

Connecting to & Preparing Data

1. Which of the following is the best reason to create a saved data source as a .TDS file
 - a. Those who wish to use the data do not have access to the underlying data.
 - b. You want to save the default field properties such as number formats and sort order.
 - c. You want to save a snapshot of the data that will not update even when the underlying data changes.
 - d. You need to apply an aggregation that takes too long when using a live connection.

Correct answer is B. TDS files save the data source metadata, such as the connection information, default field formatting, and sort order. They do not save the data source itself, and hence you should use this format if everyone who will use the data source has access to the underlying file or database defined in the connection information.

Documentation here: https://onlinehelp.tableau.com/current/pro/desktop/en-us/export_connection.html

2. When creating a union between tables, what will happen if the field names do not match?
 - a. Tableau will return an error message, explaining that you must modify the tables so that all tables used in the union have the same field names.
 - b. Tableau will automatically merge the fields, combining fields with similar field names.
 - c. Fields in the union that do not have matching field names will contain null values.
 - d. Fields with non-matching field names will be dropped, so that the union will contain only those fields with matching field names

Correct answer is C – if the field names do not match, the fields will be included in the UNION but will contain null values for the rows from the table that is missing the field.

<https://onlinehelp.tableau.com/current/pro/desktop/en-us/union.html>

3. In which of the following situations would an outer join be preferred to a blend?
 - a. when you need to include all rows from both tables, even when the join or blend criteria is not met
 - b. when the data is at different level of granularity
 - c. when you need to append rows from one table to rows in another table
 - d. when the tables are stored in different databases

A is correct because an outer join will include all rows from both tables, even when the join criteria is not met, while a blend will not include rows from the secondary table unless there is a match on the linking field.

B is not correct because blends avoid duplication when the tables are at different levels of granularity, making them preferable to joins in some situations.

C is not correct because unions are for appending rows.

D is not correct because both blends and cross database joins can combine tables stored in different databases.

4. Which of the following uses a primary and secondary data source?
 - a. Cross-database joins
 - b. Unions
 - c. Aggregation
 - d. Blending

Correct answer is D. Data blending involves a primary and secondary data source.

https://onlinehelp.tableau.com/current/pro/desktop/en-us/multiple_connections.html

5. Union the Stocks 2010-2013 and Stocks 2014 tables, and then create a chart showing average Open by year and for each Company from 2010 to 2014. In which year was the average open for Biogen Idec closest to Amazon's average open?

- a. 2011
- b. 2012
- c. 2013
- d. 2014

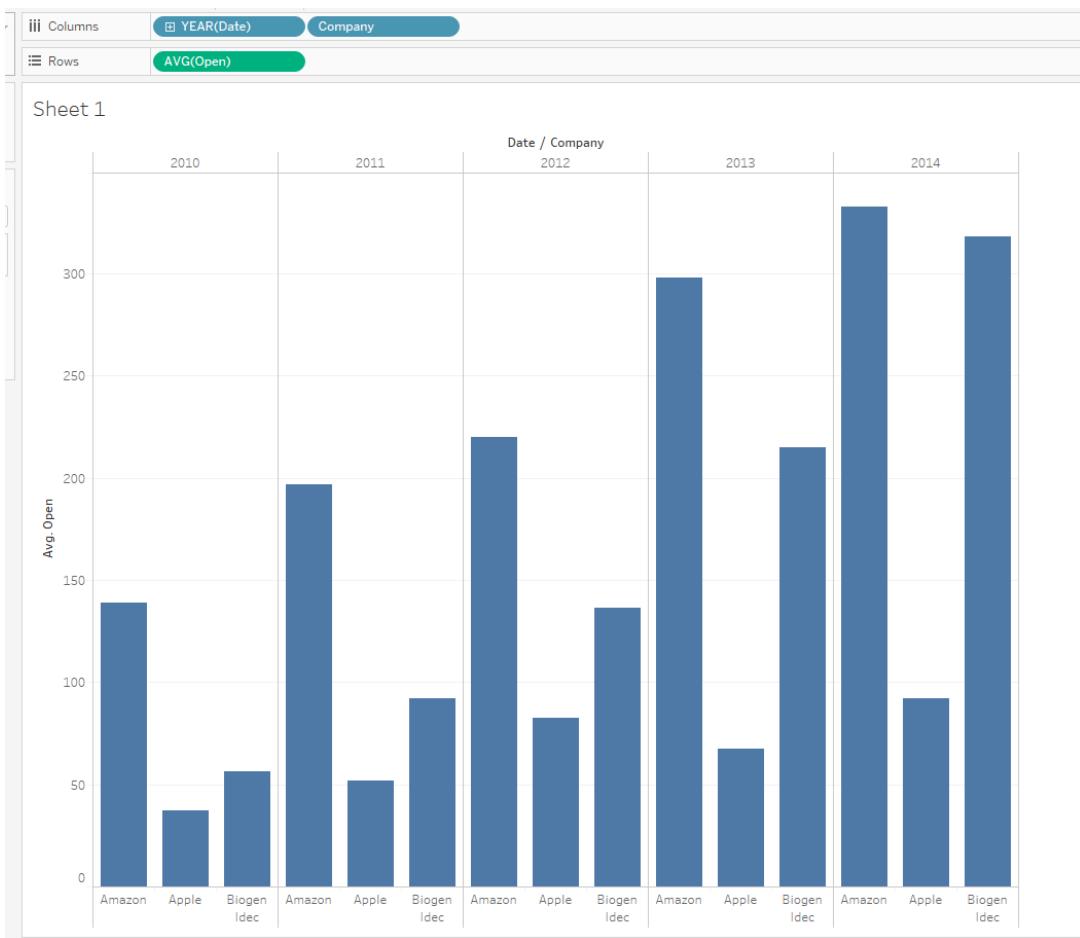
Solution: D, 2014

Create union of Stocks 2010-2013 and Stocks 2014 tables:

The screenshot shows the Tableau Public interface with the following details:

- File, Data, Window, Help** menu bar.
- Connections** section: A connection named "tds_sample_data_all" (Microsoft Excel) is selected.
- Sheets** section: Sheets listed include "Flights", "Stocks 2010 - 2013" (selected), "Stocks 2014", and "New Union".
- Stocks 2014+ (tds_sample_data_all)** view:
 - A legend indicates "Stocks 2014" (blue) and "Stocks 2010 - 2013" (orange).
 - An orange box highlights the "Stocks 2010 - 2013" sheet icon.
- Bottom pane:** A preview of the data with columns: "# Stocks 2014 Close", "Abc Stocks 2014 Company1", "# Stocks 2014 Date", "# Stocks 2014 High", "# Stocks 2014 Low", "# Stocks 2014 Number of Records", "# Stocks 2014 Open", and "# Stocks 2014 Volume".

This view shows average open by year and company. We can see Amazon and Biogen Idec are closest in 2014.



6. Blend the Southern Region table from the South_Superstore file with the Subcategory Profit Targets table from the Subcategory Profit Targets file. Add subcategory, AVG(Profit) and Profit Target to the view. For which of the following subcategories was average profit higher than the target?

