



Code Like a Snake Charmer

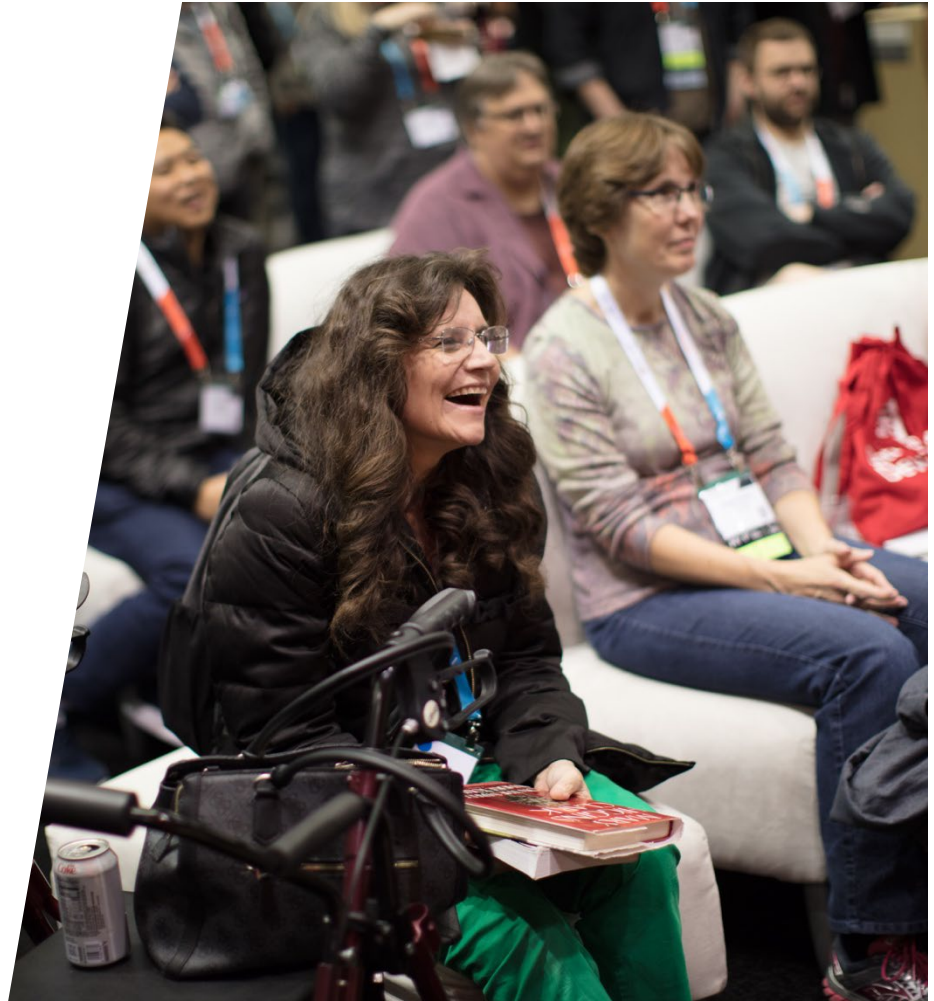
Advanced Data Modeling in Python!

Jamey Johnston, Sr. Data Scientist/Engineer





**Please silence
cell phones**



Explore everything PASS has to offer

Free Online Resources
Newsletters
PASS.org



24HOURS
of  PASS

Free online
webinar events



PASS
LOCAL
GROUPS

Local user groups
around the world



 PASS
SQLSATURDAY

Free 1-day local
training events



 PASS
MARATHON

Online special
interest user groups



PASS
VIRTUAL
GROUPS

Business analytics
training



PASS
VOLUNTEERS

Get involved



Jamey Johnston

Sr. Data Scientist/Engineer



/jameyj



@STATCowboy

Education

Texas A&M - MS in Analytics

LSU - BS in Spatial Analysis

Photographer

<http://jamey.photos>

Blog

<https://STATCowboy.com>

Code

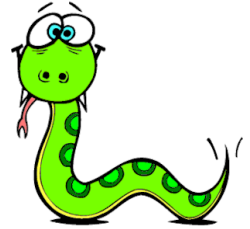
<https://github.com/STATCowboy/SnakeCharmer-Adv>

Agenda



1. Introduction to Python
2. Anaconda / IDEs
3. Data Wrangling and Visualizations
4. Functions & Classes/Objects
5. Data Science Modeling in Python
6. Flask & Docker Containers
7. Python and Microsoft

Introduction to Python



Why Python?

- Expansive Open Source Library of Data Science Tools (Giant Ecosystem)
- Easy language for new programmers
- Microsoft Support in tools like Azure Databricks, Azure Function Apps, Azure Machine Learning, SQL Server 2017+, Microsoft Machine Learning Server
- You can code on a Raspberry Pi (Who doesn't like Pi!)
- The most popular program languages (IEEE Language Rankings 2018 #1)
- Interpreted language, saves you time, no compilation and linking is necessary

Anaconda



Source: <http://www.anaconda.com>

Anaconda

<https://www.anaconda.com/download/>

Download the 64-bit Python 3.7 version (still can setup Python 2.7 environments)



You can use miniconda if you don't want the full environment.
(I have switched to it).

