

DSO 552: SQL Databases for Business Analysts

Excel as a Database Tool

Background

Medical malpractice law in the United States is derived from English common law, and was developed by rulings in various state courts. Medical malpractice lawsuits are a relatively common occurrence in the United States. The legal system is designed to encourage extensive discovery and negotiations between adversarial parties with the goal of resolving the dispute without going to jury trial. According to a recent study published in the US News and World Report the cost of medical malpractice in the United States is \$55.6 billion a year, which is 2.4 percent of annual health-care spending. An insurance company wants to develop a better understanding of its claims paid out for medical malpractice lawsuits. Its records show claim payment amounts, as well as information about the presiding physician and the claimant for a number of recently adjudicated or settled lawsuits.

The Objective of the Analysis

Using Excel's business Analytics Tools, to explore claim payment amounts, and identify factors that appear to influence the amount of the payment.

The Data

The file "case_data.xlsx" contains two sheets: (1) **MalpracticeData**, and (2) **DeductibleRates**

The **MalpracticeData** sheet contains information about the last 236 claim payments made from 01/01/2019 to 01/16/2019. The six variables in the data table are described below:

Date: Day on which the claim payment is made

Week: Week of the year

Amount_Paid: Amount of the claim payment in dollars

Severity: The severity rating of damage to the patient, from 1 (emotional trauma) to 9 (death)

Age: Age of the claimant in years

Specialty: Specialty of the physician involved in the lawsuit

Gender: Patient Gender

The DeductibleRates sheet has all the deductible rates based on physician specialty and patient's age. Deductible is the amount of expenses that must be paid out of pocket before an insurer pay any expenses.

Questions

Part 1: Data Preparation

1. (1 point) Create a new variable called "Weekday" using the Excel function (=weekday(DATE)). The weekday() function takes a date as its input, and produces an integer between 1 and 7 representing the day of the week. The day of week is saved as an integer, ranging from 1 (Sunday) to 7 (Saturday), by default in Excel.

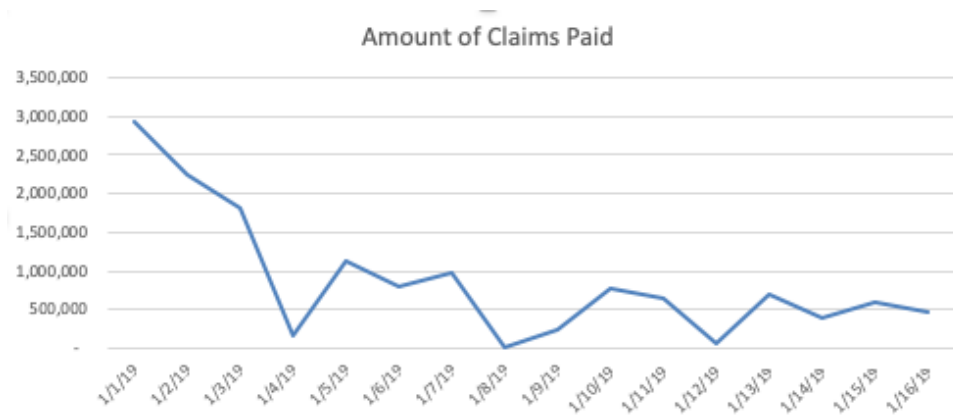
2. (2 points) Create a new variable called “Age_Group” that using the following rule.

Age_Group	Age
Child	Age < 15
Young	15 <= Age < 40
Mature	40 <= Age < 80
Old	Age >=80

