Create Aggregated Datasets in Oracle Analytics

Before You Begin

This lab shows you how to output multiple datasets from a data flow, add columns, and execute a data flow sequence. You also use the datasets output from executing the sequence to create analyses to verify that your results.

Background

In a data flow, you can add a branch to any node, except the train machine learning model node, to output multiple datasets from a single data flow. You can apply different transformations on each branch to create unique datasets.

In this tutorial, you create a sequence to execute multiple data flows using a specific order in a transaction.

You also create visualizations to verify your results.

What Do You Need?

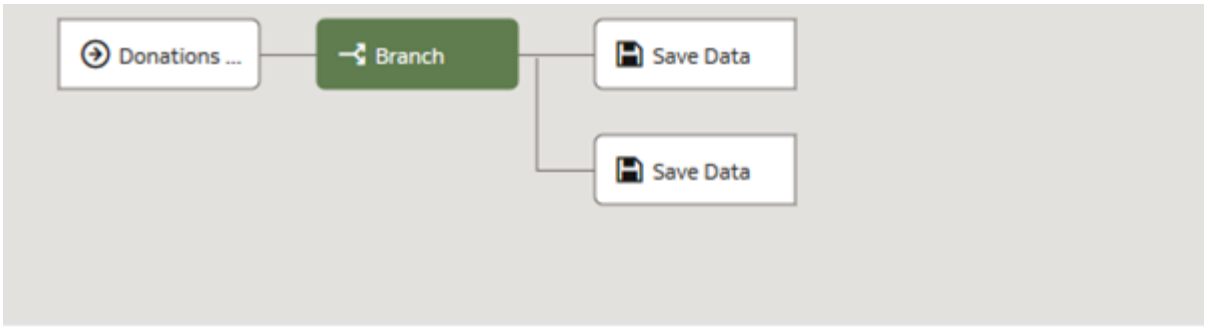
- Access to Oracle Analytics
- Access to the Donations by School dataset

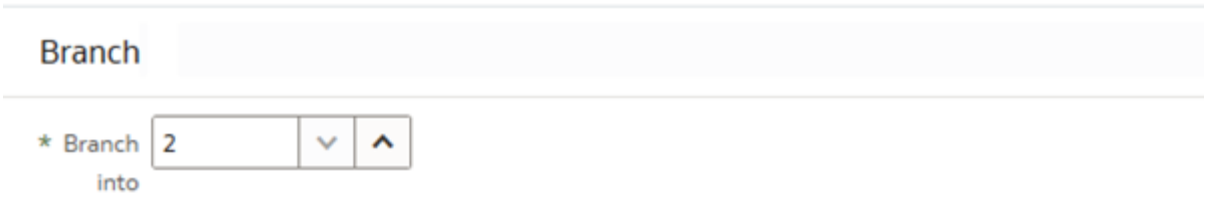
Create Multiple Datasets

In this section, you create a data flow that has a branch step so that when the data flow is executed, you have two distinct datasets.

1. Sign in to Oracle Analytics.
2. On the Home page, click **Create**, and then click **Data Flow**.
3. In Add Dataset, click **Donations by School**, and then click **Add**.
4. In Data Flow Steps, double-click **Branch**.

   The Branch step adds two Save Data steps to the data flow.

