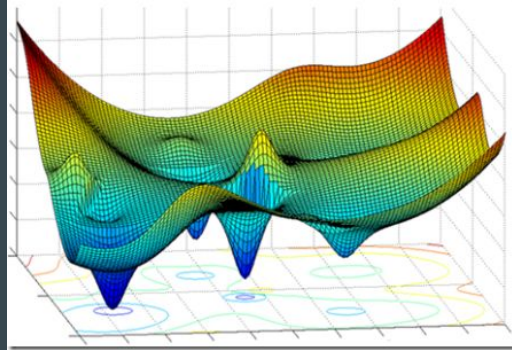


Data + Programing + Math + Curiosity = Data Science



```
31 self.file = None
32 self.fingerprints = set()
33 self.logdupes = True
34 self.debug = debug
35 self.looper = logging.getLogger(__name__)
36
37 if path:
38     self.file = open(path, 'w')
39     self.fingerprints.update(request)
40
41 @classmethod
42 def from_settings(cls, settings):
43     debug = settings.getbool('debug', False)
44     return cls(job_dir(settings), debug)
45
46 def request_send(self, request):
47     fp = self.request_fingerprint(request)
48     if fp in self.fingerprints:
49         return True
50     self.fingerprints.add(fp)
51     if self.file:
52         self.file.write(fp + os.linesep)
53
54 def request_fingerprint(self, request):
55     return request_fingerprint(request)
```



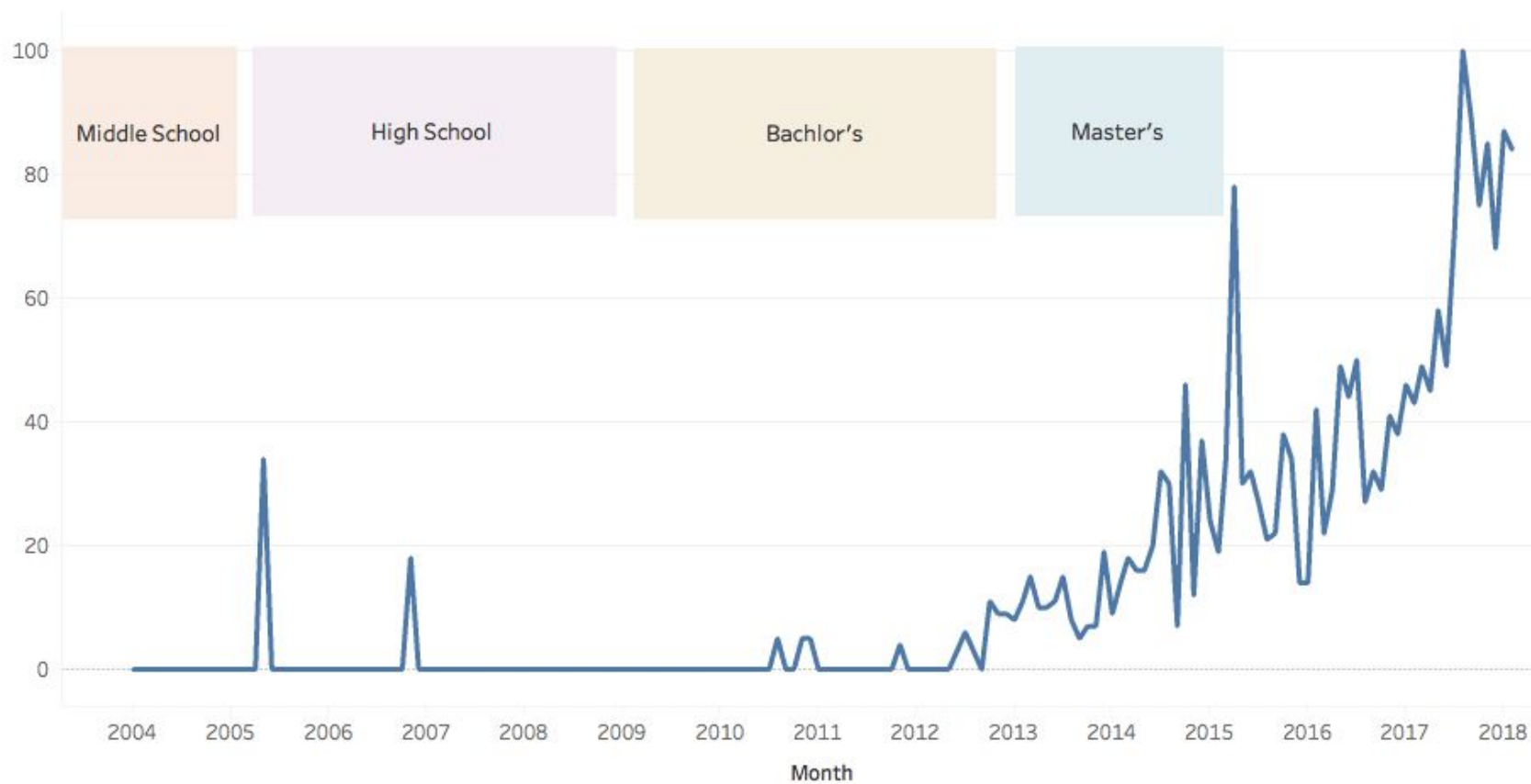
Hanna Landrus Data Scientist

Who Is This Person?

