

**La Guardia Community College**

# **DATA 203 DATA VISUALIZATION USING TABLEAU**

## **Class 4**

# FILTERS

- **What is a Filter?**
  - The process of removing certain values or range of values from a result set
- **Why use a Filter?**
  - Minimize the size of the data for efficiency purposes
  - Clean up underlying data
  - Remove irrelevant dimension members
  - Set measure or date ranges for what you want to analyze.

# FILTERS

## Types of Filter?

There are two major types of Filter

1. Granular Filter
  - Commonly used
  - Dimensions, Measures
2. Macro Filters:
  - Occasionally used
  - applied at the data source level

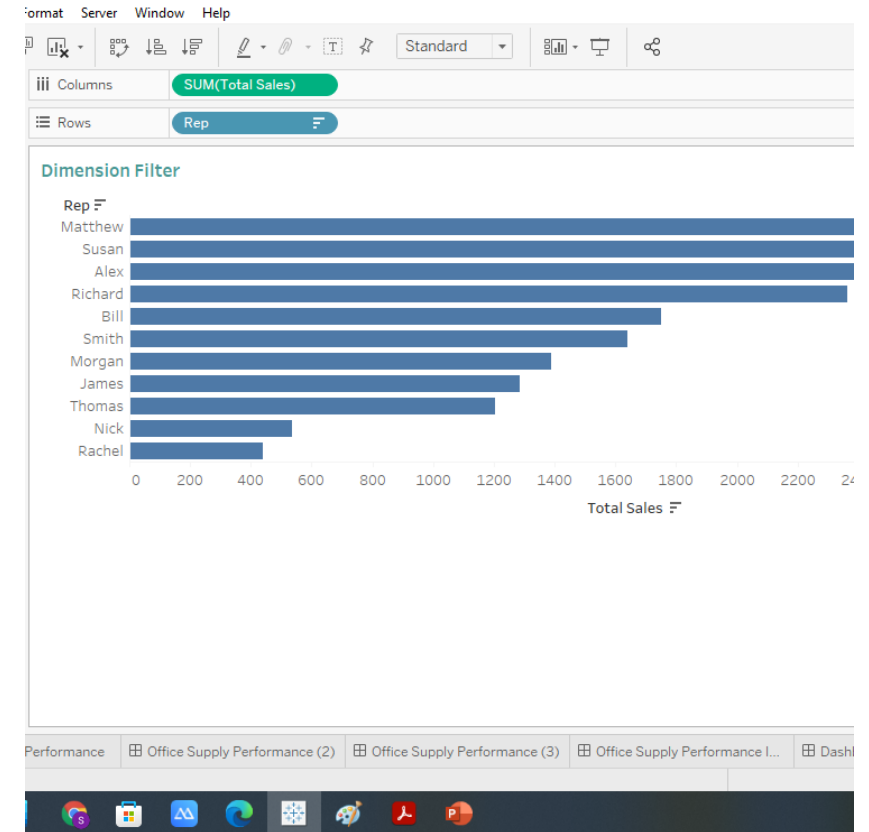
# FILTERS - Granular Filter - Dimension

## DIMENSIONS

Filters applied to Dimension Fields

### 1. CREATE A DIMENSION FILTER

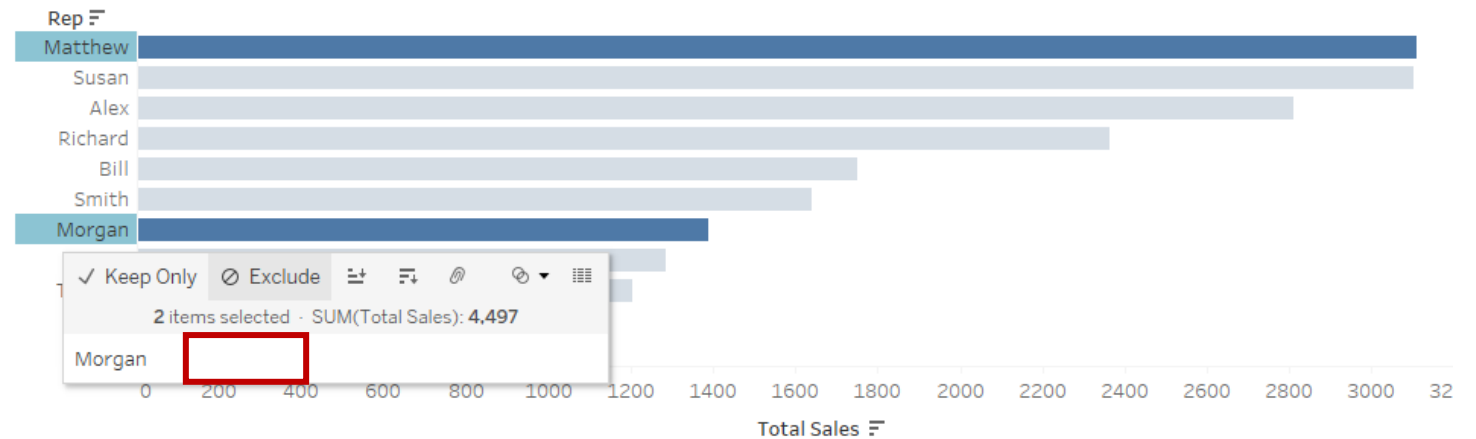
1. Open a new worksheet and name it Dimension Filter
2. Create a horizontal bar chart showing Total Sales by Rep and Sort Descending on Total Sales.



# FILTERS - Granular Filter - Dimension

- CREATE A BASIC FILTER
  - The most basic way to filter out data
    - Keep or Exclude
- Create an Exclusion Criteria to Filter
  - We want to exclude all Reps whose Names start with "M"
  - Multiselect Reps name starting with "M" (on PC, you can CTRL Click)
  - Select Exclude in the pop-up window
  - You see Keep Only, if you were creating a Keep Only Filter

Dimension Filter

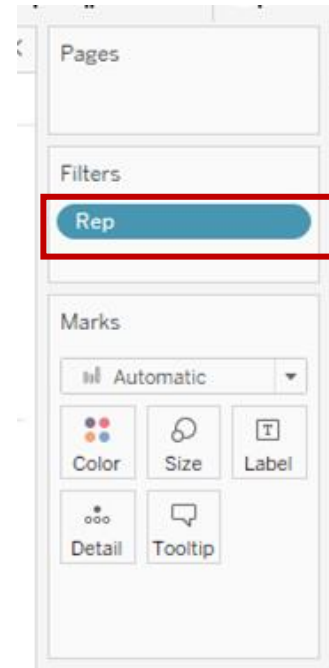


# FILTERS - Granular Filter - Dimension

---

Look at the FILTER SHELF

- Rep dimension is added to the filter shelf.
- Tableau created a Rep Filter when you excluded Rep dimension members from the view

