

To Start

```
import numpy as np
import pandas as pd
```

Create

```
pd.DataFrame(
    dict/list,
    index = None,
    columns = None)
```

create DataFrame from list or dictionary

```
df.index[ names]
```

set custom indexes

```
pd.Series(
    list/np_array/dict,
    index = None)
```

create series from the list or np_array or dictionary

Input and Output

```
pd.read_csv( 'name',
    index_col = None)
```

read csv

```
pd.read_excel(' name'
)
```

read excel

```
df.to_csv ('n ame',
    index = False)
```

save to csv

```
df.to_excel( 'name',
    'sheet_name = 'name',
    index = False)
```

save to excel

Iteration

```
for lab, row in df.iterrows():
    print(lab)
    print(row)
```

Functions/Methods

```
s.drop (ro w_i ndex, axis = 0
)
```

drop values from rows of series

```
df.drop(c ol_ name, axis = 0
)
```

drop values from columns

```
df.drop(c olu mns =[c ol_ na
mes])
```

drop columns from DataFrame

```
df.drop_d upl ica tes()
```

remove duplicate rows (only considers columns)

```
df.sort_i ndex(by = col_name
s)
```

sort by the values along an axis

```
df.sort_values(
    by = col_names,
    ascending = False)
```

order rows by values of a column high to low

```
df.ren ame (co lumns =
{'old_name':'new_name'})
```

rename the columns of a DataFrame

```
df.rank()
```

assign ranks to entries

```
pd.con cat ([d f1, df2])
```

append rows of DataFrames

```
len(df)
```

number of rows in DataFrame

```
df1.jo in (df2)
```

join two DataFrames

```
df['co l_ n ame '].u ni que()
```

return unique values from column

```
df[col_name].apply(
    func/type.method)
```

apply function to column

```
df.app ly ( f un c/t yp e.
m e thod)
```

apply function



By monika.g
cheatography.com/monika-g/

Not published yet.
Last updated 6th February, 2023.
Page 1 of 2.

Sponsored by [ApolloPad.com](https://apollopad.com)
Everyone has a novel in them. Finish Yours!
<https://apollopad.com>

Extract

<code>df[col _names]</code>	series of column
<code>df[[co l_n ames]]</code>	select column
<code>df[sta rt:end]</code>	select many columns
<code>df.loc [in dex _name]</code>	select row
<code>df.ilo c[i nde x_num]</code>	select rows
<code>df.loc [[i nde x_n ames]]</code>	select rows
<code>df.ilo c[[ind ex_ nums]]</code>	select rows
<code>df.loc [[i nde x_n ames], [col_names]]</code>	select rows and columns
<code>df.loc[:, [col_n ames]]</code>	select all rows and few columns
<code>df.iloc[:, [col_n ums]]</code>	select all rows and few columns
<code>df.head(n)</code>	select first n rows
<code>df.tail(n)</code>	select last n rows
<code>df.fil ter (regex = 'regex ')</code>	select columns whose name matches regular expression regex

Boolean Operators

<code>df[np.logica l_and(con1, con2, ...)]</code>	1 'and' 2 condition ...
<code>df.loc [con1 & con2]</code>	1 'and' 2 condition ...
<code>df[np.logical_or(con1, con2, ...)]</code>	1 'or' 2 condition ...
<code>df.loc [con1 con2]</code>	1 'or' 2 condition ...
<code>df[np.logica l_not(con)]</code>	'not' condition
<code>df.loc[~ con1]</code>	'not' condition
<code>df[var]</code>	for condition

Get DataFrame Information

<code>df.shape</code>	(rows, columns)
<code>df.index</code>	decibe index
<code>df.columns</code>	describe DataFrame columns
<code>df.info()</code>	info on DataFrame
<code>df.count()</code>	number of non_NA values for columns
<code>df.des cribe()</code>	summary statics

Math

<code>df.sum()</code>	sum of values for columns
<code>df.cum sum()</code>	cummulative sum of values
<code>df.min()</code>	minimum values for columns
<code>df.max()</code>	maximum values for columns
<code>df.med ian()</code>	median of values columns
<code>df.mean()</code>	mean of values for columns
<code>df.std()</code>	standard deviation of each object
<code>df.var()</code>	variance of each object



By monika.g

cheatography.com/monika-g/

Not published yet.

Last updated 6th February, 2023.

Page 2 of 2.

Sponsored by [ApolloPad.com](https://apollopadd.com)

Everyone has a novel in them. Finish
Yours!

<https://apollopadd.com>