- Grew up in Bend, OR
- Graduated from Pacific University with Bachelor's in Mathematics in 2013
- Graduated from Western Washington University with a Master's in Mathematics in 2015
- Moved to Seattle and have been a data scientist at a medium sized startup for two years
- Outside of data I love to climb, paint, and bake



Interest in Data Scientist vs. My Timeline





What should I do with my life?... Idk I'll just ride this horse.

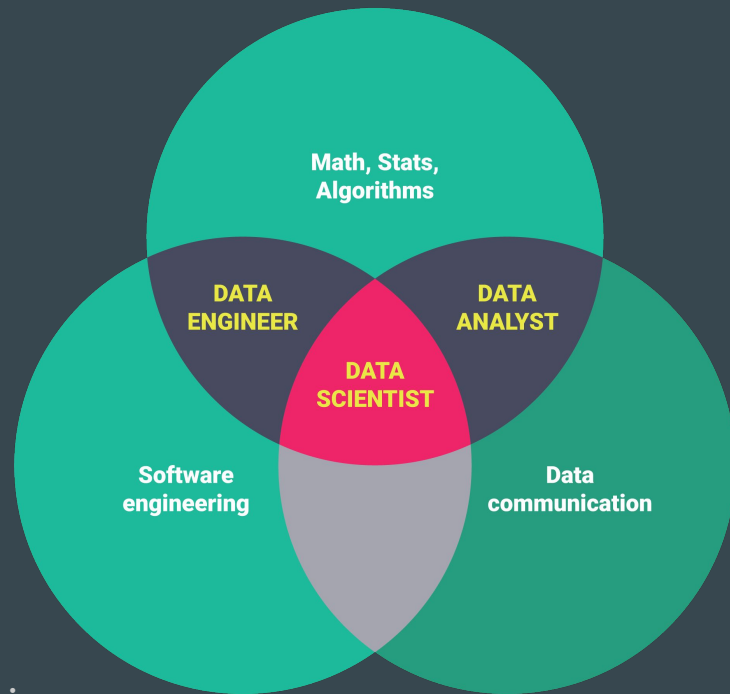
My Strategy: Pursue the things I loved but also consider what is likely to lead to a job that will engage me.

What Is a Data Scientist?

- Computers make it easy to collect lots of data (information stored on a computer)
- Data Scientist use statistics, data analysis and machine learning to provide value from data

Questions they might answer:

- What is a user likely to buy?
- How many songs does a person need to listen to in order to become a regular user?
- Is there a dog in this photo?

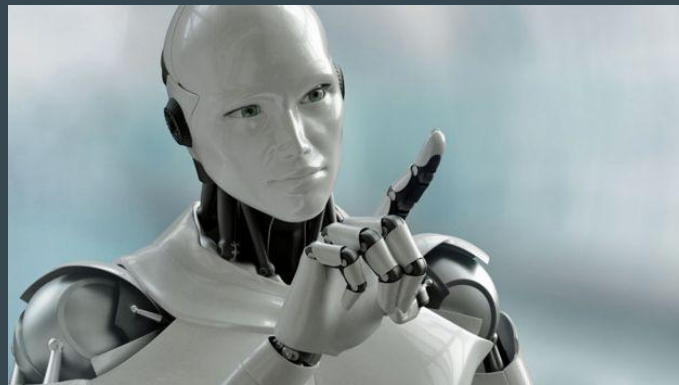


Machine Learning? Like Killer Robots and AI?

Machine Learning is a collection of techniques used to allow computers to learn from data.

Examples:

- Classification: What breed of dog is this?
- Regression: How much will this house sell for?
- Clustering: What types of people use Facebook?
- Sentiment Analysis: Is this a good or bad review?



Machine Learning Example: Classification

Question: Is this dog a chihuahua?

Dog A

Color: Black & Brown

Height: 24 in

Weight: 70 lb



Dog B

Color: Tan & White

Height: 9 in

Weight: 10 lb



Machine Learning Example: Image Recognition

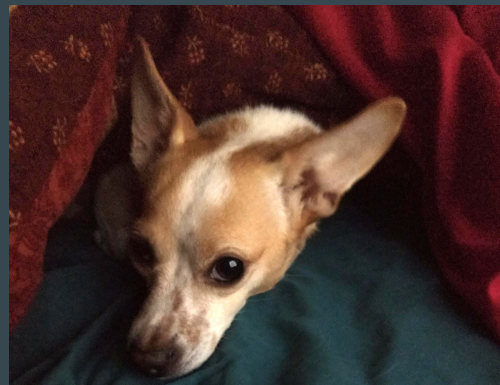
Problem: Have a computer detect what is in a photo.

