Cheatography

Pandas Less Frequent Functions Cheat Sheet

by Anoikis via cheatography.com/74320/cs/18714/

Read and Write

```
import pandas as pd
```

```
pd.rea d_c sv( fil epa th_ or_ buffer, sep=', ', names= None, index_ col =None)
```

```
df.to_ csv (pa th_ or_buf, sep, column s=None, header =True, index= True, index_ lab el= None, mode='w')
```

Numerical Features and Nans

`df.sort_values( by=feat, ascending=False)`	sort table according to the values of a columns
`df[feat].dropna()`	remove lines with a NaN value
`df[feat].isna()`	check if the value of *feat* is NaN in the table

Categorical Features

pd.get_dummies(df[*feature*])	Transform a categorical feature into dummy variables

Features Visualization

`df[feats].plot (ki nd= ['d ensity' | 'bar'], subplo ts= True, layout=(1, 2), sharex=, figsize=);`	distibution of numeric features

General Infos and Basic Statistics

`df.info()`	general infos
`df.des cribe()`	basic statistics on numerical features
`df.describe(include=[ 'object', 'bool'])`	include non-numerical features

Apply Functions

`df.apply(my_function) # ex: df.app ly( lambda x: )`	apply a function
`df['feat'] = df['feat'].map(d) # or df = df.replace({'feat': d})`	replace values in a column according to dict d

Group by

`df.gro upby(['feat']) [columns_to_keep].func()`	group by a feature
`df.gro upby(['feat']) [columns_to_keep]. agg([list_of_functions])`	group by a feature and apply several functions

Cross Tables

`pd.cro sst ab(df['feat1'], df['feat2'], normal ize=)`	confusion matrix
`df.pivot_table( ['featur es_ to_ analyze'], ['groupi ng_feat'], aggfunc='mean')`	pivot table

By **Anoikis**
cheatography.com/anoikis/

Not published yet.
Last updated 11th October, 2019.