

```

In [1]: ## How to utilise Pandas dataframe & series for data wrangling
def Snippet_112():
    print()
    print(format('How to utilise a Pandas dataframe & series for data wrangling','^82'))
    import warnings
    warnings.filterwarnings("ignore")
    # load libraries
    import pandas as pd
    # Series are one-dimensional arrays (like R's vectors)
    # Create a series of the number of floodingReports
    floodingReports = pd.Series([5, 6, 2, 9, 12])
    print(); print(floodingReports)
    # Set county names to be the index of the floodingReports series
    floodingReports = pd.Series([5, 6, 2, 9, 12], index=['Cochise County', 'Pima County',
                                                    'Santa Cruz County', 'Maricopa County', 'Yuma County'])

    print(); print(floodingReports)
    # View the number of floodingReports in Cochise County
    print(); print(floodingReports['Cochise County'])
    # View the counties with more than 6 flooding reports
    print(); print(floodingReports[floodingReports > 6])

    # Create a pandas series from a dictionary
    fireReports_dict = {'Cochise County': 12, 'Pima County': 342,
                       'Santa Cruz County': 13, 'Maricopa County': 42,
                       'Yuma County': 52}
    # Convert the dictionary into a pd.Series, and view it
    fireReports = pd.Series(fireReports_dict);
    print(); print(fireReports)
    # Change the index of a series to shorter names
    fireReports.index = ["Cochice", "Pima", "Santa Cruz", "Maricopa", "Yuma"]

    # DataFrames are like R's Dataframes
    # Create a dataframe from a dict of equal length lists or numpy arrays
    data = {'county': ['Cochice', 'Pima', 'Santa Cruz', 'Maricopa', 'Yuma'],
           'year': [2012, 2012, 2013, 2014, 2014],
           'reports': [4, 24, 31, 2, 3]}
    df = pd.DataFrame(data)
    print(); print(df)

    # Set the order of the columns using the columns attribute
    dfColumnOrdered = pd.DataFrame(data, columns=['county', 'year', 'reports'])
    print(); print(dfColumnOrdered)
    # Add a column
    dfColumnOrdered['newsCoverage'] = pd.Series([42.3, 92.1, 12.2, 39.3, 30.2])
    print(); print(dfColumnOrdered)
    # Delete a column
    del dfColumnOrdered['newsCoverage']
    print(); print(dfColumnOrdered)
    # Transpose the dataframe
    print(); print(dfColumnOrdered.T)
Snippet_112()

```

\*\*\*\*\*How to utilise a Pandas dataframe & series for data wrangling\*\*\*\*\*

```

0    5
1    6
2    2
3    9
4   12
dtype: int64

Cochise County    5
Pima County      6
Santa Cruz County 2
Maricopa County  9
Yuma County     12
dtype: int64

5

Maricopa County    9
Yuma County     12
dtype: int64

Cochise County    12
Pima County     342
Santa Cruz County 13
Maricopa County  42
Yuma County     52
dtype: int64

   county  year  reports
0   Cochice 2012     4
1     Pima 2012    24
2  Santa Cruz 2013    31
3   Maricopa 2014     2
4     Yuma 2014     3

   county  year  reports
0   Cochice 2012     4
1     Pima 2012    24
2  Santa Cruz 2013    31
3   Maricopa 2014     2
4     Yuma 2014     3

   county  year  reports  newsCoverage
0   Cochice 2012     4         42.3
1     Pima 2012    24         92.1
2  Santa Cruz 2013    31         12.2
3   Maricopa 2014     2         39.3
4     Yuma 2014     3         30.2

   county  year  reports
0   Cochice 2012     4
1     Pima 2012    24
2  Santa Cruz 2013    31
3   Maricopa 2014     2
4     Yuma 2014     3

   county  year  reports
0   Cochice 2012     4
1     Pima 2012    24
2  Santa Cruz 2013    31
3   Maricopa 2014     2
4     Yuma 2014     3

   county  year  reports
0   Cochice 2012     4
1     Pima 2012    24
2  Santa Cruz 2013    31
3   Maricopa 2014     2
4     Yuma 2014     3

```

In [ ]: