Data Analyst Training Programme

WHERE Clauses & Filtering

Introduction to Data Filtering

Why Filtering Matters

- Raw data contains everything, but you need specific answers
- Filtering reduces noise and focuses analysis
- Essential for performance on large datasets

The WHERE Clause

- Acts as a gatekeeper for your data
- Only rows meeting conditions pass through
- Applied before **SELECT** processing

Business Impact

- Find customers in specific regions
- Identify high-value transactions
- Locate products needing restocking

Comparison Operators

Equality and Inequality

- = (equals)
- <> or != (not equals)
- **> (greater than) | < (less than) | >= (greater than or equal) | <= (less than or equal)**

Data Type Considerations

- **Text**: case-sensitive, needs quotes
- **Numbers**: no quotes required
- **Dates**: format matters

Common Patterns

- Price ranges for products
- Date ranges for orders
- Exact matches for categories

Logical Operators

AND Operator

- Both conditions must be true
- Narrows results further
- Useful for multiple criteria

OR Operator

- Either condition can be true
- Broadens results
- Useful for alternatives

NOT Operator

- Excludes matching records
- Finds everything except specified criteria

Parentheses for Grouping

- Controls evaluation order
- Essential for complex logic

Pattern Matching with LIKE

LIKE Operator Purpose
- Finds partial text matches
- Essential for flexible searching
- Uses wildcards for patterns

Wildcard Characters
- % matches any number of characters
- _ matches exactly one character

Common Patterns
- Starts with: 'A%'
- Ends with: '%son'
- Contains: '%market%'
- Specific position: 'A_d%'

Case Sensitivity
- MySQL is case-insensitive by default
- Can be configured for case-sensitive matching

Advanced Filtering Techniques

IN Operator
- Match against list of values
- Cleaner than multiple **OR** conditions
- Excellent for categories or regions

BETWEEN Operator
- Range filtering made simple
- Inclusive of boundary values
- Works with numbers, dates, and text

IS NULL and IS NOT NULL
- Handle missing data properly
- NULL requires special operators
- Critical for data quality analysis

Combining Techniques
- Mix operators for complex conditions
- Logical order matters

Assignment

Complete these filtering challenges using the Northwind database:

Customer Analysis

- Find all customers from Germany, France, or UK
- List customers whose names start with 'B'
- Identify customers with missing postal codes

Product Investigation

- Show products priced between £15 and £40
- Find all products from category 1 (Beverages) or category 8 (Seafood)
- List products whose names contain 'cheese' (case-insensitive)

Order Examination

- Display orders placed in December 1996
- Find orders handled by employees 1, 3, or 5
- Show orders where customer ID is greater than 50

Advanced Challenges

- Find German customers whose names don't start with 'A'
- List expensive products (price > £50) from dairy or confection categories
- Identify orders from the last quarter of 1996 shipped by 'Speedy Express'

Until Next Week Sunday...

See you next week on Sunday, **[student name]**.

Filtering data is important for effective analysis, helping you focus on specific information and make informed decisions. Mastering the **WHERE** clause and various filtering techniques will empower you to extract valuable insights from your datasets.

Thank you, [student name].

Any Questions?