# Curso de Python básico para análisis de datos - Ejercicios clase 3

1. Indique la frase incorrecta

[ ] A diferencia de los arrays de numpy que solo tienen índices numéricos, los dataframe de pandas pueden tener índices de tipo string, fecha y otros.

[X] Puedo definir una tabla de más de dos dimensiones.

[ ] Una misma columna de un dataframe puede tener varios tipos de datos (int, string, datetime, float, etc.)

1. Verdadero o falso Este código es correcto:

aeropuertos.iloc[:, "provincia"]

Rta.: Falso

1. Complete el código. Debo eliminar las columnas

df = df.drop(["columna1", "columna2"], axis=...)

Ayuda: busque en internet o pregunte a algún chatbot si lo considera necesario

Rta.: df = df.drop(["columna1", "columna2"], axis=1)

1. Indique los fragmentos correctos (puede ser más de uno). Debo filtrar el dataframe para obtener los datos desde el año 2010 al año 2020

[X] df = df[df[“año”]>= 2010 & df[“año”]< 2021]

[ ] df = df.iloc(2010 to 2021)

[ ] df = df.filer(año beteween 2010 to 2020)

[X] df = df.query(“año >= 2010 & año <= 2020”)

1. ¿Para que sirve el método grpupby?

[ ] Para filtrar la columna de meses

[ ] Es un método estadístico de clasificación

[X] Con este método divide el dataframe en las variables seleccionadas y permite usar luego otro método para operar sobre cada división y generar un nuevo dataframe agregado

[ ] Es una operación de alta complejidad, capaz de modificar la memoria interna del disco superior de la máquina. Su poder altera el tejido espacio-temporal por lo que puede provocar mareos. Al liberar su máximo potencial, le permitirá al usuario usar el 100% de su capacidad cognitiva.

1. Verdadero o falso.

El método groupby no suele ser usado por si solo, necesita otro método auxiliar que eliga el criterio de agrupación. Por ejemplo, si ese criterio fuese la suma de los valores desagregados usaremos el método .sum(), si fuera el promedio de estos usaremos el método .mean()

Rta.: Verdadero

1. Son los dos siguientes fragmentos de código equivalentes? Fragmento A:

* indices\_expo\_agrupada\_ordenada = (
* indices\_expo\_filtrada
* .drop("Mes", axis=1).groupby("Año").sum()
* .sort\_values("suma", ascending=False))
* Fragmento B:
* indices\_expo\_agrupada\_ordenada = indices\_expo\_filtrada.drop("Mes", axis=1)  
  indices\_expo\_agrupada\_ordenada = indices\_expo\_agrupada\_ordenada.groupby("Año").sum()  
  indices\_expo\_agrupada\_ordenada.sort\_values("suma", ascending=False)
* [X] Sí, son equivalentes [ ] No son equivalentes

8. En base al archivo ovni.csv, lealo con pandas y determine lo siguiente

a. ¿De que paises son estos avistamientos? (puede usar el método .unique() )

b. Cuál es el país con mayor avistamientos según este dataset

c. Cúal es el promedio de segundos avistados del 2000 al 2010.

Subir el archivo acá, puede ser .py o .ipynb