Streaming Data Processing: Streaming Analytics and Dashboards

## Overview

Data visualization tools can help you make sense of your BigQuery data and help you analyze the data interactively. You can use visualization tools to help you identify trends, respond to them, and make predictions using your data. In this lab, you use Google Data Studio to visualize data in the BigQuery table populated by your Dataflow pipeline in the previous exercise. You **will not be accessing live streaming data** in this lab.

## Objectives

In this lab, you will perform the following tasks:

* Connect to a BigQuery data source
* Create reports and charts to visualize BigQuery data

This lab uses Google Data Studio to visualize data in BigQuery using the BigQuery connector. In subsequent tasks, you will create a data source, a report, and charts that visualize data in the sample table.

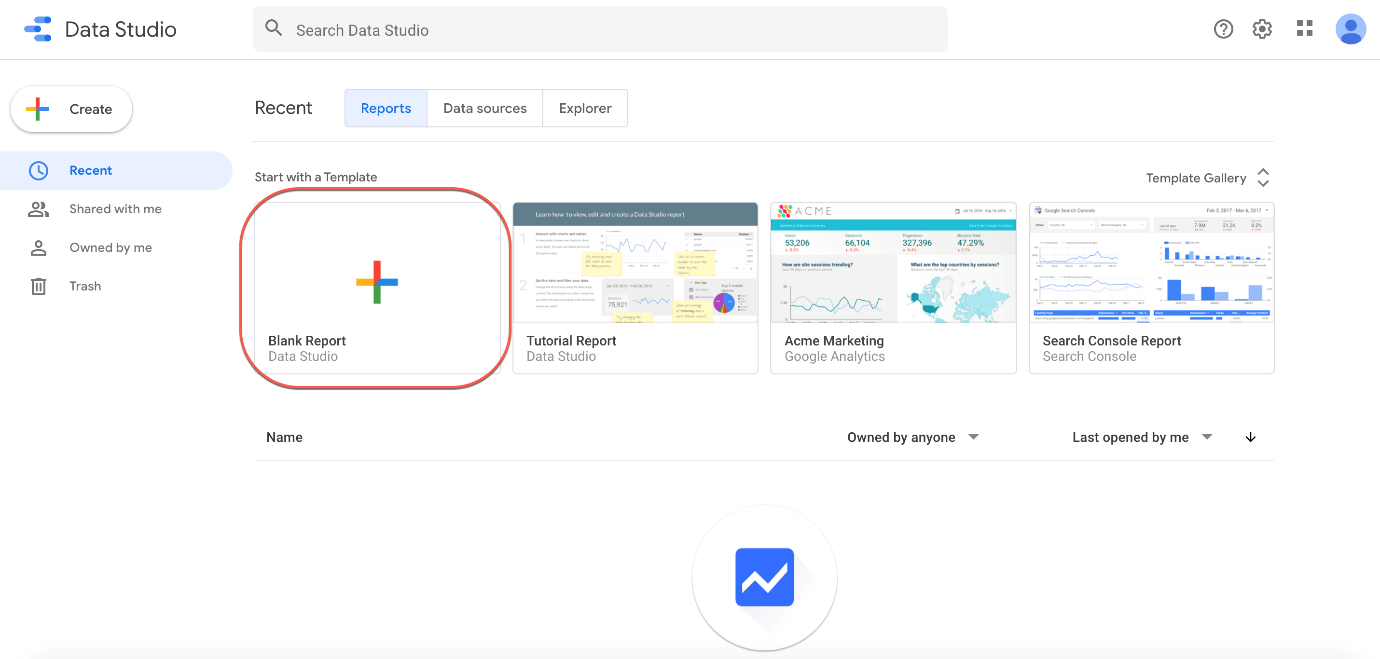
## Task 1: Creating a data source in Data Studio

1. The Google Data Studio User Inteface is accessed from outside of the Google Cloud environment. Open a new browser tab preferably in an incognito window. Navigate to: datastudio.google.com or click on this link: [Google Data Studio](https://datastudio.google.com/)

