# Data Analysis with Tableau

Tableau Learning Partner
Prepare and Analyze the Superstore Dataset

Question 1

What action(s) should you take before you start creating data visualizations and calculating statistics? Select all that apply.

1 point

**Preprocess the data.**

Turn patterns into insights.

**Explore the data.**

Communicate your insights.

Question 2

How many fields (that were not generated) were interpreted by Tableau as geographic?

1 point

**4**

Question 3

With all of the hierarchies and folders collapsed, what is at the top of the left sidebar?

1 point

**The Customer folder**

The Location hierarchy

The Measure Names dimension

The Customer ID dimension

Question 4

What is the average profit per item for the West region?

Be sure to format your average profit per item so that it is a currency showing two numbers after the decimal to represent the number of cents. Your answer below should be in the following form:

#.##

Note the following:

There is no $ symbol

There is a number to represent the dollar amount

The dollar amount is followed by a decimal

The decimal as followed by two numbers to represent the number of cents.

1 point

**8.84**

Question 5

After adding the filters, how many states are currently in the table?

Hint: Because the two filters are acting independently, your table will not yet necessarily be showing the actual top 5 states by sales in the East region, but since there is a Top 5 filter, you should not be seeing any more than 5. Together, this means that your answer should be <= 5.

1 point

**2**

Question 6

Now add the regional filter to the context so that it is applied before the Top 5 filter.

Order the table so the average Profits are in descending order.

At this point, you should have the correct table, which shows the average Profits of the top five performing States in the East region based on their average Sales.

Which of the following states appear in your table? Select all that apply.

1 point

Connecticut

**Delaware**

**District of Columbia**

Maine

Maryland

Massachusetts

New Hampshire

New Jersey

New York

Ohio

Pennsylvania

**Rhode Island**

**Vermont**

**West Virginia**


Question 7

You're now interested in learning how the sales are distributed across all orders.

Create a histogram for the Sales.

Change the bin size to 50.

Add a filter on Sales to show a maximum sales amount of $1,000.

What is the skewness of the histogram?

1 point

Negative (left)

Symmetrical

**Positive (right)**


Question 8

How many orders had sales amounts that ranged between $150 and $200?

Hint: The fourth bin from the left should represent the bin with a range of $150-200. Hover over the bin to see the count.

1 point

**497**


Question 9

You're now looking into the state sales performance and are interested in how spread out they are.

Create a box plot to show the total Sales for each state and use it to determine which value below represents the third quartile.

1 point

$2,865

$7,465

**$49,096**

$89,474

$457,688


Question 10

Return to your box plot and break it up based on the newly created Sales Region. In other words, you should have four box plots on a single visualization, one for each newly created Sales Region. Each box plot should be made up of the respective States in each Sales Region.

Which Sales Region has the smallest interquartile range?

1 point

Northeast

**Northwest**

Southwest

Southeast


Question 11

You're interested in determining the correlation between the discount and the quantity sold. You have a hunch that the higher the discount, the higher the quantity will be.

Create a scatter plot with the average Discount on the x-axis and average Quantity on the y-axis, and set the level of detail to Sub-Category.

What is the R-squared value?

1 point

-0.2354

**0.0435**

0.4712

0.8727

## Question 12

You would like to further explore the sub-categories. This time, you would like to learn more about how each Sub-Category is doing in terms of Sales.

Build a table to show both the Category and Sub-Category as rows, showing the total Sales for each.

Add a quick table calculation to show the "Percent of Total."

What percentage of the total does the "Storage" sub-category make up?

1 point

0.54%

6.51%

9.01%

**9.74%**

## Question 13

Now add column grand totals and subtotals to your table.

What is the total sales from the "Office Supplies" sub-category?

1 point

$107,532

**$719,047**

$742,000

$2,297,201

## Question 14

You'd now like to review the daily average profit in 2016.

Create a time-series plot using the Order Date and average Profit.

Filter on the year 2016.

Add a distribution band to the table with the following settings:

Value: Standard Deviation

Factors: -1, 1

Sample (not Population)

You can leave "Recalculate band for selected data" checked.

Now it's time to determine the standard deviation. The upper and lower band lines will show you the values that are one standard deviation in both directions from the mean. This means that the band represents two standard deviations, so if you divide that number in half, you can determine the standard deviation.

For example, consider an example where an upper band line is at $100 and a lower band line is at -$20. This span is $100 - (-$20) = $120. Dividing that by two will give you $60, which would be the standard deviation.

Use the lower and upper distribution band lines (hover over them to obtain the values) to determine the standard deviation. (It's okay to use a calculator!)

1 point

$24

$86

$108

**$144**


15.

Question 15

Add a forecast to your time-series plot.

What is the forecasted average profit for January 7, 2017? (Hint: Hover over January 7, 2017.)

1 point

-$125

**$16**

$107

$220