Business Intelligence Intermediate – Unit 2

Xccelerate - Data Science Immersive

Agenda

TABLEAU:

- Data Join and Data Blending
- Creating bins, parameters, sets
- Data Preparation in Tableau
- Walkthrough Exercises
- Practice makes perfect (Part 2)

Walkthrough Exercises:

Data Joining

Agenda:

- Joining Data in Tableau
- Understanding different types of Join

Connect Tableau to the Excel file provided:
BI_Intermediate_Unit2_Data_Join.xlsx
BI_Intermediate_global_superstore_2016.xlsx

Data Joining:

JOIN TYPES WITH UNION

Joining tables is a way of combining information from multiple tables on a field they share.

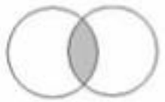
Joins are always made on a specific field (or fields)

Name	# of Siblings
Taylor	2
Alex	3
Shannon	0
Tracy	1

Name	Eye Color
Taylor	Blue
Alex	Brown
Morgan	Brown

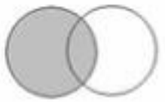
Inner Join

Name	# of Siblings	Eye Color
Taylor	2	Blue
Alex	3	Brown



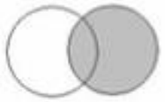
Left Join

Name	# of Siblings	Eye Color
Taylor	2	Blue
Alex	3	Brown
Shannon	0	*null*
Tracy	1	*null*



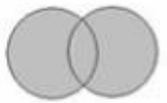
Right Join

Name	Eye Color	# of Siblings
Taylor	Blue	2
Alex	Brown	3
Morgan	Brown	*null*



Outer Join

Name	# of Siblings	Eye Color
Taylor	2	Blue
Alex	3	Brown
Shannon	0	*null*
Tracy	1	*null*
Morgan	*null*	Brown



Union follows the behavior “Union all” even if there are duplicate values for some rows.

Q1

Name	# of presentations
Kai	10
Piper	3
Lien	4

Q2

Name	# of presentations
Kai	10
Tori	7
Clair	4

Union

Name	# of presentations
Kai	10
Piper	3
Lien	4
Kai	10
Tori	7
Clair	4

Union appends new rows AND new columns if the shred field don’t match across the unioned data sources

Q1

Name	# of presentations	Expenses
Kai	10	$3,761
Piper	3	$287
Lien	4	$1,008

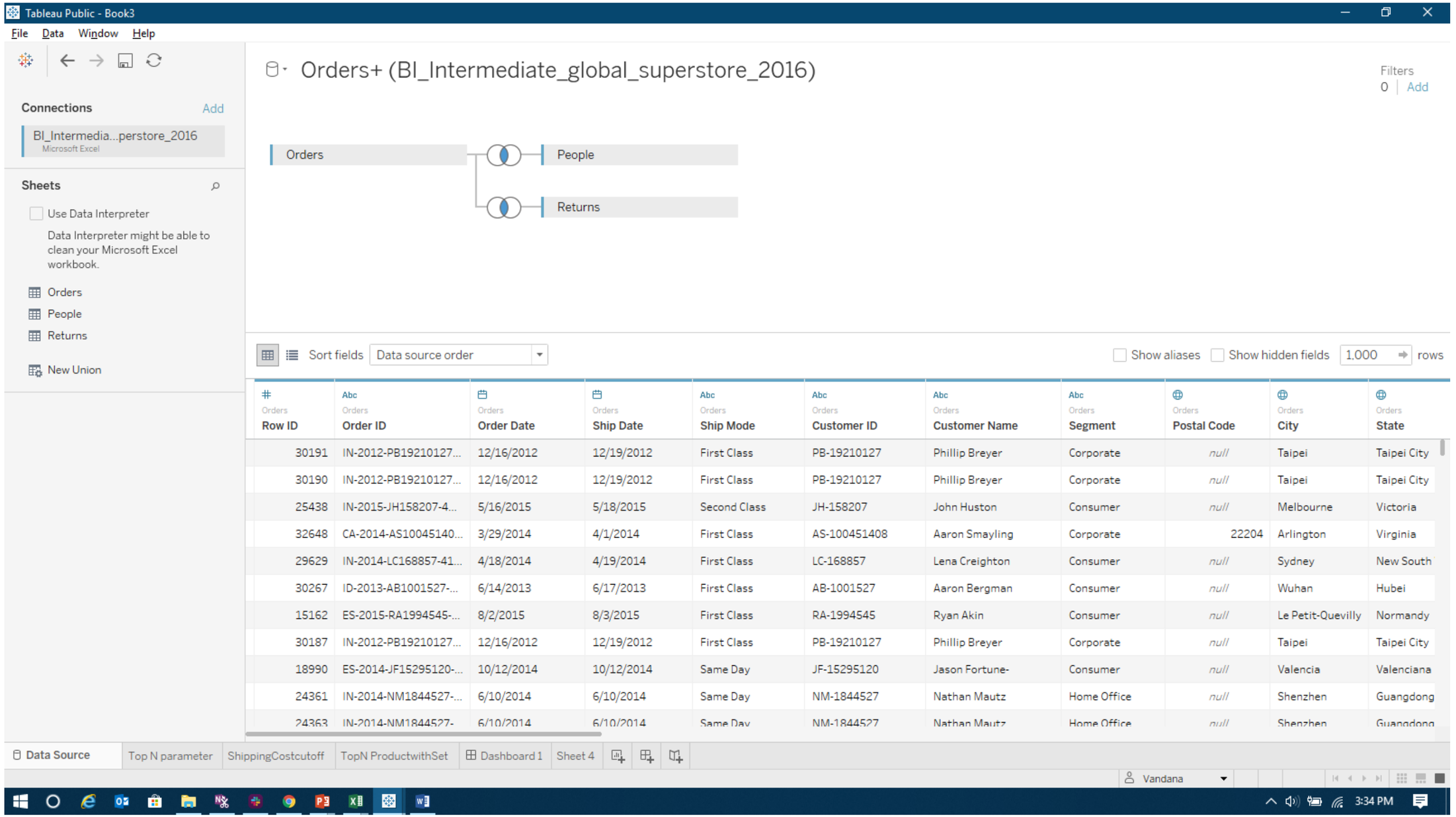
Q2

Name	# of presentations	Expense report
Kai	10	$4,024
Tori	7	$2,930
Clair	4	$987

Union

Name	# of presentations	Expenses	Expense report
Kai	10	$3,761	*null*
Piper	3	$287	*null*
Lien	4	$1,008	*null*
Kai	10	*null*	$4,024
Tori	7	*null*	$2,930
Clair	4	*null*	$987

Data Joining (Exercise):



Walkthrough Exercises:

Data Blending

Agenda:

- Blending Data in Tableau
- Understanding when to blend data

Connect Tableau to the Excel file provided:
BI_Intermediate_Unit3_Exercise_Clustering.xlsx
BI_Intermediate_Unit3_Exercise_US_Population.csv

Data Blending:

What is data bending

Data blending: method for combining data that supplements a table of data from one data source with columns of data from another data source.

For ex, you have transactional data stored in Salesforce and quota data stored in an Excel workbook. The data you want to combine is stored in different databases, and the granularity of the data captured in each table is different in the two data sources, so data blending is the best way to combine this data.

