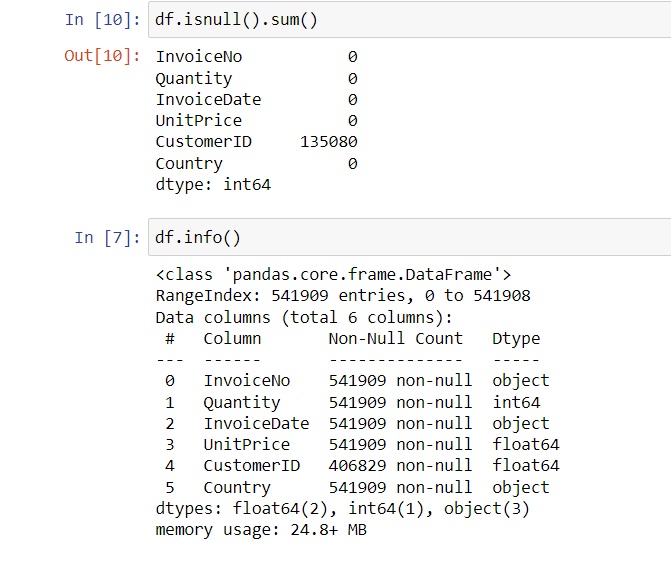
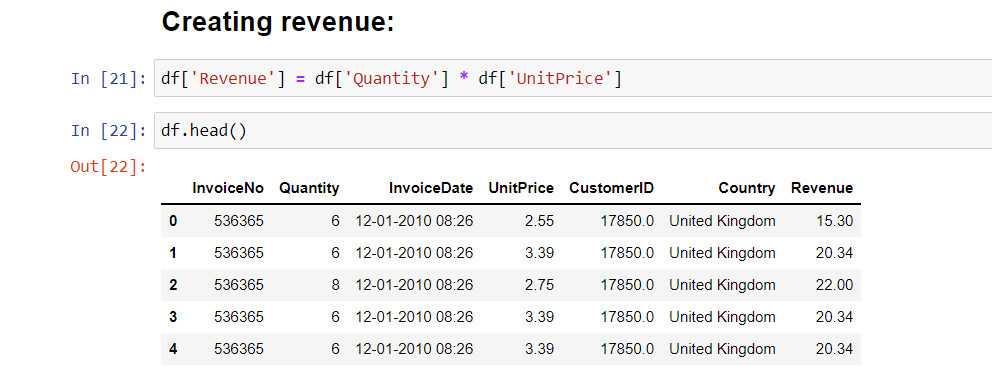
**Business Insights**

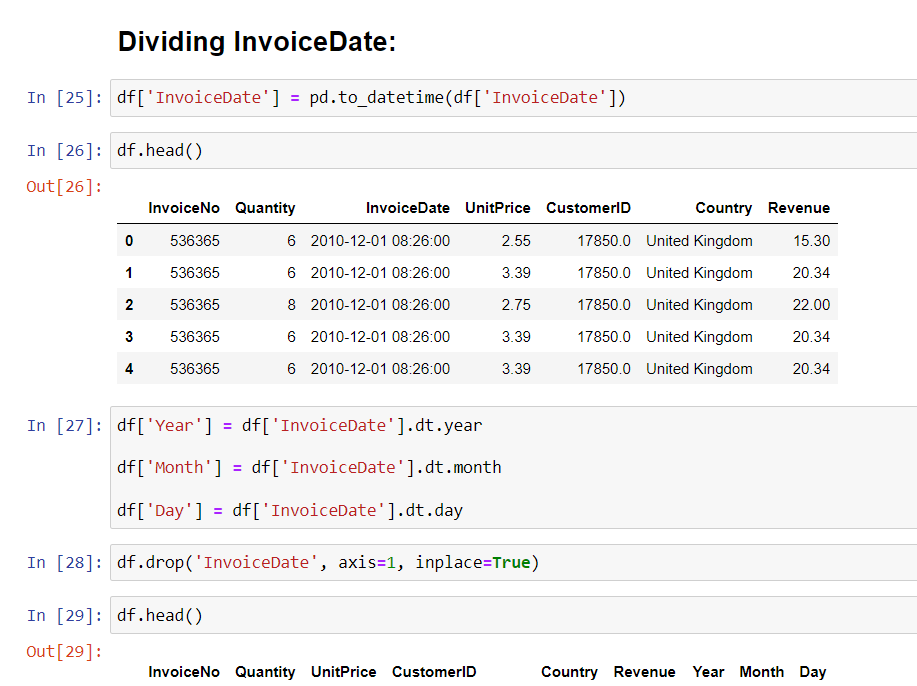
**-By Team 3**



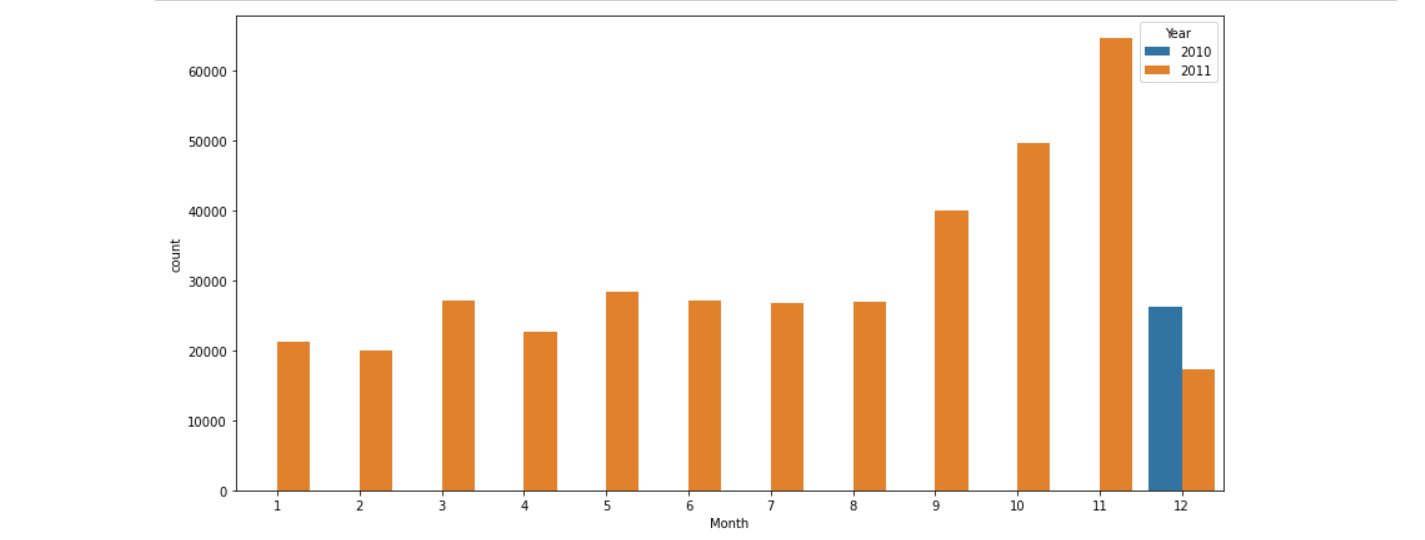
-We dropped all the null values in CustomerID because we cannot justify null value imputation in this case.



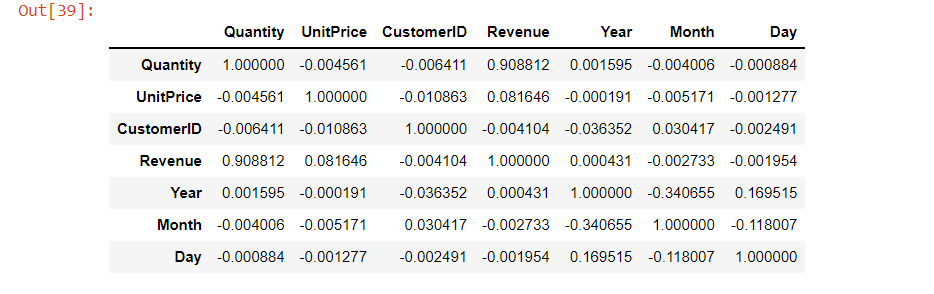
-A new column named ‘Revenue’ has been created out of the already two existing columns (Quantity and UnitPrice) to make our data more intuitive.



