GENERAL ASSEMBLY

# PYTHON 101

Jon-Cody Sokoll
Data Scientist, Deloitte

## ABOUT ME

# WECOME!

Here’s a bit about me:

**Name:** Jon-Cody Sokoll

**Background:** Rhodes College, American Institutes for Research, Deloitte

**Fun Fact:** Moonlight bread baker

# BEFORE WE DIVE IN…

Let’s talk a bit about you!

**Name**

**What brings you to GA?**

- ☐ Current activities
- ☐ Future goals

**Fun Fact**

# LEARNING OBJECTIVES

- ☐ Discover the history of Python and how it compares to other programming languages.
- ☐ Touch on fundamental Python programming techniques and tools.
- ☐ Discuss its applications in data science and the types of problems it can solve.
- ☐ See where Python programming fits into the data workflow.
- ☐ Apply your new skills to solve a real-world problem with Python.

# WHY PYTHON?

- ☐ Created for simplicity and readability
- ☐ Rapid prototyping, ease of production
- ☐ Open source, importable libraries
- ☐ Broad range of applications
- ☐ Fast growing community

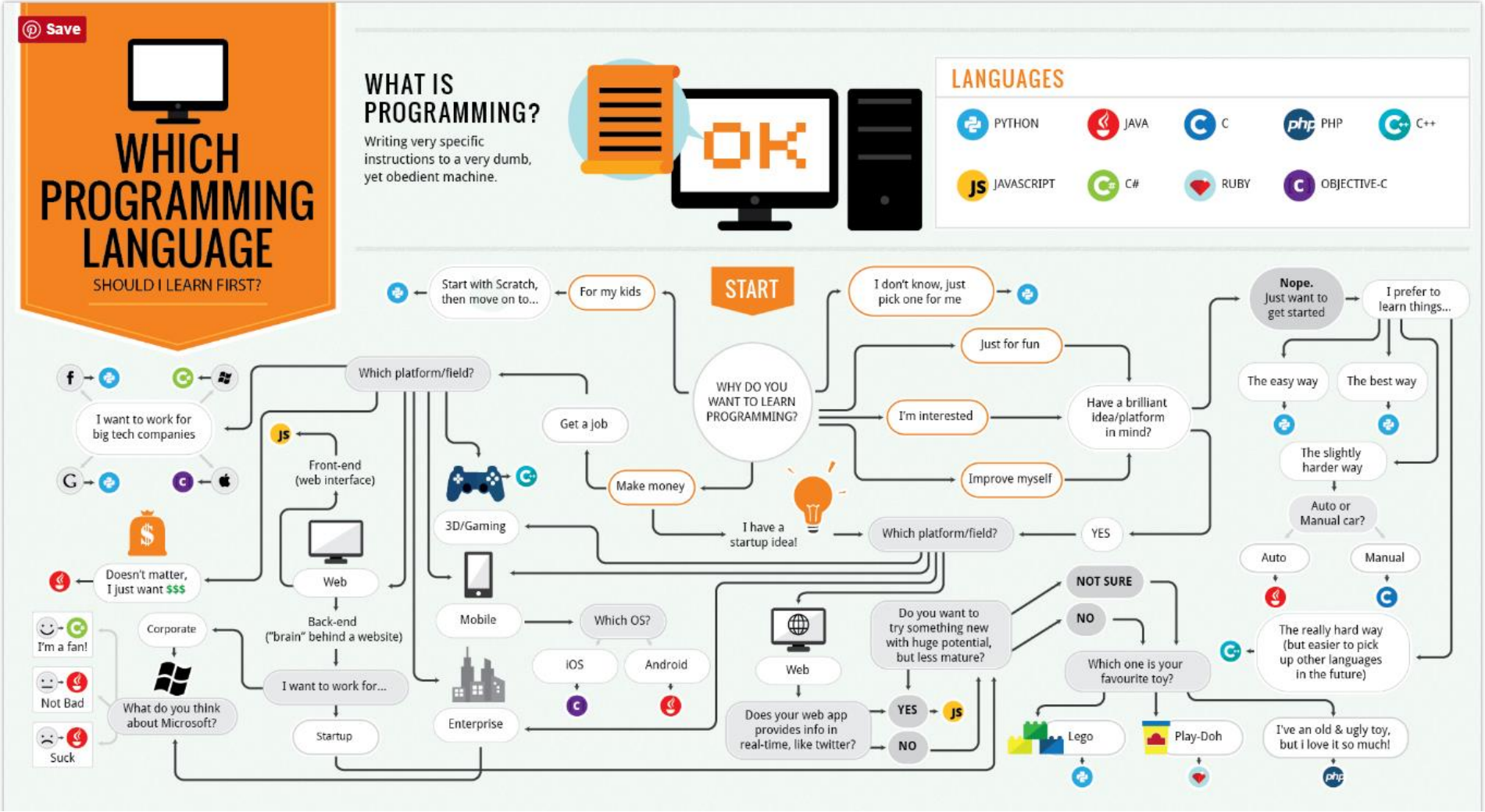
Founded by
Guido van Rossum
in 1991

## WHAT TOOLS DO DATA SCIENTISTS USE?

# ZEN OF PYTHON

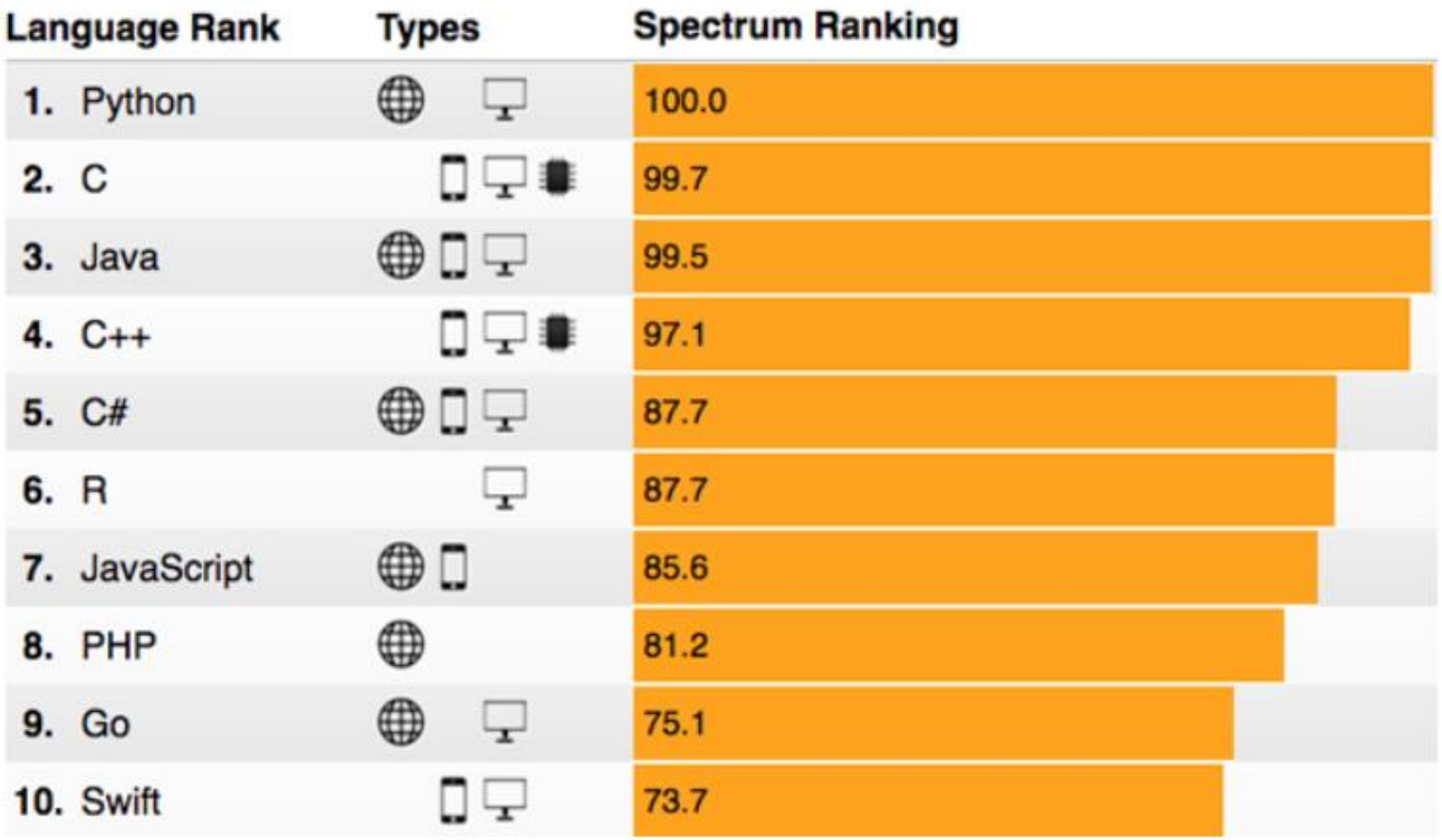
Beautiful is better than ugly.
Explicit is better than implicit.
Simple is better than complex.
Complex is better than complicated.
Flat is better than nested.
Sparse is better than dense.
Readability counts.
Special cases aren't special enough to break the rules.
Although practicality beats purity.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