<https://docs.conda.io/en/latest/miniconda.html>

Anaconda

Conda Commands

- Export Conda Environment to YAML file to build a new environment
 - `conda env export > <filename>.yaml`
 - * activate the environment to export first
- Create Conda environment from YAML file
 - `conda env create -f <filename>.yaml -n <ENV NAME>`

<https://docs.conda.io/projects/conda/en/latest/user-guide/tasks/manage-environments.html>

Packages

Popular Packages

PACKAGE	DETAILS
pandas	High performance, easy use data structures and analysis (DataFrames)
pyodbc	Open Source Python Module for ODBC data sources
matplotlib	2D Plotting library
scikit-learn	Simple tool for data mining and data analysis / statistics
numpy	N-dimensional arrays, linear algebra, random numbers
SciPy	Math, Stats, Science and Engineering package
tensorflow	Google library to train and develop ML models
keras	High-level neural networks API. Runs on top of TensorFlow, CNTK and Theano

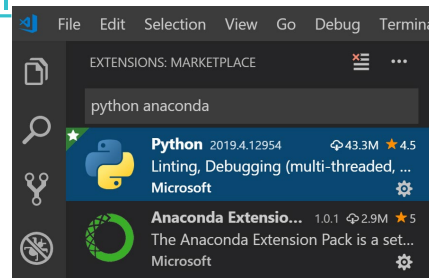
Visual Studio Code



VS Code Python Shortcuts

<https://code.visualstudio.com/docs/python/python-tutorial>

- Command Palette (CP) – `ctrl+Shift+P`
- Select Python Interpreter (in CP) – Python: Select Interpreter
- Run Selection/Line in Python Terminal – `shift+Enter`
- Install pylint for Highlighting Syntax – `conda install pylint` (run in all env)
- Install Python and Anaconda Extensions
- IPython console support in Python Interactive window



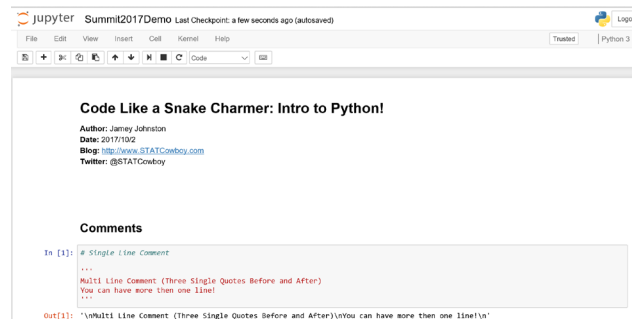
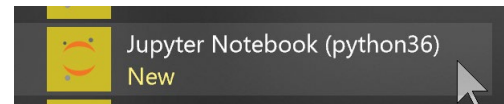
Jupyter Notebooks



Computer Code and Rich Text

<http://jupyter-notebook.readthedocs.io/en/latest/>

- Activate desired environment first
- Then to Start a Notebook – `jupyter notebook`



Pandas



Series and DataFrame

Labeled Array Data Structures

Input/output Tools (CSV, Excel, ODBC)

<http://pandas.pydata.org/pandas-docs/stable/10min.html>

NumPy



Homogeneous Multidimensional Array

N-dimensional arrays, linear algebra, random numbers

<https://numpy.org/devdocs/user/quickstart.html>

DEMO

Panda & NumPy: Data Wrangling & Visualizations



Functions

Simple Function

NOTE: non-default parameters must be first!

```
def greetSummit(year, name=None):  
    if name is not None:  
        print('Welcome to PASS Summit ', year, ', ', name, '!', sep='')  
    else:  
        print('Welcome to PASS Summit ', year, '!', sep='')
```

```
greetSummit(2019)
```

```
greetSummit(2019, 'Jamey')
```

Classes/Objects

Simple Class and Objects

```
class Database:
    def __init__(self, instance):
        self.instance = instance

    def mssql(self, version):
        print(self.instance + " is MSSQL version " + version + " and is Clustered!")

    def oracle(self):
        print(self.instance + " is Oracle and has crashed! Call Larry!")

def main():
    prod = Database("prod")
    prod.oracle()
    dev = Database("dev")
    dev.mssql("2019")

if __name__ == "__main__":
    main()
```


