INTRO TO PYTHON

WHAT YOU NEED TO KNOW TO GET STARTED

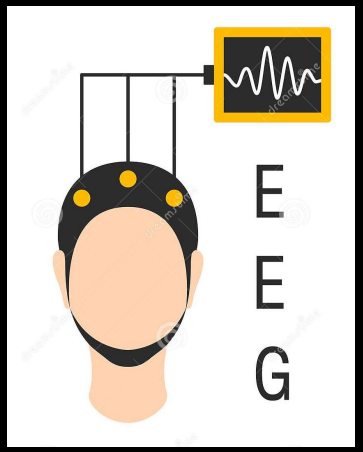
Emily Rae Sabo

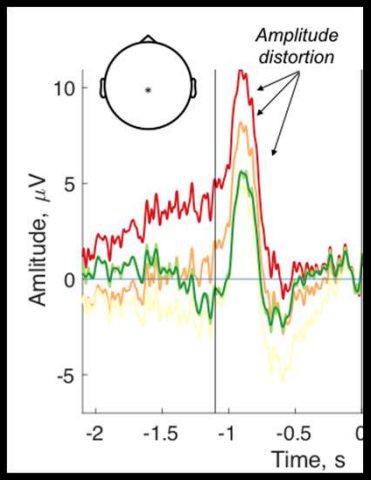
Data Camp | June 17, 2019

ABOUT ME

- PhD student of Linguistics
- There are primarily 3 kinds of data I work with

1. EEG data (brain waves)





ABOUT ME

- PhD student of Linguistics
- There are primarily 3 kinds of data I work with

2. Corpus data (transcripts of speech recordings)

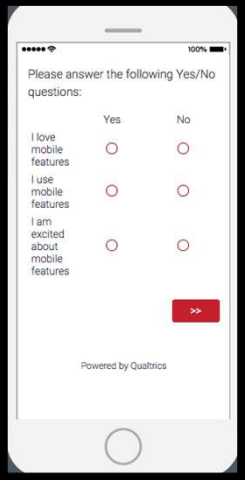
SPKR-01HASLIVED: Ecua-Andes-Imba-OtavaloOutskirt-Pe
QuitoMetro-Central
SPKR-01LIVESNOW: Ecua-Andes-Pich-QuitoMetro-Centra
SPKR-01LING: Biling-SpanDom
METADATA-End
TRANSCRIPT-Start

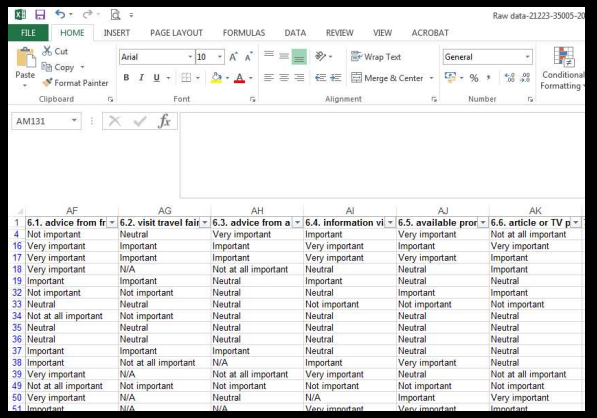
SPKR-01	00:00	*Name redacted* el 2006 en Quito, me fui a vivir--
R	00:11	Wow, hace diez años.
SPKR-01	00:13	Hace diez años, casi diez años. Sí, diez año y fui a vivir allá y ya me quedé estudiando, me quedé trabajando.
R	00:25	¿En dónde trabajaba? ¿O sigues trabajando
SPKR-01	00:28	Ahorita no.
R	00:29	¿Por la maestría?