Save
WHICH PROGRAMMING LANGUAGE
SHOULD I LEARN FIRST?
WHAT IS PROGRAMMING?
Writing very specific instructions to a very dumb, yet obedient machine.
OK
LANGUAGES
PYTHON
JAVA
C
PHP
C++
JAVASCRIPT
C#
RUBY
OBJECTIVE-C
START
WHY DO YOU WANT TO LEARN PROGRAMMING?
For my kids
Start with Scratch, then move on to...
I don't know, just pick one for me
Just for fun
I'm interested
Improve myself
Have a brilliant idea/platform in mind?
Nope. Just want to get started
I prefer to learn things...
The easy way
The best way
The slightly harder way
Auto or Manual car?
Auto
Manual
Which platform/field?
I want to work for big tech companies
Front-end (web interface)
Get a job
Make money
Doesn't matter, I just want $$$
Web
Back-end ("brain" behind a website)
I want to work for...
Corporate
I'm a fan!
Not Bad
Suck
What do you think about Microsoft?
Startup
3D/Gaming
Mobile
Which OS?
IOS
Android
Enterprise
I have a startup idea!
Which platform/field?
YES
NOT SURE
NO
Do you want to try something new with huge potential, but less mature?
Web
Does your web app provides info in real-time, like twitter?
YES
NO
Which one is your favourite toy?
Lego
Play-Doh
I've an old & ugly toy, but i love it so much!
The really hard way (but easier to pick up other languages in the future)

## WHY PYTHON?

IEEE SPECTRUM



| Language Rank | Types | Spectrum Ranking |
|---|---|---|
| 1. Python | | 100.0 |
| 2. C | | 99.7 |
| 3. Java | | 99.5 |
| 4. C++ | | 97.1 |
| 5. C# | | 87.7 |
| 6. R | | 87.7 |
| 7. JavaScript | | 85.6 |
| 8. PHP | | 81.2 |
| 9. Go | | 75.1 |
| 10. Swift | | 73.7 |

# WHY PYTHON?

## C

```
#include

int main(void)
{
    puts("Hello, world!");
}
```

## Java

```
import javax.swing.JFrame;  //Importing class JFrame
import javax.swing.JLabel;  //Importing class JLabel
public class HelloWorld {
    public static void main(String[] args) {
        JFrame frame = new JFrame();           //Creating frame
        frame.setTitle("Hi!");                 //Setting title frame
        frame.add(new JLabel("Hello, world!"));//Adding text to frame
        frame.pack();                          //Setting size to smallest
        frame.setLocationRelativeTo(null);     //Centering frame
        frame.setVisible(true);                //Showing frame
    }
}
```

## Python

```
print("Hello, world!")
```

# LEARNING OBJECTIVES

- ☑ Discover the history of Python and how it compares to other programming languages.
- ❑ Touch on fundamental Python programming techniques and tools.
- ☐ Discuss its applications in data science and the types of problems it can solve.
- ☐ See where Python programming fits into the data workflow.
- ☐ Apply your new skills to solve a real-world problem with Python.