When to use data bending

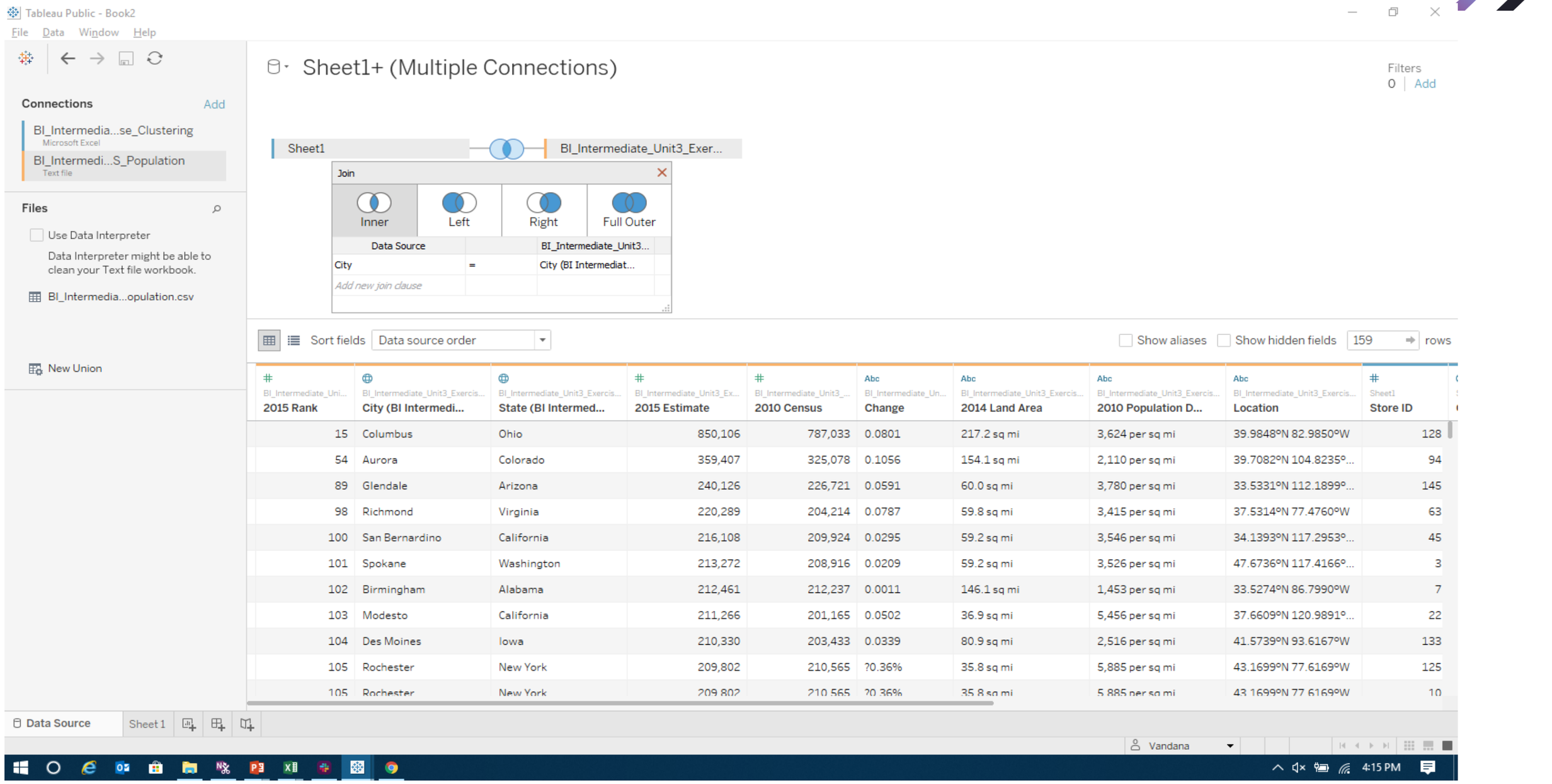
- ✓ You want to combine data from different databases that are not supported by cross-database joins
- ✓ Data is at different levels of detail
- ✓ Data needs cleaning (that is, rename columns, change column data types, create groups, use calculations, etc.)
- ✓ Joins cause duplicate data
- ✓ You have lots of data

Pre-requisites of data bending

- ✓ Data blending requires a primary data source and at least one secondary data source
- ✓ Defined relationship between the primary and secondary data sources (linking field)

Note: When you blend on a field with a high level of granularity, for example, date instead of year, queries can be slow.

Data Blending:



Walkthrough Exercises:

Exercise 1:

Maps, Scatterplots, and Dashboard

Agenda:

- Creating a Map, working with Hierarchies

Connect Tableau to the Excel file provided:
BI_Intermediate_Unit2_Ex1_Map_ofOEurope_Scatterplot.xlsx

Walkthrough Exercises:

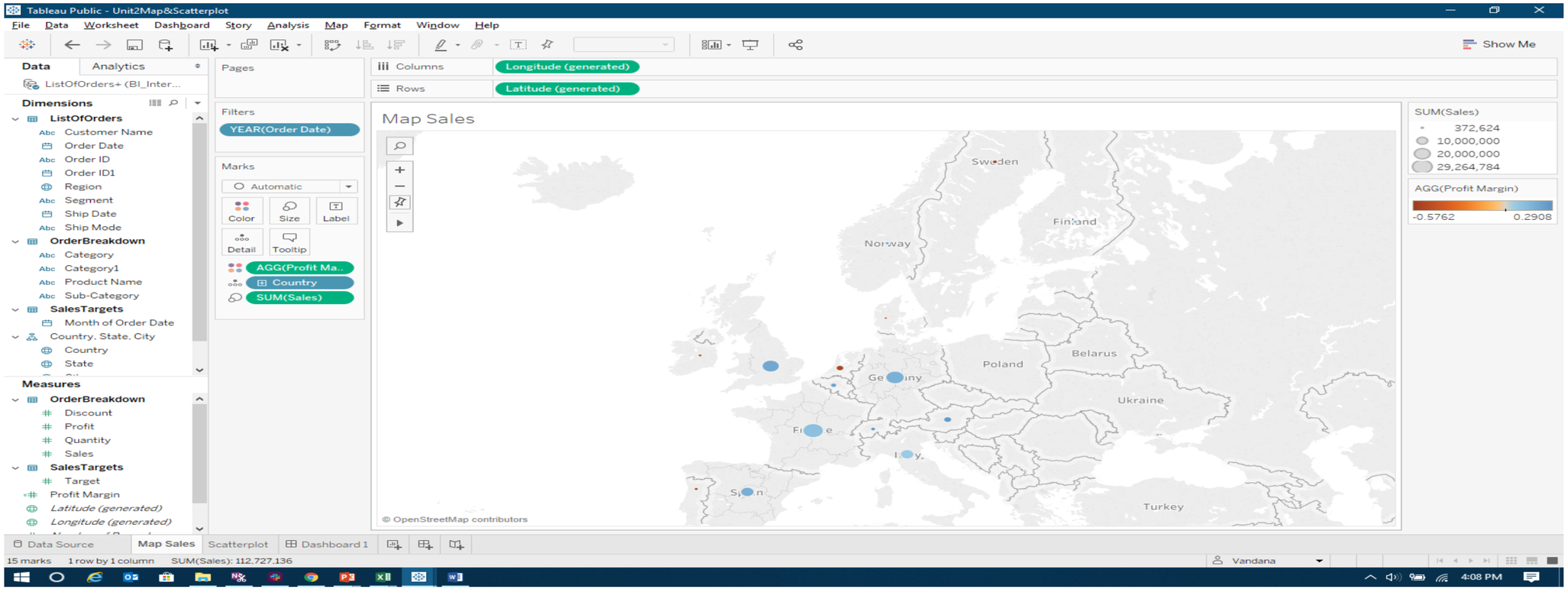
Instructions:

Creating a MAP:

Create a MAP of Europe, showcasing the Sales per Country / State / Region.

Add a calculated field and visualize it on the map:

Profit Margin = Sum of profit / Sum of sales



Walkthrough Exercises:

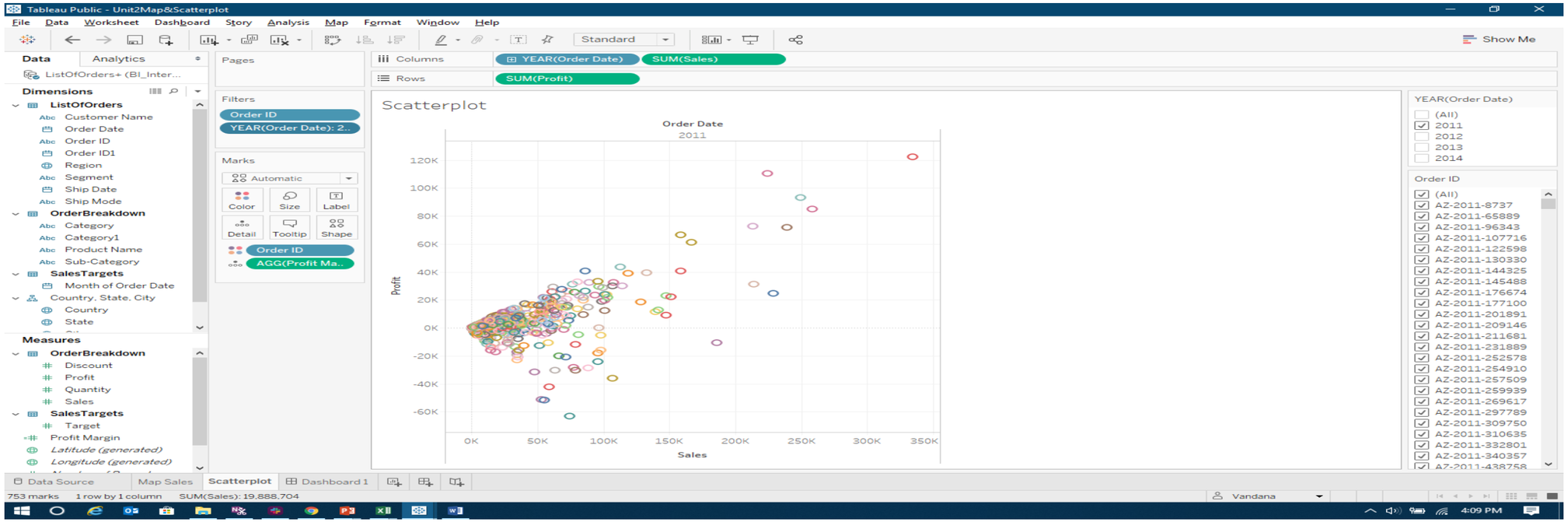
Instructions:

