Introducción al análisis de datos con Python

Nombre: _____

Fecha: _____

Puntaje: _____

1.

¿Qué devuelve el método .head() de un DataFrame de pandas?

1. Las primeras 5 filas del DataFrame
2. Las últimas 5 filas del DataFrame
3. Una muestra aleatoria de filas
4. Los nombres de las columnas

2.

El método utilizado para eliminar filas con valores faltantes en un DataFrame de pandas es:

Respuesta: _____

3.

¿Cuál es la forma correcta de seleccionar una columna llamada 'edad' de un DataFrame df?

1. df['edad']
2. df[edad]
3. df[['edad']]
4. df.get('edad')

4.

Para seleccionar filas donde la columna 'puntaje' sea mayor a 50, se utiliza:

1. df[df['puntaje'] > 50]
2. df.loc[:, df['puntaje'] > 50]
3. df.loc('puntaje > 50')
4. df.filter('puntaje > 50')

5.

El atributo que devuelve el número de filas y columnas de un DataFrame es:

Respuesta: _____

6.

¿Qué información proporciona el método .describe() aplicado a un DataFrame?

1. Estadísticas descriptivas de las columnas numéricas
2. Los tipos de datos de cada columna
3. El número de valores faltantes por columna
4. Los valores únicos de cada columna

7.

¿Cuál función de matplotlib crea una nueva figura?

1. plt.plot()
2. plt.figure()
3. plt.subplot()
4. plt.show()

8.

Para importar pandas con el alias común se escribe: import pandas as ___

Respuesta: _____

9.

¿Cuál de las siguientes NO es una forma válida de crear un DataFrame en pandas?

1. A partir de un diccionario de listas
2. A partir de una lista de diccionarios
3. A partir de un archivo CSV
4. A partir de una cadena de texto usando pd.StringData()

10.

Después de aplicar .groupby() a un DataFrame, ¿qué método devuelve directamente la media de cada grupo?

1. .mean()
2. .apply(np.mean)
3. .agg('count')
4. .transform('mean')

11.

La función de matplotlib que muestra el gráfico en pantalla es:

Respuesta: _____

12.

Para crear un gráfico de líneas con datos x e y, se utiliza:

1. plt.scatter(x, y)
2. plt.plot(x, y)
3. plt.bar(x, y)
4. plt.hist(x)

13.

En matplotlib, el parámetro 'alpha' controla:

1. El ancho de la línea
2. El color de relleno
3. La transparencia
4. El tamaño del marcador

14.

El método de pandas para renombrar columnas de un DataFrame es:

Respuesta: _____

15.

¿Cuál de los siguientes comandos fusiona correctamente dos DataFrames df1 y df2 por la columna 'id'?

1. pd.merge(df1, df2, on='id')
2. df1.merge(df2, on='id', how='cross')
3. df1.join(df2, on='id')
4. pd.concat([df1, df2], axis=1)

16.

Para añadir un título a un gráfico de matplotlib, se usa:

1. plt.xlabel('Título')
2. plt.ylabel('Título')
3. plt.title('Título')
4. plt.legend('Título')

17.

La función de pandas para leer un archivo CSV se escribe:

Respuesta: _____

18.

¿Qué devuelve el método .value_counts() aplicado a una Serie de pandas?

1. La frecuencia de cada valor único
2. La suma de todos los valores
3. La suma acumulativa de los valores
4. El promedio de los valores

19.

¿Cuál función de matplotlib se usa para crear una figura y un conjunto de subgráficos?

1. plt.subplots()
2. plt.subplot()
3. plt.figure()
4. plt.axes()

20.

En el método .plot() de pandas, el parámetro 'kind' controla:

1. El color del gráfico
2. El tipo de gráfico (línea, barra, etc.)
3. El tamaño de la figura
4. La leyenda del gráfico