

DSBA 5122: Visual Analytics

Class 14: Python

Ryan Wesslen

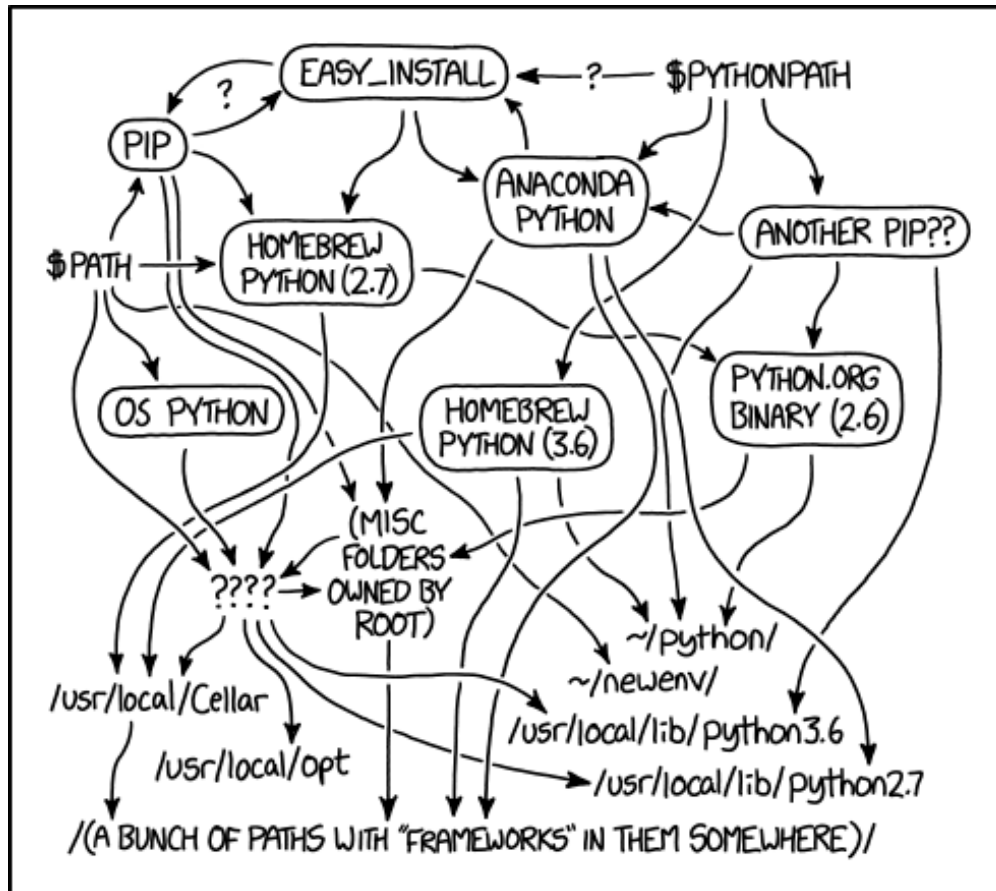
December 2, 2019

Python+R (reticulate)

Reticulate is an R package for connecting to an existing python distribution (e.g., Anaconda, virtual environment).

```
29
30 ▾ ```{r}
31 library(tidyverse)
32 flights <- read_csv("flights.csv") %>%
33   filter(dest == "ORD") %>%
34   select(carrier, dep_delay, arr_delay) %>%
35   na.omit()
36 ```
37
38 ▾ ```{python}
39 print(r.flights.head())
40 ```
41
```

Python Configurations



MY PYTHON ENVIRONMENT HAS BECOME SO DEGRADED
THAT MY LAPTOP HAS BEEN DECLARED A SUPERFUND SITE.

1. Run RStudio's ML Docker Image in one of these ways:

1. Easiest option (free, no local install): [play-with-docker](#)
2. Will take time to install (free): [local Docker](#)
3. Will cost money: [Digital Ocean Droplet](#)

This docker image also includes Machine Learning packages (for CPU) like [h2o](#), [tensorflow](#), and [keras](#). It can expand to GPU environments too!

