

Inferences & Insights

1. Out of the Total Sales of 10.9 million , 9.6 million comes from 'United Kingdom'. This accounts for 88% of the Total Sales. This could indicate either the data is imbalanced else the business chain has heavy exposure to customers from UK.

2. Maximum data being from UK , it makes us difficult to compare and get insights from other countries just on the basis of total sales. Hence we also need to consider the Average sales across different variables specifically Average sales per customer.

For Finding average sales we need the number of unique customers hence removing the data points with missing customer ID.

Top 10 countries by Total sales:

	Country	sales
0	United Kingdom	9607400.487
1	EIRE	292649.880
2	Germany	280935.631
3	France	213969.980
4	Switzerland	56143.210
5	Spain	49678.120
6	Netherlands	48611.910
7	Belgium	46550.290
8	Portugal	36149.820
9	Australia	27168.610

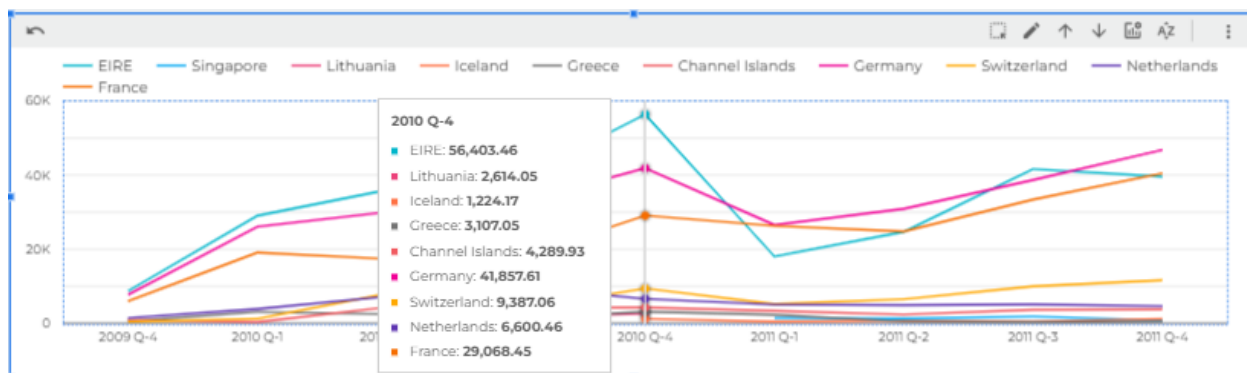
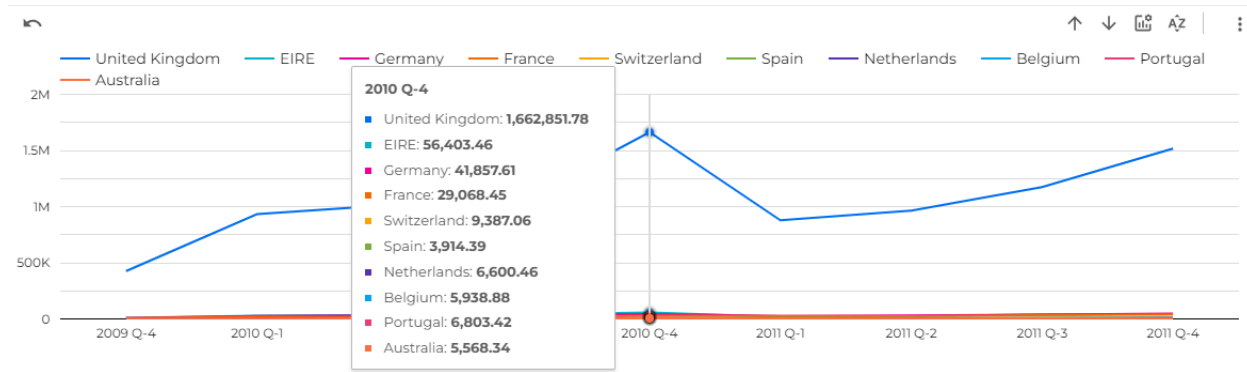
Top 10 countries by Average sales:

	Country	sales	unique_customer	avg_sales
10	EIRE	264169.710	3	88056.570000

31	Singapore	7884.610	1	7884.610000
22	Lithuania	4253.080	1	4253.080000
16	Iceland	3985.810	1	3985.810000
15	Greece	14023.710	5	2804.742000
6	Channel Islands	27079.720	10	2707.972000
14	Germany	280935.631	107	2625.566645
34	Switzerland	55558.600	22	2525.390909
24	Netherlands	48611.910	21	2314.852857
13	France	212714.640	93	2287.254194

