	Examen para Analistas de Datos	
		VERSIÓN: 1.1
		PÁGINA 1 DE 4

Fecha: 7 de junio de 2024

Instrucciones:

1. Resuelve los siguientes problemas utilizando las herramientas indicadas.
2. Asegúrate de incluir comentarios explicativos en tu código.
3. Sube tus soluciones a un repositorio de GitHub y comparte el enlace del repositorio en la entrega.

Problema 1: Resuelva en Python

Dado un rango de números enteros $[a, b]$ ($a \leq b$), tu tarea es escribir una función `suma_primos_en_rango` que tome como argumentos dos números enteros a y b y devuelva la suma de todos los números primos dentro del rango $[a, b]$ (incluidos a y b si son primos).

EJEMPLO:

`suma_primos_en_rango(10, 20).`

Test:

`test.assert_equals(suma_primos_en_rango(10,20) , 60)`

Problema 2: Realizar Datamart

Dada la siguiente información de préstamos, realizar un diagrama de clases del data mart que propondría para realizar una carga de información:

`customer_id`: Identificador único para cada cliente

`customer_age`: Edad del cliente

`customer_income`: Ingreso anual del cliente

`home_ownership`: Estado de propiedad de la vivienda (por ejemplo, ALQUILER, PROPIEDAD, HIPOTECA)

`employment_duration`: Duración del empleo en meses

`loan_intent`: Propósito del préstamo (por ejemplo, PERSONAL, EDUCACIÓN, MÉDICO, EMPRENDEDOR)

`loan_grade`: Calificación asignada al préstamo

`loan_amnt`: Monto del préstamo solicitado


`loan_int_rate`: Tasa de interés del préstamo

`term_years`: Plazo del préstamo en años

`historical_default`: Indica si el cliente tiene un historial de incumplimiento (S/N)

`cred_hist_length`: Duración del historial crediticio del cliente en años

`Current_loan_status`: Estado actual del préstamo (INCMPLIDO, NO INCMPLIDO)

 <small>CRÉDITO HIPOTECARIO NACIONAL</small>	Examen para Analistas de Datos	
		VERSIÓN: 1.1
		PÁGINA 2 DE 4

Problema 3: Análisis de datos del COVID-19

Se te proporciona un archivo CSV llamado `country_wise_latest.csv` que contiene información actualizada sobre casos confirmados, muertes y recuperaciones de COVID-19 en todo el mundo. El conjunto de datos se puede descargar desde [este enlace](#).


Tu tarea es realizar un análisis exploratorio de los datos utilizando la biblioteca `pandas` en Python. Realiza las siguientes tareas:

1. Leer el archivo CSV y cargar los datos en un `DataFrame` de `pandas`.
2. Limpiar y preprocesar los datos (por ejemplo, tratar los valores faltantes, convertir tipos de datos si es necesario, cambiar el nombre de las columnas para que sean más descriptivas).
3. Calcular estadísticas descriptivas para las columnas numéricas (por ejemplo, media, mediana, desviación estándar, etc).
4. Crear una nueva columna llamada `Active` que sea la cantidad de pacientes vivos.
5. Agrupar los datos por `Country/Region` y calcular el total de `Confirmed`, `Deaths`, `Recovered`, y `Active` para cada país.
6. Ordenar los países por el número total de casos confirmados y mostrar los 10 países con más casos confirmados.
7. Visualizar la evolución de casos confirmados, muertes y recuperaciones a lo largo del tiempo para un país específico (por ejemplo, Estados Unidos, España o Brasil) utilizando gráficos de línea. Incluye comentarios explicativos en tu código y asegúrate de utilizar las funciones y métodos de `pandas` de manera adecuada.

