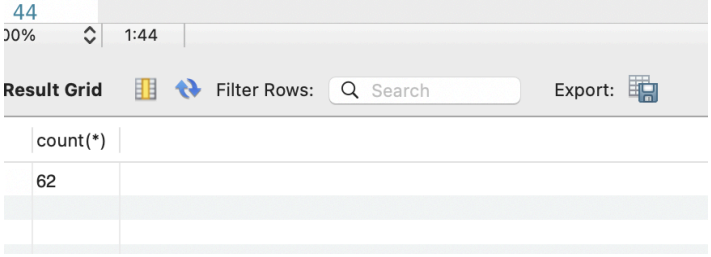
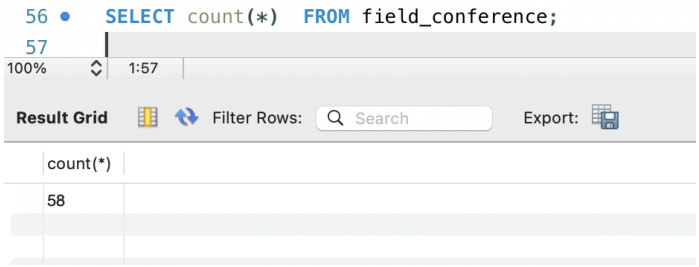
Question 1:

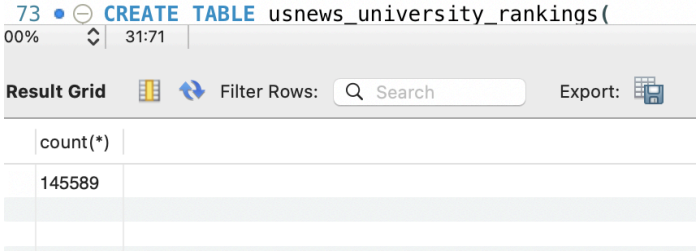
```
43 •  SELECT count(*)  FROM conference_ranking;
44
```



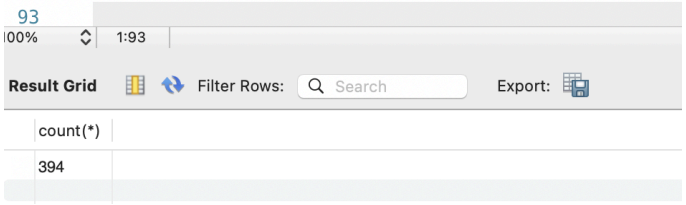

```
56 •  SELECT count(*)  FROM field_conference;
57
```



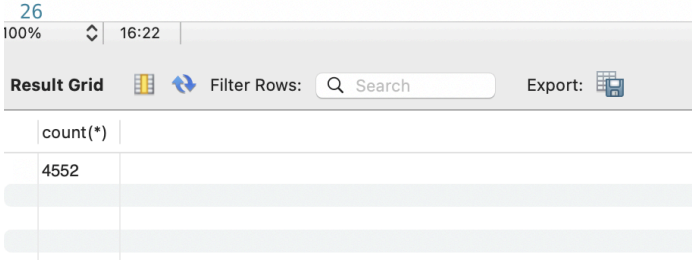

```
71 •  SELECT count(*)  FROM pub_info;
72
73 •  CREATE TABLE usnews_university_rankings(
```




```
92 •  SELECT count(*)  FROM usnews_university_rankings;
93
```




```
25 •  SELECT count(*)  FROM Author;
26
```



Question 2:

SQL Query:

```
CREATE TABLE IF NOT EXISTS Astar_Based_Rating(
      university_name VARCHAR(50) PRIMARY KEY,
      score FLOAT NOT NULL,
      ranking INT NOT NULL
    );
CREATE TABLE IF NOT EXISTS Balanced_Rating(
      university_name VARCHAR(50) PRIMARY KEY,
      score FLOAT NOT NULL,
      ranking INT NOT NULL);
CREATE TABLE IF NOT EXISTS General_Rating(
      university_name VARCHAR(50) PRIMARY KEY,
      score FLOAT NOT NULL,
      ranking INT NOT NULL);

INSERT INTO Astar_Based_Rating (university_name, score, ranking)
SELECT
a.affiliation AS university_name,
(0.4 * SUM(CASE WHEN cr.rank = 'A*' THEN p.count ELSE 0 END) +
0.4 * COUNT(DISTINCT CASE WHEN cr.rank = 'A*' THEN a.name END) +
0.1 * SUM(p.count) +
0.1 * COUNT(DISTINCT a.name)) AS score,
RANK() OVER(ORDER BY (0.4 * SUM(CASE WHEN cr.rank = 'A*' THEN p.count ELSE 0 END)
+
0.4 * COUNT(DISTINCT CASE WHEN cr.rank = 'A*' THEN a.name END) +
0.1 * SUM(p.count) +
0.1 * COUNT(DISTINCT a.name)) DESC) AS ranking
FROM
        author a
JOIN
        pub_info p ON a.name = p.name
JOIN
        conference_ranking cr ON p.conference = cr.conf_abbr
WHERE
        p.year BETWEEN 2014 AND 2024
GROUP BY
        a.affiliation;

SELECT * FROM Astar_Based_Rating
WHERE ranking <= 50
ORDER BY ranking;
```

university_name	score	ranking
Carnegie Mellon University	1037	1
University of California, San Diego	619.3	2
University of California, Berkeley	603.1	3
Massachusetts Institute of Technology	531.2	4
Georgia Institute of Technology	525.9	5
University of Washington	503.5	6
Stanford University	481.4	7
Cornell University	386.2	8
University of Pennsylvania	384.2	9
University of California, Los Angeles	378.3	10
Northeastern University	377.6	11
University of Texas at Austin	374.1	12
New York University	355.9	13
Princeton University	321.4	14
Columbia University	294	15
University of Southern California	292.2	16
Duke University	283.4	17
University of Chicago	281.2	18
University of California, Santa Barbara	235.8	19
Arizona State University	230.2	20
University of Virginia	223.7	21
University of California, Irvine	211.7	22
Johns Hopkins University	208.2	23
Harvard University	201.6	24
Texas A&M University	179.2	25
Boston University	178.6	26
North Carolina State University	177.1	27
University of California, Riverside	175.8	28
Rice University	170.2	29
Yale University	167.5	30
Northwestern University	162.1	31
University of Utah	161.9	32
University of California, Santa Cruz	153.4	33
George Mason University	153.4	33
University of California, Davis	152.5	35
University of Central Florida	151.6	36
Brown University	138.9	37
Virginia Tech	130.9	38
Oregon State University	127.7	39
Michigan State University	125.5	40
University of Notre Dame	124.7	41
University of Rochester	120.4	42
Washington University in St. Louis	104.8	43
Rochester Institute of Technology	98.2	44
University of Colorado Boulder	97.7	45
Emory University	96.4	46
Worcester Polytechnic Institute	88.2	47
University of Pittsburgh	74	48
University of California, Merced	70.9	49
Georgetown University	70.6	50
NULL	NULL	NULL

```
INSERT INTO Balanced_Rating (university_name, score, ranking)
SELECT
a.affiliation AS university_name,
(0.25 * SUM(CASE WHEN cr.rank = 'A*' THEN p.count ELSE 0 END) +
0.25 * COUNT(DISTINCT CASE WHEN cr.rank = 'A*' THEN a.name END) +
0.25 * SUM(p.count) +
0.25 * COUNT(DISTINCT a.name)) AS score,
RANK() OVER(ORDER BY (0.25 * SUM(CASE WHEN cr.rank = 'A*' THEN p.count ELSE 0
END) +
0.25 * COUNT(DISTINCT CASE WHEN cr.rank = 'A*' THEN a.name END) +
0.25 * SUM(p.count) +
0.25 * COUNT(DISTINCT a.name)) DESC) AS ranking
FROM
        author a
JOIN
        pub_info p ON a.name = p.name
JOIN
        conference_ranking cr ON p.conference = cr.conf_abbr
WHERE
        p.year BETWEEN 2014 AND 2024
GROUP BY
        a.affiliation;

SELECT * FROM Balanced_Rating
WHERE ranking <= 50
ORDER BY ranking;
```

