Filters and Aggregations

FILTERS = , !=, >, <
IN, NOT IN, BETWEEN, LIKE, NOT LIKE
SJM, MIN, MAX, COUNT
GROUP BY, HAVING
COMMENTING

Matthew Morris

Git: Morrisdata
MatthewMorris.DA@gmail.com

Previously in Data Analytics

Explain where SQL fits in the dataflow
Retrieve and filter data with basic SQL
Navigate a Relational Database

SELECT

FROM

WHERE

ORDERBY

LIMIT

FILTER

WHERE

- =, !=, >,<
- IS NULL, IS NOT NULL
- IN, NOT IN
- BETWEEN
- LIKE
- OR

GROUP BY HAVING

SELECT

FROM

WHERE

GROUP BY

HAVING

ORDER BY

LIMIT

Aggregate functions

- MIN
- MAX
- SUM
- COUNT

GROUP BY store, item

GROUP BY 1,2 DESC

HAVING AVG(sales)>100
AND COUNT(customers)>20

WHERE – filter for dimensions and measures

GROUP BY –groups dimensions when a measure is aggregated

HAVING – filter for aggregated measure

SELECT   - Fields you want to see in your results

FROM   - Table where fields come from

WHERE   - Filters for your results

GROUP BY   - Groups dimensions when using an aggregate

HAVING   - Filters aggregations

ORDER BY   - How you can sort your results

LIMIT   - Limits number of records returned

SELECT Store, (cost –sell price),   SUM(sales),

FROM sales

WHERE Category = 'Tequila'

AND units purchased >2

GROUP BY Store

HAVING SUM(sales) > 30.00

ORDER BY 3

Which products have a case cost of more than $100?

Which tequilas have a case cost of more than $100?

Which tequilas or scotch whiskies have a case cost of more than $100?

Which tequilas or scotch whiskies have a cast cost between $100 and $120?

Which whiskies of any kind cost more than $100?'

Which whiskies of any kind cost between $100 and $150?

Which products except tequilas cost between $100 and $120?

Comment up your code!

Q & A

Fall down seven times, get up eight.
– Japanese Proverb

Conclusion

Group by will need to be used when using any aggregates

Having is how you can filter those aggregates.

EXIT TICKET

CLASS : Filters and Aggregations

QUESTION:

What is the difference between GROUP BY and HAVING?