The data for this exercise has been derived from the Indeed Data Scientist/Analyst/Engineer [dataset](https://www.kaggle.com/elroyggj/indeed-dataset-data-scientistanalystengineer) on kaggle.com

**Provide the SQL queries and answers for the following questions/tasks using the data\_analyst\_jobs table you have created in PostgreSQL:**

1. How many rows are in the data\_analyst\_jobs table?

–count rows-- Select COUNT (\*) FROM data\_analyst\_jobs;

1793

1. Write a query to look at just the first 10 rows. What company is associated with the job posting on the 10th row?

Select \* FROM data\_analyst\_jobs LIMIT 10

Exxon Mobil

1. How many postings are in Tennessee? How many are there in either Tennessee or Kentucky?

TN - Select \* FROM data\_analyst\_jobs WHERE location = 'TN' **21**

TN and KY - Select \* FROM data\_analyst\_jobs WHERE location = 'TN' OR location = 'KY' **27**

1. How many postings in Tennessee have a star rating above 4?

Select \* From data\_analyst\_jobs WHERE location =’TN’ AND star\_rating >4

**3**

1. How many postings in the dataset have a review count between 500 and 1000?

Select \* FROM data\_analyst\_jobs WHERE review\_count BETWEEN 500 and 1000

**151**

1. Show the average star rating for each state. The output should show the state as state and the average rating for the state as avg\_rating. Which state shows the highest average rating?

Select AVG(star\_rating) AS avg\_rating, location AS state FROM data\_analyst\_jobs WHERE star\_rating IS NOT null GROUP BY location ORDER BY avg\_rating DESC

**Nebraska**

1. Select unique job titles from the data\_analyst\_jobs table. How many are there?

SELECT DISTINCT title FROM data\_analyst\_jobs

**881**

1. How many unique job titles are there for California companies?

WHERE location = ‘CA’

**230**

1. Find the name of each company and its average star rating for all companies that have more than 5000 reviews across all locations. How many companies are there with more that 5000 reviews across all locations?

SELECT AVG(star\_rating), title FROM data\_analyst\_jobs WHERE review\_count >5000 GROUP BY title

**41**

1. Add the code to order the query in #9 from highest to lowest average star rating. Which company with more than 5000 reviews across all locations in the dataset has the highest star rating? What is that rating?

ORDER BY star\_rating DESC

**AMERICAN EXPRESS 4.1999998090000000**

1. Find all the job titles that contain the word ‘Analyst’. How many different job titles are there?

SELECT DISTINCT titleFROM data\_analyst\_jobsWHERE title LIKE 'Analyst%'

**10**

1. How many different job titles do not contain either the word ‘Analyst’ or the word ‘Analytics’? What word do these positions have in common?

\*SELECT COUNT(title), title FROM data\_analyst\_jobsWHERE title NOT ILIKE '%Analyst%' AND title NOT ILIKE '%Analytics%'GROUP BY title\*/

**Tableau**