

Python Programming for Young Data Scientists

Kan Ouivirach
Burasakorn Sabyeying

a bit about our work..

Better Marketing Visibility

Pronto Online Dashboard Views Search Pronto Marketing Kan Ouvirach

Insights / Main Pronto Site / Keyword Rankings

Main Pronto Site prontomarketing.com | United States - English Add Keywords Keyword Settings

KEYWORD RANKINGS

HIGHLIGHTS

- Highest Keyword Rank: 1
- Keywords Ranked in Top 5: 4
- Keywords Ranked in Top 10: 9
- Average Keyword Rank: 18

Tracked Keywords (33) of 15 Keywords Limit

KEYWORD	GOOGLE RANK	SEARCH VOLUME	DIFFICULTY	MANAGE
health club web design	1 -	Low	30	[Edit]
msp web design	1 ▲	Low	24	[Edit]
internet presence management	3 -	110	47	[Edit]
small business internet marketing	4 -	210	63	[Edit]
msp marketing	6 ▲	140	34	[Edit]
msp website	7 ▼	70	55	[Edit]
health club internet marketing	8 -	Low	25	[Edit]

Pronto Online Dashboard Views Search Pronto Marketing Kan Ouvirach

Insights / Main Pronto Site / Traffic

Main Pronto Site Pronto Marketing - All Domains http://prontomarketing.com | UA-5186490-4 November 7, 2015 - December 6, 2015 Settings

Total Visitor Traffic: 11,746 Unique Visits: 6,861 Avg Time On Site: 141s Bounce Rate: 62%

TRAFFIC EVOLUTION Nov 7 2015 - Dec 6 2015

NEW VS RETURNING VISITORS Nov 7 2015 - Dec 6 2015

TOP ORGANIC KEYWORDS Nov 7 2015 - Dec 6 2015

	VISITS
(not provided)	4,083
pronto marketing	25
https://app.prontomarketing.com/services/15/details/	4
pronto	4
pronto marketing - www.gowhiteowl.com	4
business website marketing	3
how do you create campaign structure.	3
how to use long tail keyword in blog post	3
https://app.prontomarketing.com/services/	3
prontomarketing	3

Pronto Online Dashboard Views Search Pronto Marketing Kan Ouvirach

Insights / Main Pronto Site / Competitors

Main Pronto Site Add Competitor

Competitive Analysis Metrics

Domain Authority Linking Domains MozRank MozTrust

COMPETITOR DOMAIN

COMPETITOR DOMAIN	DOMAIN AUTHORITY	LINKING DOMAINS	MOZRANK	MOZTRUST
www.prontomarketing.com Primary	43	673	5.95	5.79
www.verticalxion.com	21	16	4.94	5.26
www.joomconnect.com	40	217	6.03	5.85
www.kutenda.com	33	16	4.25	5.17
www.ulistic.com	36	160	5.91	5.77
www.technologymarketingtoolkit.com	29	96	5.25	5.41
www.propelmarketing.com	48	40	4.79	5.26
www.qunb.com	36	106	5.03	5.85

Pronto Online Dashboard Views Search Pronto Marketing Kan Ouvirach

Insights / Main Pronto Site / Leads

Main Pronto Site January 1, 2015 - December 7, 2015 Sync Export to CSV Spam Trash Settings

Total Leads: 1,532

LEADS BY CHANNEL

LEADS OVER TIME

CONTACT

CONTACT	LEAD CHANNEL	CREATED	PAGE	FORM	DETAILS
jeff reale	Organic	Dec 06, 2015	Lead Generation Ebook	[ebook download] Lead Generation	[Details]
Rick Gordhamer	Organic	Dec 06, 2015	25 Website Must Haves Ebook	[ebook download] 25 Website Must Haves	[Details]
Stephen Hackett	Organic	Dec 06, 2015	Marketing Metrics Ebook	[ebook download] 6 Metrics	[Details]
Timothy	Email	Dec 06, 2015	Pricing	Pricing PDF Download	[Details]
vimal	Organic	Dec 05, 2015	25 Website Must Haves Ebook	[ebook download] 25 Website Must Haves	[Details]
Gcinithemba Fakudze	Organic	Dec 05, 2015	Marketing Metrics Ebook	[ebook download] 6 Metrics	[Details]
Roy McKenzie	Email	Dec 04, 2015	Pricing Comparison	[pricing download] Website Cost Comparison	[Details]
Famokun Emma	Organic	Dec 04, 2015	LinkedIn Company Page Ebook	[ebook download] LinkedIn Company Page	[Details]
Thao Tran	Referral	Dec 04, 2015	Ipm Report	IPM Report Request	[Details]
AVID	Organic	Dec 04, 2015	Pricing	Pricing PDF Download	[Details]

Pronto Online Dashboard Views Search Pronto Marketing Kan Ouvirach

Insights / Main Pronto Site / Call Tracking

Main Pronto Site Call Details Actions

Call Status: Received Caller Name: Virginia Call Phone Number: (540) 691-8903 Date: Dec 04, 2015 at 01:45 PM (GMT -5) Duration: 1m 58s Tracking Number: Keyword Pool (855) 997-2317 Location: Covington, VA

Call Recording Download MP3

Caller History

Dec 04, 2015 at 01:45 PM (GMT -5) [Details]

Call Notes

Add New Note Add Note

Web Source Details

Source Referrer: https://www.google.com/ Source Channel: Organic Source Keyword: (not provided) Landing Page: https://www.prontomarketing.com/2015/...

Pages Visited

https://www.prontomarketing.com/2015/01/14-tips-for-effective-email-m...

Pronto Online Dashboard Views Search Pronto Marketing Kan Ouvirach

Insights / Main Pronto Site / Reports

Main Pronto Site Summary Nov. 30, 2015 - Dec. 6, 2015

Performance Report

Summary

Visits: 2,960 Top Traffic Channel: Organic Leads: 38 Top Lead Channel: Organic

Visitor Traffic Nov. 30, 2015 - Dec. 6, 2015

In this week, your site had 2,960 visits, a **19% increase** compared with the previous week.

