DATA 203
DATA VISUALIZATION USING TABLEAU

La Guardia Community College

Class 11

LEARNING OBJECTIVES

- Data Blending

DATA BLENDING

WHAT IS DATA BLENDING

- Data Blending is a powerful feature in Tableau.
- When you have related data in multiple data sources, which you want to analyze together in a single view.
- A data blend can join data sources **after aggregation** is performed on the individual sources.

DATA BLENDING

Date Files: CoffeeChain Sales, CoffeeChain Location, CoffeeChain Product

Tableau: Open a new workbook and call it Data Blending

ADD THE PRIMARY DATA SOURCE

1. With Blend you create one data source per data file that you upload
2. First load the Coffee Chain Sales file to Tableau through the Data level Connection.
3. Rename the data source as Coffee Chain Sales
4. Highlight your Connection Name Sheet1 (CoffeeChain Sales) and rename it to CCSales

DATA BLENDING

ADD THE SECONDARY DATASOURCE

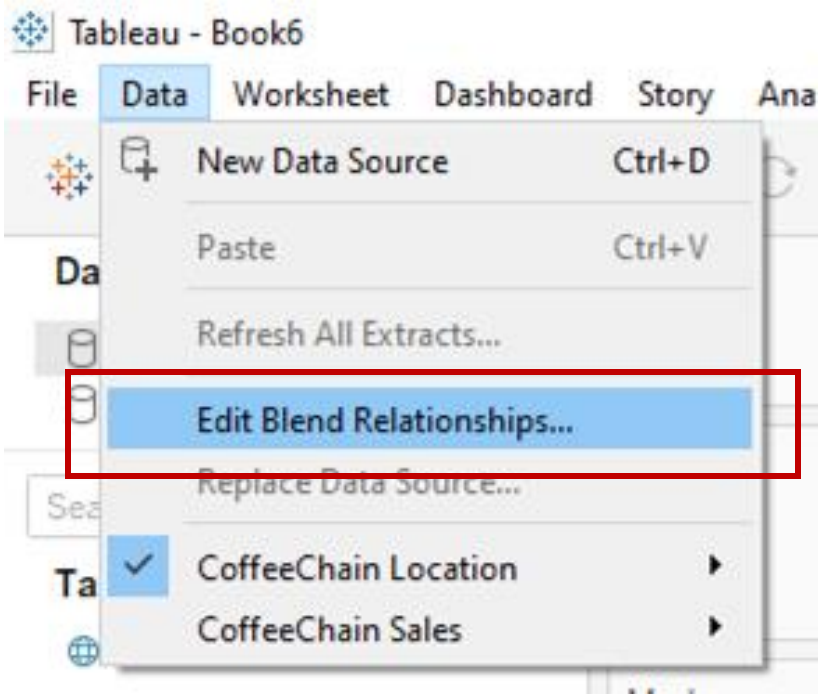
1. You can add the second data source either from the Data Connection window
2. Click on the downward arrow on the Database Icon.

1. Upload the CoffeeChain Location file.
2. Rename the data source as CCLocation.
3. Go to Sheet 1 and rename it CCSales

DATA BLENDING

BLENDING THE DATA

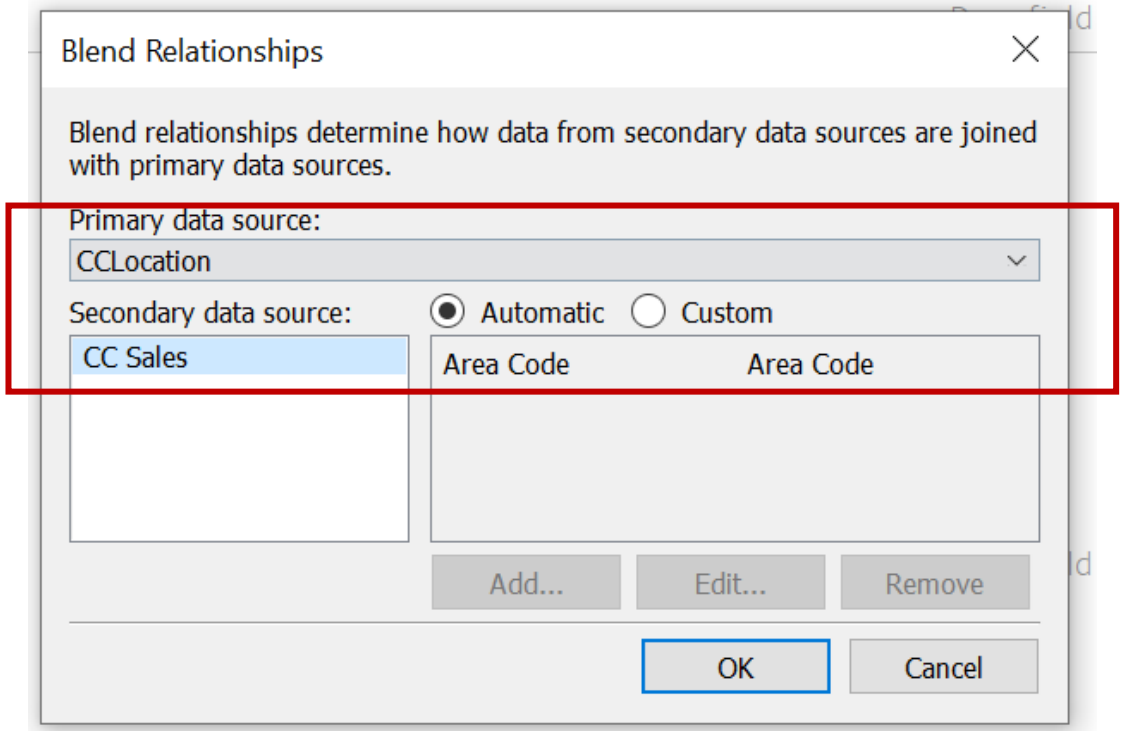
1. On your Data Panel you see two different data sources:
    1. CCSales
    2. CCLocation
2. These two data sources are connected by a common field
3. To see how Tableau links two data source
    1. Click on Data > Edit Blend Relationship



DATA BLENDING

HOW TO EDIT DATA BLEND RELATIONSHIP

1. You see the Primary & Secondary Data sources
2. Tableau by default has identified that the Area code field is common to both the Primary and Secondary data source.
3. Should you need to change this, you can click on Custom and add any additional fields that you might need too.