- a. Appliances
- b. Art
- c. Binders
- d. Bookcases

Correct Answer: A

Add a connection to the South_Superstore, then add a new data source:

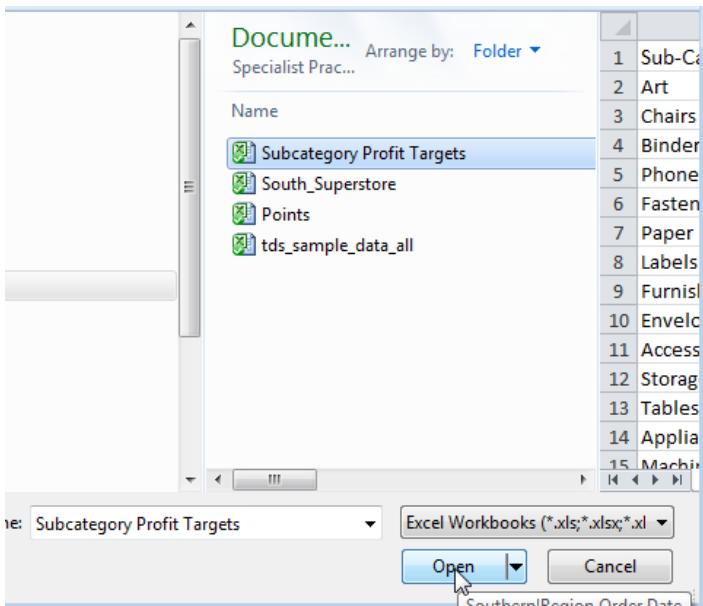
Southern Region (South_Superstore)

Southern_Region (South_Superstore)

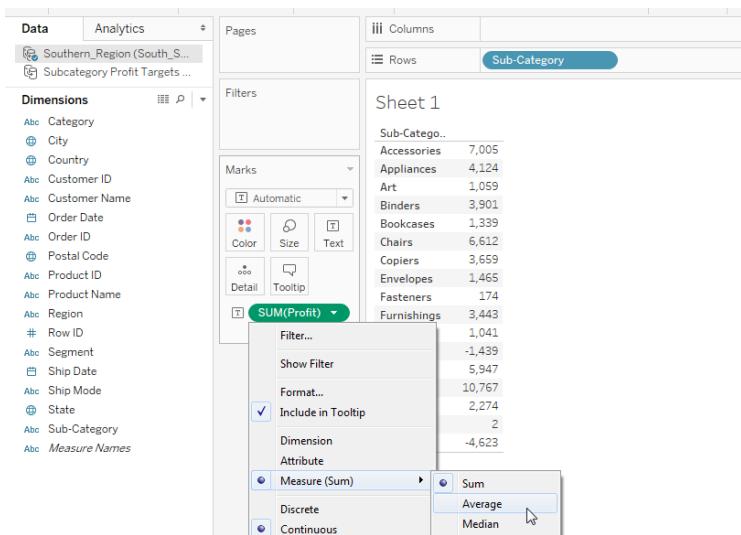
New Data Source

Southern_Region

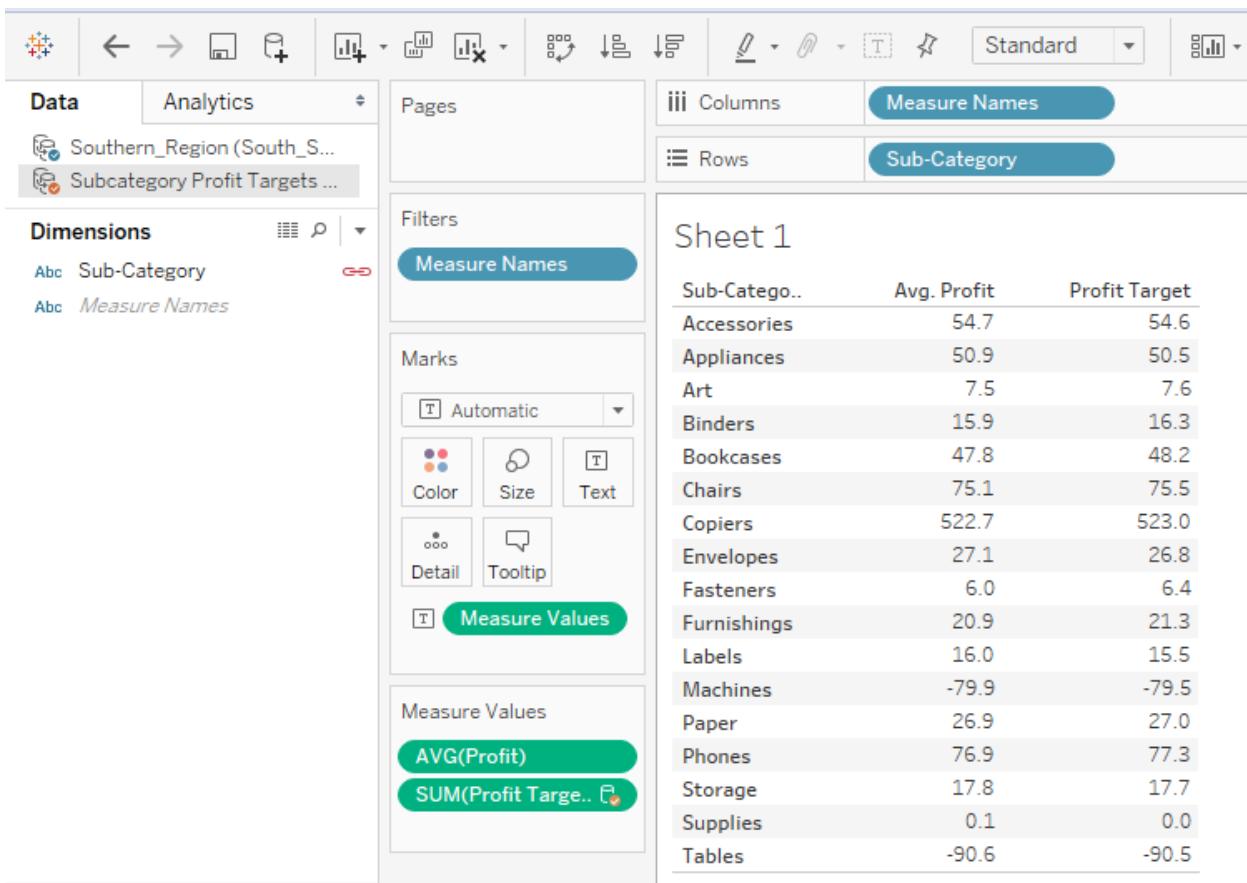
Select Subcategory Profit Targets



From Southern Region, add Subcategory and Profits, Set profits to Average



Select the Subcategory Profit Targets dat source, and add Profit Target:



7. Inner join Southern_Region with Returns using Order ID. This will limit the data to only those orders which were returned by the customer. Create a chart that shows sales by year of order date. In which year was the sales value of the returned items the lowest?
- 2013
 - 2014
 - 2015
 - 2016