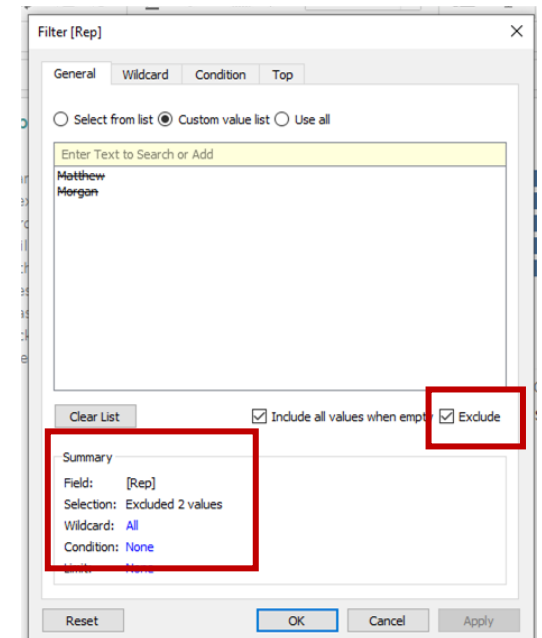


# FILTERS - Granular Filter - Dimension

---

How to Manually setup a DIMENSION FILTER

1. Right click on the Rep Filter and select Edit Filter
2. Tableau created an Exclusion Filter
3. When we chose to exclude Reps whose name started with the letter “M” from the view, as indicated by the box for Exclude being checked.
4. The Summary Box at the bottom tells you the criteria for the Rep Filter



# FILTERS - Granular Filter - Dimension

---

Click on the Wildcard tab

1. Instead of manually selecting all names that start with the letter “M”, we will leverage the “Does not start with” function
2. Check the box for Exclude.
3. Select “Does not start with”
4. Type “M” in the Match Value box.
5. Click OK

The screenshot shows the 'Filter [Rep]' dialog box with the 'Wildcard' tab selected. The 'Match value:' field contains 'M'. The 'Exclude' checkbox is checked. The 'Does not start with' radio button is selected. The 'Include all values when empty' checkbox is also checked. The 'OK' button is highlighted in blue.

Filter [Rep]

General Wildcard Condition Top

Match value: M

☒ Exclude

☐ Does not contain

☒ Does not start with

☐ Does not end with

☐ Does not match

☒ Include all values when empty

Clear

Reset OK Cancel Apply



# FILTERS -

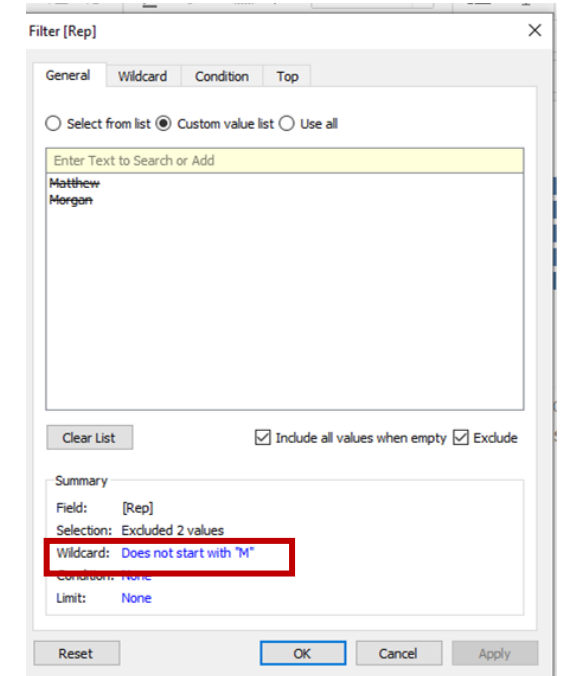
## Granular Filter -

## Dimension

---

Click on the Wildcard tab

1. Navigate back to the Summary Card on the General tab.
2. All the Rules for the filter are stated
3. Each Rule acts as AND statement
4. Dimension members must meet all criteria to be included or excluded from the view.



# **FILTERS - Granular Filter - Dimension**

---

## **Condition & Top tabs**

### **1. Condition**

1. Add Quantitative thresholds that must be met

### **2. Top**

1. Focus only on the Top or Bottom dimension members

## **HANDS ON**

- **CREATE A CONDITION FILTER WHERE THE SALES REP TOTAL SALES IS GREATER THAN \$1000.**

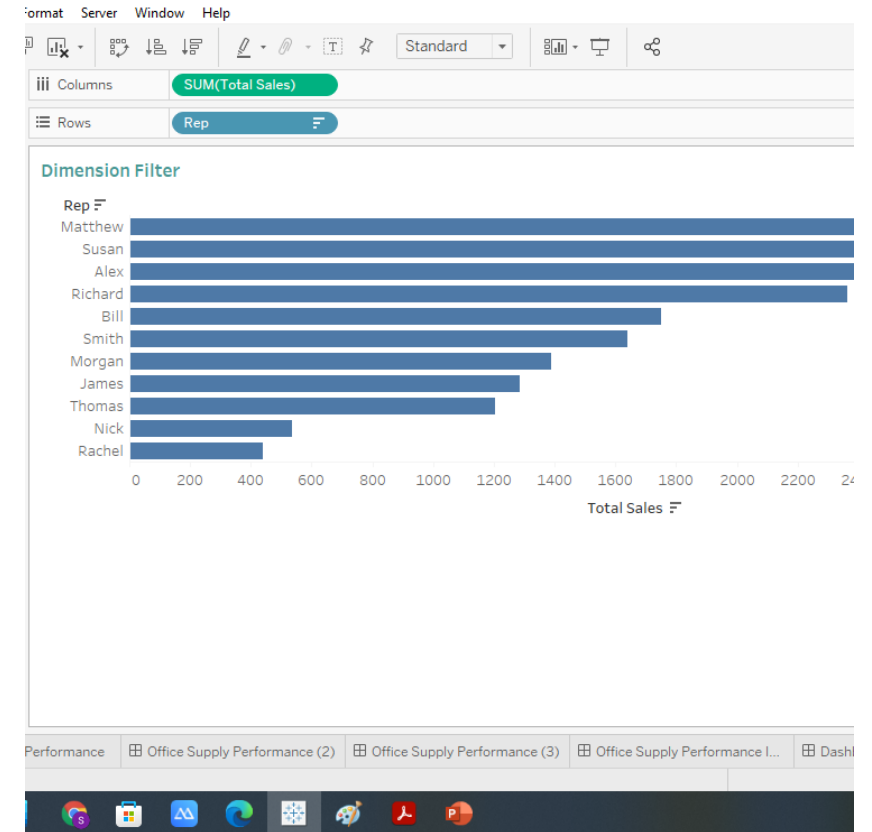
# FILTERS - Granular Filter - Measures

## 2. MEASURES

Filters applied to Measures Fields

### CREATE A MEASURES FILTER

1. Duplicate the Dimension Filter sheet.
2. Rename it Measures Filter.



# FILTERS - Granular Filter - Measures

## CREATE A TOTAL SALES FILTER

1. Drag Total Sales from the Data Pane to the Filter Shelf
2. You need to choose the **aggregation of the measure** before choosing your measure filter.
3. Choose SUM

Filter Field [Total Sales] ✕

How do you want to filter on [Total Sales]?