# LEARNING OBJECTIVES

- ☑ Discover the history of Python and how it compares to other programming languages.
- ☑ Touch on fundamental Python programming techniques and tools.
- ☐ Discuss its applications in data science and the types of problems it can solve.
- ☐ See where Python programming fits into the data workflow.
- ☐ Apply your new skills to solve a real-world problem with Python.

## WHAT IS A DATA SCIENTIST?

# MY DEFINITION

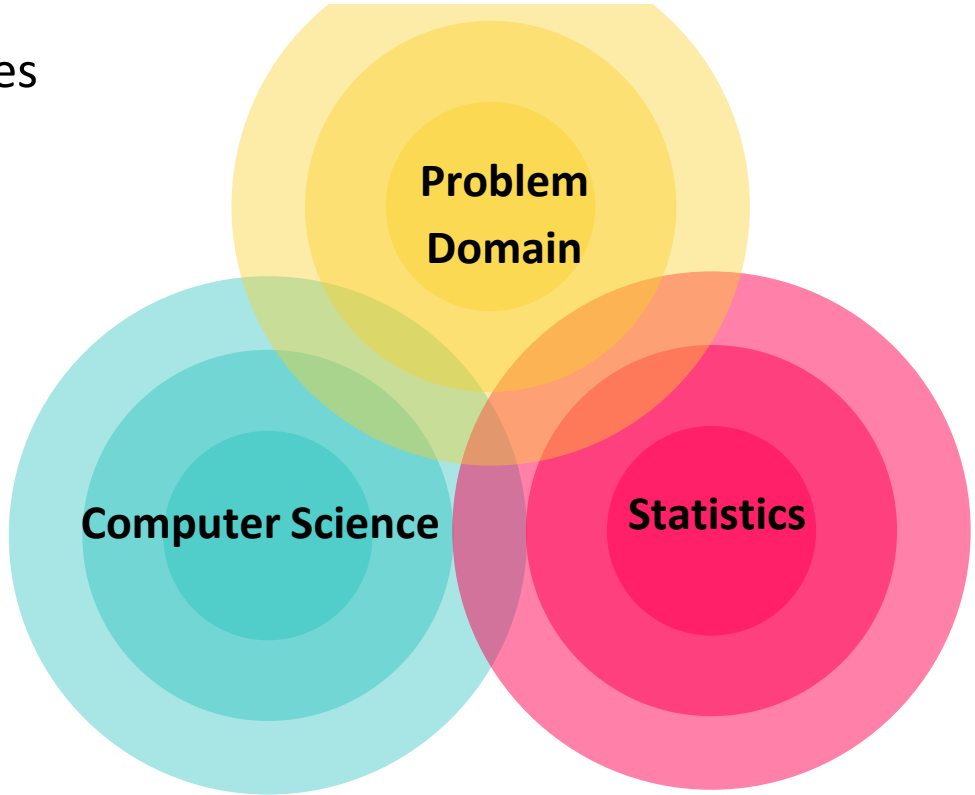
Data Scientists use **data mining** techniques to generate new insights that increase efficiencies / give businesses a competitive advantage

