

MIS 0855 Fall 2017 – Data Science *Day 1 – Introduction*

Min-Seok Pang

**Management Information Systems
Fox School of Business, Temple University
minspang@temple.edu**

Aug. 28th, 2017

Who am I?

- Min-Seok Pang, Ph.D.
- minspang@temple.edu / Speakman Hall 201A
- Office hours : Mon, Wed 10 – 11 am or by appointment
- studied at University of Michigan for a doctorate for 5 years
- from Seoul, South Korea
- had worked as a software engineer for 3 yrs

Fighting Crimes with Data (1/2)

- What does she mean by “fighting a crime is all about information?”
 - What information?
 - Specifically, what information did NYPD use to catch “Sugar”?
- What does it mean by “connecting the dots”?



<http://www.tomchaveslaw.com/wrongful-arrest-civil-rights-violations/>

Fighting Crimes with Data (2/2)

- What if NYPD does not have the real-time crime information? How would or should its detectives solve the crimes?
 - Would it be effective?
- Where can NYPD get information they need for crime-solving?



<http://www.tomchaveslaw.com/wrongful-arrest-civil-rights-violations/>

Why Information?

- What kind of information do we need?
- Why do we need information at all?
- Without information, what would happen?



<http://www.blackcareerzone.com/>
<http://www.bankingsense.com/buying-a-house-5-great-tips-for-getting-a-great-deal/>

The Telegraph

Google employee's theory that women are unsuited to tech jobs provokes immediate backlash



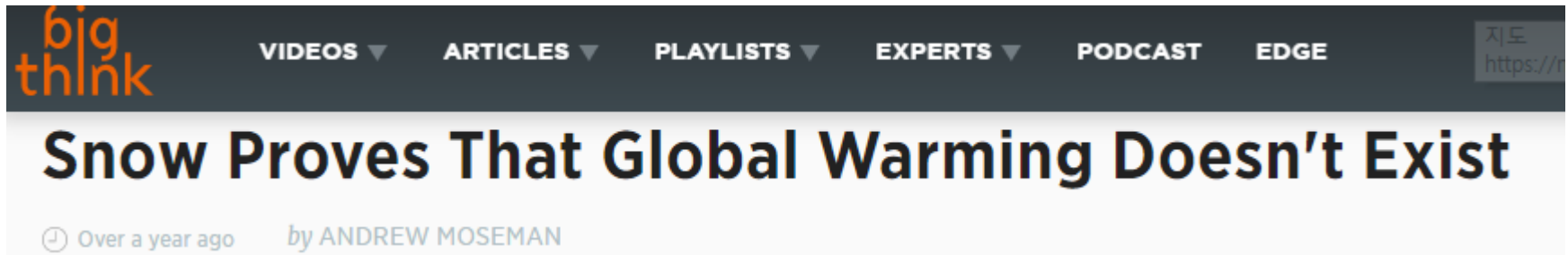
- Openness directed towards feelings and aesthetics rather than ideas. Women generally also have a stronger interest in **people rather than things**, relative to men (also interpreted as **empathizing vs. systemizing**).
 - These two differences in part explain why women relatively prefer jobs in social or artistic areas. More men may like coding because it requires systemizing and even within SWEs, comparatively more women work on front end, which deals with both people and aesthetics.



Former Google staff member Joshua Damore: The Google employee's manifesto has sparked outrage CREDIT: FACEBOOK

<http://www.telegraph.co.uk/news/2017/08/06/google-employees-theory-women-unsuited-tech-jobs-provokes-immediate/>
<https://assets.documentcloud.org/documents/3914586/Googles-Ideological-Echo-Chamber.pdf>

How do we know if global warming exists or not?



big think

VIDEOS ▼ ARTICLES ▼ PLAYLISTS ▼ EXPERTS ▼ PODCAST EDGE

지도
https://

Snow Proves That Global Warming Doesn't Exist

🕒 Over a year ago by ANDREW MOSEMAN

It was a hot few weeks for global warming deniers like George Will and Matt Drudge. That is until a winter storm slammed the East Coast and buried any chance of intelligent conversation in a pile of dirty snow.

