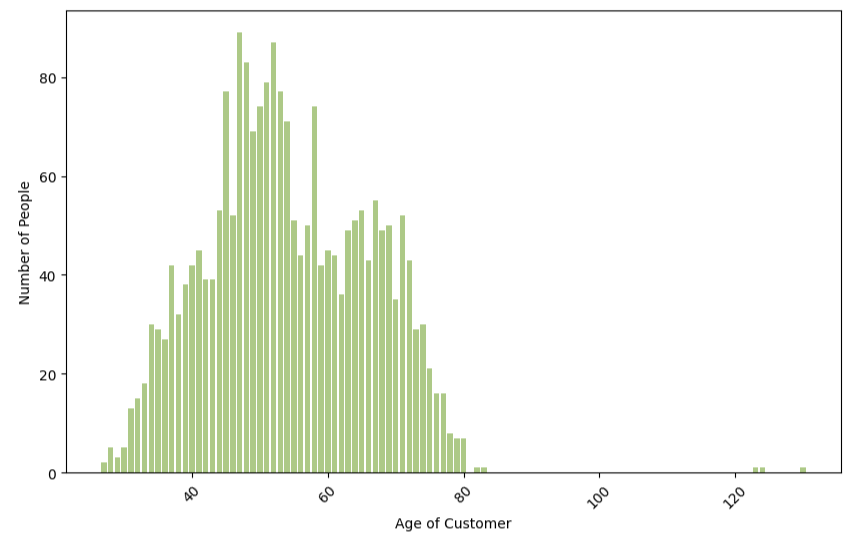


Unique columns in given data set are as follows:

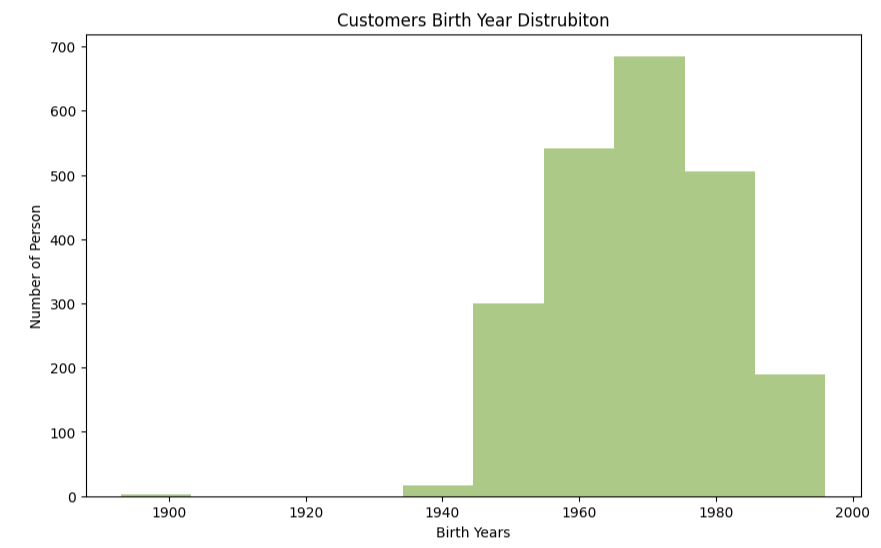
* ID
* Income
* ID\_NumWebVisitsMonth
* ID\_Year\_Birth
* ID\_Income