Creating a Scatterplot

1) Plot the customer Scatterplot with:
- SUM of Sales in Columns
- SUM of profit in Rows

2) Add a filter by Year

3) Add your Profit Margin calculated field to add more detail to the scatterplot

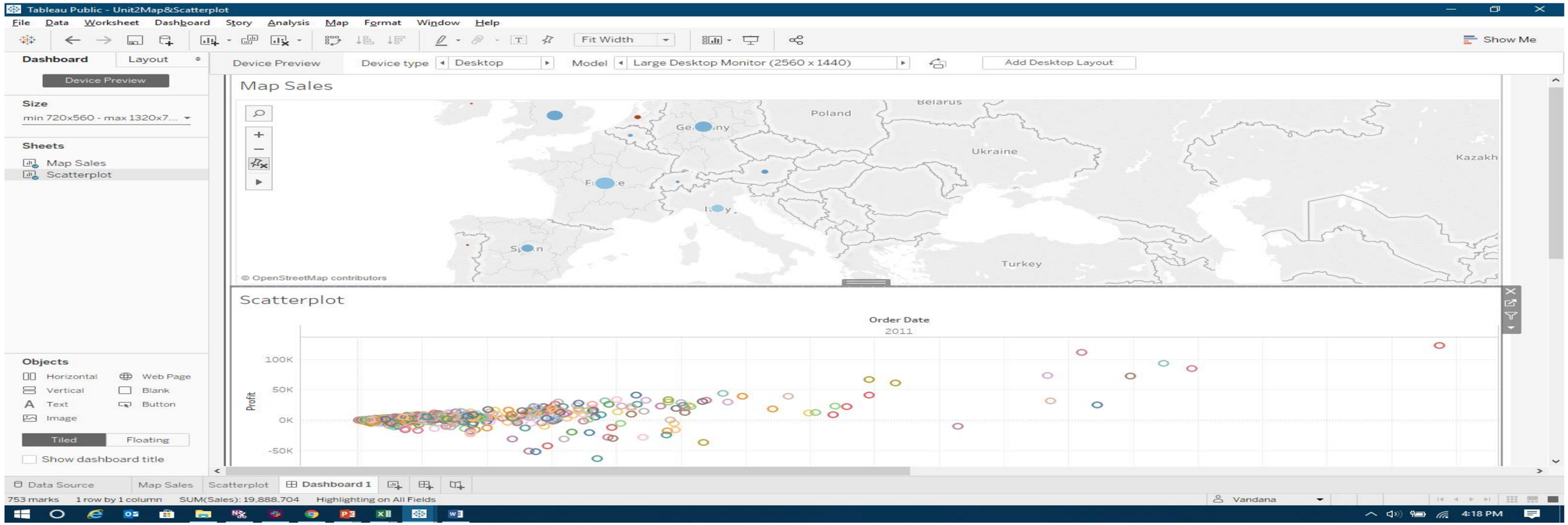


Walkthrough Exercises:

Instructions:

Create your first Dashboard:

1) Create an unified view of your analysis in a dashboard with your MAP and Scatterplot visualizations.
2) Add an interactive action: Filter
3) Add an interactive action: Highlighting



Walkthrough Exercises:

Adding Interactive action to your Dashboard:

Filter:

Filter actions send information between worksheets. Typically, a filter action sends information from a selected mark to another sheet showing related information. Behind the scenes, filter actions send data values from the relevant source fields as filters to the target sheet.

To create Action filter or Highlight in your Dashboard:

- Select Dashboard → Actions
- In the Actions dialog box, click Add action and select Filter or Highlight.
- Or select an existing action and choose Edit.

Walkthrough Exercises:

Exercise 1 **Review:**

MAP:

1) You need to create hierarchies in Tableau when there are hierarchies present in the Dimension and Tableau needs to know about them for you to build your visualization.

Walkthrough Exercises:

Exercise 2:

Create a Customer Segmentation Dashboard

Agenda:

- Creating table Calculation
- Creating bins and Distributions
- Learn about parameters
- Analyse your Segmentation Dashboard

Connect Tableau to the CSV file provided:
BI_Intermediate_Unit2_Ex2_Segmentation_Dashboard.xlsx

Mapping Geographical Roles:

Sometimes Tableau does not do a proper job at identifying locations.
Follow the following process to fix this problem and learn about location mapping.

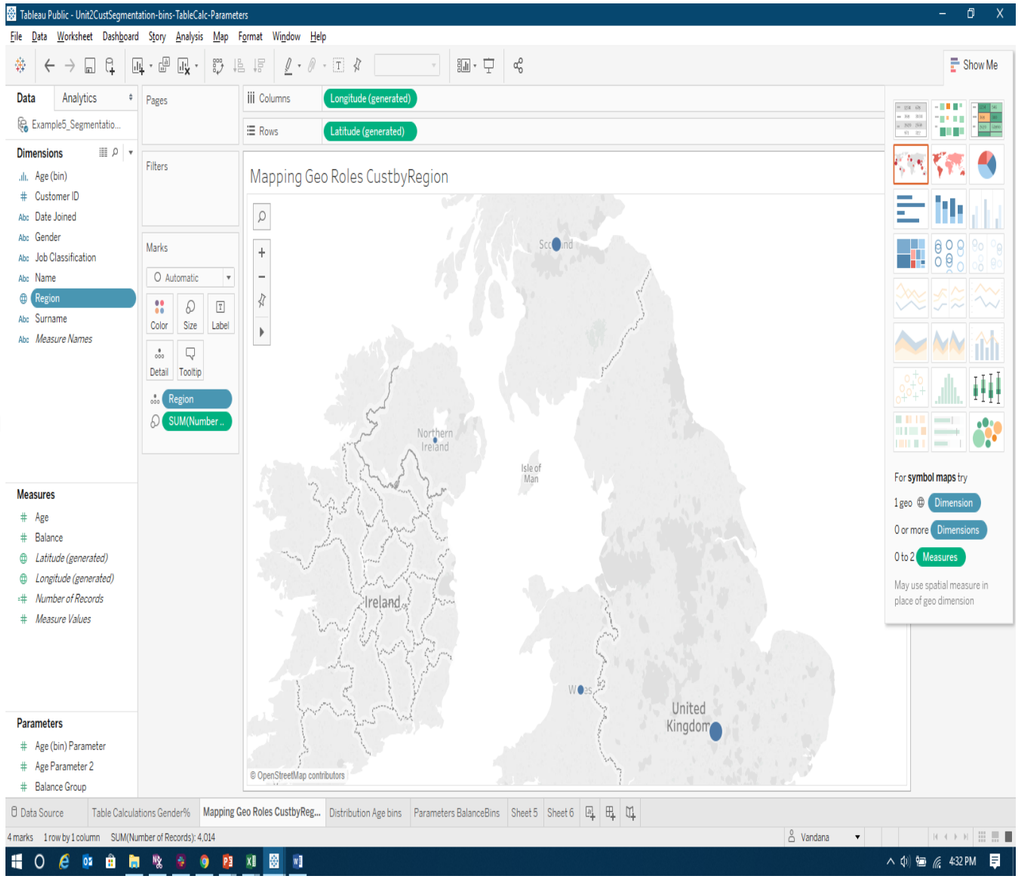
1) Right click on the Region field → Geographic Role → State/Province
Tableau recognize region as a Geographic Data Type.

2) Drag your region field into your Canvas.
You can see a map of the world but nothing else.

3) On the bottom right you can see Unknown.
You have 4 unknown values n this Geographic classification.

4) To fix this, click on Unknown → Edit Locations
You need to specify Tableau which country you are working with
In your drop down menu for Country/Region, choose UK
Click OK

Now Tableau has matched your locations.