George Will trotted out the same shoddy reasoning that conservatives have employed for years to insinuate that global warming is a lie, first igniting a throwdown on the science blogosphere and then attempting to defend himself in a followup. (For a step-by-step analysis of Will's flaws, check out Carl Zimmer). And today, once again confusing weather with climate, Matt Drudge had a laugh at politicians grounded by the massive snowstorm that prevented them from reaching a global warming conference.

Who will win the presidency?

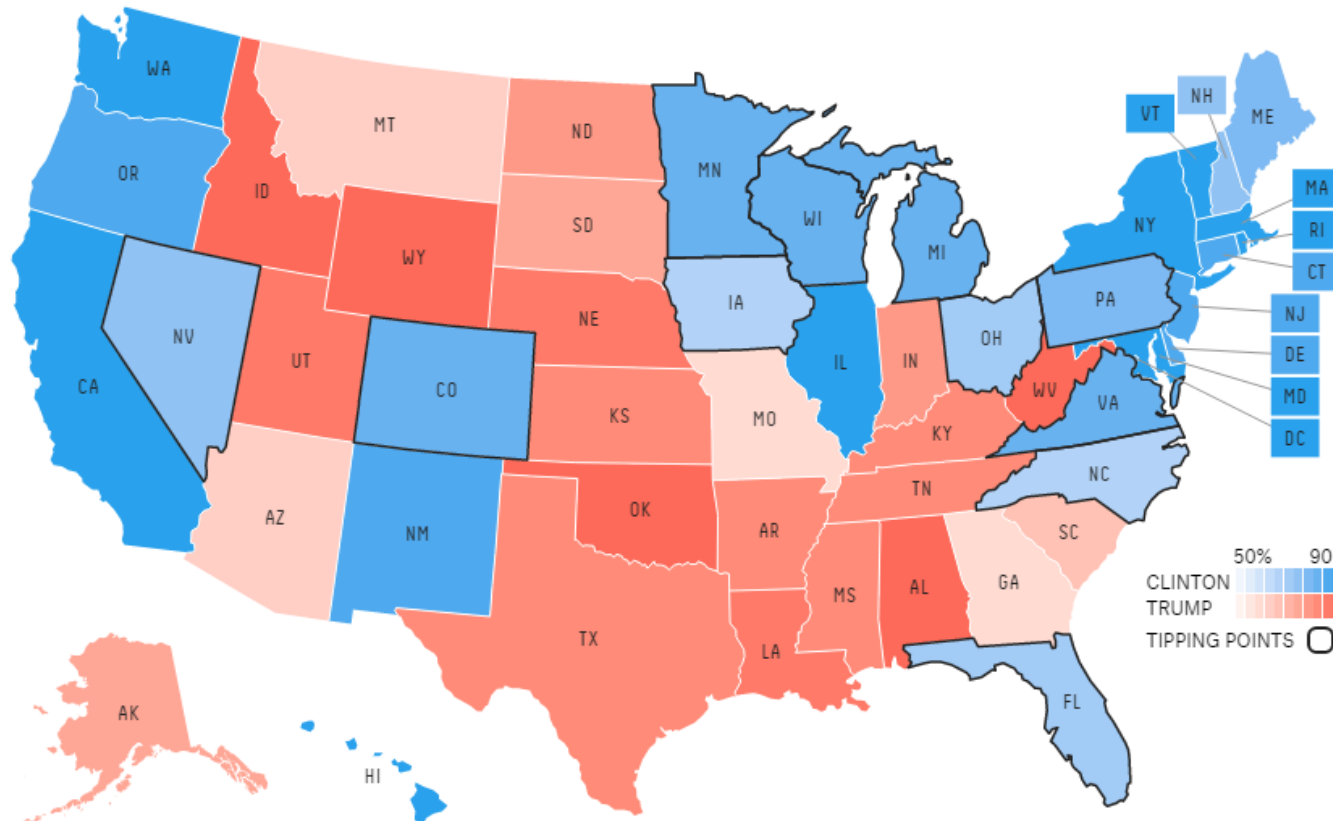


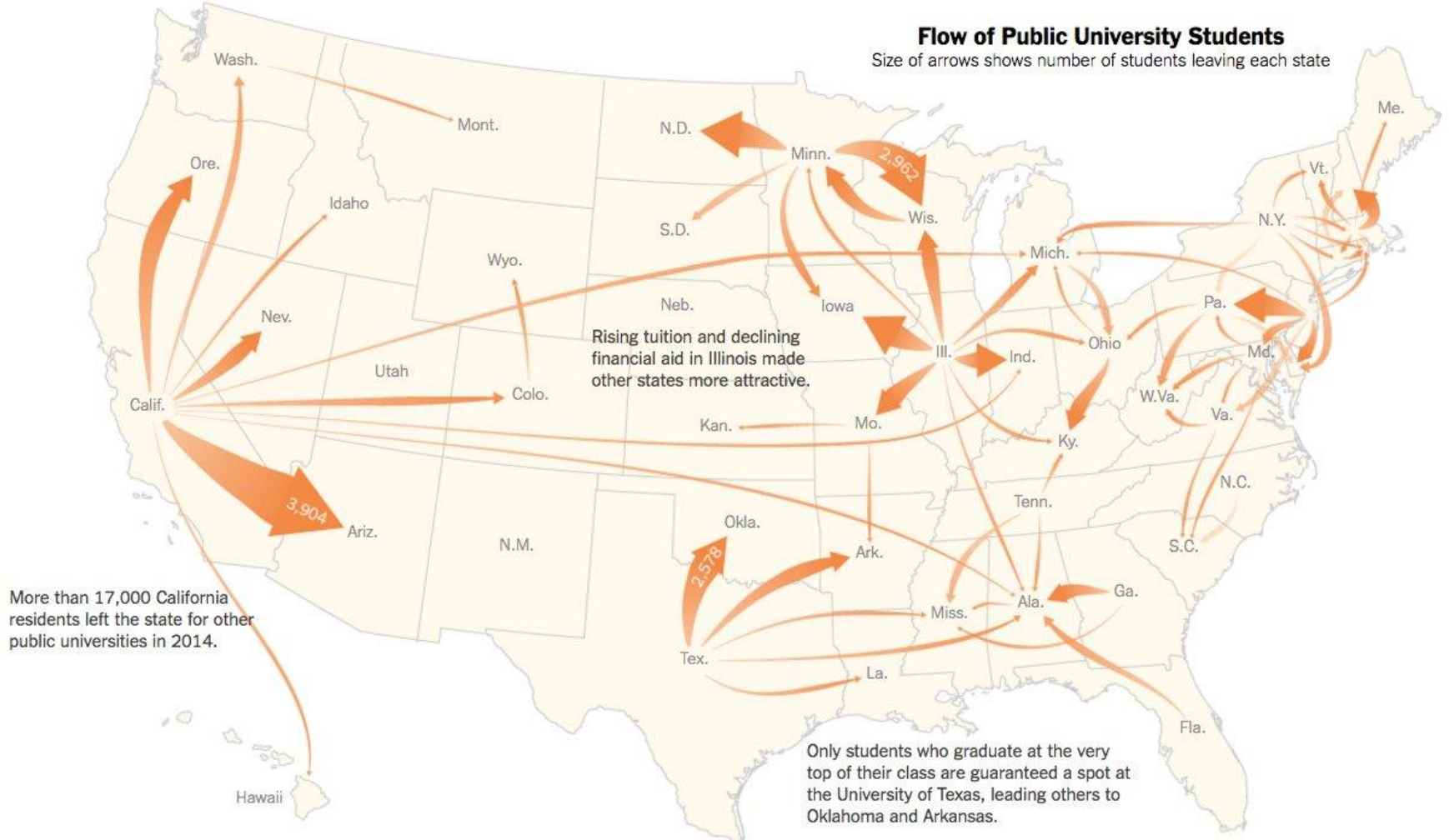
Chance of winning



Hillary Clinton
