

**Programming, Data Science, and using
it to have fun in the real world... :-)**

Who & What is a Data Scientist?

MODERN DATA SCIENTIST

Data Scientist, the sexiest job of 21st century requires a mixture of multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business. Finding a data scientist is hard. Finding people who understand who a data scientist is, is equally hard. So here is a little cheat sheet on who the modern data scientist really is.

MATH & STATISTICS

- ☆ Machine learning
- ☆ Statistical modeling
- ☆ Experiment design
- ☆ Bayesian inference
- ☆ Supervised learning: decision trees, random forests, logistic regression
- ☆ Unsupervised learning: clustering, dimensionality reduction
- ☆ Optimization: gradient descent and variants

PROGRAMMING & DATABASE

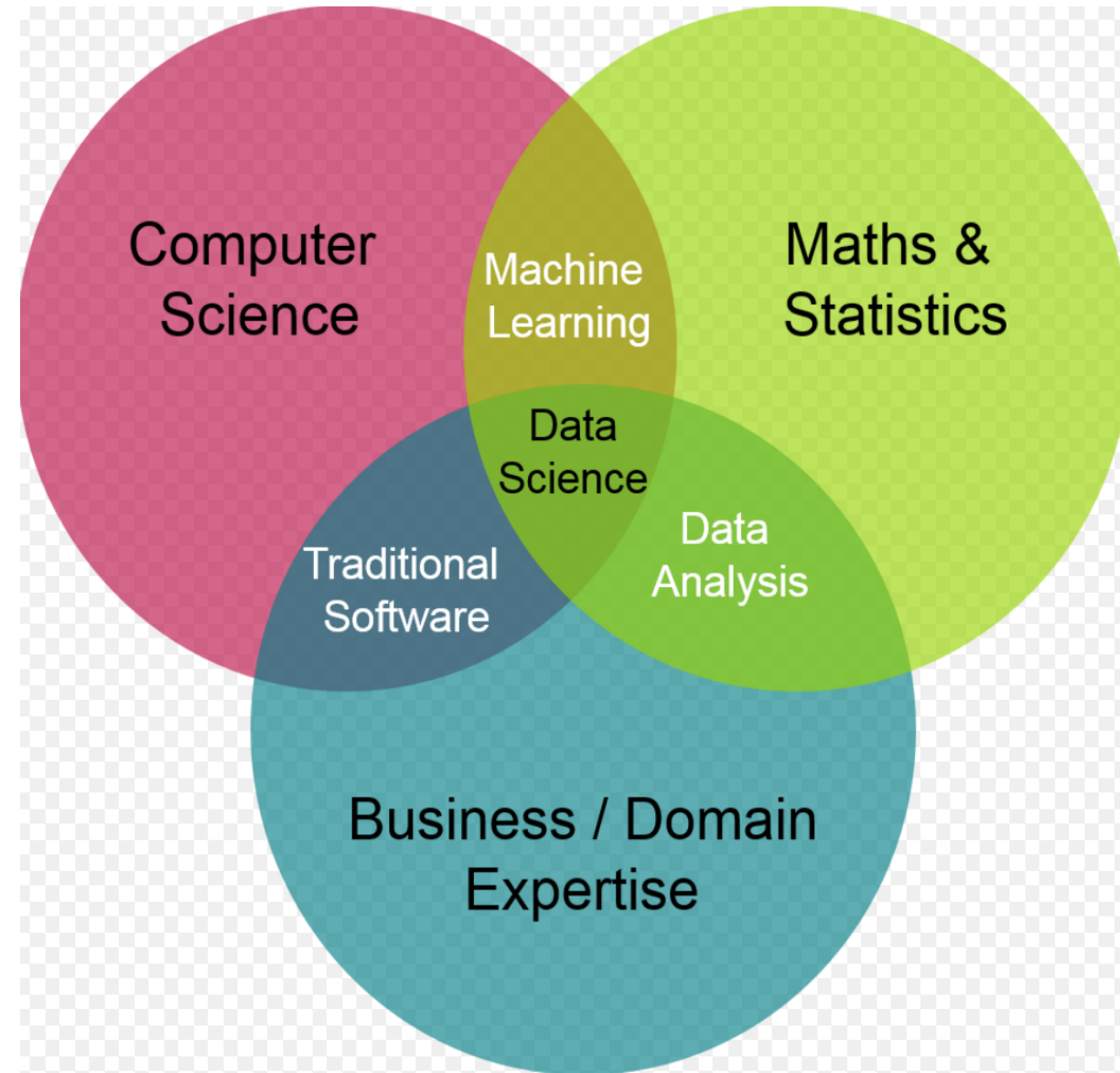
- ☆ Computer science fundamentals
- ☆ Scripting language e.g. Python
- ☆ Statistical computing package e.g. R
- ☆ Databases SQL and NoSQL
- ☆ Relational algebra
- ☆ Parallel databases and parallel query processing
- ☆ MapReduce concepts
- ☆ Hadoop and Hive/Pig
- ☆ Custom reducers
- ☆ Experience with xaaS like AWS

