**CASE STUDY # 1**

Please answer the SQL questions. There are several ways to write these SQL statement and any form ( PostgreSQL, MySQL, MS SQL) is acceptable as long as it generates the correct answers.

1. The salary table contains salary information of all employees in various departments of ABC Company. A partial view of the table is given below.

CREATE TABLE merged\_inner AS

SELECT e1.\*, e2.\*

FROM EE\_information e1

INNER JOIN EE\_Salary e2 ON e1.emp\_id = e2.emp\_id;

* 1. Write a SQL statement to find the highest salaried employee in each department.

SELECT emp\_id,

ROW\_NUMBER() OVER (PARTITION BY Department ORDER BY Yearly\_Salary DESC) AS SalaryRank

FROM

merged\_inner

WHERE

SalaryRank = 1

* 1. Write a SQL statement to calculate average salary per department.

SELECT

Department,

AVG(Yearly\_Salary) AS Average\_Salary\_Department

FROM

merged\_inner

GROUP BY

Department;

* 1. Write a SQL statement to find employees whose name starts with ‘J’.

Select \* from merged\_inner

WHERE

name LIKE ‘J%’

* 1. Write a SQL statement to find employees who have been in the company 5+ years.

SELECT

\*

FROM

merged\_inner

WHERE

DATEDIFF(YEAR, Hire\_Date, GETDATE()) >= 5;

* 1. Write a SQL statement to find the number of employees who work in Finance department

SELECT

COUNT(\*) AS Number\_Employees

FROM

merged\_inner

WHERE

Department = 'Finance';

* 1. Write a SQL statement to list employee names and their tenure in the company.

SELECT

name,

DATEDIFF(YEAR, Hire\_Date, GETDATE()) AS Tenure\_Years

FROM

merged\_inner;

* 1. Write a SQL statement to list employees with the highest tenure in each department.

WITH RankedEmployees AS (

SELECT

name,

Department,

DATEDIFF(YEAR, Hire\_Date, GETDATE()) AS TenureInYears,

ROW\_NUMBER() OVER (PARTITION BY Department ORDER BY DATEDIFF(YEAR, Hire\_Date, GETDATE()) DESC) AS RankInDepartment

FROM

merged\_inner

)

SELECT

name,

Department,

TenureInYears

FROM

RankedEmployees

WHERE

RankInDepartment = 1;

Table name: EE\_information



Table name: EE\_salary



1. Employee Benefit table contains an array of structure with Benefit code and deduction amount for each employee. DNT🡪 Dental Insurance deduction amount, HB🡪 Health Insurance Deduction Amount and FSA🡪 FSA deductiom amount. Write a SQL statement to calculate total amount of benefit deductions for each employee, e.g. John Doe’s total benefit deduction should be 200(50+150).

Table Name: EE\_Benefit



SELECT

name,

Emp\_ID,

COALESCE(JSON\_VALUE(Benefit, '$.DNT'), 0) + COALESCE(JSON\_VALUE(Benefit, '$.HB'), 0) + COALESCE(JSON\_VALUE(Benefit, '$.FSA'), 0 AS TotalBenefit

FROM

ee\_benefit;

**CASE STUDY # 2**

**Background:** ADP’s Digital team tracks behavior of web visitors on ADP.com website. The attached data gives information about web activity of our online visitors sourced from Paid Search over a month. This is a free form analysis and either R or Python can be used to run the analysis. Here are some general expectations:

* Analyze the data to identify patterns
* Identify anomalies/outlier in the data
* Try to predict which visitors are more likely to submit a lead form
* Create a power point presentation of 10-15 slides outlining the analysis process and final outcome
* Send over the power point slides and any code that was used to conduct the analysis
* Review the case study over the phone with hiring manager (15 minutes)

**Data Dictionary**