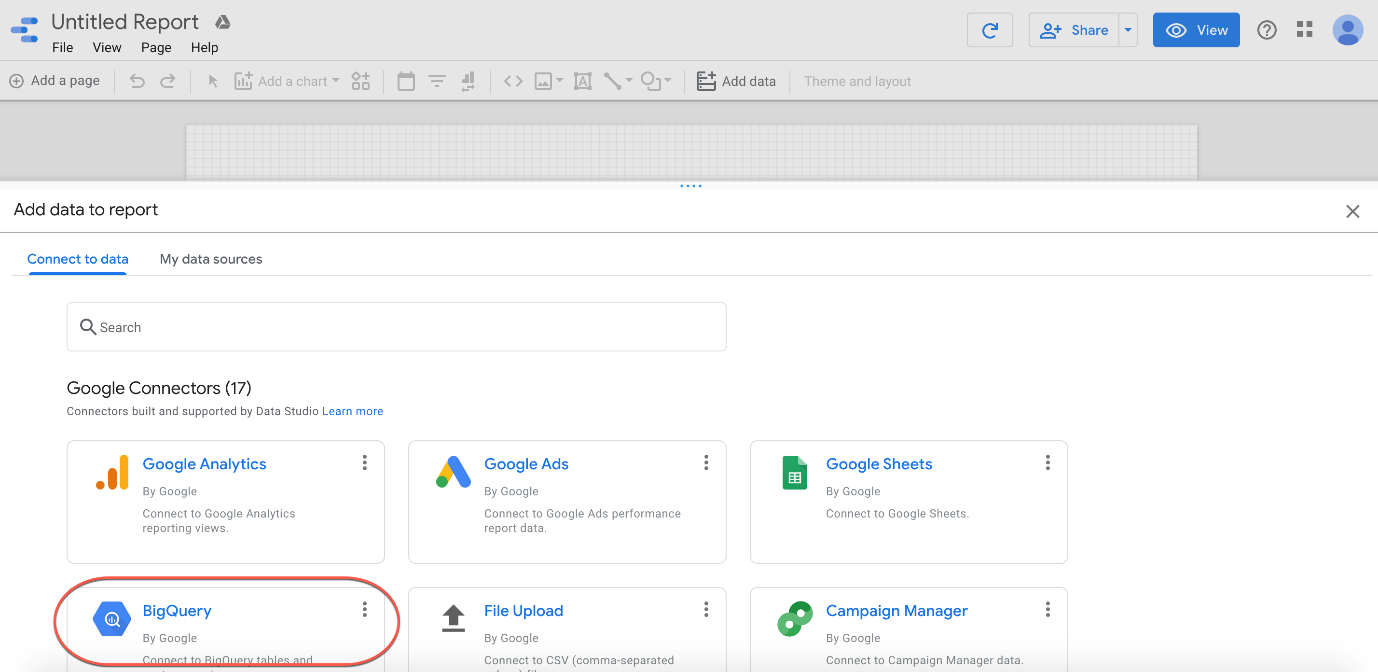
The first step in creating a report in Data Studio is to create a data source for the report. A report may contain one or more data sources. When you create a BigQuery data source, Data Studio uses the BigQuery connector.

You must have the appropriate permissions in order to add a BigQuery data source to a Data Studio report. Lab initialization steps created a BigQuery Dataset and tables for use in your exercises. The permissions applied to BigQuery datasets will apply to the reports, charts, and dashboards you create in Data Studio. When a Data Studio report is shared, the report components are visible only to users who have appropriate permissions at the data level.

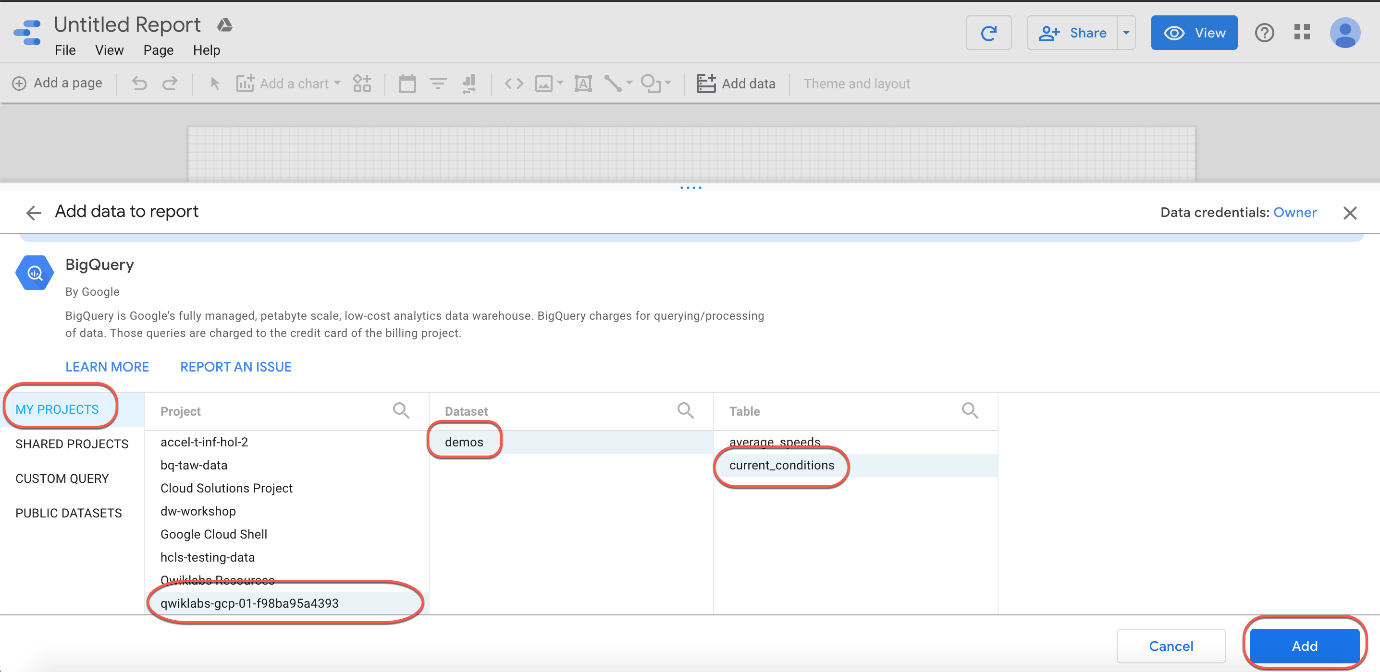
1. On the **Reports** page, in the **Start with a Template** section, click the **Blank Report** template. This starts the account setup process.