5. Trend Analysis of Sales

1. YoY Total Sales Growth: The overall trend seems to have increased from 2009 to 2010. On surface this looks good but we only have last quarter data for 2009 so this might be an incorrect conclusion. On the other hand the sales has almost remained flat from 2010 to 2011.
2. QoQ Sales growth: The QoQ sales trend gives a better idea about the whole picture. Here we can see the continuous growth is sales from Q1 to Q4, with surge in Q4. Surge in Q4 could mean many things like the bulk of products being sold could be something related to holiday/festivities as it is this time of the year. It could also be some good which are seasonal and might be used during that season of the year. Another possibility is the promotional events & campaigns might be coinciding with the high sales numbers or before the start of new year the old inventory could be sold at heavy discounts.
4. Likewise you can see the difference in growth rate of total sales might not capture the entire picture. Hence the growth rate of average sales per customer also needs to be factored. The Top 10 Countries by Total sales are not entirely the same as the Top 10 Countries by Avg. sales per customer



Similarly you can see a difference in the trends of the Total sales of product vs the Avg. Sales of the product.

Digging deep into the trend of the sales for countries, I tried to look at the growth rate of sales for countries with sales above 20000 and growth rate of sales for products with sales above 40000. This gives a good perspective about the emerging markets and SKUs for business , to focus our strategic investments and resources.

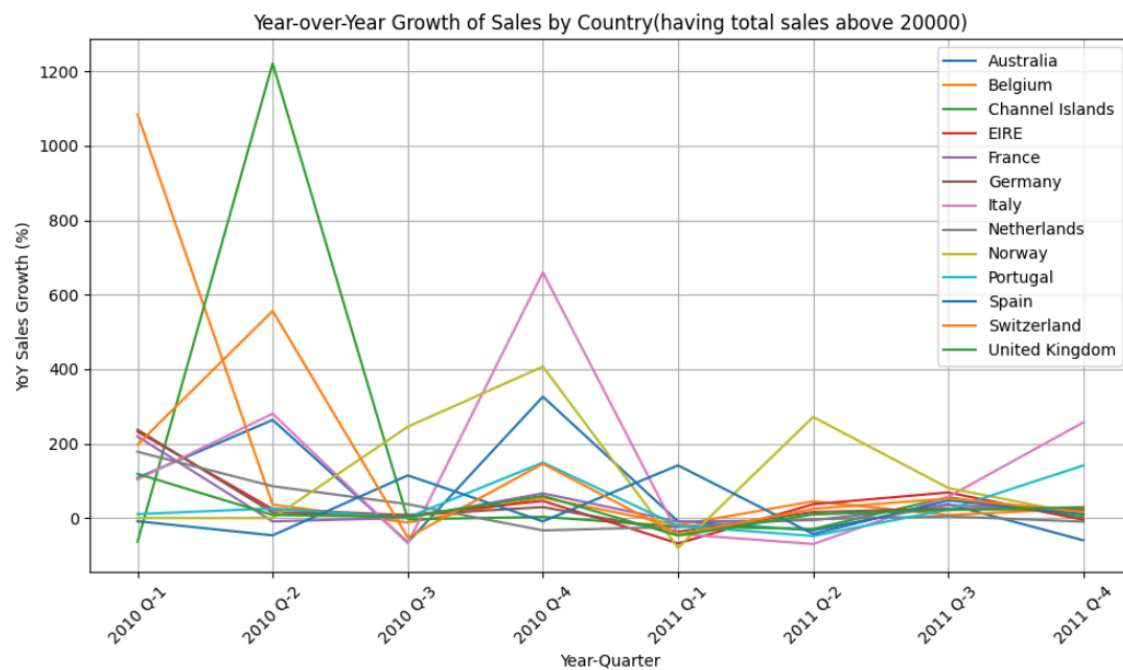
This might help to get an idea about the potential ROI from a country & the Product.

