

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

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

- Answer: 8924
- This query counts all records in the table where the term field is 'Fall 2026'.

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

```
SELECT COUNT(*) FROM admissions
WHERE us_or_international = 'International';
```

- Answer: 50.32%
- This query counts all records in the table where the 'us\_or\_international' field is 'International'.

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

```
SELECT
ROUND(AVG(gpa)::numeric, 2) AS avg_gpa,
ROUND(AVG(gre)::numeric, 1) AS avg_gre_quant,
ROUND(AVG(gre_v)::numeric, 1) AS avg_gre_verbal,
ROUND(AVG(gre_aw)::numeric, 2) AS avg_gre_aw
FROM admissions
WHERE
(gpa BETWEEN 0.1 AND 5.0)
AND (gre BETWEEN 130 AND 170)
AND (gre_v BETWEEN 130 AND 170)
AND (gre_aw BETWEEN 0.0 AND 6.0);
```

- Answer: GPA = 3.74, GRE = 164.7, GRE V = 159.8, GRE AW = 4.36
- This query selects the average of each of the GPA/GRE fields, casts them to a numeric datatype for the ROUND function, and rounds the result to two decimal places. The WHERE clause ensures that only realistic scores are included for each field. This prevents cases where users may have misplaced a decimal or inadvertently swapped two fields (e.g., entering a GRE score of 160 as the GPA).

4. What is the average GPA of American students in Fall 2026?

```
SELECT
ROUND(AVG(gpa)::numeric, 2) AS avg_gpa
FROM admissions
WHERE
(gpa BETWEEN 0.1 AND 5.0)
AND (us_or_international = 'American')
AND (term = 'Fall 2026');
```

- Answer: 3.80
- This query gets the average GPA for entries where the entered GPA is in a realistic range like the last query. Two additional AND statements ensure that the average is only computed for rows where the 'us\_or\_international' value is 'American' and the 'term' field is 'Fall 2026'.

#### 5. What percent of entries for Fall 2025 are Acceptances (to two decimal places)?

```
SELECT
ROUND(
(COUNT(*) FILTER (WHERE status = 'Accepted')::numeric /
COUNT(*) * 100),
2
) as acceptance_percentage
FROM admissions
WHERE term = 'Fall 2025';
```

- Answer: 35.98%
- Here, inside the ROUND function, the numerator is a count of all records with a FILTER condition applied to only return the count of records where the 'status' field is 'Accepted'; this is also type-casted to a numeric datatype to ensure the following division does not always return 0 due to integer division (and because ROUND requires a numeric datatype anyway). The denominator counts all records, and then the result of the overall division is multiplied by 100 to obtain the answer as a percentage. The second argument of 2 in the ROUND function ensures the result is rounded to two decimal places. The WHERE clause at the end ensures that this percentage is only computed over records where the 'term' field is 'Fall 2025'.

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

```
SELECT
ROUND(AVG(gpa)::numeric, 2) AS avg_gpa
FROM admissions
```

```
WHERE
(term = 'Fall 2026')
AND (status = 'Accepted');
```

- Answer: 3.79
- This query looks at all records where the 'term' field is 'Fall 2026' and the 'status' field is 'Accepted', and rounds the average GPA of these records to two decimal places in the same fashion as in query 4.

**7. How many entries are from applicants who applied to JHU for a master's degree in Computer Science?**

```
SELECT COUNT(*)
FROM admissions
WHERE (
llm_generated_university ILIKE '%john%hopkin%'
OR university ILIKE '%jhu%'
OR university ILIKE '%john%hopkin%'
)
AND (
llm_generated_program ILIKE '%comp%sci%'
OR program ILIKE '%comp%sci%'
OR program = 'CS'
)
AND (degree = 'Masters');
```

- Answer: 8
- This query selects the total number of entries where all of the following WHERE clauses are satisfied. The first condition checks the 'llm\_generated\_university' field for any string that has 'john' and 'hopkin' in it (case insensitive) using an ILIKE statement, OR if the raw 'university' field meets this same condition, OR if the university field has 'jhu' in it (case insensitive). The second condition checks whether the 'llm\_generated\_program' field has the substrings 'comp' and 'sci' in it, OR if the raw 'program' field satisfies the same condition, OR if the raw 'program' field simply equals 'CS'. The final condition is that the 'degree' field must equal 'Masters'.