Summary Nov. 30, 2015 - Dec. 6, 2015

Visits: 2,960 Top Traffic Channel: Organic Leads: 38 Top Lead Channel: Organic

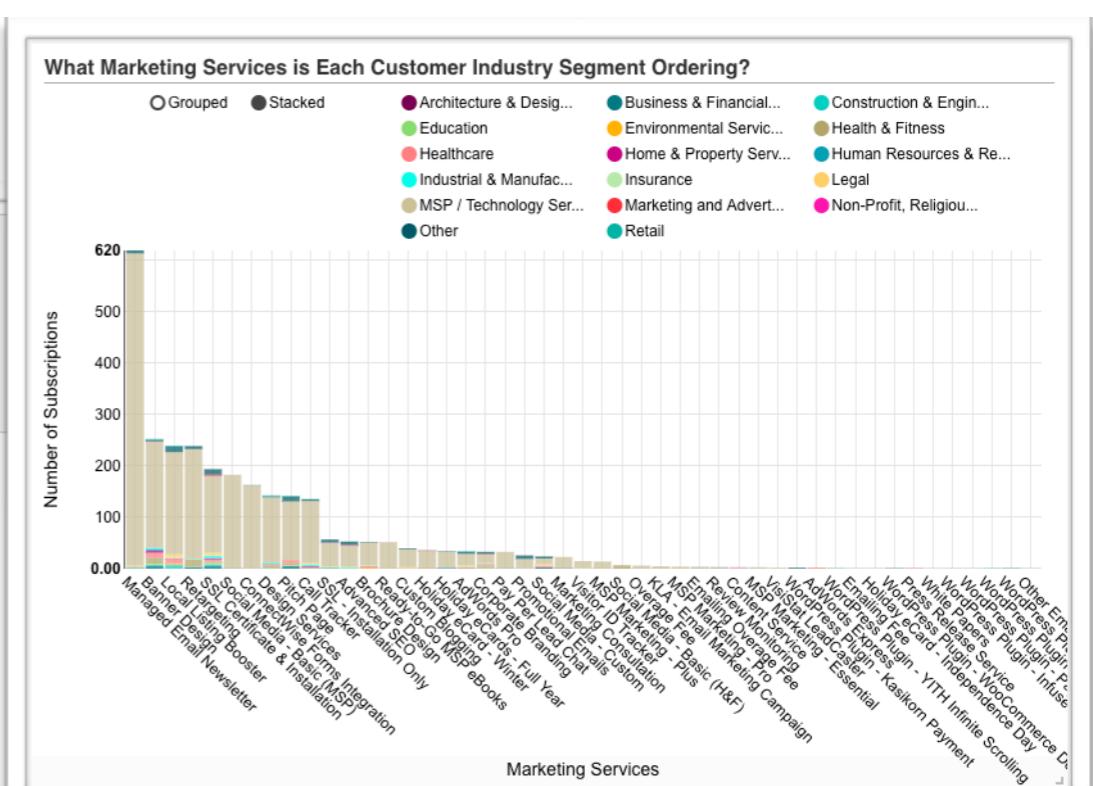
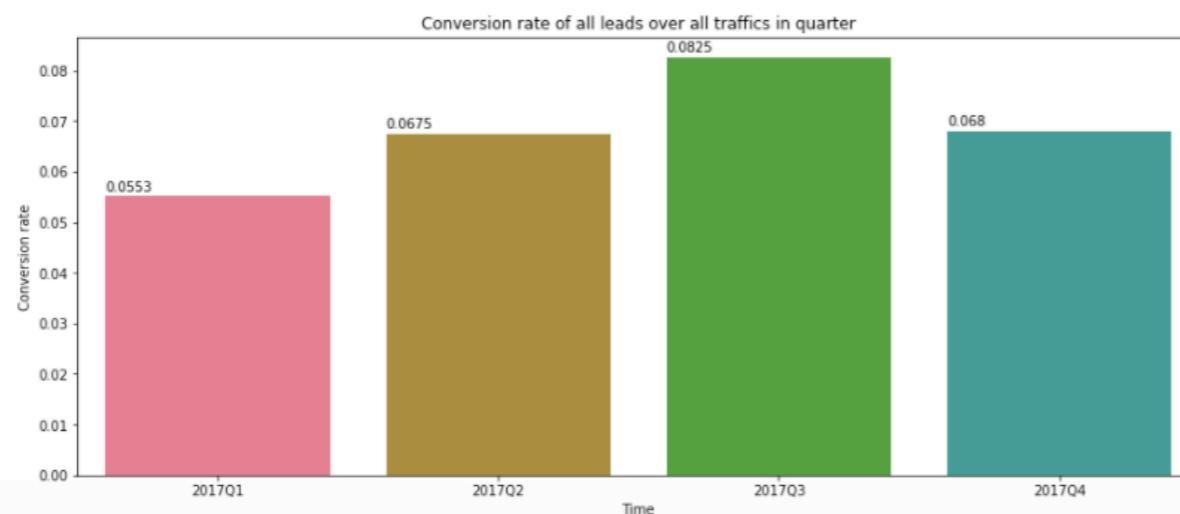
Visitor Traffic Nov. 30, 2015 - Dec. 6, 2015

32% of your all your site's visitors came from **Organic** traffic.

