Notebooks are still cool...with Jupyter

Ryan Bales (@ryanbales)
ryan@balesofdata.com

Bio

- Ryan Bales
- Live, work, and play in Cleveland (sometimes Chicago)
- Director of Analytics at DialogTech
- Python and Data Engineering Instructor at DriveIT

Purpose

You don't always need a desktop-based IDE installed locally to build software, analyze data, and solve complex problems.

In this talk let's dig deeper into:

- What is JupyterLab?
- How can I get it?
- What can I do with JupyterLab?
- JupyterLab in the cloud

What is Project Jupyter?

- Nonprofit focused on open and interactive computing
  - Open-source Software (modified BSD License)
  - Open-standards across many programming languages
- Born out of the IPython project in 2014
- Focused on the Data Science and Research communities

- 3.3 Million **PUBLIC** Jupyter notebooks on Github
  - https://github.com/parente/nbestimate/blob/master/estimate.ipynb

What's in a name?

- Jupyter Notebooks
  - **JU**lia
  - **PYT**hon
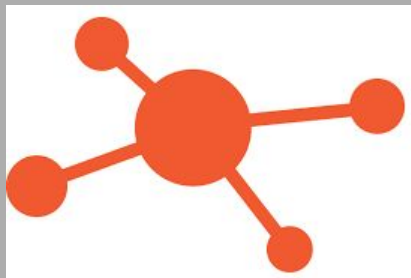 - **R**
- Homage to Galileo's notebooks
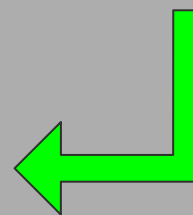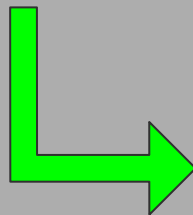
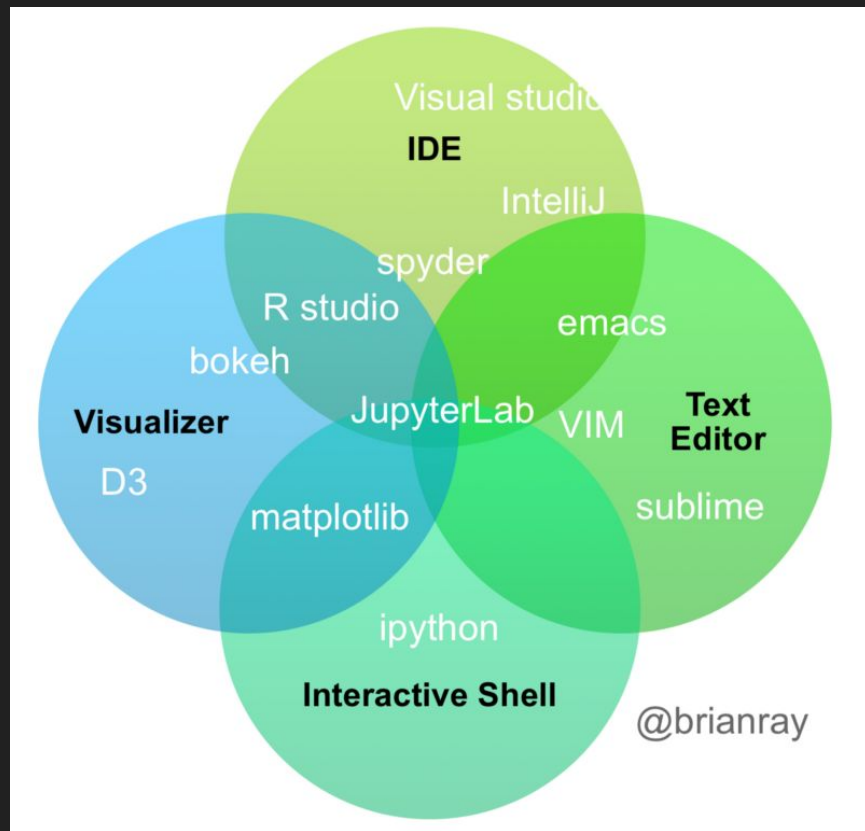# Jupyter Applications

Notebooks

LAB

HUB

Why should I use Jupyter Notebooks?

