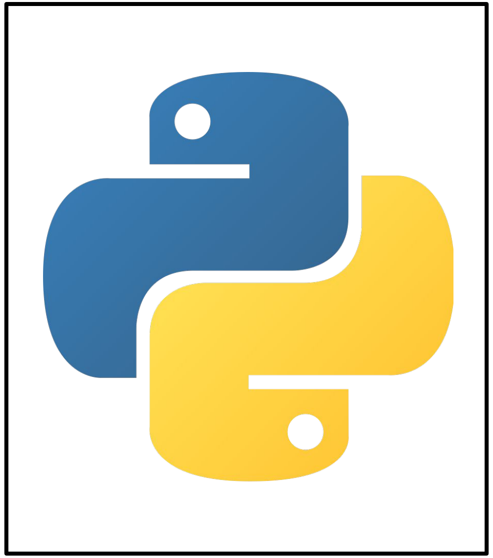
Polynote:

- first-class Scala
- Shared objects with python

https://polynote.org/

NETFLIX | OSS

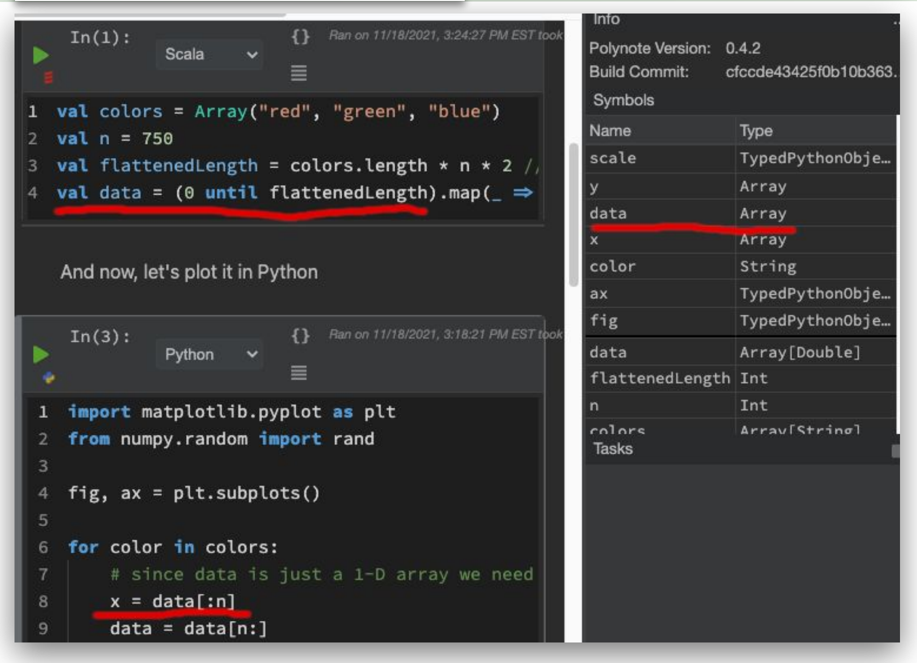
Scala



Python

Cool Stuff

Variables exist outside the language



In(1): Scala — Ran on 11/18/2021, 3:24:27 PM EST took

```scala
val colors = Array("red", "green", "blue")
val n = 750
val flattenedLength = colors.length * n * 2 //
val data = (0 until flattenedLength).map(_ ⇒
```

And now, let's plot it in Python

In(3): Python — Ran on 11/18/2021, 3:18:21 PM EST took

```python
import matplotlib.pyplot as plt
from numpy.random import rand

fig, ax = plt.subplots()

for color in colors:
    # since data is just a 1-D array we need
    x = data[:n]
    data = data[n:]
```

Info

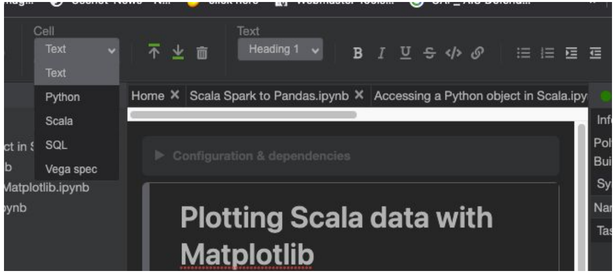
Polynote Version: 0.4.2

Build Commit: cfccde43425f0b10b363.