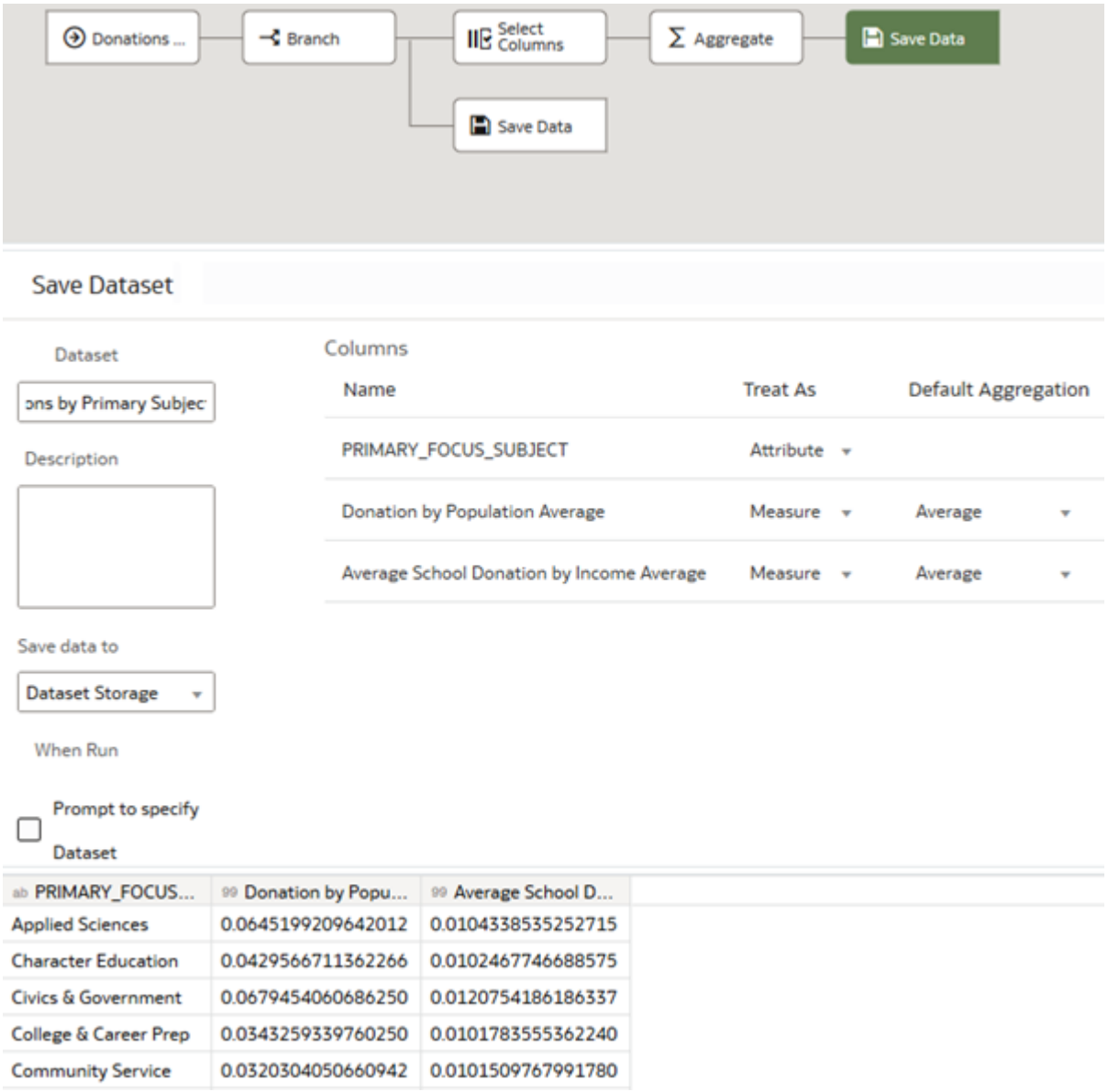




SCHOOL_ID	PRIMARY_FOCU...	RESOURCE_T...	POVERTY_LEVEL
0004ffe3558fd70d939ad522b92447c8	Applied Sciences	Technology	C: High poverty
0004ffe3558fd70d939ad522b92447c8	Literacy	Books	C: High poverty
0004ffe3558fd70d939ad522b92447c8	Literacy	Books	C: High poverty

