When you run query_data:

```
Q1 (Count Fall 2026): 6574
Q2 (% International Applicants): 50.54%
Q3 (Averages of scores):
   avg_gpa   = 3.75
   avg_gre_q = 165.47
   avg_gre_v = 160.01
   avg_gre_aw = 4.35
Q4 (Avg GPA American Fall 2026): 3.79
Q5 (% Acceptances Fall 2026): 23.76%
Q6 (Avg GPA Fall 2026 Acceptances): 3.76
Q7 (JHU Masters CS count) 6
Q8 (2026 PhD CS Acceptances at GU/MIT/Stanford/CMU): 2
Q9 (2026 PhD CS Acceptances at GU/MIT/Stanford/CMU using LLM fields): 2
Extra Q1: What are the top 3 most common universities/program names in the dataset?
   64  University of Chicago - Physics
   59  Northwestern University - Philosophy
   57  Princeton University - Computer Science
Extra Q2: How many people got rejected from JHU? 228


Process finished with exit code 0
```

Description of queries used and why:

- All queries run against a single PostgreSQL table named applicants

Fetch methods used:
- fetch_one(sql, params=()): Executes a query and returns the first column of the first row (cur.fetchone()[0]). This is used when the query returns a single value like a count, a single percentage, or a single average.
- fetch_row(sql, params=()): Executes a query and returns the entire first row (cur.fetchone()) which can contain multiple columns. This is used when multiple values are needed at once from a single row such as multiple averages in one query.
- fetch_all(sql, params=()): Executes a query and returns all rows(cur.fetchall()). This is used when the query returns a list of results, such as rankings or grouped counts.

Q1: Parameters used = %s = "Fall 2026" to match requirements. SELECT COUNT(*) = counts all matching rows. Used WHERE to filter for term = Fall 2026. Fetch_one was used because only one single value was returned.

Q2: CASE WHEN … THEN 1 ELSE 0 is used to convert "international vs not" into a numeric flag. SUM(...) counts how many are international and then divides it by COUNT(*) to give the proportion. I then multiplied it by 100 to make it a percentage. NULLIF(COUNT(*), 0) prevents division by zero. ROUND(....,2) formats to 2 decimal points for requirements. Fetch_one was used because only one single value was returned.

Q3: WHERE is used to only consider the rows where the value is missing OR keep the row if the value is between certain numbers. This was done to make sure that if a row is missing GPA

but still has valid GRE scores, they are still included for GRE averages. I included the part between certain digits so it can return a realistic range check to reduce outliers/values that can distort averages. Then the AVG(column) only computes the non-null values. ::numeric is used to support decimal rounding. And then ROUND(value, 2) is to make sure it is rounded to 2 decimal points. Fetch_row was used because it returned the entire tuple.

Q4: Parameters: %s = 'Fall 2026', %s = 'American'. Used ROUND(value, 2) to make sure the returned value is to 2 decimal points. I also used ::numeric to cast the result into a numeric type so that it can support decimal rounding. WHERE is used to filter through the parameters set. I also made sure to include where the gpa IS NOT NULL so we only have averages with real gpa values. Fetch_one was used since one average gpa value was returned.

Q5: Parameters: %s = 'Fall 2026'. This was the same pattern as Q2. CASE WHEN … THEN 1 ELSE 0 is used to convert "Accepted" into a numeric flag. SUM(...) counts how many are accepted and then divides it by COUNT(*) to give the proportion. I then multiplied it by 100 to make it a percentage. NULLIF(COUNT(*), 0) prevents division by zero. ROUND(....,2) formats to 2 decimal points for requirements. Restricted to Fall 2026. Fetch_one was used because only one single value was returned.

Q6: Parameter: %s = '%Fall 2026%', %s = '%Accepted%. Same application of ROUND(value, 2) was used. This time ILIKE was used to make matching case-insensitive. This makes the filters more flexible if there are small inconsistencies in text formatting. COALESCE(field, '') was used to prevent errors when values are NULL (treats missing text as empty strings). AVG(gpa) was used so only averages from accepted applicants with valid GPA are returned for Fall 2026. Fetch_one was used since it returned one average gpa value.

Q7: Parameters: %s = "%master%", %s = '%computer science%', %s = '%johns hopkins%', %s = '%jhu%'. ILIKE handles case differences. Filters on degree containing "master", program containing "computer science" and program containing either "johns hopkins" or "jhu" to catch common naming variations. fetch_one was used since one single count was returned.

Q8: Parameters: %s = '%2026%', %s = 'Accepted', %s = '%phd%', %s = '%computer science%', %s = '%georgetown%', %s = '%Massachusetts Institute of Technology%', %s = '%mit%', %s = '%stanford%'. %s = '%carnegie mellon%'. ILIKE used to handle case differences. This query was used to use multiple filters to match the requested subgroup. Multiple patterns were used to handle naming differences. fetch_one was used since only one count was returned.

Q9: Parameters: %s = '%2026%', %s = 'Accepted', %s = '%phd%', %s = '%computer science%', %s = '%georgetown%', %s = '%massachusetts institute of technology%', %s = '%mit%', %s = '%stanford%', %s = '%carnegie mellon%. Same concept as Q8 except this time it is using the llm_generated_program and llm_generated_university instead of the raw program. I also used COALESCE to protect against missing values so the query won't break on nulls. fetch_one was used since one count was returned.

