the CENTER for PUBLIC HUMANITI

Teachers as Scholars:
Data Science in the Humanities Classroom

Summer 2019

Jason Renn

MESSIAH

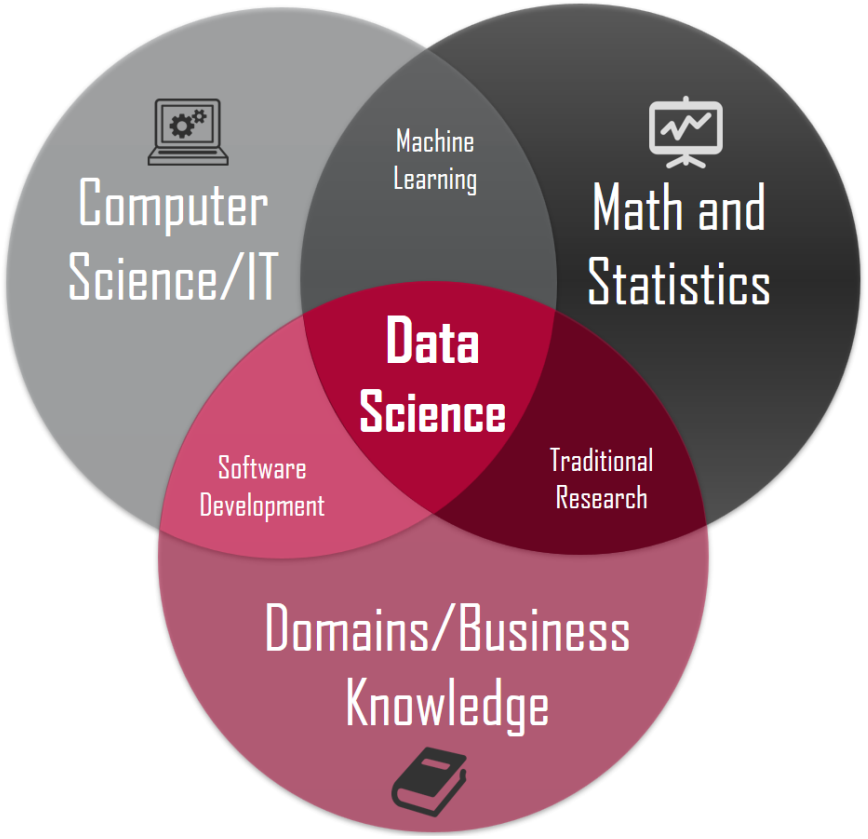
Day 1 (June 11th)

0900	Greetings and Introduction
0915-1030	Session 1: What is Data Science?
1030-1045	Coffee Break
1045-1200	Session 2: Data in News and Professional Settings – Providing for Peace
1200-1300	Lunch
1300-1345	Session 3: Quantitative Learning with Less (But Still Some) Math!
1345-1400	Coffee Break
1400-1500	Session 4: Data in My Classroom Context
1500	Adjourn

Introductions

- Name
- School
- Discipline/Scholarly Interests
- Example of data in your life

What is Data Science?



DATA

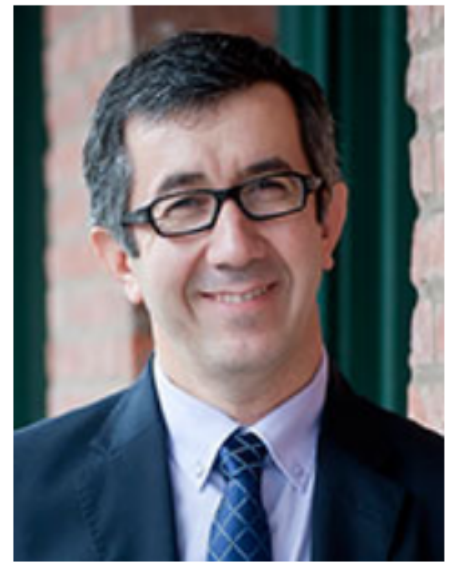
Data Scientist: The Sexiest Job of the 21st Century

The Data Economy: Why do so many analytics projects fail?

July/August 2014

Key considerations for deep analytics on big data, learning and insights.

By (l-r) Haluk Demirkan and Bulent Dal



What is big data? Big data, which means many things to many people, is not a new technological fad. In addition to providing innovative solutions and operational insights to enduring challenges and opportunities, big data with deep analytics instigate new ways to transform processes, organizations, entire industries and even society. Pushing the boundaries of deep data analytics uncovers new insights and opportunities, and "big" depends on where you start and how you proceed.

Big data is not just "big." The exponentially growing volume of data is only one of many characteristics that are often associated with big data, such as variety, velocity, veracity and others (the six Vs; see box).

