

PYTHON PANDAS QUICK THROUGH GUIDE



github.com/uzairafridi00

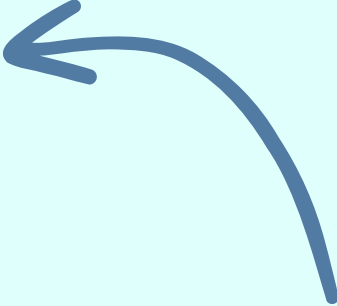
linkedin.com/in/uzair-afridi00

Introduction:

Python pandas is an open-source library widely used for data analysis.

Pandas library is used for reading and manipulating data in machine learning and data science.

```
pip install pandas
```




*pip command to install pandas
in your system*

Pandas DataFrame:

A pandas DataFrame is a 2-dimensional data array or table with rows and columns.

Create dataframe in pandas




```
1  import pandas as pd
2
3  car_dataset = {
4      'cars': ['Toyota', 'Honda', 'Suzuki'],
5      'Model': ['Gli', 'Civic', 'Swift'],
6  }
7  car_df = pd.DataFrame(car_dataset)
8  print(car_df)
```

Column Operation on DataFrame:

You can easily access the data frame columns using square brackets and also assign or update new values.

Below are some basic operations you can perform on a data frame column.




```
1  # Accessing Single Column
2  print(car_df[['cars']])
3  # Accessing Multiple Columns
4  print(car_df[['cars', 'Model']])
5  # Add New Column
6  car_df['New_Column_Name'] = [1,2,3]
7  # Delete Column
8  car_df.drop(columns=['New_Column_Name'],inplace=True)
9  # Rename Column
10 car_df.rename(columns={'Model':'model'},inplace=True)
```

Read CSV File:

A simple way to store big data sets is to use CSV files (Comma Separated Values).

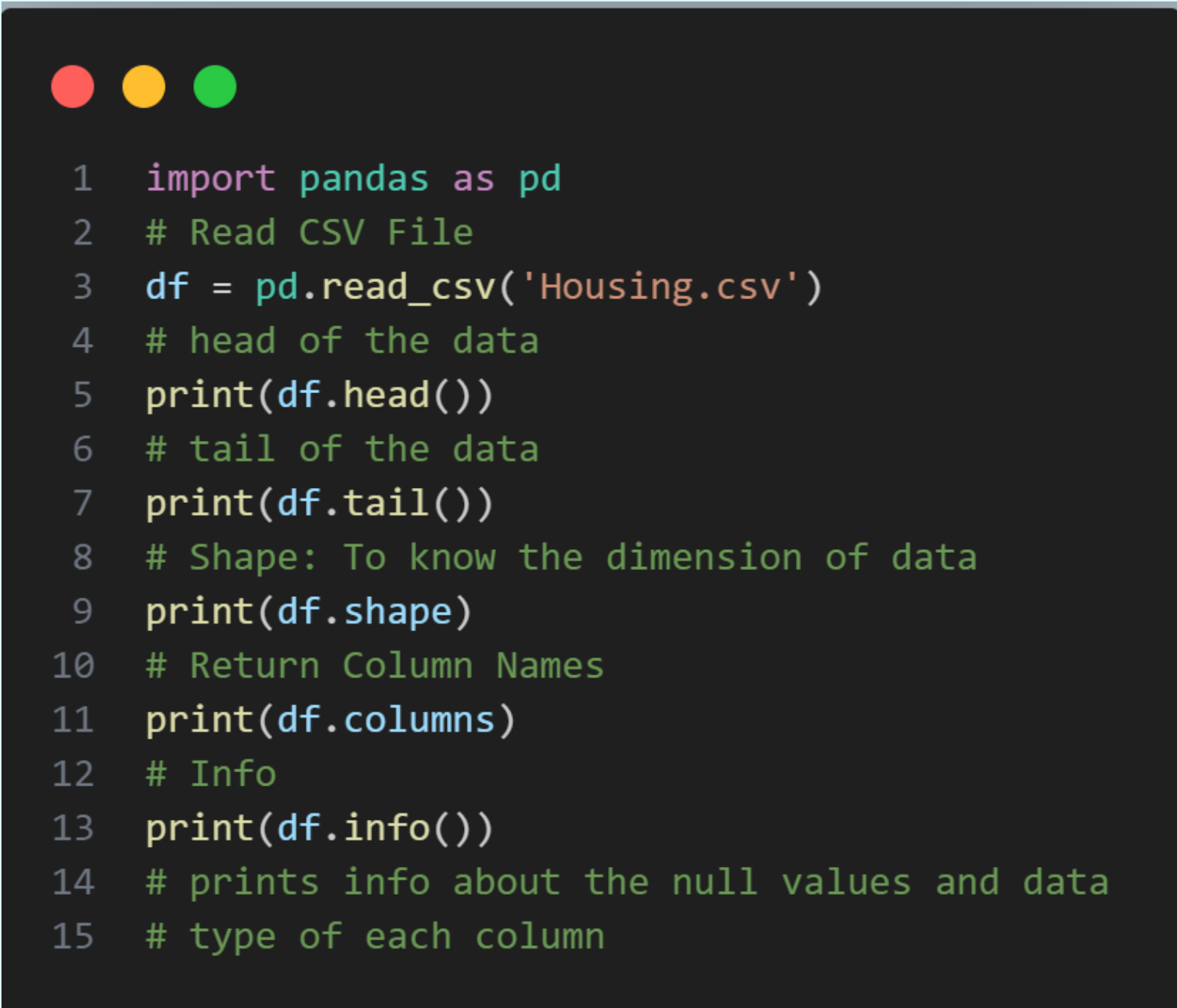
CSV files are the common files types you will use while working in Machine Learning or Data Science.



```
1  import pandas as pd
2
3  df = pd.read_csv('Housing.csv')
4  print(df)
```

Peek Into The Data:

To understand the high-level overview of data, pandas offers multiple functions and some of them are:



```
1  import pandas as pd
2  # Read CSV File
3  df = pd.read_csv('Housing.csv')
4  # head of the data
5  print(df.head())
6  # tail of the data
7  print(df.tail())
8  # Shape: To know the dimension of data
9  print(df.shape)
10 # Return Column Names
11 print(df.columns)
12 # Info
13 print(df.info())
14 # prints info about the null values and data
15 # type of each column
```

Statistical Analysis Using Pandas:

Pandas offer some functions that help you to dig deeper and find more useful insight from the data.



```
1  # describe: returns statistics measures like mean, min,  
2  # max, standard deviation and more  
3  df.describe()  
4  
5  # unique: returns all the unique values in column  
6  df['ColumnName'].unique()  
7  
8  # value_counts: returns the frequency of the values  
9  df['ColumnName'].value_counts()  
10  
11 # correlation: find the correlation among  
12 # the features respectively  
13 df.corr()
```

