



pandas DATAFRAME CHEATSHEET

A DataFrame is a two-dimensional (i.e., rows x columns) data structure. Pandas provides a number of functions to create and manipulate DataFrames.

For more Python packages related resources visit

activestate.com/learn-python

Create a DataFrame	Result										
<p>By importing a CSV file:</p> <pre>df = pd.read_csv("Report_Card.csv")</pre> <p>Using a pandas series:</p> <pre>classes = pd.Series(["Mathematics", "Chemistry", "Physics"]) grades = pd.Series([90, 54, 77]) pd.DataFrame({"Classes": classes, "Grades": grades})</pre>	<pre>Classes 0 Mathematics 1 Chemistry 2 Physics</pre>										
Slice a DataFrame	Result										
<p>Get just the Math grade:</p> <pre>Grade = df.iloc[[0],[2,3]] Or Grade = df.loc[["Mathematics"],["Classes","Grades"]]</pre>	<pre>Classes 0 Mathematics</pre>										
Delete a Column	Result										
<pre>df.drop("Grades",axis=1,inplace=True) Or df.drop(3,axis=1,inplace=True)</pre>	<pre>Classes 0 Mathematics 1 Chemistry 2 Physics</pre>										
Delete a Row	Result										
<pre>df.drop("Physics",axis=0,inplace=True) Or df.drop(2,axis=0,inplace=True)</pre>	<pre>Classes 0 Mathematics 1 Chemistry</pre>										
Access an Element	Result										
<p>Using <i>at</i>:</p> <pre>df.loc[0].at["Grades"]</pre> <p>or</p> <p>Using <i>iat</i>:</p> <pre>df.loc[0].iat[2]</pre>	<pre>90</pre>										
Append Rows	Result										
<table border="1"><thead><tr><th>Classes</th><th>Grades</th></tr></thead><tbody><tr><td>Mathematics</td><td>90</td></tr><tr><td>Chemistry</td><td>54</td></tr></tbody></table> <table border="1"><thead><tr><th>Classes</th><th>Grades</th></tr></thead><tbody><tr><td>0 Physics</td><td>77</td></tr></tbody></table> <pre>pd.concat([df1, df2])</pre>	Classes	Grades	Mathematics	90	Chemistry	54	Classes	Grades	0 Physics	77	<pre>Classes Grades 0 Mathematics 90 1 Chemistry 54 2 Physics 77</pre>
Classes	Grades										
Mathematics	90										
Chemistry	54										
Classes	Grades										
0 Physics	77										
Append Columns	Result										
<table border="1"><thead><tr><th>Classes</th></tr></thead><tbody><tr><td>0 Mathematics</td></tr><tr><td>1 Chemistry</td></tr><tr><td>2 Physics</td></tr></tbody></table> <table border="1"><thead><tr><th>Grades</th></tr></thead><tbody><tr><td>0 90</td></tr><tr><td>1 54</td></tr><tr><td>2 77</td></tr></tbody></table> <pre>pd.concat([df1, df2], axis=1)</pre>	Classes	0 Mathematics	1 Chemistry	2 Physics	Grades	0 90	1 54	2 77	<pre>Classes Grades 0 Mathematics 90 1 Chemistry 54 2 Physics 77</pre>		
Classes											
0 Mathematics											
1 Chemistry											
2 Physics											
Grades											
0 90											
1 54											
2 77											
Rename Column	Result										
<pre>df.rename(columns = {'Classes':'Subjects'})</pre>	<pre>Subjects Grades 0 Mathematics 90 1 Chemistry 54 2 Physics 77</pre>										
Replace Values	Result										
<pre>df.replace([90,54,77],['A','D','B'])</pre>	<pre>Classes Grades 0 Mathematics A 1 Chemistry D 2 Physics B</pre>										