According to Gartner Research, the worldwide market for analytics will remain the top focus for CIOs through 2017 [1]. According to Gartner, more than half of all analytics projects fail because they aren't completed within budget

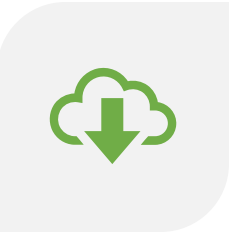
Key Terms

"BIG DATA" - VOLUME, VELOCITY, VARIETY (SCOPE, SPEED, SCALE)

QUANTITATIVE ANALYSIS

SPATIAL DATA

VALIDITY AND RELIABILITY



TEXT AS DATA

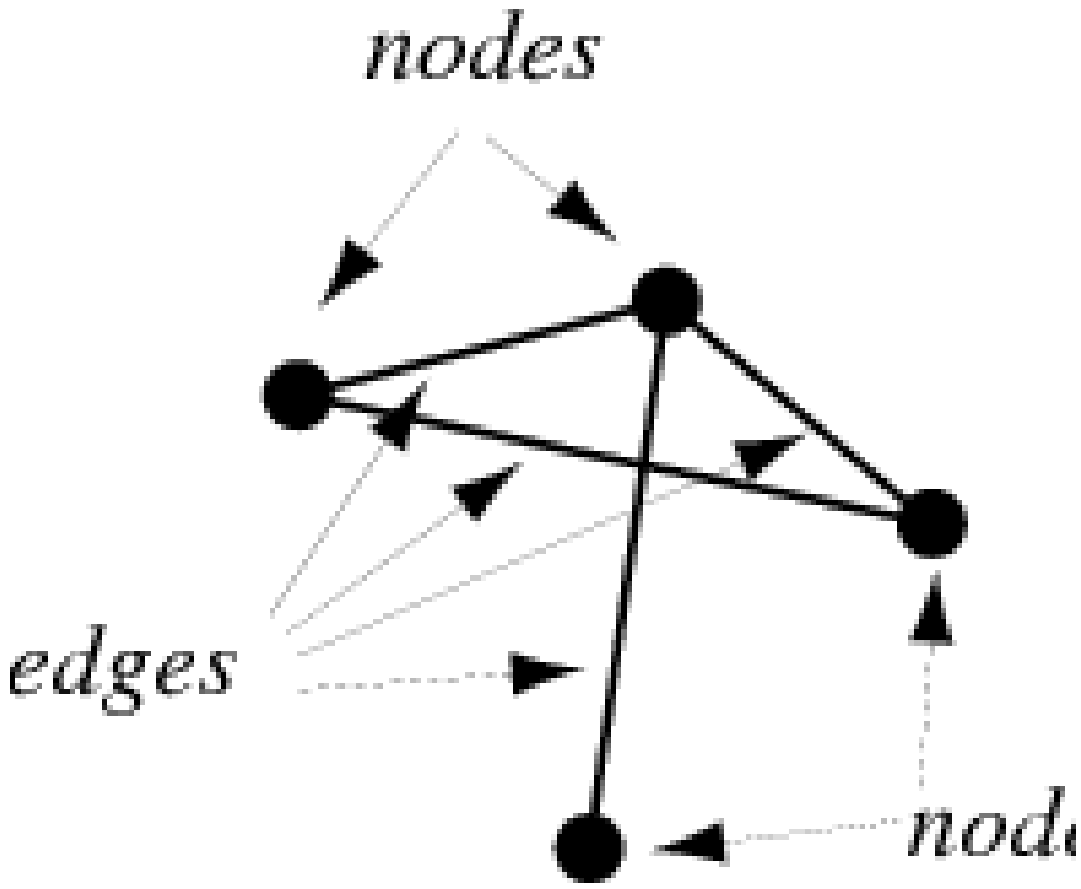
Where do you see data?

- “I think data-scientist is a sexed-up term for a statistician....Statistics is a branch of science. Data scientist is slightly redundant in some way and people shouldn’t berate the term statistician.” - Nate Silver

What we're skipping...

Social network analysis (SNA)

- Bridge
- Isolate
- Connectedness
- Centrality (node)



Application of SNA

- IMDB (https://medium.com/@iggisv9t/what-to-watch-tonight-scraping-imdb-and-visualizing-its-data-as-interactive-website-328a794498a2) - Graph here: http://iggisv9t.xyz/imdb/index.html
- Social Graph (http://curleylab.psych.columbia.edu/friendship1.html)
- Terrorism (https://www.digitaltonto.com/2013/how-the-nsa-uses-social-network-analysis-to-map-terrorist-networks/)
- Gun Homicides (http://www.chicagomag.com/city-life/December-2013/The-Small-Social-Networks-at-the-Heart-of-Chicago-Violence/)

Review

- What is Data Science and why is it important to a broad audience?
- What are the key concepts in Social Network Analysis?
- How do people apply SNA to real world issues?

Find the Data! (The Importance of Measurement)

You shouldn't trust the data until you've proven that it is trustworthy. Until you've got another independent way of backing it up, or you've asked the same question three different ways and you get the same answer three different times. Then you should feel like the data is trustworthy. But until you've understood the process by which the data has been collected and gathered … I think you should be very skeptical. Your default position should be skepticism.

- Hadley Wickham

Session 3: Quantitative Learning

In Words, not Pictures

- GIS Systems (See here and R map example)
- Sentiment Analysis (See here and twitter example)
- Data at Scale (See Cline Center News Archive)

Day 2 (June 13th)

0900	Greet and Group Updates
0915-1015	Session 5: Project 1 Discussion and Workshop for Temporal and Spatial Trends in Terrorism
1015-1030	Coffee Break
1030-1130	Session 6: Project 2 Discussion and Workshop for Text Analysis in Literature and Politics
1130-1230	Lunch
1230-1330	Session 7: Project 3 Discussion and Workshop for Armed Conflict Location & Event Data Project (ACLED)
1330-1345	Coffee Break
1345-1445	Session 8: A Lesson Plan
1445-1500	Assessment Forms and Dismissal