university_name	score	ranking
Carnegie Mellon University	1175	1
University of California, San Diego	745.75	2
University of California, Berkeley	697	3
Georgia Institute of Technology	608.25	4
University of Washington	596.5	5
Massachusetts Institute of Technology	594.5	6
Stanford University	569	7
Northeastern University	437.75	8
University of Pennsylvania	432.5	9
University of California, Los Angeles	429.75	10
Cornell University	429.25	11
University of Texas at Austin	429	12
New York University	403	13
Princeton University	366.25	14
University of Southern California	339.75	15
Columbia University	338.25	16
Duke University	323.75	17
University of Chicago	322	18
University of California, Santa Barbara	272.25	19
Arizona State University	260.5	20
Johns Hopkins University	260.25	21
University of California, Irvine	249.5	22
University of Virginia	243.5	23
Harvard University	225.75	24
University of California, Riverside	216.75	25
North Carolina State University	216.25	26
Boston University	204.25	27
Texas A&M University	199	28
University of Utah	194.75	29
University of Central Florida	193.75	30
Rice University	187.75	31
Northwestern University	187	32
George Mason University	184.75	33
University of California, Davis	181	34
Yale University	180.25	35
University of California, Santa Cruz	172.75	36
Virginia Tech	163	37
Brown University	162	38
Michigan State University	152.5	39
Oregon State University	146.75	40
University of Notre Dame	144.5	41
University of Rochester	144.25	42
University of Colorado Boulder	119.75	43
Rochester Institute of Technology	112.75	44
Washington University in St. Louis	112	45
Emory University	104.5	46
University of California, Merced	98.5	47
Worcester Polytechnic Institute	95.25	48
University of Pittsburgh	86.75	49
University of Connecticut	78.5	50
NULL	NULL	NULL

```
INSERT INTO General_Rating (university_name, score, ranking)
SELECT
a.affiliation AS university_name,
(0.1 * SUM(CASE WHEN cr.rank = 'A*' THEN p.count ELSE 0 END) +
0.1 * COUNT(DISTINCT CASE WHEN cr.rank = 'A*' THEN a.name END) +
0.4 * SUM(p.count) +
0.4 * COUNT(DISTINCT a.name)) AS score,
RANK() OVER(ORDER BY (0.1 * SUM(CASE WHEN cr.rank = 'A*' THEN p.count ELSE 0 END)
+
0.1 * COUNT(DISTINCT CASE WHEN cr.rank = 'A*' THEN a.name END) +
0.4 * SUM(p.count) +
0.4 * COUNT(DISTINCT a.name)) DESC) AS ranking
FROM
        author a
JOIN
        pub_info p ON a.name = p.name
JOIN
        conference_ranking cr ON p.conference = cr.conf_abbr
WHERE
        p.year BETWEEN 2014 AND 2024
GROUP BY
        a.affiliation;

SELECT * FROM General_Rating
WHERE ranking <=50
ORDER BY ranking;
```

university_name	score	ranking
Carnegie Mellon University	1313	1
University of California, San Diego	872.2	2
University of California, Berkeley	790.9	3
Georgia Institute of Technology	690.6	4
University of Washington	689.5	5
Massachusetts Institute of Technology	657.8	6
Stanford University	656.6	7
Northeastern University	497.9	8
University of Texas at Austin	483.9	9
University of California, Los Angeles	481.2	10
University of Pennsylvania	480.8	11
Cornell University	472.3	12
New York University	450.1	13
Princeton University	411.1	14
University of Southern California	387.3	15
Columbia University	382.5	16
Duke University	364.1	17
University of Chicago	362.8	18
Johns Hopkins University	312.3	19
University of California, Santa Barbara	308.7	20
Arizona State University	290.8	21
University of California, Irvine	287.3	22
University of Virginia	263.3	23
University of California, Riverside	257.7	24
North Carolina State University	255.4	25
Harvard University	249.9	26
University of Central Florida	235.9	27
Boston University	229.9	28
University of Utah	227.6	29
Texas A&M University	218.8	30
George Mason University	216.1	31
Northwestern University	211.9	32
University of California, Davis	209.5	33
Rice University	205.3	34
Virginia Tech	195.1	35
Yale University	193	36
University of California, Santa Cruz	192.1	37
Brown University	185.1	38
Michigan State University	179.5	39
University of Rochester	168.1	40
Oregon State University	165.8	41
University of Notre Dame	164.3	42
University of Colorado Boulder	141.8	43
Rochester Institute of Technology	127.3	44
University of California, Merced	126.1	45
Washington University in St. Louis	119.2	46
Emory University	112.6	47
Worcester Polytechnic Institute	102.3	48
University of Pittsburgh	99.5	49
University of Connecticut	87.5	50
Dartmouth College	87.5	50
NULL	NULL	NULL