<input checked="" type="radio"/>	All values
<input type="radio"/>	Sum
<input type="radio"/>	Average
<input type="radio"/>	Median
<input type="radio"/>	Count
<input type="radio"/>	Count (Distinct)
<input type="radio"/>	Minimum
<input type="radio"/>	Maximum
<input type="radio"/>	Standard deviation
<input type="radio"/>	Standard deviation (Population)
<input type="radio"/>	Variance
<input type="radio"/>	Variance (Population)
<input type="radio"/>	Attribute

# FILTERS - Granular Filter - Measures

## CREATE A TOTAL SALES FILTER

- After you choose an aggregation, select
  - Range of values
  - A minimum threshold
  - A maximum threshold
- Range of values is the best option
  - Provides more flexibility
  - Best choice if you want to control the top end and bottom end of the measure.
  - Set the bottom level value to 1000

Filter [Total Sales]

Range of values | At least | At most | Special

Range of values

1000 | 3,109.44

438.37 | 3,109.44

Show: Only Relevant Values | Include Null Values

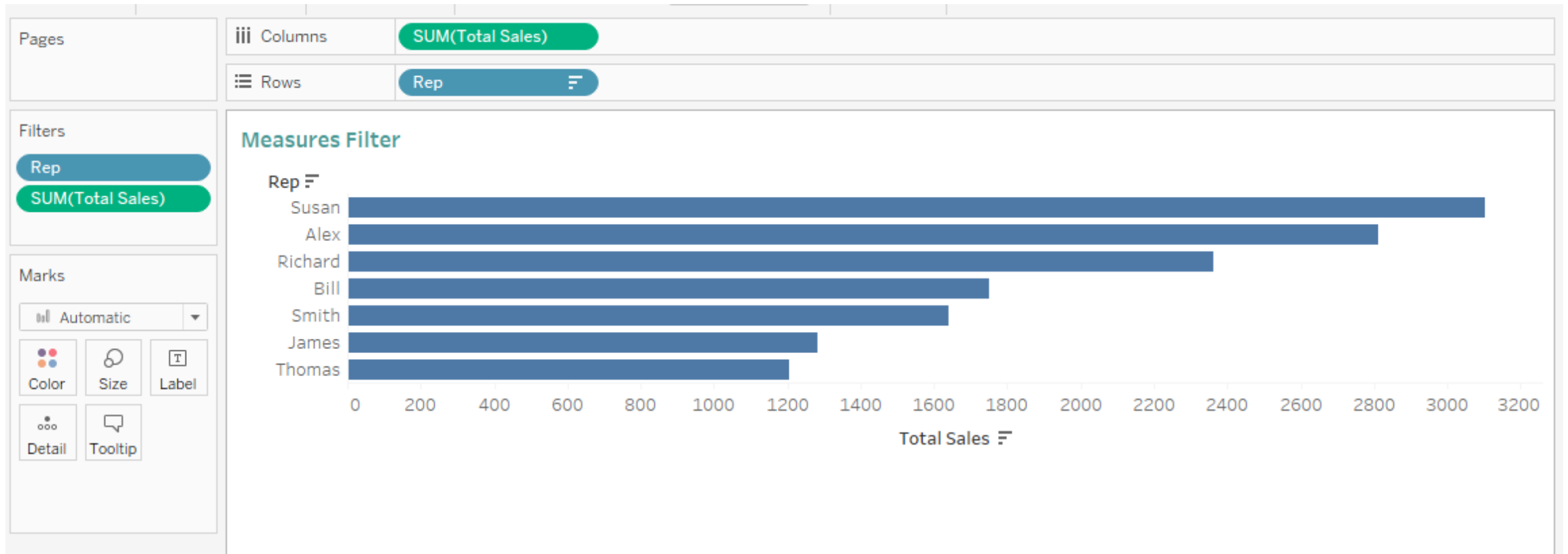
Reset | OK | Cancel | Apply

# FILTERS - Granular Filter

You have added a Dimension and Measures Filter to one chart.

## HANDS ON

Create one Dimension and One Measures filter on a field that is not part of the view.



# CALCULATIONS

- **What are Calculated Fields?**
  - Most **powerful tool** in Tableau
  - Allows you to create new data from existing data
  - Create new Dimensions such as Segments and new Measure such as ratios.
  - Used with any data type, a multitude of functions and aggregations, and logical operators.
- **ARE LIMITLESS**

# CALCULATIONS

- **Why use Calculated Fields ?**
  - To segment your data in new ways on the fly
  - To prove a concept such as a new dimension or measure
  - To filter out unwanted results for better analyses
  - To take advantage of the power of parameters
  - To calculate ratios across many different variables in Tableau.



# CALCULATIONS

- **Types of Calculations**

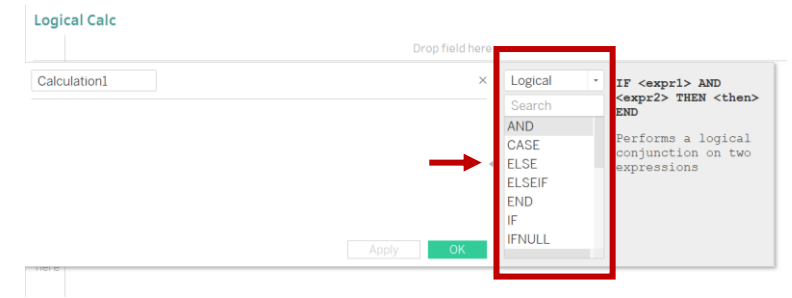
- Logical - test whether a situation is True or False
- Numeric – takes a value in a set of numbers
- Date – returns either a date or date part as a result
- String – manipulate strings or query information about a string
- Type Conversion – change data types to force a result for the desired data type
- Aggregate – performs calculation on a set of values
- User

# CALCULATIONS

---

## Logical

- Test whether a situation is True or False
- Open a new worksheet and name it Logical Calc
- Open a Calculated Field Window
- Click on the little arrow on the RHS
- Select Logical from the dropdown
- A list of all Logical functions.
- We will focus on IF/ELSE function



# CALCULATIONS

## Logical What is an “IF/ELSE” statement?

1. They are a series of IF Statements.
2. The IF Statement get evaluated in order until one of the expression is true or the END of the IF/ELSE statement is reached
3. If the end of the IF/ELSE statement is reached without a true expression, then that code block is not executed.