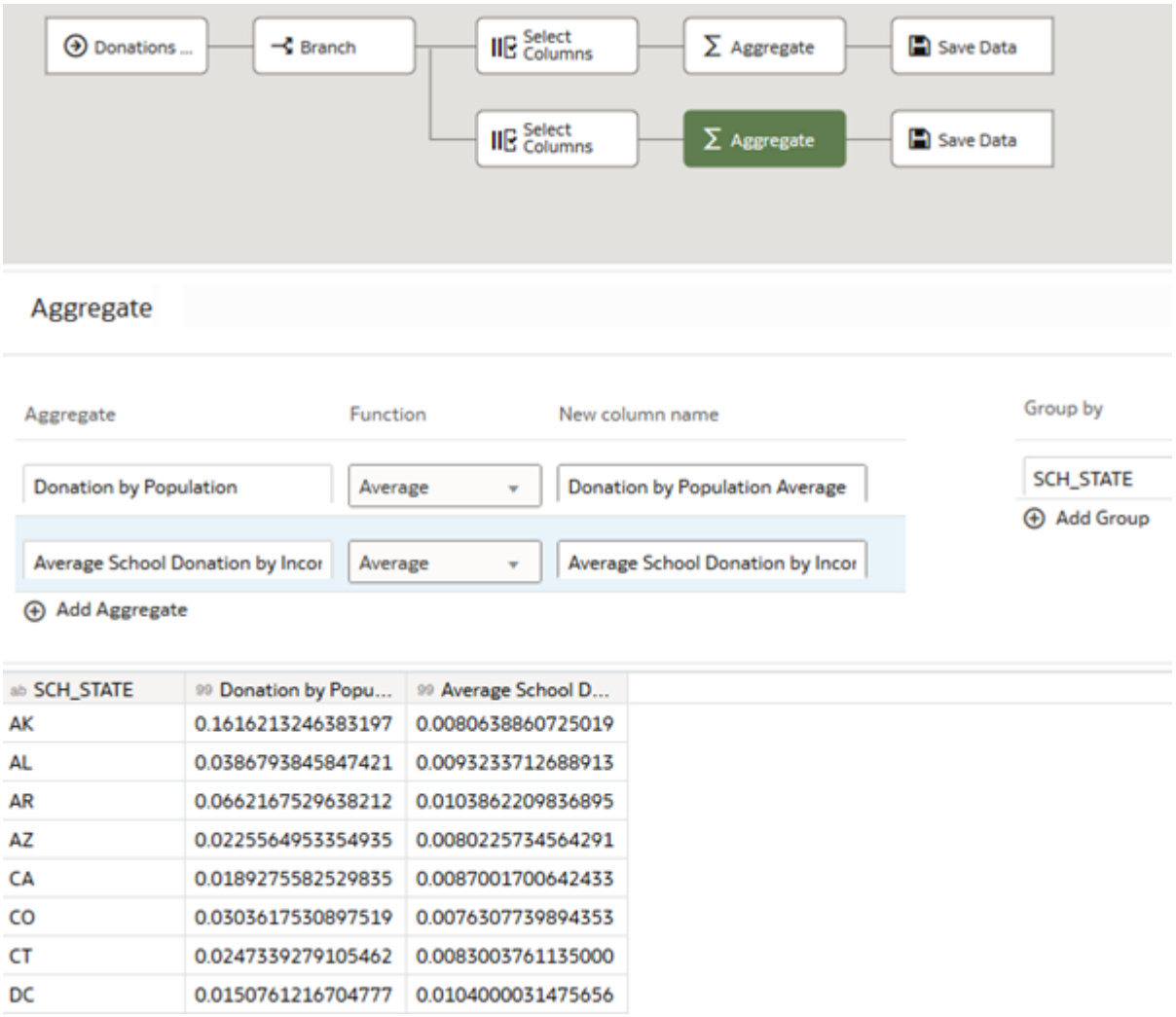
Click **Add a step** on the branch before the first Save Data step, and then select **Select Columns**. In Select Columns, click **Remove all**.

5. Hold down the **Ctrl** key, and select the following:
   - **PRIMARY_FOCUS_SUBJECT**
   - **Donation by Population**
   - **Average School Donation by Income**
6. Click **Add a step** between Select Columns and Save Data, and then click **Aggregate**. In Aggregate, select **Average** under Function in both rows.
7. Click **Save Data** on the first branch. In Save Dataset, enter Donations by Primary Subject Area in **Name**. In Save data to, select **Dataset Storage** to save the data in Oracle Analytics.