DEMO

Functions & Classes



DEMO

Data Science: Regression, Trees, XGB & Deep Learning



DEMO

Flask & Docker Containers



Python and Azure App Services

Web Apps on a Fully Managed Platform

<https://azure.microsoft.com/en-us/services/app-service/>



Python and Azure Functions

Serverless Compute Platform

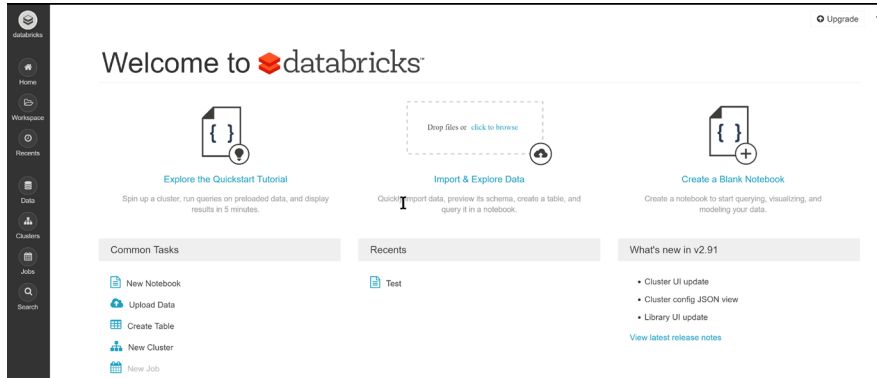


<https://azure.microsoft.com/en-us/services/functions/>

<https://azure.microsoft.com/en-us/blog/announcing-the-general-availability-of-python-support-in-azure-functions/>

Python and Azure Databricks

Workbooks



The screenshot shows the Databricks home page. At the top, it says "Welcome to databricks". Below this, there are three main action buttons: "Explore the Quickstart Tutorial", "Import & Explore Data", and "Create a Blank Notebook". Each button has a brief description of what it does. On the left side, there is a vertical sidebar with icons for Home, Workspace, Recents, Data, Clusters, Jobs, and Search. At the bottom, there are three sections: "Common Tasks" with links to "New Notebook", "Upload Data", "Create Table", "New Cluster", and "New Job"; "Recents" with a link to "Test"; and "What's new in v2.91" with a list of updates and a link to "View latest release notes".

Welcome to **databricks**

Explore the Quickstart Tutorial
Spin up a cluster, run queries on preloaded data, and display results in 5 minutes.

Import & Explore Data
Quickly import data, preview its schema, create a table, and query it in a notebook.

Create a Blank Notebook
Create a notebook to start querying, visualizing, and modeling your data.

Common Tasks

- New Notebook
- Upload Data
- Create Table
- New Cluster
- New Job

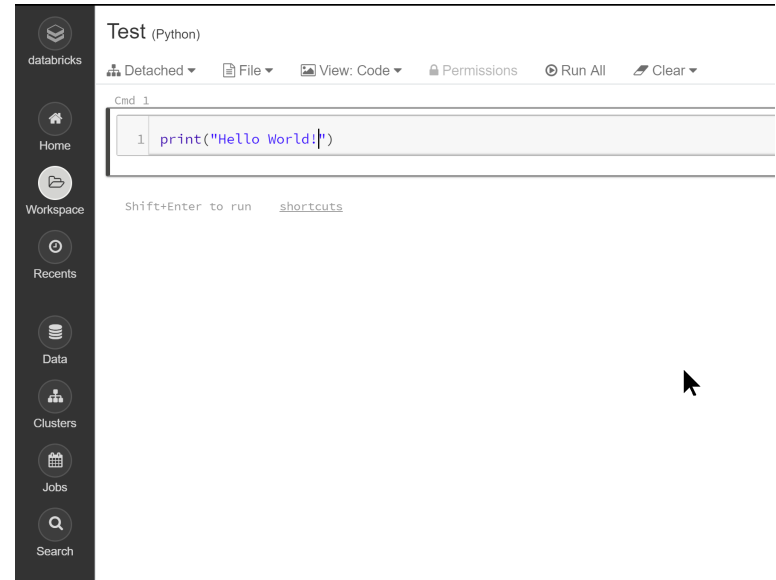
Recents

- Test

What's new in v2.91

- Cluster UI update
- Cluster config JSON view
- Library UI update

[View latest release notes](#)



The screenshot shows a Databricks notebook titled "Test (Python)". The notebook is in a "Detached" state. The code editor shows a single line of Python code: `print("Hello World!")`. Below the code editor, there is a status bar that says "Shift+Enter to run" and a link to "shortcuts". On the left side, there is a vertical sidebar with icons for Home, Workspace, Recents, Data, Clusters, Jobs, and Search.

Test (Python)

Detached File View: Code Permissions Run All Clear

Cmd 1

```
1 print("Hello World!")
```

Shift+Enter to run [shortcuts](#)

DEMO

MS & Python



References

Python Docs

<https://docs.python.org/3/reference/introduction.html>

Coursera

<https://www.coursera.org/specializations/python>

MS Academy

<https://academy.microsoft.com/en-us/professional-program/tracks/data-science/>

References

The Hitchhiker's Guide to Python!

<http://docs.python-guide.org/en/latest/>

Code Academy

<https://www.codecademy.com/catalog/language/python>

Google

<https://developers.google.com/edu/python/?hl=en>

Session Evaluations

Submit by 5pm Friday,
November 15th to
win prizes.

3 WAYS TO ACCESS



Go to PASSsummit.com



Download the GuideBook App
and search: PASS Summit 2019



Follow the QR code link on session
signage



Thank You

Jamey Johnston



@STATCowboy



jj@jameyj.com