DOMAIN KNOWLEDGE & SOFT SKILLS

- ☆ Passionate about the business
- ☆ Curious about data
- ☆ Influence without authority
- ☆ Hacker mindset
- ☆ Problem solver
- ☆ Strategic, proactive, creative, innovative and collaborative

COMMUNICATION & VISUALIZATION

- ☆ Able to engage with senior management
- ☆ Story telling skills
- ☆ Translate data-driven insights into decisions and actions
- ☆ Visual art design
- ☆ R packages like ggplot or lattice
- ☆ Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau



Job Outlook for Software Developers / Data Scientists

Spoiler alert... good

US Bureau of Labor Statistics

Quick Facts: Software Developers, Quality Assurance Analysts, and Testers	
2021 Median Pay ?	\$109,020 per year \$52.41 per hour
Typical Entry-Level Education ?	Bachelor's degree
Work Experience in a Related Occupation ?	None
On-the-job Training ?	None
Number of Jobs, 2021 ?	1,622,200
Job Outlook, 2021-31 ?	25% (Much faster than average)
Employment Change, 2021-31 ?	411,400

Quick Facts: Data Scientists	
2021 Median Pay ?	\$100,910 per year \$48.52 per hour
Typical Entry-Level Education ?	Bachelor's degree
Work Experience in a Related Occupation ?	None
On-the-job Training ?	None
Number of Jobs, 2021 ?	113,300
Job Outlook, 2021-31 ?	36% (Much faster than average)
Employment Change, 2021-31 ?	40,500

<https://www.bls.gov/ooh/computer-and-information-technology/software-developers.htm>

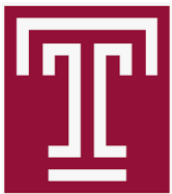
<https://www.bls.gov/ooh/math/data-scientists.htm>

About Me (Ted)

School & Work:



Michigan State



Temple



Hobbies:



Running

+



Programming

+



Family

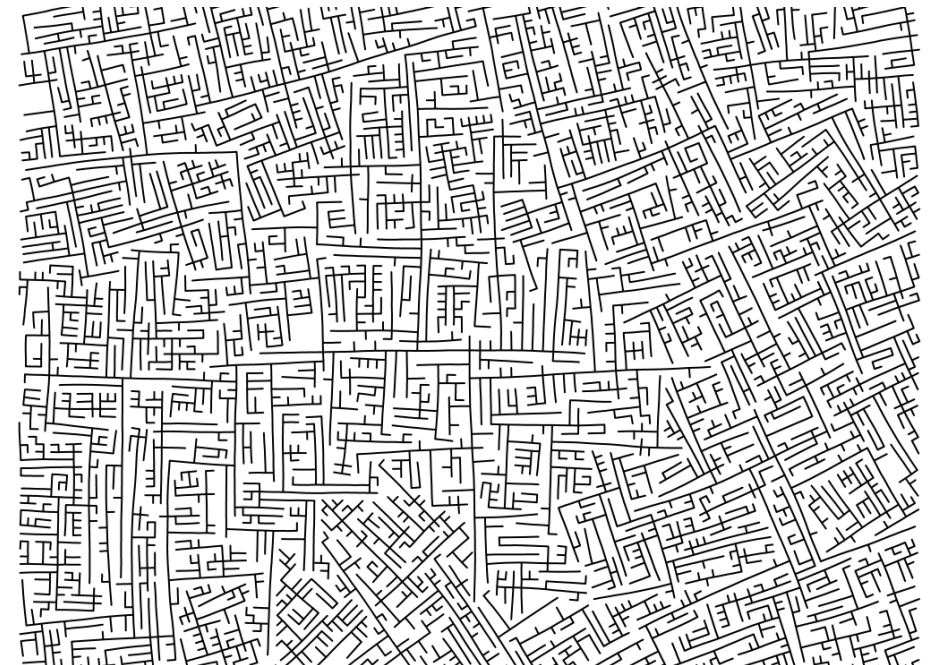
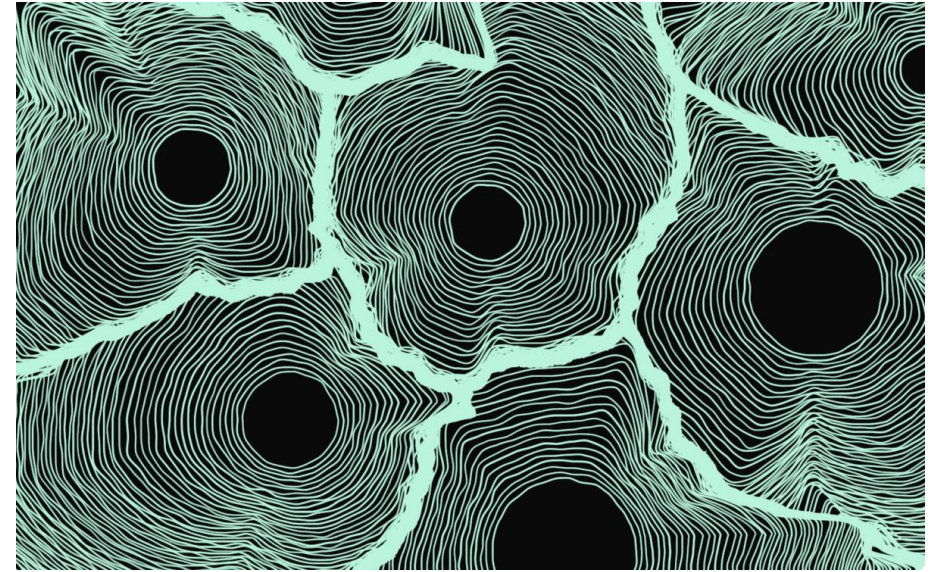
Matt Celaya

- Graduated from MIT with a degree in Economics
- Worked at ON Semiconductor as a data scientist
- Currently working at Intel as a supply chain technical analyst
- Python, R, Javascript, SQL
- I enjoy reading, chess, guitar, hiking, and gaming



Programming for Creativity

- There's a wide range of interesting applications of programming outside of business or academic spaces
- Lots of people use programming to express their creativity
 - Generative art
 - AI generated images and text
 - Data viz
 - Games!
- What are some creative projects you're interested in working on?



Tips for creative coding

- Work on projects that interest you
- Understand your own skill-level and set reasonable expectations in deciding on what to build (you don't need to build the next ChatGPT!)
- Don't limit yourself to one language or technology
- Try to understand concepts and patterns you can apply to any language or project
- Have fun! It's cliché but it's true.

Examples

- <https://mc-markov.surge.sh/>
- <https://mc-sir-model.glitch.me/>
- <https://mc-tree-generator.glitch.me/>
- <https://mc-pixel-paint.glitch.me/>
- <https://mc-maze-generator.glitch.me/>

Exploring Data – Prison Statistics

Things I like to do with programming (outside of work, and really inside as well 😊)...

Find something fun that you want to try to technically accomplish, and match it up with data which you can use to accomplish it.

In this case I wanted to draw a choropleth map which moved over time, and then find data to use to accomplish it.

So... lets play

The stuff I am presenting is posted here: <https://github.com/TedBraun08/fun>

This presentation: https://github.com/TedBraun08/fun/blob/main/DV_Slides_May_2023.pptx

Program: https://github.com/TedBraun08/fun/blob/main/PrisonStats_and_Corr.r

Output #1: <https://github.com/TedBraun08/fun/blob/main/az.png>

Output #2: https://github.com/TedBraun08/fun/blob/main/PrisonAdmitPer100k_02.gif

Output #3: https://github.com/TedBraun08/fun/blob/main/States_10.gif

Output #4: https://github.com/TedBraun08/fun/blob/main/CorrelationMatrix_01.gif