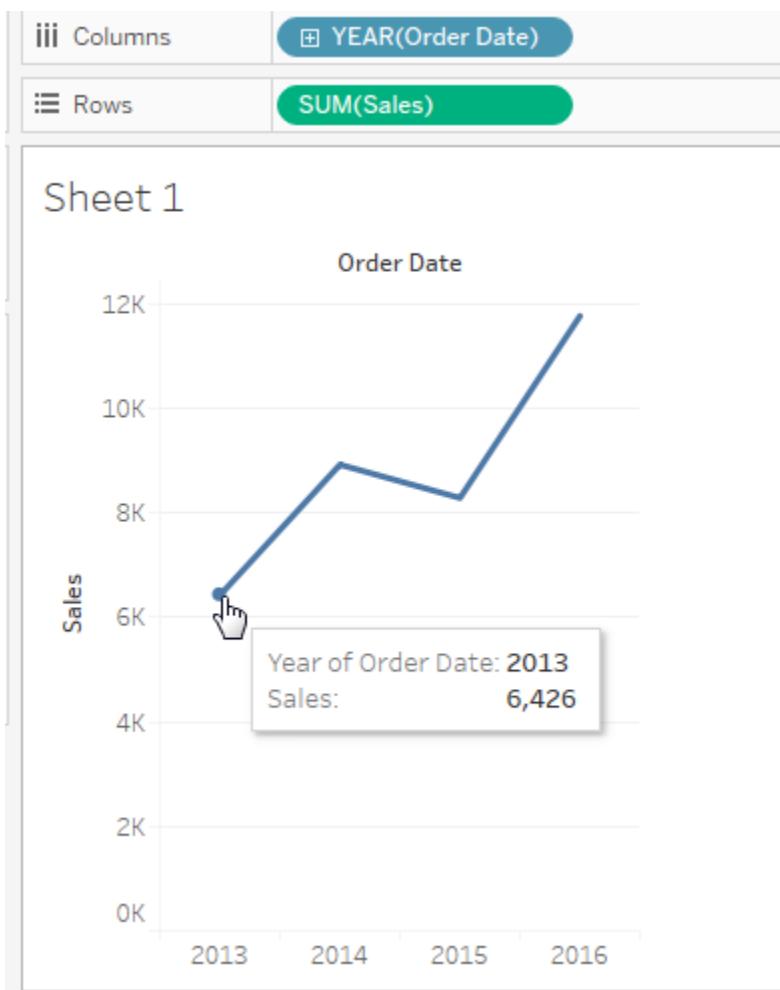
Answer: A. 2013

On the Data Source tab, create the Inner Join on Order ID

The screenshot shows the Tableau Data Source view. On the left, there's a sidebar with 'Connections' (South_Superstore, Microsoft Excel) and 'Sheets' (Returns, Southern_Region). A 'New Union' option is also present. In the main area, a join dialog is open between 'Southern_Region' and 'Returns'. The 'Join' dropdown is set to 'Inner'. Below it, 'Data Source' is 'Southern_Region' and 'Returns' is 'Returns'. The join condition is 'Order ID = Order ID (Returns)'. The resulting preview table shows data from both tables joined on their Order ID. The columns are Order ID, Order ID (Returns), Ship Date, and Sales.

Order ID	Order ID (Returns)	Ship Date	Sales
CA-2011-106054	7181	CA-2011-106054	1/6/2013
CA-2011-167199	7475	CA-2011-167199	1/6/2013
CA-2011-167199	7476	CA-2011-167199	1/6/2013
CA-2011-167199	7477	CA-2011-167199	1/6/2013
CA-2011-167199	7478	CA-2011-167199	1/6/2013
CA-2011-167199	7479	CA-2011-167199	1/6/2013
CA-2011-167199	7480	CA-2011-167199	1/6/2013

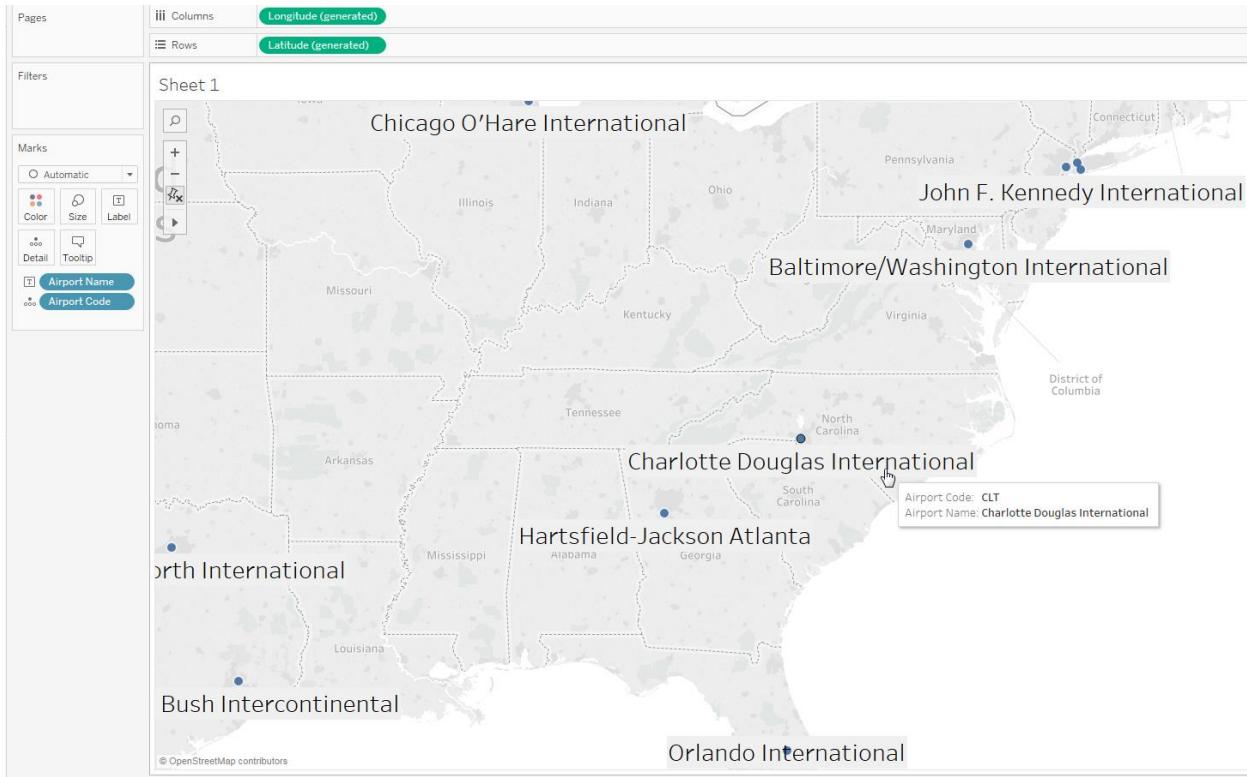
Add Sales and YEAR(Order Date) to the view. SUM(Sales) is lowest in 2013.



8. Open the Flights worksheet of the TDS Sample - Stock Prices & Flight Data workbook. Assign the geographic role for Airport Code as airport. Add the airport code to the view and drag the Airport Name to the Label mark. Which airport shown on the map is closest to Hartsfield-Jackson Atlanta (ATL)?
 - a. Chattanooga Metropolitan Airport (CHA)
 - b. Charlotte Douglas International (CTL) [highlight]

- c. Orlando International (MCO)
- d. George Bush Intercontinental (IAH)

Answer: B - Charlotte Douglas International (CTL) is closest to Hartsfield-Jackson Atlanta



Exploring & Analyzing Data

9. In the Southern Region Superstore data, the Product Category, Product Subcategory, and Product Name might be combined using which of the following:

- a. Set
- b. Group
- c. Hierarchy**
- d. Parameter

Correct Answer: C - Hierarchy. Hierarchies allow you combine similar fields while Groups and Sets combine values within a field. Parameters allow the user to change a value, which can then be used in calculations, filters, and reference lines.