Play-with-docker.com

run in bash command/terminal

```
docker run -e PASSWORD=[create_a_pwd] -p 8787:8787 rocker/ml
```

03:58:40

CLOSE SESSION

Instances  

+ ADD NEW INSTANCE

192.168.0.8
node1

bj34br7d_bj34bt7du41g00a89pm0

IP

192.168.0.8

Memory

0.94% (37.77MiB / 3.906GiB)

CPU

0.80%

SSH

ssh ip172-18-0-12-bj34br7du41g00a89plg@direct.labs.play-with-d 

DELETE

 EDITOR

```
#####
#                               #
#           WARNING!!!!        #
# This is a sandbox environment. Using personal credentials #
# is HIGHLY! discouraged. Any consequences of doing so are #
# completely the user's responsibilites.                     #
#                                                             #
# The PWD team.                                              #
#####
[node1] (local) root@192.168.0.8 ~
$ docker run -e PASSWORD=[create_a_pwd] -p 8787:8787 rocker/ml
```

3. Open RStudio via port 8787

You can do this by clicking the "Open Port" button and then typing in 8787.

FYI: username: rstudio, pwd: whatever you chose

4. Importing python packages

```
# run in RStudio console/file  
# fyi reticulate is prepackaged  
library(reticulate)  
os <- import("os")  
os$listdir(".")
```

5. Installing python packages

- Run this in RStudio *terminal*:

```
#run in rstudio terminal  
easy_install pip  
easy_install virtualenv
```

- Restart RStudio session (Session > Restart R)
- Run this in RStudio *console*:

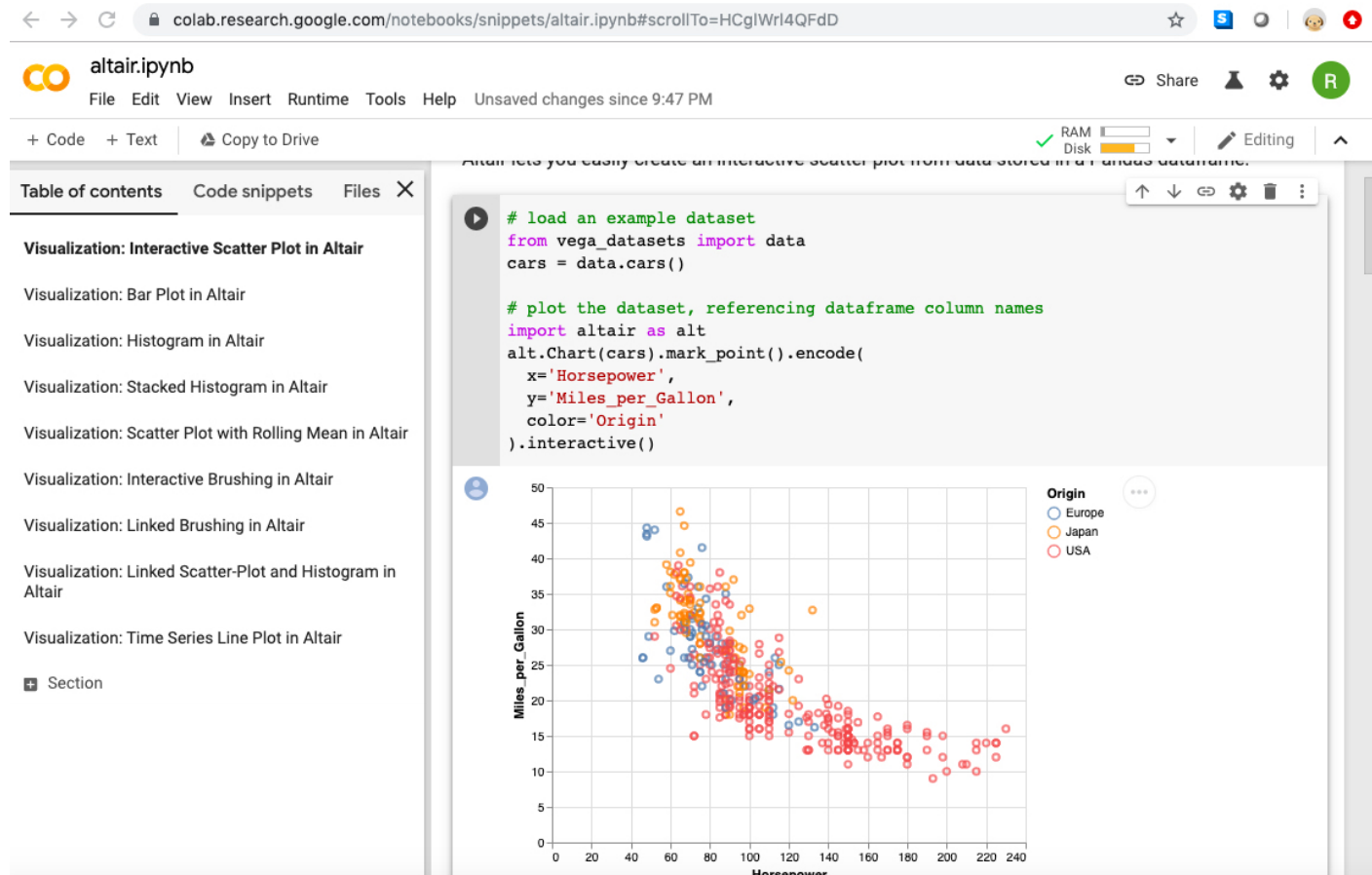
```
#run in rstudio console  
reticulate::py_install("pandas")
```