- Tell the story of your data
  - Code, Visualizations, and Markdown
- Data State
  - Load Data into your Notebook once
  - Run numerous experiments without reloading
- Share with Others
  - Commit Executed State to Github
  - Github renders Notebooks in your browser

What is JupyterLab?

- JupyterLab is the Next Generation of Jupyter Notebooks
- Extension System
  - Core Functionality is a set of extensions
- Added text editors, terminal access, data file review.
- Multi-Panel Support
- Drag and Drop / Copy Paste of Cells
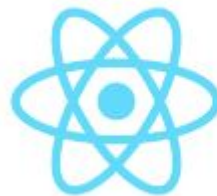
What does JupyterLab do?

JupyterLab History

- 2011 - First IPython Notebooks ships
- Feb 09, 2017 -- v0.16.0 (First documented release)
- Jan 11, 2018 -- v0.31.0 (Initial Beta release)
- Jul 26, 2018   -- v0.33.0 (Beta tag removed)
- Oct 03, 2018 -- v0.35.0 (Latest Release)
- ???   -- v1.0.0 (First General Release)
  - Due Dec 21, 2018
  - 88 Open Issues

JupyterLab Tech Stack

Ok, How do I get it?

- Install the Latest Version of Python
  - Homebrew (macOS)
  - Download from Python.org (Windows)
- Install JupyterLab
  - Anaconda
    - conda install -c conda-forge jupyterlab
  - PIP
    - *python3 -m pip install jupyter*lab
- In a terminal, execute *(jupyter lab)*
- Default browser will open to http://localhost:8888/lab

A Lap around JupyterLab

- Getting Started
  - Cell Types
  - Drag and Drop
  - Variable State
  - Visualization
- Autocomplete
- Debugging
- Magic Commands
- Panel/Tab Management

Does JupyterLab speak my language?

Can I use R?

- Yes!
- R Kernel Setup
  - Install R
  - From an R Command Prompt run
    - *install.packages(c('repr', 'IRdisplay', 'evaluate', 'crayon', 'pbdZMQ', 'devtools', 'uuid', 'digest'))*
    - *devtools::install_github('IRkernel/IRkernel')*
      - ***Not available on CRAN***
    - IRkernel::installspec()
- Demo

To the Cloud...

- Typical use cases for cloud scale for Jupyter
  - More CPU/RAM and/or a GPU
  - Unable to install software on your machine
  - Long running scripts/notebooks
- Let's look at a couple of options
  - Cloud-based Hardware
  - Hosted Services

Cloud-based Hardware

- Lots of Options - AWS, Google, Azure IBM
- Example Scenario:
  - AWS Elastic Cloud (EC2)
  - P2.xlarge (4 core CPU, 12 GB RAM, Nvidia Tesla K-80)
    - AWS Deep Learning AMI
    - $0.90/hour
  - Create AWS Resources via Terraform
  - Connect using SSH Port Forwarding

Hosted Jupyter Services

- Don't need full control over the hardware
- AWS SageMaker
    - Some control over the EC2 Instance Type/Size
    - Load data from S3
    - ml.P2.xlarge (4 core CPU, 12 GB RAM, Nvidia Tesla K-80)
    - $1.26/hour (~ 40% more than managing your own instance)
- Google Colaboratory
    - Minimal control over hardware (GPU/TPU settings available)
    - Load data from Google Drive, Google Sheets, Google Cloud bucket
    - Free?

Summary

- Ready for you and members of your team!
- Kernels available for 100+ languages
- Cloud-based and hosted solutions available
- Go build some cool notebooks and share!

Thank you

- Github: https://github.com/ryanbales/Jupyter-Notebooks-Overview
- Contact Information:
  - Ryan Bales (ryan@balesofdata.com)
    - https://twitter.com/ryanbales