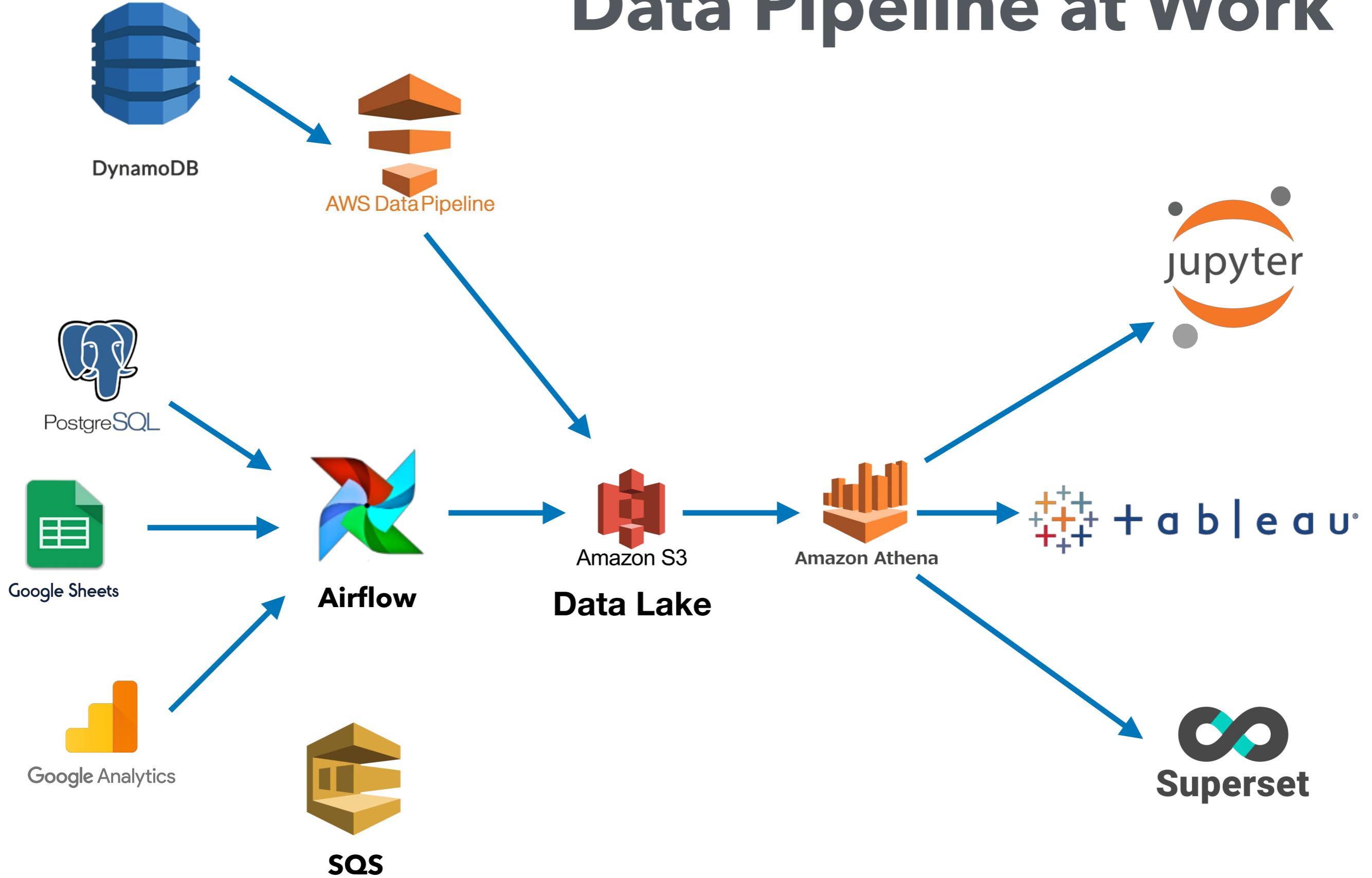
Data Science at Work

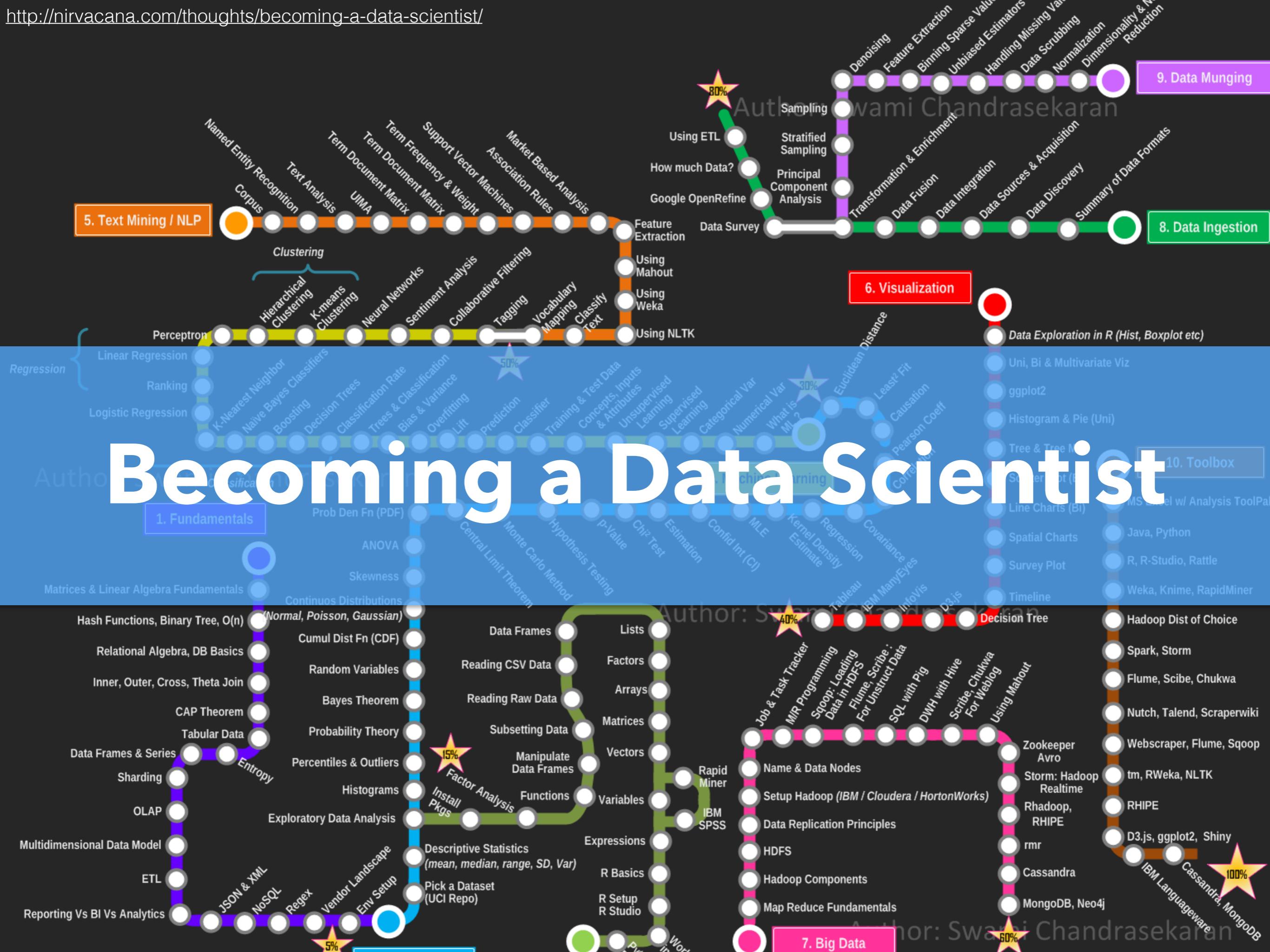


```
In [33]: plt.figure(figsize=(15, 6))
palette = sns.set_palette('husl')
df_quarter_count = df_quarter_count.round(4)
g = sns.barplot(x='year_and_month', y='conversion_rate', data=df_quarter_count, palette=palette)
g.set_xlabel('Time', ylabel='Conversion rate')
g.set_title('Conversion rate of all leads over all traffics in quarter')
for p in g.patches:
    g.annotate(str(p.get_height()), (p.get_x(), p.get_height() * 1.015));
g.set_xticklabels(g.get_xticklabels(), rotation=0);
```



Data Pipeline at Work







((Josh Wills))

@josh_wills



Following

Data Scientist (n.): Person who is better at statistics than any software engineer and better at software engineering than any statistician.

