upGrad

PLACEMENT REFRESHER PROGRAM

Session 17 - MS Excel 3
Case Study & Questions

By
Ritesh Kumar Pandey

Agenda

- MCQ / MSQ
- Case Study

Question - Amazon+Flipkart

upGrad

Flipkart and Amazon's data both have names containing author's name for any given book.

Count the number of missing values in each of these columns and assess which column should be dropped (contains more missing values)?

A) Flipkart
B) Amazon

Question - Amazon+Flipkart

upGrad

Flipkart and Amazon's data both have names containing author's name for any given book.

Count the number of missing values in each of these columns and assess which column should be dropped (contains more missing values)?

A) Flipkart
B) Amazon

Flipkart has 25 missing values in the author's column while amazon has only one.

Question - Amazon+Flipkart

upGrad

Around 25 books in the Flipkart data have missing values in the Flipkart star column. What should be the best way of treating these values?

A) Replace by Mean
B) Replace by Median
C) Replace by Either
D) Remove missing values

Question - Amazon+Flipkart

upGrad

Around 25 books in the Flipkart data have missing values in the Flipkart star column. What should be the best way of treating these values?

A) Replace by Mean
B) Replace by Median
C) Replace by Either
D) Remove missing values

The mean and the median are nearly equal, either can be used for the analysis.

Question - Amazon+Flipkart

upGrad

which of the following combinations can be treated as a unique identifier for each row?

A) ISBN
B) Book-ISBN
C) Book-Author
D) All of the above

Question - Amazon+Flipkart

upGrad

which of the following combinations can be treated as a unique identifier for each row?

A) ISBN
B) Book-ISBN
C) Book-Author
D) All of the above

All of these can be used as unique identifiers for the given dataset.

Question - Amazon+Flipkart

upGrad

Which of the following columns need to be treated for outliers before any further analysis can be done on them?

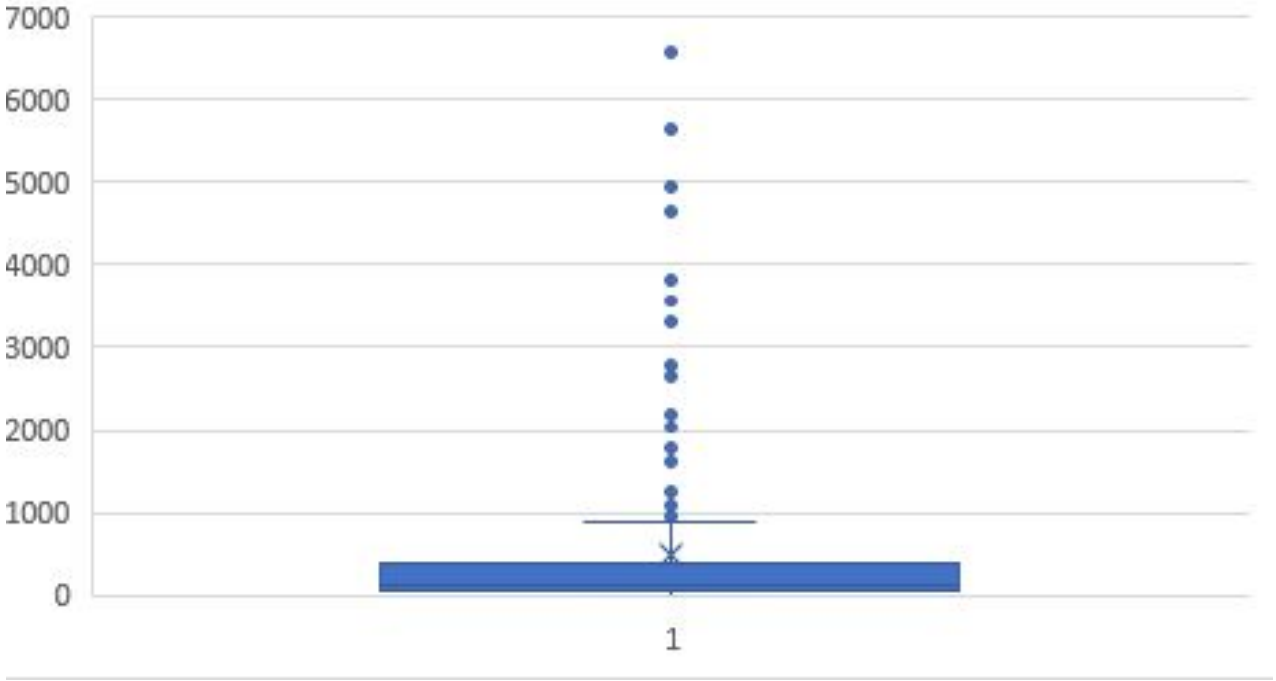
A) Review count (Amazon)
B) ISBN
C) Price (Flipkart)

Question - Amazon+Flipkart

upGrad

Which of the following columns need to be treated for outliers before any further analysis can be done on them?

