Southern New Hampshire University

DAD 220 Module 5-3
Prepared on: 3 Apr., 2022
Prepared for: Prof. Aastha Agarwal
Prepared by: Alexander Ahmann

Contents

1 Introduction

In order for our organisation to make advancements, we must make use of the technology of mechatronic computing as a means for number crunching and data analysis. In this article, I will be analysing sales data in Quantigration's database and work out descriptive statistics that can be used to give managers insights into better decision making.

2 Procedure

Before continuing with my analysis, I must give unambiguous definitions of the problem and objectives of this research. I am given a dataset called the "return merchandise authorisations" (RMA), and the ultimate goal is to apply the SQL language in producing descriptive statistics regarding the dataset.

In particular, I will be creating descriptive statistics for sales by region and returns by region. The results of this analysis must demonstrate some benefit as how our organisation will carry on regarding its operations.

It would be useful to know which U.S. state has the largest customer base. This can be accomplished with the following query:

```
SELECT State, COUNT(*)
FROM Customers
GROUP BY State
ORDER BY COUNT(*) DESC;
```

The results of the analysis reports that Massachusetts has the largest customer base with a count of 982.

2.1 Statistical analysis on sales

For an initial inquiry into the structure of the data that I will be working with regarding the sales data, I will use the following query to compare:

```
SELECT Orders.*
FROM Orders
INNER JOIN Customers ON Customers.CustomerID =
    Orders.CustomerID
WHERE Customers.State =
    Customers.State = "<state name 1>" AND
    Customers.State = "<state name 2>" AND
    ...
    Customers.State = "<state name n>";
      # for each "n" state in a region.
```

This query will be useful for myself when getting an initial peep into the scenes regarding the dataset to be investigated. Something thing of interest that can be of use to managers for good decision making is the top products being sold. The following query will allow us to work that one out:

```
SELECT DISTINCT SKU, COUNT(*)
FROM Orders
GROUP BY SKU
ORDER BY COUNT(*) DESC;
```

From these analyses, the three leading products being sold have SKU ID's of: `BAS-48-1 C`, `ENT-48-40F` and `ENT-48-10F`.

Finally, in regards to the southeast region, the following query will tell us what are the top three products being sold in said region:[1]

```
SELECT DISTINCT Orders.SKU, COUNT(*)
FROM Orders
INNER JOIN Customers
  ON Orders.CustomerID = Customers.CustomerID
WHERE
  Customers.State = "Virginia"
  OR Customers.State = "North Carolina"
  OR Customers.State = "South Carolina"
  OR Customers.State = "Georgia"
ORDER BY COUNT(*) DESC;
```

From my analysis, the only product that seems to be selling in the southeast has a SKU number of `ENT-48-10F`.

2.2 Statistical analysis on product returns

Of course, it is also important to note that returns have an effect on our business. Returned products typically imply that the customer is unhappy with their product, and managers could benefit from this information by knowing which products to investigate.

A piece of information that may be of interest to managers is the top products returned in the U.S. The following query will allow me to analyse this information:

```
SELECT DISTINCT SKU, COUNT(*)
FROM Orders
INNER JOIN Customers
  ON Orders.CustomerID = Customers.CustomerID
GROUP BY SKU
ORDER BY COUNT(*) DESC;
```

The query reports the three top products SKUs with the most returns: `BAS-48-1 C`, `ENT-48-40F` and `ENT-48-10F`.

[1] See Appendix A.1 for a list of U.S. states and their initials.

3 Summary

The following information has been worked out with their respective recommendations (when recommendations are appropriate, of course):

- The products with the SKU numbers `BAS-48-1 C`, `ENT-48-40F` and `ENT-48-10F` sell the most.
- The products with the SKU numbers `BAS-48-1 C`, `ENT-48-40F` and `ENT-48-10F` are the most likely to be returned. I recommend that managers investigate what is causing these kinds of products to be returned.

A Appendices

A.1 U.S. States and their initials, by region

DAD-220 Region Definitions (n.d.) list the following regions and individual states that are grouped into said regions:

1. U.S. States in the Southwest:

State	Initials
Arizona	AZ
New Mexico	NM
Texas	TX
Oklahoma	OK

2. U.S. States in the Southeast:

State	Initials
Arkansas	AR
Louisiana	LA
Mississippi	MS
Alabama	AL
Georgia	GA
Florida	FL
Kentucky	KY
Tennessee	TN
South Carolina	SC
North Carolina	NC
Virginia	VA
West Virginia	WV
Delaware	DE
Maryland	MD

3. U.S. States in the Northeast:

State	Initials
Pennsylvania	PA
New Jersey	NJ
New York	NY
Connecticut	CT
Rhode Island	RI
Massachusetts	MA
Vermont	VT
New Hampshire	NH
Maine	ME

4. U.S. States in the Midwest:

State	Initials
North Dakota	ND
South Dakota	SD
Kansas	KS
Nebraska	NE
Minnesota	MN
Wisconsin	WI
Iowa	IA
Missouri	MO
Michigan	MI
Indiana	IN
Illinois	IL
Ohio	OH

5. U.S. States in the West:

State	Initials
Washington	WA
Idaho	ID
Montana	MT
Oregon	OR
Wyoming	WY
Colorado	CO
Utah	UT
Nevada	NV
California	CA

References

DAD-220 Region Definitions (n.d.).