WIRTSCHAFTS
UNIVERSITÄT
WIEN VIENNA
UNIVERSITY OF
ECONOMICS
AND BUSINESS

EFMD
EQUIS
ACCREDITED

Principles of Software Programming: Input/Output

Svitlana Vakulenko, MSc.

WS 2017

Acknowledgement: slides adopted from Dr. Jürgen Umbrich GZP SS 2014 lecture slides

This Episode

- 13:00-15:45
- Files: reading, writing
- Counter
- Requests
- Map Reduce

Files Input/Output

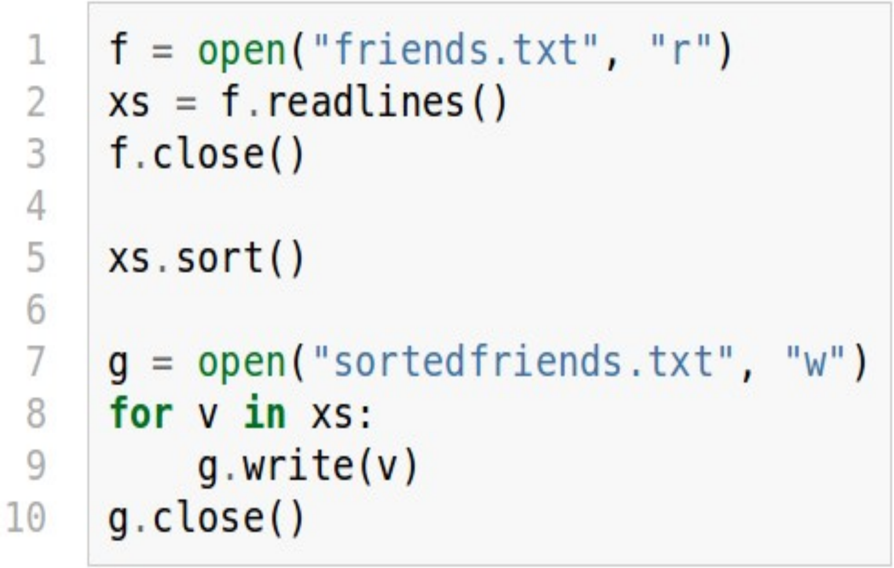
```
f = open('myTextFile.txt', "r")
lines = f.readlines()
f.close()

outFile = open('sample.txt', 'w')
outFile.write('My first output file!')
outFile.close()

with open("newfile.txt") as f:
    for line in f:
        print line
```

http://anh.cs.luc.edu/python/hands-on/3.1/handsonHtml/files.html

http://cmdlinetips.com/2011/08/three-ways-to-read-a-text-file-line-by-line-in-python/

http://www.pythonforbeginners.com/files/reading-and-writing-files-in-python

Files I/O Example



```python
f = open("friends.txt", "r")
xs = f.readlines()
f.close()

xs.sort()

g = open("sortedfriends.txt", "w")
for v in xs:
    g.write(v)
g.close()
```

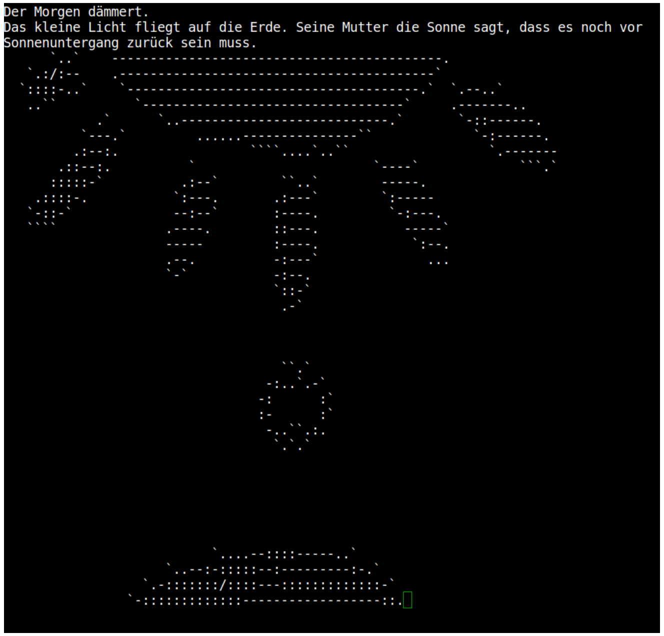
http://openbookproject.net/thinkcs/python/english3e/files.html