ABOUT ME

- PhD student of Linguistics
- There are primarily 3 kinds of data I work with

3. Survey response data (excel files)





ABOUT ME

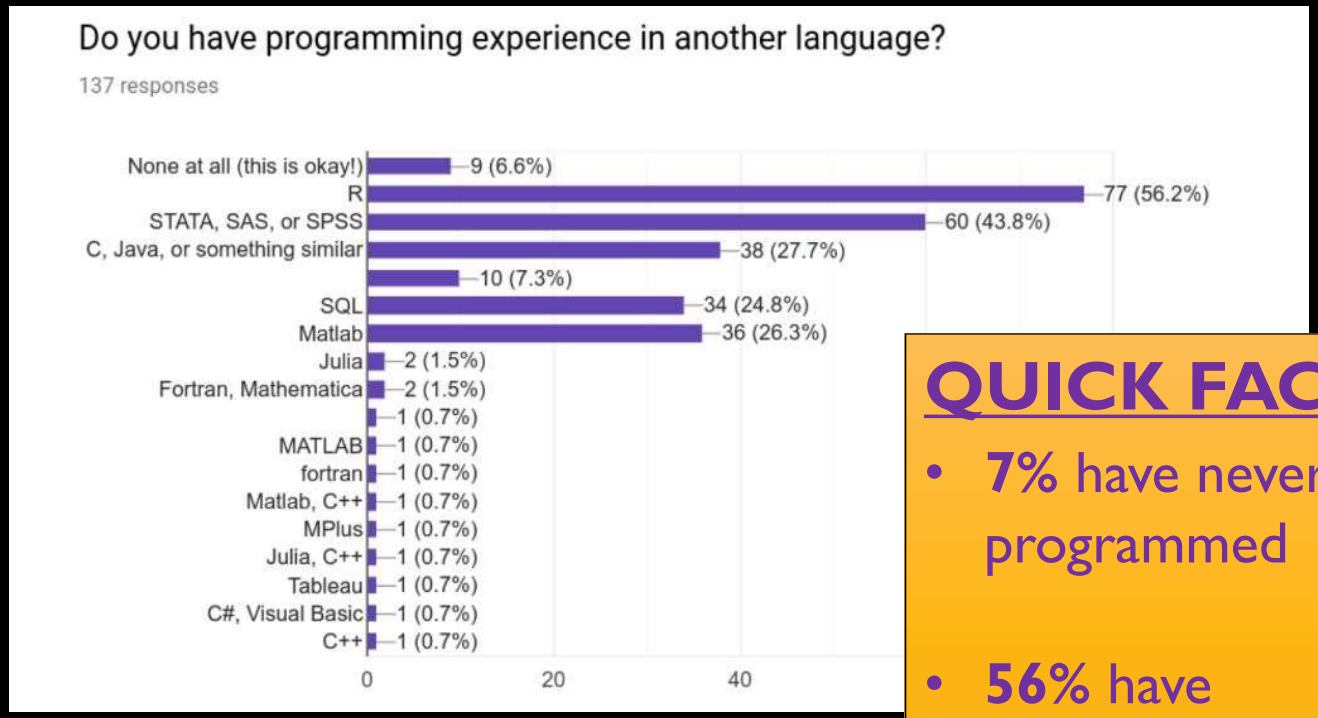
- Several entry points into coding, data analysis, and Python
- Entered this world seeing these skills not as my object of study but rather as a means to a (more efficient) end.

