Notes Analyze Data to Answer Questions

Wonnetz Phanthavong

email: wonnetz2@gmail.com.com
linkedin: https://www.linkedin.com/in/wonnetz/

Abstract

- Data Organization

- Formatting/Adjusting Data

- Data Aggregation

- Performing Data Calculations

Week 1: Introduction to Getting Organized

The 4 Phases of Analysis

1. Organize Data

2. Format/Adjust Data

3. Get Input from Others

4. Transform Data

Sorting vs. Filtering

- Sorting: Arranges the data into a meaningful order, ascending/descending, etc

- Filtering: Focuses on the important data that you want

Sorting Datasets in Spreadsheets/Excel

- Sort Sheet

- Sort Range

Sorting/Filtering Datasets in SQL

- ORDER BY
    - Defaults to ordering in ascending order, ASC

- WHERE

Week 2: Getting Started with Data Formatting

What is involved with Data Formatting?

Changing. . .
- Dates
- Strings
- Percentages
- Currencies
- etc

Useful function in Spreadsheets

- CONVERT(value, "“,”")
- CONCAT

- CONCATENATE

Data Validation

- Add dropdown lists with predetermined options

- Create custom checkboxes

- Protect structured data and formulas

Conditional Formatting

- Makes data easier to understand

- Allows you to view errors easier

Transforming Data in SQL

- in General, CAST Function (syntax)
    - CAST(expression AS typename)

- Number to a String
    - CAST(MyCount AS STRING) FROM MyTable

- String to a Number
    - CAST(MyVarCharCol AS INT) FROM MyTable

- Date to a String
    - CAST(MyDate AS STRING) FROM MyTable

- Date to DateTime
    - CAST(MyDate AS DATETIME) FROM MyTable

- SAFE_CAST Function
    - Returns a value of NULL instead of an Error

    - SAFE_CAST(expression AS typename)

Manipulating Strings in SQL

- CONCAT
    - Adds strings together to create new text strings that can be used as unique keys

- CONCAT_WS
    - A function that adds two or more strings together with a separator

- CONCAT with +
    - Adds two or more strings together using the + operator

Week 3: Aggregating Data for Analysis

What does it mean to aggregate data?
Joining or collecting pieces of data into a whole.

Trouble Shooting Questions

- How should I prioritize these issues?

- In a single sentence, what's the issue I'm facing?

- What resources can help me solve the problem?

- How can I stop this problem from happening in the future?

JOIN

- INNER JOIN
  - Joins what the two tables have in common

- LEFT JOIN
  - Joins the left table and what the two tables have in common

- RIGHT JOIN
  - Joins the right table and what the two tables have in common

- OUTER JOIN
  - Joins everything in the two tables

Week 4: Performing Data Calculations

Covers
- Formulas for Basic Calculations
- Conditional formulas that use the IF function
- The SUMPRODUCT function
- Pivot tables to organize calculations
- Queries and calculations in SQL
- Temporary tables in SQL

##SUMPRODUCT
Multiples across the row and then adds up the calculated values vertically

What can Pivot Tables do?

- Easily find out how much revenue was generated each year

- Find averages per item

- Give insights to possible trends

Calculations in SQL

- SUM

- AVG

What to do if you have a divison by zero error?

You can adjust this in the WHERE clause by filtering out the attribute that has values of zero.

Types of Data Validation

1. Data Type

- Check that the data matches the data type defined for a field.

2. Data Range

- Check that the data falls within an acceptable range of values defined for the field.

3. Data Constraints

- Check that the data meets certain conditions or criteria for a field. This includes the type of data entered as well as other attributes of the field, such as number of characters.

4. Data Consistency

- Check that the data makes sense in the context of other related data.

5. Data Structure

- Check that the data follows or conforms to a set structure.

6. Code Validation

- Check that the application code systemically performs any of the previously mentioned validations during user data input.

Temporary Tables

- WITH clause: A type of temporary table that you can query from multiples times

- SELECT INTO clause: Copies data from one table to another, but it doesn't add info to the database

- CREATE TABLE clause: Copies data from one table to another AND it adds the info to the database

Your Intermediate Guide to SQL

Intermediate Guide to SQL