Question 3:

```
Astar = '''CREATE TABLE IF NOT EXISTS Astar_Based_Rating(
       university_name VARCHAR(50) PRIMARY KEY,
       score FLOAT NOT NULL,
       ranking INT NOT NULL
     );'''
cursor.execute('''DROP TABLE IF EXISTS Astar_Based_Rating;''')
cursor.execute(Astar)
df = pd.read_csv('/Users/ingridamerica/Documents/Database/ALLDATA/star.csv')
df.to_sql('Astar_Based_Rating',connection, if_exists='replace', index=False)

balanced = '''CREATE TABLE IF NOT EXISTS Balanced_Rating(
       university_name VARCHAR(50) PRIMARY KEY,
       score FLOAT NOT NULL,
       ranking INT NOT NULL);'''
cursor.execute('''DROP TABLE IF EXISTS Balanced_Rating;''')
cursor.execute(balanced)
df = pd.read_csv('/Users/ingridamerica/Documents/Database/ALLDATA/balanced.csv')
df.to_sql('Balanced_Rating',connection, if_exists='replace', index=False)

general = '''CREATE TABLE IF NOT EXISTS General_Rating(
       university_name VARCHAR(50) PRIMARY KEY,
       score FLOAT NOT NULL,
       ranking INT NOT NULL); '''
cursor.execute('''DROP TABLE IF EXISTS General_Rating;''')
cursor.execute(general)
df = pd.read_csv('/Users/ingridamerica/Documents/Database/ALLDATA/general.csv')
df.to_sql('General_Rating',connection, if_exists='replace', index=False)

usnews_df = pd.read_sql_query("SELECT university_name, rank FROM
usnews_university_rankings WHERE rank <= 52", connection)
astar_df = pd.read_sql_query("SELECT university_name, ranking FROM
Astar_Based_Rating", connection)
balanced_df = pd.read_sql_query("SELECT university_name, ranking FROM
Balanced_Rating", connection)
general_df = pd.read_sql_query("SELECT university_name, ranking FROM General_Rating",
connection)

# Function to compute Kendall Tau Correlation
def kendall_tau_correlation(list1, list2):
   union_universities = set(list1.keys()).union(set(list2.keys()))
```

```
    n_c, n_d = 0, 0
    n = len(union_universities)
    for u_i in union_universities:
        for u_j in union_universities:
            if u_i == u_j:
                continue

            rank1_i = list1.get(u_i, float('inf'))
            rank1_j = list1.get(u_j, float('inf'))
            rank2_i = list2.get(u_i, float('inf'))
            rank2_j = list2.get(u_j, float('inf'))

            if ((rank1_i < rank1_j and rank2_i < rank2_j) or
                (rank1_i > rank1_j and rank2_i > rank2_j) or
                (rank1_i == rank1_j and rank2_i == rank2_j)
            ):
                n_c += 1
            else:
                n_d += 1

    total_pairs = n * (n - 1) / 2
    if total_pairs == 0:
        return 0  # Avoid division by zero
    return (n_c - n_d) / total_pairs

# Convert dataframes to dictionaries
usnews_rankings = dict(zip(usnews_df['university_name'], usnews_df['rank']))
astar_rankings = dict(zip(astar_df['university_name'], astar_df['ranking']))
balanced_rankings = dict(zip(balanced_df['university_name'], balanced_df['ranking']))
general_rankings = dict(zip(general_df['university_name'], general_df['ranking']))

# Compute correlations
astar_correlation = kendall_tau_correlation(astar_rankings, usnews_rankings)
balanced_correlation = kendall_tau_correlation(balanced_rankings, usnews_rankings)
general_correlation = kendall_tau_correlation(general_rankings, usnews_rankings)

# Output results
print(f"Kendall Tau Correlation (Astar vs US News): {astar_correlation}")
print(f"Kendall Tau Correlation (Balanced vs US News): {balanced_correlation}")
print(f"Kendall Tau Correlation (General vs US News): {general_correlation}")
```

Output:

```
Kendall Tau Correlation (Astar vs US News): 0.12030755314337403
Kendall Tau Correlation (Balanced vs US News): 0.061457418788410885
Kendall Tau Correlation (General vs US News): 0.06672519754170325
```