Define the Second Branch

1. In the data flow, select the Add a step node on the second branch, and then click **Select Columns**. In Select Columns, click **Remove all**.
2. Hold down the **Ctrl** key, and select the following:
    - **SCH_STATE**
    - **Donation by Population**
    - **Average School Donation by Income**
3. Click **Add a step** ⊕ on the Select Columns node, and select **Aggregate**. In Aggregate, select **Average** from Function in both rows.



SCH_STATE	Donation by Popu...	Average School D...
AK	0.1616213246383197	0.0080638860725019
AL	0.0386793845847421	0.0093233712688913
AR	0.0662167529638212	0.0103862209836895
AZ	0.0225564953354935	0.0080225734564291
CA	0.0189275582529835	0.0087001700642433
CO	0.0303617530897519	0.0076307739894353
CT	0.0247339279105462	0.0083003761135000
DC	0.0150761216704777	0.0104000031475656

4. Click **Save Data** on the second branch. In Save Dataset, enter Donations by State in **Name**. In Save data to, select **Dataset Storage**.
5. Click **Save**. Click **Run Data Flow**. After the data flow completed message appears, click **Go back**.

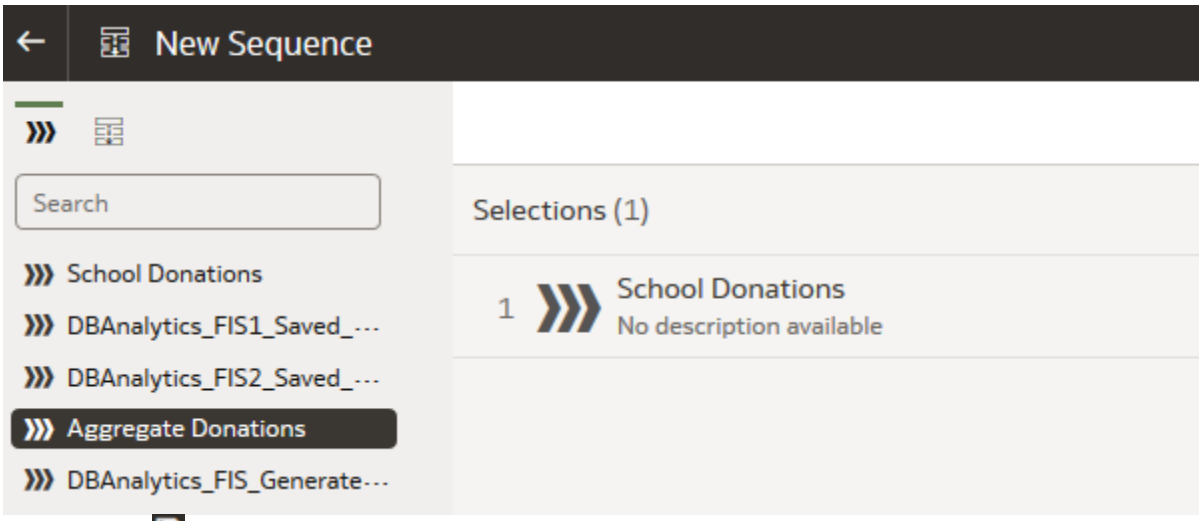
Inspect the Datasets

1. On the Home page, search for the **Donations by Primary Subject Area** dataset. Select the **Donations by Primary Subject Area** dataset, click the **Actions menu**, and then select **Inspect**.
2. In the Donations by Primary Subject Area dataset, click **Data Elements** to view the columns, and then click **Close**.
3. On the Home page, search for or select the **Donations by State** dataset. Select the **Donations by State** dataset, click the **Actions menu**, and then select **Inspect**.
4. In the Donations by State dataset, click **Data Elements** to view the columns, and then click **Close**.

