

Actividad 1

1. La primera pregunta, se cargó el archivo CSV, una vez que se tiene el dataframe se filtro la información de acuerdo con la pregunta, lo cual es el evento tornado. Con este dataframe filtrado se realiza un conteo que se ordena por default en forma descendente, lo cual nos da que los meses con más tornados son Mayo, Junio, Abril. Véase la figura 1.

```
In [48]: runfile('C:/Users/CMM/Documents/Python_semestre3/Análisis de Datos/tarea_1.py', wdir='C:/Users/CMM/Documents/Python_semestre3/Análisis de Datos')
May      184
June     159
April    150
October  134
September 121
November 108
July      97
August    90
December  68
March     66
February  55
January   16
Name: Month, dtype: int64

The three months with more tornadoes are May, June, April
```

Figura 1. Meses con tornados.

2. La segunda pregunta, es responder cual es el tercer evento más frecuente. Para ello no se filtro el dataset como en el punto anterior. Y contabilizar con respecto a la columna Event_Type, dando como resultado que el tercer evento más frecuente es Flood. Véase la figura 2.

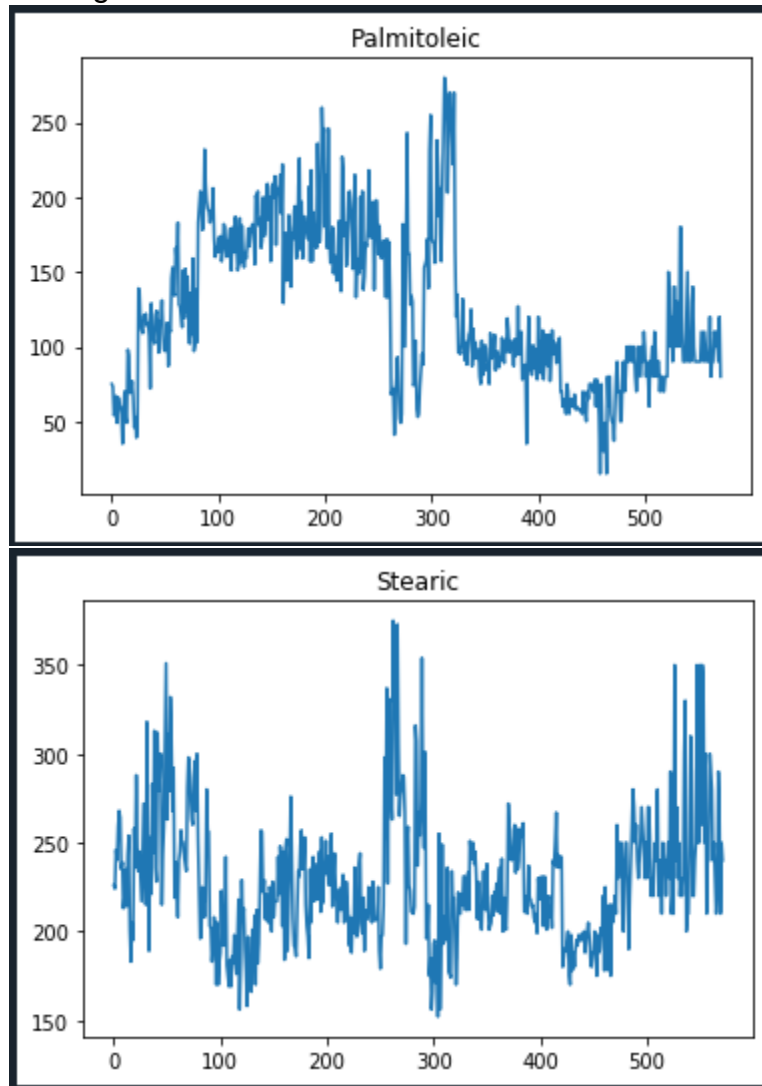
Thunderstorm Wind	14585
Hail	7861
Flood	4715
Winter Weather	4478
Flash Flood	4358
Winter Storm	3375
High Wind	2944
Drought	2410
Heavy Snow	2220
Marine Thunderstorm Wind	2090
Heavy Rain	1899
Heat	1282
Tornado	1248
Strong Wind	1021
Dense Fog	752

