⌄ Universidad Autónoma de Nuevo León

Facultad de Ciencias Físico Matemáticas

Maestría en Ciencia de Datos

Procesamiento y clasificación de datos

Tarea 2

Jorge Loredo Gonzalez

Matricula: 1850739

> Codigo

▶ ↳ **11 celdas ocultas**

Analisis

⌄ *Datos*

Se Extrae información de comntarios de post de Reddit relacionado a autos.

```
1    comments[0:10]
```

```
['My 911 because who needs practicality or a spine.',
 'A hellcat redeye charger. \nBecause I'm stupid and hate moneu',
 '[deleted]',
 'Tacoma cuz I hate power, fuel economy, and love hunting for gears.',
 'Heavily modified 1977 MGB.  Just because I refuse to be normal.',
 'M3 Competition, because I dailied an S2000 all throughout college and that was very much a "never meet your heroes" moment that was
ruining my dream car for me. Worst daily driver ever. Now with the M3 as a daily, the S2000 again feels like the baby F1 car that I
bought it to be.',
 'Life is too short to not drive something fun on the daily, but I only have room for one car at my house, so all-arounder M car it
is.',
 '[deleted]',
 'I take a tram most days.',
 'Ford Focus ST, because hot hatchback practically and speed']
```

```
1 most_common_words
```

```
[('car', 272),
 ('daily', 86),
 ('like', 82),
 ('one', 82),
 ('drive', 78),
 ('get', 63),
 ('im', 61),
 ('tesla', 58),
 ('fun', 57),
 ('dont', 57)]
```

Se utiliza la libreria nltk.sentiment para categorizar los comentarios en funcion de su sentimiento trasmitido

```
1 df[['Comentario', 'sentiment_category']]
```

	Comentario	sentiment_category
0	My 911 because who needs practicality or a spine.	Neutral
1	A hellcat redeye charger. \nBecause I'm stupid...	Negative
2	[deleted]	Neutral
3	Tacoma cuz I hate power, fuel economy, and lov...	Positive
4	Heavily modified 1977 MGB. Just because I ref...	Negative
...	...	...
482	Autopilot will keep you in the lane better on ...	Positive
483	Ok you have the ability to copy/paste. Very go...	Positive
484	My last one was a 1975 Moto Guzzi 850 police i...	Positive
485	>Why Toyota decided to stick with a 17 year ol...	Negative
486	The NHTSA instigates every case that's brought...	Positive

487 rows × 2 columns