Ex.1: Phonebook

- **load** phonebook from file
- **number** of entries in the phonebook?
- **search** phone number by name
- **insert** phone number
- **save** phonebook into a file

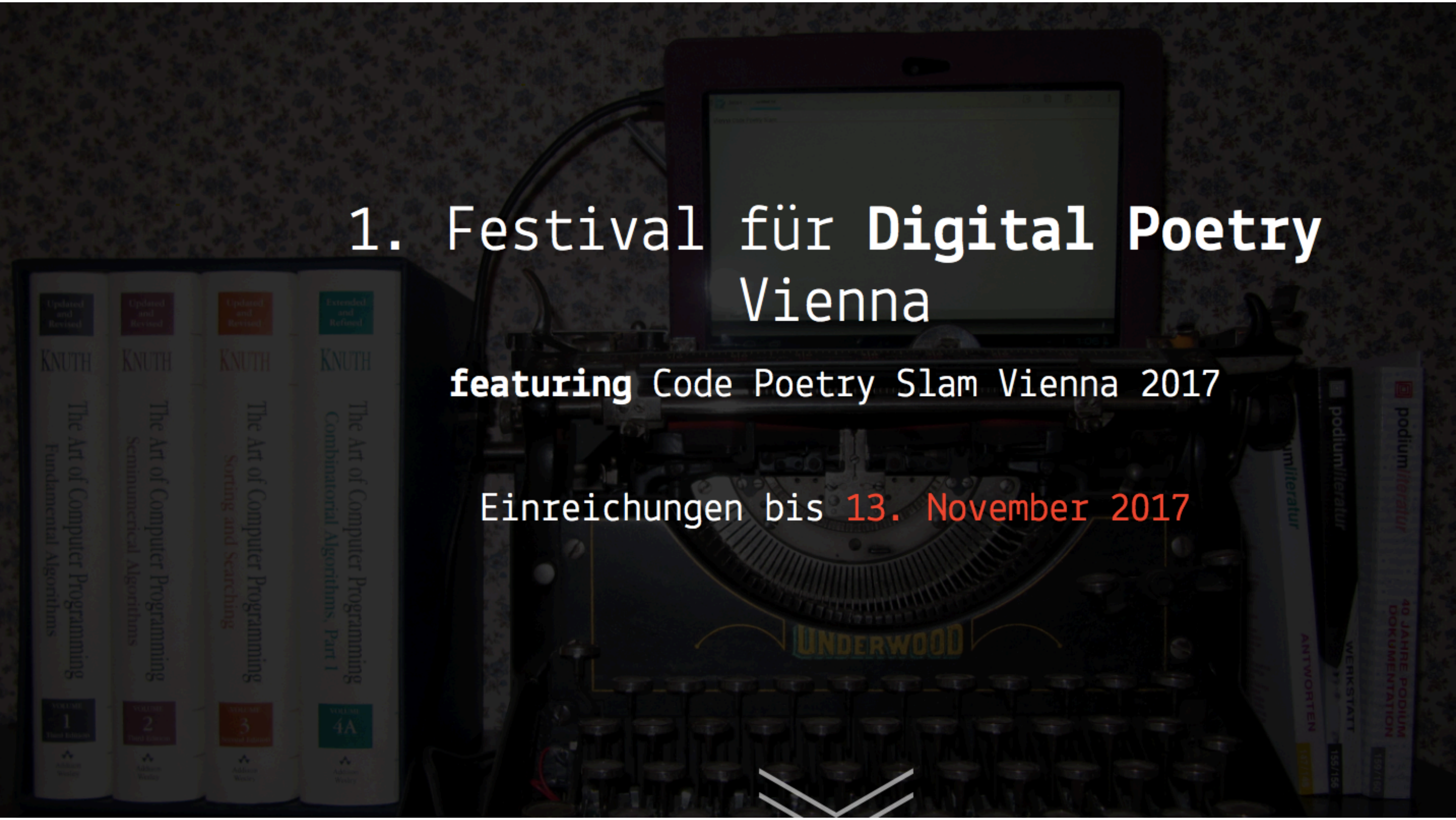
DasKleineLicht

- by Martin Szalay
- Code Poetry Slam 2016
- https://github.com/mszkb/DasKleineLicht



```
Der Morgen dämmert.
Das kleine Licht fliegt auf die Erde. Seine Mutter die Sonne sagt, dass es noch vor
Sonnenuntergang zurück sein muss.
```