10. Which chart type will best visualize the relationship between two continuous measures?

- a. Bar chart
- b. Scatter Plot**
- c. Line Chart
- d. Stacked Bar

B – Scatter Plot. Scatter plots to visualize relationships between numerical variables. Bar charts to compare data across categories. Line charts compare numeric data over time. Stacked bar charts compare numeric data over two dimensions.

https://onlinehelp.tableau.com/current/pro/desktop/en-us/help.html#buildeexamples_scatter.html?Highlight=scatter

11. Which of the following is a table calculation?

- a. SUM
- b. TOTAL**
- c. AVG
- d. STDEV

B - Of these, only TOTAL is a table calculation:

Table Calculation

TOTAL(expression)

Returns the total for the given expression.

Example: TOTAL(AVG([Profit])) = AVG(Profit) for all database rows that contribute to the partition

12. You would like to create a view showing the top 5 products by sum of sales for a given state or states. A set is used to select the top 5 products by sales. What type of filter should be used with state?

- a. Multiselect Filter
- b. Dimension Filter
- c. Context Filter**
- d. Wildcard Filter

You would like to filter on the state, and then select the top 5 products with the greatest sales from within the selected state. Therefore, you need a filter which will execute before the Set. Dimension filters execute after sets, so they are the wrong choice. If you make the filter on the state a context filter, Tableau will filter on the state, and then the set will determine the top 5 products by sales within the selected state.

More detail here: https://onlinehelp.tableau.com/current/pro/desktop/en-us/order_of_operations.html#DimtoCtx

13. Bins can be created for which type of field?

- a. Dimension
- b. Discrete
- c. Continuous measure**
- d. Date

Answer is C. A bin can be created on a continuous measure or numeric dimension

https://onlinehelp.tableau.com/current/pro/desktop/en-us/calculations_bins.html

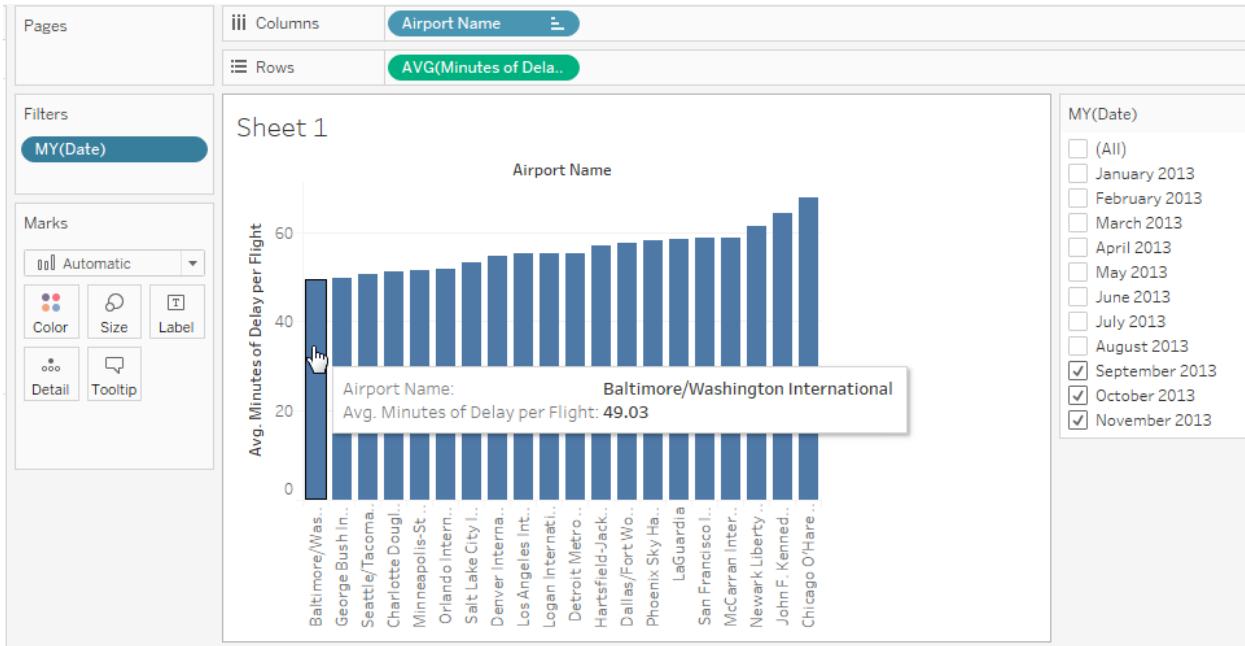
14. Using the Flights table, create a bar chart showing the average of Minutes of Delay per Flight broken down by Airport, and filtered by date to show only the last September, October, and November 2013. Which airport had the lowest average minutes of delay per flight between September and November?

- a. Baltimore / Washington International**

- b. George Bush Intercontinental
- c. Seattle/Tacoma International
- d. Chicago/O'Hare International

Answer: A - Baltimore / Washington International

Add a filter for September, October, and November 2013. Add Minutes of Delay to the view, and change the aggregation type to average. Add Airport Name to the view. Sort the chart to make it easier to see which airport has the lowest average minutes of delay.



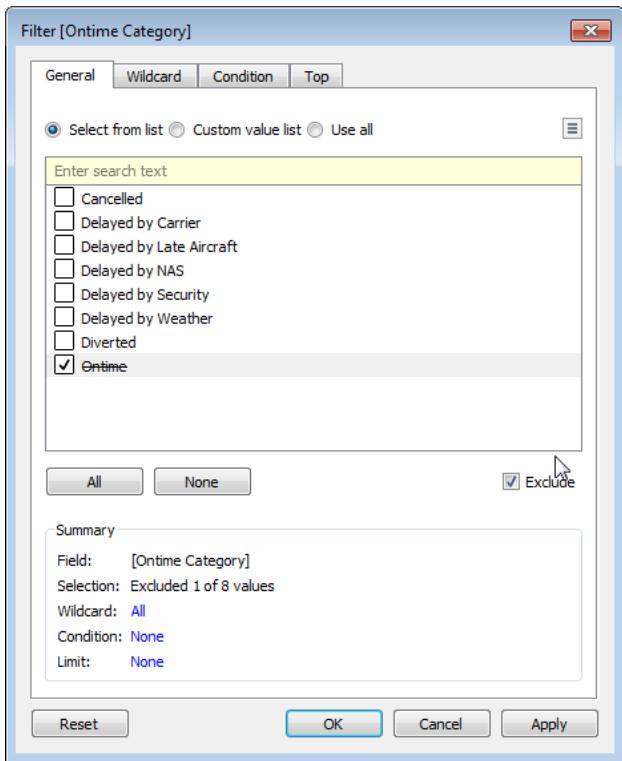
15. Using the Flights table, create a crosstab showing the sum of Flights broken out by ontime category and airport, then add grand totals to the view. Use a filter to exclude flights that were ontime. For Dallas/Fort Worth, what number of flights were not ontime, and what number were Delayed by Security?