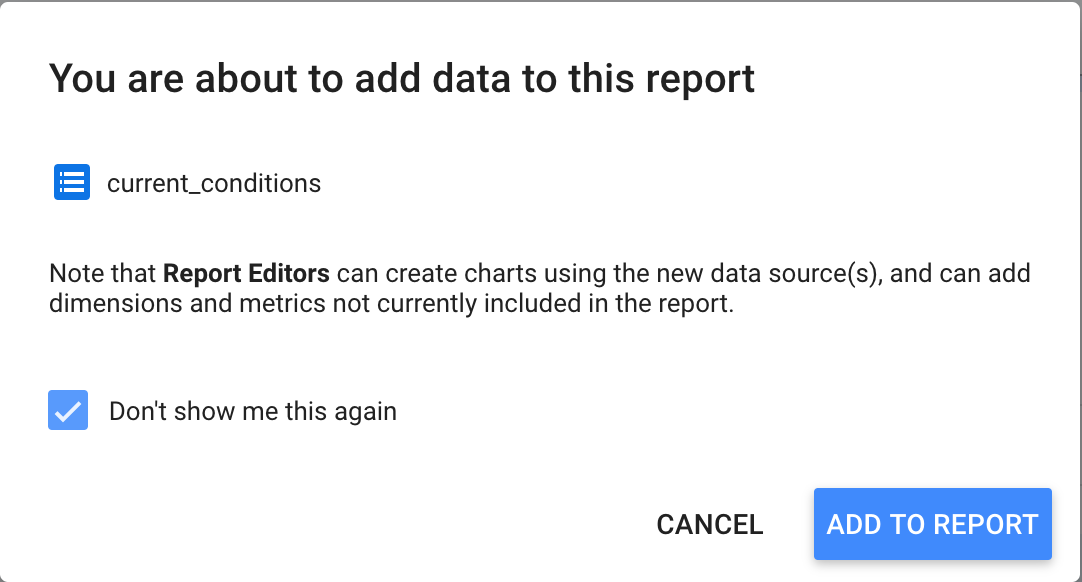
1. To get started, click on the **checkbox** to acknowledge the terms of services and click **Continue**.
2. On the **Preferences** page, select **No** for each option to receive email notifications, and click **Continue**.
3. Now that the account is initialized, you need to start the process again.
4. On the **Reports** page, in the **Start with a Template** section, click the **Blank Report** template. This time it will take you to a new page and begin an **Untitled Report**.
5. The **Add data to report** panel will load from the bottom of the page.
6. In the **Google Connectors** section, select **BigQuery**.



