

### To Start

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

### Create

```
pd.DataFrame(          create DataFrame from list or
dict/list,            dictionary
index = None,
columns = None)
```

```
df.index[ names]      set custom indexes
```

```
pd.Series(            create series from the list or
list/np_array/dict,  np_array or dictionary
index = None)
```

### Input and Output

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

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

```
df.to_csv( 'name',    save to csv
index = False)
```

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

### Iteration

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

### Functions/Methods

```
s.drop( row_index, axis = 0)  drop values from rows of
)                               series
```

```
df.drop( col_name, axis = 0)  drop values from columns
)
```

```
df.drop( columns = [ col_name] drop columns from
)                               DataFrame
```

```
df.drop_duplicates()         remove duplicate rows
(only considers columns)
```

```
df.sort_index( by = col_name  sort by the values along
)                               an axis
```

```
df.sort_values(            order rows by values of a
by = col_names,           column high to low
ascending = False)
```

```
df.rename( columns =        rename the columns of a
{'old_name': 'new_name'})  DataFrame
```

```
df.rank()                  assign ranks to entries
```

```
pd.concat( [df1, df2])     append rows of
DataFrames
```

```
len(df)                   number of rows in
DataFrame
```

```
df1.join( df2)            join two DataFrames
```

```
df['col_name'].unique()    return unique values from
column
```

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

```
df.apply( func/type.method) apply function
```



By [monika.g](https://cheatography.com/monika-g/)  
cheatography.com/monika-g/

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

Sponsored by [Readable.com](https://readable.com)  
Measure your website readability!  
<https://readable.com>

### Extract

<code>df[col_names]</code>	series of column
<code>df[[col_names]]</code>	select column
<code>df[start:end]</code>	select many columns
<code>df.loc[index_name]</code>	select row
<code>df.iloc[index_num]</code>	select rows
<code>df.loc[[index_names]]</code>	select rows
<code>df.iloc[[index_nums]]</code>	select rows
<code>df.loc[[index_names], col_names]</code>	select rows and columns
<code>df.loc[:, col_names]</code>	select all rows and few columns
<code>df.iloc[:, col_nums]</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.filter(regex = 'regex')</code>	select columns whose name matches regular expression regex

### Boolean Operators

<code>df[np.logical_and(con1, con2, ...)]</code>	1 'and' 2 condition ...
<code>df.loc[con1 &amp; 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.logical_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>	describe 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.describe()</code>	summary statistics

### Math

<code>df.sum()</code>	sum of values for columns
<code>df.cumsum()</code>	cumulative sum of values
<code>df.min()</code>	minimum values for columns
<code>df.max()</code>	maximum values for columns
<code>df.median()</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](https://monika.g)  
[cheatography.com/monika-g/](https://cheatography.com/monika-g/)

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

Sponsored by [Readable.com](https://readable.com)  
 Measure your website readability!  
<https://readable.com>