- a. 54,692 and 53
- b. 1,001 and 650,751
- c. 72,489 and 51
- d. 45,677 and 93

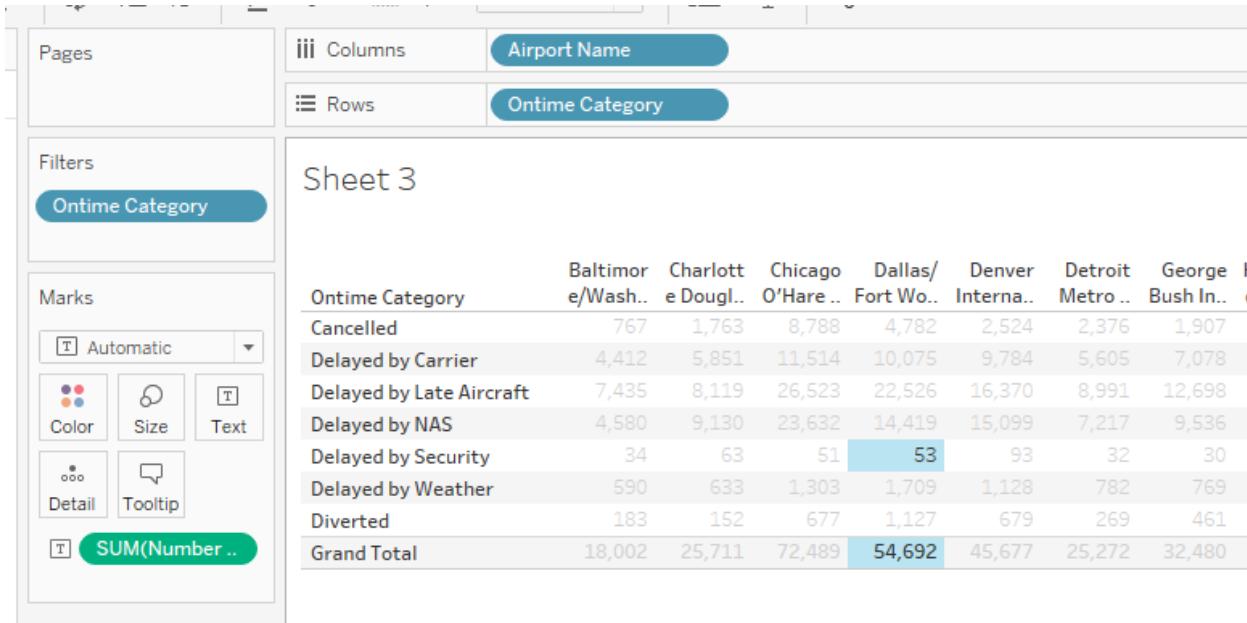
Correct Answer: A - 54,692 and 53

Create a crosstab with Airport Name and Ontime Category, showing SUM of Number of Flights. You should see Dallas/Fort Worth has 54,692 not ontime, and 53 flight delayed by security.

Exclude Ontime Flights:



Create a crosstab:



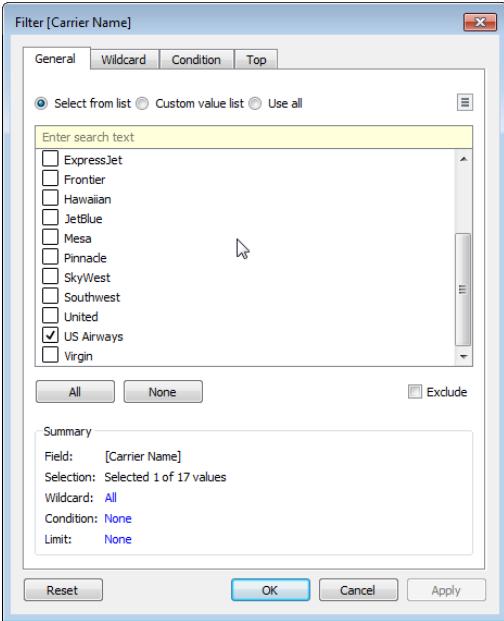
16. Using the Flights data, create a line chart showing the average minutes of delay per flight by month. Create a filter for carrier name of US Airways. Which month saw the biggest decrease in average minutes of delay per flight compared with the previous month?

- January 2013
- February 2013
- March 2013
- April 2013

Answer: C, March 2013

Add Month(Date) to the column shelf and and AVG(Minutes of Delay) to the row shelf:

Add a filter for US Airways:



Your view should now look like this:



17. Using the Stocks 2014 table, create a dual axis chart showing the daily Open and Close values. For Biogen Idec in October 2014, what was the day with the greatest difference between Open and Close?

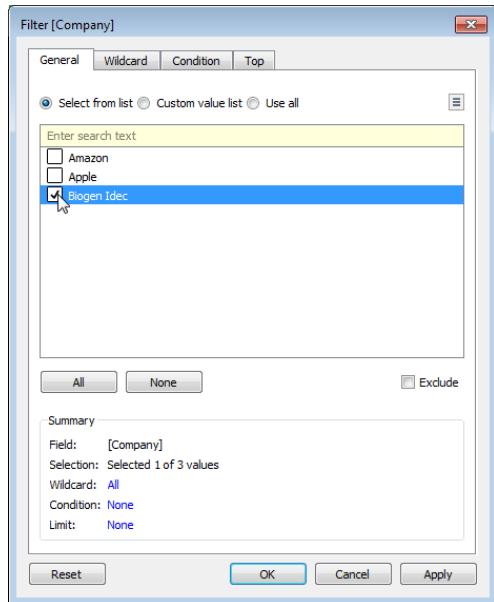
- a. October 12th
- b. October 15th

c. October 20th

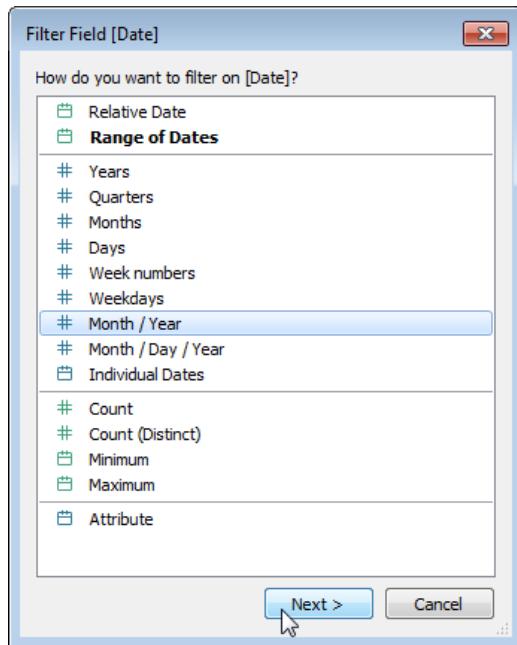
d. October 13nd

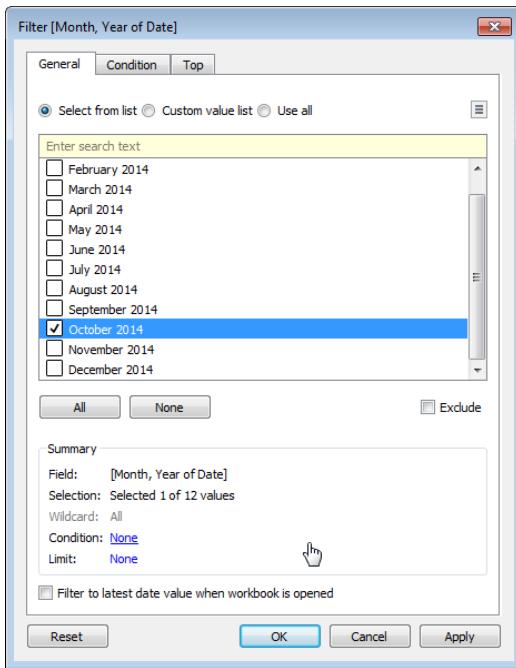
Correct Answer: D, October 13nd

Add company filter:



Add Month/Year filter:





Add a calculation for the difference between open and close:

The calculation is valid.