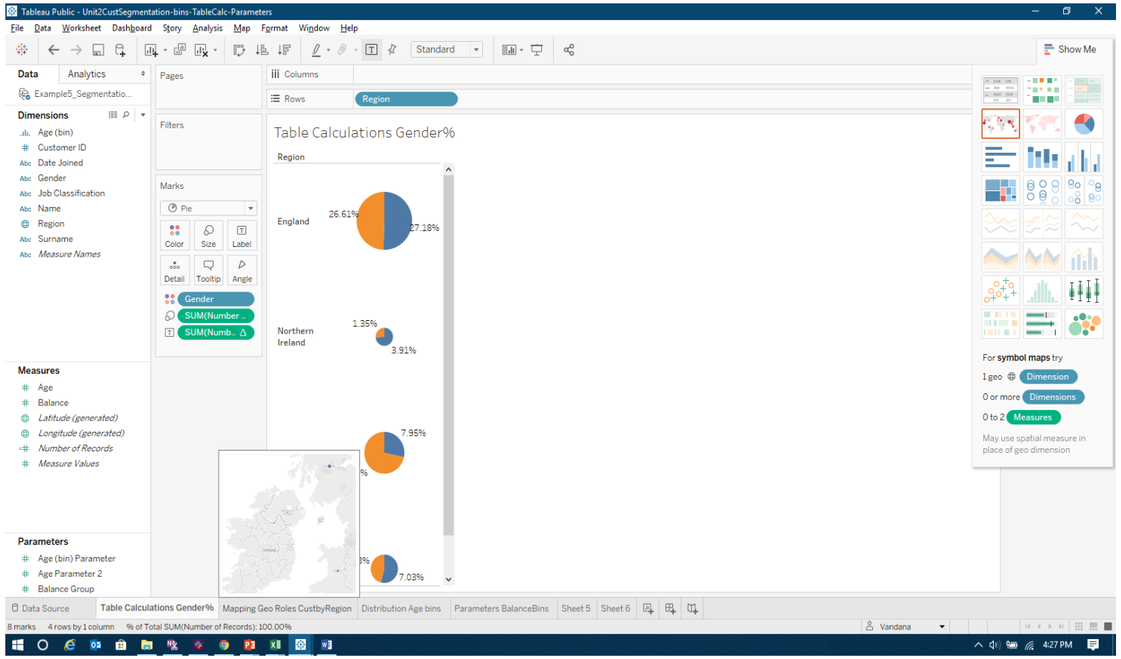
Quick intro to Table calculations:

- A table calculation is a transformation you apply to the values in a visualization. Table calculations are a special type of calculated field that computes on the local data in Tableau.

In our Pie chart, we want to display a Percentage of customers by gender rather than the number of customer by gender.

To do so, we have to use a table calculation.

- Right click onto the field that you used to label you chart
(here SUM(number of records)
→Quick Table Calculation
→Select Percent of Total



In depth table calculations: https://onlinehelp.tableau.com/current/pro/desktop/en-us/functions_functions_tablecalculation.htm

Creating Bins And Distributions:

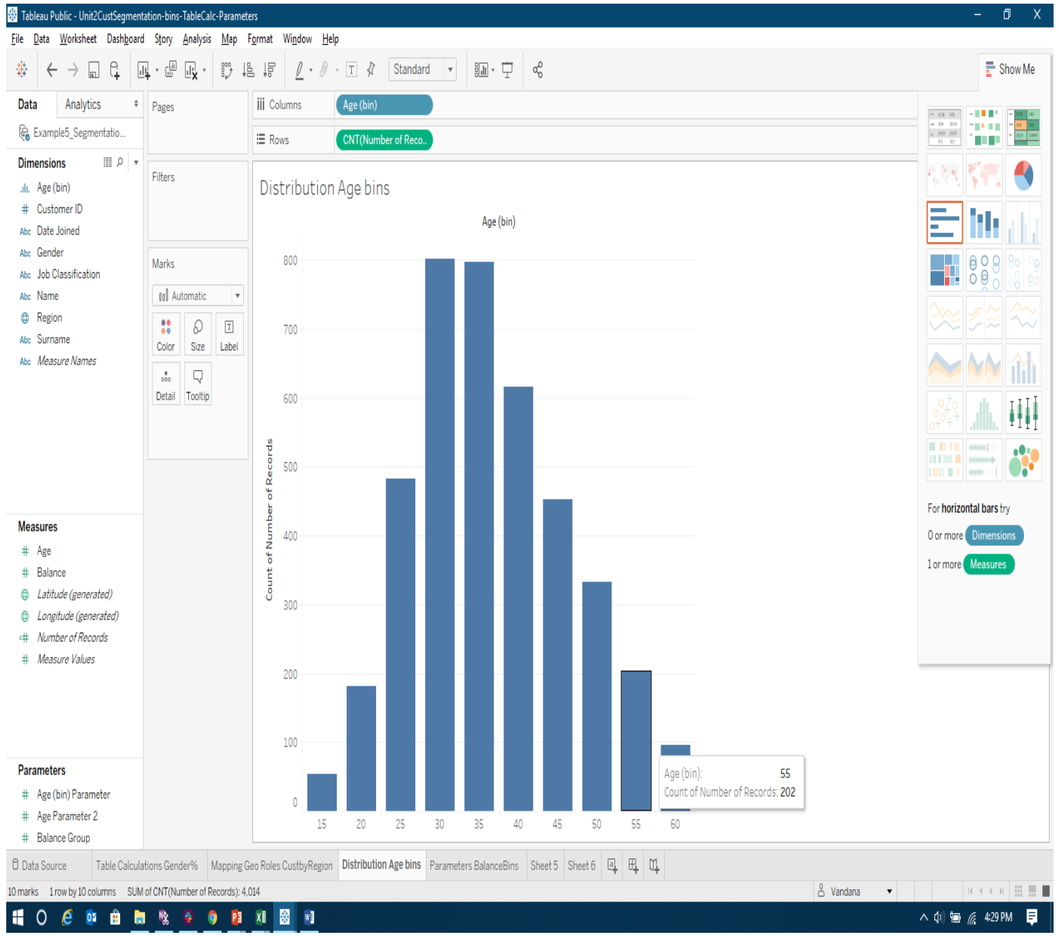
We want to create a distribution of number of customers across ages.
We don't want to see the distribution across each year, instead we want to create 5 year bins distribution.

Bins can be created only on a Measure field (variable).

To create bins for the age field :
Right click onto Age → Create → Bin
Size of bin → select 5

You just created a new categorical variable, Age(bin) should appear into your Dimensions.

You can now create your Distribution by using this new field.



Creating Parameters:

A parameter will allow you to provide a value to pass into Tableau. Parameters allow you to come up with scenarios or options that are not available in your data and create these values to put into your visualization

Here we want to create a parameter that will allow the user to adjust the size of the bins in the visualization.

To create this parameter:
1) Right click onto Measures → Create parameter

Your parameter window:
- Name your parameter (ex: balance group)
- Change Data type to Integer
- Current value : 10,000
- Allowable value : click on Range
- Minimum: 50000 – Maximum: 25000 – Step size: 5000

UNIT 2:
DATA PREPARATION IN TABLEAU

In this section, we will discuss how to prepare and clean our data. We will understand how Tableau reads Data and finally we will have a look at some of Tableau features for Data Preparation

Data Preparation

➢ What Format Your data should be in Tableau ?

Those 3 pictures provide 3 different views in Excel of the same Dataset.

	Jan-05		Feb-05		Mar-05		Apr-05		May-05		Jun-05		Jul-05	
Age	Men	Woman	Men	Woman	Men	Woman	Men	Woman	Men	Woman	Men	Woman	Men	Woman
16 to 19 years	91000	38000	90000	54000	114000	53000	86000	53000	87000	44000	74000	47000	79000	39000
20 to 24 years	175000	90000	154000	114000	137000	113000	125000	114000	106000	99000	109000	73000	110000	72000
25 to 34 years	194000	142000	176000	119000	195000	102000	191000	151000	183000	149000	162000	114000	139000	156000
35 to 44 years	201000	180000	184000	170000	211000	141000	206000	172000	173000	144000	146000	114000	164000	115000
45 to 54 years	207000	157000	186000	173000	223000	152000	181000	160000	159000	178000	171000	150000	181000	123000
55 to 64 years	101000	71000	98000	94000	115000	101000	93000	90000	87000	74000	64000	58000	56000	81000
65 years and over	33000	82000	38000	85000	32000	72000	31000	80000	22000	71000	37000	89000	36000	86000

Human Readable format

Age	Gender	Period	Unemployed
16 to 19 years	Men	Jan-05	91000
20 to 24 years	Men	Jan-05	175000
25 to 34 years	Men	Jan-05	194000
35 to 44 years	Men	Jan-05	201000
45 to 54 years	Men	Jan-05	207000
55 to 64 years	Men	Jan-05	101000
65 years and over	Men	Jan-05	33000
16 to 19 years	Women	Jan-05	38000
20 to 24 years	Women	Jan-05	90000
25 to 34 years	Women	Jan-05	142000
35 to 44 years	Women	Jan-05	180000
45 to 54 years	Women	Jan-05	157000
55 to 64 years	Women	Jan-05	71000
65 years and over	Women	Jan-05	82000
16 to 19 years	Men	Feb-05	90000
20 to 24 years	Men	Feb-05	154000
25 to 34 years	Men	Feb-05	176000
35 to 44 years	Men	Feb-05	184000
45 to 54 years	Men	Feb-05	186000
55 to 64 years	Men	Feb-05	98000
65 years and over	Men	Feb-05	38000
16 to 19 years	Women	Feb-05	54000
20 to 24 years	Women	Feb-05	114000
25 to 34 years	Women	Feb-05	119000

Tableau Preferred Format

Separate columns for every dimension that you have.
“Pivoted view” of your data.

