

```
In [1]: import pandas as pd
import numpy as np
import json
```

```
In [56]: df = pd.read_csv('tripadvisor_hotel_reviews.csv')
```

```
In [57]: df.shape
```

```
Out[57]: (20491, 2)
```

```
In [58]: df.head()
```

```
Out[58]:
```

	Review	Rating
0	nice hotel expensive parking got good deal sta...	4
1	ok nothing special charge diamond member hilto...	2
2	nice rooms not 4* experience hotel monaco seat...	3
3	unique, great stay, wonderful time hotel monac...	5
4	great stay great stay, went seahawk game aweso...	5

```
In [59]: df.tail(10)
```

```
Out[59]:
```

	Review	Rating
20481	ok price look hotel ok little run average clea...	2
20482	great choice wife chose best western quite bit...	5
20483	good bed clean convenient just night happy sta...	4
20484	deceptive staff deceptive desk staff claiming ...	2
20485	not impressed unfriendly staff checked asked h...	2
20486	best kept secret 3rd time staying charm, not 5...	5
20487	great location price view hotel great quick pl...	4
20488	ok just looks nice modern outside, desk staff ...	2
20489	hotel theft ruined vacation hotel opened sept ...	1
20490	people talking, ca n't believe excellent ratin...	2

```
In [60]: df.columns
```

```
Out[60]: Index(['Review', 'Rating'], dtype='object')
```

```
In [61]: df.rename(columns = {'Review':'Comentarios',
                             'Rating':'Calificacion'})
```

```
Out[61]:
```

	Comentarios	Calificacion
0	nice hotel expensive parking got good deal sta...	4
1	ok nothing special charge diamond member hilt...	2
2	nice rooms not 4* experience hotel monaco seat...	3
3	unique, great stay, wonderful time hotel monac...	5
4	great stay great stay, went seahawk game aweso...	5
...
20486	best kept secret 3rd time staying charm, not 5...	5
20487	great location price view hotel great quick pl...	4
20488	ok just looks nice modern outside, desk staff ...	2
20489	hotel theft ruined vacation hotel opened sept ...	1
20490	people talking, ca n't believe excellent ratin...	2

20491 rows × 2 columns

```
In [62]: hotel_5 = df[df['Rating']==5]
hotel_5.to_csv('hotel_5.csv', index=False)
```

```
In [63]: hotel_4 = df[df['Rating']==4]
hotel_4.to_csv('hotel_4.csv', index=False)
```

```
In [64]: hotel_3 = df[df['Rating']==3]
hotel_3.to_csv('hotel_3.csv', index=False)
```

```
In [65]: hotel_2 = df[df['Rating']==2]
hotel_2.to_csv('hotel_2.csv', index=False)
```

```
In [79]: hotel_1 = df[df['Rating']==1]
hotel_1.to_csv('hotel_1.csv', index=False)
```

```
In [80]: df1 = pd.read_csv('hotel_1.csv')
```

```
In [81]: df1.shape
```

```
Out[81]: (1421, 2)
```

```
In [82]: df1.head()
```

```
Out[82]:
```

	Review	Rating
0	horrible customer service hotel stay february ...	1
1	noise airconditioner-a standard, arranged stay...	1
2	bad choice, booked hotel hot wire called immed...	1
3	hated inn terrible, room-service horrible staf...	1
4	ace grunge lives does mold mildew tiny bed met...	1

```
In [ ]: //Preguntas de interes:
```

```
In [70]: df_drop = df.dropna()
```

```
In [ ]:
```