-The Division has taken place inside InvoiceDate and three different columns are created out of it: Year, Month and Day in order to check how recent the customers are.



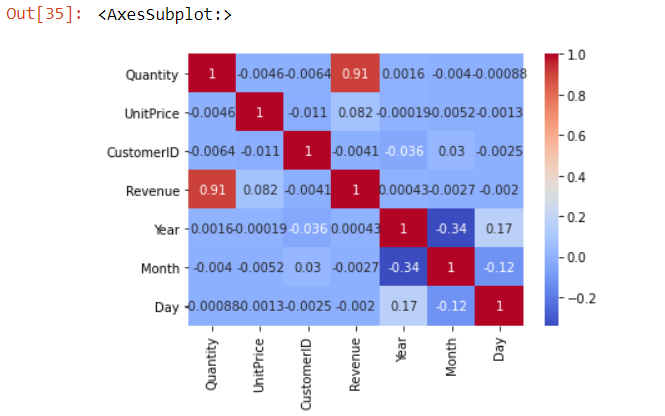
-This countplot is made to visually check the spread of the data. It shows that the data that we’re given is from the last month (December) of 2010 till the end of 2011.



-This is a Correlation table that helps identify how strongly are your given Independent Variables correlated. A strong Correlation can be negative as well as positive.

* The threshold value for a positive correlation is above 0.5, where 1 signifies strong positive correlation.
* The threshold value for a negative correlation is below

0.0, where -1 signifies a strong negative correlation.

