

QUESTIONS TO BE ANSWERED

1. Use methods of descriptive statistics to summarize data.

1.1 Which Region and which Channel seems to spend more?

1.2 Which Region and which Channel seems to spend less?

Fresh	Milk	Grocery	Frozen	Detergents_Paper	Delicatessen
514	8323	6869	529	93	1040
2083	5007	1563	1120	147	1550
4720	1032	975	5500	197	56
15218	258	1138	2516	333	204
20893	1222	2576	3975	737	3628
5909	23527	13699	10155	830	3636
2153	1115	6684	4324	2894	411
23257	1372	1677	982	429	655
7858	1110	1094	6818	49	287
9193	4885	2157	327	780	548
53205	4959	7336	3012	967	818
56083	4563	2124	6422	730	3321
8475	1931	1883	5004	3593	987
11092	5008	5249	453	392	373
31614	489	1495	3242	111	615
10379	17972	4748	4686	1547	3265
29526	7961	16966	432	363	1391
6623	1860	4740	7683	205	1693

Through the in-built tool of Microsoft Excel for data analysis we prepared a summary table of the descriptive statistics of the given numerical data.

	Fresh	Milk	Grocery	Frozen	Detergents_Paper	Delicatessen
Mean	12000.29773	5796.265909	7951.277273	3071.931818	2881.493182	1524.870455
Standard Error	602.9377463	351.8456765	453.0455118	231.4374703	227.2985424	134.4432754
Median	8504	3627	4755.5	1526	816.5	965.5
Mode	3366	1196	2062	425	918	834
Std Deviation	12647.32887	7380.377175	9503.162829	4854.673333	4767.854448	2820.105937
Smpl Variance	159954927.4	54469967.24	90310103.75	23567853.17	22732436.04	7952997.498
Kurtosis	11.53640849	24.66939775	20.91467039	54.6892807	19.00946434	170.6949393
Skewness	2.561322752	4.053754849	3.58742869	5.907985692	3.631850631	11.15158648
Range	112148	73443	92777	60844	40824	47940
Minimum	3	55	3	25	3	3
Maximum	112151	73498	92780	60869	40827	47943
Sum	5280131	2550357	3498562	1351650	1267857	670943
Count	440	440	440	440	440	440
Largest(1)	112151	73498	92780	60869	40827	47943
Smallest(1)	3	55	3	25	3	3

- Fresh item (440 records), has a mean of 12000.3, standard deviation of 12647.3, with min value of 3 and max value of 112151.]. range = max-min =112,148

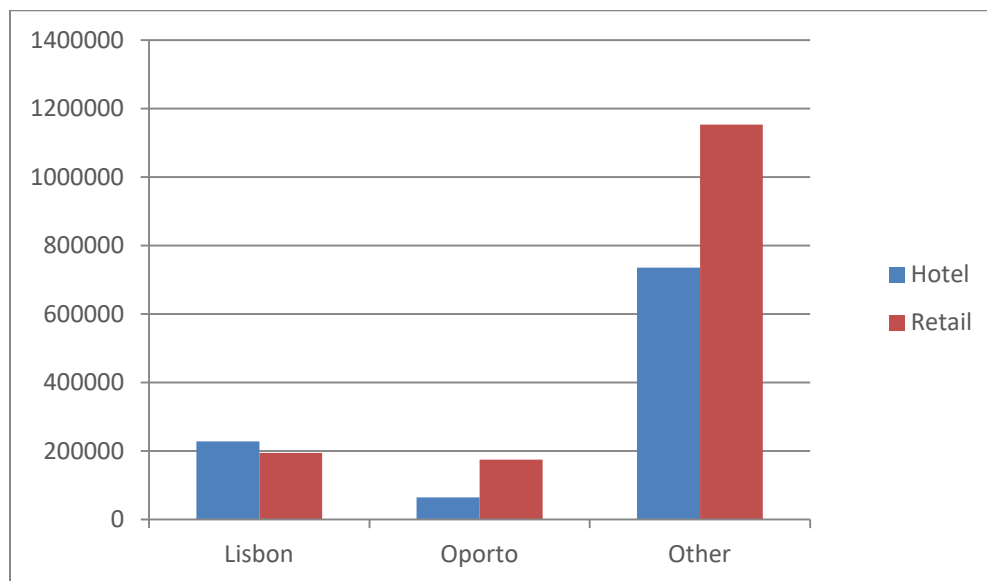
- Milk item (440 records), has a mean of 5796.27, standard deviation of 7380.38, with min value of 55 and max value of 73498. range = max-min =73443. •

