```
In [1]:
          import pandas as pd
          dataset=pd.read_csv("example.csv")
In [2]:
In [3]:
          dataset
Out[3]:
                      first_name
                   id
                                  last_name
                                                                    email
                                                                          gender
                                                                                      ip_address
                                                                                                        city
                                                                                                       Pedro
             0
                   1
                          Joseph
                                     Zaniolini
                                             jzaniolini0@simplemachines.org
                                                                             Male
                                                                                   163.168.68.132
                                                                                                    Leopoldo
             1
                   2
                          Freida
                                  Drillingcourt
                                                   fdrillingcourt1@umich.edu
                                                                           Female
                                                                                    97.212.102.79
                                                                                                        Buri
             2
                   3
                           Nanni
                                       Herity
                                                   nherity2@statcounter.com
                                                                           Female
                                                                                   145.151.178.98
                                                                                                      Claver
             3
                   4
                                                                                   25.199.143.143
                          Orazio
                                     Frayling
                                                  ofrayling3@economist.com
                                                                             Male
                                                                                                      Kungur
                                                                                                      Sainte-
             4
                   5
                         Julianne
                                    Murrison
                                                   imurrison4@cbslocal.com
                                                                           Female
                                                                                   10.186.243.144
                                                                                                    Luce-sur-
                                                                                                       Loire
                              ...
                                          ...
           995
                 996
                         Parrnell
                                                                                     159.55.55.39
                                         Jeff
                                                       pjeffrn@4shared.com
                                                                             Male
                                                                                                     Dengok
                                                                                                   Rosário do
           996
                 997
                             Pat
                                     Warstall
                                                   pwarstallro@dropbox.com
                                                                             Male
                                                                                    169.200.7.131
                                                                                                         Sul
                                   Vannuccini
           997
                 998
                           Willyt
                                                  wvannuccinirp@tuttocitta.it
                                                                          Female
                                                                                   156.149.18.155
                                                                                                  Xiaopingba
           998
                 999
                          Gaelan
                                  Petrashkov
                                               gpetrashkovrq@hostgator.com
                                                                             Male
                                                                                   219.109.169.75
                                                                                                   Bethlehem
                1000
                                                                                                  Potok Złoty
           999
                          Stirling
                                    Francesc
                                               sfrancescrr@cyberchimps.com
                                                                             Male
                                                                                    108 193 9 206
          1000 rows × 7 columns
          dataset['id'].sum()
Out[4]:
          500500
In [5]:
          dataset.sum()
Out[5]: id
                                                                               500500
                           JosephFreidaNanniOrazioJulianneLucyDyanaKassey...
          first_name
          last_name
                           ZanioliniDrillingcourtHerityFraylingMurrisonGa...
          email
                           jzaniolini@@simplemachines.orgfdrillingcourt1@...
          gender
                           MaleFemaleFemaleFemaleFemaleFemaleFemaleMa...
          ip_address
                           163.168.68.13297.212.102.79145.151.178.9825.19...
          city
                           Pedro LeopoldoBuriClaverKungurSainte-Luce-sur-...
          dtype: object
          dataset['id'].mean()
In [6]:
Out[6]:
          500.5
```

In [7]: dataset.describe()

Out[7]:

id 1000.000000 count 500.500000 mean 288.819436 std 1.000000 min 25% 250.750000 50% 500.500000 75% 750.250000 max 1000.000000

In [8]: len(dataset)

Out[8]: 1000

In [9]: dataset.head()

Out[9]:

	id	first_name	last_name	email	gender	ip_address	city
0	1	Joseph	Zaniolini	jzaniolini0@simplemachines.org	Male	163.168.68.132	Pedro Leopoldo
1	2	Freida	Drillingcourt	fdrillingcourt1@umich.edu	Female	97.212.102.79	Buri
2	3	Nanni	Herity	nherity2@statcounter.com	Female	145.151.178.98	Claver
3	4	Orazio	Frayling	ofrayling3@economist.com	Male	25.199.143.143	Kungur
4	5	Julianne	Murrison	jmurrison4@cbslocal.com	Female	10.186.243.144	Sainte-Luce- sur-Loire

In [10]: dataset.tail()

