

Programación en entornos BigData.

Primera entrega.

- En todos los gráficos se valorará que se pueda entender fácilmente la información mostrada. Incluye la utilización de etiquetas, leyendas, títulos, escalas apropiadas...
- También se valorará la calidad del código, incluyendo comentarios, y del resto del notebook.
- La entrega se realiza a través del aula virtual, completando este notebook en Jupyter. Hay que guardarlo con el nombre Entrega1_[nombre]_[apellidos].ipynb
- No hay que incluir los ficheros de datos en la entrega.
- El ejercicio 1 hay que hacerlo con los datos del fichero "insurance.csv".
- El ejercicio 2 hay que hacerlo con los datos de los ficheros "ActuacionesBomberos_2023.csv" y "ActuacionesBomberos_2024.csv". La estructura de datos es diferente entre los dos ficheros. Más información:
<https://datos.madrid.es/sites/v/index.jsp?vgnextoid=fa677996afc6f510VgnVCM1000001d4a900aRCRD&vgnnextchannel=374512b9>
- Todos los ejercicios hay que resolverlos utilizando solo Matplotlib.
- Por favor, no edites los enunciados.
- Calificación:

** Ejercicio 1: 1 punto por cada apartado. ** Ejercicio 2:

- ■ ■ apartados a, c y d: 1 punto cada uno
- ■ ■ apartados b y e: 0.5 puntos cada uno.

Carga de datos

```
In [ ]: ## no edites ni elimines esta celda !!!
# path = "../datos/"
# seguros = pd.read_csv(path + "/insurance.csv", sep = ",")
# bomberos23 = pd.read_csv(path + "/ActuacionesBomberos_2023.csv", sep = ";")
# bomberos24 = pd.read_csv(path + "/ActuacionesBomberos_2024.csv", sep = ";")
# print(seguros.head())
# print(bomberos23.head())
# print(bomberos24.head())
```

Hay que cargar los datos de seguros (insurance.csv) y de bomberos (ActuacionesBomberos_2023.csv y ActuacionesBomberos_2024.csv) en dataframes llamados seguros, bomberos23 y bomberos 24. Edita la siguiente celda para que se ajuste a tus rutas.

```
In [ ]: # seguros = pd.read_csv(path + "/insurance.csv", sep = ",")
# bomberos23 = pd.read_csv(path + "/ActuacionesBomberos_2023.csv", sep = ";")
# bomberos24 = pd.read_csv(path + "/ActuacionesBomberos_2024.csv", sep = ";")
```

Ejercicio 1

En este ejercicio utilizamos los datos de seguros.

a) Muestra un gráfico con un diagrama de barras con el número de asegurados que hay en cada región. La región estará en el eje x. (1 punto)

Introduce tu solución a partir de aquí ...

b) Muestra un gráfico con cuatro subplots, uno por cada región. En cada uno se mostrará un grafico de tarta con el porcentaje fumadores y no fumadores en esa región. (1 punto)

Introduce tu solución a partir de aquí ...

c) Muestra un gráfico con cuatro subplots, uno por cada región. En cada uno se mostrará un histograma con el precio del seguro en esa región. (1 punto)

Introduce tu solución a partir de aquí ...

d) Muestra un gráfico de violín con el precio del seguro en cada región. (1 punto)

Introduce tu solución a partir de aquí ...

e) Muestra un gráfico de cajas con el precio del seguro cada región. (1 punto)

Introduce tu solución a partir de aquí ...

f) Escribe una versión del apartado c) en la que los cuatro histogramas se muestren en un sólo gráfico 3D. (1 punto)

Ejemplos:

<https://matplotlib.org/stable/gallery/mplot3d/bars3d.html>

<https://matplotlib.org/stable/gallery/mplot3d/2dcollections3d.html>

<https://matplotlib.org/stable/tutorials/toolkits/mplot3d.html#polygon-plots>

Ejercicio 2

En este ejercicio utilizamos los datos de actuaciones de bomberos.

a) Muestra un gráfico de barras con las salidas totales por distrito en todo el año 2023. En el eje x estará el distrito. (1 punto)

Introduce tu solución a partir de aquí ...

b) Muestra un gráfico de tarta con las salidas, por tipo de salida, excepto las totales. Hay que utilizar los datos de todo Madrid para el año 2023.