Grocery item (440 records), has a mean of 7951.28, standard deviation of 9503.16, with min value of 3 and max value of 92780. range = max-min =92777.

- Frozen (440 records),has a mean of 3071.93, standard deviation of 4854.67, with min value of 25 and max value of 60869. range = max-min =60844.

- Detergents Paper (440 records), has a mean of 2881.49, standard deviation of 4767.85, with min value of 3 and max value of 40827. range = max-min =40824.
- Delicatessen (440 records), has a mean of 1524.87, standard deviation of 2820.11, with min value of 3 and max value of 47943. range =47943-3=47940.

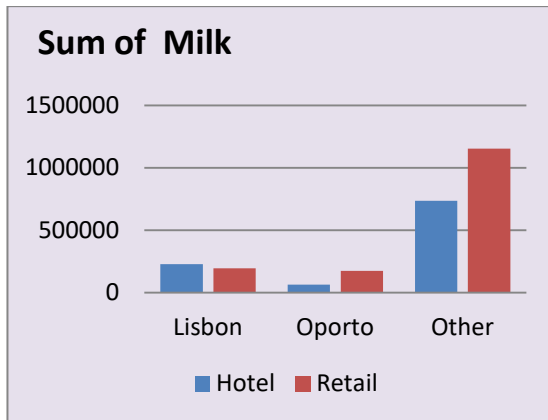
CLUSTERED COLUMN CHART



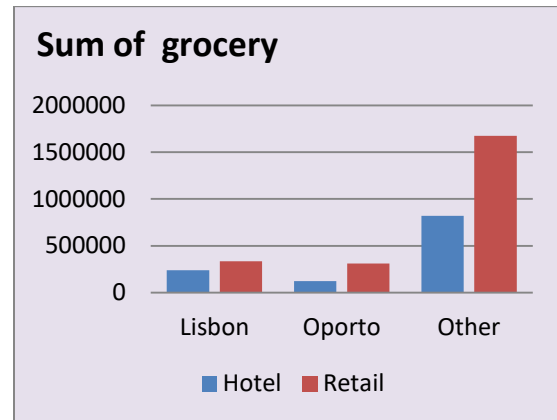
Highest spending in the Region is from Others and the highest spending in the channel is from Hotel.

Lowest spending in the region is from Oporto the lowest spending in the Channel is from Retail.

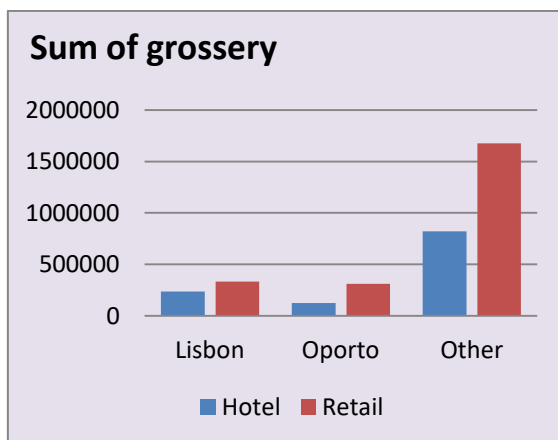
2. There are 6 different varieties of items are considered. Do all varieties show similar behaviour across Region and Channel?