83.0%

Donald Trump
16.9%

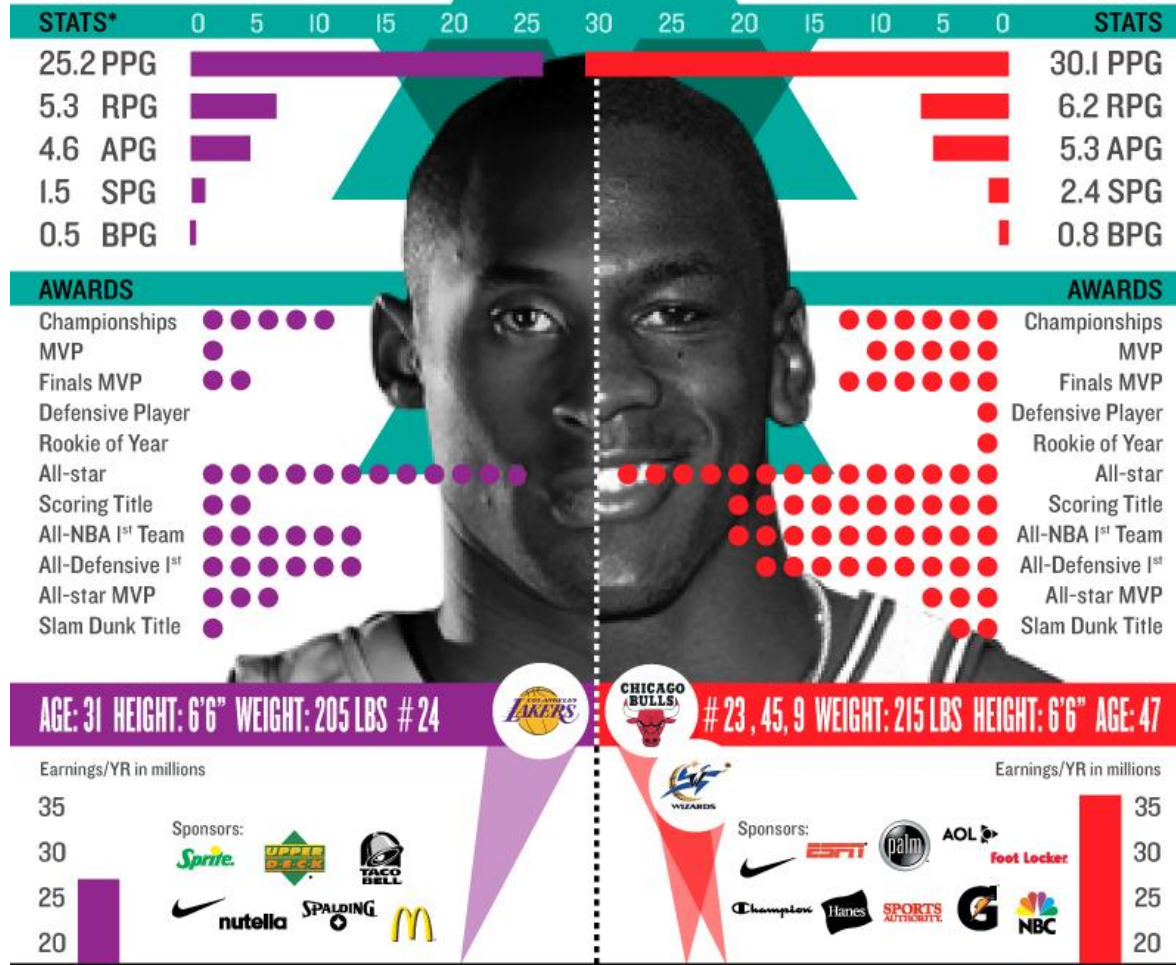




THE GREATEST OF ALL TIME:

KOBE BRYANT vs. MICHAEL JORDAN

In the basketball blogosphere, it's highly debated if Kobe can ever be considered as good as MJ, who is currently considered the greatest of all time. Have a look at how the players compare:



* - ONE AWARD. *BRYANT'S CAREER IS STILL IN PROGRESS. IMAGES: greyflannelauctions.com | nike.com | nba.com | mcdonalds.com | nutellausa.com | spalding.com | upperdeck.com | cocacola.com | facebook.com | aol.com | finashline.com | nbc.com | palm.com | gatorade.com | sportsauthority.com | hanes.com | championusa.com | footlocker.com | espn.com | SOURCES: thehoopdoctors.com | bleacherreport.com | Wikipedia | Forbes. ILLUSTRATION: saravienna.com Information provided by www.medicalinsurance.org

MEDICAL Insurance

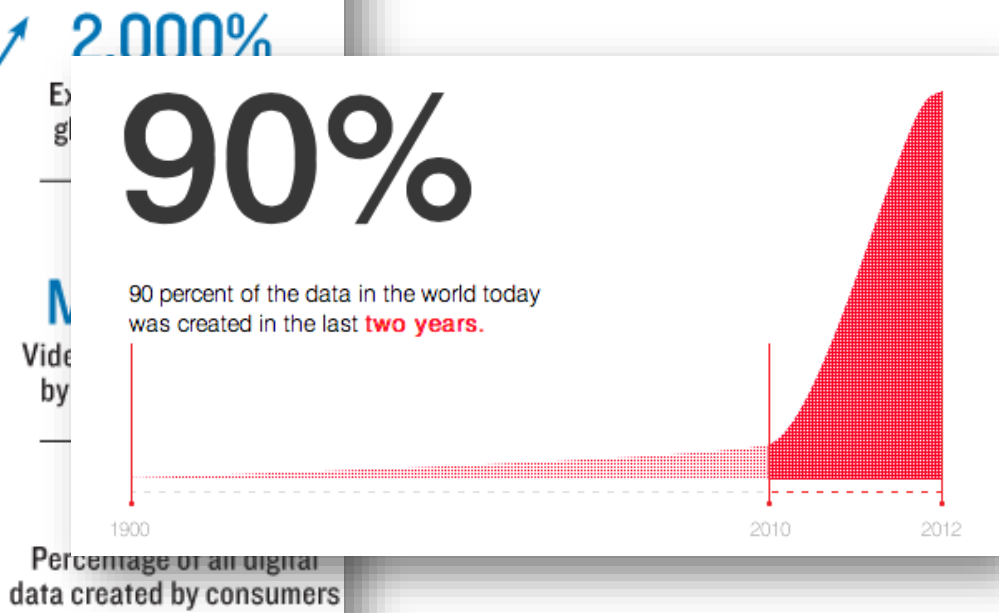
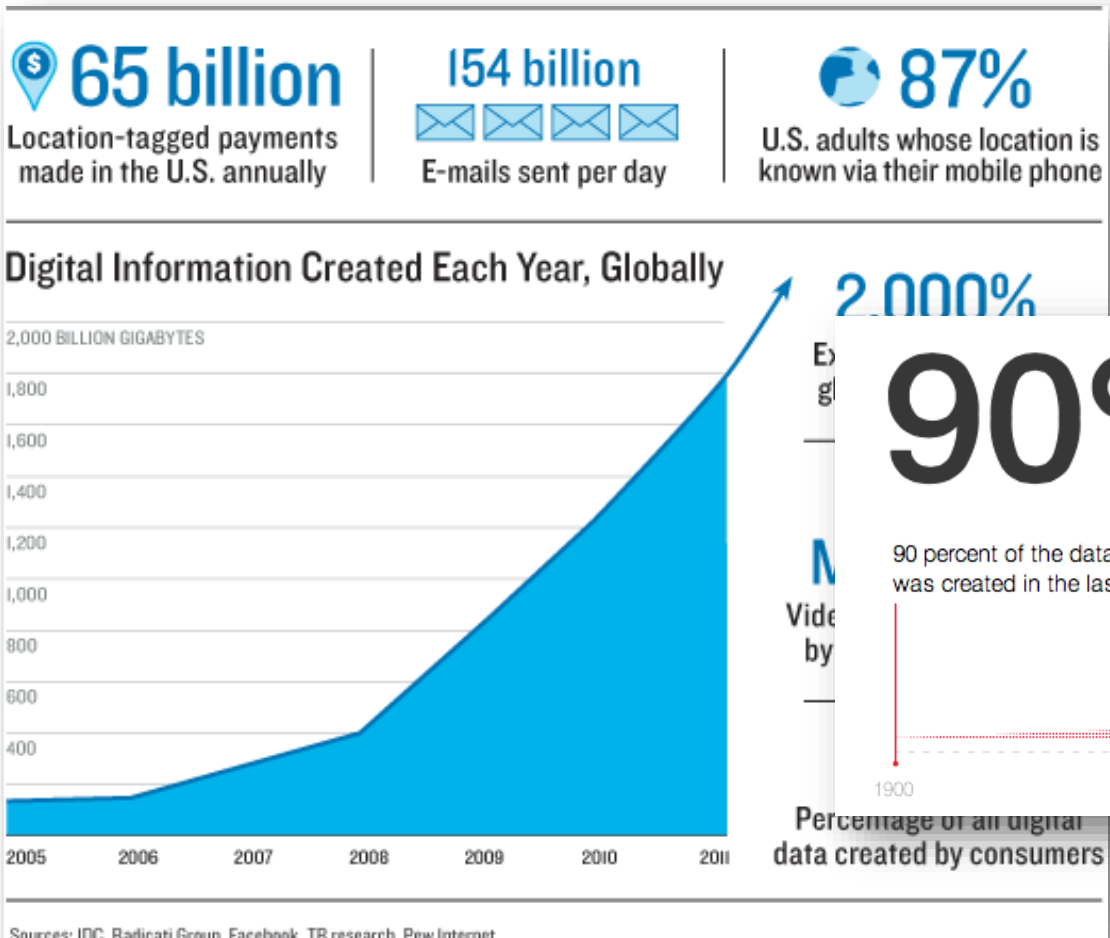
We need information and data