RETWEETS

1,474

LIKES

1,001



11:55 PM - 3 May 2012



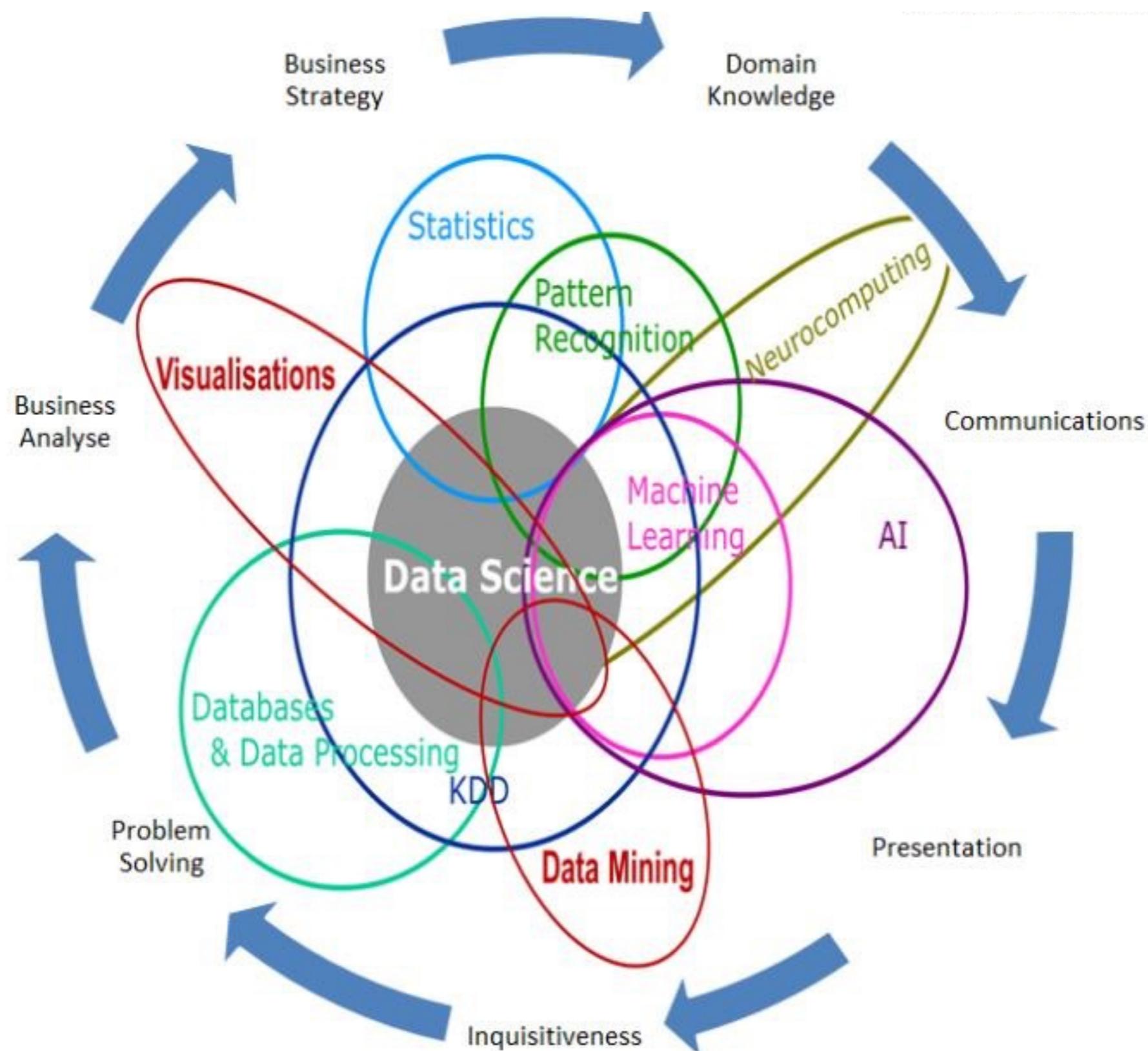
1.5K

1K

...

https://twitter.com/josh_wills/status/198093512149958656

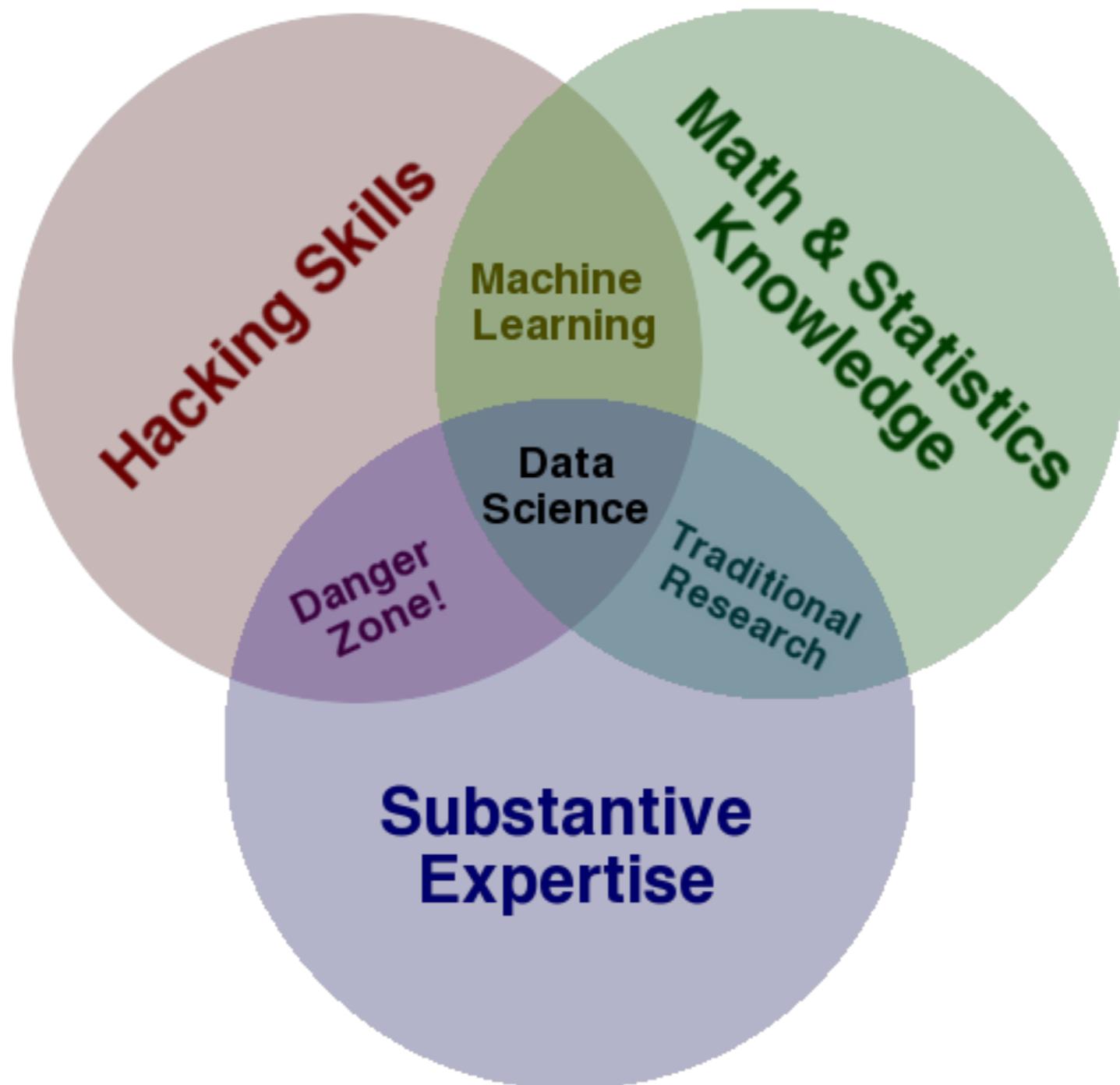
Data Science is Multidisciplinary



Bad Data Scientists

1. Don't like scientific enquiries
2. Make decisions on gut feeling
3. Like the details
4. Don't like the details
5. Have closed mind
6. Hate failure
7. Have only one fraction of skills