**8. 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?**

```
SELECT COUNT(*)
FROM admissions
WHERE (
```

```

llm_generated_university = 'Massachusetts Institute of Technology'
OR university ILIKE '%mass%institute%tech'
OR university = 'MIT'

OR llm_generated_university = 'Stanford University'
OR university ILIKE '%stanford%'
OR university = 'SU'

OR llm_generated_university = 'Georgetown University'
OR university ILIKE '%georgetown%'

OR llm_generated_university = 'Carnegie Mellon University'
OR university ILIKE '%carnegie%mellon%'
OR university ILIKE '%carnegie%melon%'
OR university = 'CMU'
)
AND (date_added BETWEEN '2026-01-01' AND '2026-12-31')
AND (
llm_generated_program ILIKE '%comp%sci%'
OR program ILIKE '%comp%sci%'
OR program = 'CS'
)
AND (degree = 'PhD');

```

- Answer: 29
- This query operates similarly to the previous query, except now the 'degree' field must be equal to 'PhD', there is an additional AND condition that restricts the 'date\_added' field to contain only dates in 2026, and there are several possible formats for the 'llm\_generated\_university' and 'university' fields separated by OR clauses to allow for entries listing any one of MIT, Stanford, Georgetown, or Carnegie Mellon. To get entries for applicants to these schools only, the 'llm\_generated\_university' fields must equal the canonical names of these schools exactly, or the 'university' field must equal the abbreviated names (e.g. 'MIT'), or the 'university' field must contain substring(s) that uniquely identify these schools (accounting for slight misspellings and variations).

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

a. LLM Fields Only

- Answer: 29

```
SELECT COUNT(*)
```

```

FROM admissions
WHERE (
llm_generated_university = 'Massachusetts Institute of Technology'
OR llm_generated_university = 'Stanford University'
OR llm_generated_university = 'Georgetown University'
OR llm_generated_university = 'Carnegie Mellon University'

)
AND (date_added BETWEEN '2026-01-01' AND '2026-12-31')
AND (
llm_generated_program ILIKE '%comp%sci%'
)
AND (degree = 'PhD');

```

b. Raw Fields Only

- Answer: 7

```

SELECT COUNT(*)
FROM admissions
WHERE (
university ILIKE '%mass%institute%tech'
OR university = 'MIT'

OR university ILIKE '%stanford%'
OR university = 'SU'

OR university ILIKE '%georgetown%'

OR university ILIKE '%carnegie%mellon%'
OR university ILIKE '%carnegie%melon%'
OR university = 'CMU'
)
AND (date_added BETWEEN '2026-01-01' AND '2026-12-31')
AND (
program ILIKE '%comp%sci%'
OR program = 'CS'
)
AND (degree = 'PhD');

```

- These two queries are almost identical to the query in question 8, except the first omits any condition that relies on the 'university' or 'program' fields, and the second omits any condition that relies on the 'llm\_generated\_university' or 'llm\_generated\_program' fields.

10. How many applicants applied for a Fall 2025 start term vs a Spring 2025 start term?

```
SELECT
COUNT(*) FILTER (WHERE term = 'Fall 2025') as fall_count,
COUNT(*) FILTER (WHERE term = 'Spring 2025') as spring_count
FROM admissions;
```

- Answer: Fall 2025: 20,023, Spring 2025: 171
- This query has two COUNT(\*) FILTER statements to return two values, where the first is the total count of all records with the 'term' field equal to 'Fall 2025' and the second is the total count of all records with the 'term' field equal to 'Spring 2025'.

