



databricks

Academy

DAWD 03-3 – Lab – Data Visualizations and Dashboards

The two fields below are used to customize queries used in this course. Enter your schema (database) name and username, and press "Enter" to populate necessary information in the queries on this page.

Schema Name:

Username:



Lesson Objective

At the end of this lesson, you will be able to:

- Create basic visualizations using Databricks SQL
 - Create a dashboard using multiple existing visualizations
-



Create a Boxplot

We can create a boxplot to highlight outliers in our data and to provide other information like quartiles, minimum, maximum, etc.

Complete the following:

1. Open a new query tab, and run the query below
2. Save the query as "Boxplot"
3. In the query results section, hover over the "+" symbol and click "Visualization"
4. Select "Box" as the visualization type
5. For "X Column" select "product_category"
6. For "Y Column" select "total_price"
7. Click the "X Axis" tab
8. Enter "Product Category" in the "Name" field to make more user friendly
9. Click the "Y Axis" tab
10. Enter "Price" in the "Name" field to make it more user friendly

Note that there are many outliers. Although we would not want to ignore these, we can hide them by completing the following step:

11. Enter "30000" for the "End Value" field
12. Click "Save" in the lower-right corner
13. Change the name of the visualization to "Price by Product"
14. Make sure the query is Saved
15. Check your work by entering your answer to the question below
16. After pressing ENTER/RETURN, green indicates a correct answer, and red indicates incorrect

Note that you can hover your mouse over the boxplot to get additional information about each plot.

```
USE hive_metastore.class_013_odg7_da_dawd;  
SELECT * FROM sales;
```

[Show Answer](#)[Copy](#)

Which product category has the highest upper fence?



Create a Funnel

The Funnel visualization shows the number of people as they progress from one step to another. In this portion of the lab, we are going to make a funnel that shows the number of customers who have progressed from sales orders to actual sales.

Complete the following:

1. Open a new query tab, and run the query below
2. Save the query as "Funnel"
3. In the query results section, hover over the "+" symbol and click "Visualization"
4. Select "Funnel" as the visualization type
5. For "Step Column" select "Step"
6. For "Value Column" select "Value"

That's it! Most of the setup in this visualization is contained in the query. The query uses a CTE and two UNION ALLS to generate the data. Really, all that's happening is we are selecting distinct customer IDs in each of the tables to determine which customers have been active in generating sales orders and sales.

7. Click "Save" in the lower-right corner
8. Click inside the tab and change the name of the visualization to "Customer Funnel"
9. Make sure the query is Saved
10. Check your work by entering your answer to the question below
11. After pressing ENTER/RETURN, green indicates a correct answer, and red indicates incorrect

```
USE hive_metastore.class_013_odg7_da_dawd;
WITH funnel AS (
  SELECT DISTINCT "Customers" AS Step, string(customer_id) FROM customers
  UNION ALL
  SELECT DISTINCT "Orders" AS Step, customer_id FROM sales_orders
  UNION ALL
  SELECT DISTINCT "Sales" AS Step, customer_id FROM sales
)
SELECT Step, count(customer_id) as Value
FROM funnel
GROUP BY Step;
```

[Show Answer](#)[Copy](#)

What is the %Max for Orders? (include the %)

 **Create a Dashboard**

Let's combine the visualizations we just made into a dashboard.

Complete the following:

1. Click "Dashboards" in the sidebar menu
2. Click the blue "Create Dashboard" button
3. Name the dashboard "Company Information" and click the "Save" button
4. Click the blue "Add" button, and click Visualization"
5. In the dropdown, select "Boxplot"
6. Click the drop down under "Select existing visualization" and select "Price by Product" (which you created earlier)
7. Change "Title" to "Prices"
8. Click the blue "Add to dashboard" button, and the Boxplot appears in your dashboard
9. Repeat steps 4-8 with the "Funnel" query ("Customer Funnel")
10. Click "Add", then "Text Box", and type "# Miscellaneous Information" in the "Edit:" field
11. Click "Add to dashboard"
12. Optional: Move the visualizations around by clicking and dragging each one
13. Optional: Resize each visualization by dragging the lower-right corner of the visualization
14. Optional: Click "Colors" to change the color palette used by visualizations in the dashboard
15. Click "Done Editing" in the upper-right corner
16. Run every query and refresh all visualizations all at once by clicking "Refresh"
17. Check your work by entering your answer to the question below
18. Hint: Hover over each Product Category to get the statistics in the Boxplot Chart
19. After pressing ENTER/RETURN, green indicates a correct answer, and red indicates incorrect

Which product category has the lowest minimum price?

© 2023 Databricks, Inc. All rights reserved.

Apache, Apache Spark, Spark and the Spark logo are trademarks of the [Apache Software Foundation](https://www.apache.org/).

[Privacy Policy](#) | [Terms of Use](#) | [Support](#)

1.2.13