Out[10]:

	id	first_name	last_name	email	gender	ip_address	city
995	996	Parrnell	Jeff	pjeffrn@4shared.com	Male	159.55.55.39	Dengok
996	997	Pat	Warstall	pwarstallro@dropbox.com	Male	169.200.7.131	Rosário do Su l
997	998	Willyt	Vannuccini	wvannuccinirp@tuttocitta.it	Female	156.149.18.155	Xiaopingba
998	999	Gaelan	Petrashkov	gpetrashkovrq@hostgator.com	Male	219.109.169.75	Bethlehem
999	1000	Stirling	Francesc	sfrancescrr@cyberchimps.com	Male	108.193.9.206	Potok Złoty

. 10 AW								pandas - Jupyter N	•0.0000	,			
In [11]:	ds= ds	dat	taset	t.iloc	[0:1]]							
Out[11]:		id	first	_name	last_	name		•	email	gende	er ip	_address	city
	0	1	,	Joseph	Za	anio l ini j	zaniolii	ni0@simplemachine	s.org	Mal	e 163.1	68.68.132	Pedro Leopoldo
In [12]:	ds												
Out[12]:		id	first	_name	last_	name		6	email	gende	er ip	_address	city
	0	1	,	Joseph	Za	inio l ini j	zaniolii	ni0@simplemachine	s.org	Mal	e 163.1	68.68.132	Pedro Leopoldo
In [13]:	ds=	pd	.Data	aFrame	(ds)								
In [14]:	ds												
Out[14]:		id	first	_name	last_	name		•	email	gende	er ip	_address	city
	0	1	,	Joseph	Za	niolini j	zaniolii	ni0@simplemachine	s.org	Mal	e 163.1	68.68.132	Pedro Leopoldo
In [15]:			terin et[da	ng ataset	.id>	500]							
Out[15]:			id	first_n	ame	last_	name			email	gender	ip_address	5
	50	0	501	An	nbur	Н	oldey	aholdeydw@xir	huane	et.com	Female	236.209.51.23	1 Sawaer
	50	1	502	Sa	bine	L	_yosik	slyosil	kdx@3	360.cn	Female	161.232.137.136	Bayt '
	50	2	503	Gla	adys	I	Bandt	gbandtd	y@ocr	n.ne.jp	Female	158.14.30.142	2 Hur
	50	3	504	P	aten	Vasyukh	ichev	pvasyukhichevdz(@illino	is.edu	Male	175.216.192.130) Nand
	50	4	505	Hu	ınter	Cre	eggan	hcreggane0@ar	mazon	.co.uk	Male	208.173.52.40) Albi
	99	5	996	Par	rnell		Jeff	pjeffrn@ ²	Ishare	d.com	Male	159.55.55.39) Der
	99	6	997		Pat	Wa	arstall	pwarstallro@d	dropbo	x.com	Male	169.200.7.13 ⁻	l Rosári
	99	7	998	V	Villyt	Vann	uccini	wvannuccinirp	@tutto	citta.it	Female	156.149.18.15	5 Xiaopir
	99	8	999	Ga	elan	Petra	shkov	gpetrashkovrq@ho	ostgato	or.com	Male	219.109.169.75	5 Bethle
	99	9	1000	Sti	rling	Fra	ncesc	sfrancescrr@cyber	chimp	s.com	Male	108.193.9.206	6 Potok 2
	500) ro	ws ×	7 colun	nns								

In [16]: dataset.rename(columns={'ip_address':'ip'})

\sim	. 4	Γ1 <i>Γ</i> '	1 .
UI.	н	l In	1 1
		L = ~ .	

	id	first_name	last_name	email	gender	ip	city
0	1	Joseph	Zaniolini	jzaniolini0@simplemachines.org	Male	163.168.68.132	Pedro Leopoldo
1	2	Freida	Drillingcourt	fdrillingcourt1@umich.edu	Female	97.212.102.79	Buri
2	3	Nanni	Herity	nherity2@statcounter.com	Female	145.151.178.98	Claver
3	4	Orazio	Frayling	ofrayling3@economist.com	Male	25.199.143.143	Kungur
4	5	Julianne	Murrison	jmurrison4@cbslocal.com	Female	10.186.243.144	Sainte- Luce-sur- Loire
995	996	Parrnell	Jeff	pjeffrn@4shared.com	Male	159.55.55.39	Dengok
996	997	Pat	Warstall	pwarstallro@dropbox.com	Male	169.200.7.131	Rosário do Sul
997	998	Willyt	Vannuccini	wvannuccinirp@tuttocitta.it	Female	156.149.18.155	Xiaopingba
998	999	Gaelan	Petrashkov	gpetrashkovrq@hostgator.com	Male	219.109.169.75	Bethlehem
999	1000	Stirling	Francesc	sfrancescrr@cyberchimps.com	Male	108.193.9.206	Potok Złoty