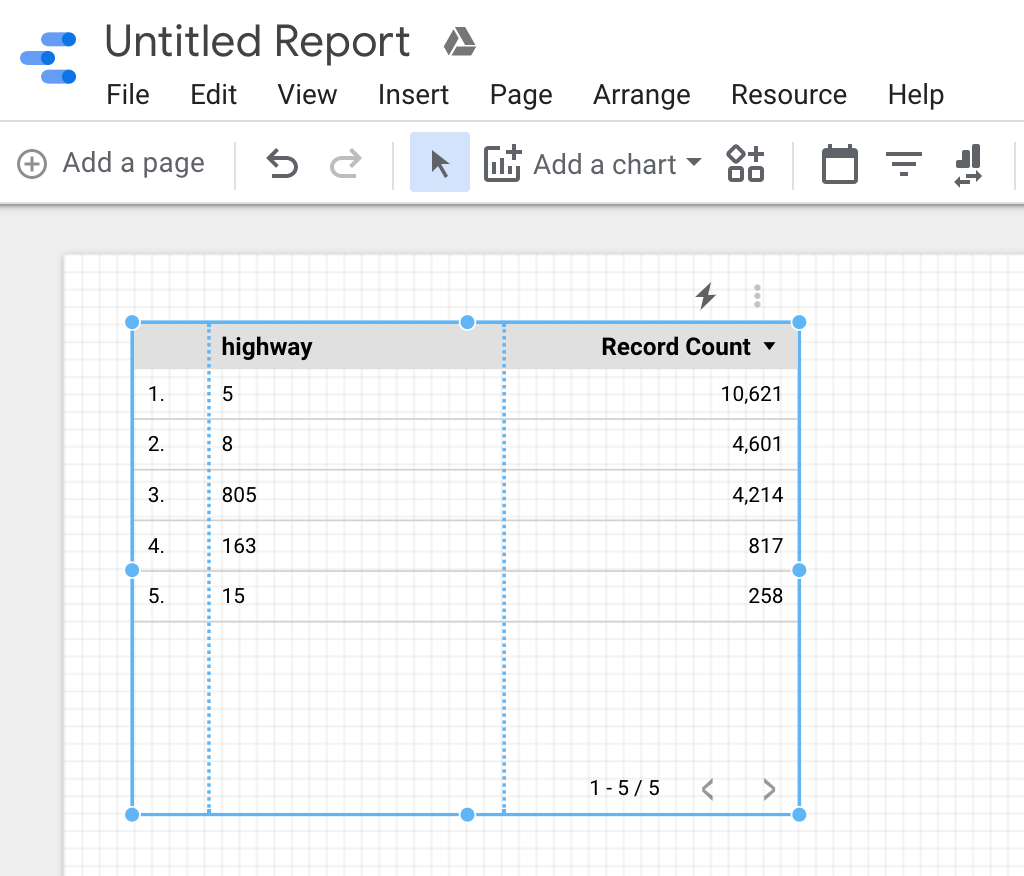
1. Click on **AUTHORIZE** for the notice that "Data Studio requires authorization to connect to your BigQuery projects."
2. If prompted, in the **Sign in** dialog, select your Qwiklabs student account.
3. If prompted, click **ALLOW** to give Google Data Studio permission to view the BigQuery resources in your lab account.
4. Then select **My Projects**.
5. In the **Project** column, click on your project name.
6. In the **Dataset** column, click on **demos**.
7. In the **Table** column, click **current\_conditions**.
8. Lastly click **Add**.



1. You will be prompted with a "You are about to add data to this report" notice. Check "Don't show me this again" and click **ADD TO REPORT**.



1. Once complete, a simple default tabular report will appear. This confirms that you can see your BigQuery data in Data Studio.



Giving Data Studio permission to Google Cloud account resources is typically a first-time activity and not something you would need to do every time you create a report.

Click Check my progress to verify the objective.