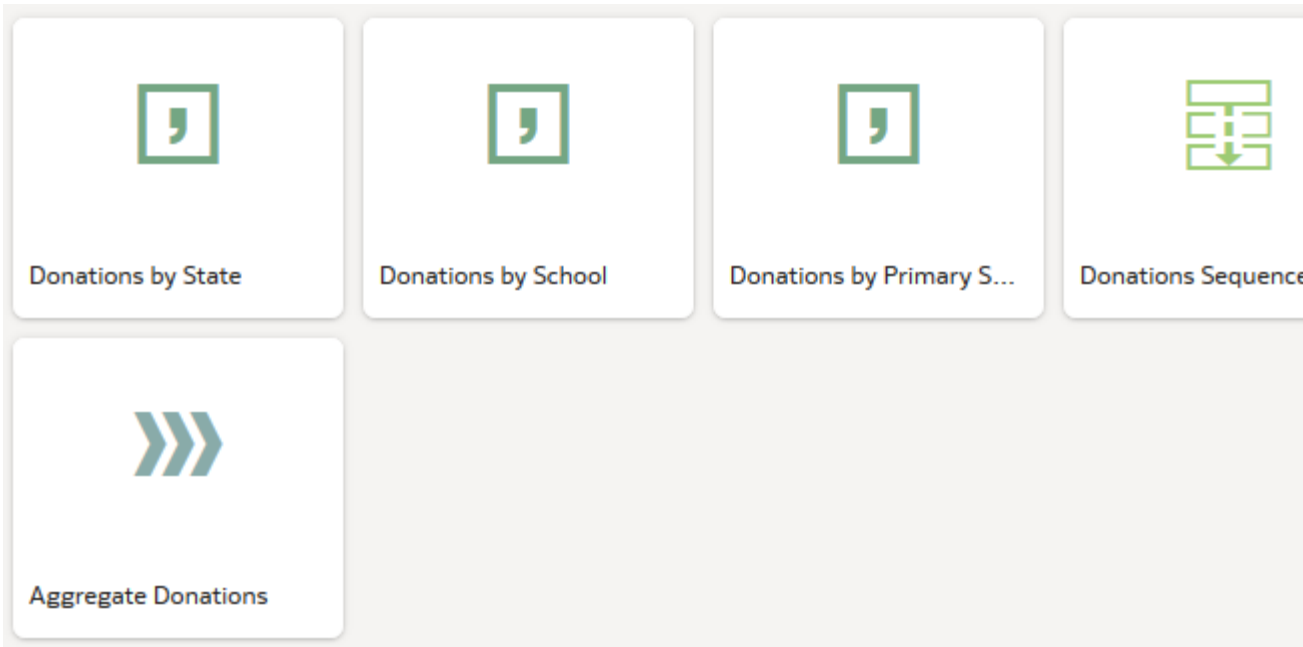
Create a Data Flow Execution Sequence

When you use dynamic data sources the data flows in the sequence, you can create a schedule to run the sequence.

1. On the Home page, click **Create**, and then click **Sequence**.
2. In the Data Flows panel, drag **Donations by School** to Selections.
3. Right-click **Aggregate Donations**, and then select **Add Data Flow**.



4. Click **Save**. In Save Sequence As, enter Donations Sequence in **Name**, and then click **OK**. Click **Execute Sequence**.
5. Click **Go back**. On the Home page, enter Donations and click **Search** to view Donations Sequence and its output.

