**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.

------------------------create table based on the data provided------------------------------------------------------

**CREATE TABLE EE\_information (**

**Emp\_ID INT,**

**Hire\_Date DATE,**

**Termination\_Date DATE,**

**Department VARCHAR(255),**

**Title VARCHAR(255),**

**Home\_State VARCHAR(2)**

**);**

**INSERT INTO EE\_information (Emp\_ID, Hire\_Date, Termination\_Date, Department, Title, Home\_State)**

**VALUES**

**(2389, '2010-12-22', NULL, 'Accounting', 'Sr. Accountant', 'NJ'),**

**(4738, '2015-07-08', NULL, 'Finance', 'Jr. Financial Anal', 'NY'),**

**(1801, '2017-06-02', '2017-12-03', 'Sales Operations', 'Operations Analy', 'PA'),**

**(8927, '2018-03-02', NULL, 'Accounting', 'Accountant', 'CT');**

**CREATE TABLE EE\_Salary (**

**Name VARCHAR(255),**

**Emp\_ID INT,**

**DOB DATE,**

**Yearly\_Salary INT**

**);**

**INSERT INTO EE\_Salary (Name, Emp\_ID, DOB, Yearly\_Salary)**

**VALUES**

**('John Doe', 2389, '1959-03-01', 90000),**

**('Jane Smith', 4738, '1992-01-02', 65000),**

**('Abby Douglas', 1801, '1979-12-01', 75000),**

**('Jim Lee', 8927, '1988-10-28', 78000);**

**CREATE TABLE merged\_inner AS**

**SELECT e1.\*, e2.Name,e2.DOB, e2.Yearly\_Salary**

**FROM EE\_information e1**

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

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

**select \* from**

**(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( day, Hire\_Date, GETDATE()) >= 5\*365.25;**

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

**SELECT**

**Department,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**

**emp\_ID, 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



**CREATE TABLE EE\_benefit (**

**Name VARCHAR(255),**

**Emp\_ID INT,**

**Benefit JSON**

**);**

**INSERT INTO EE\_benefit (Name, Emp\_ID, Benefit)**

**VALUES**

**('John Doe', 2389, '{"DNT":50.0, "HB":150.0}'),**

**('Jane Smith', 4738, '{"HB":300.0}'),**

**('Abby Douglias', 1801, '{"DNT":50.0, "HB":150.0, "FSA":125.0}'),**

**('Jim Lee', 8927, '{"FSA":55.0}');**

**select \***

**,dnt+hb+fsa as total\_expense**

**from**

**(SELECT**

**Name,**

**Emp\_ID,**

**COALESCE(JSON\_EXTRACT(Benefit, '$.DNT'),0) AS DNT,**

**COALESCE(JSON\_EXTRACT(Benefit, '$.HB'),0) AS HB,**

**COALESCE(JSON\_EXTRACT(Benefit, '$.FSA'),0) AS FSA**

**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**