da·ta min·ing

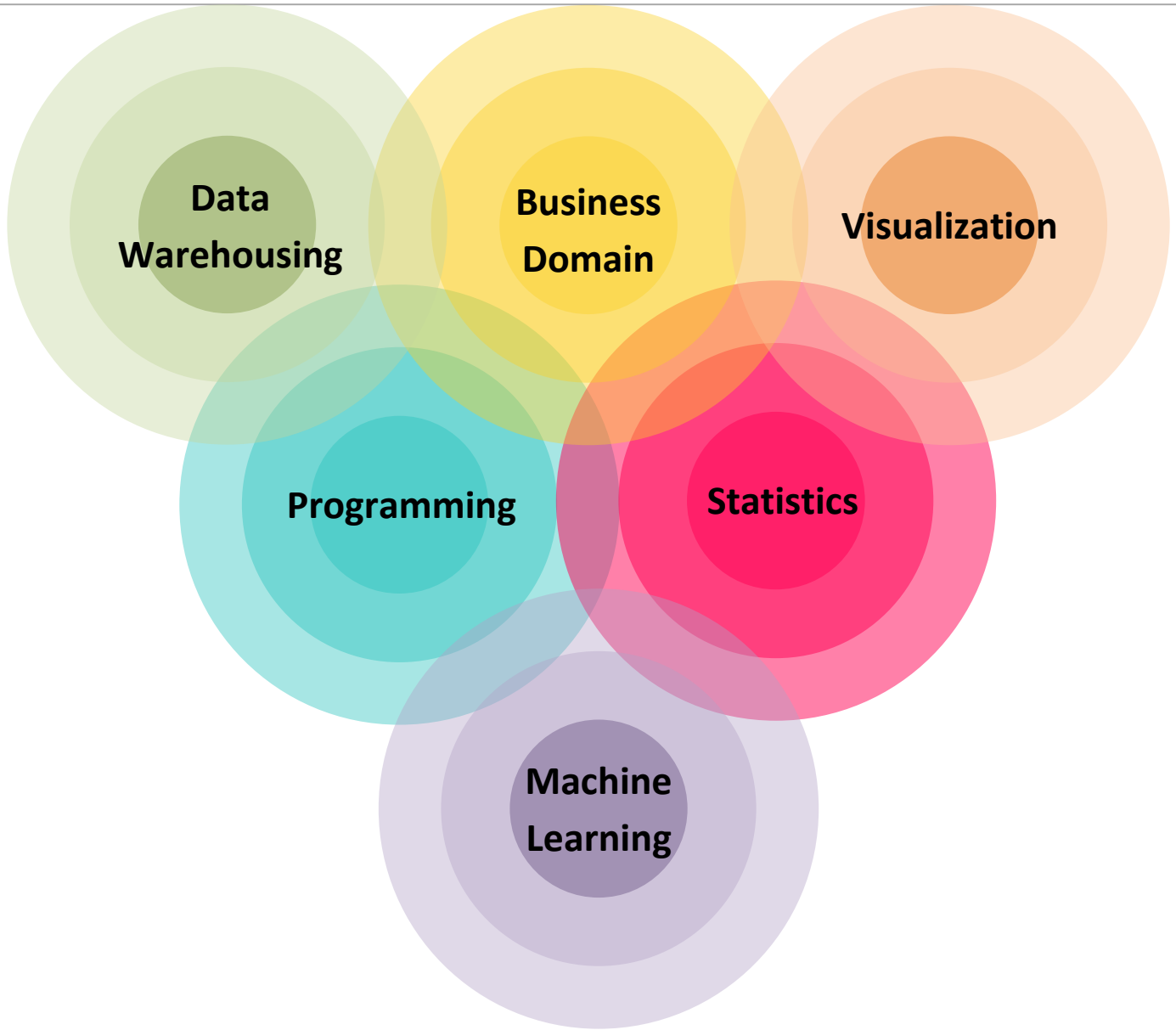
*noun* COMPUTING

the practice of examining large databases in order to generate new information.