SOLOLEARN
LEARN PLAYING. PLAY LEARNING

coursera

PYTHON
GETTING STARTED

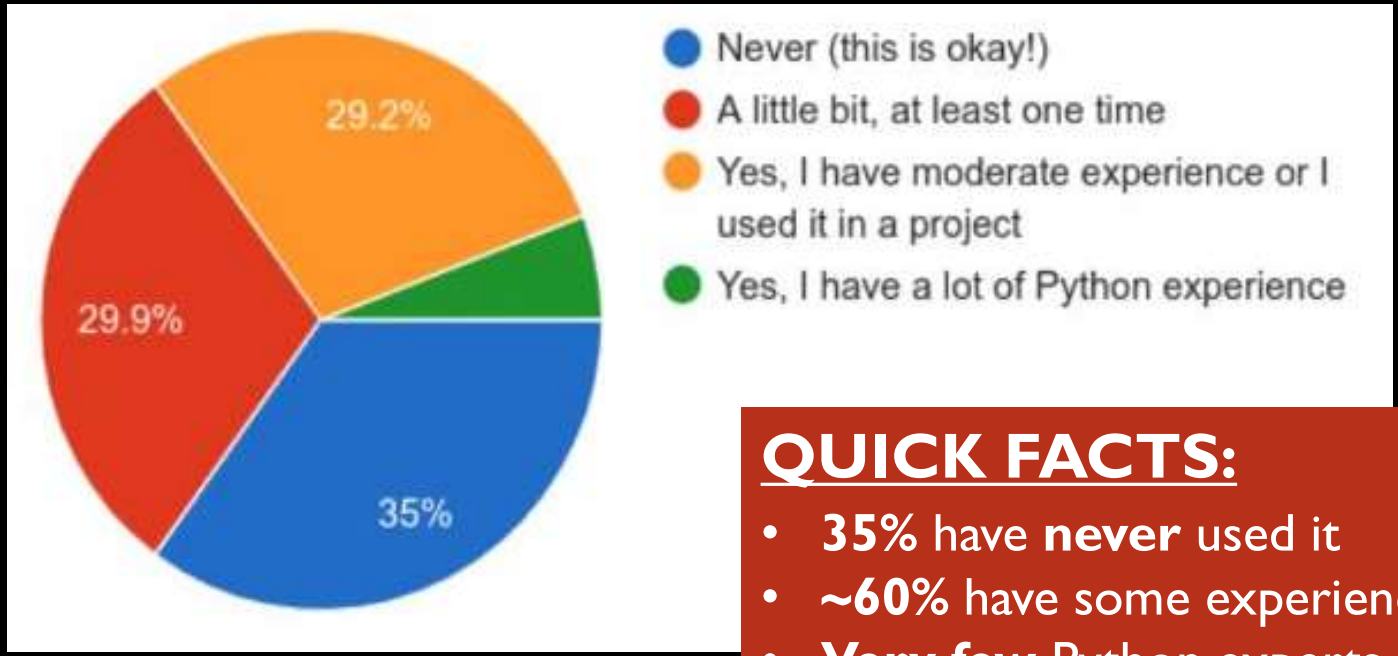
ABOUT YOUR GENERAL PROGRAMMING BACKGROUNDS



QUICK FACTS:

- **7%** have never programmed
- **56%** have programmed in R

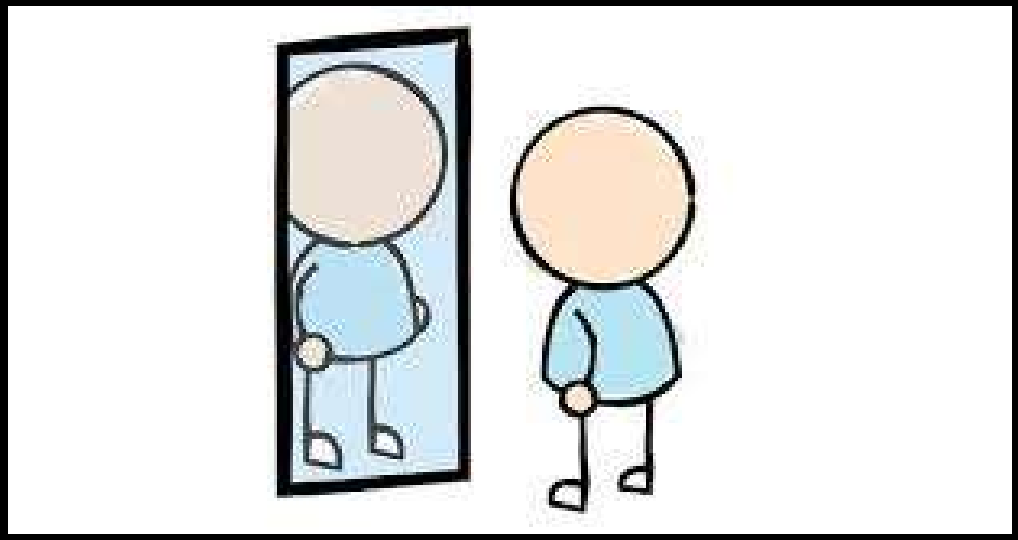
ABOUT YOUR EXPERIENCE WITH PYTHON SPECIFICALLY



QUICK FACTS:

- **35%** have **never** used it
- **~60%** have some experience
- **Very few** Python experts

FIRST, REFLECT ON YOUR OWN…



1. Why are you here today?
2. What most worries & excites you when you hear *big data analysis*?
3. What is one goal you'd like to set for yourself this week?

NOW, SHARE IN GROUPS OF 3...

1. Why are you here today?
2. What most worries & excites you when you hear *big data analysis*?
3. What is one goal you'd like to set for yourself this week?

In this session…

Basic frameworks and best practices for approaching data management, data analysis, and coding in general.