Method: Show the computer labeled images and have the computer figure out what lines and shapes images with the same label have and use that to predict the label of new photos.

Uses:

- Classify stock photos,
- Predict if tissue sample has cancer
- Find water from satellite photos

Demo: <https://cloud.google.com/vision/>



Dog



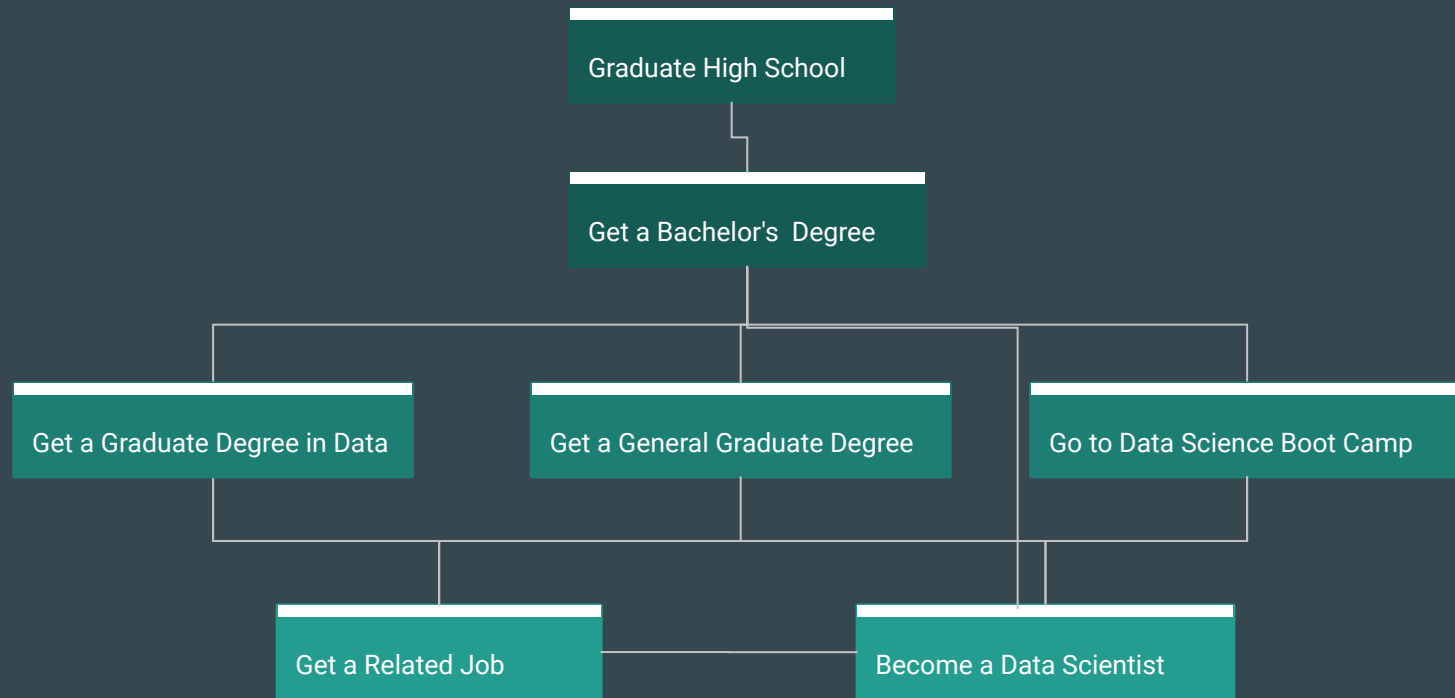
Not Dog

What Does a Data Scientist Do?

- Day to Day:
 - Clean and prepare data
 - Use statistics to analyze data
 - Develop and build machine learning models
 - Meet to share results and plan projects
- Main Tools:
 - Programing Languages: Python, R
 - Used to clean, analyze and build models with data
 - Databases: MySQL, SQL Server, Hadoop
 - Used to store and access data
 - Data Visualization: Tableau, Power BI
 - Used to visualize and sharing data and findings with others
- Demo: Show Notebook

Path To Becoming a Data Scientist

There is not one proven path yet but here are some that I have seen:



What Can You Practically Do Today?

- Learn how to program
- Work hard in math and science courses
 - Try to understand concepts rather than memorize them
 - You don't have to be fast to be good at math
- Work on communicating your thoughts and ideas with people
- Become involved with groups that interest you
- Don't be afraid to try things that are hard and fail at them

Resources - Ask your parents to join

<https://www.seattlecoderdojo.com/> - Weekly group to attend and learn to code

<https://www.codecademy.com/> - Has a Python course plus other courses like web dev

<http://www.r2d3.us/visual-intro-to-machine-learning-part-1/> - Machine Learning visualization

<https://cloud.google.com/vision/> - Image recognition demo

Questions?