

Exploratory Data Analysis: Python Tools in Memology

Dmitry Sergeyev

About me



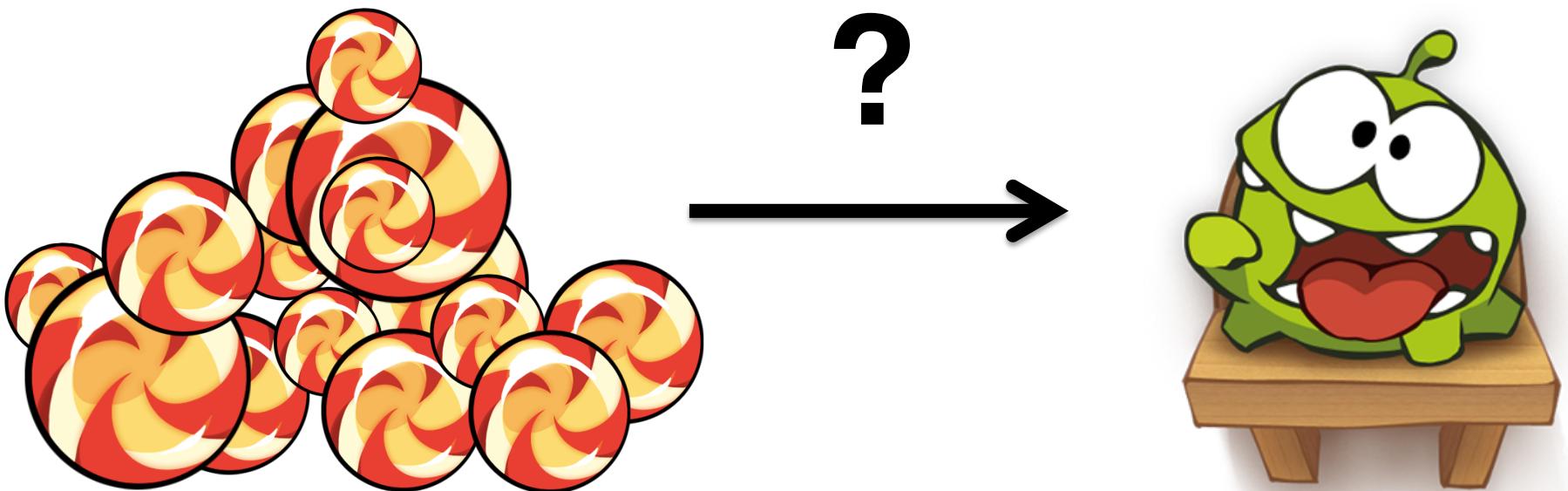
Center of Mathematical Finance
2015/2016



Zeptolab
Analyst



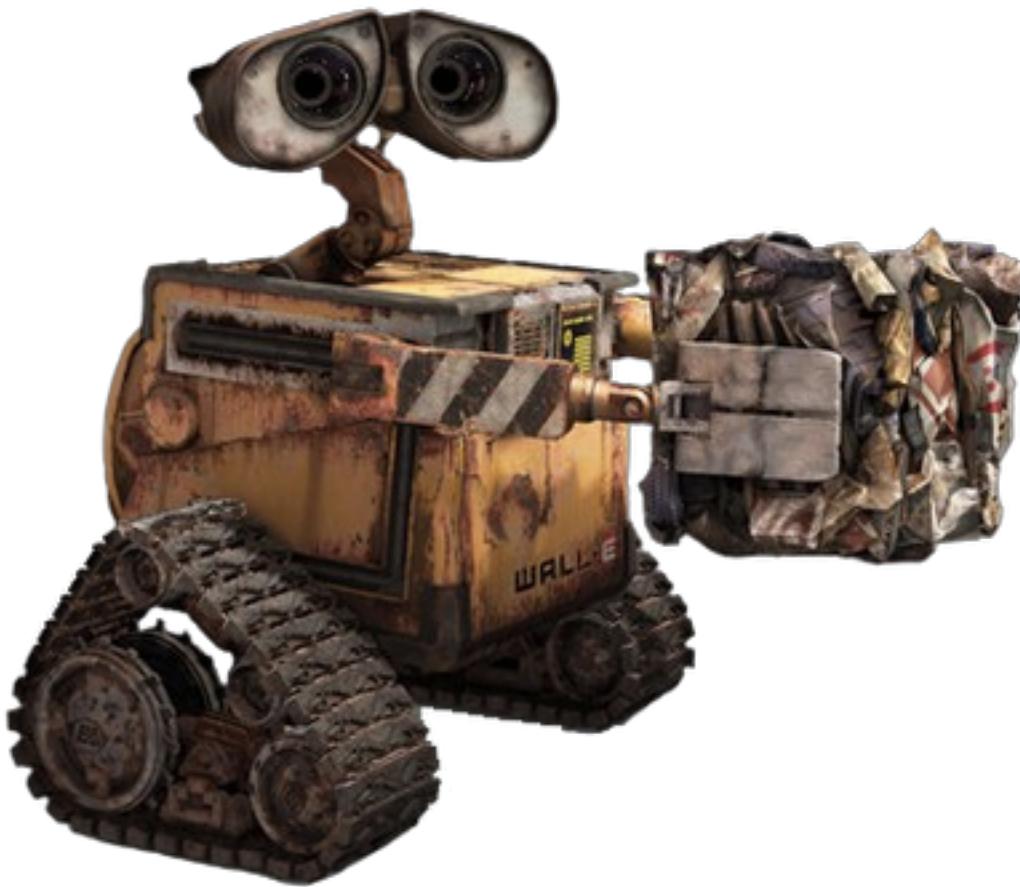
Analyst Life



Data is garbage



Data Analyst in Action



Datasets - DIY

If you are lucky enough - API

Datasets - DIY

If you are lucky enough - API

- All major Social Networks

Datasets - DIY

If you are lucky enough - API

- All major Social Networks
- Analytical databases (mostly payed)

