Breaking Data Science

Robert J. Brunner
What is Data Science?
What is Data Science?
What is Data Science?

Data science

Data Science is an interdisciplinary field about processes and systems to extract knowledge or insights from large volumes of data in various forms, either structured or unstructured, which is a continuation of some of the data analysis fields such as statistics, data mining and predictive analytics, as well as Knowledge Discovery in Databases (KDD).
What is Data Science?

The Science of *Dealing with Data*, once it has been established.


Astronomy Ph.D
Turing Award
The Problem?

Jobs

Candidates

DATA SCIENTIST
Math
Statistics
Programming
Database
Domain Knowledge
Soft Skills
Communication
Visualization
Data Science: The Topics?

The Six Divisions (David Donoho)

1. Data Exploration and Preparation
2. Data Representation and Transformation
3. Computing with Data
4. Data Modeling
5. Data Visualization and Presentation
6. Science about Data Science

https://dl.dropboxusercontent.com/u/23421017/50YearsDataScience.pdf
Data Science Incubator

Private Sector Program

Innovative Systems Laboratory

Research Park
University of Illinois Urbana Champaign

Data Science, Expert in Residence
@ Research Park, University of Illinois

Sandia National Laboratories

RESEARCH PARK
UNIVERSITY OF ILLINOIS URBANA CHAMPAIGN

CapitalOne

Ameren
Focused Energy. ForLife.

Yahoo!

John Deere

Caterpillar

Grainger®
One Universe

ML + Sims
Image Search with Uncertainty
Highway safety operations model based on Internet of Things

This paper appears in: Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference on, Issue Date: 27-29 May 2011, Written by: Lin, Guoshun; Chen, Yan; Qu, Yong
Takeaway Points

1) Defining curriculum/opportunities
2) Data Science Incubator

bigdog@illinois.edu