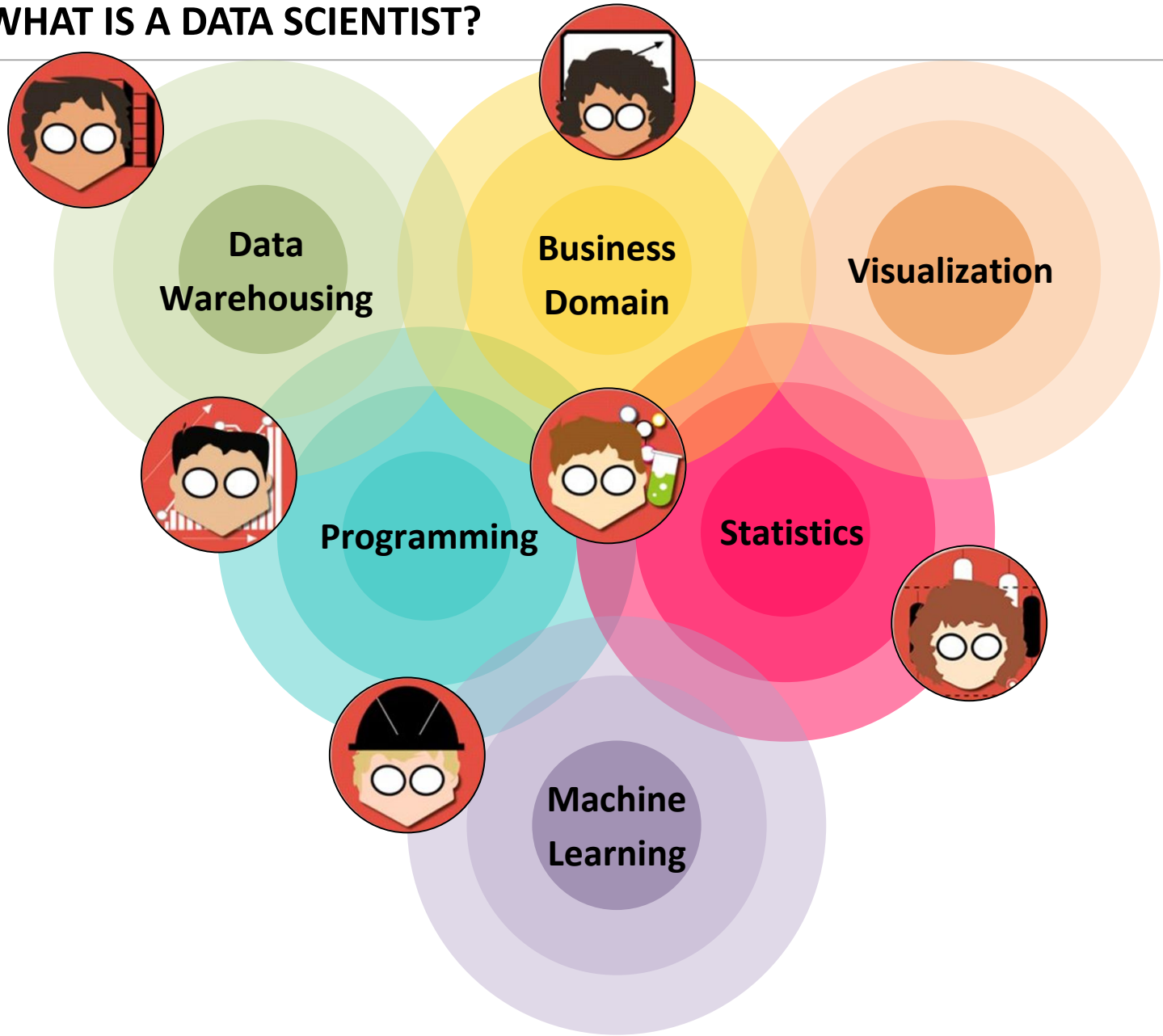
Translations, word origin, and more definitions



