Q.2-1: Create a text table for the Avg (Sales) for each subcategory using Sample Superstore? List which Sub Category is got Avg (Sale) more than $1000? -**Sample Superstore**

**Answer:** From the text table created for this question, we can understand that there are only two subcategories in Sample Superstore dataset which got an average above $1000. And those categories are: Copiers ($2198.9) and Machines ($1645.6).

Q.2-2: Create a Heat Table for the order date and Region against the Sub Category based in Count of Sales with two colours diverging that is distinguished by Sum of Profit - **Sample Superstore**

**Answer:** This heat table shows Year wise and Region wise Business Sales and Profit for Sample Superstore Dataset.

Sales: The Sales dimension has been detailed by size. It specifies the larger the size of square, the more the count of sales.

Profit: The profit has been differentiated by color. The shades of red colored square shows the loss occurred in that business. The intense the shade of red, the more is the loss. Likewise shade of Grey color shows the positive Profit. The darker shade of grey specifies that there is a large amount of profit.

* Refer following figures:

|  |  |
| --- | --- |
| q2-2.JPG | q2-2-2.JPG |

Q.2-3: Create a Highlight table for the States for the Order Date Year whose highlighting is done based on Sum of profits - **Sample Superstore**

**Answer:** The cells for this table are highlighted based on the amount of profit. If any cell in the table is highlighted by any shade of red, it means there’s some negative profit i.e. loss whereas if the cell has shade of grey then there’s some positive profit. Also the intensity of color shade shows how much profit/loss got.

* Refer following figure:

|  |
| --- |
| q2-3.JPG |

Q.2-4: Which customer is having maximum of sales in the year 2012? -**Global Superstore**

**Answer:** The treemap for this question is plotted for top 10 customers, using Order Date (Year), Customer Name and measure Sales. According to this treemap, the customer named Sean Miller has the maximum sales of 27,470 in year 2012.

Q.2-5: How much is profit share less in Pennsylvania when compared to New York? -**Sample Superstore**

**Answer:** According to the highlight table, we can see the profit for New York is 74,039 whereas the profit for Pennsylvania is -15,560 i.e. it’s facing losses.

Q.2-6: Check for the pane wise percentages of sales with Category, Sub- Category and quarter wise order date, also check for the Row wise grand totals and Column wise grand totals. - **Sample Superstore**

**Answer:**

- We can see the total Sales and sales percentage for every category, quarter wise and year wise.

-The Grand total of both sales and sales percentage for all categories and each Year & Quartile is displayed at the bottom of table.

- Also the grand total of sales and sales percentile for each category over every quartile and year is displayed at the end of all columns.

-The total percentage is divided in four quartiles for each year.

Q.3-1: Use Global Superstore. Check Which Western Country in EMEA region has least profit percentage.

**Answer**: According the Symbol map, Macedonia country has the least positive profit of 43. On the other hand Turkey is the least among all with profit of -98,447 which shows the maximum loss among all other EMEA countries.

Q.3-b: Use **“Sample Superstore. Xls”,** which state shares boarders only profit for tables

**Answer:** From the graph forplottedfor this question, all the states colored by green shades have profit for tables. Also the states colored by red shades symbolize for loss in tables. We get that Washington, Idaho, Nevada, Utah, Oklahoma, Louisiana, Mississippi, Missouri, Iowa, Minnesota, Wisconsin, Indiana, Alabama, Michigan, Virginia are the states who shares positive profit for tables only.

Q.3-c: Use **“Sample Superstore. Xls”,** which state has no data for Profits for Office Supplies

**Answer:** As per the symbol map, only Wyoming State has no data for Profit for Office Supplies. Rest all the states have some positive or negative profit for office supplies.

Q.4-1: Which Customer name & Year is having all the Product Categories sum of profit less than over-all Average profit? -**Sample Superstore**

**Answer:** For this question, we can understand that if any customer is having all the Product Categories sum of profit less than over-all Average profit is highlighted with the dark green color. E.g. Here, Aaron Smalying is having profit less than overall average profit in year 2014, likewise.