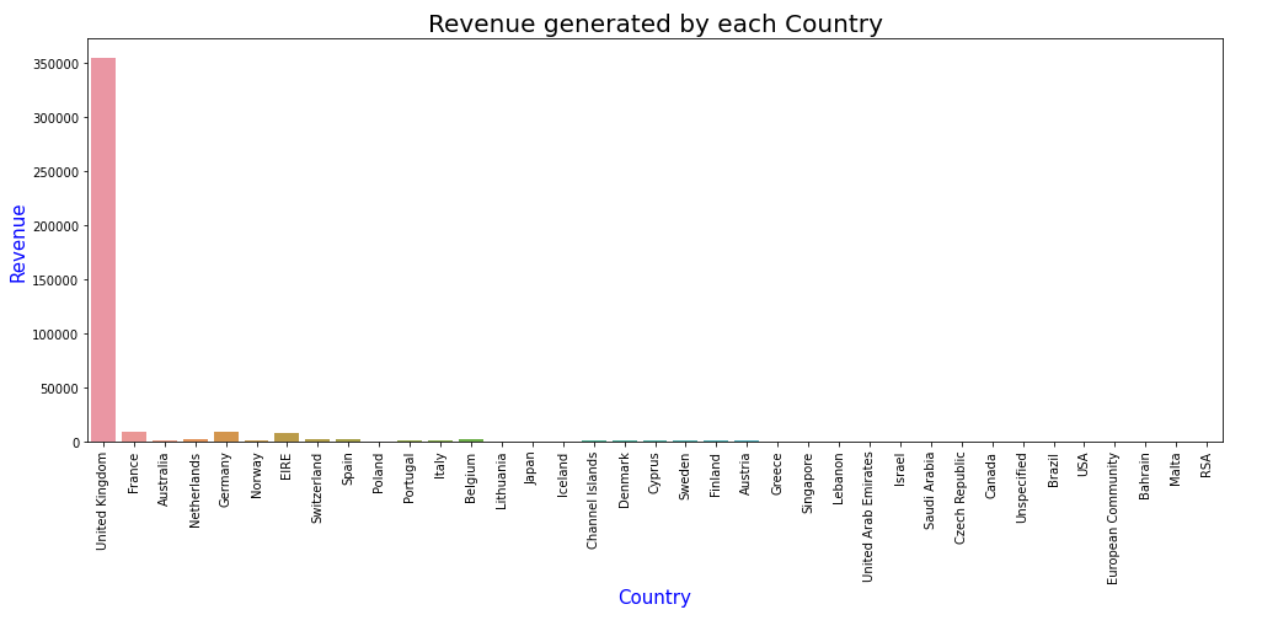


-The dark red boxes signify strong positive correlation.

* Quantity and Revenue Independent variables are highly, positively correlated.

-The dark blue boxes signify strong negative correlation

* Year and Month Independent variables are negatively correlated.

-This is a Countplot between Revenue and Country. This can help figure out which Country is generating highest and lowest Revenues.

* We can clearly see in the figure above that United Kingdom is the key Revenue generator of this E-Commerce Retail Business. This can be due to multiple factors like:

-More awareness about the Business in UK.

- The promotional activities like advertisement are only prominently reaching the masses of UK.

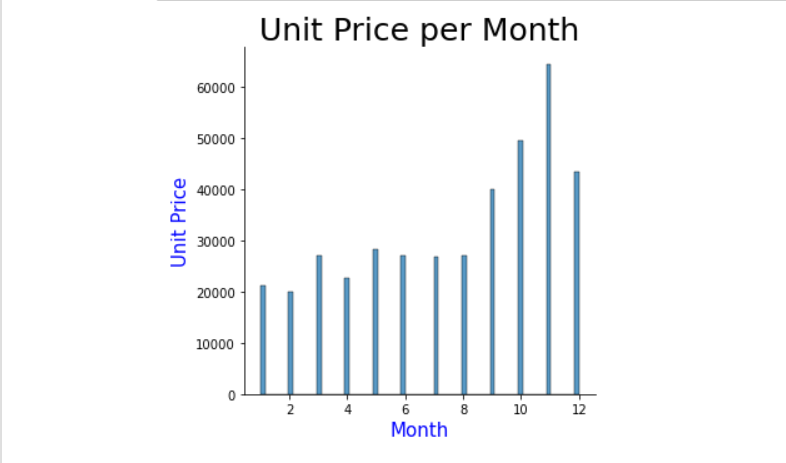
-It is possibly a homegrown business of UK itself.

-To grow their business in other countries, they can provide sponsorships, and build influencer relationship before expanding their business in any other country.