Age	Period	Men	Woman	Total
16 to 19 years	Jan-05	91000	38000	129000
20 to 24 years	Jan-05	175000	90000	265000
25 to 34 years	Jan-05	194000	142000	336000
35 to 44 years	Jan-05	201000	180000	381000
45 to 54 years	Jan-05	207000	157000	364000
55 to 64 years	Jan-05	101000	71000	172000
65 years and over	Jan-05	33000	82000	115000
16 to 19 years	Feb-05	90000	54000	144000
20 to 24 years	Feb-05	154000	114000	268000
25 to 34 years	Feb-05	176000	119000	295000
35 to 44 years	Feb-05	184000	170000	354000
45 to 54 years	Feb-05	186000	173000	359000
55 to 64 years	Feb-05	98000	94000	192000
65 years and over	Feb-05	38000	85000	123000
16 to 19 years	Mar-05	114000	53000	167000
20 to 24 years	Mar-05	137000	113000	250000
25 to 34 years	Mar-05	195000	102000	297000

Data Preparation

➢ What Format Your data should be in Tableau ?

To better interpret your data, Your source file should not contain any design formatting option (Titles in Bold, Colors for Headers…) and should present a pivoted view of your data.
Your dataset should be empty of any formula (ex: Total)

Before connecting your data source to Tableau, sometimes you will have to make the transformations in Excel directly and prepare the data to be imported later into Tableau.

Tableau however can help us interpret our data, it has some useful functionalities that will help us to transform our Data Source Format directly into Tableau.

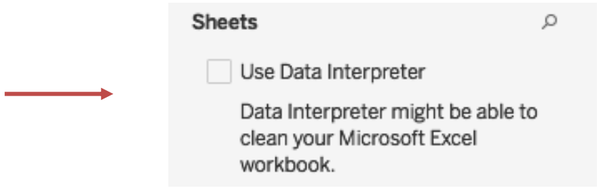
Those functionalities are:
- The Data Interpreter
- The Pivot function
- The Splitting function

Data Preparation

➢ THE DATA INTERPRETER

As you can see, the data is completely messed up. We have null values everywhere, random column names…

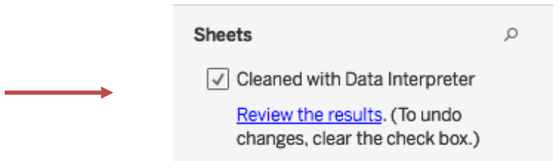
3) Select Use Data Interpreter (Top left of your Connection Manager window in Tableau.



4) As you can see, the data has changed, Tableau has done a pretty good work at interpreting our data, even if it is not perfect.

- Names of the columns are correct
- The numbers are not incorrect
- We still have the totals, which is not good.

5) Let’s try to understand what happened here. Click on review the results. A new spreadsheet is opened for us.



Data Preparation

➢ THE DATA INTERPRETER

6) Go to the second tab of your spreadsheet, it is a description of what Tableau saw.

	Estimated figures										
REGIONS/COUNTRIES	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Header
											Data
EUROPE	########	########	########	########	########	########	########	########	########	########	Data
EU 28 countries + EFTA	########	########	########	########	########	########	########	########	########	########	Data
EU 15 countries + EFTA	########	########	########	########	########	########	########	########	########	########	Data
AUSTRIA	307,915	308,594	298,182	293,697	319,403	328,563	356,145	336,010	319,035	303,318	Data
BELGIUM	480,088	526,141	524,795	535,947	476,194	547,340	572,211	486,737	486,065	482,939	Data
DENMARK	148,819	156,936	162,686	150,199	112,454	153,858	170,036	170,763	182,086	189,051	Data
FINLAND	148,161	145,700	125,608	139,669	90,574	111,968	126,123	111,251	103,455	106,236	Data
FRANCE	2,118,042	2,045,745	2,109,672	2,091,369	2,302,398	2,251,669	2,204,229	1,898,760	1,790,456	1,795,885	Data
GERMANY	3,319,259	3,467,961	3,148,163	3,090,040	3,807,175	2,916,259	3,173,634	3,082,504	2,952,431	3,036,773	Data
GREECE	269,728	267,669	279,745	267,295	219,730	141,501	97,680	58,482	58,694	71,218	Data
ICELAND	18,060	17,129	15,942	9,033	2,113	3,106	5,038	7,902	7,274	9,536	Data
IRELAND	171,742	178,484	186,325	151,607	57,453	88,446	89,911	79,498	74,367	96,344	Data
ITALY	2,244,108	2,335,462	2,494,115	2,161,359	2,159,465	1,961,580	1,749,740	1,403,010	1,304,648	1,360,293	Data
LUXEMBOURG	48,517	50,837	51,332	52,359	47,265	49,726	49,881	50,398	46,624	49,793	Data
NETHERLANDS	465,196	483,999	504,300	499,980	387,699	482,531	555,812	502,544	416,717	387,835	Data
NORWAY	109,907	109,164	129,195	110,617	98,675	127,754	138,345	137,967	142,151	144,202	Data
PORTUGAL	206,488	194,702	201,816	213,389	161,013	223,464	153,404	95,309	105,921	142,826	Data
SPAIN	1,528,877	1,634,608	1,614,835	1,161,176	952,772	982,015	808,051	699,589	722,689	855,308	Data
SWEDEN	274,301	282,766	306,794	253,982	213,408	289,684	304,984	279,899	269,599	303,948	Data
SWITZERLAND (+FL)	266,770	269,421	284,674	288,525	266,018	294,239	318,958	328,139	307,885	301,942	Data
UNITED KINGDOM	2,439,717	2,344,864	2,404,007	2,131,795	1,994,999	2,030,846	1,941,253	2,044,609	2,264,737	2,476,435	Data
EUROPE NEW MEMBERS	1,056,340	1,140,950	1,305,088	1,309,842	864,307	846,145	827,224	794,622	789,162	899,633	Data
BULGARIA*	25,956	36,455	43,521	45,143	22,869	16,257	19,250	19,419	19,352	20,359	Data
CROATIA	70,541	78,775	82,664	88,265	44,918	38,587	41,561	31,360	27,802	33,997	Data
CYPRUS	17,687	18,639	22,878	22,241	14,981	14,088	13,480	10,123	7,000	7,794	Data
CZECH REPUBLIC	151,699	156,686	174,456	182,554	167,708	169,580	173,595	174,009	164,736	192,314	Data
ESTONIA	19,640	25,363	30,912	24,579	9,946	10,295	17,070	19,424	19,694	20,861	Data
HUNGARY	198,982	187,676	171,661	153,278	60,189	43,476	45,094	53,059	56,139	67,476	Data
LATVIA	10,467	14,234	21,606	22,217	7,515	7,970	13,234	10,665	10,636	12,452	Data
LITHUANIA	16,602	25,582	32,771	19,831	5,367	6,365	10,980	12,165	12,163	14,503	Data
MALTA	6,552	6,745	6,240	5,423	5,894	4,056	5,428	5,884	5,749	6,451	Data
POLAND	207,007	224,728	277,427	319,190	276,220	315,855	277,427	272,719	289,913	327,709	Data
ROMANIA	214,967	247,411	312,533	285,506	116,016	94,441	81,709	66,436	57,710	70,172	Data
SLOVAKIA	56,916	59,084	59,700	70,040	74,717	64,033	68,203	69,268	66,000	72,249	Data
SLOVENIA	59,324	59,578	68,719	71,575	57,967	61,142	60,193	50,091	52,268	53,296	Data

7) Let's help Tableau by transforming the data in Excel so that it can interpret it better.