`abs([Open] - [Close])`

Open

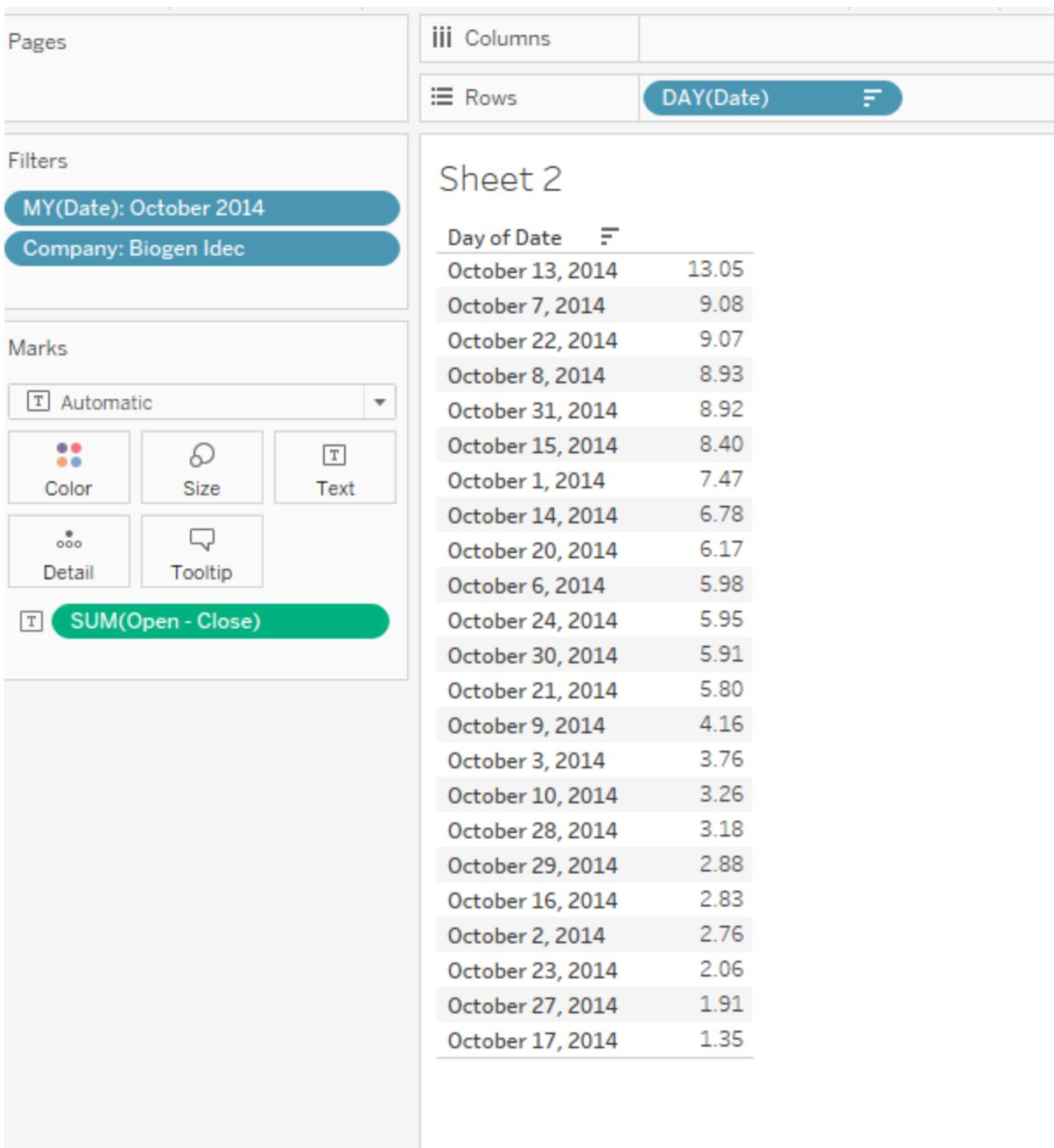
Enter search text

Data type: Float

- All
- ABS
- ACOS
- AND
- ASCII
- ASIN
- ATAN
- ATAN2
- ATTR
- Avg
- CASE
- CEILING
- CHAR
- COLLECT
- CONTAINS
- CORR
- COS

Number of records

Add date and the new calculation, then sort:



18. Using the Flights data, determine the average minutes of delay per flight by the month (using Date). Using a table calculation, find the % difference in average minute of delay per flight between from June 2013 compared with May 2013.

- a. 14.11%
- b. -8.2%
- c. 59.5%
- d. 60%

Answer: A – 14.11%

Determine the average minutes of delay per flight by the month:

Pages

iii Columns YEAR(Date) QUARTER(Date) MONTH(Date)

Rows

Sheet 2

Date
2013

Q1			Q2			Q3			Q4	
January	February	March	April	May	June	July	August	Septem..	October	Novem..
58.65	58.65	59.66	64.81	59.15	67.50	62.76	55.67	59.09	50.09	58.74

Marks

Automatic

Color Size Text

Detail Tooltip

AVG(Minutes of Delay per Flight)

The difference from May to June is $67.5 - 59.15 = 8.35$. Percent difference is $8.35 / 59.15 = 14.11\%$

Use a table calculation to determine the % change compared with the previous month:

Pages

iii Columns YEAR(Date) QUARTER(Date) MONTH(Date)

Rows

Sheet 2

Date
2013

Q1			Q2			Q3			Q4	
January	February	March	April	May	June	July	August	Septem..	October	Novem..
58.65	58.65	59.66	64.81	59.15	67.50	62.76	55.67	59.09	50.09	58.74

Marks

Automatic

Color Size Text

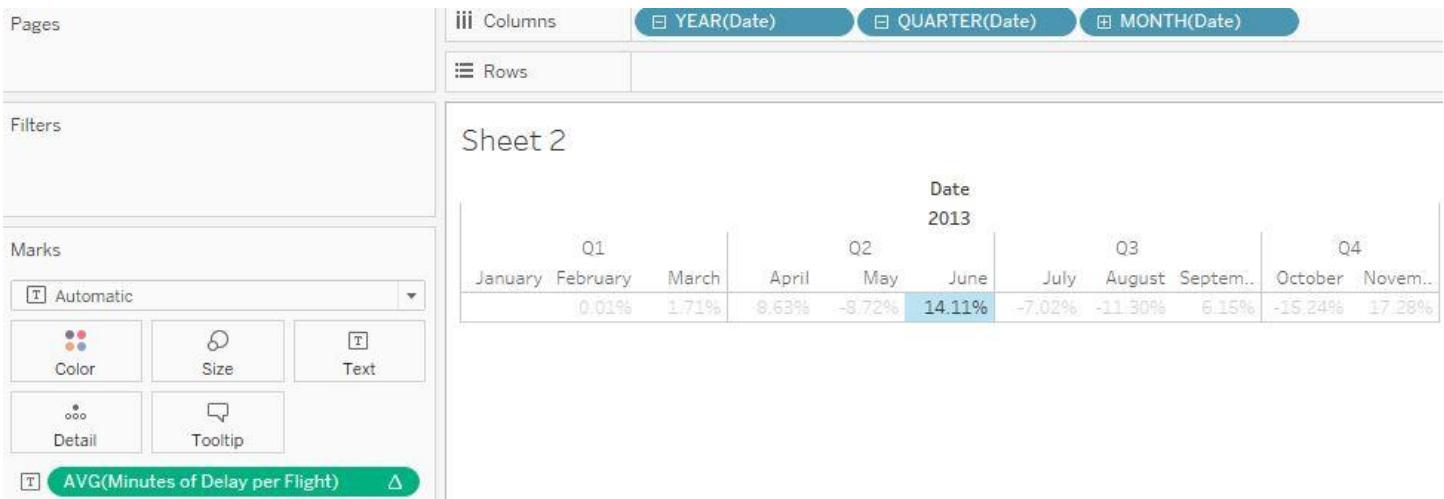
Detail Tooltip

AVG(Minutes of Delay per Flight)

