Python for Data Science

Introduction to Pandas

Python for Data Science

- Pandas is an open source library built on top of NumPy
- It allows for fast analysis and data cleaning and preparation
- It excels in performance and productivity.
- It also has built-in visualization features.
- It can work with data from a wide variety of sources.

Python for Data Science

- You'll need to install pandas by going to your command line or terminal and using either

```
conda install pandas
```

```
pip install pandas
```

Python for Data Science

- Series
- DataFrames
- Missing Data
- GroupBy
- Merging,Joining,and Concatenating
- Operations
- Data Input and Output