

Dev Python Junior Precia

Esta es la prueba de Precia para sus desarrollares Python Junior, reconocemos que quienes responden tiene acceso a diferentes fuentes de conocimiento. Por favor, céntrate en dar respuestas de forma que demuestres tu capacidad investigativa y personalidad como desarrollador.

* Obligatoria

Calidad de código

En la actualidad existe una variedad herramientas que pueden generar código funcional con algunas frases, también existe la posibilidad de apropiar nuevos conocimientos por medio de blogs u otros, por lo que el valor de un buen desarrollador está en crear código seguro, legible, mantenible y con manejo de errores, que se resume en el Zen de Python y los PEPs.

Sabemos que estás en formación, así que aplica todas las buenas practicas que conoces para responder el siguiente ejercicio. Por favor usar Python 3.9.

1. El fichero

<https://aprendeconalf.es/docencia/python/ejercicios/soluciones/pandas/titanic.csv> contiene información sobre los pasajeros del Titanic. Escribir un programa con los siguientes requisitos:

1. Generar un DataFrame con los datos del fichero.
2. Mostrar por pantalla las dimensiones del DataFrame, el número de datos que contiene, los nombres de sus columnas y filas, los tipos de datos de las columnas, las 10 primeras filas y las 10 últimas filas
3. Mostrar por pantalla los datos del pasajero con identificador 148.
4. Mostrar por pantalla las filas pares del DataFrame.
5. Mostrar por pantalla los nombres de las personas que iban en primera clase ordenadas alfabéticamente.
6. Mostrar por pantalla el porcentaje de personas que sobrevivieron y murieron.
7. Mostrar por pantalla el porcentaje de personas que sobrevivieron en cada clase.
8. Eliminar del DataFrame los pasajeros con edad desconocida.
9. Mostrar por pantalla la edad media de las mujeres que viajaban en cada clase.
10. Añadir una nueva columna booleana para ver si el pasajero era menor de edad o no.
11. Mostrar por pantalla el porcentaje de menores y mayores de edad que sobrevivieron en cada clase.

Requerimientos técnicos:

- Usar librerías pandas. Si no sabes pandas, no te preocupes, puedes usar google o cualquier herramienta; recuerda lo importante es mostrar la calidad de tu código.

Puntos extra:

- Usar la librería logging a cambio de la función print.

Enviar respuestas al correo: rparra@precia.co

Escriba su nombre a continuación:

Sebastian Camilo Tovar Pedraza

Conocimiento de base de datos relacionales: MySQL

Como Junior en Precia tendrás que conectarte constantemente a bases de datos relacionales, el motor mas usado en la organización es MySQL. Por favor responde las siguientes preguntas.

2. ¿Cuáles son las diferencias entre los tipos de datos CHAR y VARCHAR? *

El tipo de dato CHAR tiene tamaño fijo, mientras que el VARCHAR tiene tamaño variable, lo que resulta en consultas más eficientes para los tipos CHAR. Su caso de uso también varia: los VARCHAR se usan para nombres o descripciones, al ser datos de longitud variable.

3. ¿Cuál es mejor tipo para almacenar un precio? *

Según el tipo de moneda y la precisión, pueden ser DECIMAL o NUMERIC.

4. Escriba un query MySQL que permita insertar la siguiente fecha '20220220' (2022-02-20) en la columna 'valuation_date' de la siguiente tabla, por favor mantenga el formato original de la fecha.

```
CREATE TABLE us_bussines_days(  
    valuation_date DATE NOT NULL,  
    PRIMARY KEY(valuation_date)  
); *
```

```
INSERT INTO us_bussines_days (valuation_date)  
VALUES (STR_TO_DATE('20200220', '%Y%m%d'));
```

5. ¿Cuál puede ser la posible causa del siguiente mensaje de error: "ERROR 1062 (23000): Duplicate entry 'EURUSD' for key 'PRIMARY'"? *

La columna definida como PRIMARY KEY solo permite tener valores unicos, asi que es posible que ya existe una entrada con el mismo valor para esa columna.