Extra_Q1: WHERE … IS NOT NULL AND … was used so it can remove programs that are missing or empty. This prevents meaningless categories from being in the results. GROUP BY groups all rows that share the same program value. Each unique program/university name becomes its own group. This allows us to count how frequently each program appears. COUNT(*) AS n counts the number of applicants within each program group. The n represents the frequency of that program in the dataset. ORDER BY n DESC sorts the grouped results from the most frequent to least frequent. LIMIT 3 makes sure to only give the 3 most common programs fetch_all was used since the query returned multiple rows and each row contained (program,count), so we need a list of rows.

Extra_Q2: Parameters: %s = 'Rejected', %s = '%johns hopkins%', %s = '%jhu%'. Filters rejected applicants from JHU. Used common patterns Johns Hopkins and JHU. COALESCE(status, '') avoids issues if status is missing. ILIKE was used to handle case differences. fetch_one was used since one count was returned.

Def main() is where all these results were printed.

Website view when you open link:

Grad School Cafe Data Analysis

Pull Data   Update Analysis

How many entries have applied for Fall 2026?
Answer: Applicant count: 6574

What percentage are International (not American/Other)?
Answer: Percent International: 50.54%

What is the average GPA, GRE, GRE V, GRE AW of applicants who provided these metrics?
Answer: Avg GPA: 3.75, Avg GRE: 165.47, Avg GRE V: 160.01, Avg GRE AW: 4.35

What is the average GPA of American students in Fall 2026?
Answer: Avg GPA American: 3.79

What percent of Fall 2026 entries are Acceptances?
Answer: Acceptance percent: 23.76%

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

What is the average GPA of Fall 2026 applicants who are Acceptances?
Answer: Avg GPA Acceptances: 3.76

How many entries are from applicants who applied to JHU for a masters in Computer Science?
Answer: Count: 6

How many 2026 acceptances are for GU/MIT/Stanford/CMU PhD in CS?
Answer: Count: 2

How many 2026 acceptances are for GU/MIT/Stanford/CMU PhD in CS using LLM Generated fields?
Answer: Count using LLM fields: 2

What are the top 3 most common universities/program names in the dataset?
Answer: [('University of Chicago - Physics', 64), ('Northwestern University - Philosophy', 59), ('Princeton University - Computer Science', 57)]

How many people got rejected from JHU?
Answer: 228

Once you click Pull Data button:

Grad School Cafe Data Analysis

Pull Data   Update Analysis

Pulling new data... please wait.

How many entries have applied for Fall 2026?
Answer: Applicant count: 6574

What percentage are International (not American/Other)?
Answer: Percent International: 50.54%

What is the average GPA, GRE, GRE V, GRE AW of applicants who provided these metrics?
Answer: Avg GPA: 3.75, Avg GRE: 165.47, Avg GRE V: 160.01, Avg GRE AW: 4.35

What is the average GPA of American students in Fall 2026?
Answer: Avg GPA American: 3.79

What percent of Fall 2026 entries are Acceptances?
Answer: Acceptance percent: 23.76%

After Pull Data stops running:

Grad School Cafe Data Analysis

Pull Data   Update Analysis

Pull complete! Click 'Update Analysis' to refresh results.

How many entries have applied for Fall 2026?
Answer: Applicant count: 6574

What percentage are International (not American/Other)?
Answer: Percent International: 50.54%

What is the average GPA, GRE, GRE V, GRE AW of applicants who provided these metrics?
Answer: Avg GPA: 3.75, Avg GRE: 165.47, Avg GRE V: 160.01, Avg GRE AW: 4.35

What is the average GPA of American students in Fall 2026?
Answer: Avg GPA American: 3.79

What percent of Fall 2026 entries are Acceptances?
Answer: Acceptance percent: 23.76%

When you click Update Analysis:

Grad School Cafe Data Analysis

Pull Data   Update Analysis

Analysis updated.

How many entries have applied for Fall 2026?
Answer: Applicant count: 9735

What percentage are International (not American/Other)?
Answer: Percent International: 50.18%

What is the average GPA, GRE, GRE V, GRE AW of applicants who provided these metrics?
Answer: Avg GPA: 3.75, Avg GRE: 165.46, Avg GRE V: 160.08, Avg GRE AW: 4.30

What is the average GPA of American students in Fall 2026?
Answer: Avg GPA American: 3.80

What percent of Fall 2026 entries are Acceptances?
Answer: Acceptance percent: 26.25%

Grad School Cafe Data Analysis

What is the average GPA of Fall 2026 applicants who are Acceptances?
Answer: Avg GPA Acceptances: 3.77

How many entries are from applicants who applied to JHU for a masters in Computer Science?
Answer: Count: 6

How many 2026 acceptances are for GU/MIT/Stanford/CMU PhD in CS?
Answer: Count: 18

How many 2026 acceptances are for GU/MIT/Stanford/CMU PhD in CS using LLM Generated fields?
Answer: Count using LLM fields: 2

What are the top 3 most common universities/program names in the dataset?
Answer: [('University of Chicago - Physics', 68), ('Stanford University - Physics', 64), ('University of Washington - Computer Science', 63)]

How many people got rejected from JHU?
Answer: 232