→Open your Excel file and do the following:

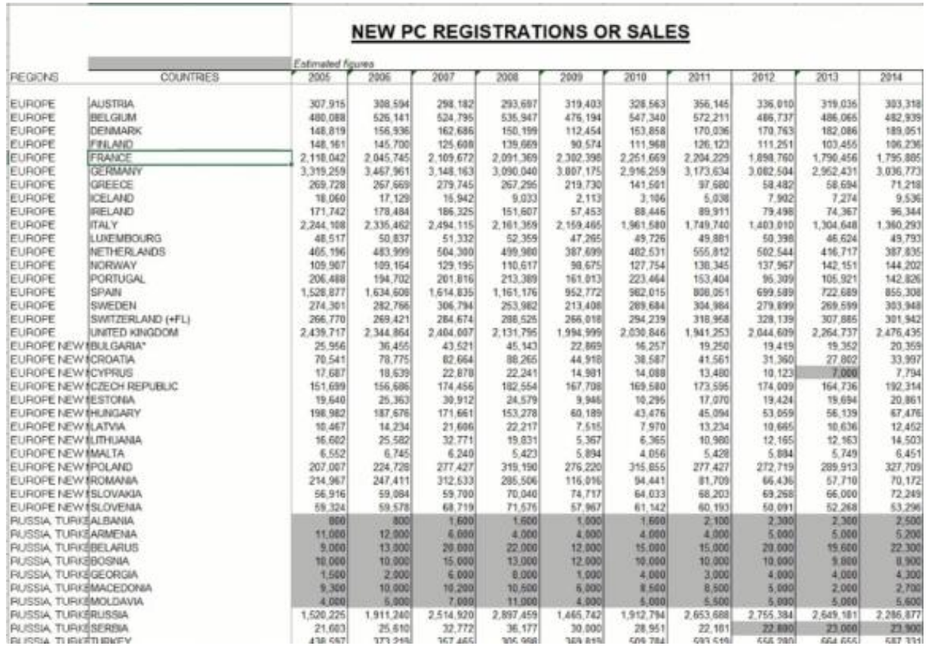
→Create a new column on the left of your table

→Name this column Regions

→Fill in this column with the Region names

→ Get rid off all your total rows

→Save your work



NEW PC REGISTRATIONS OR SALES

		Estimated figures									
REGIONS	COUNTRIES	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
EUROPE	AUSTRIA	307,915	308,594	298,182	293,697	319,403	328,563	356,145	336,010	319,036	303,318
EUROPE	BELGIUM	480,088	526,141	524,795	535,947	476,194	547,340	572,211	486,737	486,065	482,939
EUROPE	DENMARK	148,819	156,936	162,686	150,199	112,454	153,858	170,036	170,763	182,086	189,051
EUROPE	FINLAND	148,161	145,700	125,608	139,669	90,574	111,968	126,123	111,251	103,455	106,236
EUROPE	FRANCE	2,118,042	2,045,745	2,109,672	2,091,369	2,302,398	2,251,669	2,204,229	1,898,760	1,790,456	1,795,885
EUROPE	GERMANY	3,319,259	3,467,961	3,148,163	3,090,040	3,807,175	2,916,259	3,173,634	3,082,504	2,952,431	3,036,773
EUROPE	GREECE	269,728	267,669	279,745	267,295	219,730	141,501	97,680	58,482	58,694	71,218
EUROPE	ICELAND	18,060	17,129	15,942	9,033	2,113	3,106	5,038	7,902	7,274	9,536
EUROPE	IRELAND	171,742	178,484	186,325	151,607	57,453	88,446	89,911	79,498	74,367	96,344
EUROPE	ITALY	2,244,108	2,335,462	2,494,115	2,161,359	2,159,465	1,961,580	1,749,740	1,403,010	1,304,648	1,360,293
EUROPE	LUXEMBOURG	48,517	50,837	51,332	52,359	47,265	49,726	49,881	50,398	46,624	49,793
EUROPE	NETHERLANDS	465,196	483,999	504,300	499,980	387,699	482,531	555,812	502,544	416,717	387,835
EUROPE	NORWAY	109,907	109,164	129,195	110,617	98,675	127,754	138,345	137,967	142,151	144,202
EUROPE	PORTUGAL	206,488	194,702	201,816	213,389	161,013	223,464	153,404	95,309	105,921	142,826
EUROPE	SPAIN	1,528,877	1,634,608	1,614,835	1,161,176	952,772	982,015	808,051	699,589	722,689	855,308
EUROPE	SWEDEN	274,301	282,766	306,794	253,982	213,408	289,684	304,984	279,899	269,599	303,948
EUROPE	SWITZERLAND (+FL)	266,770	269,421	284,674	288,525	266,018	294,239	318,958	328,139	307,885	301,942
EUROPE	UNITED KINGDOM	2,439,717	2,344,864	2,404,007	2,131,795	1,994,999	2,030,846	1,941,253	2,044,609	2,264,737	2,476,435
EUROPE NEW	BULGARIA*	25,956	36,455	43,521	45,143	22,869	16,257	19,250	19,419	19,352	20,359
EUROPE NEW	CROATIA	70,541	78,775	82,664	88,265	44,918	38,587	41,561	31,360	27,802	33,997
EUROPE NEW	CYPRUS	17,687	18,639	22,878	22,241	14,981	14,088	13,480	10,123	7,000	7,794
EUROPE NEW	CZECH REPUBLIC	151,699	156,686	174,456	182,554	167,708	169,580	173,595	174,009	164,736	192,314
EUROPE NEW	ESTONIA	19,640	25,363	30,912	24,579	9,946	10,295	17,070	19,424	19,694	20,861
EUROPE NEW	HUNGARY	198,982	187,676	171,661	153,278	60,189	43,476	45,094	53,059	56,139	67,476
EUROPE NEW	LATVIA	10,467	14,234	21,606	22,217	7,515	7,970	13,234	10,665	10,636	12,452
EUROPE NEW	LITHUANIA	16,602	25,582	32,771	19,831	5,367	6,365	10,980	12,165	12,163	14,503
EUROPE NEW	MALTA	6,552	6,745	6,240	5,423	5,894	4,056	5,428	5,884	5,749	6,451
EUROPE NEW	POLAND	207,007	224,728	277,427	319,190	276,220	315,855	277,427	272,719	289,913	327,709
EUROPE NEW	ROMANIA	214,967	247,411	312,533	285,506	116,016	94,441	81,709	66,436	57,710	70,172
EUROPE NEW	SLOVAKIA	56,916	59,084	59,700	70,040	74,717	64,033	68,203	69,268	66,000	72,249
EUROPE NEW	SLOVENIA	59,324	59,578	68,719	71,575	57,967	61,142	60,193	50,091	52,268	53,296
RUSSIA, TURKEY	ALBANIA	800	800	1,600	1,600	1,000	1,600	2,100	2,300	2,300	2,500
RUSSIA, TURKEY	ARMENIA	11,000	12,000	6,000	4,000	4,000	4,000	4,000	5,000	5,000	5,200
RUSSIA, TURKEY	BELARUS	9,000	13,000	20,000	22,000	12,000	15,000	15,000	20,000	19,600	22,300
RUSSIA, TURKEY	BOSNIA	10,000	10,000	15,000	13,000	12,000	10,000	10,000	10,000	9,800	8,900
RUSSIA, TURKEY	GEORGIA	1,500	2,000	6,000	8,000	1,000	4,000	3,000	4,000	4,000	4,300
RUSSIA, TURKEY	MACEDONIA	9,300	10,000	10,200	10,500	6,000	8,500	8,500	5,000	2,000	2,700
RUSSIA, TURKEY	MOLDAVIA	4,000	5,000	7,000	11,000	4,000	5,000	5,500	5,000	5,000	5,600
RUSSIA, TURKEY	RUSSIA	1,520,225	1,911,240	2,514,920	2,897,459	1,465,742	1,912,794	2,653,688	2,755,384	2,649,181	2,286,877
RUSSIA, TURKEY	SERBIA	21,603	25,610	32,772	36,177	30,000	28,951	22,181	22,890	23,000	23,900
RUSSIA, TURKEY	TURKEY	438,567	373,215	367,465	306,998	369,819	509,784	593,519	566,280	664,655	587,331

Data Preparation

➢ THE DATA INTERPRETER

8) Open the newly updated Excel File into Tableau.

We have a much better result right away.

That is how the data interpreter works, it can save you some time, it cuts out this headers for you, ignores blank rows and so on.
It doesn’t do all the work but at the same time it can help to go in the right direction if you import and check the results of your data interpreter.


Sort fields Data source order

REGIONS	COUNTRIES	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
EUROPE	AUSTRIA	307,916.00	308,594	298,182	293,697	319,403	328,563	366,145	336,010	319,035	303,318
EUROPE	BELGIUM	480,088.00	526,141	524,795	535,947	476,194	547,340	672,211	486,737	486,065	482,939
EUROPE	DENMARK	148,819.00	156,936	162,686	150,199	112,454	153,858	170,036	170,763	182,086	189,061
EUROPE	FINLAND	148,161.00	145,700	125,608	139,669	90,574	111,968	126,123	111,251	103,455	106,236
EUROPE	FRANCE	2,118,042.00	2,045,745	2,105,672	2,091,369	2,302,398	2,251,669	2,204,229	1,898,760	1,790,456	1,795,885
EUROPE	GERMANY	3,319,259.00	3,467,961	3,148,163	3,090,040	3,807,175	2,916,259	3,173,634	3,082,504	2,952,431	3,036,773
EUROPE	GREECE	269,728.00	267,669	279,745	267,295	219,730	141,501	97,680	58,482	58,694	71,218
EUROPE	ICELAND	18,060.00	17,129	15,942	9,033	2,113	3,106	5,038	7,902	7,274	9,536
EUROPE	IRELAND	171,742.00	178,484	186,325	151,607	57,453	88,446	89,911	79,456	74,367	96,344
EUROPE	ITALY	2,244,108.00	2,335,462	2,494,115	2,161,359	2,159,465	1,961,580	1,749,740	1,403,010	1,304,648	1,360,293
EUROPE	LUXEMBOURG	48,517.00	50,837	51,332	52,359	47,265	49,726	49,881	50,898	46,624	49,793
EUROPE	NETHERLANDS	465,196.00	483,999	504,300	499,980	387,699	482,531	555,812	502,544	416,717	387,835
EUROPE	NORWAY	109,907.00	109,164	129,195	110,617	98,675	127,754	138,345	137,967	142,151	144,202
EUROPE	PORTUGAL	206,488.00	194,702	201,816	213,389	161,013	223,464	153,404	95,309	105,921	142,826
EUROPE	SPAIN	1,528,877.00	1,634,608	1,614,835	1,161,176	952,772	982,015	808,051	699,589	722,689	855,308
EUROPE	SWEDEN	274,301.00	282,766	306,794	253,982	213,408	289,684	304,984	279,899	269,599	303,948

Data Preparation

➢ THE PIVOT FUNCTION

If we look at our previous file imported into Tableau, we know that Tableau did a good job at interpreting the result, especially after a little bit of transformation into Excel.

However, as discussed earlier, we can see that this source is still not perfect for being machine readable.
Each year is in a separate column.

Ideally we would like to have 4 columns in total:
REGIONS / COUNTRIES / YEAR / SALES

In order to achieve this result, we will need to create a pivot view.
Tableau provide us this tool and it is very easy to perform.


Sort fields Data source order

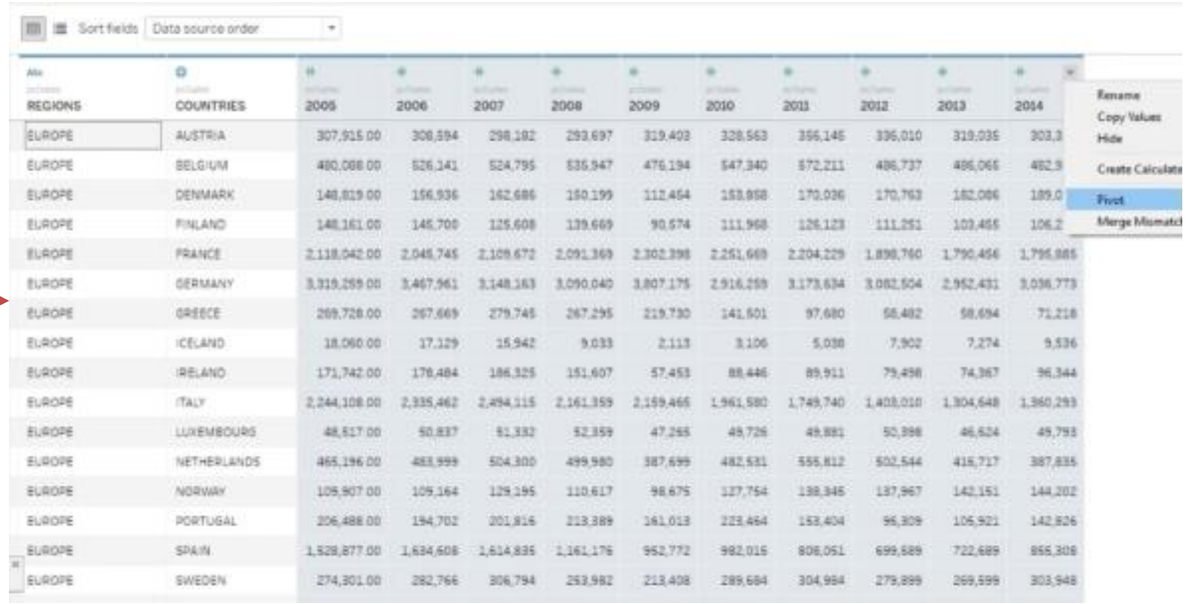
REGIONS	COUNTRIES	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
EUROPE	AUSTRIA	307,915.00	308,594	298,182	293,697	319,403	328,563	366,145	336,010	319,035	303,318
EUROPE	BELGIUM	480,088.00	526,141	524,795	535,947	476,194	547,340	572,211	486,737	486,065	482,939
EUROPE	DENMARK	148,819.00	156,936	162,686	160,199	112,454	153,850	170,036	170,763	182,086	189,061
EUROPE	FINLAND	148,161.00	145,700	125,608	139,669	90,574	111,968	126,123	111,251	103,455	106,236
EUROPE	FRANCE	2,118,042.00	2,045,745	2,109,672	2,091,369	2,302,398	2,251,669	2,204,229	1,898,760	1,790,456	1,795,885
EUROPE	GERMANY	3,319,259.00	3,467,961	3,148,163	3,090,040	3,807,175	2,916,259	3,173,634	3,082,504	2,952,431	3,036,773
EUROPE	GREECE	269,728.00	267,669	279,745	267,295	219,730	141,501	97,680	58,482	58,694	71,218
EUROPE	ICELAND	18,060.00	17,129	15,942	9,033	2,113	3,106	5,038	7,902	7,274	9,536
EUROPE	IRELAND	171,742.00	178,484	186,325	151,607	57,453	88,446	89,911	79,498	74,367	96,344
EUROPE	ITALY	2,244,108.00	2,335,462	2,494,115	2,161,359	2,159,465	1,961,580	1,749,740	1,403,010	1,304,648	1,360,293
EUROPE	LUXEMBOURG	48,517.00	50,837	51,332	52,359	47,265	49,726	45,881	50,398	46,624	49,793
EUROPE	NETHERLANDS	465,196.00	483,999	504,300	499,980	387,699	482,531	555,812	502,544	416,717	387,835
EUROPE	NORWAY	109,907.00	109,164	129,195	110,617	98,675	127,754	138,345	137,967	142,151	144,202
EUROPE	PORTUGAL	206,488.00	194,702	201,816	213,389	161,013	223,464	153,404	95,309	105,921	142,826
EUROPE	SPAIN	1,528,877.00	1,634,608	1,614,835	1,161,176	952,772	982,015	808,051	699,589	722,689	855,308
EUROPE	SWEDEN	274,301.00	282,766	306,794	253,982	213,408	289,684	304,984	279,899	269,599	303,948

Data Preparation

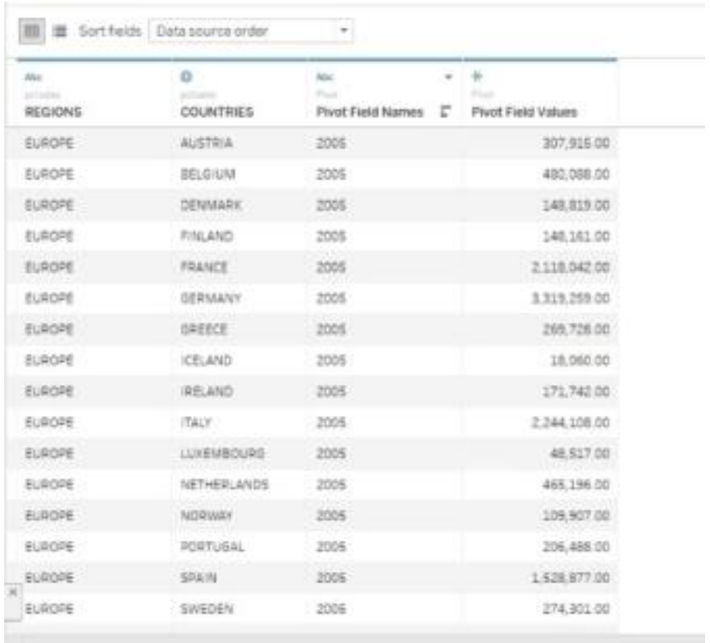
➢ THE PIVOT FUNCTION

To create a pivot in Tableau:

1) Select all the column you want to pivot.
Here the columns 2015 to 2014.
Right click and select → Pivot.