- Filter...
- Show Filter
- Format...
- Include in Tooltip
- Dimension
- Attribute
- Measure (Average)
- Discrete
- Continuous
- Edit in Shelf
- Add Table Calculation...
- Quick Table Calculation
- Remove

Running Total
 Difference
Percent Difference
 Percent of Total

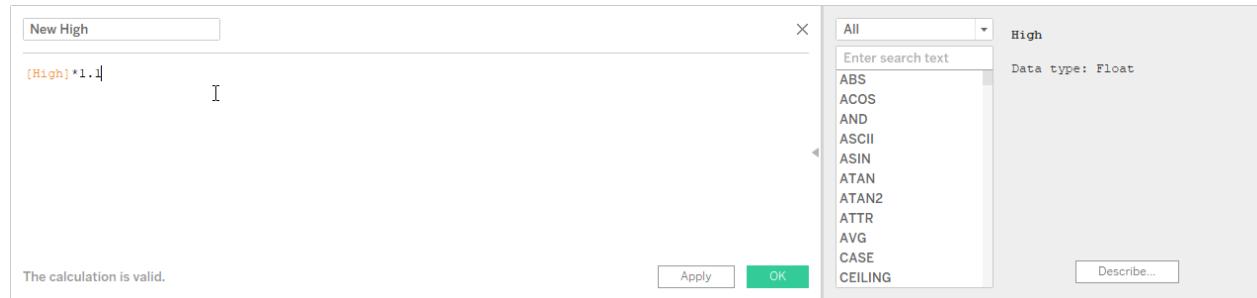
You should now be able to see the % change in average minutes of delay from May to June:



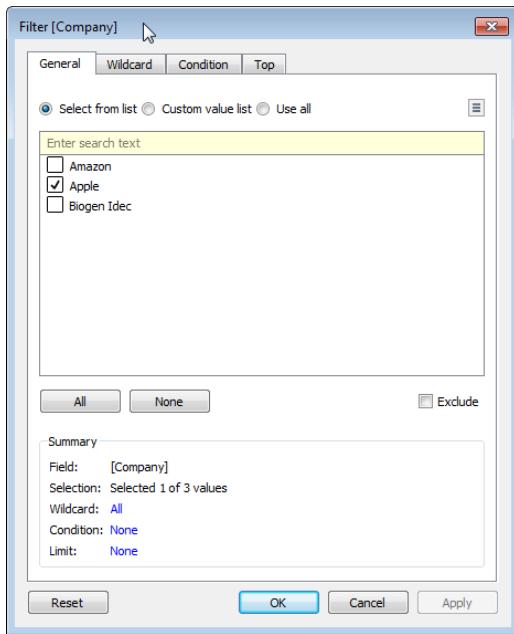
19. Analysts hope that Apple stock price will increase by 10% in 2015. Using the Stocks 2014 worksheet, create a calculated field called 2015 High that multiplies the value of High by 1.1. Use MAX aggregation to determine the maximum value of New High for Apple Stock.

- a. 131.7
- b. 23,439
- c. 119.8
- d. 448.9

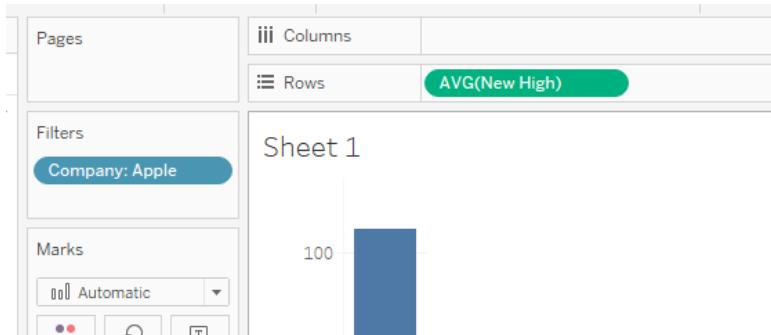
Create a calculated field called “New High”



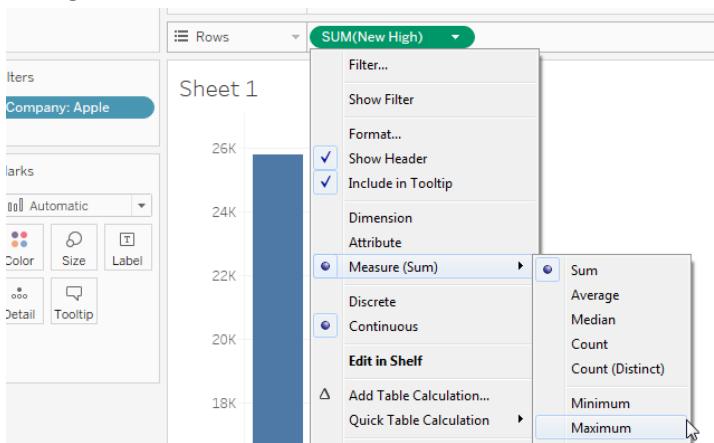
Add a filter for Apple:



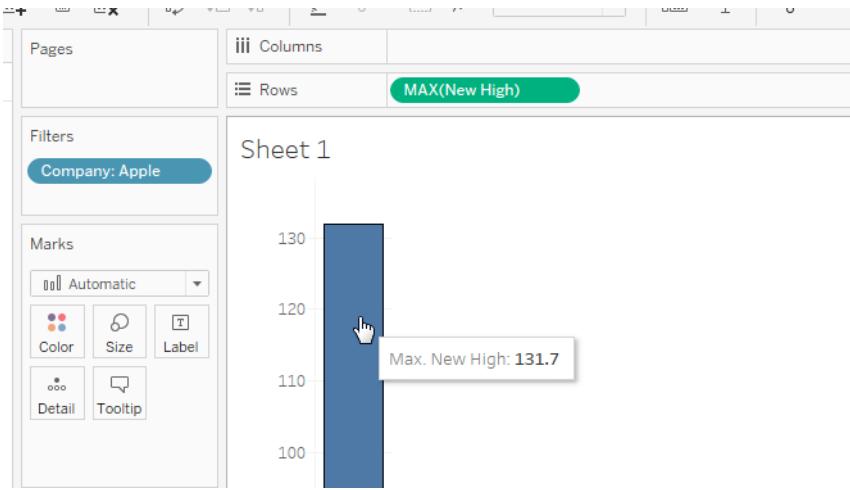
Add New High to the view:



Change the calculation from SUM to MAX:



View the MAX value of New High:



Sharing Insights

20. The most important items on a dashboard should be positioned where?

- a. Center
- b. Upper Left**
- c. Upper Right
- d. Bottom

Answer: B – Upper Left. Place the most important view at the top of your dashboard, or in the upper left corner. When looking at a dashboard, your eye is usually drawn to that corner first.