1000 rows × 7 columns

```
In [17]: df=dataset.drop_duplicates()
    print(df,id(df),id(dataset))
```

0 1 2 3 4	id - 1 2 3 4 5	first_name Joseph Freida Nanni Orazio Julianne	last_name Zaniolini Drillingcourt Herity Frayling Murrison	email gender \ jzaniolini@simplemachines.org Male fdrillingcourt1@umich.edu Female nherity2@statcounter.com Female ofrayling3@economist.com Male jmurrison4@cbslocal.com Female	\
995	996	Parrnell	Jeff	pjeffrn@4shared.com Male	
996	997	Pat	Warstall	<pre>pwarstallro@dropbox.com Male</pre>	
997	998	Willyt	Vannuccini	wvannuccinirp@tuttocitta.it Female	
998	999	Gaelan	Petrashkov	<pre>gpetrashkovrq@hostgator.com Male</pre>	
999	1000	Stirling	Francesc	sfrancescrr@cyberchimps.com Male	
0 1	163.1	o_address 58.68.132 12.102.79	Pedro Le	city opoldo Buri	
2	145.1	51.178.98		Claver	
3	25.19	9.143.143		Kungur	
4	10.18	5.243.144	Sainte-Luce-sur	-Loire	
	150	 .55.55.39		Dongok	
995 996		200.7.131	Rosário	Dengok	
997		49.18.155		pingba	
998		9.169.75		hlehem	
999		193.9.206		Złoty	
222	100.	193.9.200	PULUK	21019	

[1000 rows x 7 columns] 2181652110480 2181651126048

```
In [18]: df=dataset.dropna()
    print(df,id(df),id(dataset))
```

	id -	first_name	last_name	email	gender	\
0	1	Joseph	Zaniolini	jzaniolini0@simplemachines.org	Male	
1	2	Freida	Drillingcourt	fdrillingcourt1@umich.edu	Female	
2	3	Nanni	Herity	nherity2@statcounter.com	Female	
3	4	Orazio	Frayling	ofrayling3@economist.com	Male	
4	5	Julianne	Murrison	jmurrison4@cbslocal.com	Female	
• •		• • •	• • •	•••	• • •	
995	996	Parrnell	Jeff	pjeffrn@4shared.com	Male	
996	997	Pat	Warstall	pwarstallro@dropbox.com	Male	
997	998	Willyt	Vannuccini	wvannuccinirp@tuttocitta.it	Female	
998	999	Gaelan	Petrashkov	gpetrashkovrq@hostgator.com	Male	
999	1000	Stirling	Francesc	sfrancescrr@cyberchimps.com	Male	
	i	o_address		city		
0		58.68.132	Pedro Le			
1		12.102.79	i cai o Le	Buri		
2		51.178.98		Claver		
3		9.143.143		Kungur		
4		5.243.144	Sainte-Luce-sur	•		
	10.10		Jainee Lace Jai			
 995	150	 .55.55.39		Dengok		
996		200.7.131	Rosário	•		
997		49.18.155		ppingba		
998		9.169.75		:hlehem		
999		193.9.206		z Złoty		
フフブ	T00.	193.3.200	POLOR	. 210ty		

[1000 rows x 7 columns] 2181651133056 2181651126048

In [19]: dataset.fillna(value="shiva")

$\Omega_{\rm L}$	ı+ l	[19]	١ ٠
OU		120	1 4

	id	first_name	last_name	email	gender	ip_address	city
0	1	Joseph	Zaniolini	jzaniolini0@simplemachines.org	Male	163.168.68.132	Pedro Leopoldo
1	2	Freida	Drillingcourt	fdrillingcourt1@umich.edu	Female	97.212.102.79	Buri
2	3	Nanni	Herity	nherity2@statcounter.com	Female	145.151.178.98	Claver
3	4	Orazio	Frayling	ofrayling3@economist.com	Male	25.199.143.143	Kungur
4	5	Julianne	Murrison	jmurrison4@cbslocal.com	Female	10.186.243.144	Sainte- Luce-sur- Loire
995	996	Parrnell	Jeff	pjeffrn@4shared.com	Male	159.55.55.39	Dengok
996	997	Pat	Warstall	pwarstallro@dropbox.com	Male	169.200.7.131	Rosário do Sul
997	998	Willyt	Vannuccini	wvannuccinirp@tuttocitta.it	Female	156.149.18.155	Xiaopingba
998	999	Gaelan	Petrashkov	gpetrashkovrq@hostgator.com	Male	219.109.169.75	Bethlehem
999	1000	Stirling	Francesc	sfrancescrr@cyberchimps.com	Male	108.193.9.206	Potok Złoty