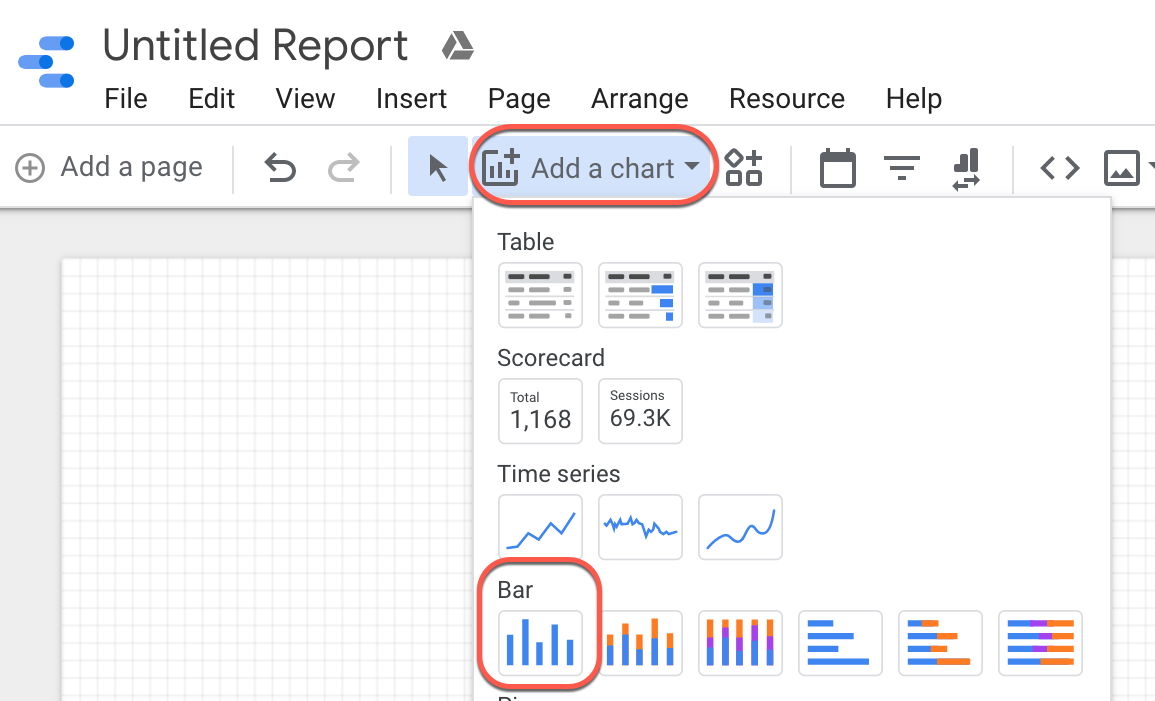
Creating a data source

Check my progress

## Task 2: Creating a bar chart using a calculated field

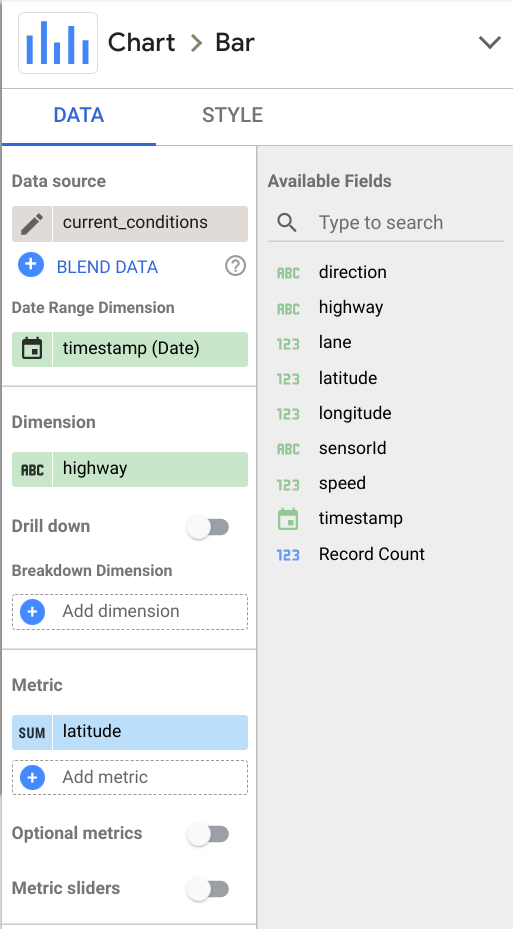
Once you have added the **current\_conditions** data source to the report, the next step is to create a visualization. Begin by creating a bar chart. The bar chart will display the total number of vehicles captured for each highway.

1. (Optional) At the top of the page, click **Untitled Report** to change the report name. For example, type **<PROJECTID>-report1-yourname**.
2. Delete the pre-populated tabluar report. You can do this by simply selecting it and pressing delete.
3. Next, from the **Add a chart** menu select the first **Bar chart**.

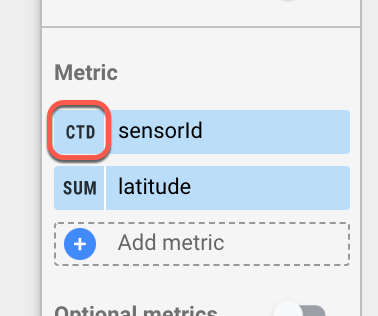
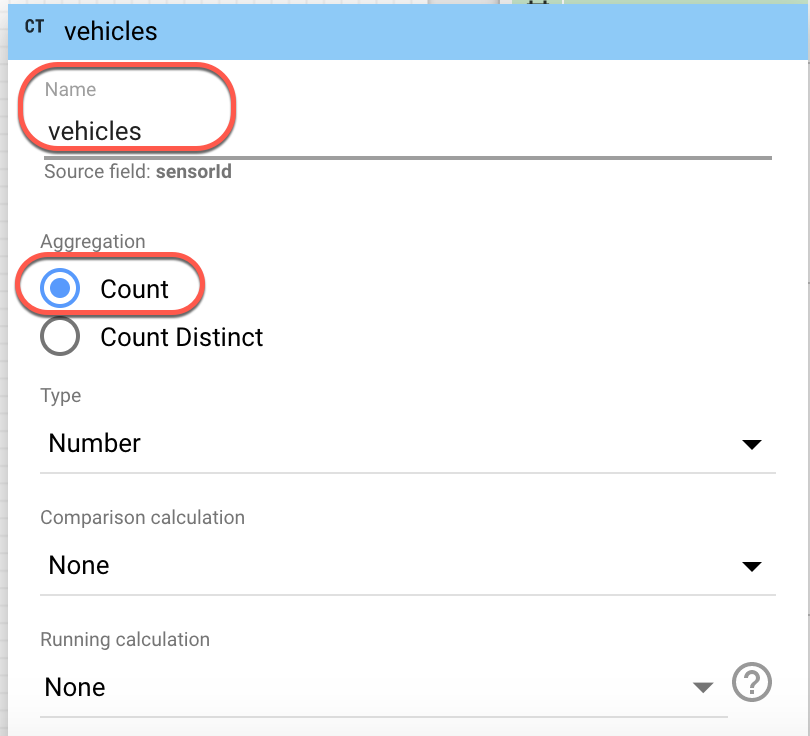


