**DAB203 - Final Project**

Section 1: *SQL and Excel for Data Analytics and Visualization*

*Description:* You are hired as a Data Analyst at ABC Company Ltd., Windsor ON and your manager has assigned you the task to mine the data set from year 1985 to 2005 and solve key business questions.

Section 2: *Purpose and Need*

Management team has asked for the insights regarding the following questions, so that appropriate decisions could be made on the hiring policies and hiring processes to achieve the goal of gender-neutral workplace.

* 1. *Key questions* –
     1. Management team want to know about the breakdown of the Male and Female employee per year from 1990 to 2002?
        1. Use stacked column chart to display your results, where *Employee Count* is on the Y-axis and *Years* on the X-Axis
     2. What is the number of Male Managers and the Female Managers for each year from 1990 to 2002? Ignore the manager’s department categorization.
        1. Plot the results on the Line chart, where *Employee Count* is on the Y-axis and *Years* on the X-Axis
     3. What is average salary of the Male employee and Female Employee in the entire company per year from 1990 to 2002?
        1. Plot the results on the Line chart or Clustered Column chart, where *Employee Salary* is on the Y-axis and *Years* on the X-Axis
     4. What is the average salary of the Male and Female employee per department?
        1. Plot the results on a graph of your choice.

Section 3: Implementation Method

3.1 *Tools required for this project*

3.1.1 Data Mining and Data Analysis tool – *Microsoft SQL server 2019 express edition*

3.1.2 Data Visualization – *Microsoft Excel*

3.2 *Data Set Details*

3.2.1 Download the *ProjectDAB203.zip* file from Black Board under section *Content > Final\_Project*. The zip file contains 6 data files and one SQL file to load the data–

3.2.2 *Data Files* –

3.2.2.1 employees.csv

3.2.2.2 departments.csv

3.2.2.3 dept\_emp.csv

3.2.2.4 dept\_manager.csv

3.2.2.5 salaries.csv

3.2.2.6 titles.csv

3.2.3 *SQL to Load data files*–

3.2.3.1 ProjectLoadSQLServer.sql

Section 4: Timeline

The deadline to submit project including SQLs queries, query outputs and visualization (graphs) in word or excel format is **Monday December 20th,2021 by 4:00 pm**.

--End--