1000 rows × 7 columns

4

In [20]: pd.isnull(dataset)

Out[20]:

	id	first_name	last_name	email	gender	ip_address	city
0	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False
995	False	False	False	False	False	False	False
996	False	False	False	False	False	False	False
997	False	False	False	False	False	False	False
998	False	False	False	False	False	False	False
999	False	False	False	False	False	False	False

1000 rows × 7 columns

In [21]: # viewing data
df.head(4)

~ .	4.1	F 0 4 T
Οι	ıτ	21

_	city	ip_address	gender	email	last_name	first_name	id	
_	Pedro Leopoldo	163.168.68.132	Male	jzaniolini0@simplemachines.org	Zaniolini	Joseph	1	0
i	Bur	97.212.102.79	Female	fdrillingcourt1@umich.edu	Drillingcourt	Freida	2	1
r	Clave	145.151.178.98	Female	nherity2@statcounter.com	Herity	Nanni	3	2
r	Kungu	25.199.143.143	Male	ofrayling3@economist.com	Frayling	Orazio	4	3

In [22]: df.tail(4)

0+	เวา
out	I ZZ I

city	ip_address	gender	email	last_name	first_name	id	
Rosário do Sul	169.200.7.131	Male	pwarstallro@dropbox.com	Warstall	Pat	997	996
Xiaopingba	156.149.18.155	Female	wvannuccinirp@tuttocitta.it	Vannuccini	Willyt	998	997
Bethlehem	219.109.169.75	Male	gpetrashkovrq@hostgator.com	Petrashkov	Gaelan	999	998
Potok Złoty	108.193.9.206	Male	sfrancescrr@cyberchimps.com	Francesc	Stirling	1000	999

In [23]: df.dtypes

Out[23]: id

id int64
first_name object
last_name object
email object
gender object
ip_address object
city object
dtype: object

In [24]: df.columns

In [25]: df.sort_values(by=['id'])

_			
()	тι	1 /5 1	
Ou	C I		

	id	first_name	last_name	email	gender	ip_address	city
0	1	Joseph	Zaniolini	jzaniolini0@simplemachines.org	Male	163.168.68.132	Pedro Leopoldo
1	2	Freida	Drillingcourt	fdrillingcourt1@umich.edu	Female	97.212.102.79	Buri
2	3	Nanni	Herity	nherity2@statcounter.com	Female	145.151.178.98	Claver
3	4	Orazio	Frayling	ofrayling3@economist.com	Male	25.199.143.143	Kungur
4	5	Julianne	Murrison	jmurrison4@cbslocal.com	Female	10.186.243.144	Sainte- Luce-sur- Loire
995	996	Parrnell	Jeff	pjeffrn@4shared.com	Male	159.55.55.39	Dengok
996	997	Pat	Warstall	pwarstallro@dropbox.com	Male	169.200.7.131	Rosário do Sul
997	998	Willyt	Vannuccini	wvannuccinirp@tuttocitta.it	Female	156.149.18.155	Xiaopingba
998	999	Gaelan	Petrashkov	gpetrashkovrq@hostgator.com	Male	219.109.169.75	Bethlehem
999	1000	Stirling	Francesc	sfrancescrr@cyberchimps.com	Male	108.193.9.206	Potok Złoty

1000 rows × 7 columns

4

In [26]: df.sort_values(by=['email'])

Out[26]:

	ip_address	gender	email	last_name	first_name	id	
Palaic	209.213.159.128	Male	aairds63@drupal.org	Airds	Albrecht	220	219
D	14.20.177.82	Female	aalletthy@accuweather.com	Allett	Abbie	647	646
	0.192.253.81	Male	aantonsenq8@whitehouse.gov	Antonsen	Aaron	945	944
	50.148.243.117	Female	aayceic@mail.ru	Ayce	Alex	661	660
Е	218.209.50.211	Male	abaker9w@spiegel.de	Baker	Aaron	357	356
	153.15.188.46	Male	yhurranoh@flickr.com	Hurran	Yorke	882	881
	115.219.51.8	Male	ytattershallmm@spiegel.de	Tattershall	Yuma	815	814
1	193.255.184.166	Male	ytockellb8@moonfruit.com	Tockell	York	405	404
(207.96.93.51	Female	zmacmenemyd1@seattletimes.com	MacMenemy	Zorah	470	469
	108.125.27.193	Male	zsparrowhawkmw@vimeo.com	Sparrowhawk	Zelig	825	824

