**Introduction about Pandas**

What is Pandas?

Pandas is a Python library used for working with data sets.

It has functions for analyzing, cleaning, exploring, and manipulating data.

The name "Pandas" has a reference to both "Panel Data", and "Python Data Analysis" and was created by Wes McKinney in 2008.

## Why Use Pandas?

Pandas allows us to analyze big data and make conclusions based on statistical theories.

Pandas can clean messy data sets, and make them readable and relevant.

Relevant data is very important in data science.

What Can Pandas Do?

Pandas gives you answers about the data. Like:

* Is there a correlation between two or more columns?
* What is average value?
* Max value?
* Min value?

Pandas are also able to delete rows that are not relevant, or contains wrong values, like empty or NULL values. This is called *cleaning* the data.

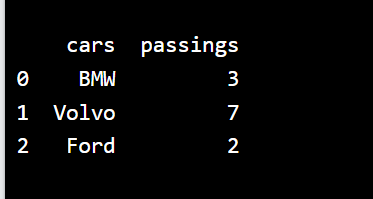
## What is a Data Frame?

A Pandas Data Frame is a 2-dimensional data structure, like a 2 dimensional array, or a table with rows and columns.

# create one dataframe

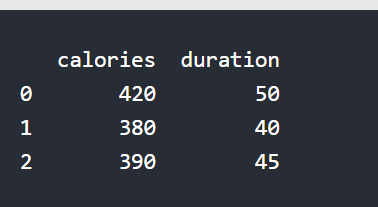
import pandas  
  
mydataset = {  
  'cars': ["BMW", "Volvo", "Ford"],  
  'passings': [3, 7, 2]  
}  
  
myvar = pandas.DataFrame(mydataset)  
  
print(myvar)

**Output:**

****

#Create a simple Pandas DataFrame with object of pandas

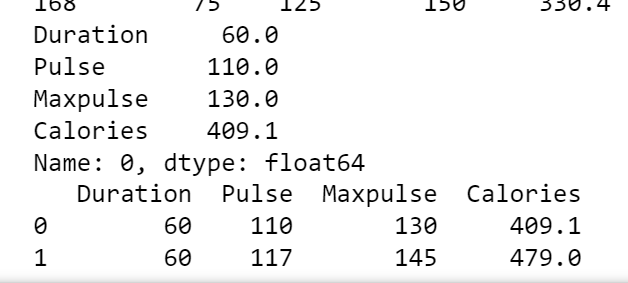
import pandas as pd  
  
data = {  
  "calories": [420, 380, 390],  
  "duration": [50, 40, 45]  
}  
  
#load data into a DataFrame object:  
df = pd.DataFrame(data)  
  
print(df)

****

Pandas use the loc attribute to return one or more specified row(s)

print(df.loc[0])

print(df.loc[[0, 1]])

****

Load the CSV into a Data Frame:

import pandas as pd

df = pd.read\_csv("C:/Users/sraba/Downloads/data.csv")

print(df.to\_string())

print(df.loc[0])

print(df.loc[[0, 1]])

**------------------------------------------------------------------------------------**

import pandas as pd  
  
pd.options.display.max\_rows = 2 # default it will show 60  
  
df = pd.read\_csv("C:/Users/sraba/Downloads/data.csv")  
  
print(df)