

DAD 220 Module Five Major Activity Template

Overview

Review the scenario for this activity in the guidelines and rubric. Then complete the steps below as you work through the directions for this activity. Replace the bracketed text with your screenshots and responses to the prompts. Size each screenshot and its explanation to fit approximately one-quarter of the page with the description written below the screenshot. Review the Template Screenshot Example linked in the guidelines and rubric for this assignment to see an example of how screenshots for your assignment should look.

Create a Database

- **Write SQL commands** that capture specific, usable data that can be used in your analysis.
 - **Reference notes for this step:** Import the CSV file into MySQL Table. Use the following line terminators when importing: `\r\n`. Do not use IGNORE 1 LINES for data that does not have column headers in the first row.
- **Analyze the results of queries** to identify specific information that can be presented in your summary.
 - Sales by region:
 - Analyze sales data by state to determine where the company has the largest customer base.

-Massachusetts has the largest customer base with 982.

```

mysql> select count(*), State from customers Group by State Order by count(*);
+-----+-----+
| count(*) | State |
+-----+-----+
| 982      | MA |
| 1        | NY |
| 1        | NC |
| 1        | FL |
| 1        | CA |
+-----+-----+
52 rows in set (0.05 sec)

mysql>

```

Here I analyzed sales data by state to determine what state has the largest customer base.

- Analyze the data to determine the top three products sold in the United States.

- Top 3 products sold in the United States are:

- 8385 - BAS-48-1 C
- 6186 - ENT-48-40F
- 4329 - ENT-48-10F

The screenshot shows a Codio IDE interface. The left pane displays the file tree for a project named 'DAD-220 SQL Lab Environment'. The middle pane shows a list of 52 US states with their corresponding IDs. The right pane shows the terminal output of a SQL query.

Filetree: JCASSELLO, DAD-220 SQL Lab Environment

- .settings
- create
- CREATE
- customers.csv
- describe
- FleetMaintenanceRecords.csv
- HRandIS-Employees.csv
- MajorLabActivity4.csv
- mysqlsampledatabase.sql
- orders.csv
- quit
- README.md
- rma.csv
- use
- USE

FleetMaintenance... Terminal

713	South Carolina
724	Georgia
725	New Jersey
728	Colorado
731	Nebraska
732	Kansas
746	Ohio
755	Nevada
756	Oklahoma
757	Texas
758	Illinois
759	Maine
759	Utah
760	Michigan
770	California
771	New Hampshire
772	Florida
773	Rhode Island
776	North Dakota
777	Maryland
782	Missouri
784	Arizona
785	North Carolina
791	Washington
791	New York
792	Hawaii
793	Vermont
795	Wyoming
800	Louisiana
801	Minnesota
807	South Dakota
811	Indiana
811	Pennsylvania
814	Iowa
818	New Mexico
819	Kentucky
820	Montana
822	Wisconsin
824	Delaware
830	Connecticut
831	Tennessee
834	Mississippi
838	Idaho
838	Alabama
842	Oregon
843	West Virginia
854	Arkansas
982	Massachusetts

52 rows in set (0.04 sec)

```
mysql> select count(*), SKU from Orders group by SKU order by count(*) desc;
```

count(*)	SKU
8385	BAS-48-1 C
6186	ENT-48-40F
4329	ENT-48-10F
4285	BAS-08-1 C
4275	ENT-24-10F
4178	ADV-24-10C
4174	ADV-48-10F
2152	ENT-24-40F
34	BAS-24-1 C

9 rows in set (0.09 sec)

```
mysql>
```

Here I analyzed the data to determine the top three products sold in the United States.

