SQL Questions, Queries, and Explanations

Question 1

Prompt

How many entries do you have in your database who have applied for Fall 2026?

SQL

```
SELECT COUNT(*)
FROM applicant
WHERE term = 'Fall 2026';
```

Explanation

• Counts matching rows using COUNT(*).
• Filters to the Fall 2026 application term.

Question 2

Prompt

What percentage of entries are from international students (not American or Other) (to two decimal places)?

SQL

```
SELECT ROUND(
COUNT(*),
2
COUNT(CASE WHEN us_or_international NOT IN ('American', 'Other') THEN 1 END) * 100.0 /
)
FROM applicant;
```

Explanation

• Computes a percentage by dividing a conditional count by total rows, then rounds to 2 decimals.
• Treats any value other than 'American' or 'Other' as international for this calculation.

Question 3

Prompt

What is the average GPA, GRE, GRE V, GRE AW of applicants who provide these metrics?

SQL

```
SELECT
ROUND(AVG(gpa)::numeric, 2) AS avg_gpa,
ROUND(AVG(gre)::numeric, 2) AS avg_gre,
ROUND(AVG(gre_v)::numeric, 2) AS avg_gre_v,
ROUND(AVG(gre_aw)::numeric, 2) AS avg_gre_aw
FROM applicant
WHERE gpa IS NOT NULL OR gre IS NOT NULL OR gre_v IS NOT NULL OR gre_aw IS NOT NULL;
```

Explanation

• Calculates averages with AVG(...) and rounds to 2 decimals for readability.

Question 4

Prompt

What is their average GPA of American students in Fall 2026?

SQL

```
SELECT ROUND(AVG(gpa)::numeric, 2) AS avg_gpa
FROM applicant
WHERE us_or_international = 'American' AND term = 'Fall 2026';
```

Explanation

• Calculates averages with AVG(...) and rounds to 2 decimals for readability.

Question 5

Prompt

What percent of entries for Fall 2026 are Acceptances (to two decimal places)?

SQL

```
SELECT ROUND(
```

```
2
COUNT(CASE WHEN status = 'Accepted' THEN 1 END) * 100.0 / COUNT(*),
)
FROM applicant
WHERE term = 'Fall 2026';
```

Explanation

- Computes a percentage by dividing a conditional count by total rows, then rounds to 2 decimals.
- Filters to the Fall 2026 application term.
- Filters to accepted applications only.

Question 6

Prompt

What is the average GPA of applicants who applied for Fall 2026 who are Acceptances?

SQL

```
SELECT ROUND(AVG(gpa)::numeric, 2) AS avg_gpa
FROM applicant
WHERE term = 'Fall 2026' AND status = 'Accepted';
```

Explanation

- Calculates averages with AVG(...) and rounds to 2 decimals for readability.
- Filters to the Fall 2026 application term.
- Filters to accepted applications only.

Question 7

Prompt

How many entries are from applicants who applied to JHU for a masters degrees in Computer Science?

SQL

```
SELECT COUNT(*)
FROM applicant
WHERE (program LIKE '%Computer Science%' OR llm_generated_program LIKE '%Computer Science%')
AND degree = 'Masters'
AND (
program LIKE '%Johns Hopkins%'
OR program LIKE '%JHU%'
OR llm_generated_university LIKE '%Johns Hopkins%'
OR llm_generated_university LIKE 'John%'
OR llm_generated_university LIKE '%JHU%'
);
```

Explanation

- Counts matching rows using COUNT(*).
- Matches Computer Science-related programs using LIKE '%Computer Science%'.
- Uses LLM-generated normalization fields to make matching more robust to messy input text.

Question 8

Prompt

How many entries from 2026 are acceptances from applicants who applied to Georgetown University, MIT, Stanford University, or Carnegie Mellon University for a PhD in Computer Science?

SQL

```
SELECT COUNT(*) AS phd_cs_acceptances_count
FROM applicant
WHERE term LIKE '%2026%'
AND status = 'Accepted'
AND program LIKE '%Computer Science%'
AND degree LIKE '%PhD%'
AND (
program LIKE '%Georgetown%'
OR program LIKE '%MIT%'
OR program LIKE '%Massachusetts Institute of Technology%'
OR program LIKE '%Stanford%'
OR program LIKE '%Carnegie Mellon%'
OR program LIKE '%CMU%'
```



```
);
```

Explanation

- Counts matching rows using COUNT(*).

• Filters to accepted applications only.
• Matches Computer Science-related programs using LIKE '%Computer Science%'.

Question 9

Prompt

Do you numbers for question 8 change if you use LLM Generated Fields (rather than your downloaded fields)?

SQL

```
SELECT COUNT(*) AS phd_cs_acceptances_count
FROM applicant
WHERE term LIKE '%2026%'
AND status = 'Accepted'
AND llm_generated_program LIKE '%Computer Science%'
AND degree LIKE '%PhD%'
AND (
llm_generated_university LIKE '%Georgetown%'
OR llm_generated_university LIKE '%MIT%'
OR llm_generated_university LIKE '%Massachusetts Institute of Technology%'
OR llm_generated_university LIKE '%Stanford%'
OR llm_generated_university LIKE '%Carnegie Mellon%'
OR llm_generated_university LIKE '%CMU%'
);
```

Explanation

• Counts matching rows using COUNT(*).
• Filters to accepted applications only.
• Matches Computer Science-related programs using LIKE '%Computer Science%'.
• Uses LLM-generated normalization fields to make matching more robust to messy input text.

Question 10

Prompt

Which top 5 universities have the highest acceptance rates in Fall 2026?

SQL

```
SELECT
llm_generated_university AS university,
ROUND(COUNT(CASE WHEN status = 'Accepted' THEN 1 END) * 100.0 / COUNT(*), 2) AS
acceptance_rate
FROM applicant
WHERE term = 'Fall 2026'
```



```
GROUP BY llm_generated_university
ORDER BY acceptance_rate DESC
LIMIT 5;
```

Explanation

• Computes a percentage by dividing a conditional count by total rows, then rounds to 2 decimals.
• Filters to the Fall 2026 application term.
• Filters to accepted applications only.
• Uses LLM-generated normalization fields to make matching more robust to messy input text.
• Aggregates by university, then computes and ranks acceptance rates.
• Sorts results and limits the output to the top N rows.

Question 11

Prompt

What are the top 10 distinct programs applied for Fall 2026?

SQL

```
SELECT DISTINCT program
FROM applicant
WHERE term = 'Fall 2026'
ORDER BY program
LIMIT 10;
```

Explanation

• Returns unique values using DISTINCT.
• Filters to the Fall 2026 application term.
• Sorts results and limits the output to the top N rows.