Session 4 – Significance of data joins & blends in visual analytics

Case study six: data joining & blending using different data sources

> **Dataset description: Sales in retail and e-commerce domains**
>
> **Download link:** click here
>
> **Names of the datasets: sample superstore sales and coffee chain.mdb**
>
> **Source: Tableau**

Scenario: As a data analyst you work with data from different sources. In this case study identify the common fields in the datasheets or tables to join them. We will use joining and blending methods to connect datasets and analyze data.

Information reference:

SQL Joins

https://www.w3schools.com/sql/sql_join.asp

1. Load **coffee chain** (MS Access database) and **sample superstore** (Excel data file)

2. Within the coffee chain database, how would you connect different data tables?

3. Identify the primary keys which could be used for joining data tables?

4. Create a data blend between **sample superstore** and **coffee chain** database

5. Explain the revenue trends from **coffee chain** and **sample superstore** datasets and write down insights

Session 4 – Significance of data joins & blends in visual analytics

Case study seven: creating dual axis charts and performing data blending

Dataset description: Amazing market EU

Data Source: superdatascience

Download link: click here

Scenario: Perform exploratory analysis to derive insights on the target revenues and the actual revenues generated by a store. The data comes from different worksheets hence level of detailed information which exists under each variable is different. We will use data blending and data joining methods for connecting various data sheets.

6. First create an inner join to connect data sheets "list of orders" and "orders break down" in the **Amazing Market EU** dataset

7. Secondly create a data blend between data sheets "list of orders", "orders break down" and "sales targets"

8. You can use the edit relationships option in tableau to map dimensions in the two data sheets

Session 4 – Significance of data joins & blends in visual analytics
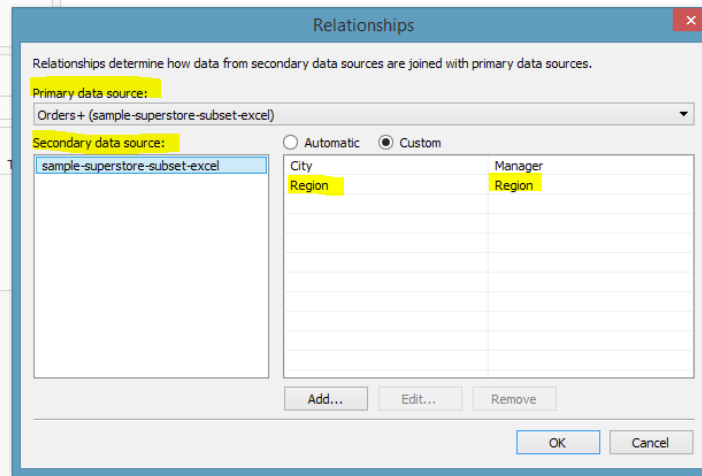
9. What is the reason to blend dimensions with data sheet "sales targets"?

10. Create a bar graph with year wise revenues for different category of products

11. Now integrate target revenue measure into a dual axis chart

12. **Advanced concept**: create a calculated field, name it as "difference in revenue" for calculating the difference in actual revenues obtained and revenue targets

13. Use the newly calculated "difference in revenue" measure, try to see if the sales targets have been met for each product category

14. Prepare a summary of your observations

Hint how the final visualization of part six should look like

Session 4 – Significance of data joins & blends in visual analytics