Cheatography

Essential Pandas Cheat Sheet

by Cheto via cheatography.com/162578/cs/34032/

Loading Pandas

Import Pandas Module with the alias *pd*

```
import pandas as pd
```

Creating Dataframes From Files

From a csv file

```
df = pd.rea d_c sv( 'fi le.c sv')
```

From a python dictionary

```
df = pd.Dat aFr ame.fr om_ dic t(< dic t>)
```

Displaying Dataframe Info

Display first five rows in dataframe

```
df.head()
```

Display last five rows in dataframe

```
df.tail()
```

Show all column names

```
df.columns
```

Show all object types in dataframe

```
df.dtypes
```

Show statistics for all int and float columns

```
df.des cribe()
```

Show statistics for 'object' type columns

```
df.des cri be( inc lud e=' obj ect')
```

Show number of rows and columns

```
df.shape
```

Display True for each NaN value, False otherwise

```
df.isn ull()
```

Display a table with the number of NaN values for each column

```
df.isn ull ().s um()
```

Updating

Delete all rows containing *NaN* values in the *df* Dataframe

```
df.dro pna (in pla ce= True)
```

Delete *'col_name'* column

```
df.dro p(' col _name', axis=1)
```

Example of a calculated column

```
df['ne w_col'] = df['co l_1'] + df['co l_2']
```

Update the entire column to value *<value>*

```
df['ne w_col'] = <va lue>
```

Update the cell at *(a,b)* to *<value>*

```
df.ilo c[a,b] = <va lue>
```

Update (or creates) *'col_a'* with the result of lambda function applied to *'col_b'*

```
df['co l_a'] = df['co l_b '].a pp ly( <lambda
functi on>)
```

Filtering Columns

Indexing with iloc

Displays the entire row indexed n

```
df.iloc[n]
```

Displays the element in row *n* & column *m*

```
df.iloc[n, m]
```

Displays a slice of rows: from row *a* to row *b*

```
df.ilo c[a:b]
```

Displays rows *a* to *b* only in the columns *c* to *d*

```
df[a:b, c:d]
```

Indexing with loc

Shows all rows indexed with `<in d>`

```
df.loc [<i nd>]
```

Manipulating Dataframes

Create a copy of the dataframe

```
new_df = df.copy()
```

Set *'column_name'* as the index

```
df.set _in dex ('c olu mn_ name', inplac e=True)
```

Delete / Output

Output to *csv* file

```
df.to_ csv ('o utp ut.c sv')
```

Output to *json* file

```
df.to_ json()
```

Output to *html* file

```
df.to_ html()
```

Delete a Dataframe

```
del df
```

Display an entire column as a series

```
df['co lum n_n ame']
```

Display all columns in the given list

```
df[['c ol_1', 'col_2', ... 'col_n' ]]
```

Show all unique elements in 'column_name'

```
df['co lum n_n ame '].u ni que()
```

Filtering Rows

Display all rows satisfying <condition>

```
df[<co ndi tio n>]
```

Display all rows where `df['co l_n ame'] == <va lue>`

```
df[df[ 'co l_n ame'] == <va lue >]
```

Show all rows satisfying both conditions

```
df[(<c ond iti on_ 1>]) & (<c ond iti on_ 2>)]
```

By **Cheto**
cheatography.com/cheto/

Published 6th September, 2022.
Last updated 6th September, 2022.