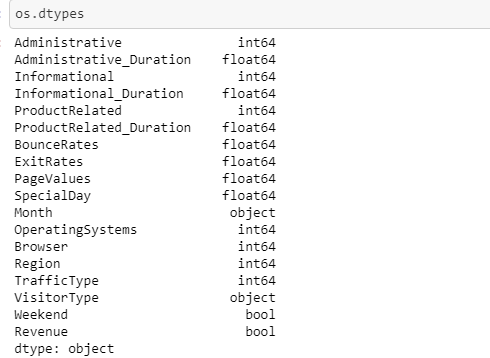
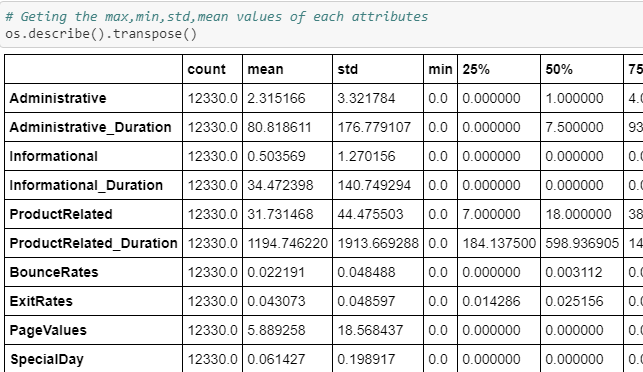
We identified that the parameters provided were also of “object”type so we could not use them for our analysis , we had to convert the parameters to category type by encoding.

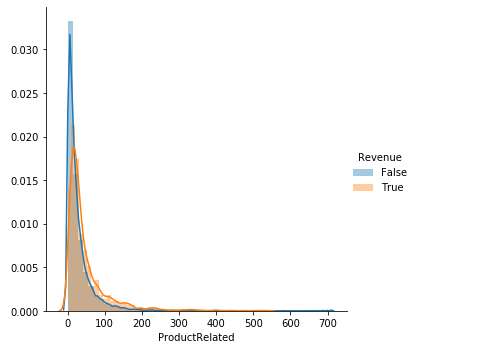


Analyzing the statistics of each individual parameters helped us understand the distribution of the data values.

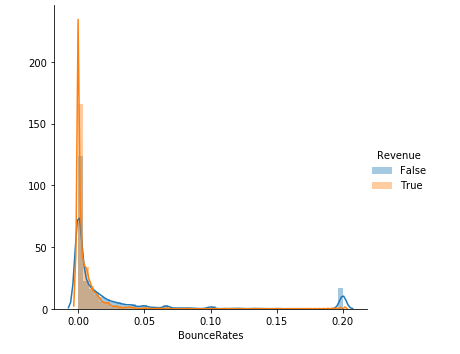


We tried Univariant Analysis by taking each individual parameter against the target. i.e. Revenue.

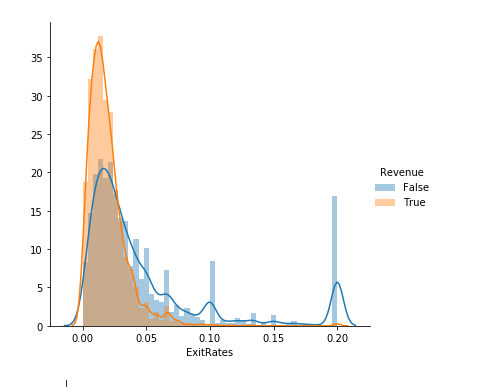
This helped us identify the variance and affect of individual parameter on revenue.



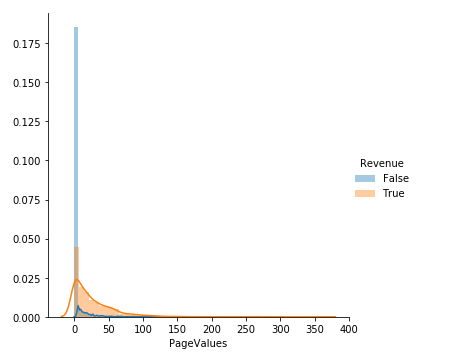
We identified that value of ProductRelated web page affects the revenue. But since the data has more False revenues and is overlapping it is difficult to correctly infer anything from this graph. But it was identified the mostly positive revenues lied between values 0- 100



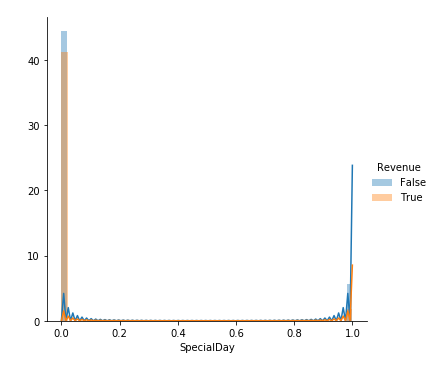
Again, the bounce rate of a customer plays a major role, the lesser the bounce rate, more the positive revenue.



