

Objetivo

Esta práctica tiene como objetivo el tratamiento de datos, en este caso WhastApp, con el fin de poder sacar información relevante del agregado de los mismos.

URL: <https://github.com/juanquer/DatosNoEstructurados>

Funciones

Además de las funciones ya incluidas en el código original, se ha extendido el mismo con algunas funciones propias:

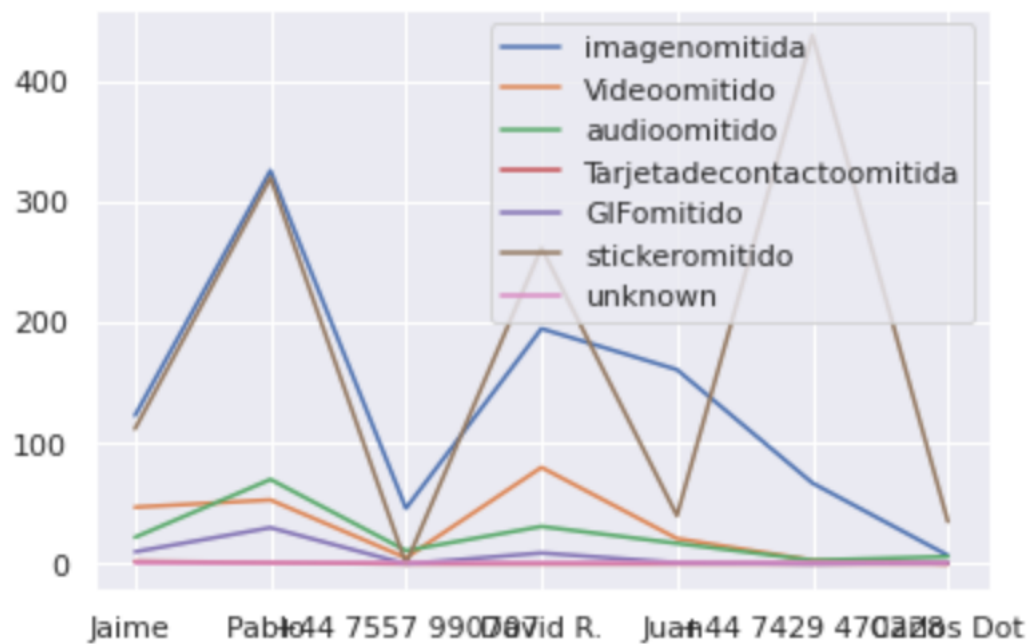
- **num_chars:** para recoger los metadatos del usuario, entre los que encontramos el número de caracteres escritos por los usuarios
- **get_latewords:** el objetivo de este método, es el de conseguir un WordCloud de las palabras escritas a altas horas de la noche.
- **increment_users_attachment_count:** esta función, permite llevar un recuento de los attachments enviados por los miembros del chat

Gráficos propios

La siguiente figura representa una tabla resumen de toda la media (imágenes, videos, audios, GIFs y stickers) enviada por cada uno de los participantes del grupo.

	imagenomitida	Videomitido	audiomitido	Tarjetadecontactomitida	GIFomitido	stickeromitido	unknown
Jaime	123.0	47.0	22.0	2.0	10.0	112.0	0.0
Pablo	326.0	53.0	70.0	1.0	30.0	320.0	0.0
+44 7557 990707	46.0	5.0	11.0	0.0	0.0	0.0	0.0
David R.	195.0	80.0	31.0	0.0	9.0	262.0	0.0
Juan	161.0	21.0	17.0	0.0	1.0	40.0	0.0
+44 7429 470228	67.0	3.0	3.0	0.0	0.0	438.0	0.0
Carlos Dot	7.0	0.0	6.0	0.0	1.0	35.0	0.0

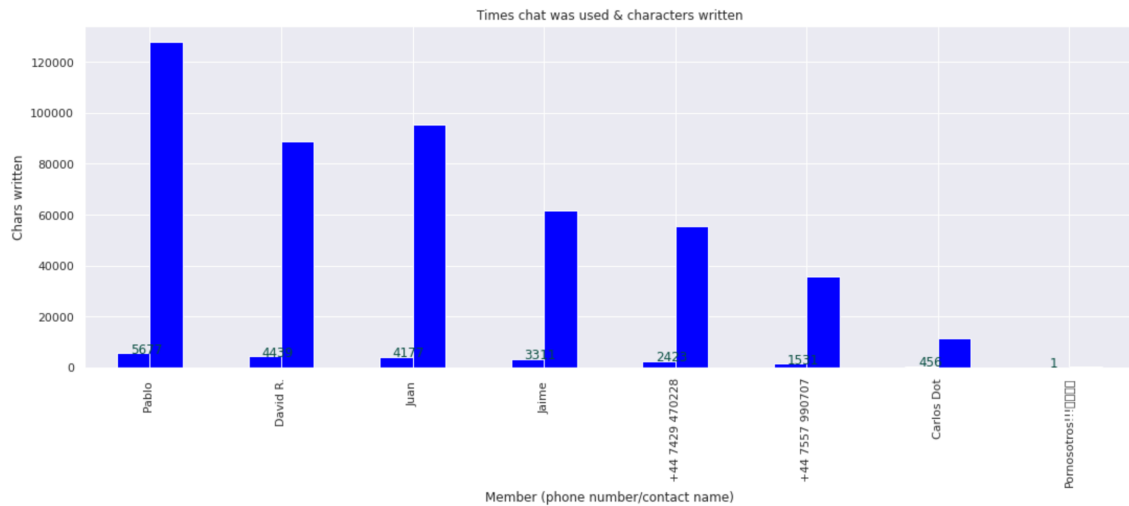
Ahora se busca plasmar toda la información anterior sobre un gráfico de líneas, donde se puede ver de forma más intuitiva qué personas son más propensas a enviar ciertos archivos de media.