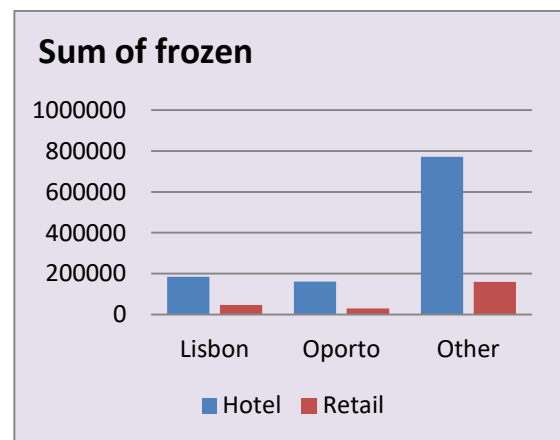
Based on the plot, milk item are sold more in the Retail channel



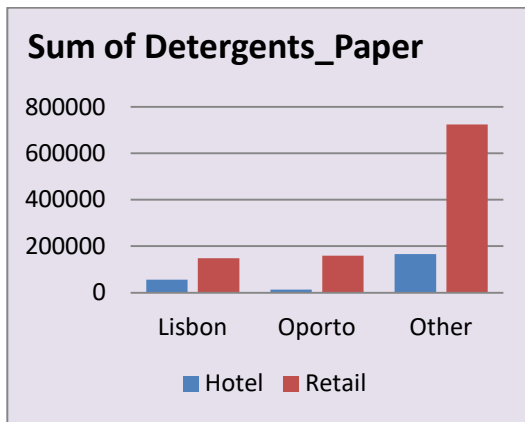
Based on the plot, Fresh item are sold more in the Hotel channel



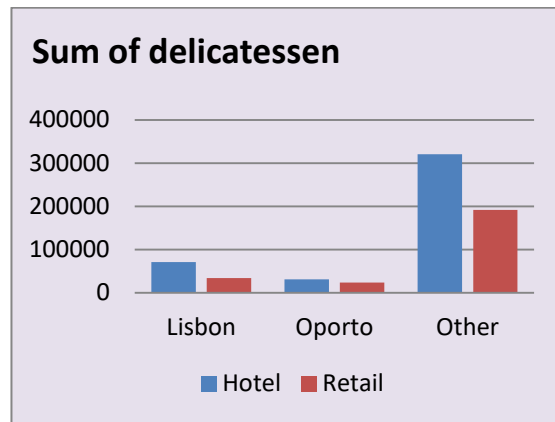
Based on the plot, grocery item are sold more in the Retail channel



Based on the plot, frozen item are sold more in the Hotel channel



Based on the plot, Detergents Paper is sold more in the Retail channel



Based on the plot, delicatessen is sold more in the Hotel channel

3. On the basis of the descriptive measure of variability, which item shows the most inconsistent behaviour? Which items shows the least inconsistent behaviour?

DESCRIPTIVE MEASURE		
	Standard Deviation	Variance
Fresh	12647.32887	159954927.4
Milk	7380.377175	54469967.24
Grocery	9503.162829	90310103.75
Frozen	4854.673333	23567853.17
Detergents_Paper	4767.854448	22732436.04
Delicatessen	2820.105937	7952997.498

Fresh items have the highest Standard deviation So they show the most inconsistent behaviour.

Delicatessen items have smallest Standard deviation So they show the most consistent behaviour.

4. Which columns in the dataset are of type Categorical and which are of type Numerical ?

data

1. **Categorical**
 - Channel
 - Region

2. **Numerical**
 - FRESH
 - MILK
 - GROSSERY
 - FROZEN
 - DETERGENT PAPER
 - DELICATESEEN

5. On the basis of this report, as a strategist what are the recommendations you can provide for expanding the application product line?

- Since demand for fresh items are comparatively higher in all the regions, the adequate supply of these items should be ensured.
- Lisbon consists of more high spending customers than Oporto. Therefore, wholesaler should focus on distribution of products which are preferred by the hotels of Lisbon.
- Spending on milk is comparatively low so quality of milk products could be improved to encourage customers to purchase more milk products.
- More research and survey can be done on more preferred frozen items to increase frozen item sales.