Problema 4: Consultas SQL para un banco con cuentas monetarias, cheques y préstamos

Imagina que estás trabajando con la base de datos de un banco que maneja cuentas monetarias, cheques y préstamos. La base de datos tiene las siguientes tablas:

1. clientes :
 - `id_cliente` (INTEGER): identificador único del cliente
 - `nombre` (VARCHAR): nombre del cliente
 - `direccion` (VARCHAR): dirección del cliente
 - `telefono` (VARCHAR): número de teléfono del cliente
2. cuentas :
 - `id_cuenta` (INTEGER): identificador único de la cuenta
 - `id_cliente` (INTEGER): identificador del cliente al que pertenece la cuenta (clave foránea que hace referencia a `clientes.id_cliente`)
 - `saldo` (DECIMAL): saldo actual de la cuenta
3. cheques :
 - `id_cheque` (INTEGER): identificador único del cheque
 - `id_cuenta` (INTEGER): identificador de la cuenta desde la que se emite el cheque (clave foránea que hace referencia a `cuentas.id_cuenta`)
 - `monto` (DECIMAL): monto del cheque
 - `fecha_emision` (DATE): fecha en la que se emitió el cheque

 <small>CRÉDITO HIPOTECARIO NACIONAL</small>	Examen para Analistas de Datos	
		VERSIÓN: 1.1
		PÁGINA 3 DE 4

beneficiario (VARCHAR): nombre del beneficiario del cheque

4. prestamos :

id_prestamo (INTEGER): identificador único del préstamo

id_cliente (INTEGER): identificador del cliente que recibió el préstamo (clave foránea que hace referencia a clientes.id_cliente)


monto (DECIMAL): monto del préstamo

fecha_otorgamiento (DATE): fecha en la que se otorgó el préstamo

plazo_meses (INTEGER): plazo del préstamo en meses

Tu tarea es realizar consultas SQL para obtener la siguiente información:

1. Obtener una lista de todos los clientes con sus respectivas cuentas, mostrando las columnas nombre , direccion , telefono , id_cuenta y saldo . Ordenar los resultados por el nombre del cliente.
2. Calcular el saldo total en el banco sumando los saldos de todas las cuentas.
3. Encontrar los 5 clientes con los saldos más altos en sus cuentas, mostrando las columnas nombre, id_cuenta y saldo. Ordenar los resultados por el saldo de mayor a menor.
4. Obtener una lista de todos los cheques emitidos en el último mes, mostrando las columnas id_cheque , nombre (del cliente que emitió el cheque), monto , fecha_emision y beneficiario . Ordenar los resultados por la fecha de emisión de más reciente a más antigua.
5. Encontrar el monto total de cheques emitidos por cada cliente en el último mes, mostrando las columnas nombre y monto_total_cheques . Ordenar los resultados por el monto total de cheques de mayor a menor.
6. Calcular el monto total de préstamos otorgados en el último año.
7. Encontrar los 10 clientes con los mayores montos de préstamos otorgados en el primer cuatrimestre del año.

 <small>CRÉDITO HIPOTECARIO NACIONAL</small>	Examen para Analistas de Datos	
		VERSIÓN: 1.1
		PÁGINA 4 DE 4

Acuerdo de Confidencialidad del Examen

Al comenzar este examen, usted acepta que no compartirá, distribuirá ni divulgará de ninguna manera el contenido del examen, incluyendo, pero no limitándose a preguntas, problemas, ejercicios, respuestas y soluciones, ya sea total o parcialmente, en forma escrita, verbal o electrónica.

Al aceptar este acuerdo de confidencialidad, usted comprende y acepta lo siguiente:

1. La propiedad intelectual y el contenido del examen están protegidos por derechos de autor y son propiedad exclusiva del examinador o de la entidad que administra el examen.
2. No se permite la reproducción, distribución, divulgación o uso no autorizado del contenido del examen, ya sea total o parcialmente.
3. Cualquier violación de este acuerdo de confidencialidad puede resultar en acciones legales y en consecuencias académicas y/o profesionales, incluyendo, pero no limitándose a la descalificación del examen, la revocación de certificaciones, la suspensión o expulsión de instituciones académicas, y la pérdida de oportunidades de empleo o promoción.