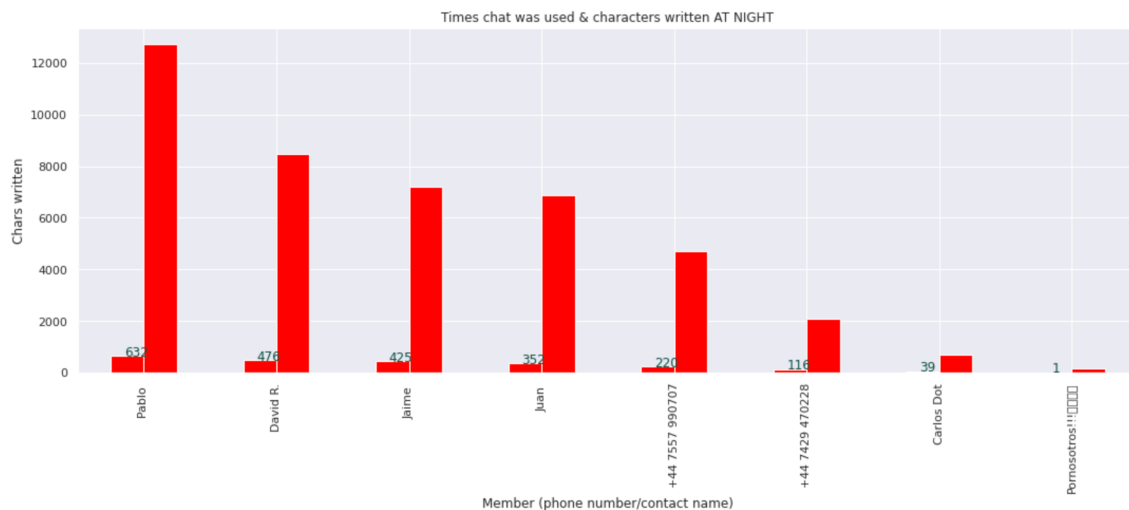
A través de este gráfico se pueden ver las interacciones de los miembros del grupo. El número de caracteres enviados por cada miembro y las veces que ha usado el chat.

member	chat_count	chars_count
Pablo	5677	127795
David R.	4439	88644
Juan	4177	95330
Jaime	3311	61502
+44 7429 470228	2423	55279
+44 7557 990707	1531	35779
Carlos Dot	456	11406
Pornosotros!!!	1	143

La idea de esta tabla anterior se refleja a través del siguiente gráfico de barras, representando las dos columnas: tanto los usos como los caracteres



El siguiente gráfico es un spin off del anterior, ya que se representa la misma información, pero filtrada a un rango de horas específico (por la noche) y así analizar cual es el miembro que más usa el móvil de fiesta.



Misma idea pero desagregando (drill out) por el rango de horas que componen la noche.

member	+44 7429 470228	29	25	1		3	2		56	
	+44 7557 990707	65	22	14	23	10	8	13	65	
	Carlos Dot	20					12		7	
	David R.	88	89	9	24	7	14	8	9	228
	Jaime	136	63	23	21	1	10	12	1	158
	Juan	74	18	3	11	7	4	1	14	220
	Pablo	165	70	21	23	16	14	11	13	299
	Pornosotros!!!!									1
	0	1	2	3	4	5	6	7	23	
	hora									

Finalmente se acaba la sesión nocturna analizando cual es el porcentaje de uso diurno respecto al uso por la noche.



day	Friday	69	32	8	9	3	2	1	10	9	46	72	185	181	125	166	205	242	229	389	290	183	213	249	139
	Monday	61	17	6	5	4	5	6	2	6	54	72	114	140	176	181	111	162	153	157	254	150	176	108	83
	Saturday	136	30	7	62	24	37	5	3	3	38	52	98	206	286	162	142	193	221	225	256	179	246	180	136
	Sunday	96	79	20	4	1	15		7	1	4	16	132	235	412	290	140	140	169	219	215	95	280	144	107
	Thursday	37	22	14	8	2			4	6	13	83	191	183	107	194	261	194	187	278	350	287	338	246	191
	Tuesday	52	39	14	10	3	2	32	3	9	21	69	164	167	271	304	225	232	302	163	239	258	218	181	223
	Wednesday	101	64	3	2		3	1	2	10	10	69	83	163	148	108	140	154	106	157	150	153	215	184	102
		00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23
		hour																							