QoQ Growth rate of sales for countries (with sales above 20000):

invoice_yr_quarter	2009 Q-4	2010 Q-1	2010 Q-2	2010 Q-3	2010 Q-4	2011 Q-1	2011 Q-2	2011 Q-3	2011 Q-4
Country									
Australia	NaN	108.1 61399	263.73028 7	-64.9254 45	326.2191 43	-8.82148 0	-32.8534 03	37.3826 68	-59.7900 31
Belgium	NaN	1084. 45224 5	36.224935	-12.1030 73	54.72078 6	-18.5945 83	44.90637 2	8.26324 7	25.32075 3
Channel Islands	NaN	-63.71 9655	1221.5698 92	-3.46862 1	3.309573	-22.5546 80	-29.2211 84	55.9387 12	2.868067

EIRE	NaN	232.5 12760	23.880528	7.437289	45.98933 2	-68.0921 70	37.05931 8	68.8076 61	-4.92551 2
France	NaN	219.3 76098	-9.024348	0.568425	66.39505 1	-9.48753 7	-5.78930 8	34.7687 55	21.19280 5
Germany	NaN	237.4 41681	15.052386	7.726978	29.52451 9	-36.5127 87	16.19546 5	25.2434 01	21.11307 4
Italy	NaN	103.8 54227	280.47538 7	-67.8103 72	660.2503 62	-42.6398 67	-69.6737 38	57.0688 12	257.1170 78
Netherlands	NaN	178.1 57224	85.562372	37.25795 5	-33.1450 04	-23.4865 45	-3.15291 9	4.59128 1	-9.70429 5
Norway	NaN	0.000 000	0.000000	245.4200 10	406.9913 97	-79.9670 35	271.5688 06	80.8523 21	8.789180
Portugal	NaN	10.69 7118	25.783915	1.556524	149.2204 79	-20.8955 20	-48.2239 25	22.9184 39	141.3796 93
Spain	NaN	-8.519 268	-46.44598 3	114.6691 58	-8.89161 0	141.5377 62	-42.7663 19	47.3561 76	10.40598 9
Switzerland	NaN	197.4 02439	556.42842 5	-52.3344 11	146.0417 85	-44.3607 48	24.71715 9	52.6988 69	16.86573 0
United Kingdom	NaN	118.8 84821	8.455313	2.952166	59.64988 5	-47.1993 56	9.883190	21.6428 00	29.17465 5

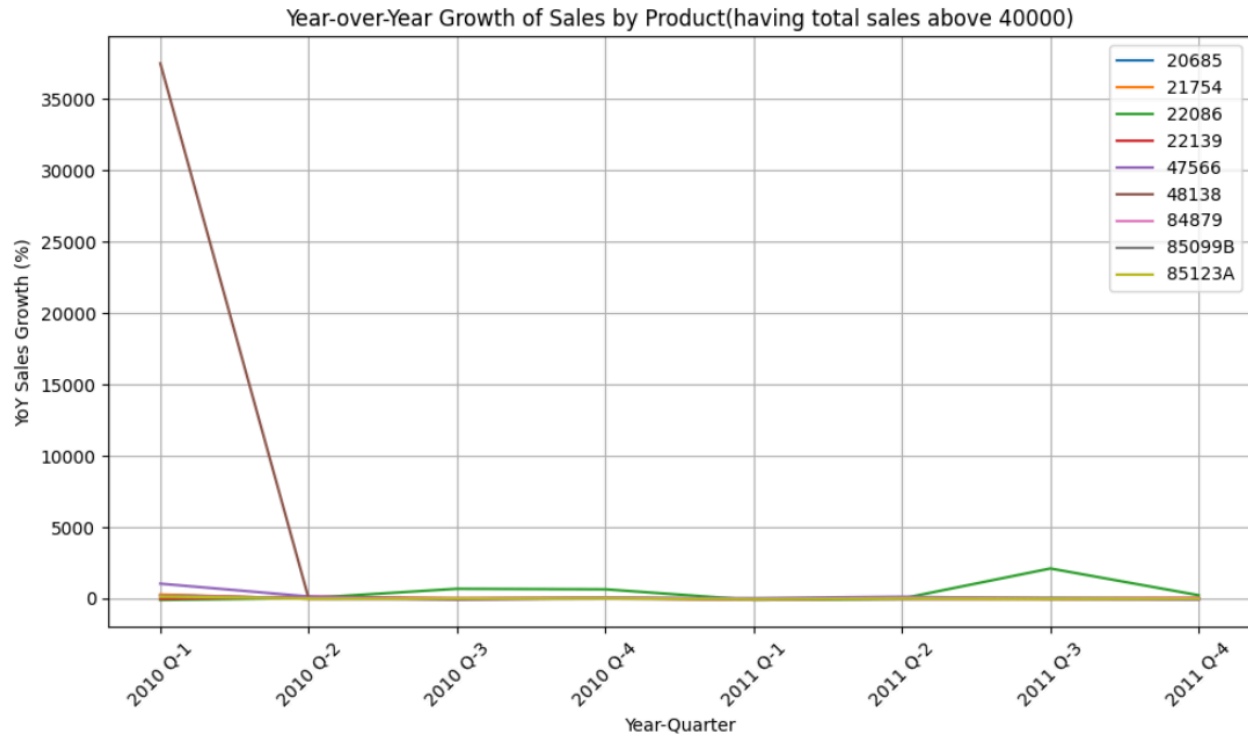
Trend of the growth rate of sales for countries (with sales above 20000)::