1. When you have more than 2 IF Statements

```
IF X = 1 THEN ACTION A ELSE  
IF X = 2 THEN ACTION B ELSE  
IF X = 3 THEN ACTION C  
END
```

2. When you have only 2 IF Statements

```
IF X = 1 THEN ACTION A ELSE  
ACTION B  
END
```

# CALCULATIONS

## Logical

### What is an "IF/ELSE" statement?

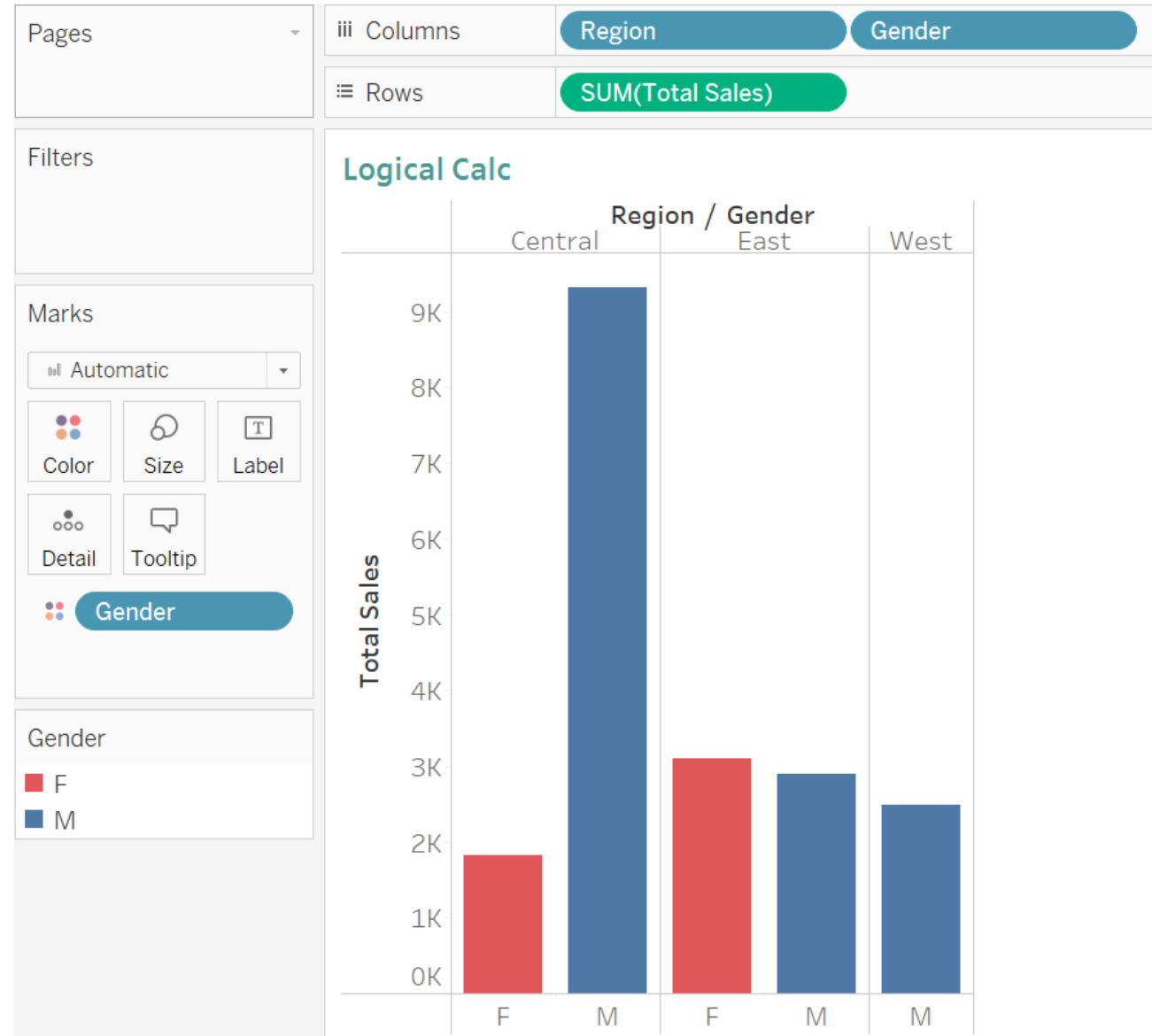
1. Name your calculation Gender  
IF [Rep] = 'Alex' then "M"  
ELSEIF [Rep] = 'Bill' then "M"  
ELSEIF [Rep] = 'James' then 'M'  
ELSEIF [Rep] = 'Matthew' then 'M'  
ELSEIF [Rep] = 'Nick' then 'M'  
ELSEIF [Rep] = 'Smith' then 'M'  
ELSEIF [Rep] = 'Richard' then 'M'  
ELSEIF [Rep] = 'Thomas' then 'M'  
ELSEIF [Rep] = 'Rachel' then 'F'  
ELSEIF [Rep] = 'Morgan' then 'F'  
ELSEIF [Rep] = 'Susan' then 'F'  
END

```
IF [Rep] = 'Alex' then "M"
ELSEIF [Rep] = 'Bill' then "M"
ELSEIF [Rep] = 'James' then 'M'
ELSEIF [Rep] = 'Matthew' then 'M'
ELSEIF [Rep] = 'Nick' then 'M'
ELSEIF [Rep] = 'Smith' then 'M'
ELSEIF [Rep] = 'Richard' then 'M'
ELSEIF [Rep] = 'Thomas' then 'M'
ELSEIF [Rep] = 'Rachel' then 'F'
ELSEIF [Rep] = 'Morgan' then 'F'
```

The calculation is valid.

# CALCULATIONS

- Logical
- What is an “IF/ELSE” statement?
1. Total Sales by Gender by Region
  2. Column – Add Region and Gender
  3. Row – Add Total Sales
  4. Color the bars by Gender



# CALCULATIONS

---

- **Numeric**

- Open a new worksheet and name it Numeric Calc
- Open a Calculated Field Window and name it Max Date
- Check all the Numeric functions
- We will focus on Max function

MAX( [Order Date])

Max Date

×

MAX( [Order Date])

The calculation is valid.

