

Jupyter, Numpy, Pandas

Jupyter Notebooks & Markdown
Further Review of Python & Pandas

Matthew Morris

Git: Morrisdata

MatthewMorris.DA@gmail.com



Jupyter Notebooks & Mark Down



Visit KDnuggets.com for more cheat sheets and additional learning resources.

Markdown CheatSheet



Markdown is a lightweight markup language for creating formatted text using a plain-text editor.

Headings

Heading 1

Heading1

Heading2

Heading2

Heading 3

Heading3

Heading ID

Heading3 {#custom-id}

Text

Bold

KDnuggets
>>> KDnuggets

Italic

KDnuggets
>>> KDnuggets

Blockquote

> KD stands for Knowledge Discovery

>>> | [KD stands for Knowledge Discovery](#)

Strikethrough

→Data Science is Dying Career←

>>> Data Science is Dying Career←

Subscript

H-2-O

>>> H₂O

Superscript

a²*a + b²a

>>> a² * b²

Highlight

Highlight == Important words==

>>> Highlight **important words**

Footnote

Here's a sentence with a footnote. [^1]
[^1]: This is the footnote.

>>> Here's a sentence with a footnote. 1
>>> 1. This is the footnote. ^1

Link

[KDnuggets](https://www.kdnuggets.com/)

>>> [KDnuggets](#)

Lists

Ordered List

1. Data Science

2. Data Analytics
3. Machine Learning

Unordered List

- Data Science
- Data Analytics
- Machine Learning

Definition List

term
: definition

Task List

- [x] Prepare the draft
- [] Send for the review
- [] Publish

Code

Inline code

filter(function, sequence)

Fenced Code Block

```
ages = [20, 33, 44, 66, 78, 92]
filtered_ages = filter(lambda a: a > 50, ages)
print(list(filtered_ages))
```

Extended Syntax

Horizontal Rule

>>> _____

Image

!alt text(kdn-logo.jpg)



Table

Name	Grade
Abid	C
Nisha	B

Name	Grade
Abid	C
Nisha	B

Emoji

I am the best sunglasses-face

>>> I am the best 😎

Latex

\$\$V_{sphere} = 4\pi r^3/3\$\$

$$V_{sphere} = \frac{4}{3}\pi r^3$$

Subscribe to KDnuggets News

.00 .00 .00 .00
→.0 ←.0 →.0 ←.0

Abid Ali Awan | 2022

http://localhost:8889/notebooks/03_ipython_notebook_tutorial.ipynb#

Terms

Data Frame- Structures data into a two-dimensional table of Columns and rows.

Data Table – Columns, and rows that contain specific values



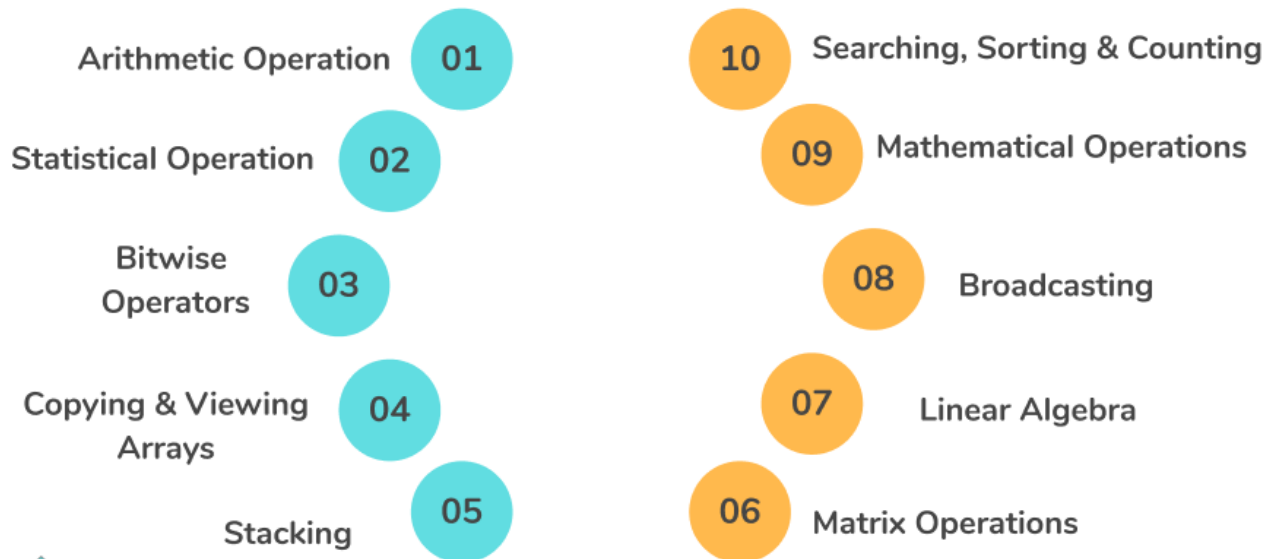
Terms

DATA FRAME	DATA TABLE
Used to build tables and matrices	Used for complex data structures and big data
20x slower than a data table	Faster even then spark in many cases
Can work with structured or semi structured data	Is often an output of a Data Frame and works with modeled structured data.
Does not need a table	Dependent on a Data Frame or prepped organization

Pandas and Numpy

PANDAS	NUMPY
When we have to work on Tabular data , we prefer the pandas module.	When we have to work on Numerical data , we prefer the numpy module.
The powerful tools of pandas are Data frame and Series .	Whereas the powerful tool of <i>numpy</i> is Arrays .
<i>Pandas</i> consume more memory .	<i>Numpy</i> is memory efficient .
<i>Pandas</i> have a better performance when the number of rows is 500K or more .	<i>Numpy</i> has a better performance when number of rows is 50K or less .
Indexing of the <i>pandas</i> series is very slow as compared to <i>numpy</i> arrays.	Indexing of <i>numpy</i> arrays is very fast .
<i>Pandas</i> have 2d table object called DataFrame .	<i>Numpy</i> is capable of providing multi-dimensional arrays .
It was developed by Wes McKinney and was released in 2008.	It was developed by Travis Oliphant and was released in 2005.
It is used in a lot of organizations like Kaidee, Trivago, Abeja Inc. , and a lot more.	It is being used in organizations like Walmart Tokopedia, Instacart, and many more.
It has a higher industry application.	It has a lower industry application.

Uses Of NumPy



Terms

Indexing numpy

https://www.pythonlikeyoumeanit.com/Module3_IntroducingNumpy/BasicIndexing.html

In your own words what is the difference numpy and pandas?

Mention the different types of Data Structures in Pandas

What are the significant features of the pandas Library?

Define Series in Pandas?

What are the different ways in which a series can be created?

How do you read Text Files with Pandas?

Tasks

Use Pandas to connect to Chipotle data
Use Chipotle data to rebuild Pandas lecture
Leverage outputs to support Chipotle Project

Matthew Morris

Git: Morrisdata

MatthewMorris.DA@gmail.com