**Test de Conocimiento SQL Y Python IBM**

Objetivos:

1. Medir conocimiento de capacidad analítica.
2. Medir conocimiento de sentencias SQL.
3. Medir conocimiento de Python.

Procedimiento:

Se entregan dos archivos pilotos.csv y vuelos.csv de los cuales se obtienen datos con los siguientes esquemas.

Vuelos

root

|-- Codigo Piloto: integer (nullable = true)

|-- Aerolínea: integer (nullable = true)

|-- Origen: string (nullable = true)

|-- Destino: string (nullable = true)

|-- Minutos de retraso: integer (nullable = true)

|-- OnTime: string (nullable = true)

Pilotos

root

|-- Codigo Piloto: integer (nullable = true)

|-- Piloto: string (nullable = true)

**SENTENCIAS SQL**

Para medir el conocimiento en SQL puede realizar la carga de los archivos CSV en Python o un ambiente con base de datos SQL con el fin de que pueda ejecutar las sentencias SQL, también puede utilizar SPARK.SQL para lo cual se deja una guía de como configurarlo en cualquier IDE.

Se deben contestar las siguientes preguntas:

¿Qué aerolínea tiene más vuelos?

¿Qué Origen se repite más?

¿Desde donde vuela más la aerolínea 8?

¿Hacia dónde vuela más la aerolínea 4?

¿Qué piloto vuela más?

**PYTHON**

Para la prueba con Python se deben realizar los siguientes pasos

1. En primer instancia se deben cargar los archivos vuelos y pilotos a un dataframe
2. Agregar en la hoja Vuelos un campo para el nombre del piloto
3. Insertar el nombre del piloto
4. Descartar/marcar los registros donde Origen y Destino sean iguales.
5. Agregar comentario en ONTIME, si el tiempo en valor absoluto es menor o igual a 30 A, si es esta entre 30 y 50 B, si es mayor que 50 C.
6. ¿Quién es el piloto que tiene más vuelos A?
7. ¿Qué aerolínea tiene más vuelos C?
8. ¿Para qué aerolínea vuela Hung Cho?
9. ¿Cuántos vuelos A, B, C tiene Chao Ma?

Entregables:

1. Para SQL un archivo .sql con el procedimiento y las queries que le hicieron llegar a la solución.
2. Para SQL comentar el análisis y resultados en el mismo archivo .sql
3. Para Python archivo con los resultados obtenidos de los pasos b) al e) en formato Excel.
4. Para Python un archivo .py con las sentencias usadas para llegar a las respuestas y en el mismo archivo comentar las respuestas.

Enviar en el tiempo estipulado por el presentador de la entrevista.