

## Lesson 3 - Exercise 3.1: Joins and Data Extract

### Business Scenario

Rollin Ford, the regional manager of a retail chain, is in process of re-designing distribution and production strategies. He wants to look at profit for each state along with names of managers. He also wants to see the number of products returned by customers in each state. Combine the data that is stored across different tables and create a visualization to satisfy the above requirement. In order to enhance the performance, create an extract for this data source. Rollin wants refresh to be happen incrementally based on date.

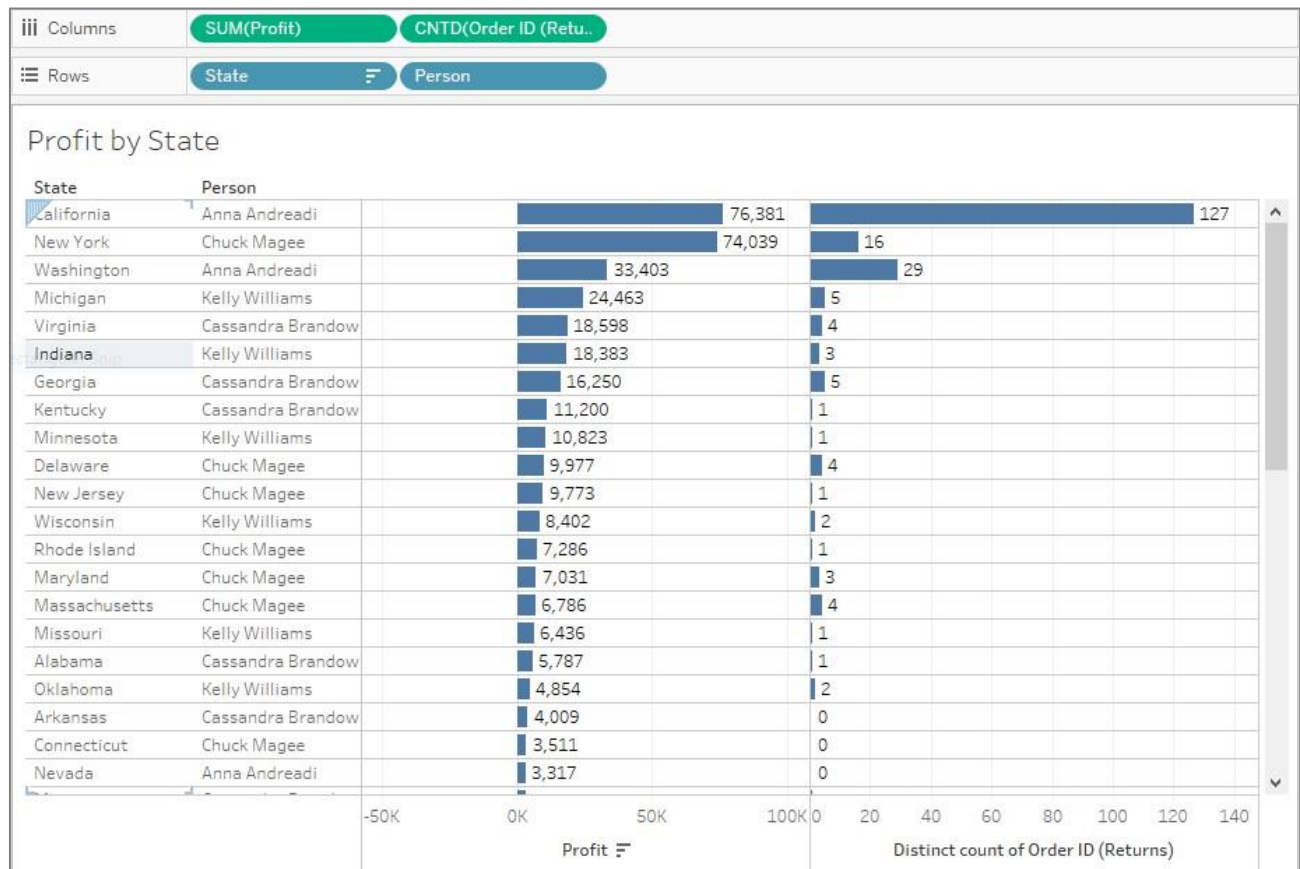
Using the view, answer the following questions:

- Which is the least profitable state?
- Customers from which state have returned the highest number of products?

### Overview:

- Combine the Orders, Returns, and People tables.
- Use inner join to combine Orders and People table.
- Use left join between Orders and Returns table.
- Create the required view.
- Enable mark labels.
- Create an incremental refresh extract using the order date field.

The result should resemble the following image:



## Detailed Instructions

1. Open **Tableau 10.x** (You can go to Start Menu -> All Programs -> Tableau 10.x).
2. On the home page, under **Connect**, under **To a file**, click **Excel**.
3. Browse and connect to the **Sample - Superstore** Excel file.
4. Drag and drop **Orders** and **People** tables to the canvas area.
5. Make sure inner join is selected.
6. Next, place **Returns** table in the canvas area and select **Left Join**.
7. To create a view, go to **Sheet 1**.
8. From Dimensions, locate **State** and **Person** and place on rows.
9. From Measures, locate **Profit** and place on columns.
10. On toolbar, use the **Sort** icon to sort the view in descending order.
11. To count the number of orders placed, locate **Order ID (Returns)** from Dimensions and place on columns.
12. Select aggregation as **Count Distinct**.  
(Right click on Order ID (Returns) in Columns -> Measure -> Count Distinct)

13. To create an extract, under the data pane, right click the data source and select **Extract Data**.
14. In the Extract Data window, select the **Incremental Refresh** check box.
15. From the drop-down, select **Order Date**.
16. Click **Extract**.
17. Specify a path for the extract.
18. Rename the worksheet as **Profit by State**.
19. Save the workbook as **Lesson 3 - Exercises**.

**Keep the workbook open to perform next exercise.**

### **Answers:**

Based on the view, these are the answers to the questions in the problem statement:

- Which is the least profitable state?  
Answer - **Texas**
- Customers from which state have returned the highest number of products?  
Answer - **California**