# WHAT IS A DATA SCIENTIST?



# WHAT IS A DATA SCIENTIST?

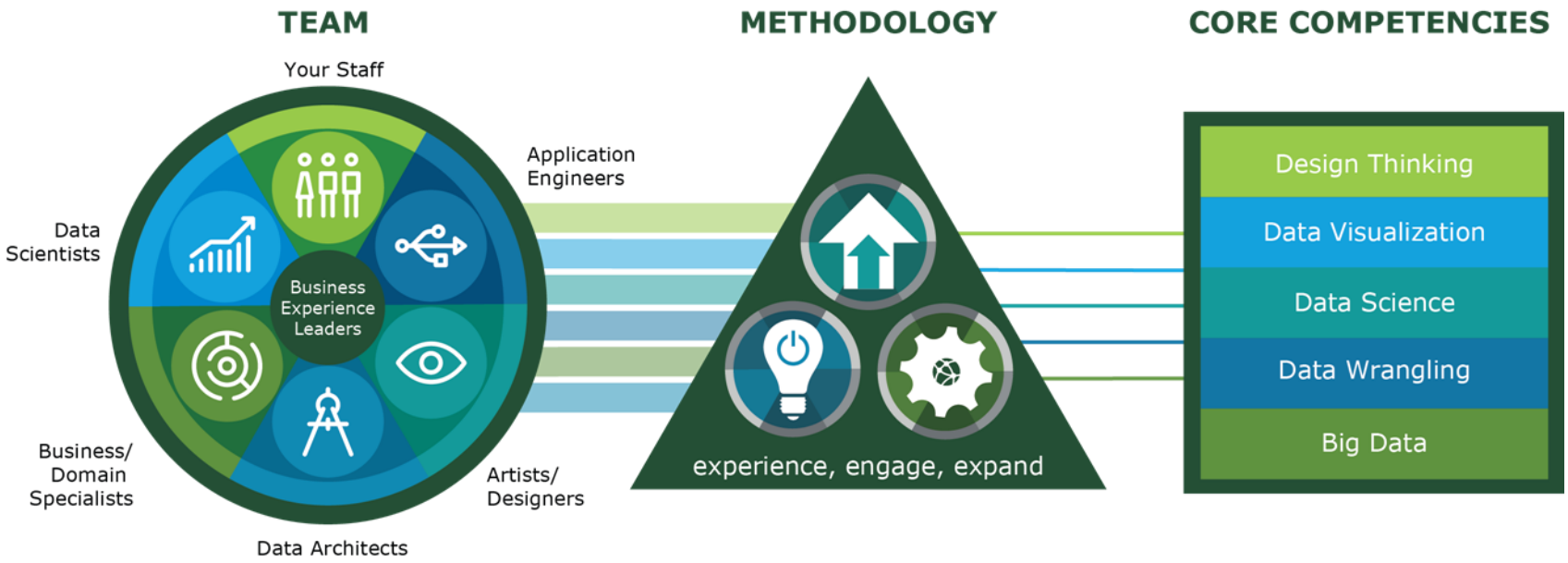


## WHAT IS A DATA SCIENTIST?

# ROLE:

## WHO DO DATA SCIENTISTS WORK WITH?

# TEAM




Deloitte. InsightStudio

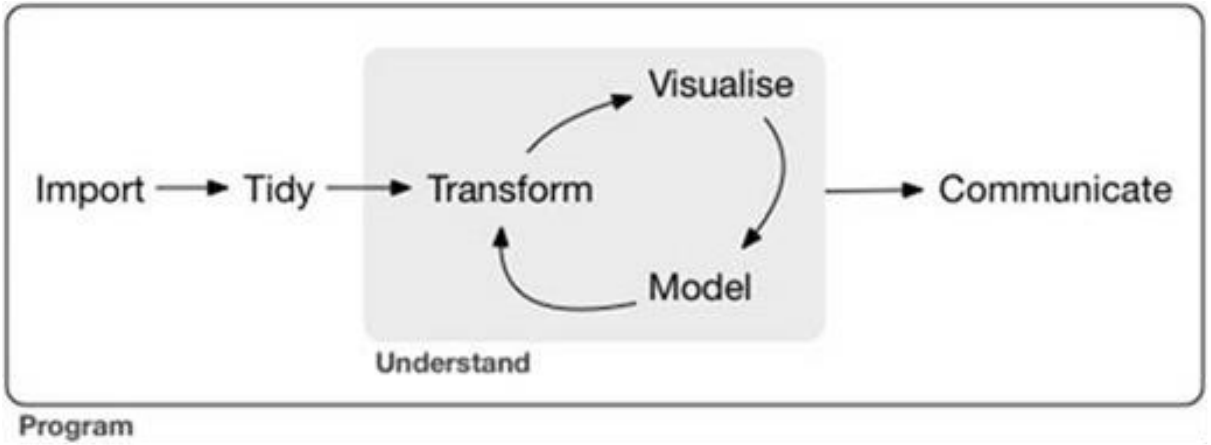
# VALUE ADD

- ☐ Predicting the bad
- ☐ Identifying the good
- ☐ Automating existing processes

| | **Categorical** | **Continuous** |
|---|---|---|
| **Supervised** | Classification | Regression |
| **Unsupervised** | Clustering | Dimension Reduction |

# WORKFLOW

- ☐ Define the problem / question
- ☐ Identify and collect data
- ☐ Explore and prepare data
- ☐ Build and evaluate model
- ☐ Communicate results