codepoetry.at



1. Festival für **Digital Poetry** Vienna

featuring Code Poetry Slam Vienna 2017

Einreichungen bis 13. November 2017

http://code-poems.com

Ex.2: Alice in Wonderland

http://www.lewiscarroll.org/2012/05/29/new-alice-in-wonderland-poster-from-prospero-art/

http://www.gutenberg.org/files/11/11-0.txt

Requests

```
import requests

r = requests.get(‘https://api.github.com/events')

print(r.text)
```

http://docs.python-requests.org/en/master/user/quickstart/

Counter

- Find the ten most common words

```
from collections import Counter

>>> cnt = Counter()
>>> for word in ['red', 'blue', 'red', 'green', 'blue', 'blue']:
...     cnt[word] += 1
>>> cnt
Counter({'blue': 3, 'red': 2, 'green': 1})


>>> c = Counter(['eggs', ‘ham'])
>>> c[‘bacon’]


>>> Counter('abracadabra').most_common(3)
[('a', 5), ('r', 2), ('b', 2)]
```

https://docs.python.org/2/library/collections.html

Regex

```
>>> import re
```

```
>>> phonePattern = re.compile(r'^(\d{3})-(\d{3})-(\d{4})$')
>>> phonePattern.search('800-555-1212').groups()
('800', '555', ‘1212')
```

```
>>> re.sub(r'\sAND\s', ' & ', 'Baked Beans And Spam',
flags=re.IGNORECASE)
'Baked Beans & Spam’
```

```
>>> re.split('\W+', 'Words, words, words.')
['Words', 'words', 'words', '']
```

http://www.diveintopython.net/regular_expressions/phone_numbers.html

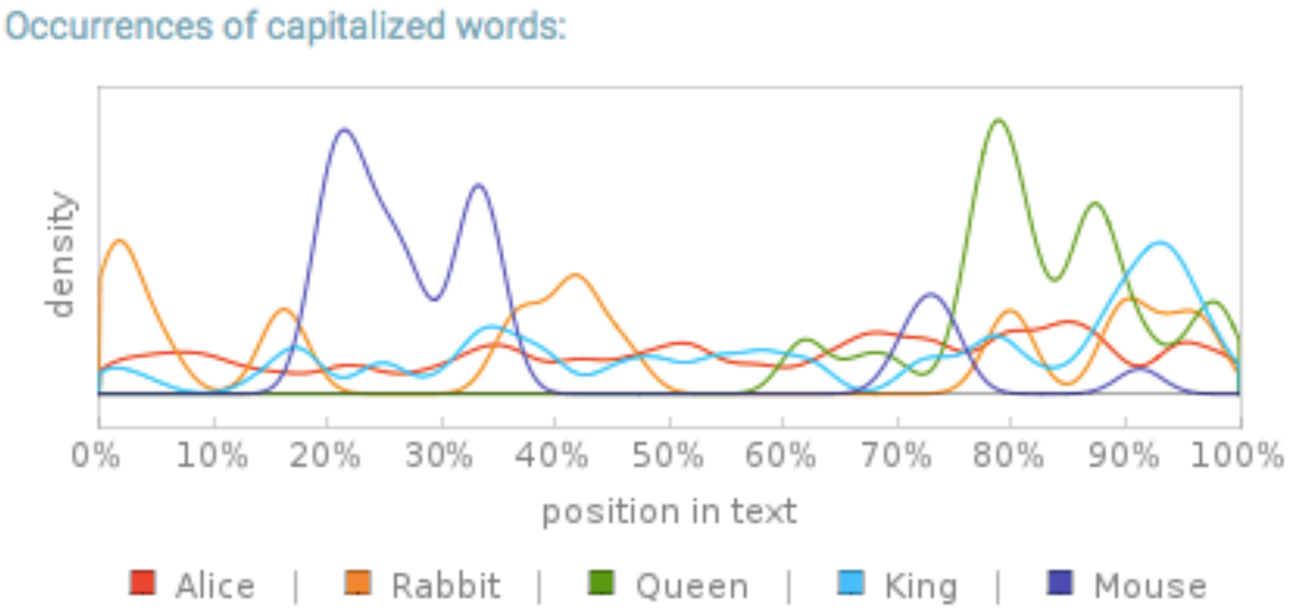
https://docs.python.org/2/library/re.html