1. In the **Bar chart** properties window, on the **Data** tab, notice the value for **Data Source** (**current\_conditions**) and the default values for **Dimension** and **Metric**.
2. If **Dimension** is not set to **highway**, then change **Dimension** to **highway**. In the **Dimension** section, click the existing dimension and in the **Dimension picker**, select **highway**.
3. In the **Metric** section, click **Add metric** and add **latitude**.
4. In the **Metric** section, mouse over **Record Count** and click the **(x)** to remove it.

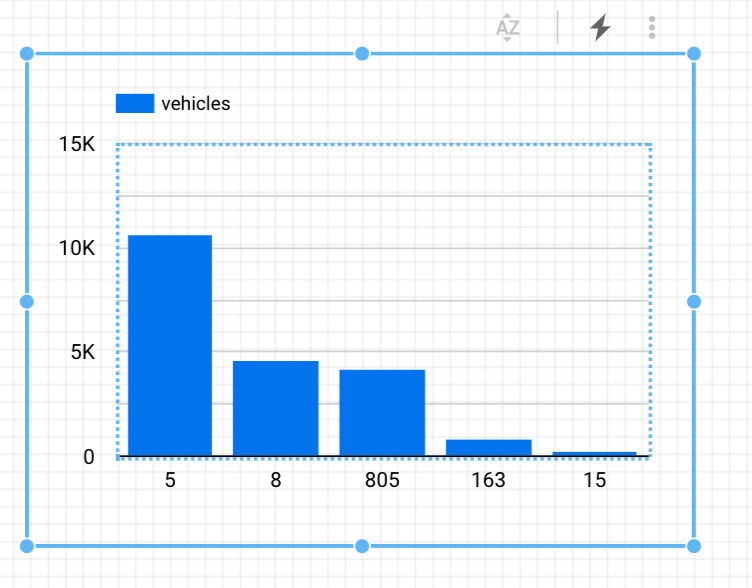
**Example**:



1. To gain insight on vehicle volume you need to add a metric for each vehicle detected.
2. In the **Metric** section, click **Add metric** and add **sensorId**.
3. A count distinct for this column is automatically created. This metric set as a count distinct does not give you a true sense of traffic volume. Click on the **CTD** text and on the popup window choose **Count** . Type the name **vehicles** in the name box. Click in the report space off the popup to close it. The change is saved automatically.

1. In the **Metric** section, mouse over **latitude** and click the **(x)** to remove it.
2. The Dimension should be set to highway and the Metric should be set to sensorId. Notice the chart is sorted in Descending order by default. The highway with the most vehicles is displayed first.



1. To enhance the chart, change the bar labels. In the **Bar chart** properties window, click the **STYLE** tab.
2. In the **Bar chart** section, check **Show data labels**.
3. The total number of vehicles is displayed above each bar in the chart.

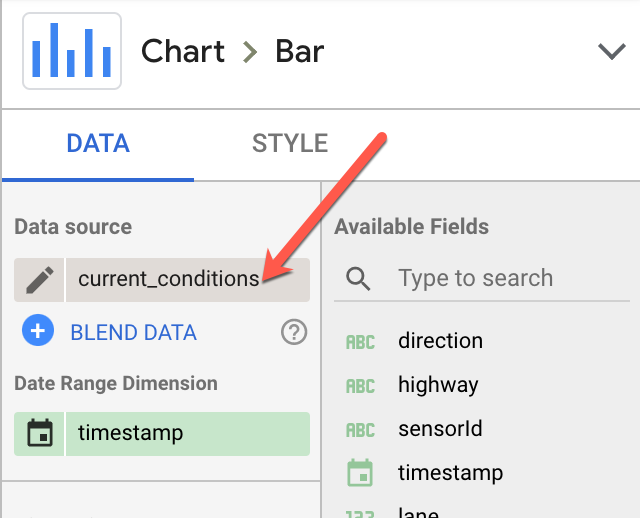
## Task 3: Creating a chart using a custom query

You may find that it is easier to work with an existing query to produce the desired reports and visualizations in Data Studio. The Custom Query option lets you leverage BigQuery's full query capabilities such as joins, unions, and analytical functions.

