

Data Science and Data Engineering

Data Science for Everyone

Instructor – Raja Iqbal

- CEO & Chief Data Scientist
- <Goodness #1>
- <Goodness #2>
- <Goodness #3>

Instructor – Phuc Duong

- Sr. Data Engineer
- Jack of all trades developer:
 - Full-stack web developer, data engineer, data scientist
- Recovering accountant:
 - Used to perform risk analysis to plan audit hours, financial forecasting, and minimizing cost of wrong predictions
- Loves teaching and learning new technologies.

Instructor – Dave Langer

- VP of Data Science
- 20+ years in technology:
 - Roles in development, architecture, & BI/DW/analytics.
 - Last job – Sr. Director, BI & Analytics @ Microsoft.
- Hooked on Data Science 5 years ago:
 - Extensive background in data and analytics.
 - Learned Machine Learning from 2nd place Netflix Prize winner.
 - #1 Data Scientist on YouTube.
- Joined Data Science Dojo to democratize Data Science.

About Data Science Dojo

- Started in August 2014
- 60+ bootcamps, workshops and corporate trainings
- [2700+ attendees](#)
- [400+ companies](#)
- 10 countries

Common Perception

- 6-12 week bootcamps
- Ph.D. with strong math background
- Mathematics
 - Linear algebra, probability, statistics, optimization, graph theory
- Engineering
 - Programming, distributed and cloud computing

Unlike popular opinion, we
believe that

getting started with data science is easy!

CURRICULUM

Success Metrics

At the end of the bootcamp, **each and every attendee** should be able to:

- Explore an unknown dataset
- Build, evaluate, fine-tune, and deploy a predictive model as a service
- Build an end-to-end big data pipeline to handle large volumes of data – even if the data is arriving in real time

What if I am unable to?

We will work with you to make sure that it happens.

We will need your full commitment for this.

Logistics

- 8:00 am – 6:00 pm daily
- Please make sure that your computers are ready
- Course material
 - Books
 - Git repository.
- Request
 - Keep the session interactive
 - Social media, email...

Post-Bootcamp

- Office hours
 - Schedule to be announced
- Mentored Kaggle project participation
- More topics covered through webinars
- Certification for participation
- LinkedIn group for alumni

Agenda for Today

Before we get started:

- Quiz/Exercise on R fundamentals

Session I: Data exploration, visualization and feature engineering

Session II: Introduction to predictive modeling

Session III: Decision Tree Learning and building your first predictive model

Please Introduce Yourself

- Name.
- Company/Role.
- Analytics/Data Science Background?
- What you are hoping to get out of the week!

Let's Eat!

- We want to get to know each one of you better over dinner:
 - Totally optional.
- Specifics:
 - Due to class size, we split up into groups each evening.
 - We target reservations for 7:30 PM (19:30).
 - We prioritize restaurants within walking distance of the venue.
 - Data Science Dojo picks up the check!
- There will be daily signup sheets:
 - Each day is first-come.
 - Please respect space limitations and signup for only 1 day.
 - Please signup by 1:00 PM each day.
 - Please let us know about dietary restrictions – or suggestions!

R PROGRAMMING QUIZ