Data Science Venn Diagram



Data Science Process in Organization

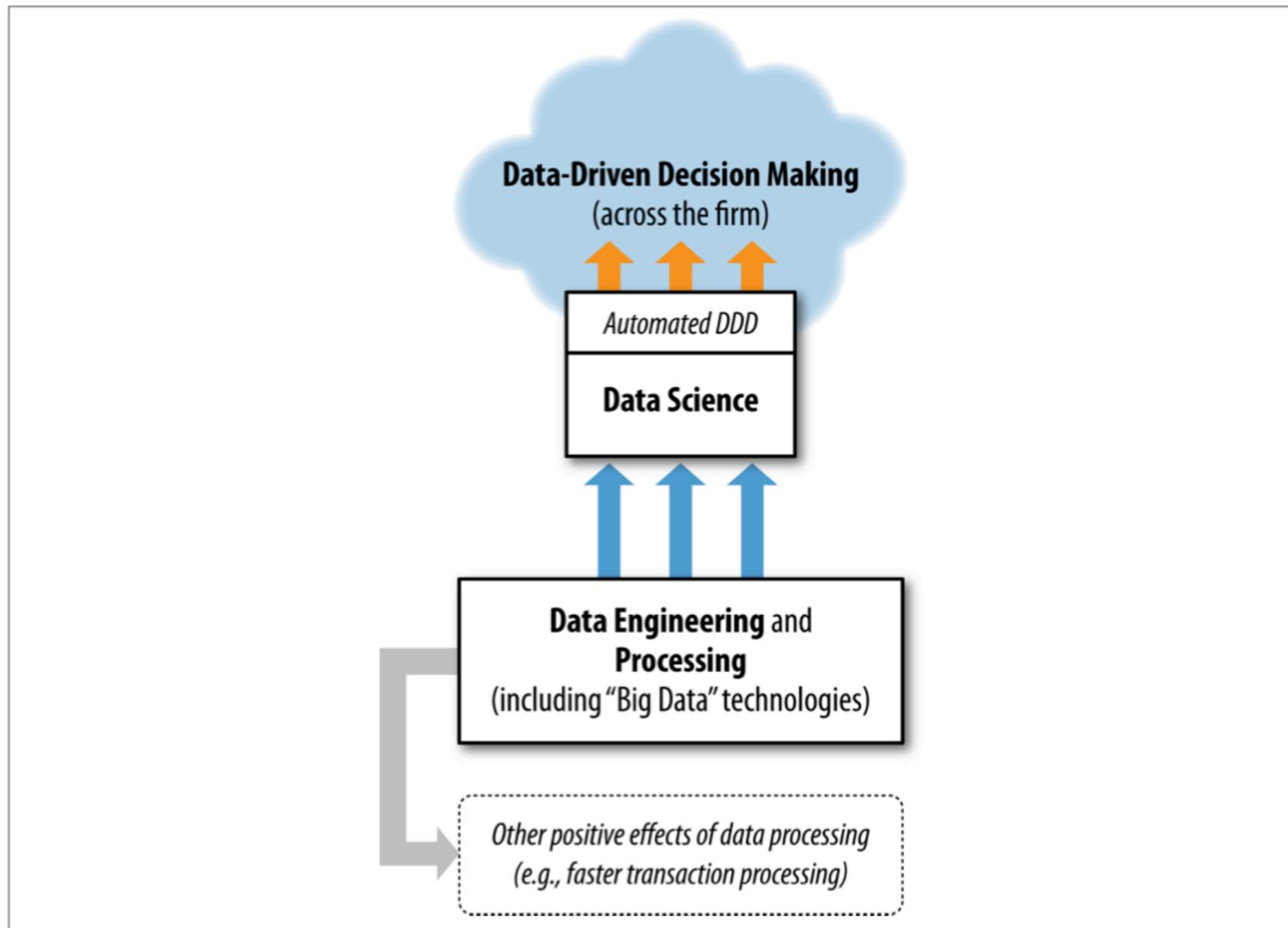
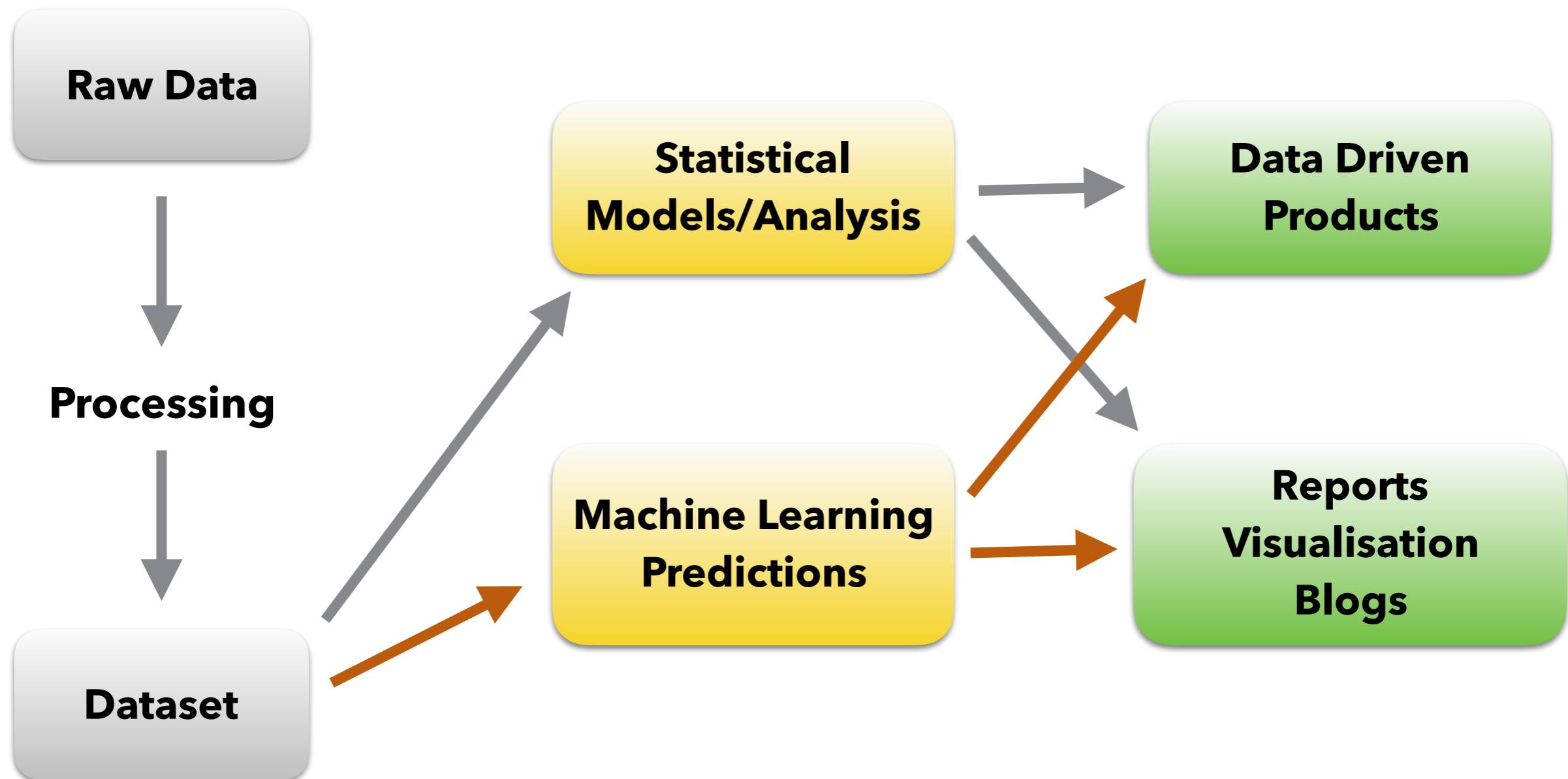


Figure 1-1. Data science in the context of various data-related processes in the organization.