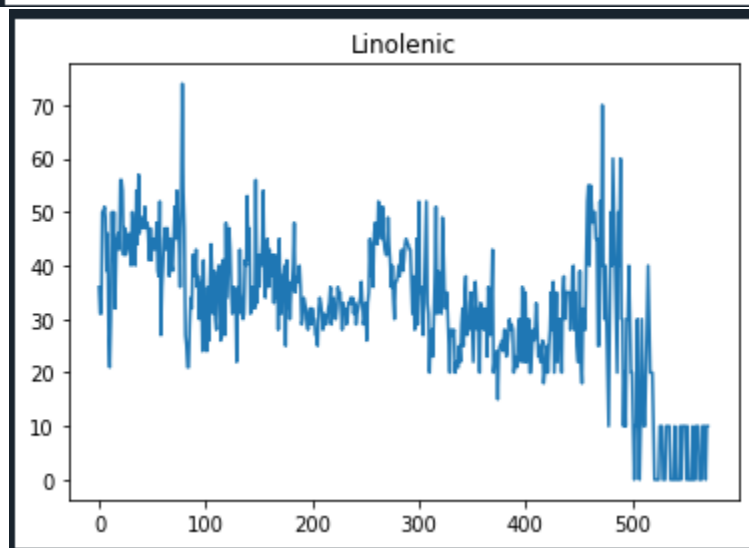
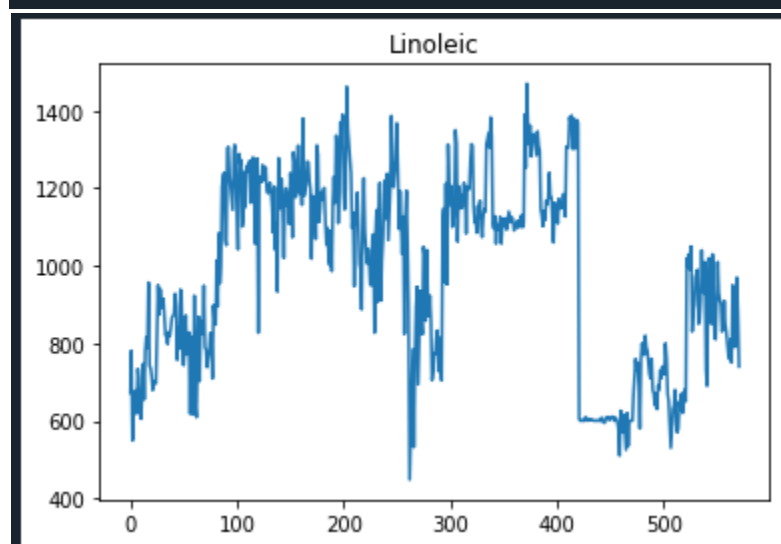
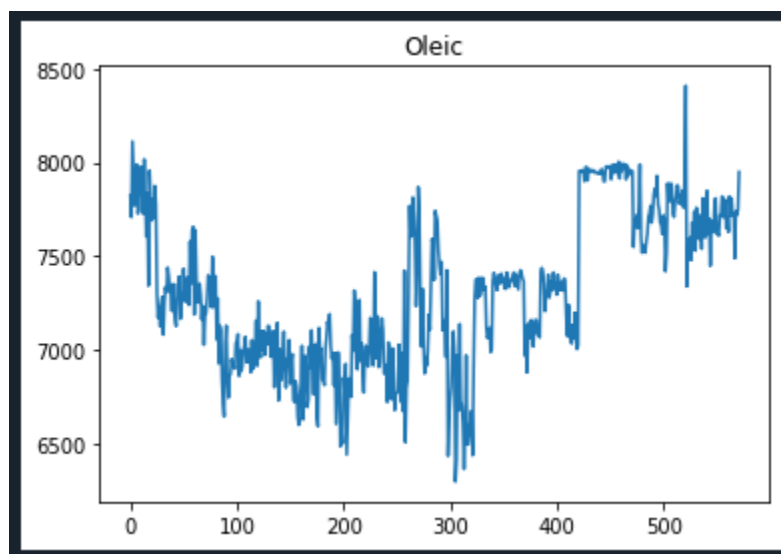
Figura 2. Tipos de eventos.

3. La tercera pregunta es con el archivo hw.csv, el cual es otra base de datos. Se pide identificar el tipo de datos. En la primera columna son datos de tipo String, para las columnas región, área y palmitic son enteros, para las

columnas palmitoleic, stearic, oleic, linoleic, linolenic, arachidic, eicosenic los datos son de tipo string puesto que contienen comillas, a pesar de esto al utilizar los datos en Python, este los toma como tipo string a todos.