1000 rows × 7 columns

 \blacktriangleleft

```
In [27]:
          # viewing data
          df[0:3]
Out[27]:
              id first_name
                             last_name
                                                            email gender
                                                                             ip_address
                                                                                                 city
                                                                                               Pedro
           0
              1
                               Zaniolini jzaniolini0@simplemachines.org
                                                                          163.168.68.132
                     Joseph
                                                                     Male
                                                                                            Leopoldo
               2
                            Drillingcourt
                                            fdrillingcourt1@umich.edu
                                                                   Female
                                                                           97.212.102.79
                                                                                                 Buri
                     Freida
           2
              3
                                            nherity2@statcounter.com Female
                                                                         145.151.178.98
                                                                                               Claver
                      Nanni
                                 Herity
In [29]: df["ip_address"]
Out[29]: 0
                  163.168.68.132
                   97.212.102.79
          2
                  145.151.178.98
          3
                  25.199.143.143
                  10.186.243.144
          995
                     159.55.55.39
          996
                   169.200.7.131
          997
                  156.149.18.155
          998
                  219.109.169.75
          999
                   108.193.9.206
          Name: ip_address, Length: 1000, dtype: object
In [31]: df.iloc[2,3]
Out[31]: 'nherity2@statcounter.com'
In [36]: df1=df[["id",'email']]
          df2=df[["ip_address","email"]]
```

```
pd.merge(df1,df2,on='email')
In [35]:
         KeyError
                                                     Traceback (most recent call last)
         Input In [35], in <cell line: 1>()
          ----> 1 pd.merge(df1,df2,on='email')
         File ~\AppData\Roaming\Python\Python310\site-packages\pandas\core\reshape\mer
         ge.py:107, in merge(left, right, how, on, left_on, right_on, left_index, right
         t_index, sort, suffixes, copy, indicator, validate)
              90 @Substitution("\nleft : DataFrame or named Series")
              91 @Appender(_merge_doc, indents=0)
              92 def merge(
             (\ldots)
             105
                      validate: str | None = None,
              106 ) -> DataFrame:
                      op = MergeOperation(
          --> 107
             108
                          left,
             109
                          right,
             110
                          how=how,
             111
                          on=on,
                          left_on=left_on,
             112
                          right on=right on,
             113
                          left_index=left_index,
             114
             115
                          right_index=right_index,
                          sort=sort,
             116
             117
                          suffixes=suffixes,
             118
                          copy=copy,
                          indicator=indicator,
             119
             120
                          validate=validate,
             121
                      )
             122
                      return op.get result()
         File ~\AppData\Roaming\Python\Python310\site-packages\pandas\core\reshape\mer
         ge.py:700, in MergeOperation. init (self, left, right, how, on, left on, r
         ight on, axis, left index, right index, sort, suffixes, copy, indicator, vali
         date)
             693 self. cross = cross col
              695 # note this function has side effects
              696 (
             697
                      self.left_join_keys,
                      self.right join keys,
             698
             699
                      self.join_names,
          --> 700 ) = self._get_merge_keys()
              702 # validate the merge keys dtypes. We may need to coerce
              703 # to avoid incompatible dtypes
             704 self._maybe_coerce_merge_keys()
         File ~\AppData\Roaming\Python\Python310\site-packages\pandas\core\reshape\mer
         ge.py:1097, in _MergeOperation._get_merge_keys(self)
            1095 if not is rkey(rk):
            1096
                      if rk is not None:
          -> 1097
                          right_keys.append(right._get_label_or_level_values(rk))
            1098
                      else:
            1099
                          # work-around for merge_asof(right_index=True)
                          right_keys.append(right.index)
            1100
```

```
File ~\AppData\Roaming\Python\Python310\site-packages\pandas\core\generic.py:
1840, in NDFrame._get_label_or_level_values(self, key, axis)
    1838     values = self.axes[axis].get_level_values(key)._values
    1839     else:
-> 1840         raise KeyError(key)
         1842         # Check for duplicates
         1843         if values.ndim > 1:
KeyError: 'email'
```

In []: