```
import numpy as np
import pandas as pd
# I dataframe, nella realtà, sono molto grandi. Questo significa che
non si filtra in base
# alla posizione, ma si filtra in base ad una condizione. I filtri ci
permettono di selezionare
# le righe in base a delle condizioni applicate alle colonne
# Vediamo questi filtri nel dettaglio, considerando il nostro solito
Df tips
tips = pd.read_csv("tips.csv")
tips.head()
  total bill tip
                       sex smoker day time size
price_per_person
       16.99 1.01 Female
                               No
                                  Sun
                                       Dinner
                                                  2
8.49
                                                  3
       10.34 1.66
                      Male
                               No
                                  Sun
                                       Dinner
3.45
2
       21.01 3.50
                      Male
                               No
                                   Sun
                                       Dinner
                                                  3
7.00
       23.68 3.31
                      Male
                               No
                                   Sun
                                       Dinner
                                                  2
11.84
       24.59 3.61 Female
                               No Sun
                                                  4
                                       Dinner
6.15
          Payer Name
                             CC Number Payment ID
  Christy Cunningham 3560325168603410
                                         Sun2959
1
      Douglas Tucker 4478071379779230
                                         Sun4608
2
      Travis Walters
                      6011812112971322
                                         Sun4458
3
    Nathaniel Harris 4676137647685994
                                         Sun5260
        Tonya Carter 4832732618637221
4
                                         Sun2251
# VOGLIAMO ESTRARRE TUTTE LE RIGHE I CUI TOTAL BILL SONO > 40
mask = tips["total bill"] > 40
tips[mask]
    total bill
                  tip
                          sex smoker
                                      day
                                             time size
price_per_person
                 6.73
         48.27
                         Male
                                  No
                                      Sat
                                           Dinner
```

12.07	7						
95	40.17	4.73	Male	Yes	Fri	Dinner	4
10.04	4						
102	44.30	2.50	Female	Yes	Sat	Dinner	3
14.77							
142	41.19	5.00	Male	No	Thur	Lunch	5
8.24	40 17	F 00	NA 7		•	ъ.	6
156	48.17	5.00	Male	No	Sun	Dinner	6
8.03 170	50.81	10.00	Male	Yes	Sat	Dinner	3
16.94		10.00	Mate	163	Jac	DTIIIIEI	3
182	45.35	3.50	Male	Yes	Sun	Dinner	3
15.12		3.30	11000	103	Jun	DIMICI	J
184	40.55	3.00	Male	Yes	Sun	Dinner	2
20.27	7						
197	43.11	5.00	Female	Yes	Thur	Lunch	4
10.78					_		
212	48.33	9.00	Male	No	Sat	Dinner	4
12.08	3						
	Payer	Namo	CC	Number	Daymo	n+ TD	
59	Brian (559645382		-	t8139	
95	Aaron Ber	-	18002661			i9628	
102	Heather (•	37977111			t6240	
142	Eric And	drews 4	135653176	1046453	Thu	r3621	
156	Ryan Gonz	zales 3	352315148	2063321	Su	n7518	
170	Gregory (47385096			t1954	
182	Jose Pai		111220755			n2337	
184	Stepher		354779822			n5140	
197	Brooke		554490220			r9313	
212	Alex Willia	anison	0/021	.8815212	29	t4590	

VOGLIAMO ESTRARRE TUTTE LE RIGHE IL CUI SESSO SIA PARI A M

mask = tips["sex"] == "Male"
tips[mask]

tota	l_bill	tip	sex	smoker	day	time	size
<pre>price_per</pre>	_person	\					
1	10.34	1.66	Male	No	Sun	Dinner	3
3.45							
2	21.01	3.50	Male	No	Sun	Dinner	3
7.00							
3	23.68	3.31	Male	No	Sun	Dinner	2
11.84							
5	25.29	4.71	Male	No	Sun	Dinner	4
6.32							

```
6
          8.77 2.00 Male
                               No Sun
                                        Dinner
                                                   2
4.38
. .
236
          12.60 1.00 Male
                              Yes
                                   Sat
                                        Dinner
                                                   2
6.30
237
         32.83 1.17 Male
                              Yes
                                   Sat
                                                   2
                                        Dinner
16.42
239
         29.03 5.92 Male
                               No
                                   Sat
                                        Dinner
                                                   3
9.68
241
                                                   2
         22.67 2.00 Male
                              Yes
                                   Sat
                                        Dinner
11.34
242
         17.82 1.75 Male
                               No
                                               2
                                   Sat
                                        Dinner
8.91
            Payer Name
                               CC Number Payment ID
1
        Douglas Tucker 4478071379779230
                                            Sun4608
2
        Travis Walters 6011812112971322
                                            Sun4458
3
       Nathaniel Harris 4676137647685994
                                            Sun5260
5
            Erik Smith
                                            Sun9679
                         213140353657882
6
    Kristopher Johnson 2223727524230344
                                            Sun5985
236
         Matthew Myers
                        3543676378973965
                                            Sat5032
237
          Thomas Brown 4284722681265508
                                            Sat2929
239
         Michael Avila 5296068606052842
                                            Sat2657
241
            Keith Wong
                                            Sat3880
                        6011891618747196
242
          Dennis Dixon
                           4375220550950
                                              Sat17
[157 rows x 11 columns]
# VOGLIAMO ESTRARRE LE RIGHE IN CORRISPONDENZA DELLE QUALI total bill
> 30 E sex=="Male"
mask = (tips["total bill"]>30) & (tips["sex"]=="Male")
tips[mask]
                  tip
                        sex smoker
                                     day time
     total bill
                                                  size
price per person
23
         39.42
                 7.58
                       Male
                                No
                                     Sat Dinner
9.86
39
         31.27
                                     Sat Dinner
                 5.00
                       Male
                                No
10.42
44
         30.40
                 5.60
                       Male
                                No
                                     Sun Dinner
7.60
47
         32.40
                 6.00
                       Male
                                     Sun Dinner
                                No
8.10
                 3.00 Male
56
         38.01
                               Yes
                                     Sat Dinner
                                                     4
```

9.50		40.07	6 70			6 1	р.	
59 12.0	7	48.27	6.73	Male	No	Sat	Dinner	4
83	,	32.68	5.00	Male	Yes	Thur	Lunch	2
16.3	4							
95	1	40.17	4.73	Male	Yes	Fri	Dinner	4
10.04 112	4	38.07	4.00	Male	No	Sun	Dinner	3
12.69	9	30.07	4.00	Tid CC	110	Juli	DIMICI	J
141		34.30	6.70	Male	No	Thur	Lunch	6
5.72								
142		41.19	5.00	Male	No	Thur	Lunch	5
8.24 156		48.17	5.00	Male	No	Sun	Dinner	6
8.03		40.17	3.00	Mate	NO	Suii	DIIIIEI	U
167		31.71	4.50	Male	No	Sun	Dinner	4
7.93								
170		50.81	10.00	Male	Yes	Sat	Dinner	3
16.9	4	21 05	2 10		\ <u>/</u>	6	ъ.	2
173 15.92	2	31.85	3.18	Male	Yes	Sun	Dinner	2
175	_	32.90	3.11	Male	Yes	Sun	Dinner	2
16.4	5	32.30	3.11	Tid CC	103	Juli	DIMICI	
179		34.63	3.55	Male	Yes	Sun	Dinner	2
17.3	2							
180		34.65	3.68	Male	Yes	Sun	Dinner	4
8.66		4F 2F	2 50	Mala	Vaa	C	Dinner	2
182 15.1	2	45.35	3.50	Male	Yes	Sun	Dinner	3
184	_	40.55	3.00	Male	Yes	Sun	Dinner	2
20.2	7	10133	3.00	114 00	103	Sun	DIIIICI	_
187		30.46	2.00	Male	Yes	Sun	Dinner	5
6.09				_				
207		38.73	3.00	Male	Yes	Sat	Dinner	4
9.68 210		30.06	2.00	Male	Yes	Sat	Dinner	3
10.0	2	30.00	2.00	nate	165	Sat	DIIIIEI	ی
212	_	48.33	9.00	Male	No	Sat	Dinner	4
12.0	8							
237	_	32.83	1.17	Male	Yes	Sat	Dinner	2
16.4	2							
			Payer N	ame	(^C Niimł	oer Paymen	+ TD
23			e Peter		35425840			t239
39		Mr. Bra			60115258	3510698		6373
44			odd Coo			3467612		2274
47	-		mes Bar		35520025			9677
56 59	Jame	s Chris [.]			3497936 35064539			8903
59 Brian Ortiz 6596453823950595 Sat8139								

```
83
             Daniel Murphy
                            5356177501009133
                                               Thur8801
95
             Aaron Bentley
                             180026611638690
                                                Fri9628
112
                Jeff Lopez
                            3572865915176463
                                                 Sun591
141
            Steven Carlson
                            3526515703718508
                                               Thur1025
142
              Eric Andrews
                            4356531761046453
                                               Thur3621
156
             Ryan Gonzales
                            3523151482063321
                                                Sun7518
167
            Michael Lawson
                            3566285921227119
                                                Sun3719
170
             Gregory Clark
                            5473850968388236
                                                Sat1954
               Scott Perez
173
                            3577115550328507
                                                Sun9335
175
           Nathan Reynolds
                             370307040837149
                                                Sun5109
179
              Brian Bailev
                             346656312114848
                                                Sun9851
                                                Sun7544
180
          James Hebert DDS
                                676168737648
182
              Jose Parsons
                            4112207559459910
                                                Sun2337
184
               Stephen Cox
                            3547798222044029
                                                Sun5140
187
             David Barrett
                            4792882899700988
                                                Sun9987
207
             Ricky Ramirez
                             347817964484033
                                                Sat4505
210
             Shawn Mendoza
                              30184049218122
                                                Sat8361
212
           Alex Williamson
                                676218815212
                                                Sat4590
237
              Thomas Brown 4284722681265508
                                                Sat2929
# Possiamo anche vedere un esempio in cui si utilizza l'operatore or
|. Questo significa che,
# date due condizioni, otterremo true se è vera solo la prima, solo la
seconda, o entrambe.
# Quindi, supponiamo di voler selezionare tutte le righe in
corrispondenza delle quali il sesso sia
# F OPPURE smoker sia "Yes"
mask = (tips["sex"] == "Female") | (tips["smoker"] == "Yes")
tips[mask]
                          sex smoker
     total bill
                  tip
                                       dav
                                              time size
price_per_person \
0
          16.99 1.01 Female
                                  No
                                       Sun
                                           Dinner
                                                       2
8.49
4
          24.59 3.61 Female
                                  No
                                       Sun
                                           Dinner
6.15
          35.26 5.00 Female
11
                                  No
                                       Sun
                                           Dinner
8.82
14
          14.83 3.02 Female
                                           Dinner
                                  No
                                       Sun
7.42
16
          10.33 1.67 Female
                                  No
                                       Sun Dinner
                                                       3
3.44
. .
237
          32.83 1.17
                         Male
                                 Yes
                                       Sat Dinner
                                                       2
```

```
16.42
          35.83 4.67 Female
                                  No
                                       Sat Dinner
238
                                                      3
11.94
240
          27.18 2.00 Female Yes
                                       Sat
                                          Dinner
                                                       2
13.59
241
          22.67 2.00
                         Male
                                Yes
                                       Sat
                                          Dinner
                                                       2
11.34
243
          18.78 3.00 Female
                                 No Thur Dinner
                                                       2
9.39
             Payer Name
                                CC Number Payment ID
     Christy Cunningham 3560325168603410
0
                                             Sun2959
4
           Tonya Carter 4832732618637221
                                             Sun2251
                                             Sun6686
11
           Diane Macias 4577817359320969
14
          Vanessa Jones
                           30016702287574
                                             Sun3848
16
       Elizabeth Foster 4240025044626033
                                             Sun9715
237
           Thomas Brown 4284722681265508
                                             Sat2929
238
        Kimberly Crane
                                             Sat9777
                             676184013727
        Monica Sanders 3506806155565404
240
                                             Sat1766
             Keith Wong 6011891618747196
241
                                             Sat3880
243
        Michelle Hardin 3511451626698139
                                             Thur672
[147 rows x 11 columns]
# POSSIAMO ANCHE LAVORARE SU UNA SOLA COLONNA: VEDIAMO OUI UN ESEMPIO.
# CONSIDERIAMO LA VARIABILE time. PER PRIMA COSA VEDIAMO QUALI SONO LE
OCCORRENZE DI QUESTA VARIABILE
# CATEGORICA
tips['time'].value_counts()
time
Dinner
          176
Lunch
           68
Name: count, dtype: int64
# SONO SOLO DINNER O LOUNCH, L'ESEMPIO NON VERREBBE BENE
# PROVIAMO CON LA VARIABILE day:
tips['day'].value counts()
```

```
day
Sat
       87
Sun
       76
Thur
       62
Fri
       19
Name: count, dtype: int64
# VOGLIAMO ESTRARRE LE RIHJE PER LE QUALI LA COLONNA day ASSUME I
VALORI Sun oppure Fri
mask = (tips["day"]=="Sun") | (tips["day"]=="Fri")
tips[mask]
    total bill
                tip sex smoker day time size
price per person \
         16.99
                1.01 Female
                                No
                                    Sun Dinner
8.49
         10.34 1.66
                       Male
                                    Sun Dinner
1
                                No
                                                   3
3.45
2
         21.01 3.50
                       Male
                                No
                                    Sun Dinner
7.00
```

Male

Male

Male

Female

10.09 2.00 Female Yes Fri

Douglas Tucker 4478071379779230

Tonya Carter 4832732618637221

Mary Rivera 5343428579353069

Jason Lawrence 3505302934650403

Nathaniel Harris 4676137647685994

Sun

Fri

Fri

Fri

Fri

CC Number Payment ID

No

No

Yes

No

Yes

Yes

3560325168603410

6011812112971322

Dinner

Lunch

Lunch

Lunch

Lunch

Lunch

Sun2959

Sun4608

Sun4458

Sun5260

Sun2251

Fri6624

Fri6014

Sun Dinner

2

3

2

2

2

23.68 3.31

8.58 1.92

15.98 3.00

13.42 1.58

Christy Cunningham

24.59 3.61 Female

16.27 2.50 Female

Payer Name

Travis Walters

3

11.84 4

6.15 ... 222

8.58 223

5.33

224

6.71 225

8.14 226

5.04

0

1

2

3

4

222

223

```
224
     Ronald Vaughn DVM 341503466406403
                                          Fri5959
        Whitney Arnold 3579111947217428
225
                                          Fri6665
226
            Ruth Weiss 5268689490381635
                                          Fri6359
[95 rows x 11 columns]
# VEDIAMO SE È VERO CHE RIMANGONO SOLO LE OCCORRENZE Sun E Fri:
solo sun fri = tips[mask]
solo_sun_fri["day"].value_counts()
day
Sun
      76
Fri
      19
Name: count, dtype: int64
# CVD
# UN ALTRO MODO PER UTILIZZARE OR | QUANDO SI LAVORA SU CONDIZIONI
APPARTENENTI ALLA STESSA VARIABILE
# E' L'OPERATORE isin. RIVEDIAMO L'ULTIMO ESEMPIO
tips
    total bill tip sex smoker day time size
price_per_person \
         16.99 1.01 Female
                               No
                                    Sun Dinner
                                                   2
8.49
         10.34 1.66
                       Male
                               No
                                    Sun Dinner
1
3.45
         21.01 3.50
                       Male
                                    Sun Dinner
                               No
                                                   3
7.00
3
         23.68 3.31
                       Male
                               No
                                    Sun Dinner
                                                   2
11.84
         24.59 3.61 Female
                               No
                                    Sun Dinner
4
6.15
. . .
         29.03 5.92
                       Male
                               No
239
                                    Sat Dinner
                                                3
9.68
240
         27.18 2.00 Female Yes
                                    Sat Dinner
                                                   2
13.59
         22.67 2.00
                       Male
                              Yes
                                                   2
241
                                    Sat Dinner
11.34
242
         17.82 1.75
                       Male
                               No
                                    Sat Dinner
8.91
```

243 9.39	18.7	78 3.00	Female	No	Thur	r Dinner	2
0 1 2 3	Christy (Dougl Travi	Payer Nam Cunningha Las Tucke is Walter iel Harri nya Carte	m 35603 r 44780 s 60118 s 46761	32516860 37137977 31211297	3410 9230 1322 5994	Sun445	9 8 8 0
239 240 241 242 243	Monio k Der	 nael Avil ca Sander Keith Won nnis Dixo lle Hardi	a 52966 s 35068 g 60118 n 43	96860605 30615556 39161874 37522055 45162669	5404 7196 0950		7 6 0 7
	rows x 11			13102003	0133	1110107	_
	= tips["(mask]	-				time	size
-	total_bil e_per_pers	son \		smoker	day		
0 8.49	16.9	99 1.01	Female	No	Sun	Dinner	2
1	10.3	34 1.66	Male	No	Sun	Dinner	3
3.45	21.0	01 3.50	Male	No	Sun	Dinner	3
7.00 3	23.6	8 3.31	Male	No	Sun	Dinner	2
11.84 4	24.5	59 3.61	Female	No	Sun	Dinner	4
6.15							
							_
222	8.5	58 1.92	Male	Yes	Fri	Lunch	1
8.58 223	15.9	98 3.00	Female	No	Fri	Lunch	3
5.33	13.4	12 1.58	Male	Yes	Fri	Lunch	2
6.71 225	16.2	27 2.50	Female	Yes	Fri	Lunch	2
8.14	10.2	., 2130	7 Cilia CC	103		Lancii	
226 5.04	10.0	9 2.00	Female	Yes	Fri	Lunch	2
0	F Christy (Payer Nam Cunningha		CC Nui 32516860		Payment I Sun295	

```
1
         Douglas Tucker 4478071379779230
                                             Sun4608
2
         Travis Walters 6011812112971322
                                             Sun4458
3
       Nathaniel Harris 4676137647685994
                                             Sun5260
4
           Tonya Carter 4832732618637221
                                             Sun2251
222
         Jason Lawrence 3505302934650403
                                             Fri6624
223
            Mary Rivera 5343428579353069
                                             Fri6014
224
      Ronald Vaughn DVM 341503466406403
                                             Fri5959
        Whitney Arnold 3579111947217428
225
                                             Fri6665
             Ruth Weiss 5268689490381635
226
                                             Fri6359
[95 rows x 11 columns]
solo_sun_fri = tips[mask]
solo_sun_fri["day"].value_counts()
day
Sun
       76
Fri
       19
Name: count, dtype: int64
# Il risultato è lo stesso
```