4. Las gráficas individuales de las columnas 5 a 11, para esto se eliminaron los caracteres especiales con los que contaba cada columna, las cuales eran comillas, posteriormente se convirtieron los valores en valores numéricos para dar valor al eje y, y en el eje x se tomaron los valores de renglón. Véase figura 3.





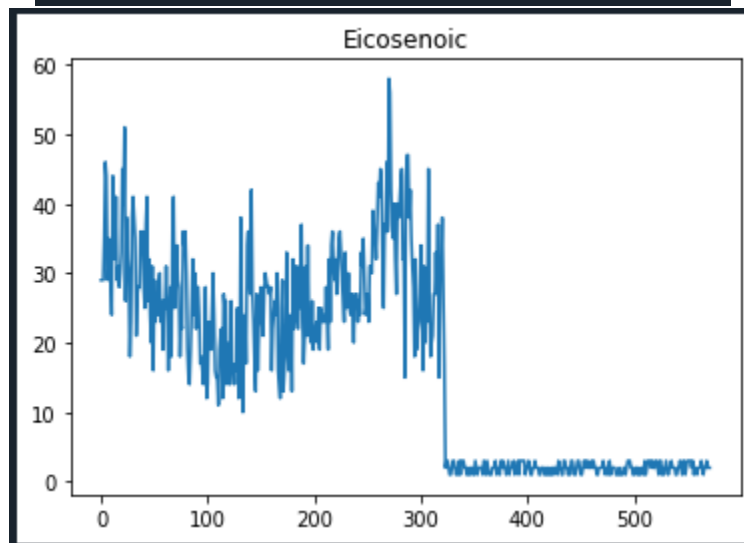
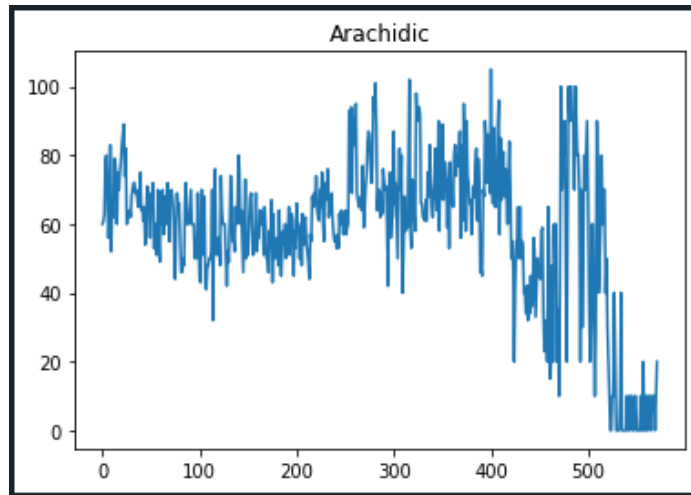


Figura 3. Gráficas por columna (columna 5 a 11).

5. Por último, graficar los datos de la columna 11 vs los datos de la columna 5. Véase la figura 4.

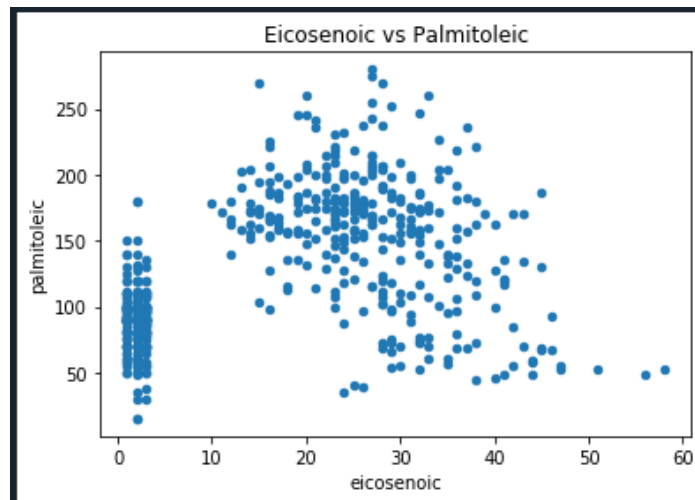


Figura 4. Eicosenoic vs Palmitoleic.