- Analyze the data to determine the top three products sold in the southeastern region of the United States.
 - Southeastern states to include in your analysis: Virginia, North Carolina, South Carolina, and Georgia

-The top 3 products sold in the southeastern region of the United States are:

504 - BAS-48-1 C

337 - ENT-48-40F

257 - BAS-08-1 C

```
C Code Project File Edit View Tools Education Help Configure ... Project: test10000 ... Configure ...
DAD-220 SQL Lab Environment
mysql> select count(*), SKU from Orders group by SKU order by count(*) desc;
+-----+-----+
| count(*) | SKU |
+-----+-----+
| 8282 | BAS-48-1 C |
| 6118 | ENT-48-40F |
| 4287 | ENT-48-10F |
| 4285 | BAS-08-1 C |
| 4270 | ENT-24-10F |
| 4178 | ADV-24-10C |
| 4174 | ADV-48-10F |
| 2152 | ENT-24-40F |
| 34 | BAS-24-1 C |
+-----+-----+
9 rows in set (0.04 sec)

mysql> select count(SKU), SKU from Orders left join customers on Orders.CustomerID=customers.CustomerID where State = 'Virginia' or State = 'North Carolina' or State = 'South Carolina' or State = 'Georgia' group by SKU order by count(SKU) desc;
+-----+-----+
| count(SKU) | SKU |
+-----+-----+
| 504 | BAS-48-1 C |
| 337 | ENT-48-40F |
| 257 | BAS-08-1 C |
| 265 | ADV-08-10F |
| 247 | ENT-48-10F |
| 243 | ADV-24-10C |
| 235 | ENT-24-10F |
| 143 | ENT-24-40F |
| 1 | BAS-24-1 C |
+-----+-----+
9 rows in set (0.06 sec)

mysql>
```

Here I analyzed the top three products sold in the southeastern region of the United States.

- Returns by region:
 - Analyze the data to determine the top three products returned in the United States.

-The top 3 products returned in the United States are:

8282 - BAS-48-1 C

6118 - ENT-48-40F

4287 - ENT-48-10F

```

mysql> select count(*), SKU from Orders group by SKU order by count(*) desc;
+-----+-----+
| count(*) | SKU |
+-----+-----+
| 6380 | BAS-48-1 C |
| 6180 | ENT-48-40F |
| 4220 | ENT-08-10F |
| 4205 | BAS-08-1 C |
| 4270 | ENT-24-10F |
| 4170 | ADW-24-10C |
| 4174 | ADW-48-10F |
| 2102 | ENT-24-40F |
| 34 | BAS-24-1 C |
+-----+-----+
9 rows in set (0.09 sec)

mysql> select count(SKU), SKU from Orders left join customers on Orders.CustomerID=customers.CustomerID where State = 'Virginia' or State = 'North Carolina' or State = 'South Carolina' or State = 'Georgia' group by SKU order by count(SKU) desc;
+-----+-----+
| count(SKU) | SKU |
+-----+-----+
| 504 | BAS-48-1 C |
| 337 | ENT-48-40F |
| 257 | BAS-08-1 C |
| 250 | ADW-48-10F |
| 247 | ENT-48-10F |
| 243 | ADW-24-10C |
| 236 | ENT-24-10F |
| 143 | ENT-24-40F |
| 1 | BAS-24-1 C |
+-----+-----+
9 rows in set (0.06 sec)

mysql> select Orders.SKU, count(*) from orders inner join RM on RM.OrderID = Orders.ID group by Orders.SKU order by count(*) desc;
ERROR 1140 (42002): Table 'QuantificationRM.orders' doesn't exist

mysql> select Orders.SKU, count(*) from orders inner join RM on RM.OrderID = Orders.OrderID
-> group by Orders.SKU
-> order by count(*) desc;
+-----+-----+
| SKU | count(*) |
+-----+-----+
| BAS-48-1 C | 6382 |
| ENT-48-40F | 6118 |
| ENT-48-10F | 4287 |
| BAS-08-1 C | 4268 |
| ENT-24-10F | 4231 |
| ADW-48-10F | 4174 |
| ADW-24-10C | 4122 |
| ENT-24-40F | 3123 |
| BAS-24-1 C | 33 |
+-----+-----+
9 rows in set (0.10 sec)

mysql>

```

Here I analyzed the top three products returned in the United States.

