



Herramientas computacionales: el arte de la analítica

TC1002S.222

Fecha de entrega: 11 de mayo del 2022

Actividad : “Obtención de estadísticas descriptivas”

Nombre : Andrea Samantha Aguilar Ramírez

Matrícula: A01656200

Profesor: Sergio Ruíz Loza

Análisis de los resultados :

Para comenzar me di cuenta de que existen ciertos datos/variables que no son del todo funcionales dentro del dataset, como por ejemplo la descripción de usuario o en algunos caso la ubicación ya que no se indica ninguna región del mundo en específico y se hace uso de emoticones que representan planetas, estrellas o banderas.

De igual manera me parece que la variable 'is_retweet' podría ser de uso, sin embargo, para este caso todas fueron falsas entonces al ser la misma respuesta para todas la variable puede ser eliminada.

En la variable 'user_name' el usuario con mayor frecuencia dentro de los resultados fue Global.Pandemic.NET con una frecuencia de 312.

Por otra parte otro de los datos que pude analizar es desde qué dispositivos se generan los tweets obteniendo así que : Twitter Web App: 22974, Twitter for Android: 16758, Twitter for iPhone: 15824, Twitter for iPad: 1884. Lo cuál nos permite ver que de los dispositivos más utilizados que son los smartphones y las computadoras portátiles y tablets.

Así mismo mediante la variable de región se puede observar que los países con mayor participación en los tweets fueron India y los Estados Unidos Americanos. La siguiente variables que se analizó fue la cantidad de cuentas verificadas, donde la cantidad de cuentas que si se encontraban verificadas era mucho menor que las que no obteniendo que las cuentas verificadas era el 12.5% (9354) y el 87.5% (65082) para las cuentas no verificadas.

Para finalizar, creo que podemos tomar una idea general después de obtener estos datos y es que las redes sociales no siempre son una fuente de información veredigna, ya que tenemos usuarios sin ubicación lo que afectaría la noticia e incluso dentro del análisis también se puede apreciar que algunos tweets no contienen información relevante para la pandemia de covid 19, si no más datos o anécdotas individuales sobre el aislamiento o algunos otros temas como higiene.