11. What is the average acceptance rate of Master's programs vs PhD programs for American applicants who do and do not self-report their GPA?

```
SELECT
CASE
WHEN gpa IS NOT NULL THEN 'Reported GPA'
ELSE 'No GPA'
END AS gpa_status,
COUNT(*) as total_apps,
ROUND(
(COUNT(*) FILTER (WHERE status = 'Accepted')::numeric / COUNT(*) * 100),
2
) as acceptance_rate
FROM admissions
WHERE degree = 'Masters'
AND us_or_international = 'American'
GROUP BY gpa_status;
```

- Answer:

Program	With GPA	No GPA
Master's	68.48%	54.23%
PhD	27.14%	22.63%

- The CASE statement in this query tags all records with a 'gpa\_status' equal to either 'Reported GPA' or 'No GPA' depending on whether the GPA exists or is null, respectively. The WHERE condition in this query restricts the SELECT to only look at records where 'degree' is 'Masters' and 'us\_or\_international' is 'American', and the GROUPBY separates the remaining records into two separate groups based on their

‘gpa\_status’. The average acceptance rate of these two groups is calculated using the same method as the query in question 5. This same query was run again with ‘PhD’ instead of ‘Masters’ in the WHERE clause to get the same information for PhD program applicants as well.

## SCREENSHOTS

Analysis Webpage:

### TheGradCafe Analysis

Push Data

Pulls any new GradCafe results, appends them to your local files, and loads them into the database.

Update Analysis

Refreshes the analysis summary using the most up-to-date data currently available.

#### Overview

Total number of records: 42865

- How many applicants in the database applied for Fall 2026?  
Fall 2026 Applicants: 8924
- What percentage of entries are from international students?  
International applicants: 50.32%
- What is the average GPA of American students in Fall 2026?  
Average GPA: 3.80
- What percent of entries for Fall 2025 are Acceptances?  
Fall 2025 Acceptance Rate: 35.98%
- What is the average GPA of accepted applicants for the Fall 2026 term?  
Average GPA: 3.79
- How many applicants applied to JHU for a Master's degree in Computer Science?  
JHU CS Master's Applicants: 8
- How many entries added in 2026 are acceptances from applicants who applied to Georgetown University, MIT, Stanford University, or Carnegie Mellon University for a PhD in Computer Science?  
2026 CS PhD Acceptances from (Georgetown/MIT/Stanford/MIT/Carnegie Mellon): 29  
Same information relying on LLM fields only: 29  
Same information relying on raw fields only: 7

#### Average Metrics

Fall 2025 Acceptance Rate: 35.98%

- What is the average GPA of accepted applicants for the Fall 2026 term?  
Average GPA: 3.79
- How many applicants applied to JHU for a Master's degree in Computer Science?  
JHU CS Master's Applicants: 8
- How many entries added in 2026 are acceptances from applicants who applied to Georgetown University, MIT, Stanford University, or Carnegie Mellon University for a PhD in Computer Science?  
2026 CS PhD Acceptances from (Georgetown/Stanford/MIT/Carnegie Mellon): 29  
Same information relying on LLM fields only: 29  
Same information relying on raw fields only: 7

#### Average Metrics

Metric	Average
GPA	3.74
GRE (Quant)	164.7
GRE (Verbal)	159.8
GRE (AW)	4.36

#### Additional Analysis

- How many applicants applied for a Fall 2025 start term vs. a Spring 2025 start term?  
Fall 2025: 20023  
Spring 2025: 171
- What is the average acceptance rate of Master's programs vs PhD programs for American applicants who do and don't report their GPA?  
Acceptance Rates (American Applicants)

Program	With GPA	No GPA
Master's	68.48%	54.23%
PhD	27.14%	22.63%

query\_data.py Console Output:

There are 42865 entries

There were 8924 Fall 2026 applicant entries.

There were 50.32% International applicant entries.

Average metrics - GPA: 3.74 | GRE: 164.7 | GRE V: 159.8 | GRE AW: 4.36

Fall 2026 American Applicant Avg GPA: 3.80

Fall 2025 Acceptance Rate: 35.98%

Avg GPA of Fall 2026 Acceptances: 3.79

Number of JHU Master's CS applicants: 8

2026 PhD Acceptances from Georgetown, MIT, Stanford, and Carnegie Mellon: 29

Same info relying on LLM generated fields only: 29

Same info relying on raw fields only: 7

Fall 2025 Applicants: 20023

Spring 2025 Applicants: 171

Master's program acceptance rates with/without reported GPA...

With GPA: 68.48% | No GPA: 54.23%

PhD program acceptance rates with/without reported GPA...

With GPA: 27.14% | No GPA: 22.63%