

How To Get Started in Data Science

ABIGAIL HADDAD



Learn skills



Do projects



Be able to talk about them

Who I Am

Kenyon Economics

Public Policy Ph.D.

Lead Data Scientist

https://github.com/abigailhaddad

abigail.Haddad@gmail.com



Agenda

What do you do in a 'data science' job?

How much math do you need to know?

Python vs. R

How to learn the basics

Job searching

Preparing for interviews

Writing good-enough code

You will always have to learn new skills

What Do You Do In A 'Data Science' Job?

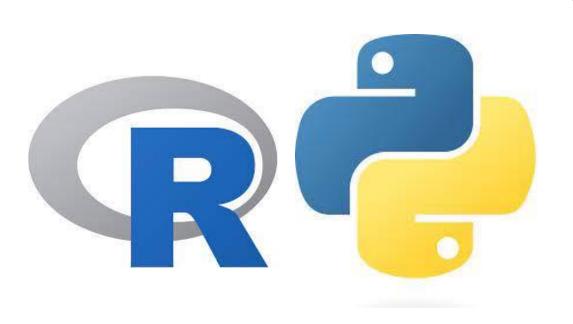
- Getting data
- Cleaning data
- Analysis and modelling
- Visualization
- Managing code
- Talking to people
- Writing



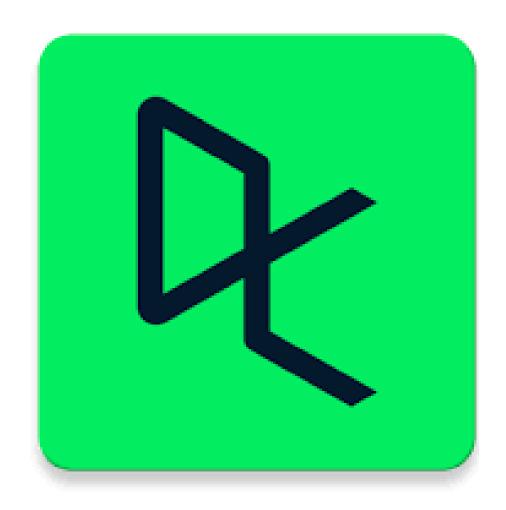
How Much Math Do You Need To Know?

Algebra, statistics, probability – but really well

Should I learn Python or R?



Yes.



How Do I Learn The Basics?

- Boot camps
- Classes
- Certificates
- Books like Learn Python the Hard Way
- Practice with leetcode/hackerrank

Be careful with master's degrees



Do Projects

- -Anything you're interested in
- -Good code
- -GitHub







ACCURATE

TRANSPARENT

REPEATABLE

Good-Enough Code

```
ror_mod = modifier_ob.
  mirror object to mirror
mirror_mod.mirror_object
 peration == "MIRROR_X":
mirror_mod.use_x = True
"Irror_mod.use_y = False
 !rror_mod.use_z = False
 _operation == "MIRROR_Y"
 lrror_mod.use_x = False
 !!rror_mod.use_y = True
 lrror_mod.use_z = False
  operation == "MIRROR_Z";
  rror_mod.use_x = False
  rror_mod.use_y = False
  rror_mod.use_z = True
  election at the end -add
   ob.select= 1
    er ob.select=1
   ntext.scene.objects.action
   "Selected" + str(modified
    rror ob.select = 0
  bpy.context.selected ob
   ata.objects[one.name].se
  int("please select exaction
  --- OPERATOR CLASSES ----
      mirror to the selecter
    ject.mirror mirror x
  ext.active_object is not
```

Practices for Achieving That

Automating everything – doing all the steps in your code

Version control

Documentation

Isolation/containerization

Modular code/packaging your code

Test cases

Coding standards

Code review

Apply to jobs like this early in your career

Qualifications

- Ability to program using SAS, R, Python, or Stata for data management and analysis
- Exceptional problem-solving abilities with an understanding of statistical methods
- Strong interpersonal and communication skills
- Ability to produce careful and detail-oriented work
- Ability to work in a team environment

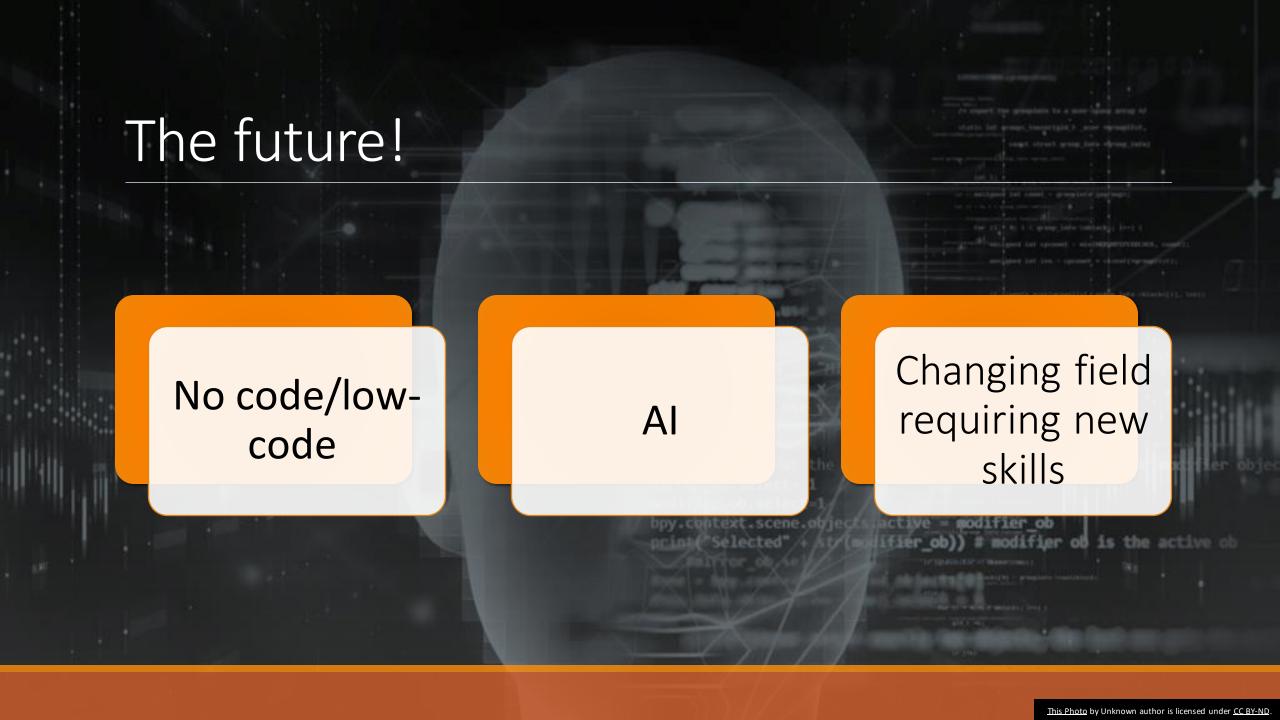
Education Requirements

BA/BS in statistics, biostatistics, economics, mathematics, quantitative social sciences, or related field required; MA/MS preferred.

Interview Failure Modes

Listing skills you don't have

Not being able to talk about what you did





Questions

abigail.Haddad@gmail.com