Conocimiento de AWS

Una buena oportunidad laboral es la que te permite crecer, la que valoriza tu perfil profesional, y como tú, Precia también está en constante crecimiento. Actualmente estamos migrando nuestra infraestructura a AWS (<https://aws.amazon.com/es/what-is-aws/>), no te preocupes sino sabes que es, lo que queremos identificar es tu capacidad de adquirir nuevos conocimientos e investigar.

Las cosas que te preguntaremos a continuación te las enseñaremos, por lo que aprovecha las respuestas que te encuentres como una oportunidad para conocer lo que sería tu diario vivir en Precia.

Responde las siguientes preguntas, puedes usar cualquier medio disponible, pero por favor agrega la fuente que usaste al final de cada respuesta, si es conocimiento que posees decláralo al momento de escribir la fuente.

6. Defina en sus palabras que es una AWS lambda *

Es un servicio de AWS usado para la construcción de arquitecturas Event-driven. No hay que gestionar ningún servidor y el código se ejecuta con algún evento preestablecido.

Fuente: conocimiento que poseo.

7. ¿Como usar librerías Python no nativas en AWS lambda? *

1. Crear un directorio local bajo el nombre de "python" e instalar adentro la librería usando pip.
2. Crear un archivo tipo zip dentro del directorio python. El path debe ser como: Package_Name/python/ todos_los_archivos
3. Ingresar a AWS Lambda Function y entrar a la sección "Layers" en el menú de la izquierda. Seleccionar el botón "Create Layer". Dentro definir un nombre, subir el archivo zip y seleccionar "Create".
4. Ir a la función Lambda objetivo y seleccionar el botón "Add a layer", luego "Custom layers", seleccionar la capa que acabas de crear y seleccionar el botón "Add".
5. Importar la librería y desplegar la función Lambda. Probar la función.

Fuentes: <https://awstip.com/how-to-add-external-python-libraries-to-aws-lambda-baa8cb8dd7e1>
<https://docs.aws.amazon.com/lambda/latest/dg/creating-deleting-layers.html>

8. Defina en sus palabras que es una AWS S3 bucket *

Sirve para almacenar archivos que será llamado "objetos", como carpetas/directorios en un sistema operativo. Cada objeto cuenta con 3 componentes: un identificador único, su contenido e información de metadata (nombre del objeto, URL y tamaño). Su objetivo es facilitar el almacenamiento y manejo los archivos dentro del entorno Cloud.

Fuente: <https://www.techtarget.com/searchaws/definition/AWS-bucket>

9. ¿Que es el ciclo de vida de un AWS bucket S3? *

Es el conjunto de reglas que definen las acciones que el S3 aplicar a un grupo de objetos, pueden ser de transición (pasar a otra clase de almacenamiento) o de vencimiento (eliminarlos según un tiempo establecido).

Fuentes: <https://docs.aws.amazon.com/AmazonS3/latest/userguide/object-lifecycle-mgmt.html>

10. Defina en sus palabras que son los procesos ETL *

Es el proceso para centralizar la información en una data warehouse u a partir de múltiples fuentes. Sus siglas indican las etapas del proceso: extraer información, transformarla (limpiar, depurarla, darle formato) y cargarla en algún sistema objetivo.

Fuente: conocimiento que poseo.

11. ¿Cual es el servicio AWS por excelencia para desplegar procesos ETL? *

El más usado es AWS Glue. También se puede usar AWS Data Pipeline y AWS Kinesis.

Fuentes: https://hevodata.com/learn/aws-etl-tools/#Top_8_AWS_ETL_Tools
<https://www.fivetran.com/learn/best-aws-etl-tools>
<https://airbyte.com/top-etl-tools-for-sources/best-aws-etl-tools>

Este contenido no está creado ni respaldado por Microsoft. Los datos que envíe se enviarán al propietario del formulario.



Microsoft Forms