Grouping Data with SQL

Sadie St. Lawrence, MSA
Data Scientist, VSP

UC**DAVIS**
EXTENSION

Learning Objectives

Perform some additional aggregations using the GROUP BY and HAVING clauses

Discuss how NULLs are or aren't affected by the GROUP BY and HAVING clauses

Use the GROUP BY and ORDER BY clauses together to better sort your data

Grouping Data

Learn how to group data in order to summarize subsets of data

New clauses GROUP BY; HAVING

How to aggregate on a particular value

Grouping Example

Counts customers after group on region
rather than counting the whole table

```
SELECT
Region
,COUNT(CustomerID) AS total_customers
FROM Customers
GROUP BY Region;
```

Additional GROUP BY Information

GROUP BY clauses can contain multiple columns

Every column in your SELECT statement must be present in a GROUP BY clause, except for aggregated calculations

NULLs will be grouped together if your GROUP BY column contains NULLs

Slide 5: Additional GROUP BY Information

HAVING Clause – Filtering for Groups

WHERE does not work for groups

WHERE filters on rows

Instead use HAVING clause to filter for groups

Slide 6: HAVING Calause – Filtering for Groups

Grouping Example

```
SELECT
CustomerID
,COUNT (*) AS orders
FROM Orders
GROUP BY CustomerID
HAVING COUNT (*) >=2;
```

WHERE vs. HAVING

WHERE filters before data is grouped

HAVING filters after data is grouped

Rows eliminated by the WHERE clause will not be a included in the group

Slide 8: WHERE vs. HAVING

ORDER BY with GROUP BY

ORDER BY sorts data

GROUP BY does not sort data

```
SELECT SupplierID
,COUNT(*) AS Num_Prod
FROM Products
WHERE UnitPrice >= 4
GROUP BY SupplierID
HAVING COUNT (*) >=2;
```

Slide 9: ORDER BY with GROUP BY