Symbols

Name	Type
scale	TypedPythonObje…
y	Array
data	Array
x	Array
color	String
ax	TypedPythonObje…
fig	TypedPythonObje…
data	Array[Double]
flattenedLength	Int
n	Int
colors	Array[String]

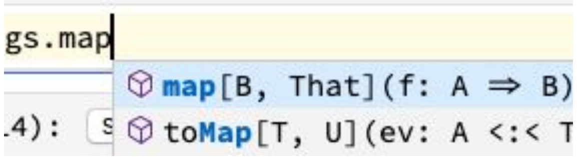
Tasks



Cell
Text
Text
Python
Scala
SQL
Vega spec
Heading 1
B
I
U
Home
Scala Spark to Pandas.ipynb
Accessing a Python object in Scala.ipy
Configuration & dependencies
Matplotlib.ipynb

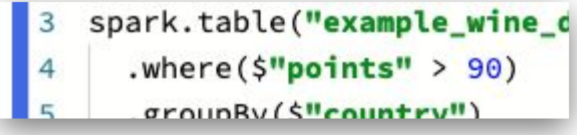
Plotting Scala data with Matplotlib

Nice Editing



```
gs.map
map[B, That](f: A ⇒ B)
toMap[T, U](ev: A <:< T
```

Shockingly good auto complete

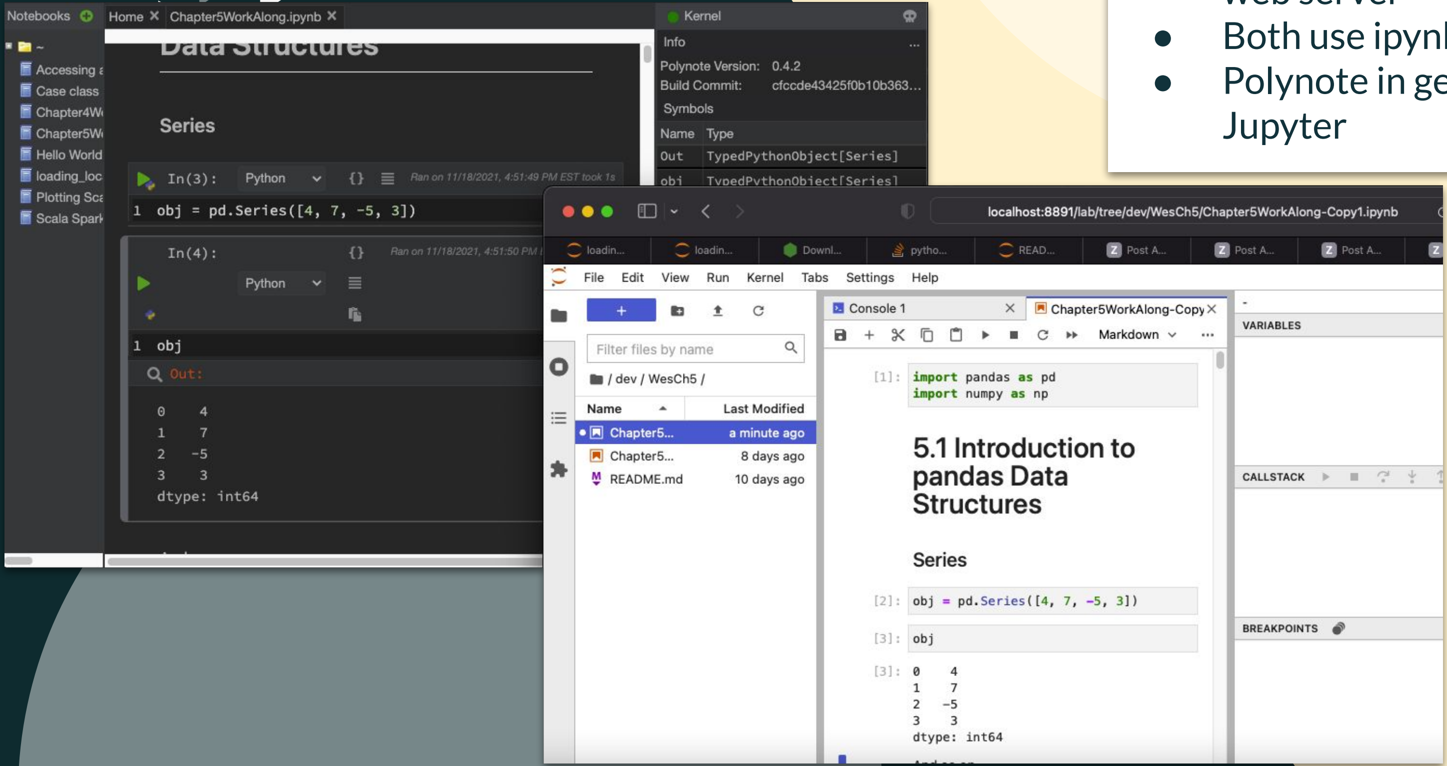


```scala
spark.table("example_wine_d
  .where($"points" > 90)
  .groupBy($"country")
```

Built for Spark

How are Polynote and Jupyter similar

- Both are OSS software (with permissive licenses)
- Both retain a cell/output paradigm
- Both work by setting up a simple web server
- Both use ipynb file formats
- Polynote in general feels a lot like Jupyter



Polynote Version: 0.4.2
Build Commit: cfccde43425f0b10b363...

Symbols

Name	Type
Out	TypedPythonObject[Series]
obj	TypedPythonObject[Series]

In(3): Python — Ran on 11/18/2021, 4:51:49 PM EST took 1s

```
obj = pd.Series([4, 7, -5, 3])
```

In(4): Python — Ran on 11/18/2021, 4:51:50 PM

```
obj
```

Out:

```
0    4
1    7
2   -5
3    3
dtype: int64
```

localhost:8891/lab/tree/dev/WesCh5/Chapter5WorkAlong-Copy1.ipynb

Name	Last Modified
Chapter5...	a minute ago
Chapter5...	8 days ago
README.md	10 days ago

```
[1]: import pandas as pd
     import numpy as np
```

5.1 Introduction to pandas Data Structures

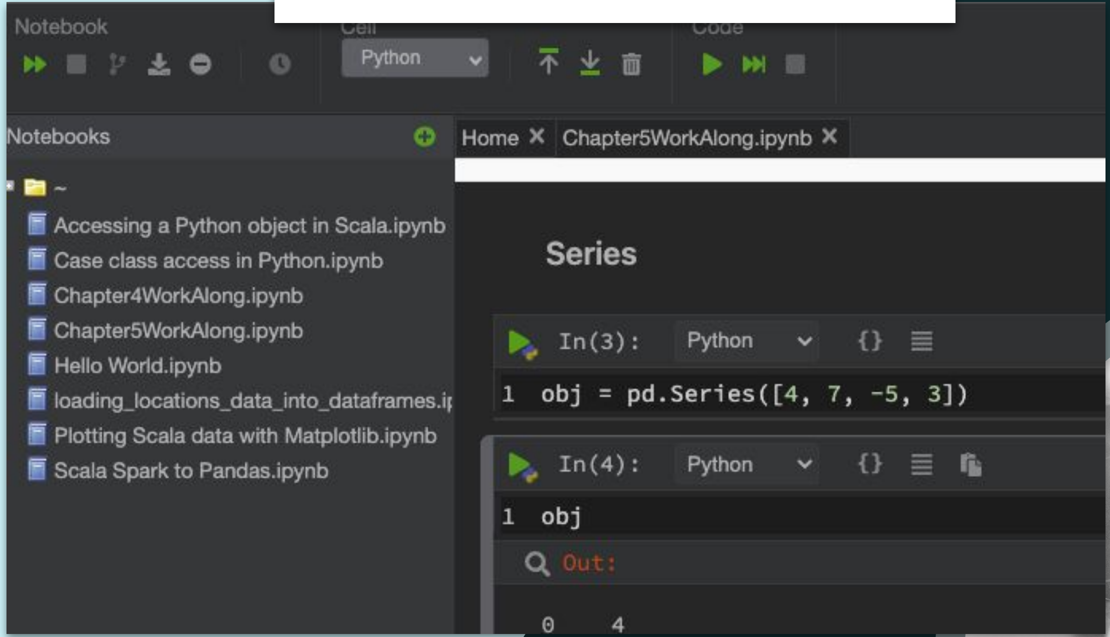
Series

```
[2]: obj = pd.Series([4, 7, -5, 3])

[3]: obj

[3]: 0    4
     1    7
     2   -5
     3    3
     dtype: int64
```

How is Polynote different than Jupyter?