DATA BLENDING

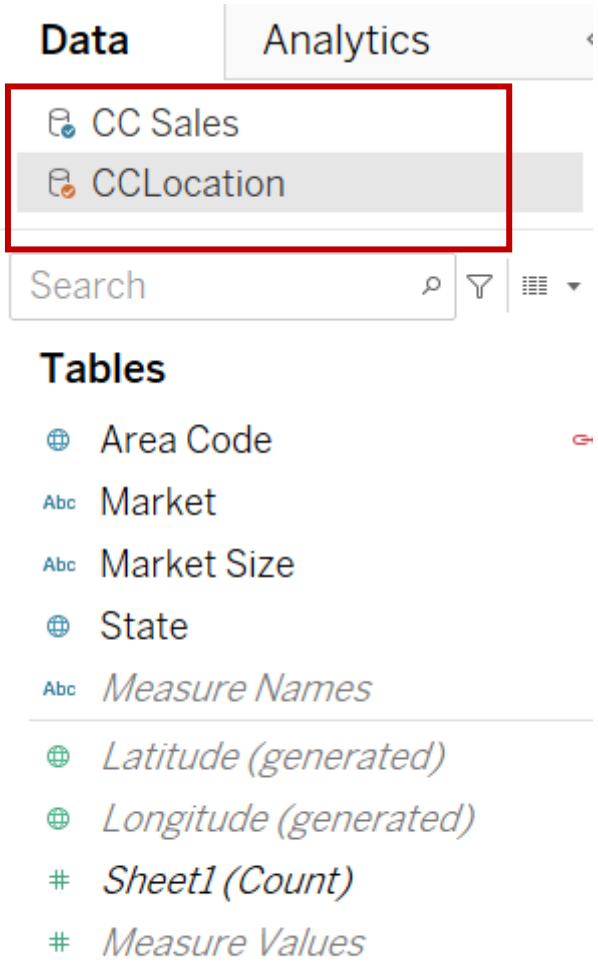
Creating visualization with Blended Data Sources

1. Add Sales from CC Sales to the Row Shelf

1. A **blue check** next to the CC Sales data source is identified as **a primary data source**

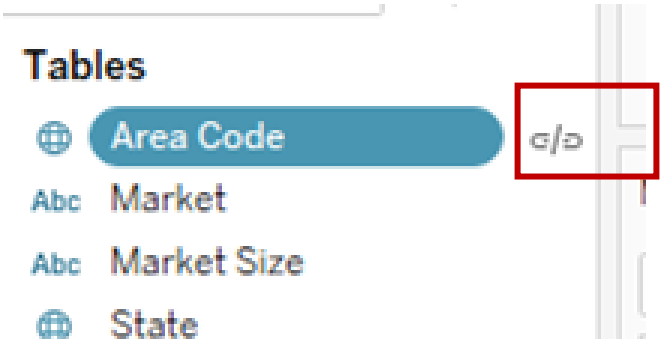
1. Add Area Code from CC Location to the Column Shelf.

1. You will see an **orange check** next to CC Location as its identified as a **secondary data source**



DATA BLENDING

1. You will see a **red link (active link)** next to the Area Code in data source CC Location.
2. This indicates that Area Code is related to your primary data source CC Sales.
3. At times you will see a **gray link (inactivate).**
4. You will have to click on the gray link to change to red link
5. This is on a per worksheet basis.
6. Click on Sheet 2. What do you see?



DATA BLENDING

RESULTS OF BLENDED DATA SOURCE

1. Column Shelf - add Product ID from the CC Sales
2. Row Shelf - add Area Code from CC Location
3. Marks Text Card – add Sales from CCSales
4. The secondary data source is joined to the primary data source by a **LEFT JOIN**
5. Primary Table is CCSales where we have sales values for 13 Product IDs.
   - All Product IDs are going to be retained by default
   - Aggregation is on the primary data source
   - All Products are not being sold in all Area Codes or all Markets.

	Product Id												
Area C..	1	2	3	4	5	6	7	8	9	10	11	12	13
203		2,532			1,557	1,186			897	930	484		348
206		963	2,802	1,094	787	1,731		3,092	197		796	1,187	1,171
209	464	1,471			289	1,190		247	1,128	90	276	187	201
210		532			856	123			84				
212					213		1,358		647	112	621		556
213		2,253		1,054		1,702		198	957	119	290		168
214		546	283			334		383	101				
216	255		122		1,220	225		97				200	
217		1,168	638		1,191	478			190		716		
224		463			1,210			436	178	262			

DATA BLENDING

HANDS ON EXERCISE

- Add a 3rd data source CoffeeChain Product.
- Rename it CC Product
- Create a visualization showing sales by Product in the different market
- What are your findings?

TIME: 15 minutes

DATA BLENDING

HANDS ON EXERCISE