DATA SCIENCE 101

TAKEAWAYS

We covered a lot today! Here are the top five things to remember:

1. Data science is a set of tools and techniques used to extract useful information from data.
2. A data scientist requires skills in programming, math & statistics, business acumen and communication.
3. The data science workflow takes us from problem identification through to results, requiring data acquisition, parsing, mining and modeling.
4. Open source tools such as Python (among others) enable data scientists to tackle the workflow above.
5. Continuous learning is an important part of the data scientist role.

VOCAB

KEY TERM	DEFINITION
Data Science	*A set of tools and techniques used to extract useful information from data.*
Algorithm	*A set of steps to accomplish a task. In terms of computer science, it is a formal way of describing very precisely how to carry out certain computational tasks.*
Machine Learning	*A branch of artificial intelligence concerned with the construction and study of systems that can learn from data.*
Mathematical Modeling	*The description of a system using mathematical concepts and language helping in explaining the system and make predictions.*

RESOURCES

Want to dig deeper? Check out some of these resources:

- Data Analysis with Open Source Tools, P. K. Jannert
- Data Science for Business, F. Provost and T. Fawcett
- Pattern Recognition and Machine Learning, C. Bishop
- Data Science and Analytics with Python, J Rogel-Salazar
- Andrew Ng's Machine Learning Class on Coursera link
- MIT's Artificial Intelligence course link
- DataTau: Like Hacker News, but for data
- MachineLearning on reddit: Very active subreddit
- Quora's Machine Learning section: Lots of interesting Q&A
- KDnuggets: Data mining news, jobs, classes and more
- Data Science courses at General Assembly!