2) Simple, easy and quick.
Now you have Regions / Countries / Pivot field names (year) / pivot field values (Sales)



Data Preparation

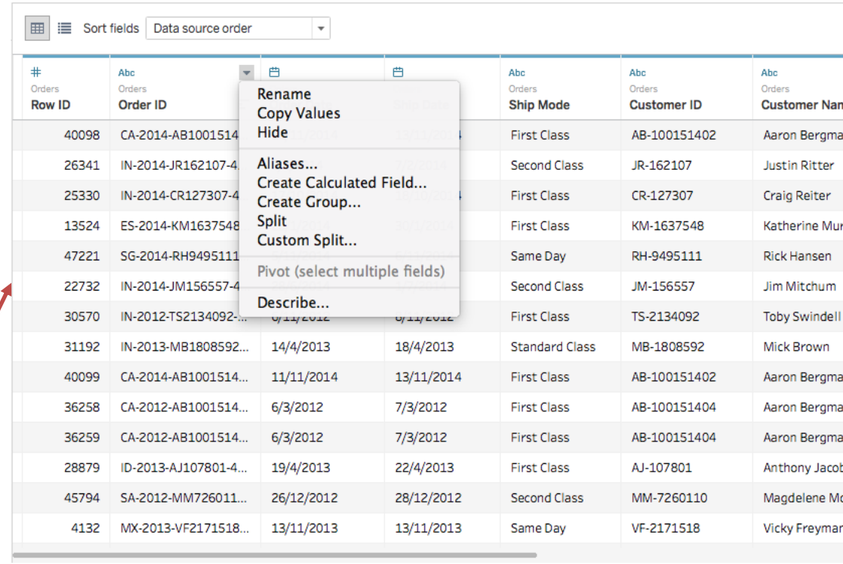
➢ SPLITTING FUNCTION

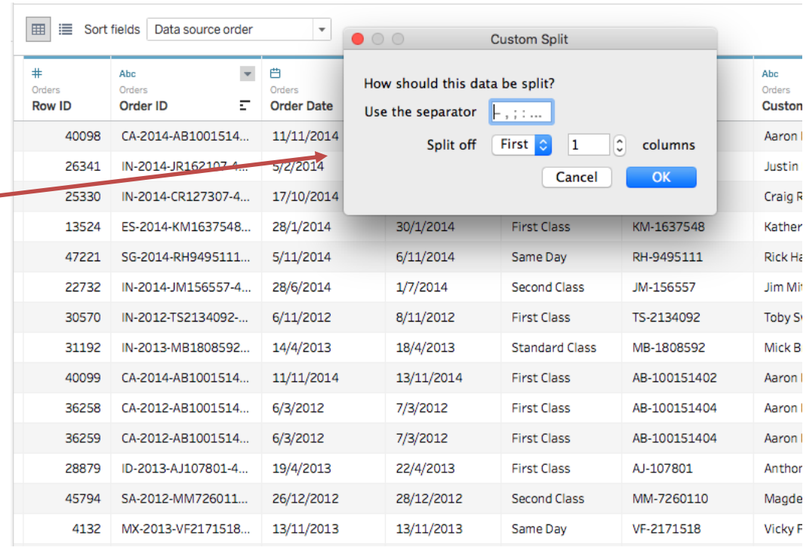
The SPLIT features allows you to split one column into multiple columns.

Can be handy in some situation, for example if you have customer data and you have the full name of the customers into one column and you want to create two separate column of First and Last name.

As the pivot function, the split function is very easy to use. Just click on the little icon right next to your column name, select custom split.

A pop up window appear and ask you how you would like to split your column

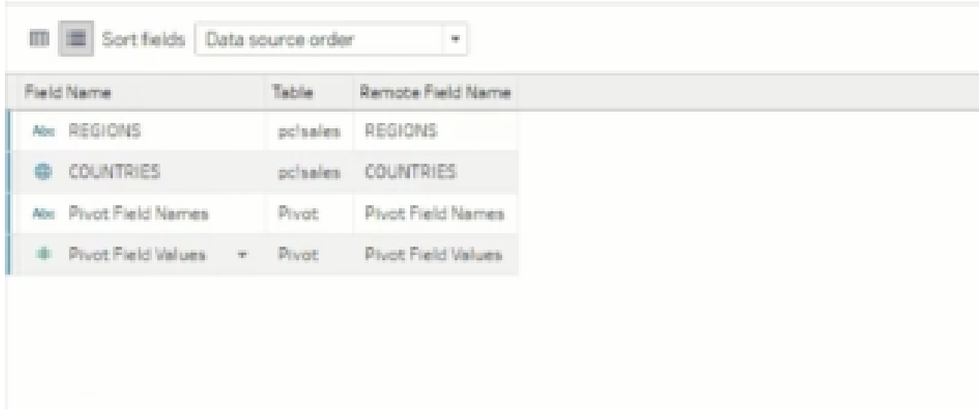




Data Preparation

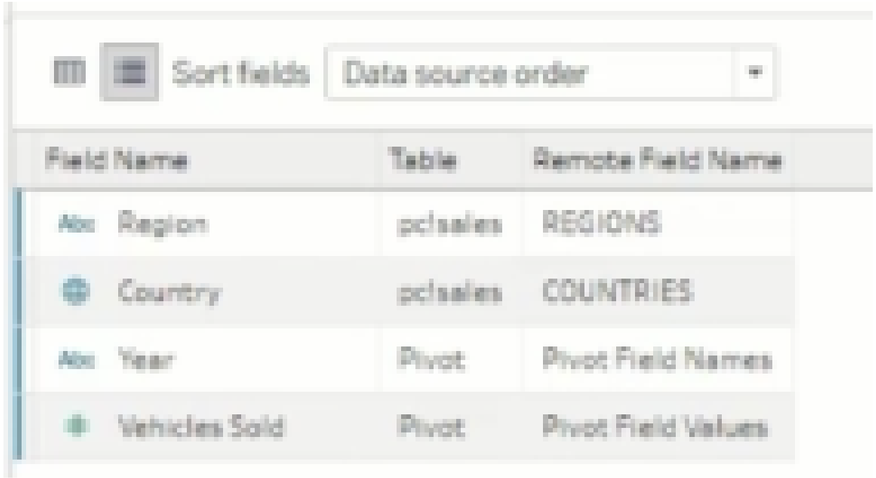
➢ METADATA GRID

Use it to have a look at the Metadata of your datasets.
You can use this view and perform some few adjustment.



Let’s rename our columns that comes from our Pivot created previously.

1) Click on the metadata grid icon.
2) Double click on each field you want to rename.
Rename each field with their proper column name.



You can do a lot of other things, don’t hesitate to explore further.

PRACTISE MAKES PERFECT – EXERCISE1

Create a Customer Segmentation Dashboard

Connect Tableau to the CSV file provided:
BI_Intermediate_Unit2_Ex2_Segmentation_Dashboard.csv

Practice Makes Perfect

Instructions:

1) Create a Map that displays the number of Customers per region.
Hint: use the “Number of Records” field to display the number of customers in your visualization

2) Create a Pie chart that display the number of customers by gender.
Use table calculation to change your labels from numbers of customers to percentage.

3) Create a Distribution (histogram) of number of customers by age.
Create age bins to showcase the number of customers distributed across a 5 years bins.

More on the next slide

Practice Makes Perfect

Instructions:

4) Create a new Distribution of your number of customers by balance.
- Create balance bins of 10,000.
- Use table calculation to display your number of records as Percentages.
- Create a parameter that will allow the user to adjust the size of the bins in the visualization.

5) Create a parameter on your distribution of number of customers by age.
This time create a list of values, not a range.
List should be 1, 5 and 10.
Don’t forget to link the parameter to your Age(bin) field)

6) Create a Tree Map that showcase number of records of your customers by job classification

7) Create A Customer Segmentation Dashboard using the 5 previous graphs.
Make sure your parameters appears and are connected to your 2 distribution charts.
Enhance your dashboard, Create title, format you charts with useful color scales.

Practice Makes Perfect

Instructions:

8) Make your Dashboard Interactive.
Create Action filters for all your graphs.

9) Analyse your Customer Dashboard.
What is happening in your different regions of UK?

PRACTISE MAKES PERFECT – EXERCISE2

Create a Dashboard for Company's performance

Connect Tableau to the excel file provided:
Global Superstore Dataset

Practice Makes Perfect (Dashboard)

Exercise: Time for analysis

Given your **Global Superstore Dataset:**

You are asked to create a Dashboard and Storyline to provide insights about the company's key Performance worldwide for the 2016 period.

You will have to create multiple charts and graphs and provide an interactive visual analysis that yields insights.

Based on your analysis and Dashboard, create a storyline for presentation and provide business recommendations.

You will be asked to present your Analysis to the Top Management.

This Visual analysis should focus on High level Performance of the company worldwide.

Your findings/visualizations may be shared internally, hence it is important to strive for clarity, simplicity and consistency.