- Run this to install a demo MNIST model:

```
#run in rstudio console  
download.file("https://gist.github.com/wesslen/aec5666a3a29238dd651eccf1425ae20",  
             destfile = "mnist-keras.R")
```


Colab Notebooks

<https://colab.research.google.com/>



Think of it as RStudio.Cloud for Python. Even has GPU's!

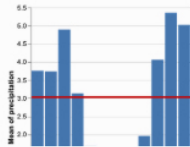
Altair: An Interactive Grammar of Graphics

Altair is a python package for Vega/Vega-Lite, which are an interactive grammar of graphics. What is that?

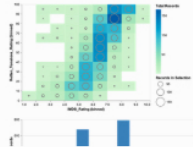
Jake VanderPlas' Altair Tutorial at PyCon 2018

<https://altair-viz.github.io/gallery/index.html>

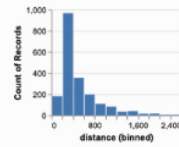
Interactive Charts



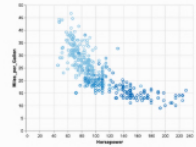
Interactive Average



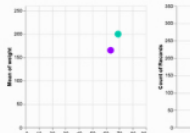
Interactive Chart
with Cross-Highlight



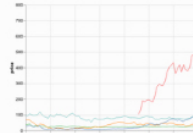
Interactive Crossfilter



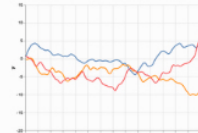
Interactive
Rectangular Brush



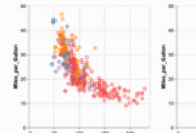
Interactive Scatter
Plot and Linked
Layered Histogram



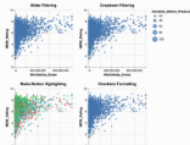
Multi-Line Highlight



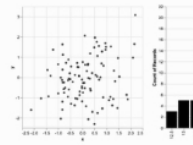
Multi-Line Tooltip



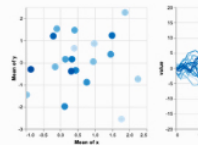
Multi-panel Scatter
Plot with Linked
Brushing