## WHAT TOOLS DO DATA SCIENTISTS USE?

# KEY TOOLS

**Querying / Collection**



- Database Queries
- Web Scraping
- API Calls

## WHAT TOOLS DO DATA SCIENTISTS USE?

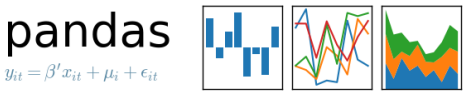
# KEY TOOLS

### Querying / Collection



- Database Queries
- Web Scraping
- API Calls

### Manipulation / Modeling



- Munging and wrangling
- Merging and enhancing
- Building Models

## WHAT TOOLS DO DATA SCIENTISTS USE?

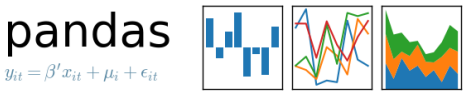
# KEY TOOLS

### Querying / Collection



- Database Queries
- Web Scraping
- API Calls

### Manipulation / Modeling



- Munging and wrangling
- Merging and enhancing
- Building Models

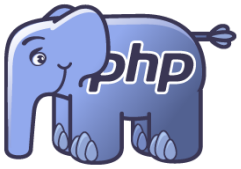
### Exploration / Visualization




- Exploratory analysis
- Plotting and graphing
- Dashboard creation

# OTHER TOOLS

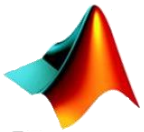
## Querying / Collection



- Database Queries
- Web Scraping
- API Calls

## Manipulation / Modeling




- Munging and wrangling
- Merging and enhancing
- Building Models

## Exploration / Visualization



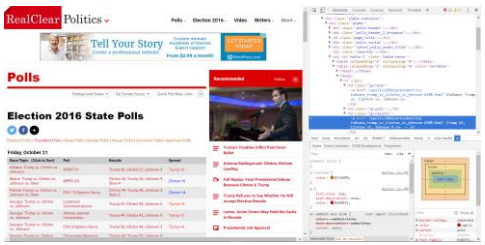



- Exploratory analysis
- Plotting and graphing
- Dashboard creation

## WHAT TOOLS DO DATA SCIENTISTS USE?

# EXAMPLE PROJECT

### Querying / Collection




```
import pandas as pd     # how python creates dataframes
import requests         # how python goes onto the internet!
from bs4 import BeautifulSoup # how python makes sense of html code

r = requests.get('http://www.realclearpolitics.com/epolls/latest_polls/state/')
b = BeautifulSoup(r.text)
```

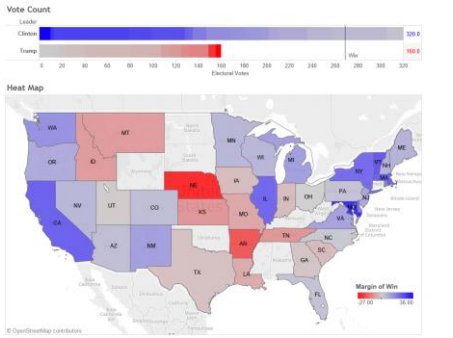
- Web Scraping

### Manipulation / Modeling

```
data = []
for item in results:
    data.append(str(b.findAll('a')[item]))

data_state = []
data_scores = []

for item in data:
    data_state.append(item[32:34].upper())
    data_scores.append(item.split('>')[1].split('<')[0])

data_clinton = []
data_trump = []

for item in data_scores:
    for items in item.split(', '):
        if items.split(' ')[0] == 'Clinton':
            data_clinton.append(items.split(' ')[1])
        elif items.split(' ')[0] == 'Trump':
            data_trump.append(items.split(' ')[1])
```

- Munging and wrangling

### Exploration / Visualization




- Dashboard creation

# CODING

# Q & A

# LEARNING OBJECTIVES

- ☑ Discover the history of Python and how it compares to other programming languages.
- ☑ Touch on fundamental Python programming techniques and tools.
- ☑ Discuss its applications in data science and the types of problems it can solve.
- ☑ See where Python programming fits into the data workflow.
- ☐ Apply your new skills to solve a real-world problem with Python.

# CODING