Count words

```python
 # Download the content
content = urllib.urlopen('http://bit.ly/thewonderfulwizard').read()

# Clean the content a little
content = re.sub('\s+', ' ', content)  # condense all whitespace
content = re.sub('[^A-Za-z ]+', '', content)  # remove non-alpha chars
words = content.split()

# Start counting
word_count = Counter(words)

# The Top-N words
print("The Top {0} words".format(n))
for word, count in word_count.most_common(n):
    print("{0}: {1}".format(word, count))
```

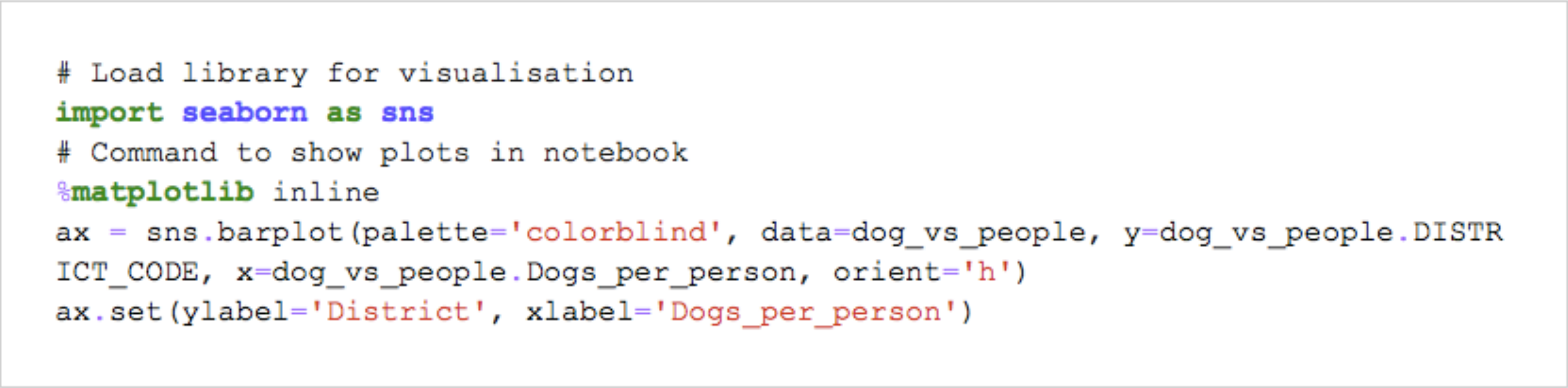
https://gist.github.com/bradmontgomery/4717521

WolframAlpha

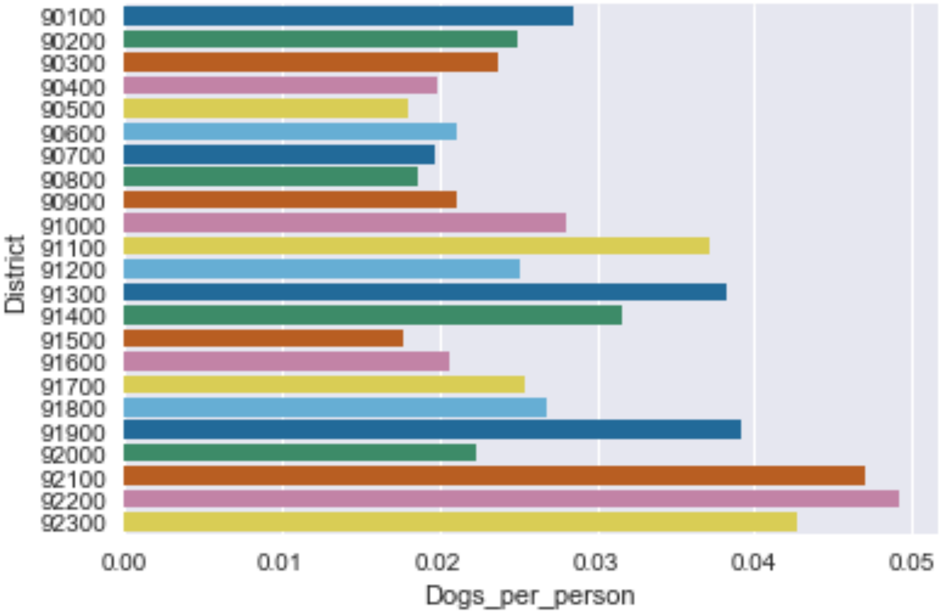


http://www.wolframalpha.com/input/?fileinput=%2FFileUpload%2FAliceInWonderland.txt&i=+&fileinputtype=file&examplefile=1