The Good

- Very nice highlighting and code completion
- Supports Scala as a main language
- Supports Spark directly
- Share variables with Python
- Immutable principles (cannot initialize a variable in a cell below its first appearance)



Accessing a Python object in Scala.ipynb
Case class access in Python.ipynb
Chapter4WorkAlong.ipynb
Chapter5WorkAlong.ipynb
Hello World.ipynb
loading_locations_data_into_dataframes.i
Plotting Scala data with Matplotlib.ipynb
Scala Spark to Pandas.ipynb

Series

In(3): Python

```
obj = pd.Series([4, 7, -5, 3])
```

In(4): Python

```
obj
```

Out:

```
0    4
```

The Bad

- Can be very slow to load
- Noticeable latency in starting cell execution
- Bad file browser
- Some small glitches
- Lacks polish

The Ugly

- Some of the polish lack can feel extreme
- No easy installer (may be related to my Java version)
- Does not support newest version of Java LTS
- Very difficult install for Java 17

Why was Polynote created?

Polynote was created so there was a notebook that ran Scala nativelyto access Spark databases to get data into Python. With the ability to share variable and objects between the two, it makes Python and Scala good friends. Get data with Scala; analyze data in Python.

Why would one use Polynote?

Polynote would be a good choice for anyone familiar with Scala wanting to do data engineering from a Spark database.

It can natively run Scala to access the database and hand that data directly over to Python for analysis and graphing.

If one wants to share the notebook, the order enforcing can make sure a sloppy mess isn't being handed off (notebooks will always be able to run from top to bottom and give the same output).

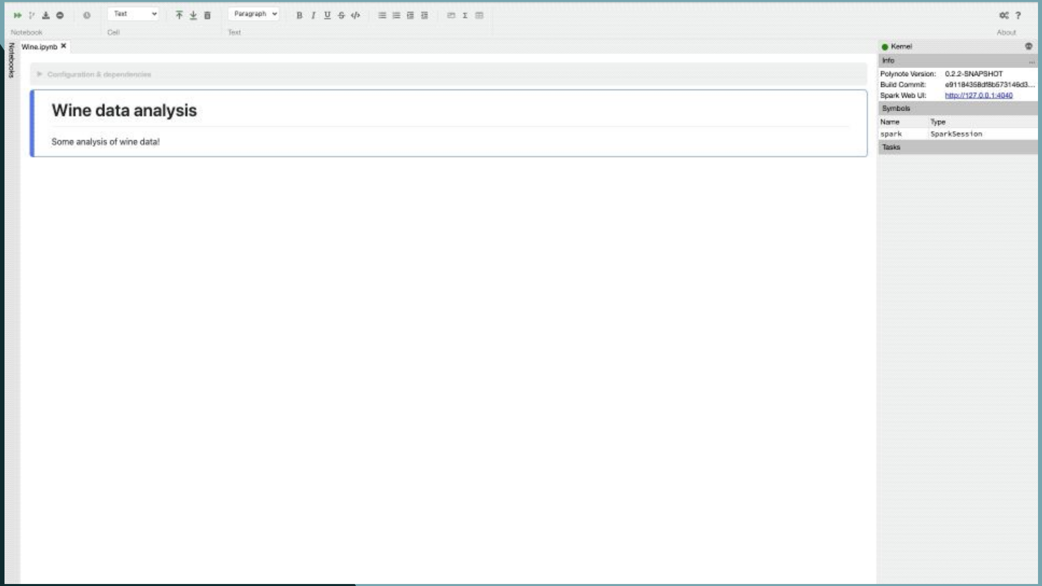
Will I use Polynote?

TL;DR I sure hope!

In its current state and with my current level of knowledge in Scala and Spark (zero) I do not find the lack of polish worth dealing with in exchange for some of the other perks.

I will definitely be trying the next version when it’s released if it is compatible with Java 17 though, and will also give it a shot when we are working with Spark (if I can get the Spark connection to work).

I believe it has potential to be better than Jupyter with some more maturity, and is quite likely better already for the specific needs it was meant to address.



Wine data analysis

Some analysis of wine data!

Polynote Version: 0.2.2-SNAPSHOT
Build Commit: e91184358df8b573146d3...
Spark Web UI: http://127.0.0.1:4040

Name	Type
spark	SparkSession