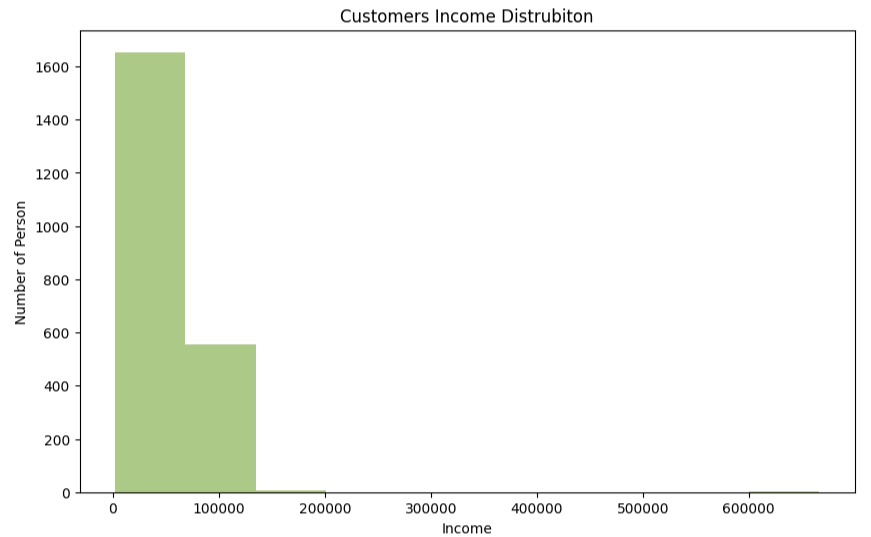
**Q5.**

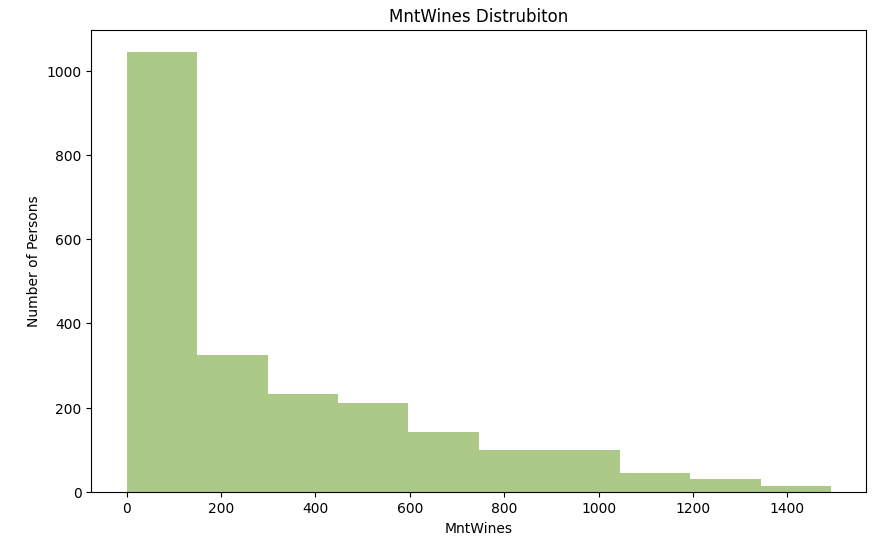


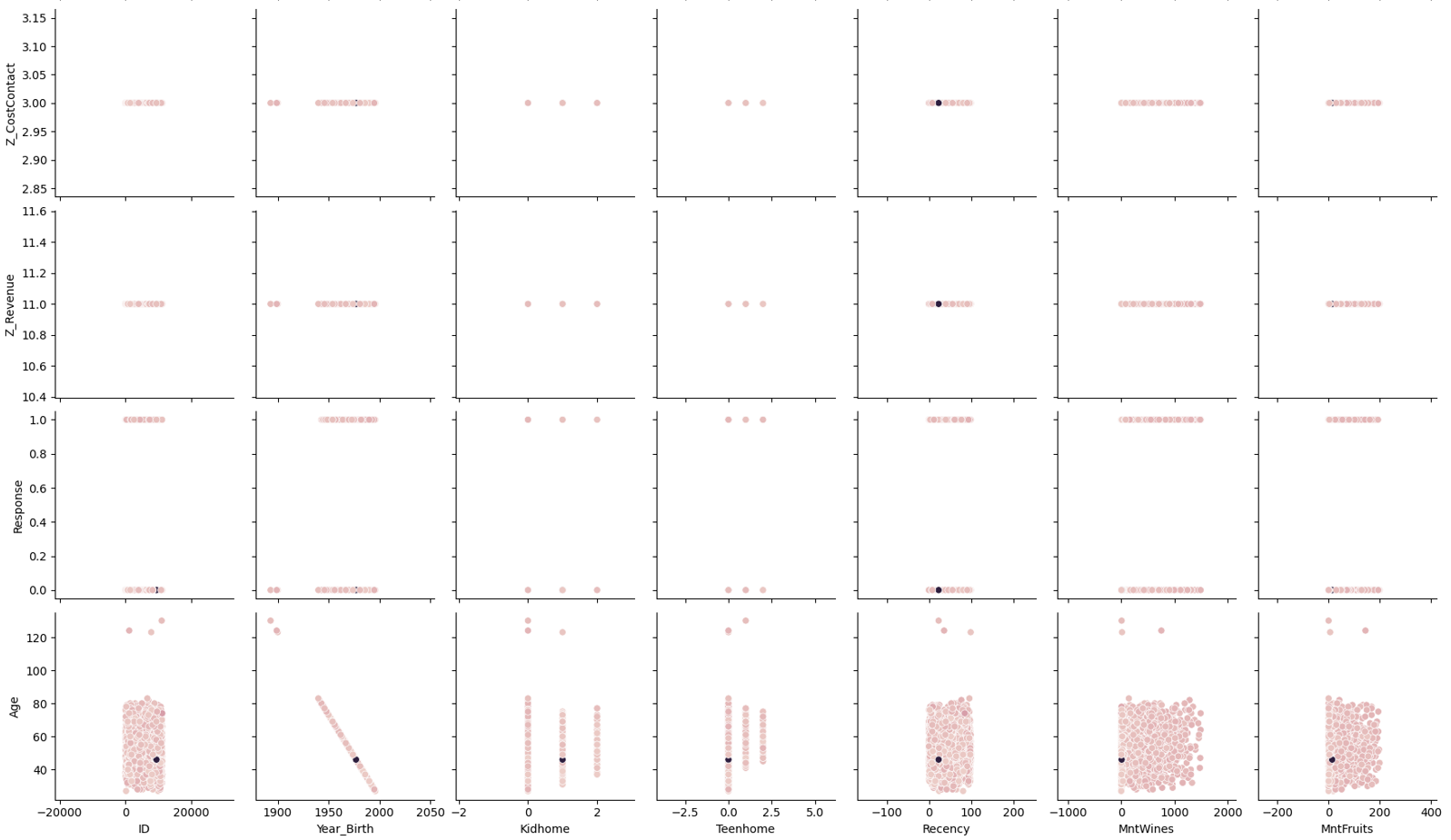


**Q6.**

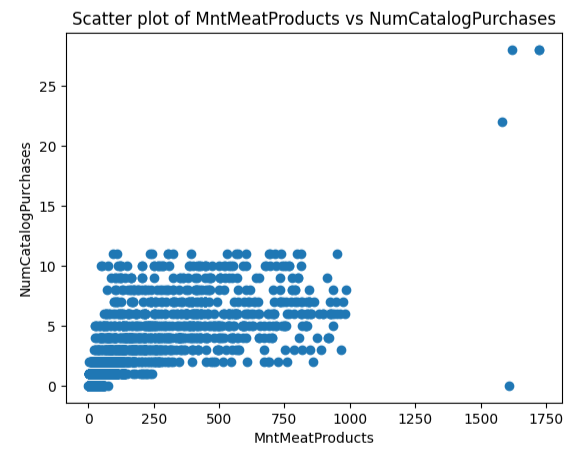


