

Prueba técnica SIMEM 2024

El Sistema de Información del Mercado de Energía Mayorista (SIMEM) es el primer portal de datos abiertos del sector eléctrico colombiano, este portal de datos consiste en un mecanismo centralizado que reúne y unifica, tanto variables del mercado de energía para la toma de decisiones por parte de participantes del mercado como para otro tipo de extracciones por parte de terceros interesados. El SIMEM fue implementado y puesto en operación por XM S.A. E.S.P. desde septiembre de 2023. Actualmente el SIMEM cuenta con más de 234 conjuntos de datos que han sido estandarizados, transformados y almacenado de acuerdo con los lineamientos definidos por los órganos rectores de la política de datos abiertos para Colombia. Con base en lo anterior, es por esto que el proceso encargado de administrar la plataforma debe garantizar la calidad de los datos y la oportunidad en su publicación.

Considerando lo anterior, en este ejercicio Usted revisará algunos datos que deben ser publicados en el SIMEM pero previamente deben ser transformados, evaluado para garantizar los principales atributos de las guías de datos abiertos.

Recomendaciones para la prueba.

Tenga en cuenta los siguientes criterios de evaluación:

- Se dará mayor puntuación a soluciones que tengan un mayor elemento de programación.
- Se dará mayor puntuación a soluciones en lenguajes de programación como: R, Python, Scala, Matlab o similares.
- Separador decimal (.) y no hay separador de miles.
- Formato de fecha: AAAA-MM-DD
- Son archivos Excel en formato XLSX
- **Respuestas a la prueba sin que se relacione el código fuente o archivo en el que fue desarrollada (pbix, xlsx, etc.) recibirá una calificación equivalente a 0.0**
- Enviar sus respuestas únicamente a los correos jgaviria@xm.com.co y jcuartas@xm.com.co se les enviará un correo de confirmación cuando se reciba la prueba en el buzón indicado.
Importante: En caso de que decida usar algún servicio de almacenamiento en la nube para compartir las respuestas de la prueba tengan en cuenta que se la fecha de última de modificación deberá coincidir con el plazo de la prueba. Recuerde dar acceso público a los archivos compartidos para poder evaluarlo.
- Un valor vacío o faltante en los archivos de Generación o precio puede ser interpretado como cero.

Archivos suministrados:

1. Generacion_(kWh)_YYYY: Contiene la generación real de los recursos de generación en (kWh) siendo YYYY el año al que hace referencia cada archivo

2. Listado_Recursos_Generacion.xlsx: Contiene la información básica de las plantas de generación registradas ante el Centro Nacional de Despacho e indica la empresa que la representa
3. Precio_Bolsa_Nacional_(\$kwh)_YYYY: Contiene el precio de bolsa nacional horario expresado en (\$/kWh) siendo YYYY el año al que hace referencia cada archivo.

Pregunta 1 (15%). Unifique los registros del archivo Generacion_(kWh)_2022.xlsx con los del archivo Generacion_(kWh)_2023.xlsx en un único archivo. **Recuerde los criterios de evaluación al resolver este punto.**

Pregunta 2 (20%). Utilizando la información en el archivo Listado_Recursos_Generacion.xlsx, construya una tabla con la generación diaria por agente; exporte esta tabla a un archivo csv llamado *pregunta2_tabla.csv*. De la información contenida en esta tabla, elija tres agentes y presente en una gráfica de líneas la serie de generación diaria de cada uno (**tres series en un mismo gráfico**); exporte este gráfico en el formato que considere pertinente bajo el nombre *pregunta2_grafico.jpg* (o *.png*, *.gif*, etc.).

Pregunta 3 (20%). Construya una gráfica donde se presente la generación promedio mensual por tipo (tercera columna de los archivos de generación); luego construya una segunda gráfica donde se presente la generación mensual promedio sin ningún tipo de diferenciación. Exporte estos gráficos en el formato que considere pertinente bajo el nombre *pregunta3_grafico1.jpg* (o *.png*, *.gif*, etc.) y *pregunta3_grafico2.jpg* (o *.png*, *.gif*, etc.).

Pregunta 4 (20%). Construya una tabla donde se indique el porcentaje de participación de cada tipo de fuente de energía (cuarta columna de los archivos de generación) respecto al total de la generación mensual, la tabla deberá tener la siguiente estructura:

Mes	PorcentajeAgua	PorcentajeGas	PorcentajeBiogas	Etc..
2022-01-01				
2022-02-01				
2022-03-01				
...				
2023-12-01				

Exporte esta tabla a un archivo llamado *pregunta4_tabla.csv*.

Pregunta 5 (25%). Construya una función que envíe automáticamente correos electrónicos en la que sus parámetros de entrada sean los siguientes:

- Una lista de correos
- Un asunto
- Una lista de rutas que dirijan a archivos que podrán ser adjuntados
- Un texto que sea el cuerpo del correo

Con base en lo anterior, la función debe enviar un correo que tenga como destinatario al primer elemento de la lista de correos y en copia a las demás direcciones, como asunto la frase enviada en los parámetros, como cuerpo el texto enviado como parámetro, y como adjuntos los archivos de la lista de rutas.

En caso de que se traten de adjuntar más de cinco archivos la función deberá generar un mensaje de error indicando que se debe reducir el número de archivos adjuntar.

Importante: Asuma que la librería para enviar correos está dada, lo valioso de este punto es identificar su habilidad para resolver el problema mediante funciones. Es deseable que, al crear la función, la documente y utilice anotaciones para definir los tipos de variables que recibe.

Bonus (opcional)

Utilice los archivos denominados *Precio_Bolsa_Nacional_(\$kwh)_YYYY* y con base en estos construya un modelo que le permita pronosticar el precio promedio diario de Bolsa. Explique las razones por las cuales seleccionó el modelo, realice un diagnóstico de los residuales y señale si se cumplen los criterios para confiar en los pronósticos que arroja su modelo.

Recuerde considerar criterios como parsimonia y simplicidad para elegir el modelo que utilizará.