Visualisation



```
# Load library for visualisation
import seaborn as sns
# Command to show plots in notebook
%matplotlib inline
ax = sns.barplot(palette='colorblind', data=dog_vs_people, y=dog_vs_people.DISTR
ICT_CODE, x=dog_vs_people.Dogs_per_person, orient='h')
ax.set(ylabel='District', xlabel='Dogs_per_person')
```



https://www.communidata.at/notebook/dogs-in-vienna-part-2-open-data-integration-tutorial/

DataFrame

```
import pandas as pd # CSV file processing

data_path = 'https://www.wien.gv.at/finanzen/ogd/hunde-wien.csv'

data = pd.read_csv(data_path, delimiter=';', skiprows=1,
thousands=',', encoding=‘latin-1’)
```

Constructing DataFrame from a dictionary.

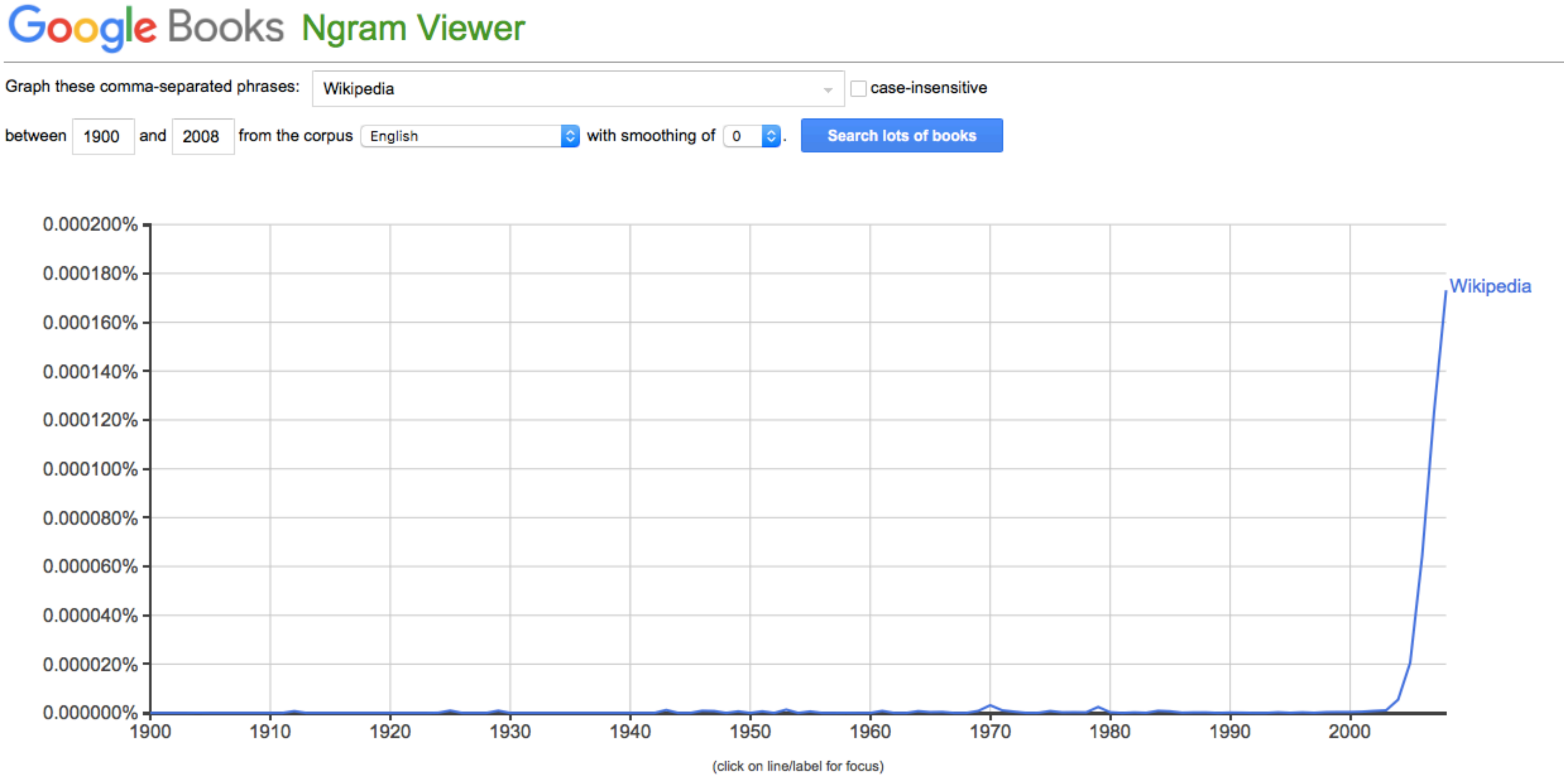
```
>>> d = {'col1': [1, 2], 'col2': [3, 4]}
>>> df = pd.DataFrame(data=d)
>>> df
   col1  col2
0     1     3
1     2     4
```