- Analyze the data to determine the top three products returned in the northwestern region of the United States.
 - Northwestern states to include in your analysis:
Washington, Oregon, Idaho, and Montana

- The top 3 products returned in the northwestern region of the United States are:

697 - BAS-48-1 C

534 - ENT-48-40F

379 - BAS-08-1 C

```

mysql> select count(Orders.SKU), count(*) from Orders inner join RMK on RMK.OrderID = Orders.OrderID group by Orders.SKU order by count(*) desc;
ERROR 1146 (42S02): Table 'QuantigrationRMK.orders' doesn't exist

mysql> select count(Orders.SKU), count(*) from Orders inner join RMK on RMK.OrderID = Orders.OrderID
    > group by Orders.SKU
    > order by count(*) desc;
+-----+-----+
| SKU          | count(*) |
+-----+-----+
| BAS-48-1 C   | 8282     |
| ENT-48-40F   | 4138     |
| ENT-48-10F   | 4287     |
| BAS-08-1 C   | 4249     |
| ENT-24-10F   | 4231     |
| ENT-48-10F   | 4124     |
| ADV-24-10C   | 4122     |
| ENT-24-40F   | 2122     |
| BAS-24-1 C   | 53       |
+-----+-----+
8 rows in set (0.18 sec)

mysql> select count(Orders.SKU), SKU from Orders left join customers on Orders.CustomerID = customers.CustomerID where state = 'Washington' or State = 'Oregon' or State = 'Idaho' or State = 'Montana' group by SKU order by count(Orders.SKU) desc;
+-----+-----+
| count(Orders.SKU) | SKU          |
+-----+-----+
| 540               | BAS-48-1 C   |
| 387               | ENT-48-40F   |
| 389               | BAS-08-1 C   |
| 292               | ENT-48-10F   |
| 289               | ENT-24-10F   |
| 289               | ADV-24-10C   |
| 258               | ADV-24-10F   |
| 149               | ENT-24-40F   |
+-----+-----+
8 rows in set (0.09 sec)

mysql> select count>Returns.SKU), as ReturnCount, Returns.SKU
    > from Returns
    > left join customers on Returns.CustomerID = customers.CustomerID
    > where customers.State in ('Washington', 'Oregon', 'Idaho', 'Montana')
    > group by Returns.SKU
    > order by ReturnCount desc;
ERROR 1146 (42S02): You have an error in your SQL syntax; check the manual that corresponds to your MySQL server version for the right syntax to use near 'as ReturnCount, Returns.SKU' from Returns

mysql> select count(Returns.CustomerID) as ReturnCount
    > from Returns
    > inner join RMK on Returns.CustomerID = RMK.OrderID
    > where customers.State in ('Washington', 'Oregon', 'Idaho', 'Montana')
    > group by Returns.SKU
    > order by ReturnCount desc;
+-----+-----+
| SKU          | ReturnCount |
+-----+-----+
| BAS-48-1 C   | 497         |
| ENT-48-40F   | 534         |
| BAS-08-1 C   | 379         |
| ENT-24-10F   | 374         |
| ENT-48-10F   | 362         |
| ADV-48-10F   | 357         |
| ADV-24-10F   | 306         |
| ENT-24-40F   | 198         |
+-----+-----+
8 rows in set (0.09 sec)

mysql>

```

Here I analyzed the data to determine the top three products returned in the northwestern region of the United States.

- **Write a report** to the Quantigration product manager that explains your findings in a way nontechnical stakeholders can understand and use.
 - Provide an effective summary of your findings from the analysis of the sales and returns data you captured. Write at least one paragraph for each of your responses to the prompts below.
 - Sales data by region: Provide a well-written summary of findings from your analysis in Step Two, Part A.
 - In analyzing the sales data by region, I found that the top three products sold across the United States were BAS-48-1 C, ENT-48-40F, and ENT-48-10F, with 8,385, 6,186, and 4,239 units sold respectively. Further drilling down into specific regions, particularly in the southeastern states of Virginia, North Carolina, South Carolina, and Georgia, I discovered similar trends. In this region, BAS-48-1 C emerged as the most popular product, followed by ENT-48-40F and BAS-08-1 C, with 504, 337, and 257 units sold respectively.
 - Returns data by region: Provide a well-written summary of findings from your analysis in Step Two, Part B.
 - In examing returns data by region, I identified the top three products returned in the northwestern region in the United States, including Washington, oregon, Idaho, and Montana. BAS-48-1 C, ENT-48-40F, and BAS-08-1 C emerged as the most returned products,

with 697, 534, and 379 returns respectively. Understanding return patterns in specific regions allows me to investigate potential issues such as product quality and customer satisfaction.