A) Review count (Amazon)
B) ISBN
C) Price (Flipkart)



On analysing the box plot, you can see that this variable has the maximum number of outliers and hence needs to be treated before any further analysis.

Question - Marks

upGrad

Create a column named ‘Count’ that maintains the number of missing values in each row of the ‘Marks’ data set given above. What is the sum of the cells of the ‘Count’ column?

A) 33
B) 22
C) 44
D) 55

Question - Marks

upGrad

Create a column named ‘Count’ that maintains the number of missing values in each row of the ‘Marks’ data set given above. What is the sum of the cells of the ‘Count’ column?

A) 33
B) 22
C) 44
D) 55

Use ‘=COUNTIF(A2:F2,"")’ to get the number of missing values in the second row. Now, using the similar function, calculate the number of missing values for the remaining rows. Later, calculate the sum of all the cells in the Count column. Therefore, the sum of all the cells in the Count column is 44.

Question - Marks

upGrad

This question is in continuation of the previous question. Once you are done with the previous question, remove all the rows with five missing values.

You created the Count column in the previous question. Now, what is the sum of the Count column after removing the rows with five missing values?

A) 44
B) 39
C) 29
D) 34

Question - Marks

upGrad

This question is in continuation of the previous question. Once you are done with the previous question, remove all the rows with five missing values.

You created the Count column in the previous question. Now, what is the sum of the Count column after removing the rows with five missing values?

A) 44
B) 39
C) 29
D) 34

Once you are done with the previous question, follow the points given below:

- By applying the filter, get all the rows with the value of the missing value count as 5.
- Now, remove all the rows that you get after performing the step above.
- Next, remove the filter to get all the rows.
- Then, calculate the sum of all the cells of the Count column.

There are two rows with five missing values. After deleting these rows, the sum of all the cells in the Count column will be reduced by 10, which then becomes 34.

Question - Customers

upGrad

The given data set 'customer' has a column 'Cust_id', which has values Cust_1, Cust_2 and so on.

Remove the repeated 'Cust_' from the column Cust_id so that the column Cust_id has just numbers like 1, 2, 3 and so on. Next, find the sum of all the elements of the column Cust_id.

A) 17,79,028
B) 16,79,028
C) 1,67,902
D) 6,79,028

Question - Customers

upGrad

The given data set 'customer' has a column 'Cust_id', which has values Cust_1, Cust_2 and so on.

Remove the repeated 'Cust_' from the column Cust_id so that the column Cust_id has just numbers like 1, 2, 3 and so on. Next, find the sum of all the elements of the column Cust_id.

A) 17,79,028
B) 16,79,028
C) 1,67,902
D) 6,79,028

Select all the elements of the column ‘Cust_id’ and then click ‘Ctrl+H’ (replace). Now, replace ‘Cust_’ with the empty string. Next, calculate the sum of the column ‘Cust_id’. Hence, the answer is 16,79,028.

MCQ

upGrad

Which Excel feature is used to create a one-variable data table?

A. Scenario Manager
B. Data Validation
C. Goal Seek
D. Data Table

MCQ

upGrad

Which Excel feature is used to create a one-variable data table?

A. Scenario Manager
B. Data Validation
C. Goal Seek
D. Data Table

Case Study - Telcom

upGrad

Open the telcom dataset and answer the following questions:

1. Explain the dataset
2. Copy the Phone Number and International Plan Columns in some another worksheet of the same csv file and then delete the International Plan column from the original worksheet
3. Find out the total minutes, total number of call and total charge for each user
4. Find out the top 10% total number of call
5. Apply a color scale in total charge such that highest charge is green in color and lowest charge is red in color
6. Find out the churn rate of each state
7. Fill in the International Plan of each user
8. What other insights you can find from the dataset

THANK YOU