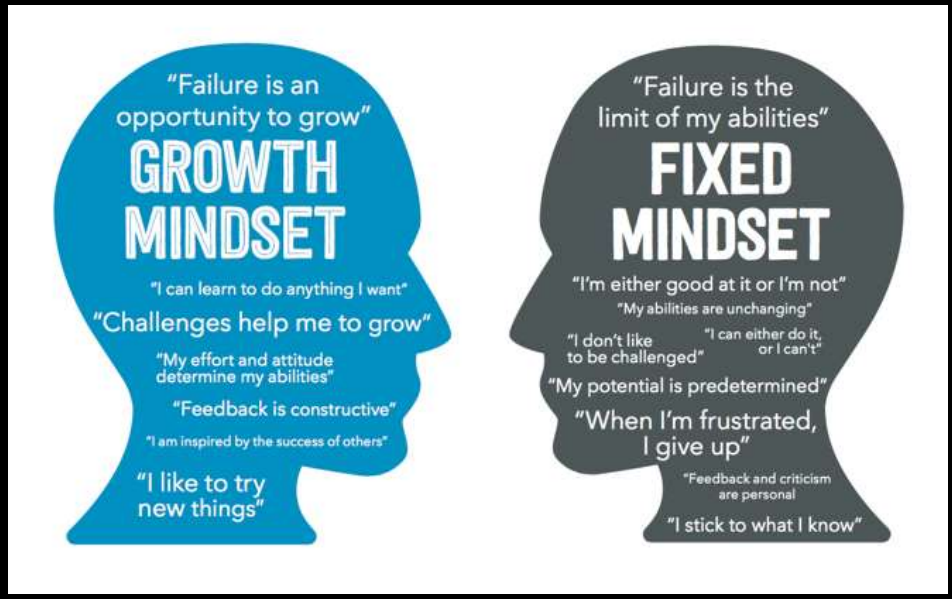
Break down the basic elements of code.
(variables, expressions, data structures, Python syntax)

Jot down new terms (throughout all of DataCamp)

What counts as data?

- This depends largely on your discipline and your research program.
- Don't underestimate the effort that may need to go into transforming your raw data into the form you'll eventually need it to be in.

You'll be the best data analyst and programmer you can be if you adopt a Growth Mindset.



Taking a long time to figure something out doesn't make you dumb. It makes you persistent – and crucially, more likely to get the job done.

Working with big data in Python is a task that's simultaneously creative and computational.

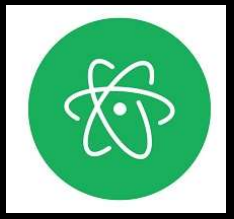
- Creative? → There are so many ways to accomplish any given task.
- Computational? → There are systematic rules & principles to guide your decisions.

- Marr's 3 levels of computation provides a helpful framework for coding, data management & data analysis

Functional level | **Algorithmic level** | **Implementation level**

Workflow can vary wildly when it comes to data management & data analysis.

That is, it's not uncommon for an individual researcher to bounce around between different DBMSs, software programs, and programming languages in a given project…or day.



THE BASIC BUILDING BLOCKS FOR YOUR PYTHON CODE ARE EXPRESSIONS & VARIABLES.

Here's an expression:

```
>>> 2 + 2
4
```

Here are some variables:

```
>>> a = 1 #integer
>>> b = 1.1 #float
>>> c = "cat" #string
>>> d = True  #Boolean (T/F)
```

Here's an expression with variables:

```
>>> a = 1 #define variable
>>> a + 5 #expression adding 5
6
```

1. What outputs do you expect?
2. What is each line of code doing?

```
[1]: a = 1

[2]: print(a)
     ?

[3]: a + 5
[3]: ?

[7]: a = 4

[8]: b = 4/2

[9]: print(b)
     ?
```

1. What output do you expect?
2. What is each line of code doing?

```
[34]: MyLastName = 'Sabo'

[35]: print MyLastName
```

?

1. What output do you expect?
2. What is each line of code doing?

```
[1]: DadsAge = "60"

[2]: MomsAge = "58"

[3]: AgeGap = DadsAge - MomsAge
```

?

THERE ARE 3 BASIC KINDS OF PYTHON CODE.

