tabular_data

June 12, 2021

1 Tabular Data and Data Formats

[1]:
```
# create a data frame using the counstructor pandas.DataFrame( data, index,␣
↪columns, dtype, copy)
```

[2]:
```
#Pandas does the following
#Provides a mechanism to load data objects from different formats
#Creates efficient data frame objects with default and customized indexing
#Reshapes and pivots date sets
#Provides efficient mechanisms to handle missing data
#Merges, groups by, aggregates, and transforms data
#Manipulates large data sets by implementing various functionalities such as␣
↪slicing, indexing, subsetting, deletion, and insertion
#Provides efficient time series functionality
```

1.1 Pandas Series

[3]:
```
# A series is a one-dimensional labeled array
```

[4]:
```
import pandas as pd #imports pandas library
import numpy as np #imports numpy library
data = np.array([90,75,50,66]) #creates an array
s = pd.Series(data,index=['A','B','C','D']) #assigns labels to the array
print (s)
```

```
A    90
B    75
C    50
D    66
dtype: int64
```

[5]:
```
data = {'Ahmed' : 92, 'Ali' : 55, 'Omar' : 83}
s = pd.Series(data,index=['Ali','Ahmed','Omar'])
print (s)
```

```
Ali      55
Ahmed    92
Omar     83
dtype: int64
```

1.2 Pandas Data Frame

[6]:
```
# A data frame is a two-dimensional data structure
```

[7]:
```
import pandas as pd
data = [['Robert',42],['Ahmed',35],['Ali',17],['Omar',25]]
DataFrame1 = pd.DataFrame(data,columns=['Name','Age'])
print (DataFrame1)
```

```
     Name  Age
0  Robert   42
1   Ahmed   35
2     Ali   17
3    Omar   25
```

[8]:
```
DataFrame1[1:] #retrieve data from a data frame starting from index 1 up to the␣
↪end of rows.
```

[8]:
```
     Name  Age
1   Ahmed   35
2     Ali   17
3    Omar   25
```

[9]:
```
#create a data frame using a dictionary.
```

[10]:
```
import pandas as pd
data = {'Name':['Robert','Ahmed', 'Ali', 'Omar','Salwa'],'Age':[42,35,17,25,30]}
dataframe2 = pd.DataFrame(data, index=[100, 101, 102, 103,104])
print (dataframe2)
```

```
       Name  Age
100  Robert   42
101   Ahmed   35
102     Ali   17
103    Omar   25
104   Salwa   30
```

[16]:
```
pd.DataFrame.from_dict(dict([("A", [1, 2, 3]), ("B", [4, 5, 6])]))
```

[16]:
```
   A  B
0  1  4
1  2  5
2  3  6
```