INTERVIEW MINI-PROJECT: EMPLOYEE DASHBOARD

Purpose:
Test candidate's ability to:
- Write code to perform simple data manipulation
- Access data via an API
- Create a simple dashboard with most relevant information
- Analyze and explain key findings

Included Datasets:
(see 'explanation' tab within each .xlsx file for explanation of variables)
- Employee_Roster_Data.xlsx
- Email_Data.txt
- hours.xlsx
- skills.xlsx

Instruction:
- Read below Problem Statement and familiarize yourself with the attached datasets
- Create a root folder with the name "InterviewProject_<YourName>" (replace <YourName> with your actual full name).
- Inside the root folder include a single folder per task (task1, task2, task3, task4.doc)
- Please use Python as programming language.
- Upload all the files to your github personal account
- Send to armen.morcecian@rga.com the Git URL including all the files before the deadline.

Deadline: Sunday, March 7, 2021 (early submission is ok too!)

Submission Checklist:
Remember to include in the right folder the following items:
- Copy of formatted Employee_Roster, Hours and Skills tables (exports)
- Copy of all the code files used to perform this project.
- Power BI file.
- Document with your own analysis.

Problem Statement:
Company ABC has over 1,300 employees located in four offices around the globe. ABC would like to understand what factors, if any, correlate with employee performance. ABC would like to you to use email communication data as a proxy for employee collaboration and worked hours and skills to understand performance.

ABC has data on its employees contained within the Employee_Roster_Data.xlsx file, employee email data contained within Email_Data.txt, employee worked hours on each client within hours.csv and employee skills data within skills.xlsx (see each file for explanation of variables).

- Each row within Employee_Roster_Data describes attributes about an individual employee
- Each row within Email_Data describes an individual email sent by a person (either an employee or person not employed at ABC).
- Each row within hours describes the hours spent by an employee in admin tasks and on each different client. Also includes the target total hours.
- Each row within skills describes a skill level for a given employee.

You will see within Employee_Roster_Data that each employee has a unique Email_ID which corresponds to the from_ID and to_ID within Email_Data. If an ID within Email_Data does not appear within the Email_ID column of Employee_Roster_Data, it is because that individual person is not an employee at ABC (e.g. a vendor sending an email to an employee, or an employee sending an email to a customer).

The below tasks describe in detail what Company ABC would like you to do.

Task #1 - Format Data

Salary information is contained within **Employee_Roster_Data.xlsx** but is in different currencies. Please convert all salaries to USD using the **API** from http://fixer.io/. You should use the exchange rates as of **April 3, 2017** (do not use today's exchange rate). Important note: You must access the exchange rate data using Python; do not simply access the API from your web-browser and manually copy the exchange rate results. Please import the resulting dataset into a table in MS SQL Server Express Edition.

Task # 1 Answer:
Please create a folder inside "InterviewProject_<YourName>" with the name "task1" including the .py /.sql files you used to get data from the API and the upload script from Python into SQL Server.

Task #2 - Upload and transform data

Upload all the data into SQL Server Express Edition and create the data model. Please follow these steps:

- **Create a table** in SQL Server for each data source.
- **Upload** all the remaining files into MS SQL Server Express Edition.
- **Clean** each data source. Remove, if any, useless columns and special characters.
- Create a unified **dimension** for each attribute: Create a single and unified table for each attribute to represent all the possible values for an attribute in all the different data sources (For example, Users, Departments, Titles, etc.).
- **Alter** the existing tables by adding the following **fields using SQL Server**. Please create them inside the table you think they will fit:
    - Total Working Hours by Employee: 40 hours per week.
    - Utilization: % difference between the client hours and the total working hours minus admin hours
    - Client time: % difference between client hours and the total working hours
    - Admin Time: % difference between admin hours and the total working hours
    - Skills Level:
        - 0: wants to learn.
        - 1-2: Heavy supervision
        - 3-4: Light supervision
        - 5: Expert
- **Export** the final versions of hours, skills and Employee_Roaster_Data tables.

Task # 2 Answer:
Please create a folder inside ""InterviewProject_<YourName>"" with the name "task2" including the .py /.sql files with the scripts used to create the tables, upload the files and transform the data. Also, include the export of hours, skills and employee_roaster_data tables.

Task #3 - Create the dashboard

Now your work is to bring together in a single dashboard all the available information to show and explain what data says about all the assumptions made.

In order to do so, please follow these steps:

1. Create a single or multiple SQL Server **data connections** in Power BI.
2. Transform and format data in Power BI using **Power Query**.
3. Build a **data model** in power BI to represent the same data model built in SQL Server.
4. Create **calculated columns/measures** to fit reporting needs.
5. Create and design a **dashboard** in Power BI (It can have multiple tabs) to represent the most important findings.

Task # 3 Answer:
Please create a folder inside ““InterviewProject_<YourName>”” with the name “task3” including the SQL files used to transform the data and the Power BI file with the resulting dashboard.

Task #4 - Data Analysis

Now that the report is in place, company leadership wants to know **your thoughts** about the most important findings:

- In your opinion, which are the key findings about the employees?
- Is there any correlation between collaboration and worked hours?
- Based on findings, which are the right next steps?
- Do you recommend some additional data set to enrich the analysis?

Task # 4 Answer:
Please create a file called task4.doc inside the ““InterviewProject_<YourName>”” folder with the name task4 including your answer and any supporting image/screenshot for each question.