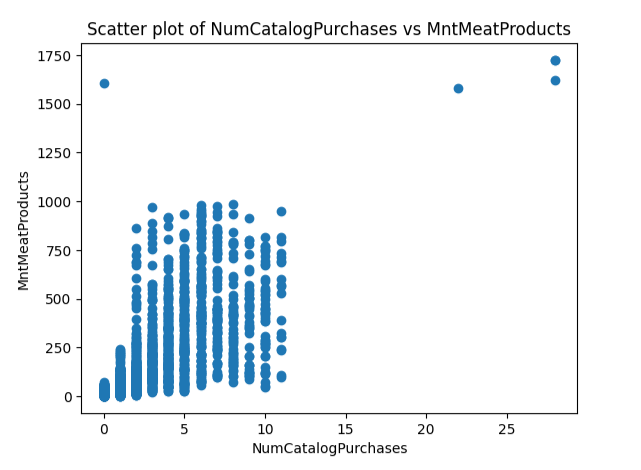


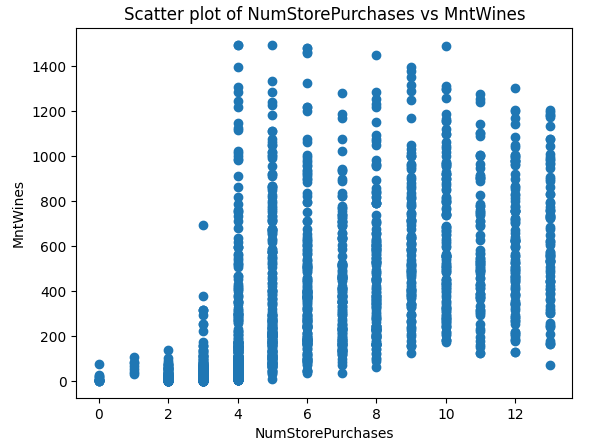


**Q7.**

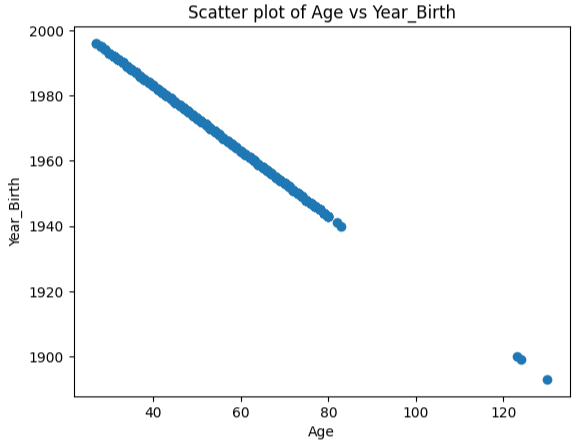
**Q8.**

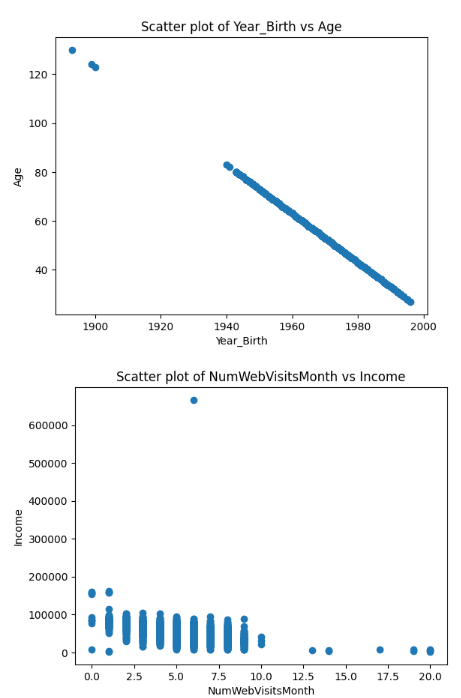
****

****

****

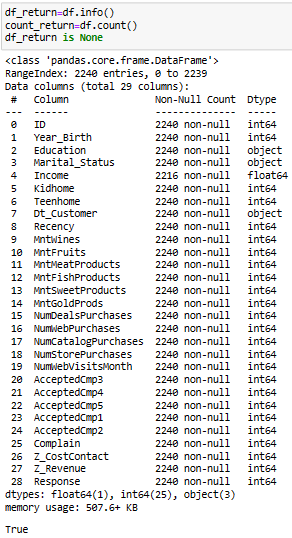