https://pandas.pydata.org/pandas-docs/stable/generated/pandas.DataFrame.html

https://www.communidata.at/notebook/dogs-in-vienna-part-2-open-data-integration-tutorial/

Google Ngram Viewer

- ngrams - sequence of *n* items from a sequence of text
- over 5 million books published up to 2008 (up to 5-grams)



Google Books Ngram Viewer

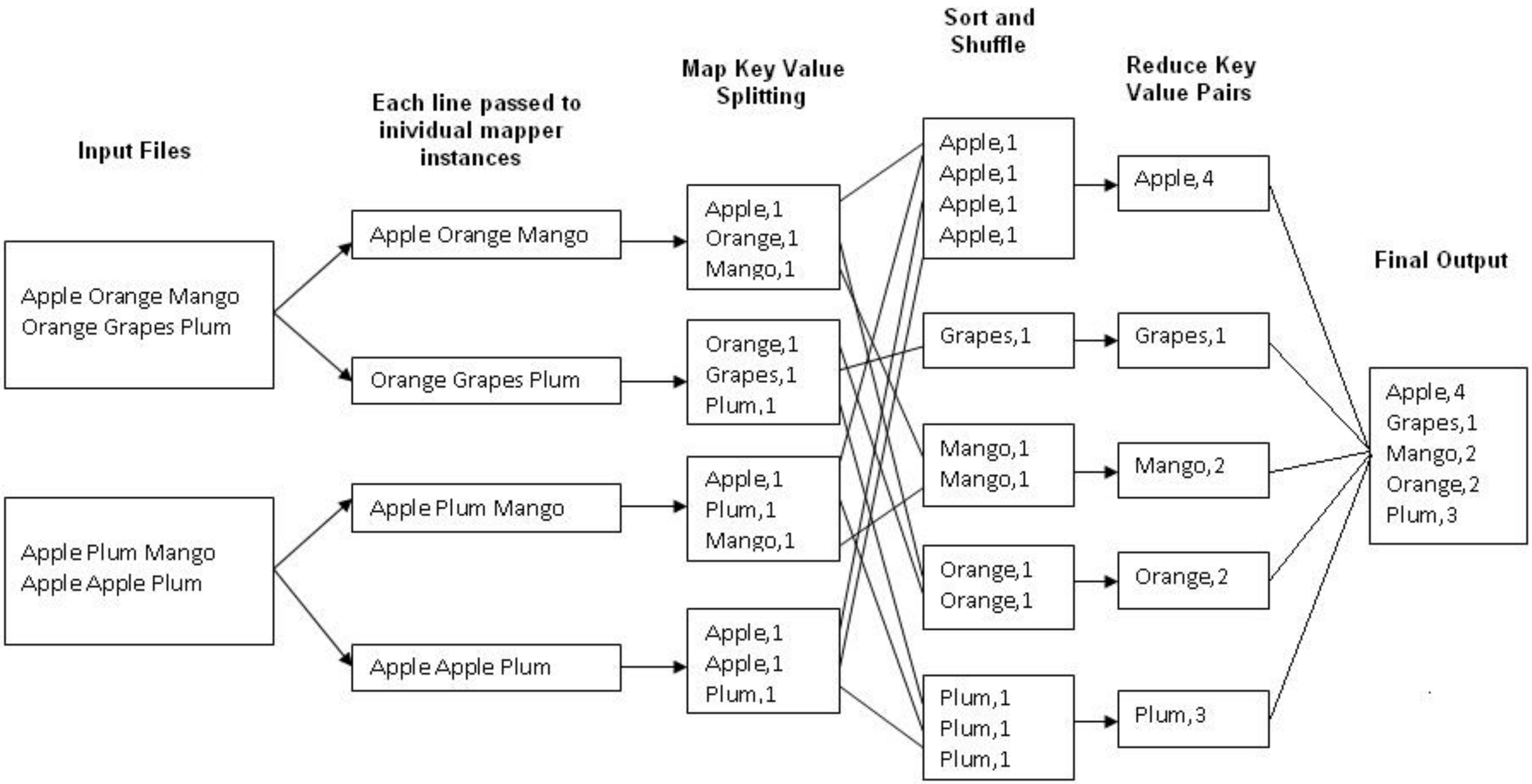
Graph these comma-separated phrases: Wikipedia □ case-insensitive

