**➜ ds\_hk\_1 git:(gh-pages) ✗** git pull

Already up-to-date.

TO GET IDENTITY MATRIX In [**1**]: from numpy import eye

Dot(x,eye(22)

Dot(X.T, y) (get

**➜ ds\_hk\_1 git:(gh-pages) ✗** ipython

Python 2.7.5 |Anaconda 1.8.0 (x86\_64)| (default, Oct 24 2013, 07:02:20)

Type "copyright", "credits" or "license" for more information.

IPython 1.1.0 -- An enhanced Interactive Python.

? -> Introduction and overview of IPython's features.

%quickref -> Quick reference.

help -> Python's own help system.

object? -> Details about 'object', use 'object??' for extra details.

In [**3**]: from numpy import array, dot

In [**4**]: from scipy.linalg import inv

In [**5**]:

In [**5**]: X = array([ [1, 1], [1, 2], [1, 3], [1, 4] ])

In [**6**]: y = array([ [1], [2], [3], [4] ])

In [**7**]: k = dot(X.T, y)

In [**8**]: n = inv(dot(X.T, X))

In [**9**]: coef\_ = dot(n, k)

In [**10**]:

Last login: Wed Feb 12 20:08:54 on ttys000

**➜ ~** cd desktop

**➜ desktop** cd ds\_hk\_1

**➜ ds\_hk\_1 git:(gh-pages) ✗** ls

DS\_HK\_1.sublime-project **finals**

Lesson 2 notes.docx lesson4.docx

Project1.sublime-workspace **lessons**

SELECT CategoryName.docx readme.md

cs229-linalg.pdf ronald

**data**

**➜ ds\_hk\_1 git:(gh-pages) ✗** cd data

**➜ data git:(gh-pages) ✗** curl https://gist.github.com/podopie/47ad0637ca6a3890eea9/raw/ce7027bc987b7a5754b1c6e85ba97f5397caa78d/nyagg.csv\# cleaned up aggregation nytimes data > mammals.csv

% Total % Received % Xferd Average Speed Time Time Time Current

Dload Upload Total Spent Left Speed

100 10102 0 10102 0 0 5600 0 --:--:-- 0:00:01 --:--:-- 5602

curl: (6) Could not resolve host: cleaned

0 0 0 0 0 0 0 0 --:--:-- --:--:-- --:--:-- 0curl: (6) Could not resolve host: up

curl: (6) Could not resolve host: aggregation

curl: (6) Could not resolve host: nytimes

curl: (6) Could not resolve host: data

**➜ data git:(gh-pages) ✗**

Last login: Wed Feb 12 19:43:09 on ttys001

**➜ ~** from numpy import eye

from: can't read /var/mail/numpy

**➜ ~** ipythn

zsh: command not found: ipythn

**➜ ~** ipython

Python 2.7.5 |Anaconda 1.8.0 (x86\_64)| (default, Oct 24 2013, 07:02:20)

Type "copyright", "credits" or "license" for more information.

IPython 1.1.0 -- An enhanced Interactive Python.

? -> Introduction and overview of IPython's features.

%quickref -> Quick reference.

help -> Python's own help system.

object? -> Details about 'object', use 'object??' for extra details.

In [**1**]: from numpy import eye

In [**2**]: X = array([ [1, 1], [1, 2], [1, 3], [1, 4] ])

---------------------------------------------------------------------------

NameError Traceback (most recent call last)

<ipython-input-2-9ce82b8c3e50> in <module>()

----> 1 X = array([ [1, 1], [1, 2], [1, 3], [1, 4] ])

NameError: name 'array' is not defined

In [**3**]: from numpy import array, dot

In [**4**]: from scipy.linalg import inv

In [**5**]:

In [**5**]: X = array([ [1, 1], [1, 2], [1, 3], [1, 4] ])

In [**6**]: y = array([ [1], [2], [3], [4] ])

In [**7**]: k = dot(X.T, y)

In [**8**]: n = inv(dot(X.T, X))

In [**9**]: coef\_ = dot(n, k)

In [**10**]: k x

File "<ipython-input-10-1adf247b1cf6>", line 1

k x

^

SyntaxError: invalid syntax

In [**11**]: k X

File "<ipython-input-11-49fab88ef8ef>", line 1

k X

^

SyntaxError: invalid syntax

In [**12**]: curl https://gist.github.com/tijptjik/2f6307ee4d192ce8927b

File "<ipython-input-12-f1efb1a7343a>", line 1

curl https://gist.github.com/tijptjik/2f6307ee4d192ce8927b

^

SyntaxError: invalid syntax

In [**13**]: import pandas as pd

In [**14**]:

In [**14**]: import matplotlib.pyplot as plt

In [**15**]: n

Out[**15**]:

array([[ 1.5, -0.5],

[-0.5, 0.2]])

In [**16**]: k

Out[**16**]:

array([[10],

[30]])

In [**17**]:

In [**8**]: cat mammals.csv

"Name","body","brain"

"Arctic fox",3.385,44.5

"Owl monkey",0.48,15.5

"Mountain beaver",1.35,8.1

"Cow",465,423

"Grey wolf",36.33,119.5

"Goat",27.66,115

"Roe deer",14.83,98.2

"Guinea pig",1.04,5.5

"Verbet",4.19,58

"Chinchilla",0.425,6.4

"Ground squirrel",0.101,4

"Arctic ground squirrel",0.92,5.7

"African giant pouched rat",1,6.6

"Lesser short-tailed shrew",0.005,0.14

"Star-nosed mole",0.06,1

"Nine-banded armadillo",3.5,10.8

"Tree hyrax",2,12.3

"N.A. opossum",1.7,6.3

"Asian elephant",2547,4603

"Big brown bat",0.023,0.3

"Donkey",187.1,419

"Horse",521,655

"European hedgehog",0.785,3.5

"Patas monkey",10,115

"Cat",3.3,25.6

"Galago",0.2,5

"Genet",1.41,17.5

"Giraffe",529,680

"Gorilla",207,406

"Grey seal",85,325

"Rock hyrax-a",0.75,12.3

"Human",62,1320

"African elephant",6654,5712

"Water opossum",3.5,3.9

"Rhesus monkey",6.8,179

"Kangaroo",35,56

"Yellow-bellied marmot",4.05,17

"Golden hamster",0.12,1

"Mouse",0.023,0.4

"Little brown bat",0.01,0.25

"Slow loris",1.4,12.5

"Okapi",250,490

"Rabbit",2.5,12.1

"Sheep",55.5,175

"Jaguar",100,157

"Chimpanzee",52.16,440

"Baboon",10.55,179.5

"Desert hedgehog",0.55,2.4

"Giant armadillo",60,81

"Rock hyrax-b",3.6,21

"Raccoon",4.288,39.2

"Rat",0.28,1.9

"E. American mole",0.075,1.2

"Mole rat",0.122,3

"Musk shrew",0.048,0.33

"Pig",192,180

"Echidna",3,25

"Brazilian tapir",160,169

"Tenrec",0.9,2.6

"Phalanger",1.62,11.4

"Tree shrew",0.104,2.5

"Red fox",4.235,50.4

In [**9**]: mammals = pd.read\_csv('mammals.csv')

In [**10**]: plt.scatter(mammals['body'], mammals['brain'])

Out[**10**]: <matplotlib.collections.PathCollection at 0x106b53b10>

In [**11**]: plt.show()

! allows you to run in a systems shell in python

In [**33**]: !curl https://gist.github.com/podopie/47ad0637ca6a3890eea9/raw/ce7027bc987b7a5754b1c6e85ba97f5397caa78d/nyagg.csv > nytimes.csv

In [**50**]: ls

**class**/ mammals.csv nytimes.csv

In [**51**]: from numpy import exp

In [**52**]: mammals['exp\_body']=exp(mammals['body'])

In [**53**]: mammals[

....:

....: ^C

KeyboardInterrupt

In [**53**]: mammals['exp\_brain']=exp(mammals['brain'])

In [**54**]: plt.scatter(mammals['exp\_body'], mammals['exp\_brain'])

Out[**54**]: <matplotlib.collections.PathCollection at 0x1086a2c10>

In [**55**]: plt.show()

Plt.hist(nytimes[‘age’], bins=range(0,1,10000,1000)

plt**.**hist(mammals['body'], bins**=**range(0, 10000, 100))

Classwork 1

Save mammals as a dataframe

Help(mammals)

Log\_reg = linear\_model.LinearRegression(1)

Fit the data (same as in class)

Log\_body = [[x]for x in mammals [‘log\_body’].values]

Log\_regr.fit(log\_body,log\_brain)

Log\_brain = mammals [‘log\_brain’].values

Exp(log\_bldy), exp(log\_regr.predict(log\_body)), color = ‘red’)