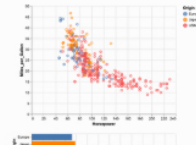
Multiple Interactions



Scatter Plot and
Histogram with
Interval Selection



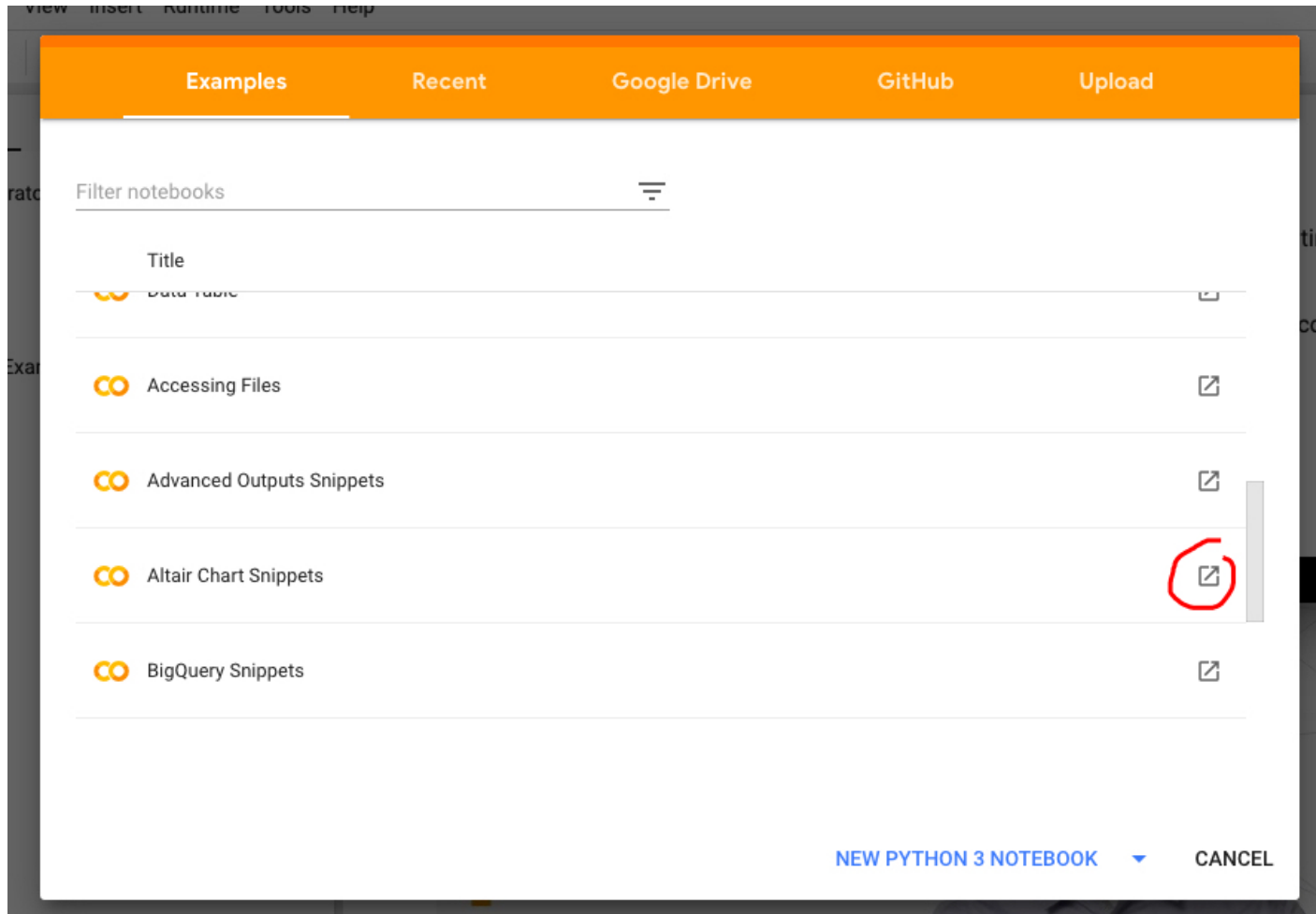
Selection Detail
Example



Selection Histogram

Altair Gallery

How to run Altair via a Colab Notebook



Run Multiple Tutorials

- Go to "GitHub" tab. In the URL, paste "<https://github.com/altair-viz/altair-tutorial.git>" then press Return (or Enter)
- Change Repository to "altair-viz/altair_tutorials"
- Open Notebooks

Can change to other Repositories like "altair-viz/altair_notebooks"

Links

1. Docker training
2. Altair tutorial jupyter notebooks and video
3. Yihui's thoughts on R-Python notebooks
4. Python **reticulate** configuration
5. RStudio **reticulate** cheatsheet