between 1900 and 2008 from the corpus English with smoothing of 0. Search lots of books

https://books.google.com/ngrams/graph?content=Wikipedia&year_start=1900&year_end=2020&corpus=15&smoothing=0&share=&direct_url=t1%3B%2CWikipedia%3B%2Cc0

https://en.wikipedia.org/wiki/Google_Ngram_Viewer https://en.wikipedia.org/wiki/N-gram

MapReduce



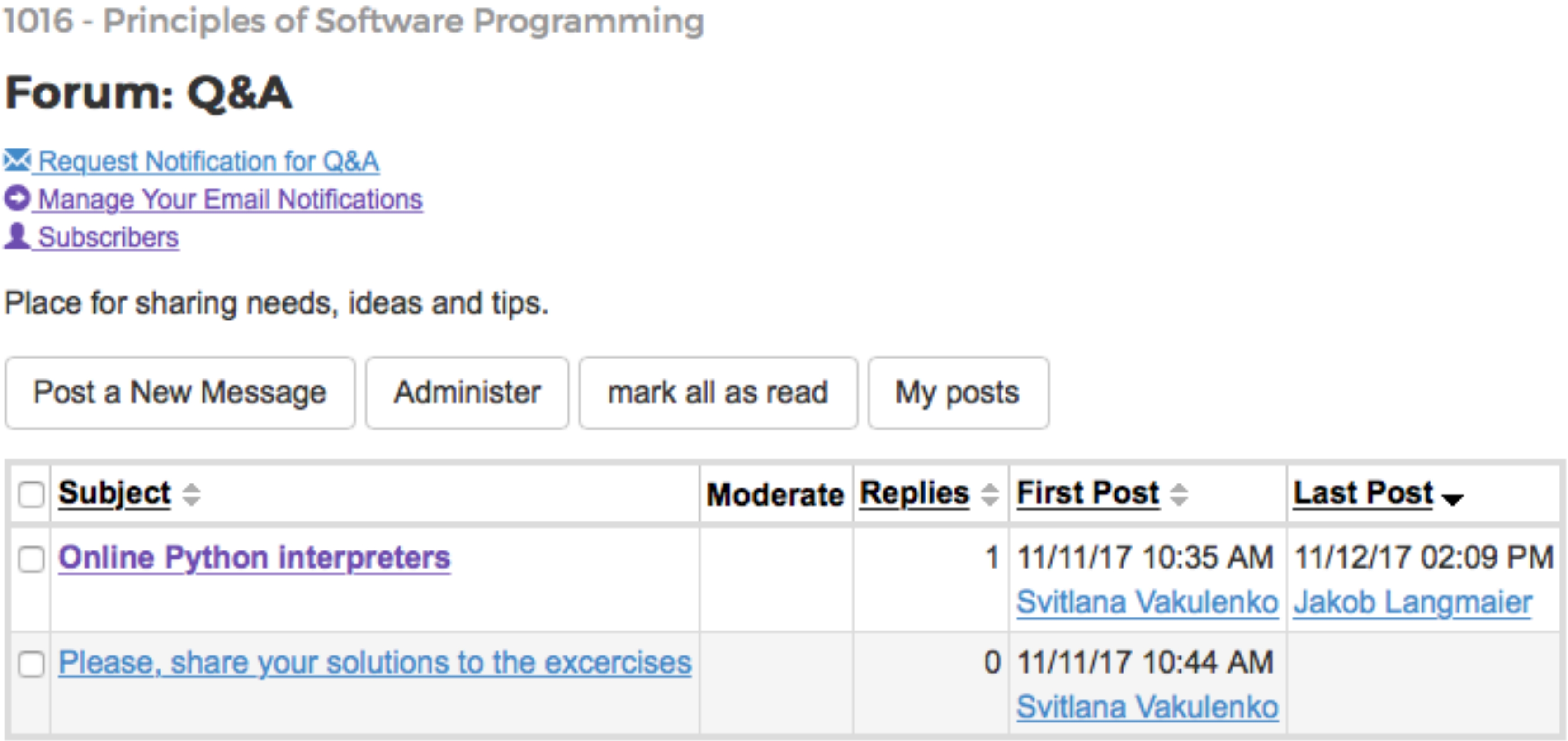
https://upload.wikimedia.org/wikipedia/commons/0/03/WordCountFlow.JPG