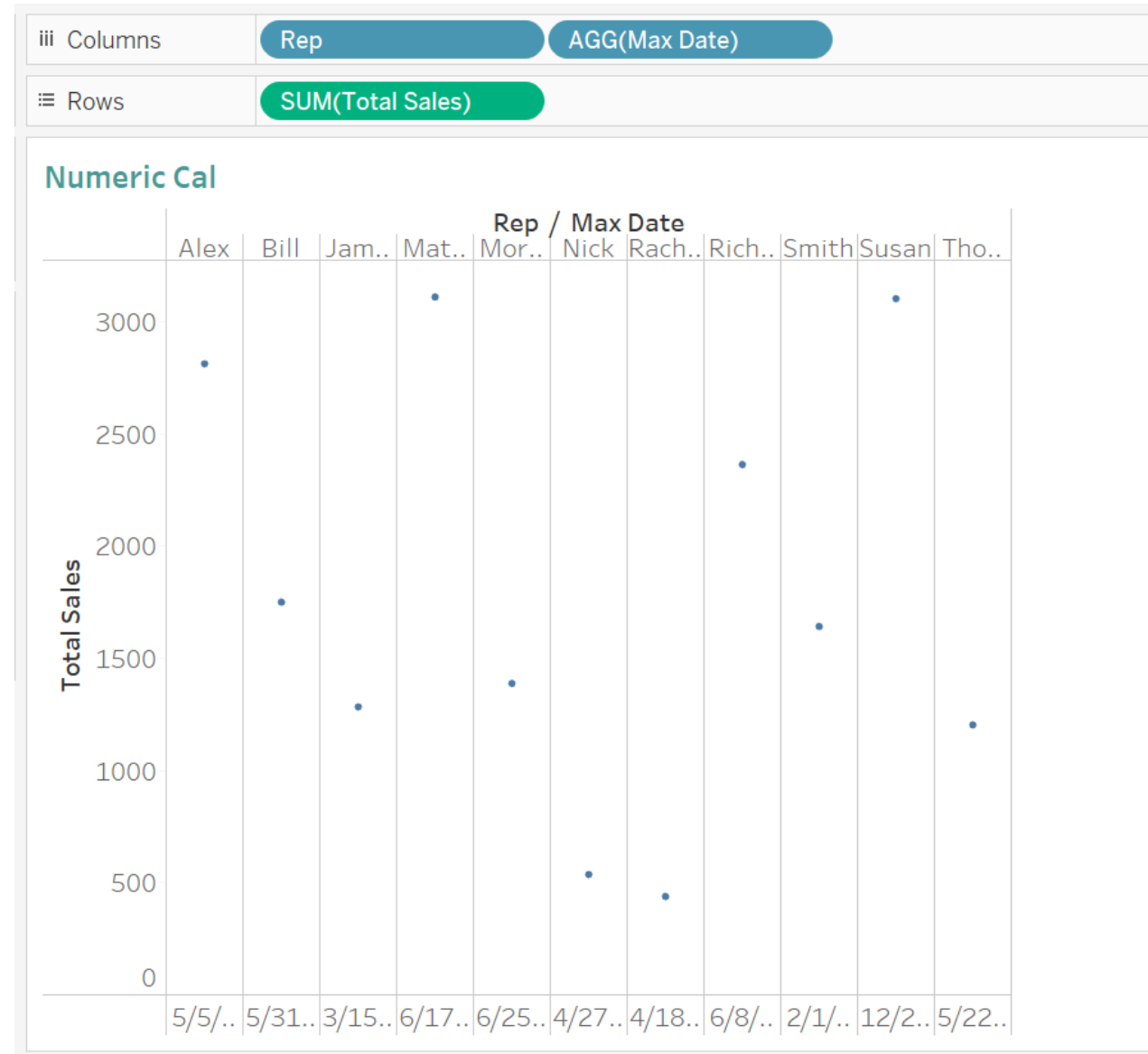
Apply

OK

# CALCULATIONS

- Name

1. What is the last Total Sales by each Rep
2. Column – Add Rep and MaxDate calculation
3. Row – Add Total Sales

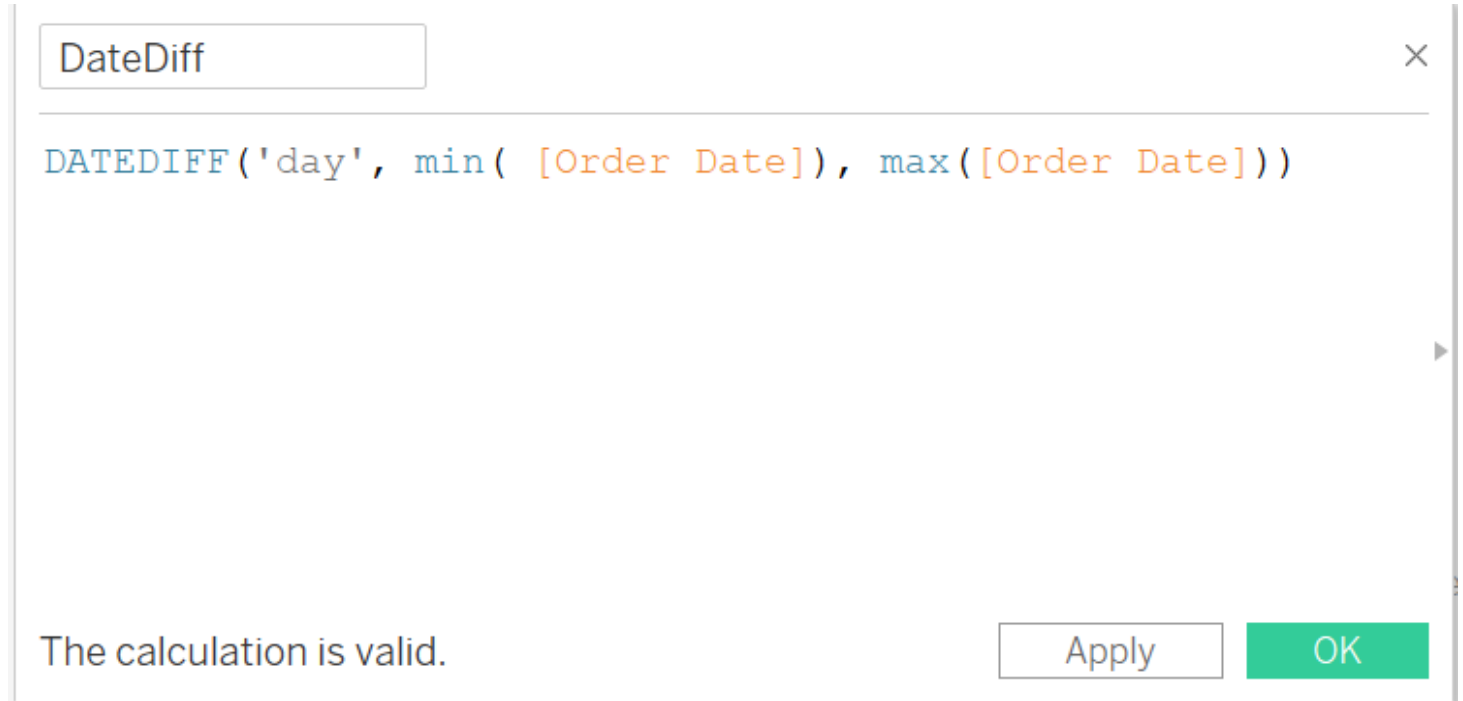


# CALCULATIONS

---

- **Date**
  - Open a new worksheet and name it Date Calc
  - Open a Calculated Field Window and name it DateDiff
  - Check all the Date functions
  - We will focus on DateDiff function
    - Takes the difference between two dates
    - We will take the difference between minimum and maximum Order Date

**DATEDIFF('day', min( [Order Date]), max([Order Date]))**

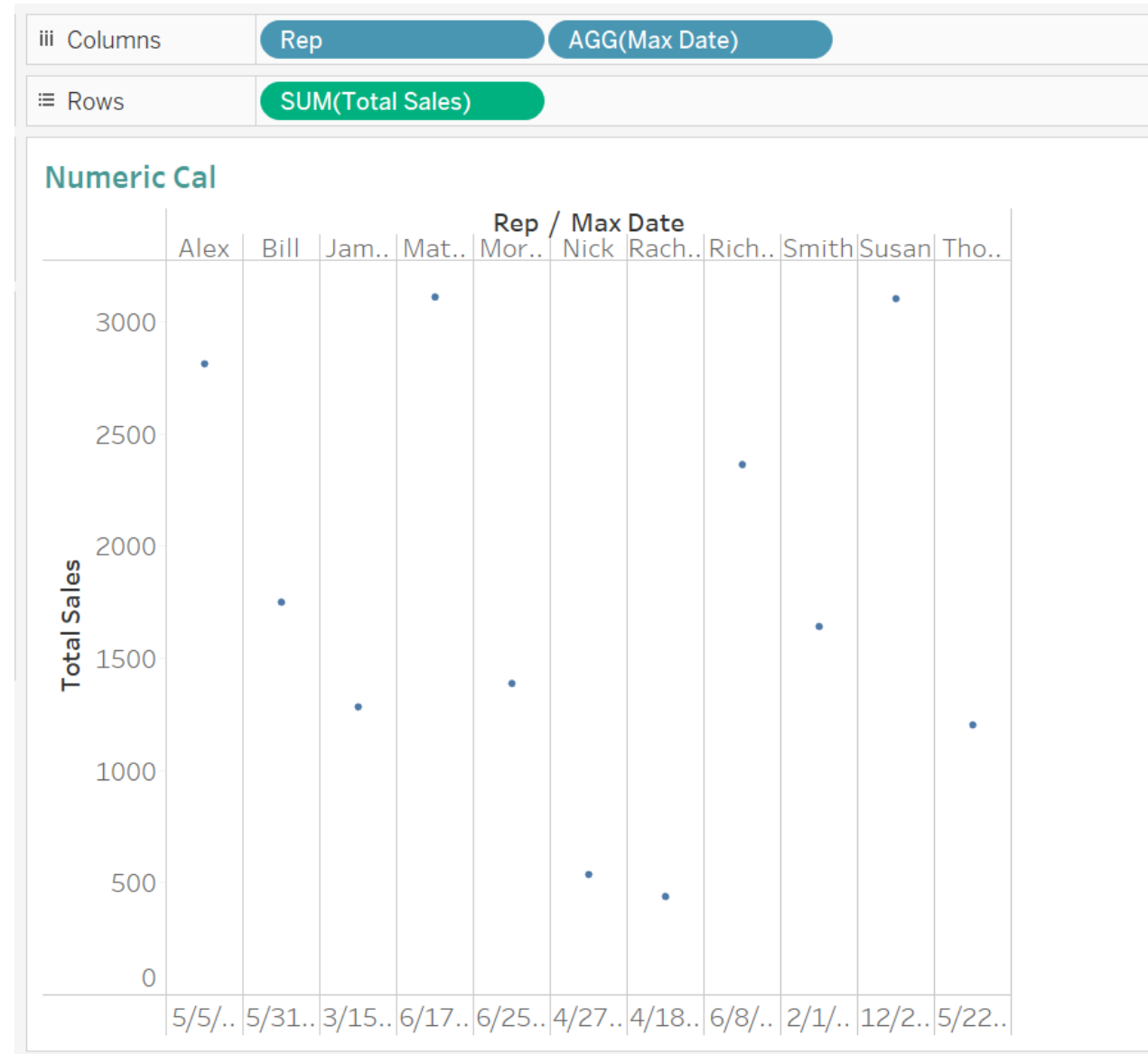




# CALCULATIONS

- Date

1. What is the last Total Sales by each Rep
2. Column – Add Rep and MaxDate calculation
3. Row – Add Total Sales



# CALCULATIONS

---

- **String**

- Open a new worksheet and name it String Calc
- Open a Calculated Field Window and name it GenderFull
- Check all the String functions
- We will focus on Replace function
  - Replaces any occurrence of the substring and replaces it with the replacement string characters.

**REPLACE**( [Gender], 'M', "Male")

GenderFull

×

REPLACE ( [Gender], 'M', "Male")|

The calculation is valid.