**Q9.**

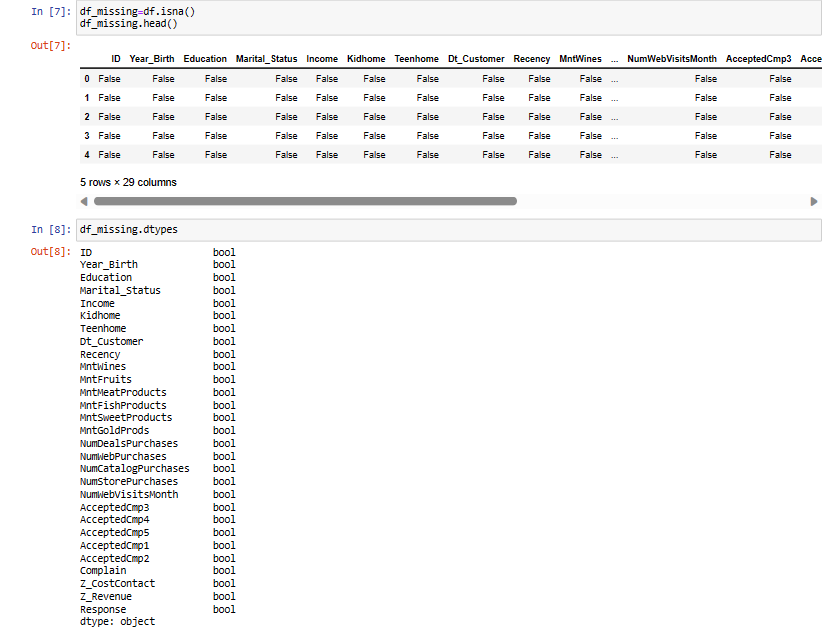
****

****

**Q10.**

* Knowing about the datatypes of the given dataset



* Converting data into Boolean form for checking missing values