Jon Robert N. Boland – Module 3

SQL Data Analysis

Query Responses

```
1. Applicant count:   33
2. International percentage:   56.75%
3. Average GPA:   3.74, Average GRE:268.29,
Average GRE V: 158.79, Average GRE AW: 8.58
4. Average US Fall 2024 GPA:   3.63
5. Percentage of Fall 2024 Acceptances:   63.64%
6. Average GPA of Fall 2024 Acceptances:   3.58
7. Entries for JHU Computer Science Applicants:   13
```

Before getting into specific queries, I should specify that many of them rely on a VIEW of all entries for a given semester. In this application, Fall 2024 is referenced many times. Rather than stacking 3+ WHERE conditionals, I figured it would be more efficient to have a VIEW that did the semester filtering to clean up the SQL.

1. How many entries do you have in your database who have applied for Fall 2024?

SELECT COUNT(*) FROM applicants WHERE term = %s;

I selected this query because it's a standard SQL function for counting entries with a simple WHERE statement. I made it into a function with a variable for future flexibility.

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

SELECT DISTINCT({field}) FROM {table};
SELECT COUNT({field}) FROM {table};

SELECT COUNT(*) FROM {table} WHERE {field} = %s;

For this query, I first generated a VIEW of all entries for a given semester as stated above. This view is passed into each {table} variable. The DISTINCT query first generates a list of all unique entries. The COUNT query then gets a total of all entries. The third statement, COUNT with a conditional, is part of an iteration that goes through each DISTINCT field and computes its count. This is then divided by the total count to get a percentage of each value. For the assignment, I then simply looked up the "International" key in the resultant dictionary.

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

SELECT AVG({field}) FROM {table}

Again, the SQL AVG function did most of the work for me. This wasn't conditional, so I passed in the basic table without a view. Fortunately, AVG organically ignores NULL fields.

4. What is their average GPA of American students in Fall 2024?

SELECT AVG({field}) FROM {table} WHERE {conditional_column} = %s;

Similar to question 3, with a WHERE statement where I specify American students. The fall 2024 VIEW is passed in lieu of the table to provide that specific filter.

5. What percent of entries for Fall 2024 are Acceptances (to two decimal places)?

SELECT COUNT(*) FROM {table}

SELECT COUNT(*) FROM {table} WHERE status LIKE %s

The first statement counts all applicants. The second uses a wildcard '%Accepted' to look for accepted applicants. Since I loaded the data, I know that every accepted applicant starts with Accepted and then has the accepted date following it. I t hen divide to get the percentage! Again, a Fall 2024 VIEW is passed in for the filter.

6. What is the average GPA of applicants who applied for Fall 2024 who are Acceptances?

SELECT AVG({field}) FROM {table} WHERE {conditional_column} LIKE %s

Once more, the Fall 2024 VIEW comes in handy! The passed in field to AVG is GPA, and the WHERE LIKE wildcard is '%Accepted' again.

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

SELECT COUNT(*) FROM {table} WHERE program LIKE %s AND program LIKE %s;

The first WHERE LIKE wildcard is '%Johns Hopkins%' and the second is '%Computer Science%'. I looked through the database, and any reference to Johns Hopkins is formatted similarly to how I state it. The same goes for the Computer Science degree.