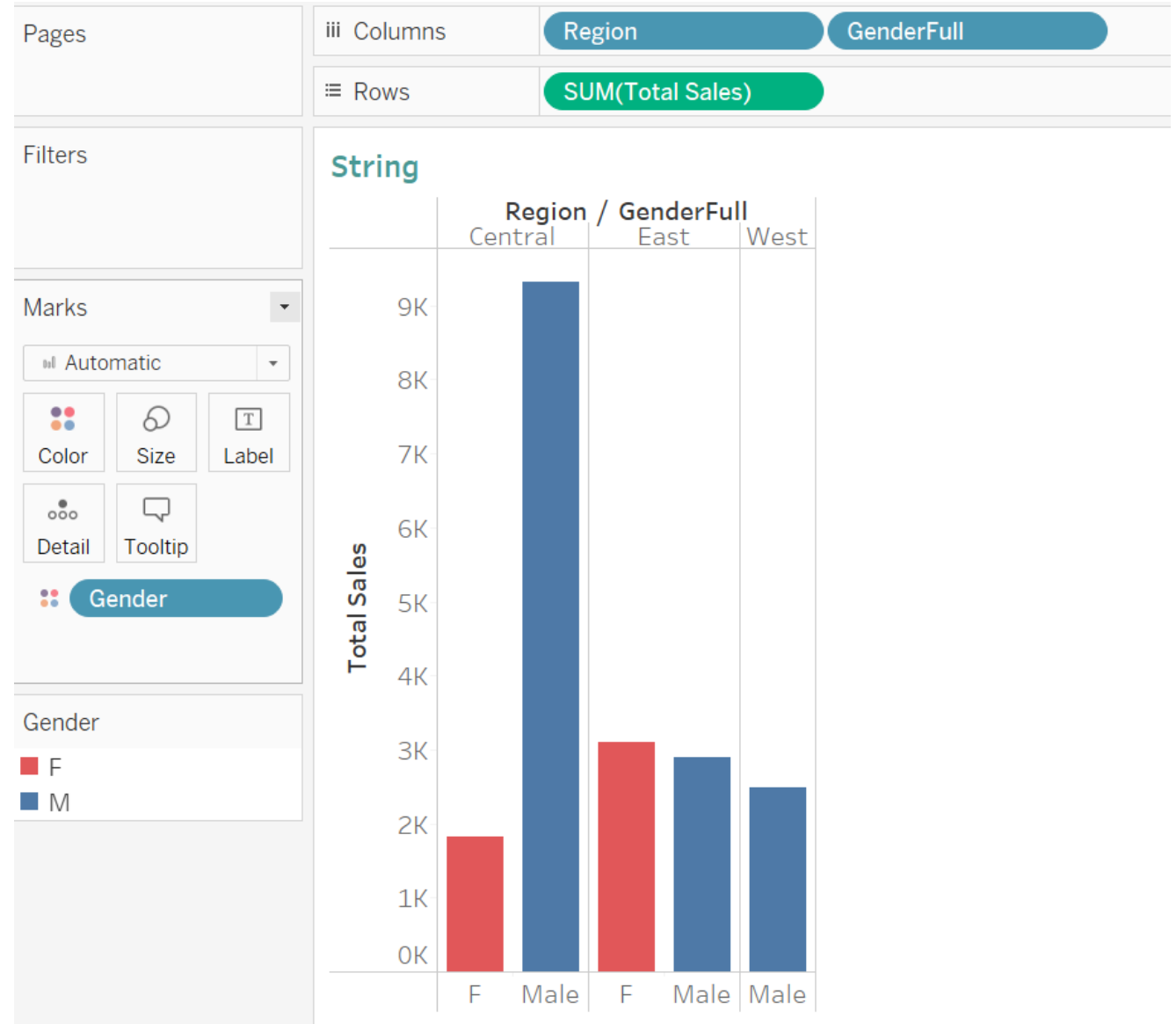
Apply

OK

# CALCULATIONS

- **String**

1. Make a duplicate of Logical Cal
2. Rename the worksheet String
3. Column – Replace Gender with GenderFull



# CALCULATIONS

---

- **Aggregate**

1. **Count of Total Types of Items sold**
  1. Open a new worksheet and name it **Aggregate**.
  2. Name the calculated field **Total Type of Item**
  3. We want to **count a distinct number** of Items offered and will focus on CountD

CountD ([Item] )

Total Item Type

CountD ([Item] )

The calculation is valid.