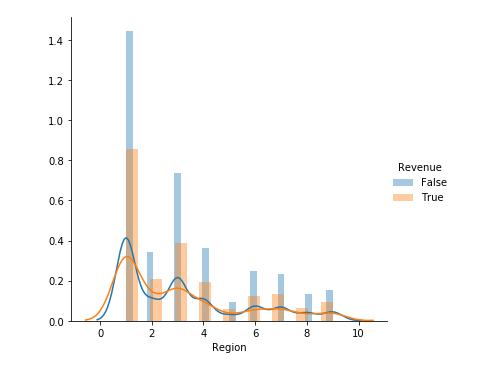
Exit rate trends show that exit rates between 0.00 to 0.05 are the maximum revenue generating values.



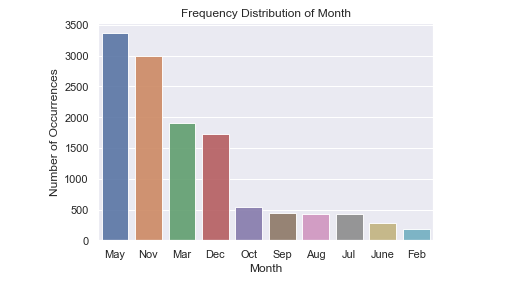
Page value trends indicates that positive revenue is generated on page value as 0.



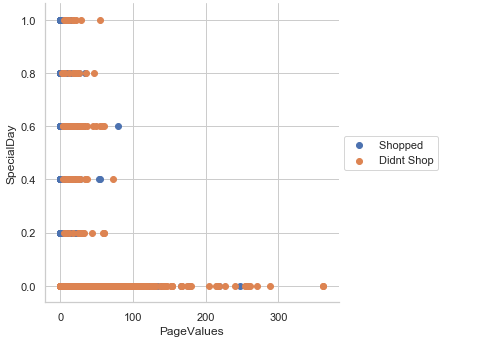
SpecialDay value trends indicates that positive revenue is generated on SpecialDay value as 0. Which means that it is somehow un-related as per this data.



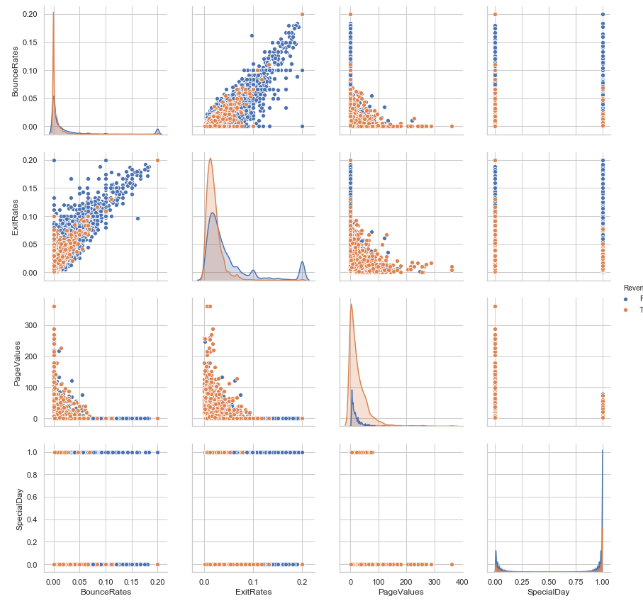
Region values graph shows that maximum revenues are generated for values 1-4.



May,November , Mar and December are the most revenue generating months and Users tend to shop more in these months.



We also tried to understand the cumulative effect and dependency of Special day and Page Values features on Revenue. But there does not seem to have a direct dependency of these features on each other.



Then from the 3D pair plot graphs, we tried to infer the same dependency of 4 features Bounce Rate, Exit Rate, Special Day and Page Values on each other.