


> Load stuff

 ↳ 3 cells hidden

✓ Load data

```
# load dataframe
df = pd.read_excel("Lab2_data.xlsx")
df
```



| | id_text | text |
|------|---------|---------------------------------------------------|
| 0 | 0 | while parking at a doctors office, i pulled in... |
| 1 | 0 | Once i had parked i reied to say sorry as i co... |
| 2 | 0 | It was after all a genuine mistake on my park ... |
| 3 | 0 | but this person was in such a foul mood i coul... |
| 4 | 1 | My friend told me I am an idiot. |
| ... | ... | ... |
| 2836 | 503 | none of us fell and i started walking again th... |
| 2837 | 503 | but the lady started screaming, addressing me ... |
| 2838 | 504 | I am talking to a saleswoman in a store and sh... |
| 2839 | 504 | I feel hurt and helpless. |
| 2840 | 504 | I want to leave. |

2841 rows x 2 columns

✓ Get agency through BERTAgent

```
# run BERTAgent ... it's very slow without GPU
df['agency'] = my_bertagent().extract_ig(df['text'])
# show the result
df
```

➔ /usr/local/lib/python3.10/dist-packages/huggingface_hub/utils/_auth.py:94: UserWarning: The secret `HF_TOKEN` does not exist in your Colab secrets.
To authenticate with the Hugging Face Hub, create a token in your settings tab. You will be able to reuse this secret in all of your notebooks.
Please note that authentication is recommended but still optional to access private repositories.
warnings.warn(

Evaluating: 100%

143/143 [18:03<00:00, 5.27s/it]

| | id_text | text | agency |
|------|---------|---------------------------------------------------|-----------|
| 0 | 0 | while parking at a doctors office, i pulled in... | -0.191778 |
| 1 | 0 | Once i had parked i reied to say sorry as i co... | -0.055492 |
| 2 | 0 | It was after all a genuine mistake on my park ... | -0.190038 |
| 3 | 0 | but this person was in such a foul mood i coul... | -0.335492 |
| 4 | 1 | My friend told me I am an idiot. | -0.190367 |
| ... | ... | ... | ... |
| 2836 | 503 | none of us fell and i started walking again th... | 0.265351 |
| 2837 | 503 | but the lady started screaming, addressing me ... | -0.095687 |
| 2838 | 504 | I am talking to a saleswoman in a store and sh... | -0.093656 |
| 2839 | 504 | I feel hurt and helpless. | -0.568953 |
| 2840 | 504 | I want to leave. | -0.229226 |

2841 rows × 3 columns

✓ Get additional markers via (simulated) LIWC

```
# chosen markers
mark_list = ['anx', 'posemo', 'negemo', 'swear', 'anger']
# check they exist
lw0 = my_liwc("LIWC2007_English100131.dic")
lw0.is_category(mark_list)
# initialize columns
for i in mark_list:
    df[i] = 0.0
# run LIWC
for i in tqdm(range(len(df))):
    v = lw0.get_categories(mark_list, df.loc[i, 'text'])
    for j in range(len(mark_list)):
        df.loc[i, mark_list[j]] = v[j]
# show result
df
```



100%

2841/2841 [00:56<00:00, 60.97it/s]

| | id_text | text | agency | anx | posemo | negemo | swear | anger |
|------|---------|---------------------------------------------------------|-----------|-----|----------|----------|-------|----------|
| 0 | 0 | while parking at a doctors office, i pulled in... | -0.191778 | 0.0 | 0.041667 | 0.000000 | 0.0 | 0.000000 |
| 1 | 0 | Once i had parked i reied to say sorry as i co... | -0.055492 | 0.0 | 0.024390 | 0.048780 | 0.0 | 0.024390 |
| 2 | 0 | It was after all a genuine mistake on my park ... | -0.190038 | 0.0 | 0.000000 | 0.133333 | 0.0 | 0.000000 |
| 3 | 0 | but this person was in such a foul mood i coul... | -0.335492 | 0.0 | 0.062500 | 0.000000 | 0.0 | 0.000000 |
| 4 | 1 | My friend told me I am an idiot. | -0.190367 | 0.0 | 0.111111 | 0.111111 | 0.0 | 0.111111 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 2836 | 503 | none of us fell and i started walking | 0.265351 | 0.0 | 0.000000 | 0.000000 | 0.0 | 0.000000 |

lw0.categories()



shehe
they
ipron
article
verb
auxverb
past
present
future
adverb
preps
conj
negate
quant
number
swear
social
family
friend
humans
affect
posemo
negemo
anx

mental
certain
inhib
incl
excl
percept
see
hear
feel
bio
body
health
sexual
ingest
relativ
motion
space
time
work
achieve
leisure
home
money
relig
death
assent
nonfl
filler