(1) Sequential code

Run every single line

(2) Conditional code

Run a line if *blah blah blah*

(3) Iterative code

Run a line over and over

1. What output do you expect?
2. What is each line of code doing?

```
[8]: Score = 99

[9]: if Score >= 90:
         print('You got an A!')
     else:print('No A for you.')
```

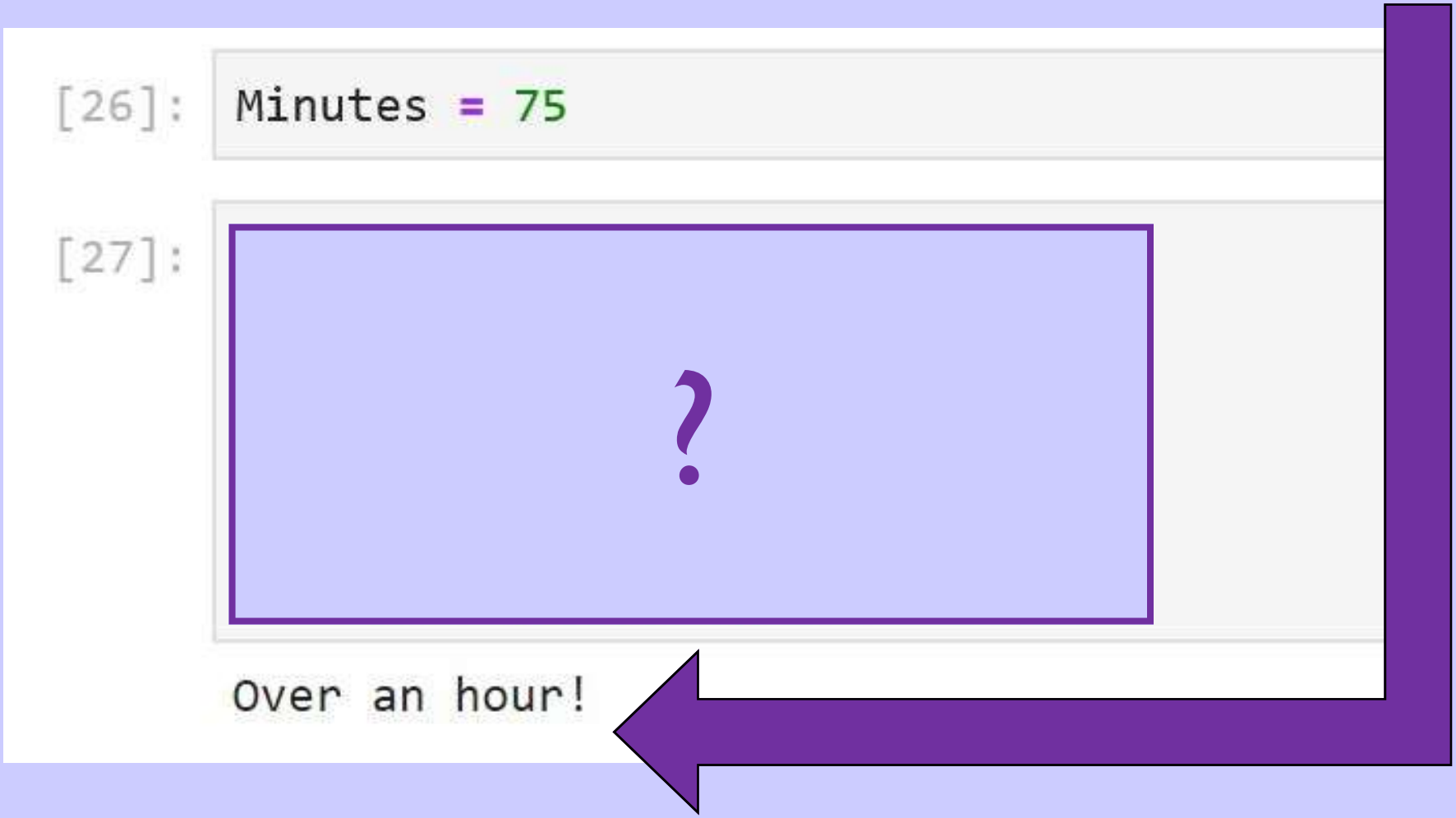
?

