

Data Analysis Challenge for Summer Internship

Data must be processed into knowledge. Having huge data stored is of no use if we have not mined knowledge from it. Hence, I've tried to mine and visually display the huge data. Having graphs helps the business understand, what's hidden in data. This analysis can be used to take important business decisions. I've applied many data mining algorithm on the data, like forecasting and market basket analysis. So that we can find the hidden pattern in the data and can estimate the things like future sales.

I've transformed the data to make 6 visualizations

1) Forecasting Engine of Sales: The graph plotted here has a shadow and a solid line. The solid line is the actual sales that has been captured and the shadow represents the forecast range. The time between December to June has been considered as test data

2)Distinct Customer Total: This graph has distinct customers in every country for each year it also has total unique customers of both the countries and the year

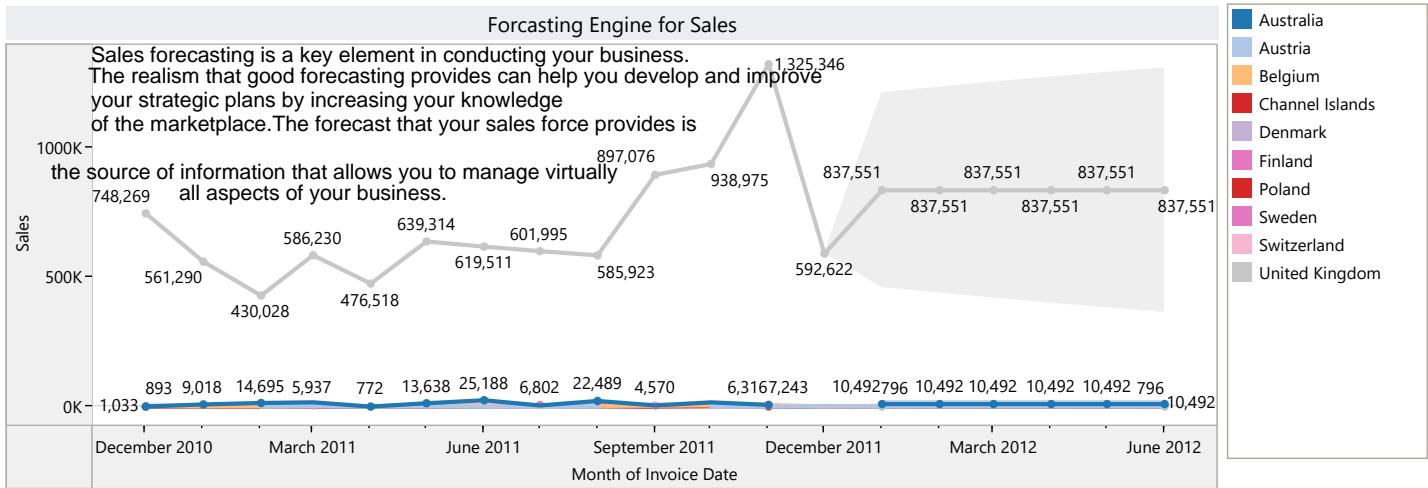
3)Cluster of sales: The sales of each countries are analyzed, and the countries were grouped according to the sales. The countries with highest sales were grouped together and then the cluster with medium sales were grouped together and the companies with low sales were grouped together. Hence, we can analyze.

4)Sales of products: This visualization helps in find out the trend of sales of each product.

5) Market Basket Analysis: In this Model, I have come up with analysis that if you buy a certain group of items, you are more (or less) likely to buy another group of items. Here I've chosen Base product as 2 Picture book Egg Easter Chicks. From the graph, we can see that more than 885 invoices have 2 Picture Eggs Easter Chicks and Jumbo Bag Baroque black and white together. That is, these two products have been brought together 885 times. With this type of analysis, we get to know the shopping trends of the customers. Market basket analysis can be used in deciding the location and promotion of goods inside a store. If, as has been observed, purchasers of Barbie dolls have are more likely to buy candy, then high-margin candy can be placed near to the Barbie doll display. Customers who would have bought candy with their Barbie dolls had they thought of it will now be suitably tempted

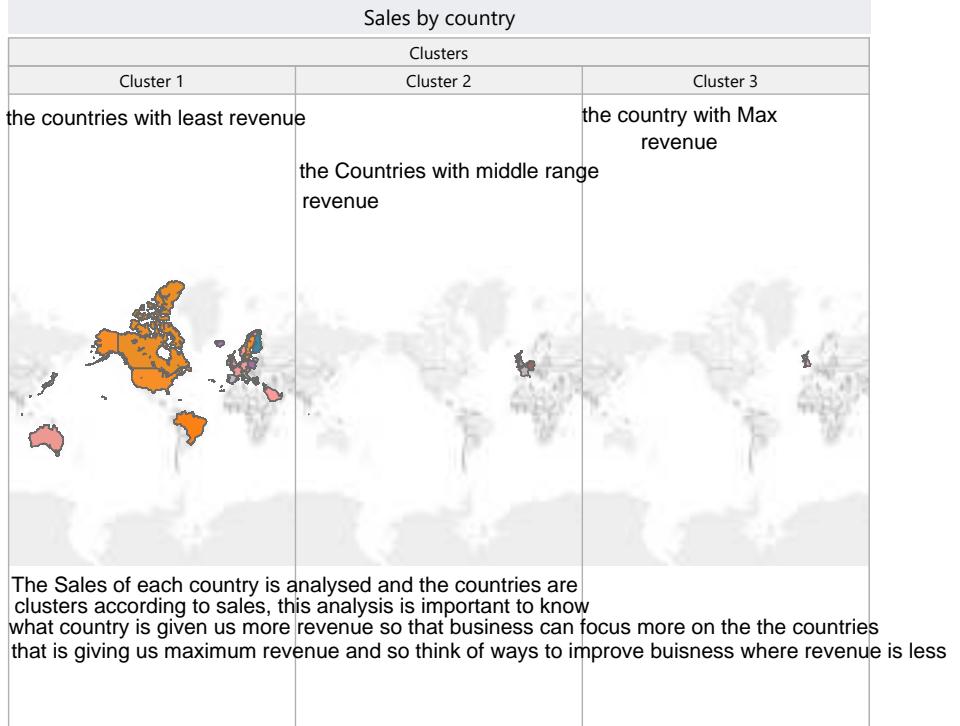
6) This Graph Shows what product is trending in which country. This Trend Analysis is important to apply promotions to the products that are not doing well in a country. We can also analyze the trends of a

The shadow Represents the forecast area and the thick line represents the actual sales values



Distinct Customer Total

Country	Invoice Date		
	2010	2011	Grand To..
United Kingdom	871	3,835	3,950
Germany	21	93	95
France	15	84	87
Spain	4	30	31
Belgium	4	25	25
Switzerland	2	21	21
Portugal	6	18	19
Italy	3	13	15
Finland	1	12	12
Austria	2	10	11
Norway	1	10	10
Australia	3	9	9
Channel Islands	1	8	9
Denmark	1	9	9
Netherlands	2	8	9
Cyprus	1	8	8
Japan	3	8	8
Sweden	1	8	8
Poland	1	6	6
Canada		4	4
Greece		4	4
Israel	1	3	4
Unspecified		4	4
USA		4	4
EIRE	3	3	3



This graph gives us the list of the countries with count of distinct customers

Distinct customer with respect to countries and year, with distinct number of customer grand total. This will give us the the count ,So that the business can focus on how to attract more customers

