

SQL Project Questions based on UNEG Recruitment Dataset

Section 1: Basic Data Retrieval

1. List all candidates and their departments.
2. Retrieve the top 10 candidates based on their average Net score (Net_AVG).
3. Find all candidates from the Pharmacy department.
4. Show candidates who scored above 25 in Content_A.
5. Display the names and departments of candidates who had no penalty.
6. Find the number of candidates assessed.
7. Show all candidates whose Grammar_AVG is above 20.

Section 2: Filtering and Conditions

1. List departments and how many candidates came from each.
2. Retrieve candidates who have a negative penalty (Penalty_AVG less than 0).
3. Display candidates who scored above 80 in their Net_AVG.
4. Find candidates whose Content_A score is greater than Structure_A score.

Section 3: Aggregations and Group Analysis

1. Calculate the overall average of Content_AVG, Structure_AVG, Style_AVG, and Grammar_AVG.
2. Group candidates by Department and compute the average Net_AVG for each department.

SQL Project Questions based on UNEG Recruitment Dataset

3. Identify departments with an average Net_AVG greater than 85.

Section 4: Ranking and Advanced Filters

1. Create a ranking of candidates based on their Net_AVG, highest to lowest.
2. Identify candidates who improved significantly ($\text{Content_B} > \text{Content_A}$ by at least 5 points).
3. Find the candidate with the highest Penalty_B deduction.

Section 5: Calculations and New Metrics

1. For each candidate, calculate a new Total_Avg_Score as the mean of Content_AVG, Structure_AVG, Style_AVG, and Grammar_AVG. Then rank them.
2. Find out if penalties impacted results by comparing Net_AVG before and after penalties for candidates with penalties.

Section 6: View Creation for Reporting

1. Write a query that creates a view showing: Name, Department, Total_Avg_Score, Penalty_AVG, and Net_AVG for future analysis.

Bonus Ideas to Expand This Project

- Create a dashboard showing best performing departments.
- Build a penalty impact report showing which penalties hurt candidates the most.
- Predict if a candidate is likely to pass a threshold score (e.g., 85) based on component scores (simple logistic regression after SQL!).