Alternatively, you can leverage BigQuery's full query capabilities by creating a view. A view is a virtual table defined by a SQL query. You can query data in a view by adding the dataset containing the view as a data source.

When you specify a SQL query as your BigQuery data source, the results of the query are in table format, which becomes the field definition (schema) for your data source. When you use a custom query as a data source, Data Studio uses your SQL as an inner select statement for each generated query to BigQuery. For more information on custom queries in Data Studio, consult the [online help](https://support.google.com/datastudio/?hl=en#topic=6267740).

1. To add a bar chart to your report that uses a custom query data source:
2. From the **Add a chart** menu select the first **Bar chart**.
3. In the **Bar chart** properties window, on the **Data** tab, notice the value for Data Source (current\_conditions) and the default values for Dimension and Metric are the same as the previous chart. In the **Data Source** section, click on the **current\_conditions** data source. At the bottom of the pane choose **ADD DATA**.



1. Under **Google Connectors**, select **BigQuery**.
2. Select **CUSTOM QUERY** in the first grouping.
3. For **Billing Project**, select your project.
4. Type the following in the **Enter custom query** window and replace the <PROJECTID> with your Project ID.

SELECT max(speed) as maxspeed, min(speed) as minspeed,

avg(speed) as avgspeed, highway

FROM `<PROJECTID>.demos.current\_conditions`

group by highway

Copied!

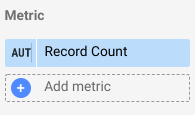
content\_copy

This query uses max/min/avg functions to give you the corresponding speed for each highway.

1. Click **ADD**.
2. When prompted, click **ADD TO REPORT**.

Data Studio may be unable to determine the appropriate Dimension and Metrics for the chart. This requires you to adjust the graph options.

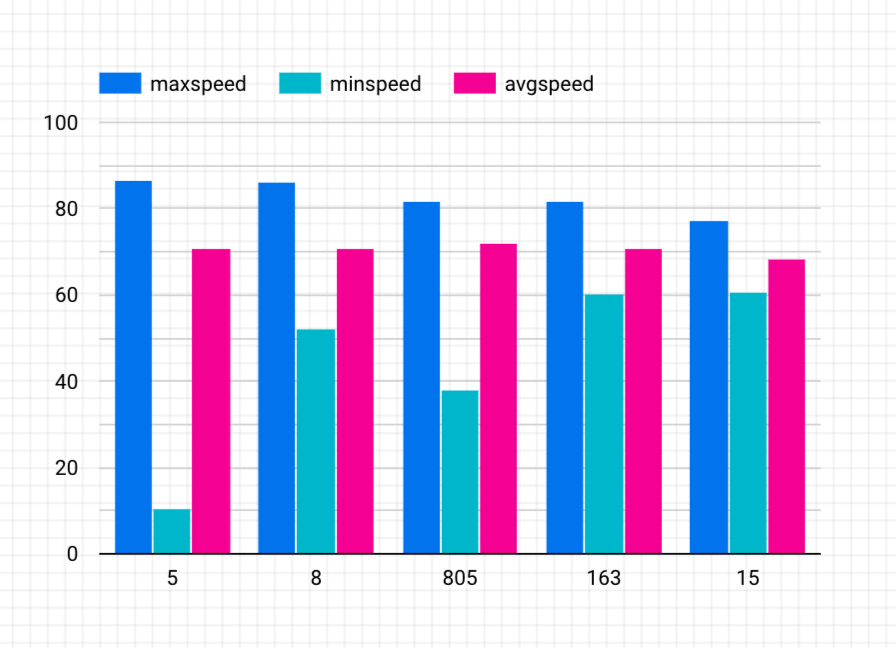
1. In the **Bar chart** properties, on the **Data** tab, in the **Metric** section, click **Record count**.