QoQ Growth rate of sales for Products (with sales above 35000):

invoice_yr_quarter	2009 Q-4	2010 Q-1	2010 Q-2	2010 Q-3	2010 Q-4	2011 Q-1	2011 Q-2	2011 Q-3	2011 Q-4
StockCode									
20685	NaN	257.320048	-12.806613	6.908850	7.845790	-31.960714	-18.205894	4.801095	-6.602984
21754	NaN	286.082084	-18.233536	20.464485	5.452019	-34.677425	-19.801938	-23.800507	-9.343390
22086	NaN	-96.549184	34.593072	685.198979	648.066767	-98.601661	2.481149	2109.256774	232.573269
22139	NaN	-18.556611	131.581871	-0.498256	32.644337	-59.620668	41.110476	15.335444	42.875227
47566	NaN	1044.255277	143.700023	-75.942688	48.789407	25.201437	130.621047	-18.321450	-64.815568
48138	NaN	37512.000000	-1.037100	-40.800126	-3.005274	-23.834725	-11.668146	-21.117851	-16.251437
84879	NaN	193.138731	22.559279	3.566443	58.479591	-49.230521	41.516420	0.746961	10.703345
85099B	NaN	176.675919	18.750621	-13.381712	39.271663	-30.065621	26.431902	37.456980	-14.438841
85123A	NaN	131.206391	-1.470695	-1.147340	25.851469	-46.813883	15.225653	-20.538774	-9.123320

Trend of the growth rate of sales for Products(with sales above 40000)::

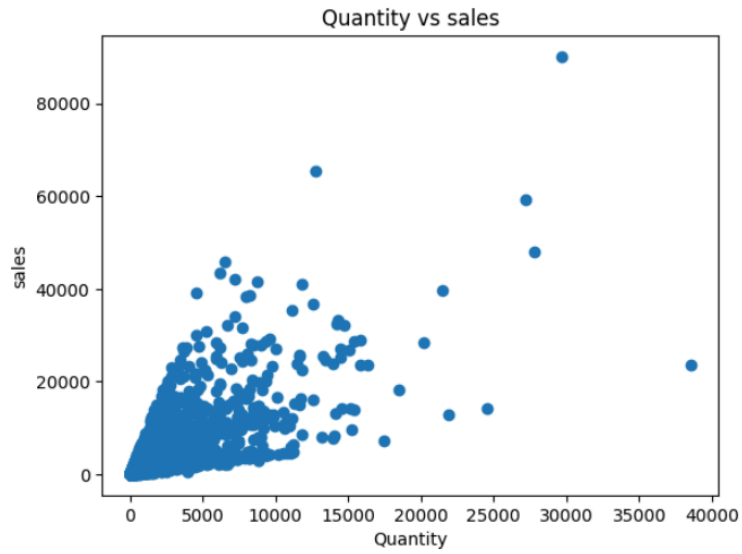


4. Relationship between numerical features:

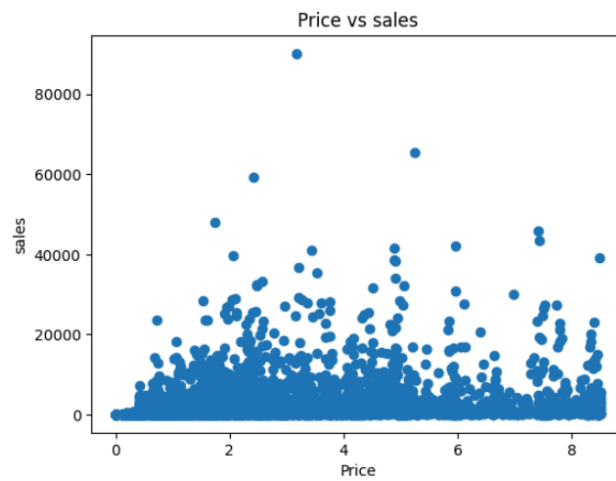
- a. Quantity vs Price : As you can see the distribution of data, it is right skewed data which shows that the items bought in high quantity are usually lower priced and vice versa.



- b. Quantity vs sales: This follows a linear relationship



- c. Price vs sales: These features follow gaussian distribution, though not a perfect one but somewhat related.



3.