1. What output do you expect?
2. What is each line of code doing?

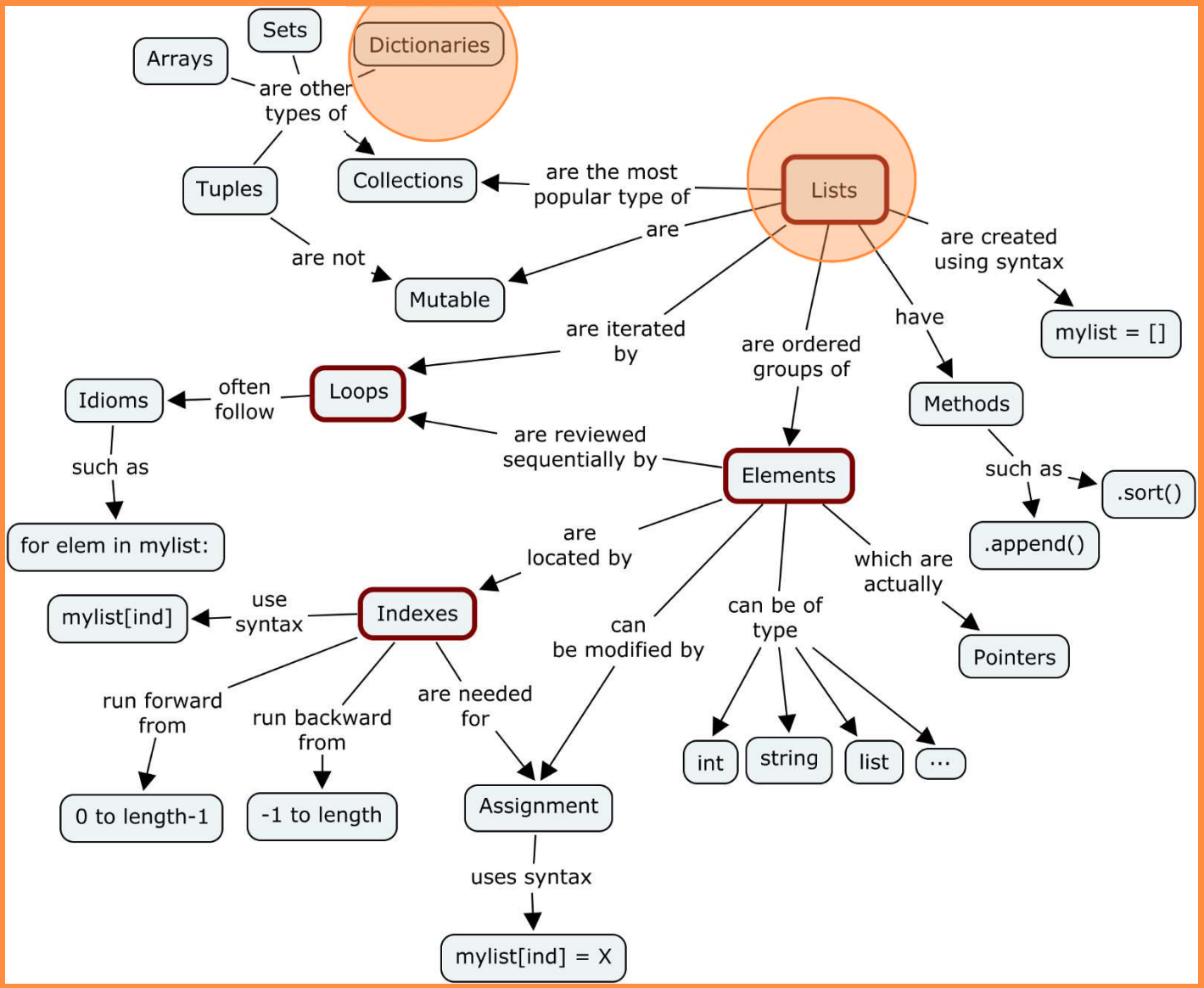
```
[14]: Minutes = 45

[15]: if Minutes >= 60:
          print('Over an hour!')
```

?

Then, how might you write the code do get the desired output?



BASIC DATA STRUCTURES IN PYTHON



IN CONCLUSION

When using Python to work with big data:

1. **Growth mindset** is the key to success.
2. Let **Marr's 3 levels of analysis** guide you.
3. Remember the **2 basic building blocks**.
4. Remember there are **3 kinds of code**.

Have fun! Enjoy getting to be computational & creative.


Thank you!
Emily Rae Sabo
@StandupLinguist
LANGUAGE MATTERS
The University of Michigan