What do Data Scientists do?

What do Data Scientists do?



As a Data Scientist, We Need

1. Good problem
2. Good data
3. Good **TOOLs**



Solving Data Science using Good TOOLS

Beautiful is better than ugly.
Explicit is better than implicit. Simple
is better than complex. Complex is better
than complicated. Flat is better than
nested. Sparse is better than dense.
Readability counts. Special cases aren't
special enough to
break the rules.

Although practicality beats purity. Errors should never
pass silently. Unless explicitly silenced. In the face of
ambiguity, refuse the temptation to guess. There should be one
— and preferably only one — obvious way to do it. Although that
way may not be obvious at first *unless you're Dutch*. Now is
better than never. Although never is often better than right
now. If the implementation is hard to explain, it's a bad
idea. If the implementation
is easy to explain, it
may be a good idea.
Namespaces are
one honking great
idea — let's do
more of those!

Beautiful is better than ugly.
Explicit is better than implicit. Simple
is better than complex. Complex is better
than complicated. Flat is better than
nested. Sparse is better than dense.
Readability counts. Special cases aren't
special enough to
break the rules.
Although practicality beats purity. Errors should never
pass silently. Unless explicitly silenced. In the face of
ambiguity, refuse the temptation to guess. There should be one
— and preferably only one — obvious way to do it. Although that
way may not be obvious at first *unless you're Dutch*. Now is
better than never. Although never is often better than right
now. If the implementation is hard to explain, it's a bad
idea. If the implementation
is easy to explain, it
may be a good idea.
Namespaces are
one honking great
idea — let's do
more of those!

Why?

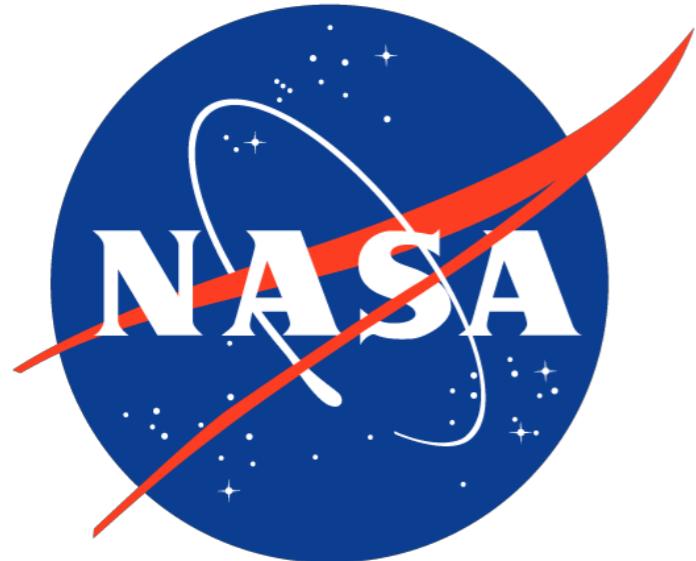
TM

Popular



DISQUS

Google



Pinterest



Easy



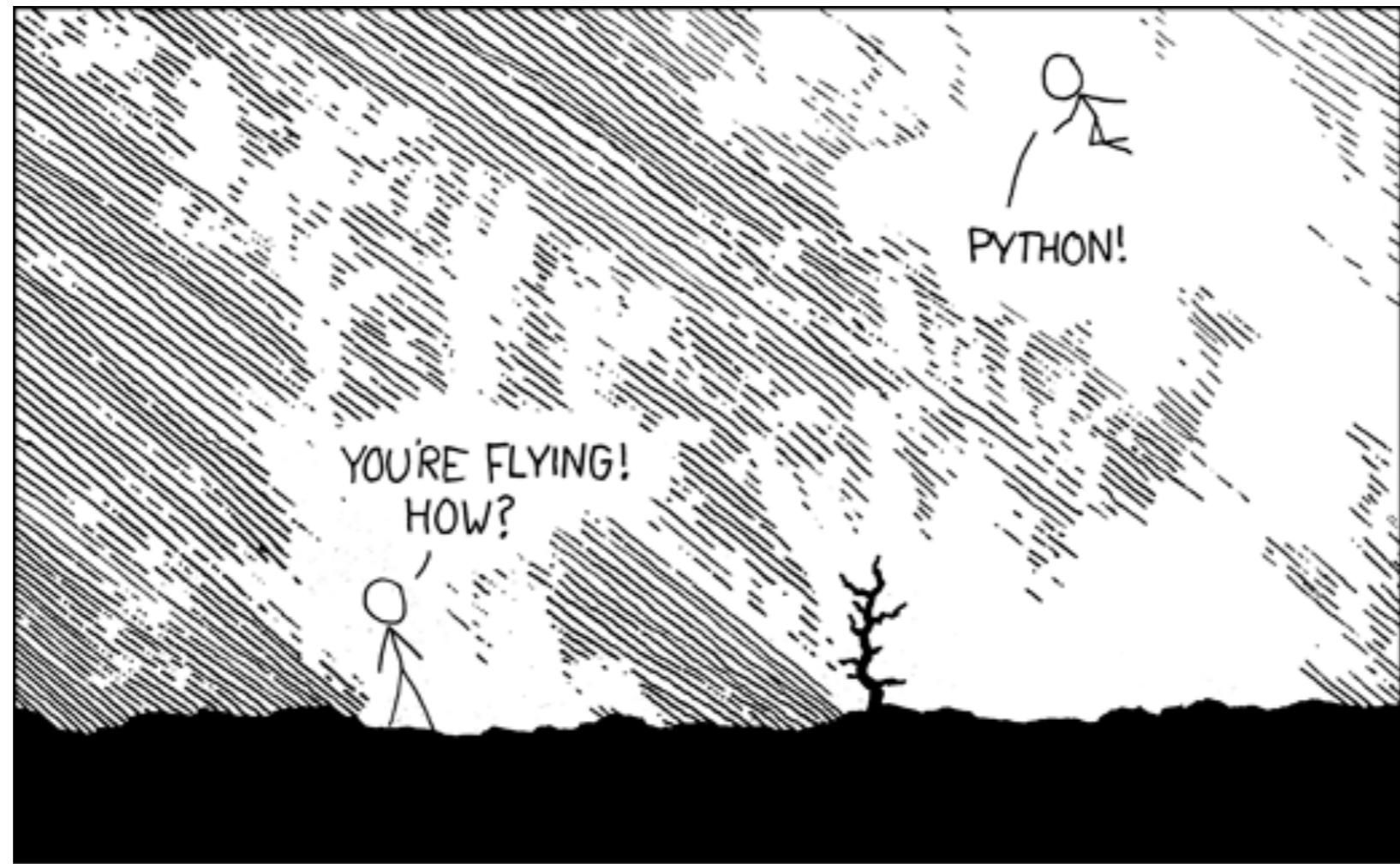
Ja...:

```
public class HelloWorld {  
    public static void main(String[], args) {  
        System.out.println("Hello, World!");  
    }  
}
```

Python:

```
print("Hello, World")
```





I LEARNED IT LAST NIGHT! EVERYTHING IS SO SIMPLE!

