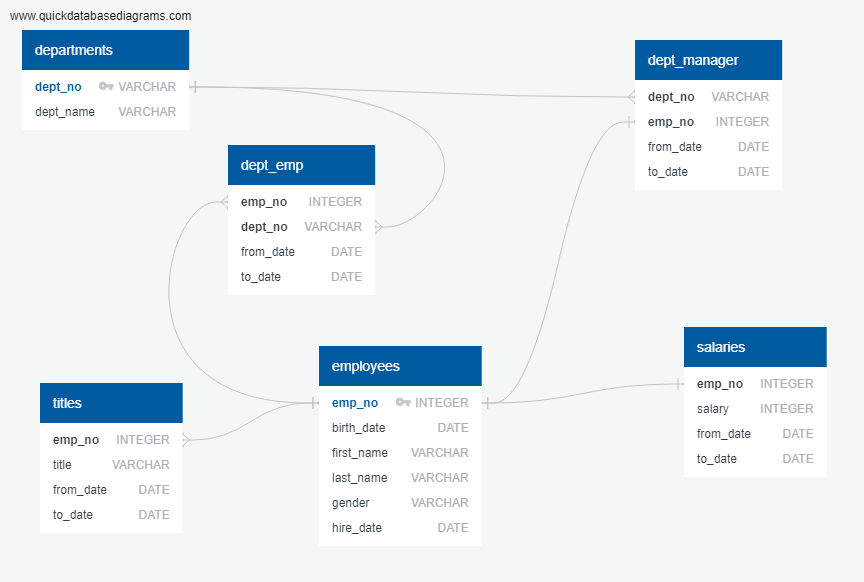
**## ABSTRACT**

In this project, I am aiming to make an employee database that will use data from CSV files. In this database, we are tasked with creating an ERD that represents the proper constraints and table structure. After the ERD is made and or employee database structure is planned out, I then created the tables for database and run a sequence of queries to extract key information from our Employee Database. After the queries have been ran, we will then use SQL Alchemy to connect to our Database and run a visualization that shows us average salary vs title.

**## ERD**



**## SQL QUERIES**

**# Check to see if my tables imported properly**

SELECT \* FROM departments;

SELECT \* FROM employees;

SELECT \* FROM dept\_emp;

SELECT \* FROM dept\_manager;

SELECT \* FROM salaries;

SELECT \* FROM titles;

**# 1. List the following details of each employee: employee number, last name, first name, gender, and salary.**

CREATE VIEW Question1 as

SELECT e.emp\_no, e.last\_name, e.first\_name, e.gender, s.salary

FROM employees as e

INNER JOIN salaries as s

ON (e.emp\_no = s.emp\_no);

SELECT \* FROM Question1;

--DROP VIEW Question1;

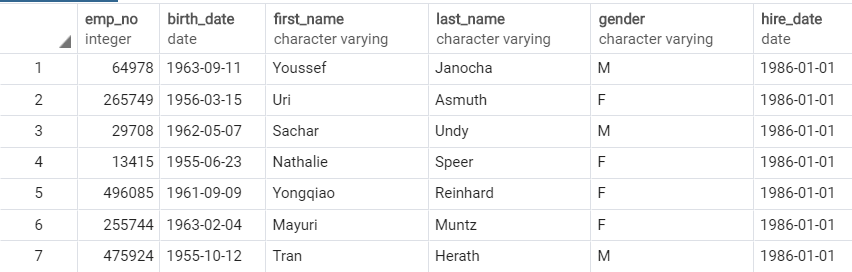


**# 2. List employees who were hired in 1986.**

SELECT \* FROM employees

WHERE hire\_date between '1986-01-01' and '1986-12-31'

ORDER BY hire\_date ASC;



**# 3. List the manager of each department with the following information: department number, department name, the manager's employee number, last name, first name, and start and end employment dates.**

CREATE VIEW Question3Join as

SELECT d.dept\_no, d.dept\_name, dm.emp\_no, dm.from\_date, dm.to\_date

FROM departments as d

INNER JOIN dept\_manager as dm

ON (d.dept\_no = dm.dept\_no);

CREATE VIEW Question3 as

SELECT q3j.dept\_no, q3j.dept\_name, q3j.emp\_no, e.last\_name, e.first\_name, q3j.from\_date, q3j.to\_date

FROM Question3Join as q3j

INNER JOIN employees as e

ON (q3j.emp\_no = e.emp\_no);

SELECT \* FROM Question3;

--DROP VIEW Question3Join;

--DROP VIEW Question3;



**# 4. List the department of each employee with the following information: employee number, last name, first name, and department name.**

CREATE VIEW Question4Join as

SELECT d.dept\_no, d.dept\_name, de.emp\_no, de.from\_date, de.to\_date

FROM departments as d

INNER JOIN dept\_emp as de

ON (d.dept\_no = de.dept\_no);

CREATE VIEW Question4 as

SELECT e.emp\_no, e.last\_name, e.first\_name, q4j.dept\_name

FROM Question4Join as q4j

INNER JOIN employees as e

ON (q4j.emp\_no = e.emp\_no);

SELECT \* FROM Question4;

--DROP VIEW Question4Join;

--DROP VIEW Question4;



**# 5. List all employees whose first name is "Hercules" and last names begin with "B."**

SELECT \* FROM employees

WHERE first\_name = 'Hercules' and last\_name LIKE 'B%';



**# 6. List all employees in the Sales department, including their employee number, last name, first name, and department name.**

CREATE VIEW Question6Join as

SELECT d.dept\_no, d.dept\_name, de.emp\_no, de.from\_date, de.to\_date

FROM departments as d

INNER JOIN dept\_emp as de

ON (d.dept\_no = de.dept\_no);