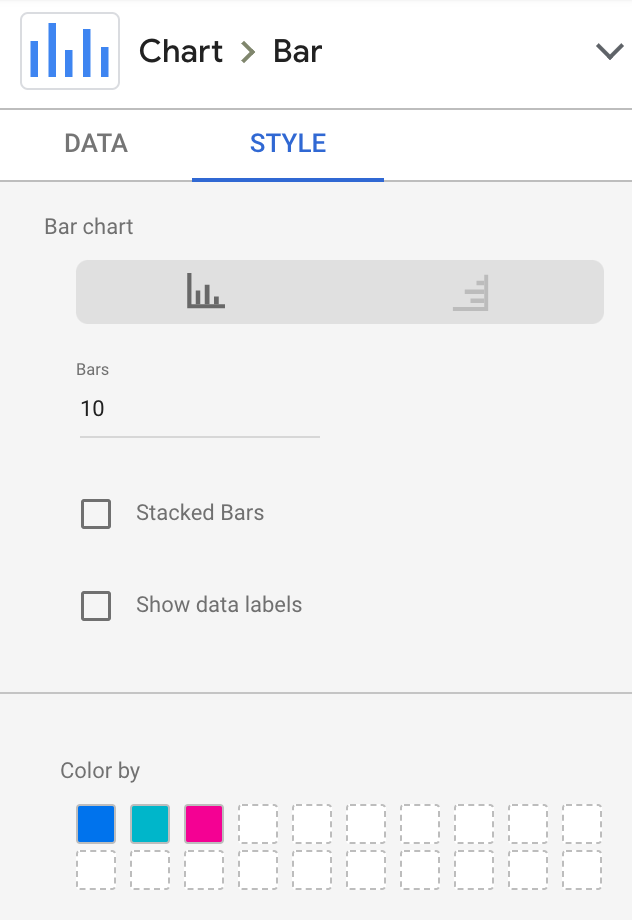
1. In the **Metric picker,** select **maxspeed**.
2. In the **Metric** section, click **Add metric**.
3. In the **Metric picker**, select **minspeed**.
4. In the **Metric** section, click **Add metric**.
5. In the **Metric picker**, select **avgspeed**.
6. Remove the metric other than maxspeed, minspeed and avgspeed, if exist.

Your chart now displays the maximum speed, minimum speed, and average speed for each highway.

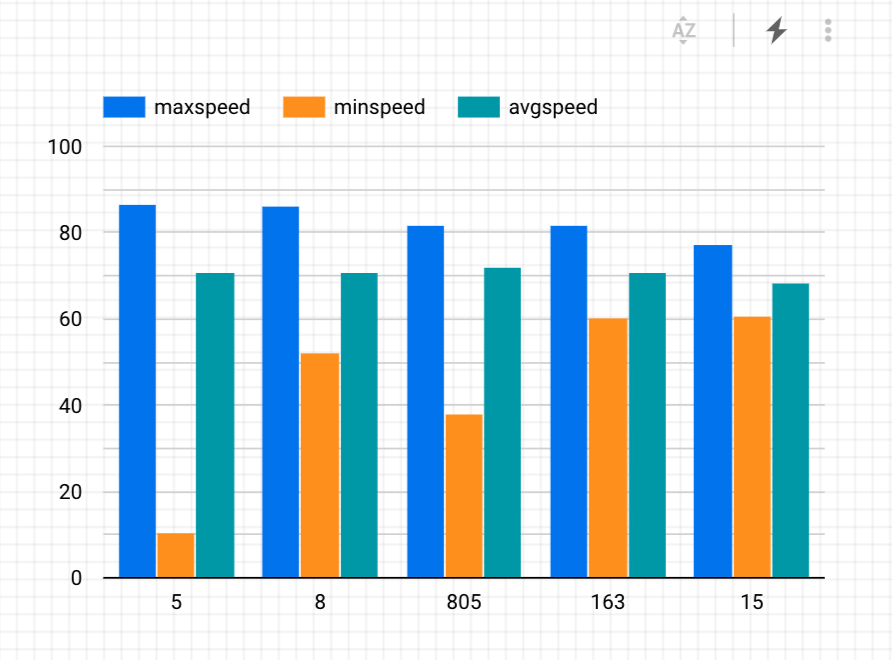
Notice each bar has a default color based on the order the metrics were added to the chart.



1. For readability, change the chart styles. In the **Bar chart** properties, click the **Style** tab.



1. In the **Color By** section, click on the boxes to select different colors.



Click Check my progress to verify the objective.

Create a chart using a custom query

Check my progress

## Task 4: Viewing your query history

You can view queries submitted via the BigQuery Connector by examining your query history in the BigQuery web interface. Using the query history, you can estimate query costs, and you can save queries for use in other scenarios.

### **Open BigQuery Console**

1. In the Google Cloud Console, select **Navigation menu** > **BigQuery**.

The **Welcome to BigQuery in the Cloud Console** message box opens. This message box provides a link to the quickstart guide and lists UI updates.

1. Click **Done**.
2. In the bottom panel the first item in the list will be **Personal history**. On your initial visit to the page the query history should appear on the bottom right underneath the Query editor pane. If it is not loaded click the **Personal history** link.
3. The list of queries is displayed with the most recent queries first. Click on any Query to view details on the query such as Job ID and Bytes Processed.



## End your lab