HELLO WORLD IS JUST

`print "Hello, world!"`

I DUNNO...
DYNAMIC TYPING?
WHITESPACE?

COME JOIN US!
PROGRAMMING IS FUN AGAIN!
IT'S A WHOLE NEW WORLD UP HERE!

BUT HOW ARE YOU FLYING?

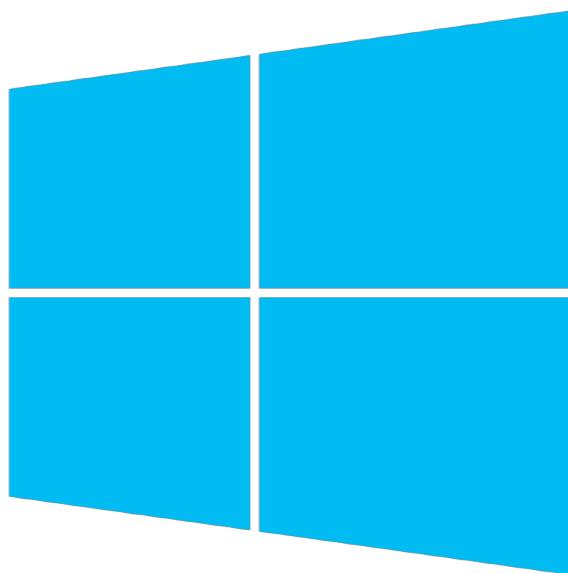
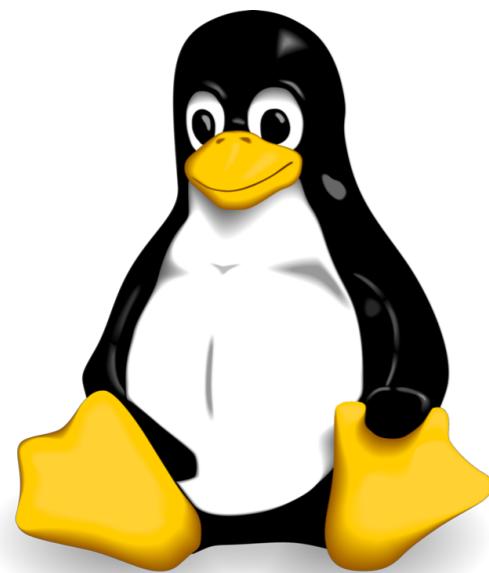
I JUST TYPED
`import antigravity`

THAT'S IT?

... I ALSO SAMPLED EVERYTHING IN THE MEDICINE CABINET FOR COMPARISON.

BUT I THINK THIS IS THE PYTHON.

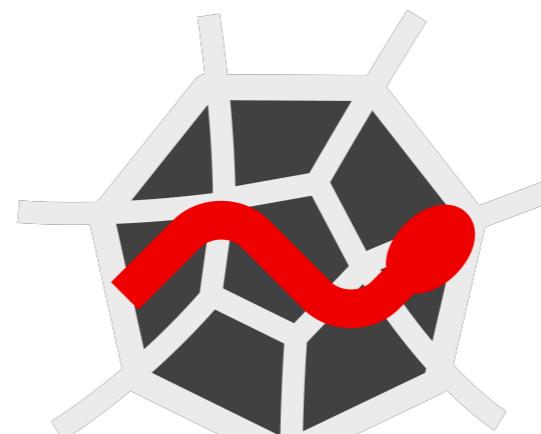
Flexible



Coding Environments and Editors



PyCharm



spyder

and more..

“Python = Potential”

-JEREMY ACHIN



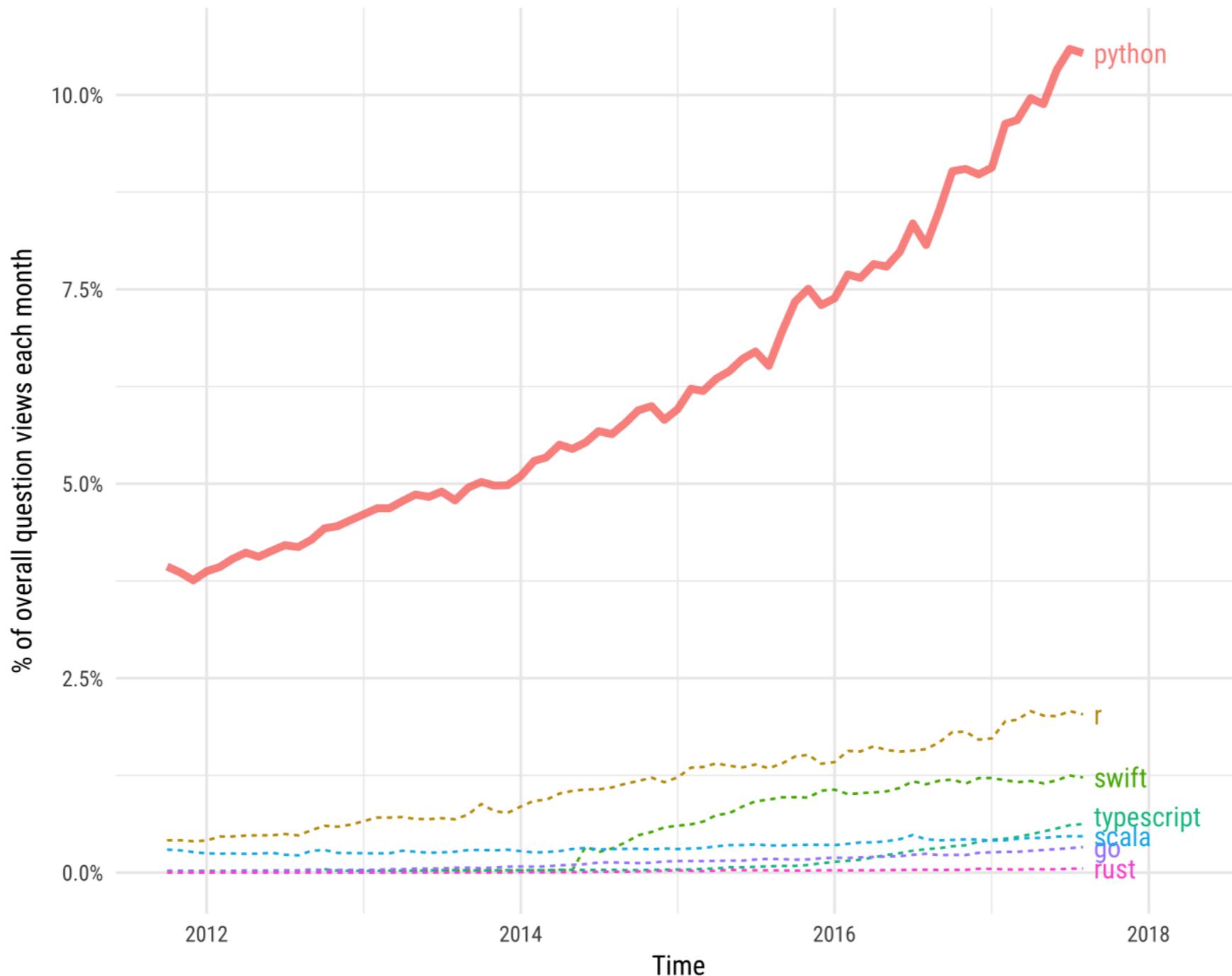
<https://www.youtube.com/watch?v=CoxjADZHUQA>