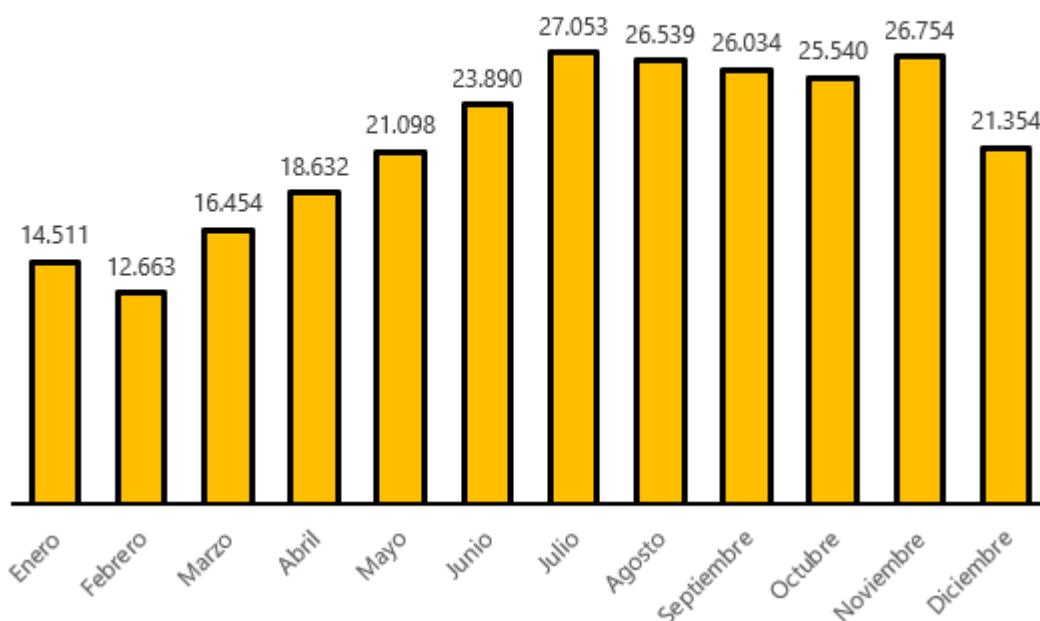
(0.5 puntos)

Introduce tu solución a partir de aquí ...

c) Muestra un gráfico gráfico de barras con las salidas totales mes a mes para 2023 para todo Madrid. En el eje x estarán los meses ordenados. Similar a:

Gráfico de barras

PlantillasPyme

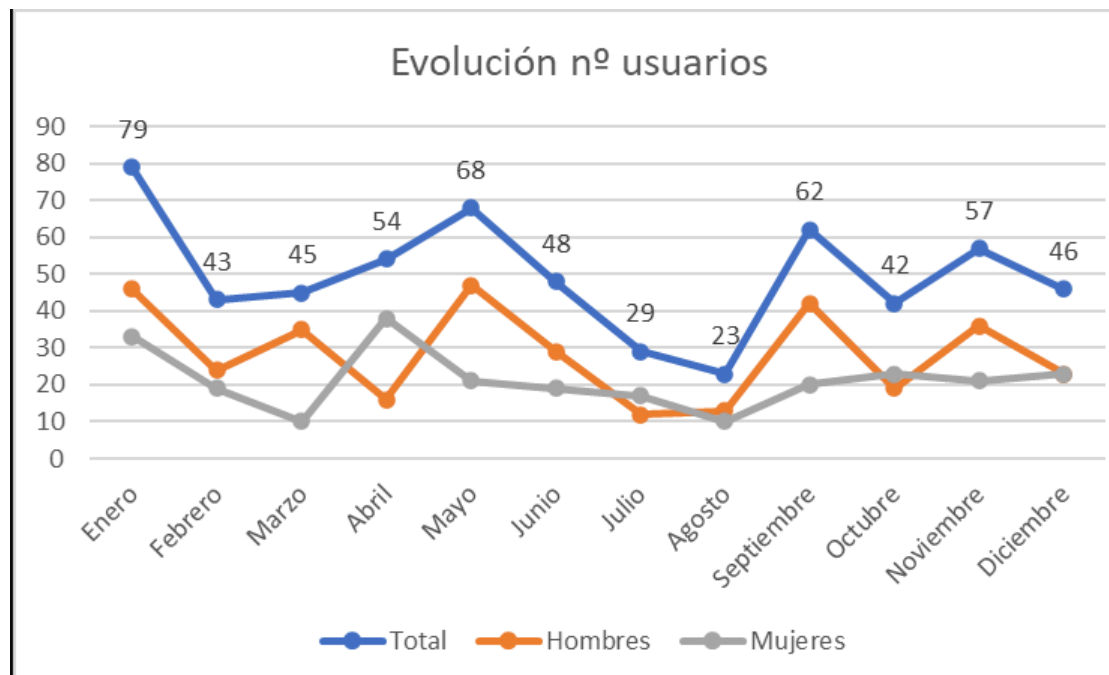


(1 punto)

Introduce tu solución a partir de aquí ...

d) Muestra un gráfico de línea con las salidas totales mes a mes con los datos de tres distritos para 2023.

Similar a:



Introduce tu solución a partir de aquí ...

e) Muestra un gráfico como el anterior, pero añadiendo los meses disponibles para 2024. En este caso el eje x irá desde enero de 2023 hasta septiembre de 2024. Se valorará que la solución acepte ficheros de 2024 con más meses, según se vayan publicando.

(0.5 puntos)

Introduce tu solución a partir de aquí ...