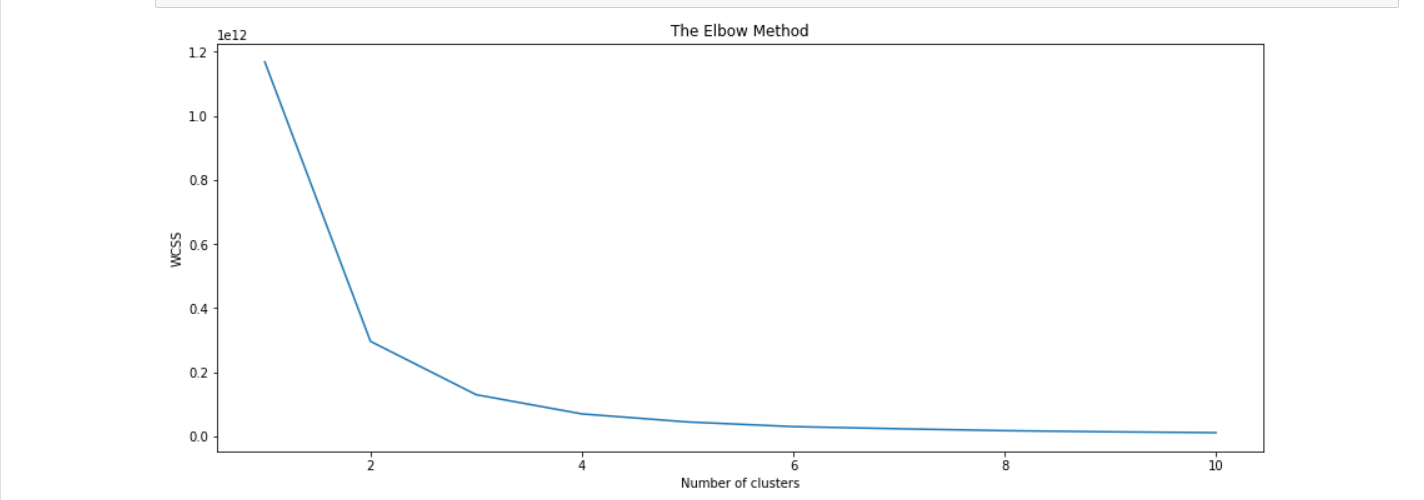


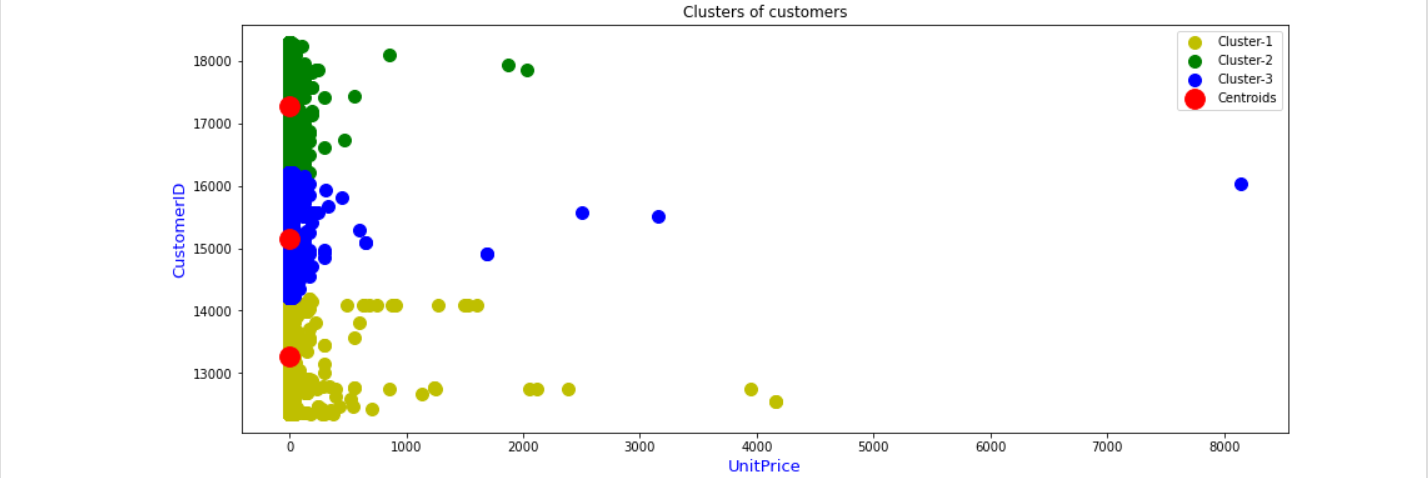
-This is a distribution plot between Revenue and Month to identify in which month did maximum Revenue generate.

* We can see that maximum Revenue was generated in the month of November. This can be due to multiple fall festivals prominently celebrated in the UK (it is the only significant buyer), like, Thanksgiving, Guy Fawkes Day etc.



-This is another distribution plot made to recognize maximum UnitPrice in a month

-This is an Elbow plot that helps determine the value of K.



-This is a scatter plot of 3 clusters chosen out of CustomerID, it indicates: the more the UnitPrice increases, the more likely the customer is going to Churn.