21. You have a dashboard and you want to add a feature so that selecting an element on one of the worksheets on the dashboard causes a change on another worksheet in the same dashboard. What features will do this? (select all that apply)

- a. Filter Action**
- b. Url Action
- c. Highlight Action**
- d. Drilldown Action

Answer: A & C. A filter action can filter items on another worksheet, and a highlight action can highlight items on another worksheet.

22. What steps should you take to add a phone view to a dashboard?

- a. Instruct those viewing the dashboard on mobile to put "m." in front of the dashboard URL to get the mobile-friendly view.
- b. Select Layout, then adjust the dimensions until they are appropriate for a phone (750x1334)
- c. On the "Show Me" menu, click the Phone icon
- d. Select Device Preview, then select Phone for Device Type, then click the button labeled "Add Phone Layout"**

Answer: D. You create a phone view by following the steps listed.

https://onlinehelp.tableau.com/current/pro/desktop/en-us/dashboards_dsd_create.html

Understanding Tableau Concepts

23. You start with a blank worksheet and add a continuous field to the row shelf. How will this new field change your view?

- a. It will add a horizontal axis
- b. It will add a horizontal header
- c. It will add a vertical axis
- d. It will add a vertical header

Answer: C - vertical axis

https://onlinehelp.tableau.com/current/pro/desktop/en-us/help.html#datafields_typesandroles.html?Highlight=discrete

24. You have a scatter plot showing sales on the x-axis and profit on the y-axis, with each dot representing a different product subcategory. You want to show how the points on the scatter plot vary based on Net Profit, a third continuous variable. Which would be most appropriate? (Select all that apply)

- a. Tooltip
- b. Label
- c. Size
- d. Color

Answer C & D: Size would be most appropriate, because with size it is clear that larger values are represented by larger points on the plot. This is easier for the eye to process than understanding which color or size is associated with the larger values than it is to read a label or tooltip value on each point in the scatter plot. The [Visual Analysis Best Practices Guidebook](#) explains, “a rule of thumb is to put the most important data on the X- or Y- axis and less important data on color, size, or shape.”

25. If you have bar chart with a single bar, but you want to subdivide that bar into multiple categories, what should you add to your view?

- a. A measure
- b. A dimension
- c. A discrete field
- d. A continuous field

The answer is B – you need to add a dimension to your view to subdivide the bar.

26. Which of the following fields will have AGG in front of it when you add it to the view?

- a. A calculated field with the formula SUM([Profit])
- b. A measure called Profit
- c. A date field
- d. A calculated field with the formula 5 * [Profit]

Answer: A, since this is a measure that has an aggregation already built in, when you add it to the view you will see AGG in front of the field name

27. A histogram shows the distribution of _____ data by creating bins that are _____.

- a. numeric, continuous
- b. discrete, continuous

c. continuous, discrete

d. date, discrete

Correct answer is C. A histogram uses bins to subdivide a continuous measure into discrete bins.

28. Which of the following functions can be used on string to create a measure?

a. SUM

b. COUNTD

c. +

d. DATEDIFF

Answer B. Of these functions, only COUNTD will take a string and return a measure.

29. You have SUM([Minutes of Delay]) in your view. If aggregate measures is unchecked, what will the view show instead of SUM([Minutes of Delay])?

a. Minutes of Delay

b. ATTR([Minutes of Delay])

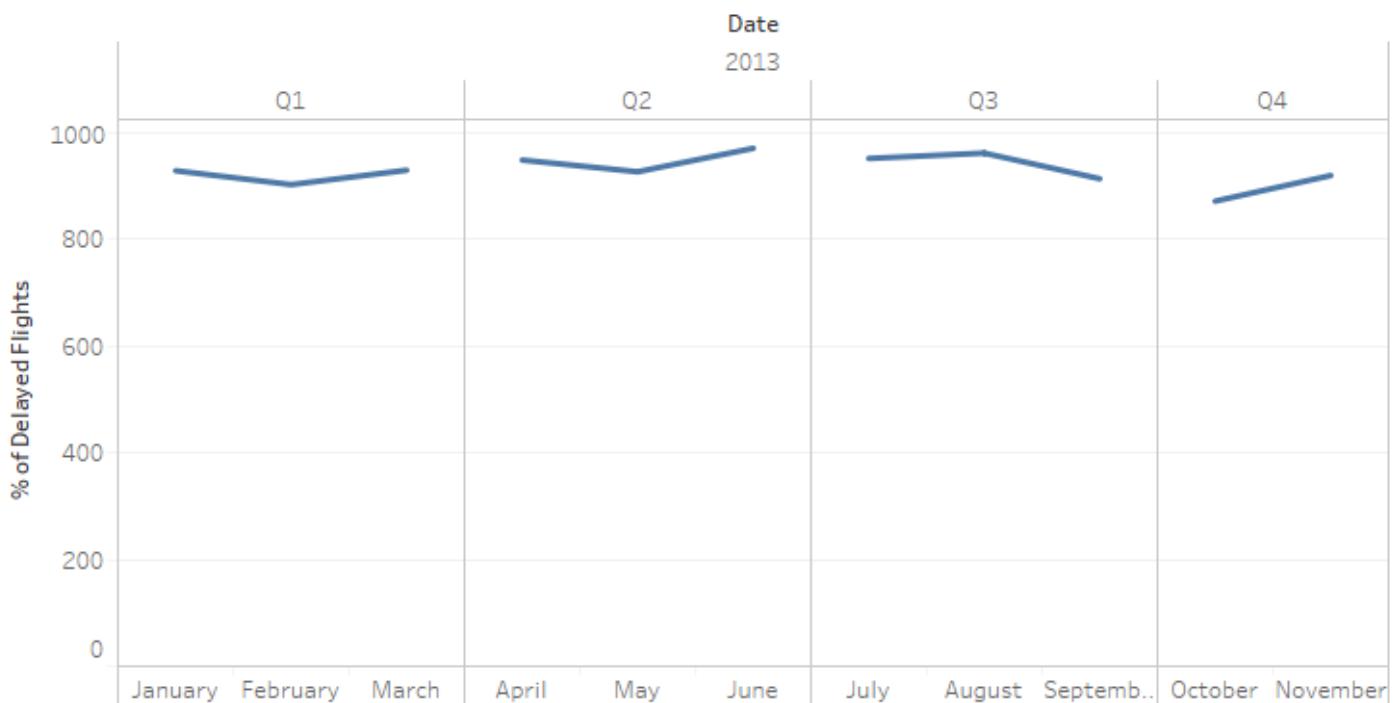
c. AGG([Minutes of Delay])

d. Minutes of Delay (bin)

A – disaggregating will show the underlying values. The aggregation will be removed, so we will see just “Minutes of Delay” https://onlinehelp.tableau.com/current/pro/desktop/en-us/help.html#calculations_aggregation.html#DisAggData

30. What field types are used in this view?

Sheet 6



A. Discrete for the Date, continuous measure for % of Delayed Flights

B. Continuous for the x and y axis

C. String for the x-axis, continuous measure for the y-axis

D. String for the x-axis, continuous measure for the y-axis

- A. The date is discrete, because otherwise the line would not be segmented
% of Delayed Flights is used as an axis, so it must be a continuous measure.