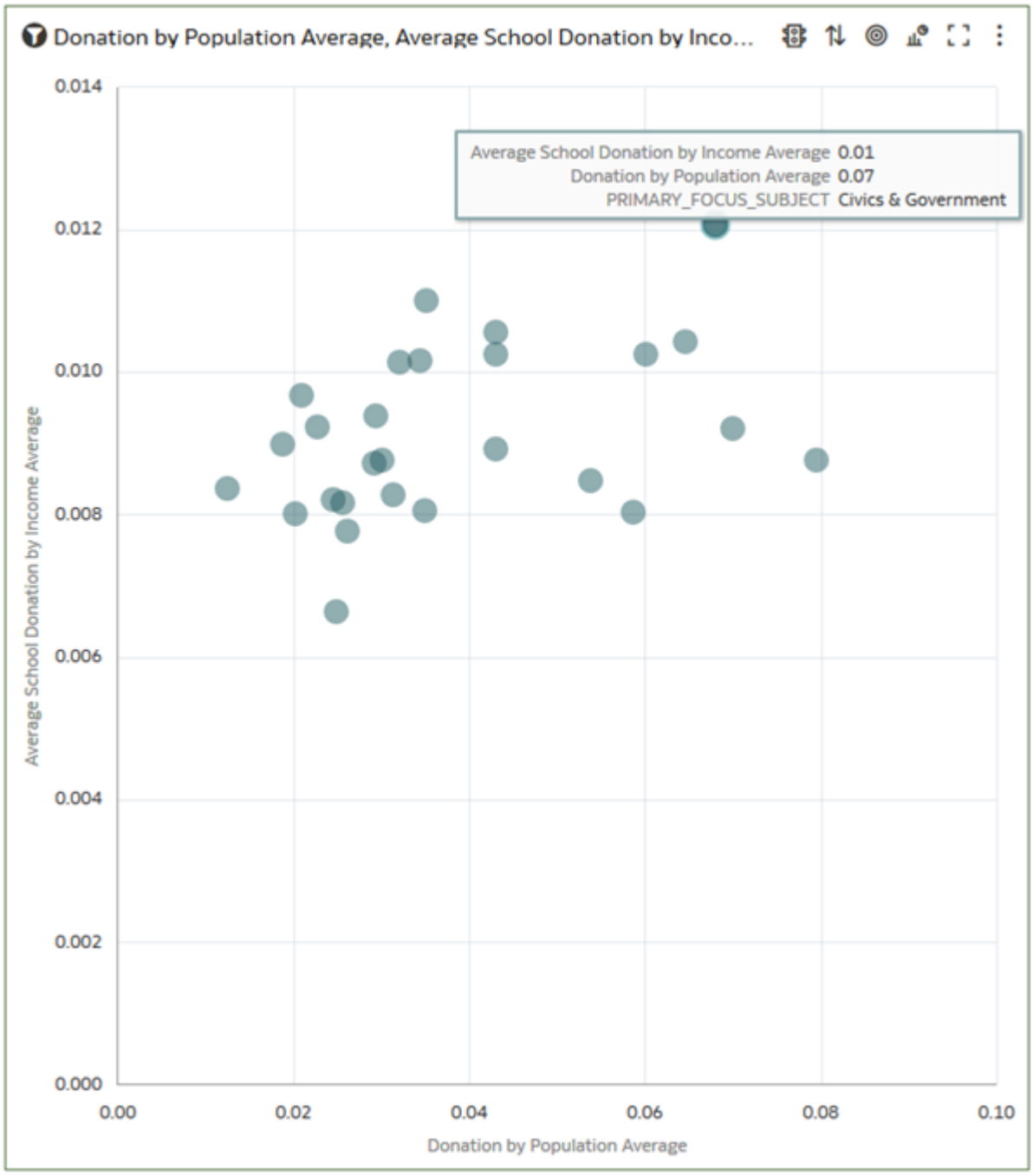


Create Visualizations

1. On the Home page, select **Donations by Primary Subject Area** dataset, click the **Actions menu**, and then select **Create Workbook**.
2. In the Data Elements pane, hold down the **Ctrl** key, and then select the following:

- **PRIMARY_FOCUS_SUBJECT**
- **Donation by Population Average**
- **Average School Donation by Income Average**

3. Drag the data elements to the canvas.



4. Click **Go back** ⇐.
5. On the Home page, select the **Donations by State** dataset, click the **Actions menu** ⋮, and then select **Create Workbook**.
6. In the Data Elements pane, hold down the **Ctrl** key, and then select the following:
   - **SCH_STATE**
   - **Donation by Population Average**
   - **Average School Donation by Income Average**

7. Drag the data elements to the canvas.