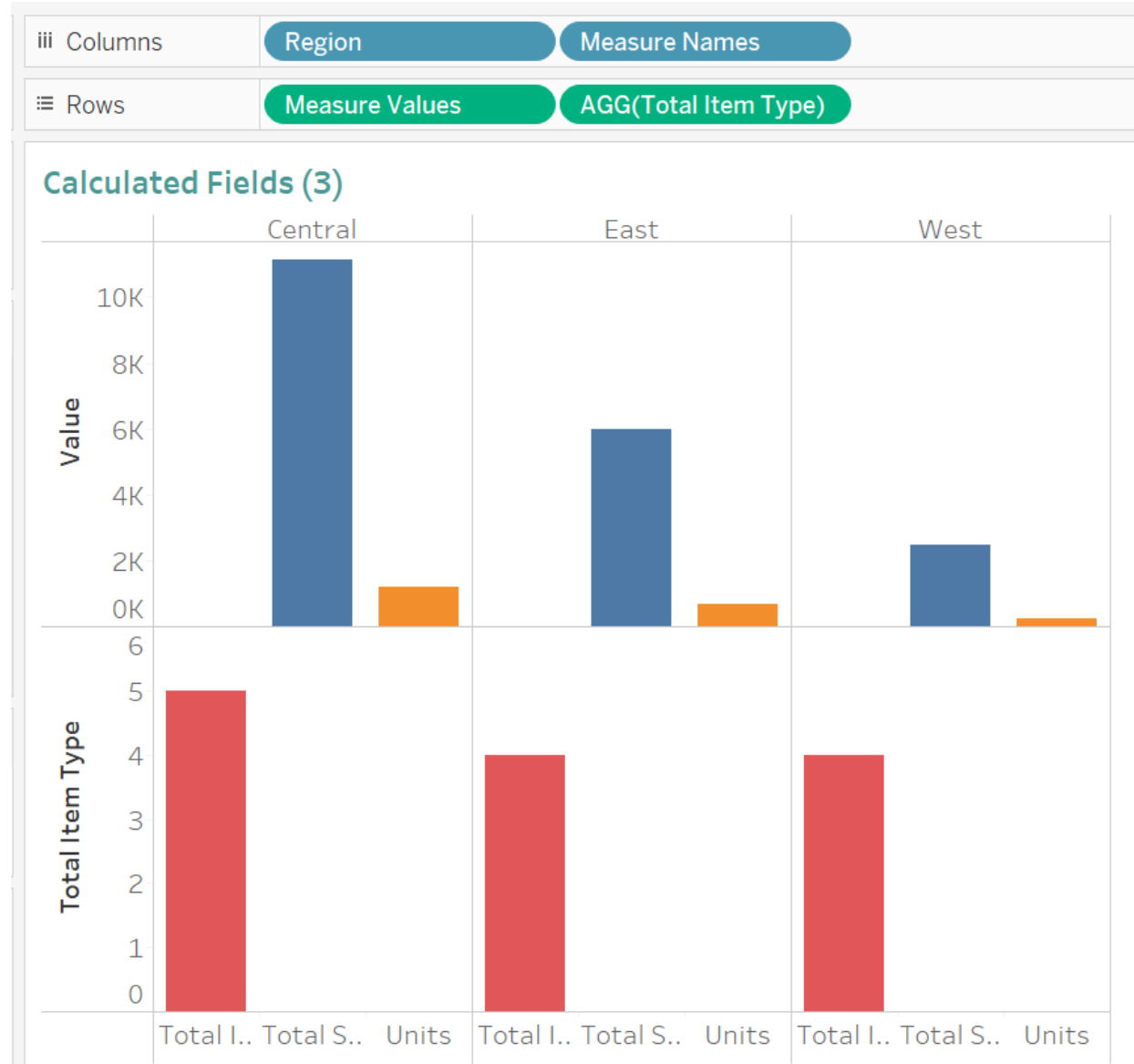
Apply

OK

# CALCULATIONS

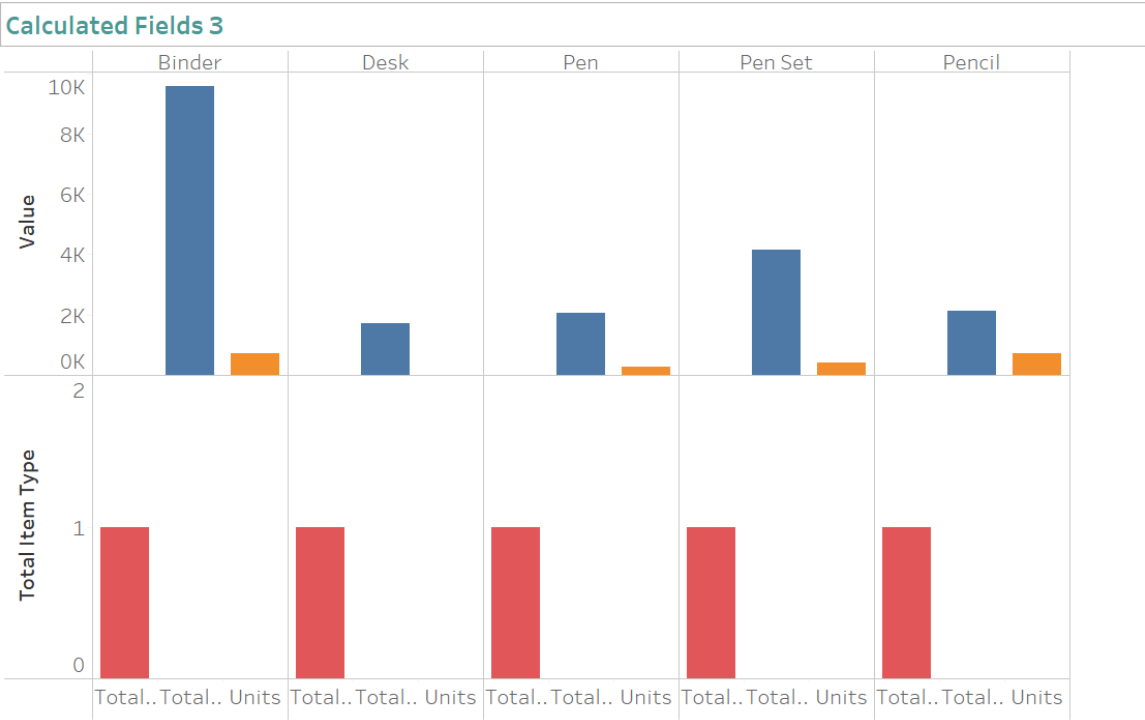
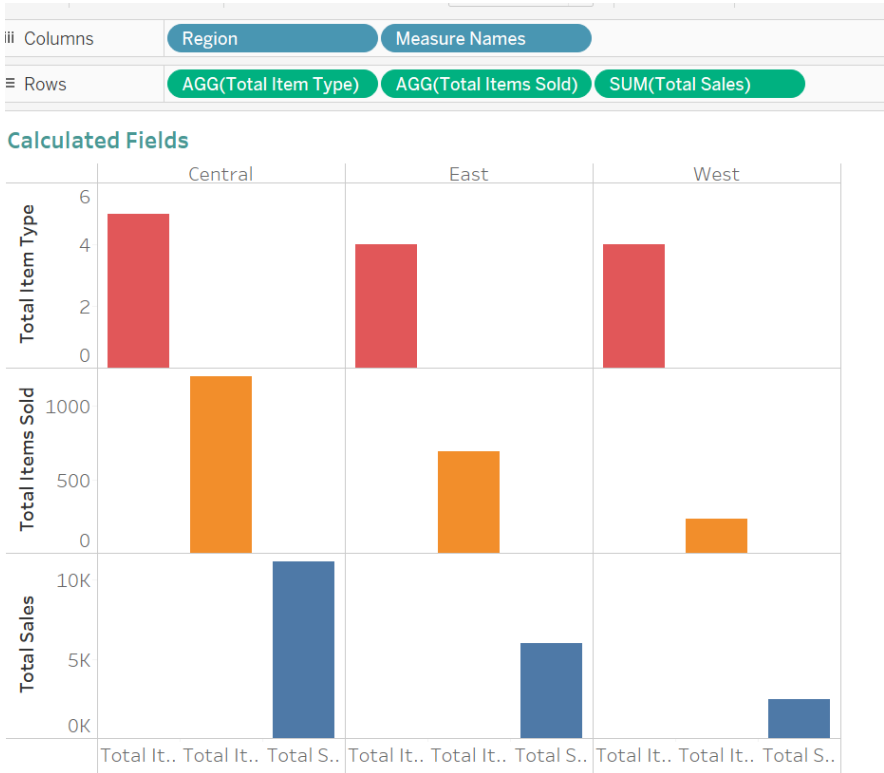
- **Aggregate**

1. Columns – Add Region and Measure Names
2. Rows - Add Measure Values, Total Item Types
3. Change color of Total Item Type and Total Units Sold



# CALCULATIONS

- Aggregate
- Couple of alternate views



# CLASS HANDS ON

## Use Data File - TFL Bus Safety

- Create one filter for a Dimension and one filter for a Measure
- Create a filter for a field that is not part of the view
- Create one calculation of each type that we covered.
- Incorporate your Filters and Calculated Fields in the worksheet that you created for TFL Bus Safety in Assignment .
- Incorporate filters and calculated fields in your final Dashboard.