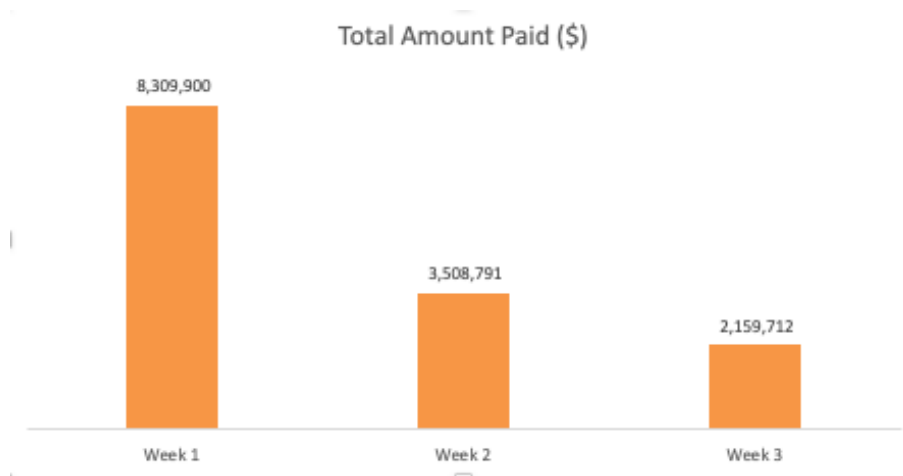
3. (2 points) Create a new variable called “Deductible_Rate”. Find the appropriate deductible rate using the table in the DeductibleRates sheet provided for each claim (Hint: use vlookup).
4. (1 point) Create a new variable called “Claim_Amount”. Use the following formula to create this variable.
Claim_Amount = Amount_Paid / (1 – Deductible Rate)
5. (1 point) Create a new variable called “Deductible”. Use the following formula to create this variable.
Deductible = Claim_Amount – Amount_Paid

Part 2: Data Analysis

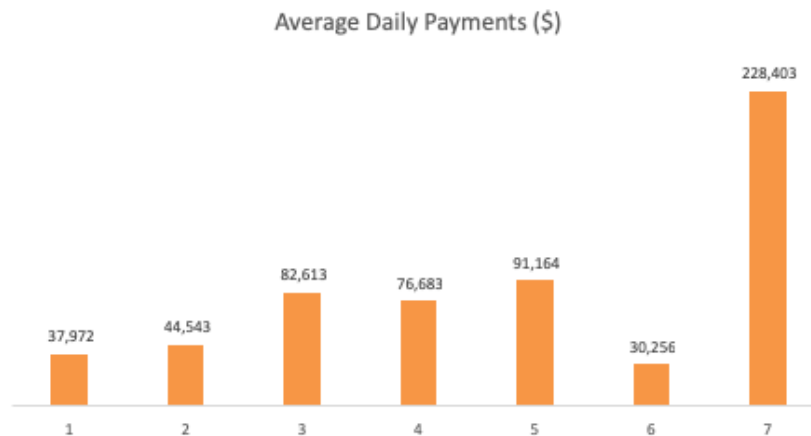
6. (1 point) Which day out of the 16 days in the data set corresponds to the largest total **Amount_Paid**? Report the day and the corresponding value in dollars.
7. (1 point) Using the results from the previous question, create a copy of the following graph to show the claim amounts paid over the 16 days:



8. (1 point) Which week has the lowest total amount paid by the company? What is the corresponding value?
9. (1 point) Using the results from the previous question, create a copy of the following graph to show the distribution of the total claim amounts over the weeks.



10. (1 point) Which day of the week in the given data set corresponds to the largest average Claim_Amount? What is the corresponding amount in dollars.
11. (1 point) Using the results from the previous question, create a copy of the following graph to show the distribution of the average claim amounts over the week days.



Part 3: Dashboard

12. (2 points) Create an interactive dashboard to understand the claims amounts paid by the company. Your dashboard should have the three graphs created earlier, and the slicer variables are **Age_Group** and **Gender**.

Amount Paid Dashboard

Age_Group

Child

Mature

Old

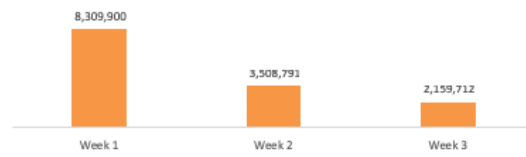
Young

Gender

Female

Male

Total Amount Paid (\$)



Amount of Claims Paid



Average Daily Payments (\$)