Datasets - DIY

If you are lucky enough - API

- All major Social Networks
- Analytical databases (mostly payed)
- Some huge sites

Datasets - DIY

If you are not that lucky – brace yourselves

Datasets - DIY

If you are not that lucky – brace yourselves

- Write your own parsers

Datasets - DIY

If you are not that lucky – brace yourselves

- Write your own parsers
- Collect the data you need

Datasets - DIY

If you are not that lucky – brace yourselves

- Write your own parsers
- Collect the data you need
- Face constant challenges and limitations

Datasets - DIY

If you are not that lucky – brace yourselves

- Write your own parsers
- Collect the data you need
- Face constant challenges and limitations
- Error 400, 401, 403, 404, ...

Datasets - DIY

What is good about that?

- Exploration

Datasets - DIY

What is good about that?

- Exploration
- You get the data you really want

Datasets - DIY

What is good about that?

- Exploration
- You get the data you really want
- You might find something really cool and write an article about that!

Datasets - DIY

What is good about that?

- Exploration
- You get the data you really want
- You might find something really cool and write an article about that!
- Last, but not least – Python can hadle it all

Python vs R

Python vs R

- 2 years experience on R
- 3 years – Python

Python vs R

- 2 years experience on R
- 3 years – Python
- Now completely switched to latter

Python vs R

- 2 years experience on R
- 3 years – Python
- Now completely switched to latter
- But still visualisation is awesome in R

Memology

A little fun project about memes

CHEEZburger

I Can Has | Pop Culture | FAIL Blog | Memebase | Geek Universe | Loquillo | Know Your Meme

About Rules Chat Random Activity

Welcome! [Login](#) or [signup now!](#)

Search the database... Advanced Search Protips

HOME MEMES IMAGES VIDEOS FORUMS BLOG EPISODES

Also Trending: Five Nights at Freddy's Yahoo Male Purge Lawsuit Bear In The Big Blue House DURR PLANT Trigger Power Rangers Donald Trump



All Images

+ Add an Image

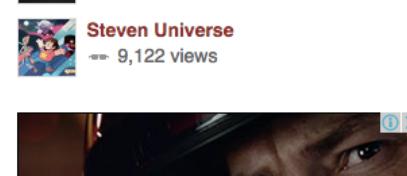
Browsing all 1,083,332 images

Share this Gallery 268

Like us on Facebook! Like 791K

[All](#) [Trending](#) [Most Commented](#) [Most Favorited](#) [Most Liked](#) [Least Liked](#) [Most Viewed](#)

Sorted by: **Newest** [Oldest](#) [Comments](#) [Favorites](#) [High Score](#) [Low Score](#) [Views](#)



Today's Top Image Galleries

Pokémon
18,707 views

JoJo's Bizarre Adventure
11,184 views

Overwatch
9,729 views

Steven Universe
9,122 views

Memology

A little fun project about memes

```
<!DOCTYPE html>
<html xmlns:fb="http://www.facebook.com/2008/fbml" xmlns="http://www.w3.org/1999/xhtml">
  ▶<head>...</head>
  ▷<body class="cols guest photos photos_index listings production" data-controller-action="photos-
index" data-hide-ads="false" data-layout="desktop" data-nsfw="false">
    <div id="cboxOverlay" title="Overlay" style="display: none;"></div>
    ▶<div id="colorbox" class="padding-bottom: 57px; padding-right: 28px; display: none;">
    ...</div>
    ▷<div id="container">
      ▷<div class="longcat wrap">
        ▶<div class="blueBar">...</div>
        ▶<div class="rel" id="omgkym">...</div>
        ▷<div class="x_brb6 c" id="content">
          ▷<div class id="maru">
            ▶<div class="c" id="section_header_wrap">...</div>
            <div id="flashes"> </div>
            ▶<nav>...</nav>
            ▶<div class="search_sort_options images">...</div>
            ▷<div class="infinite" id="photo_gallery" style="position: relative; height: 1550px;">
              ▷<div class="item" style="position: absolute; left: 0px; top: 0px;">
                ▷<a class="photo cboxElement" data-colorbox-url="/photos/1178374-five-nights-at-freddy-
s/colorbox" href="/photos/1178374-five-nights-at-freddy-s" rel="photo_gallery" target=
"_blank">
                   = $0
                ▶<div class="info abs">...</div>
              </a>
            </div>
          </div>
        </div>
      </div>
    </div>
  </body>
</html>
```

Memology

A little fun project about memes

- Parsed data from KnowYourMeme.com

Memology

A little fun project about memes

- Parsed data from KnowYourMeme.com
- 365 pages, 40 memes on each

Memology

A little fun project about memes

- Parsed data from KnowYourMeme.com
- 365 pages, 40 memes on each
- 14 600 data points for research purposes

Memology

A little fun project about memes

- Parsed data from KnowYourMeme.com
- 365 pages, 40 memes on each
- 14 600 data points for research purposes
- Descriptions, views, origins, etc...

Memology

A little fun project about memes

- How to write your very own parsers - Next time
- Now – let's look at basic data exploration tools in Python