Schedule

	Topics	Dates
1	Course Overview, Introduction Python	Monday 10/30/17
2	Structured & Object-oriented paradigms	Friday 11/03/17
3	Data Structures: List, Set, Dictionary	Monday 11/06/17
4	Version Control, Project Structure	Wed 11/08/17
5	Files: Input/Output	Friday 11/10/17
6	**Debugging: Exceptions, Assertions**	Monday 11/13/17
7	Recap*	Wed 11/15/17
8	Trees, Recursion, Sort&Search*	Friday 11/17/17

01:00 PM - 03:45 PM D2.0.031 Workstation-Raum
*01:00 PM - 04:00 PM

Forum



1016 - Principles of Software Programming

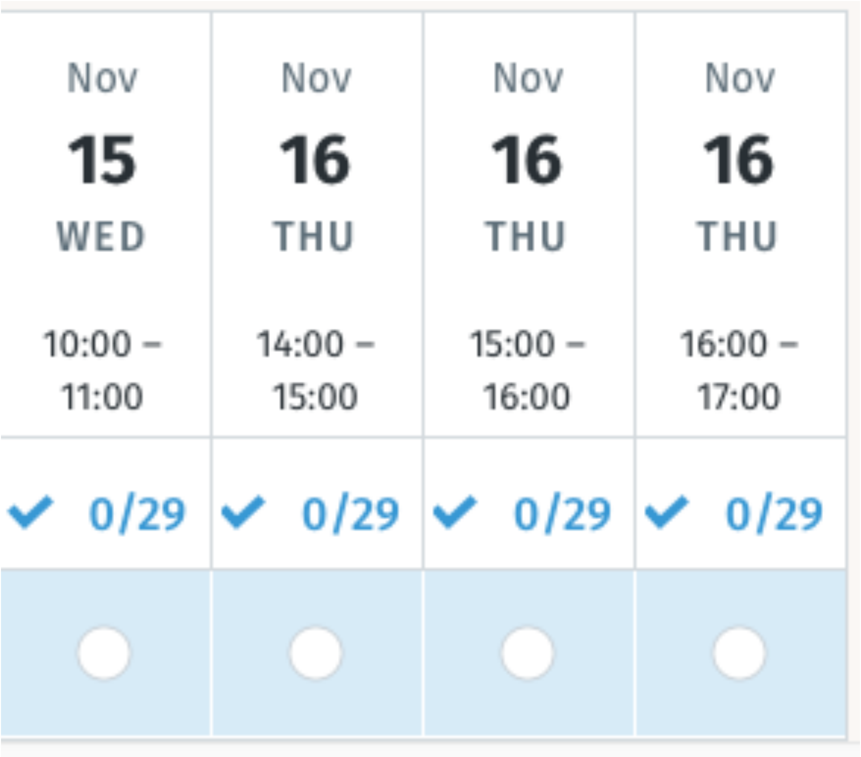
Forum: Q&A

Request Notification for Q&A
Manage Your Email Notifications
Subscribers

Place for sharing needs, ideas and tips.

□	Subject	Moderate	Replies	First Post	Last Post
□	Online Python interpreters		1	11/11/17 10:35 AM Svitlana Vakulenko	11/12/17 02:09 PM Jakob Langmaier
□	Please, share your solutions to the excercises		0	11/11/17 10:44 AM Svitlana Vakulenko	

Open data exploration task



Nov 15 WED 10:00 – 11:00	Nov 16 THU 14:00 – 15:00	Nov 16 THU 15:00 – 16:00	Nov 16 THU 16:00 – 17:00
✓ 0/29	✓ 0/29	✓ 0/29	✓ 0/29
○	○	○	○

https://doodle.com/poll/fmmphgnvi7ezsyuw