* Refer following fig:

|  |
| --- |
| q4-1.JPG |

Q.4-2: What is the Maximum of Life Expectancy Female for the region Africa & year 2012? - **World Indicators**

**Answer:** According to bar graph plotted for this question, is the Maximum of Life Expectancy Female for the region Africa & year 2012 is 78 years.

Q.5-2: Which Region is having Sum of Energy Usage>1000000 and sum of Population 65+>10? - **World Indicators.**

**Answer:** By observing the graph, we get that by following criteria for this question Asia region has the most energy usage upto 34.55% followed by Americans (29.79%), Europe (25.10%), Africa (4.73%), Middle East and then Oceania with (1.27%) energy usage.

Q.6-1: Draw a trend line for profit as a linear function of sales only for product technology? -**Sample Superstore**

**Answer:** According to the trend line drawn using linear function for sales and profit for Technology, we get that there’s a positive relation between the sales and profit for Product technology. The coefficients for this trend line are -9.71792 as B0 and 0.195423 as B1. But the trend line is not capturing most of the data. Also it has the R-Squared value as 0.255272 which is much less 1. So we can say that linear function for this model is not much suitable.

Q.6-2: Create a histogram showing the number of Sales using Sales Bins of $1000. Which bins have profit ratios of more than 25%? -**Global Superstore**

**Answer:** As perthe bar plot for this question, there are total 7 bins which have profit ratio more than 25%, each has sales count as 23, 5, 3, 1, 1, 1, 1 respectively. The max profit ratio for this problem found to be of 147.29% whereas least profit ratio found was of 76.54% which has the maximum sales count of 23 among all.

Q.6-3: Using “**Sample Superstore”**, use order sheet create a histogram showing the number of orders using sales bins of $1000.

**Answer:** According to worksheet for this question, the bin having sales less than 1K $ sales has the maximum no. of orders (i.e. 9,526) and has total sales of 1,307,032, followed by 1k bin with total 328 orders, 2k bin with 77 orders, 3k bin with 23 orders, 4k bin with 21 orders and so on.

Q.6-4: Using **“Global Superstore**”, use the orders sheet, build a scatter plot showing the sum of sales on the x-axis and sum of profits on the y axis for all products (Product name). What is the equation for linear regression for products in Technology?

**Answer:** By the scatter plot plotted for this question, we can say there is a positive relation between sales and profit. Increase in sales will also increase the profit.

After plotting the regression graph, the equation we got for linear regression for products in technology is: Profit = 0.203737\*Sales + (-29.8652)

Here, coefficients for regression line are as B0 = -29.8652 and B1 = 0.203737.

Q.12-2: Use **“Sample Superstore”.** What is the Sales Forecast Estimate for the month of September 2018?

**Answer:** We had a data for years 2012 to 2015, and using that we made forecasting for next years from 2016 to 2018 starting from exact actual date. The lighter green colored line represents the actual trend line whereas the dark green shade of line represents the forecast for next three years. For September 2018, we estimated that there will be sales of 646,750$ approximately.

Forecasting model used for this question is built using trend and season as ‘Additive’ with 95% of confidence interval. We can check the quality of model using ‘Describe Forecast’ option. Inside this option, we can check the summary of model, which also specifies the quality of model we have built. Along with summary, there’s other option called ‘Model’ inside which we can have all the information about the model like it’s trend and season, quality metrics and soothing coefficients.

Here for our model, the MASE i.e. Mean Absolute Scaled Error is 0.26, which is an accuracy measure. It means the values we are predicting are 0.26 times different from any random walk models which are still acceptable. The less MASE is the better is the model.

Q.13: Create a Pie Chart using regions and sum of sales, sort the pie in ascending order, increase the size in the view and label them with Count of Quantity and Sum of Profits- **Sample superstore**

**Answer:**  According to Pie chart, we conclude that the West region has the most no. of quantity (3,203) with profit of 108,418 where as South region has the least no. of quantities (1,620) with profit of 46,749. On the other hand East region has the second largest no. of quantity with profit of 91,523 and Central region has the third largest no. of quantity with profit of 39,706.