Ejercicios:

**1) Preguntas teóricas**

● ¿Qué es Python y por qué es popular en el análisis de datos? (0,5)

● ¿Para qué se usa la librería de Numpy?¿Y qué ventaja nos ofrece? (0,5)

● ¿Para qué nos sirve Pandas?¿Y qué tipos de objetos maneja? (0,5)

● ¿Qué es Matplotlib y cual es su uso?. Mencione un escenario real en donde lo

usaría. (0,5)

**2) Ejercicio de algoritmia (3)**

Estás trabajando en un programa que procesa listas, pero te enfrentas al desafío de

eliminar los elementos duplicados de una lista, manteniendo el orden original.

Especificaciones de la Función:

Nombre de la función:

● El nombre de la función debe ser: eliminar\_duplicados

Parámetros:

● lista (list): Una lista que puede contener elementos duplicados.

Retorno:

● list: Una nueva lista que contiene los elementos de la lista original sin

duplicados, conservando el orden original.

Ejemplo

● La función eliminar\_duplicados toma la lista [1, 2, 3, 4, 4, 5, 6, 6, 7] como

entrada y devuelve una nueva lista [1, 2, 3, 4, 5, 6, 7] sin elementos duplicados,

manteniendo el orden original.

**3) Ejercicio de Pandas (2,5)**

Se proporcionará una lista de calificaciones de los estudiantes y su tarea es realizar

análisis estadístico y visualización de datos.

calificaciones = [

{"nombre": "Juan", "matematicas": 85, "ciencias": 90,

"historia": 75},

{"nombre": "María", "matematicas": 70, "ciencias": 80,

"historia": 85},

{"nombre": "Pedro", "matematicas": 95, "ciencias": 75,

"historia": 90},

{"nombre": "Ana", "matematicas": 80, "ciencias": 85, "historia":

80},

{"nombre": "Luis", "matematicas": 75, "ciencias": 70,

"historia": 95},

{"nombre": "Sofía", "matematicas": 90, "ciencias": 85,

"historia": 75},

{"nombre": "Carlos", "matematicas": 85, "ciencias": 90,

"historia": 80},

{"nombre": "Elena", "matematicas": 70, "ciencias": 75,

"historia": 85},

{"nombre": "Javier", "matematicas": 80, "ciencias": 85,

"historia": 90},

{"nombre": "Laura", "matematicas": 75, "ciencias": 70,

"historia": 95},

{"nombre": "Diego", "matematicas": 90, "ciencias": 85,

"historia": 75},

{"nombre": "Paula", "matematicas": 85, "ciencias": 90,

"historia": 80},

{"nombre": "Carmen", "matematicas": 70, "ciencias": 75,

"historia": 85}

]

● Calcular el promedio de calificaciones para cada asignatura y mostrarlo.

● Encuentra a los estudiantes que tienen las calificaciones más altas en cada

asignatura y mostralos junto con sus respectivas calificaciones.

● Calcular el porcentaje de estudiantes que aprobaron cada asignatura (con una

calificación igual o superior a 60) y mostrar los resultados.

● Crear un DataFrame que incluya dos columnas una para las asignaturas y otra

para la frecuencia de cada calificación.

**4) Ejercicio de Gráficos (2,5)**

Basándote en el DataFrame creado, realiza un análisis visual de las calificaciones de

los estudiantes en las diferentes asignaturas. Utiliza un gráfico de barras donde el eje

x represente los nombres de los estudiantes y el eje y represente las calificaciones

obtenidas en cada asignatura.

**Desarrollo**

1)

a) Python es un lenguaje de programación de alto nivel, que fue creado en 1991, es popular en el análisis de datos porque es adecuado hacer representaciones visuales.

b) La librería numpy se usa para los cálculos numéricos y análisis de datos.

ventajas:

* Bastante ordenado
* Amplia gama de opciones como algebra lineal
* Las matrices de NumPy son mucho mejores
* Permite crear operaciones matemáticas de alta complejidad.
* c) Pandas no sirve para el análisis de datos y la manipulación de estructuras tabulares. Los tipos de datos que maneja son:
* Series: Estructura de una dimensión.
* DataFrame: Estructura de dos dimensiones (tablas).
* Panel: Estructura de tres dimensiones (cubos).

d) Matplotlib es una librería su uso es para la visualización datos.