CREATE VIEW Question6 as

SELECT e.emp\_no, e.last\_name, e.first\_name, q6j.dept\_name

FROM Question6Join as q6j

INNER JOIN employees as e

ON (q6j.emp\_no = e.emp\_no);

SELECT \* FROM Question6

WHERE dept\_name = 'Sales';

--DROP VIEW Question6Join;

--DROP VIEW Question6;



**# 7. List all employees in the Sales and Development departments, including their employee number, last name, first name, and department name.**

CREATE VIEW Question7Join as

SELECT d.dept\_no, d.dept\_name, de.emp\_no, de.from\_date, de.to\_date

FROM departments as d

INNER JOIN dept\_emp as de

ON (d.dept\_no = de.dept\_no);

CREATE VIEW Question7 as

SELECT e.emp\_no, e.last\_name, e.first\_name, q7j.dept\_name

FROM Question7Join as q7j

INNER JOIN employees as e

ON (q7j.emp\_no = e.emp\_no);

SELECT \* FROM Question7

WHERE dept\_name = 'Sales' or dept\_name = 'Development';

--DROP VIEW Question7Join;

--DROP VIEW Question7;

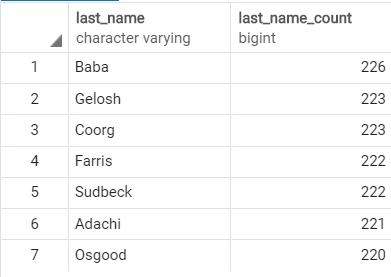


**# 8. In descending order, list the frequency count of employee last names, i.e., how many employees share each last name.**

SELECT last\_name, count(last\_name) as "last\_name\_count" FROM employees

GROUP BY last\_name

ORDER BY "last\_name\_count" DESC;



**# Bonus... "Search your ID number." You look down at your badge to see that your employee ID number is 499942.**

CREATE VIEW BonusJoin as

SELECT d.dept\_no, d.dept\_name, de.emp\_no, de.from\_date, de.to\_date

FROM departments as d

INNER JOIN dept\_emp as de

ON (d.dept\_no = de.dept\_no);

CREATE VIEW BonusJoin2 as

SELECT e.emp\_no, e.last\_name, e.first\_name, e.gender, e.hire\_date, bj.dept\_name, bj.from\_date, bj.to\_date

FROM BonusJoin as bj

INNER JOIN employees as e

ON (bj.emp\_no = e.emp\_no);

CREATE VIEW Bonus as

SELECT bj2.emp\_no, bj2.last\_name, bj2.first\_name, bj2.gender, s.salary, bj2.dept\_name, bj2.hire\_date

FROM BonusJoin2 as bj2

INNER JOIN salaries as s

ON (bj2.emp\_no = s.emp\_no);

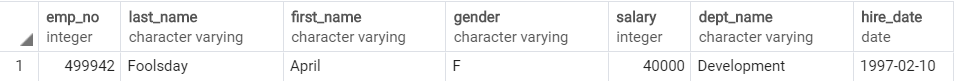
SELECT \* FROM Bonus

WHERE emp\_no = 499942;

--DROP VIEW BonusJoin;

--DROP VIEW BonusJoin2;

--DROP VIEW Bonus;



**# Build table for the bonus "Average Salary vs. Title" bar chart**

CREATE VIEW salary\_v\_title as

SELECT SUM(s.salary), tl.title, COUNT(tl.title)

FROM salaries as s

INNER JOIN titles as tl

ON (s.emp\_no = tl.emp\_no)

GROUP BY tl.title;

CREATE VIEW sal\_v\_tl\_bonus as

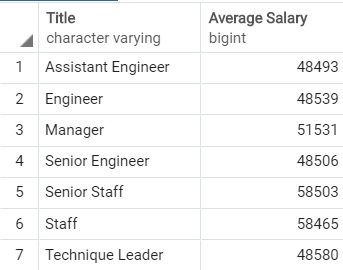
SELECT tl.title as "Title", (tl.sum/ tl.count) as "Average Salary"

FROM salary\_v\_title as tl;

SELECT \* FROM sal\_v\_tl\_bonus;

--DROP VIEW salary\_v\_title;

--DROP VIEW sal\_v\_tl\_bonus;



**## Average Salary vs. Title visualization**

**(Visualizations of the steps before the chart included in my Jupyter Notebook file)**

**# Import dependencies**

from sqlalchemy import create\_engine

import psycopg2

import pandas as pd

from db\_config import pg\_pswd

import matplotlib.pyplot as plt

**# Create engine and connection to connect to DB**

engine = create\_engine(f'postgresql+psycopg2://postgres:{pg\_pswd}@localhost:5432/SQL\_Employee\_Database')

conn = engine.connect()

**# Create DF out of the sal\_v\_tl\_bonus table**

salary\_v\_title = pd.read\_sql("SELECT \* FROM sal\_v\_tl\_bonus;", conn)

**# Check salary\_v\_title DB**

salary\_v\_title

**# Check datatypes**

salary\_v\_title.dtypes

**# Create and plot Salary vs. Title bar chart**

salary\_v\_title\_bar\_chart = salary\_v\_title.plot.bar(x='Title', y='Average Salary', rot=0, figsize = (15, 10))

**# Plot title and lables**

plt.xlabel("Title", fontsize = 17)

plt.ylabel("Average Salary", fontsize = 17)

plt.title(f"Average Salary vs. Title", fontsize = 30, fontweight="bold")

**# Export picture the Salary vs. Title bar chart to a png file**

plt.savefig("output\_data/Average\_Salary\_vs\_Title\_bar\_chart.png")