THE INCREDIBLE GROWTH OF PYTHON

<https://stackoverflow.blog/2017/09/06/incredible-growth-python/>

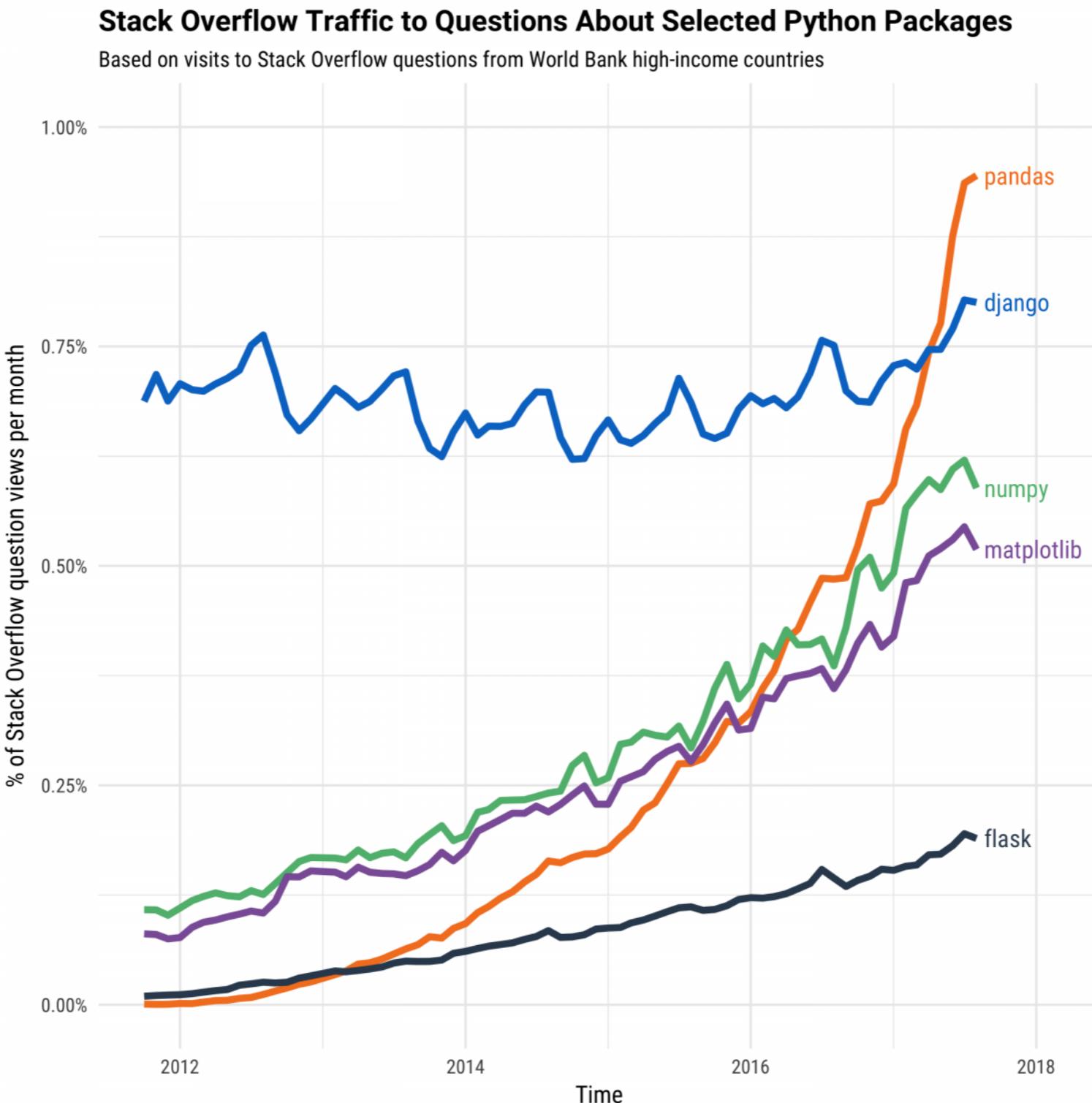
Python compared to smaller, growing technologies

Based on question traffic in World Bank high-income countries



WHY IS PYTHON GROWING SO QUICKLY?

<https://stackoverflow.blog/2017/09/14/python-growing-quickly/>



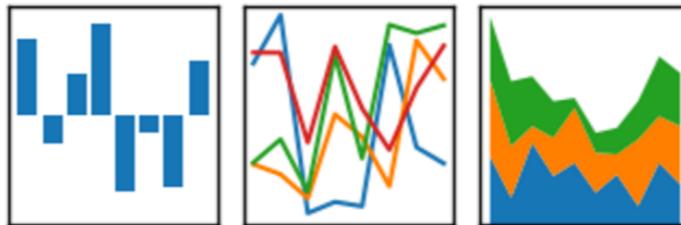
A giant panda is climbing a tree, its black and white fur contrasting with the brown bark. The background is a soft-focus green and blue.

Pandas

Python Data Analysis Library

pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$



[overview](#) // [get pandas](#) // [documentation](#) // [community](#) // [talks](#)

Python Data Analysis Library

pandas is an open source, BSD-licensed library providing high-performance, easy-to-use data structures and data analysis tools for the [Python](#) programming language.

pandas is a [NUMFocus](#) sponsored project. This will help ensure the success of development of *pandas* as a world-class open-source project.

A Fiscally Sponsored Project of



0.18.1 Final (May 3, 2016)

This is a minor release from 0.18.0 and includes a large number of bug fixes along with several new features, enhancements, and performance improvements. We recommend that all users upgrade to this version.

VERSIONS

Release

0.18.1 - May 2016

[download](#) // [docs](#) // [pdf](#)

Development

0.18.2 - July 2016

[github](#) // [docs](#)

Previous Releases

0.18.0 - [download](#) // [docs](#) // [pdf](#)

0.17.1 - [download](#) // [docs](#) // [pdf](#)

0.17.0 - [download](#) // [docs](#) // [pdf](#)

0.16.2 - [download](#) // [docs](#) // [pdf](#)

0.16.1 - [download](#) // [docs](#) // [pdf](#)

0.16.0 - [download](#) // [docs](#) // [pdf](#)

0.15.2 - [download](#) // [docs](#) // [pdf](#)

0.15.1 - [download](#) // [docs](#) // [pdf](#)

0.15.0 - [download](#) // [docs](#) // [pdf](#)

0.14.1 - [download](#) // [docs](#) // [pdf](#)

<http://pandas.pydata.org/>

panel data
(multi-dimensional data)

Our Goals in this Course!

- Take data and produce insights
- Build an exploratory data visualizations



Learning Data Science

- Learn to love data
- Learn by doing
- Learn to communicate insights
- Learn from experts and peers
- Learn constantly the degree of difficulty



ANACONDA[®]

<https://www.continuum.io/downloads>