











Descripción DataUranga

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Longitud de párrafo | Num. de palabras por párrafo | Num. de puntos por párrafo | Num. de comas | Num. de punto y coma | Num. de dos puntos |
| Count | 81.000.000 | 81.000.000 | 81.000.000 | 81.000.000 | 81.000.000 | 81.000.000 |
| mean | 6.938.272 | 90.604.938 | 4.098.765 | 5.481.481 | 234.568 | 469.136 |
| std | 4.433.807 | 63.260.311 | 3.337.682 | 5.148.085 | 597.319 | 881.567 |
| min | 1.000.000 | 4.000.000 | 0 | 0 | 0 | 0 |
| 25% | 3.000.000 | 37.000.000 | 2.000.000 | 1.000.000 | 0 | 0 |
| 50% | 7.000.000 | 88.000.000 | 4.000.000 | 5.000.000 | 0 | 0 |
| 75% | 9.000.000 | 127.000.000 | 6.000.000 | 7.000.000 | 0 | 1.000.000 |
| max | 21.000.000 | 258.000.000 | 20.000.000 | 26.000.000 | 3.000.000 | 5.000.000 |

Descripción DataJorgito

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Longitud de párrafo | Num. de palabras por párrafo | Num. de puntos por párrafo | Num. de comas | Num. de punto y coma | Num. de dos puntos |
| Count | 189.000.000 | 189.000.000 | 189.000.000 | 189.000.000 | 189.000.000 | 189.000.000 |
| mean | 4.386.243 | 54.169.312 | 2.153.439 | 3.328.042 | 571.429 | 100.529 |
| std | 2.169.425 | 29.574.696 | 1.513.328 | 2.608.967 | 990.072 | 301.503 |
| min | 1.000.000 | 1.000.000 | 0 | 0 | 0 | 0 |
| 25% | 3.000.000 | 32.000.000 | 1.000.000 | 1.000.000 | 0 | 0 |
| 50% | 4.000.000 | 55.000.000 | 2.000.000 | 3.000.000 | 0 | 0 |
| 75% | 6.000.000 | 72.000.000 | 3.000.000 | 5.000.000 | 1.000.000 | 0 |
| max | 12.000.000 | 166.000.000 | 11.000.000 | 15.000.000 | 6.000.000 | 1.000.000 |

Descripción Data Mondriguin

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Longitud de párrafo | Num. de palabras por párrafo | Num. de puntos por párrafo | Num. de comas | Num. de punto y coma | Num. de dos puntos |
| Count | 569.000.000 | 569.000.000 | 568.000.000 | 567.000.000 | 569.000.000 | 569.000.000 |
| mean | 4.319.859 | 46.926.186 | 1.998.239 | 2.957.672 | 381.371 | 159.930 |
| std | 4.029.772 | 38.486.162 | 1.797.999 | 3.270.757 | 1.095.961 | 412.068 |
| min | 1.000.000 | 0 | 0 | 0 | 0 | 0 |
| 25% | 1.000.000 | 10.000.000 | 1.000.000 | 0 | 0 | 0 |
| 50% | 4.000.000 | 38.000.000 | 1.000.000 | 2.000.000 | 0 | 0 |
| 75% | 6.000.000 | 75.000.000 | 3.000.000 | 5.000.000 | 0 | 0 |
| max | 70.000.000 | 169.000.000 | 11.000.000 | 25.000.000 | 14.000.000 | 3.000.000 |

Code Mondri

#Importamos las librerías que ocuparemos durante el proceso.

%matplotlib inline

import pandas as pd

import matplotlib.pyplot as plt

#Cargamos los datos con la función read excel de pandas

data = pd.read\_excel('UranguitaE.xlsx') #Contiene la estilometría del corpus de Emilio Uranga

data2 = pd.read\_excel('JorgitoE.xlsx') #Contiene la estilometría del corpus de Jorge Joseph Piedra

data3 = pd.read\_excel('MondriE.xlsx') #Contiene la estilometría del Mondrigo

#EMILIO URANGA

#data.plot()

#data.LongdParrafo.plot(color = 'm')

#plt.title('Longitud de Párrafo')

#data.NumPalabrasxParrafo.plot(color = 'm')

#plt.title('Numero de Palabras por Parrafo')

#data.NumdPuntosxParrafo.plot(color = 'm')

#plt.title('Numero de Puntos por parrafo')

#data.NumdComas.plot(color = 'm')

#plt.title('Numero de Comas')

#data.NumdPuntoComa.plot(color = 'm')

#plt.title('Numero de Punto y Coma')

#data.NumdDosPuntos.plot(color = 'm')

#plt.title('Numero de Dos Puntos')

#data.describe()

#JORGE JOSEPH PIEDRA

#data2.plot()

#data2.describe()

#data2.LongdParrafo.plot(color ='k')

#plt.title('Longitud de Párrafo')

#data2.NumPalabrasxParrafo.plot(color ='k')

#plt.title('Numero de Palabras por Parrafo')

#data2.NumdPuntosxParrafo.plot(color ='k')

#plt.title('Numero de Puntos por parrafo')

#data2.NumdComas.plot(color ='k')

#plt.title('Numero de Comas')

#data2.NumdPuntoComa.plot(color ='k')

#plt.title('Numero de Punto y Coma')

#data2.NumdDosPuntos.plot(color ='k')

#plt.title('Numero de Dos Puntos')

#MONDRIGO

#data3.plot()

#data3.describe()

#data3.LongdParrafo.plot(color='y')

#plt.title('Longitud de Párrafo')

#data3.NumPalabrasxParrafo.plot(color='y')

#plt.title('Numero de Palabras por Parrafo')

#data3.NumdPuntosxParrafo.plot(color='y')

#plt.title('Numero de Puntos por parrafo')

#data3.NumdComas.plot(color='y')

#plt.title('Numero de Comas')

#data3.NumdPuntoComa.plot(color='y')

#plt.title('Numero de Punto y Coma')

#data3.NumdDosPuntos.plot(color='y')

#plt.title('Numero de Dos Puntos')

#data2[data2.LongdParrafo == 1] #Prueba para el párrafo de longitud 1 en J

#data[data.LongdParrafo == 1] #Prueba para el párrafo de Longitud 1 en EU

#pd.crosstab(data2.LongdParrafo, data2.NumPalabrasxParrafo)

#fig = plt.figure(figsize=(30, 10))

#data2.NumPalabrasxParrafo[data2.LongdParrafo ==1].value\_counts(normalize = True).plot(kind='barh', alpha=0.5, color='red')

#plt.title('Numero de palabras que JJP usa en parrafos con longitud 1')

#plt.show()

pd.crosstab(data.LongdParrafo, data.NumPalabrasxParrafo)