- to make wise, informed decisions,
- to make credible, convincing arguments,
- to make the future (correctly),
- and to deliver your message more effectively.
- anything else?



<https://obamawhitehouse.archives.gov/blog/2016/02/05/open-data-empowering-americans-make-data-driven-decisions>

Data is Exploding!



<http://www.klcommunications.com/blog/anonymity-in-a-big-data-world/>
<https://www.novell.com/events/data-governance-webinars/>

ANALYTICS

Data Scientist: The Sexiest Job of the 21st Century

by **Thomas H. Davenport** and **D.J. Patil**

FROM THE OCTOBER 2012 ISSUE

Calling All Number Crunchers

In a survey, more than 1,000 executives rated obstacles to implementing a big-data strategy

Security 51%



Budget 47%



Lack of talent 41%



Integration with existing systems 35%



Limited options from vendors 33%



DATA: ACCENTURE

<https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century/>
<http://www.bloomberg.com/news/articles/2015-06-04/help-wanted-black-belts-in-data>

Data Scientist

ROLES	Data scientist, business analyst, Chief Data Officer, analytics manager, DevOps
MINIMUM EDUCATION	Master's degree recommended
RELEVANT AREAS OF STUDY	Statistics, analytics, mathematics, computer science, engineering, physics
TECHNOLOGY SKILLS	Data analytics, algorithms, neural networks, machine learning, artificial intelligence
POTENTIAL COMPANIES TO WORK FOR	Academia, healthcare, SMBs, large enterprises, PR/Advertising/Marketing agencies, research firms, technology companies, manufacturing, retail.
NATIONAL MEDIAN SALARY	\$93,146

Big Data, Big Paycheck

Median salary for analytics professionals and those specifically within data science, by level of experience.



Note: Data do not include managers Source: Burtch Works

The Wall Street Journal

<http://www.cio.com/article/3057574/careers-staffing/it-career-roadmap-data-scientist.html>

<http://www.wsj.com/articles/SB10001424052702304819004579489541746990638>

Course Goal (1/2)

- After taking this course, you are expected to be able to
- make sense of the world through data by
 - by selecting the right data,
 - by preparing the data for analysis,
 - by deriving insights and knowledge that are useful for decision making,
 - and by effectively communicating analyses and insights via visualization.

Course Goal (2/2)

- After taking this course, you are expected to be able to
- be able to “do it” by
 - hands-on work with real tools (Tableau Desktop, Piktochart, Excel)
 - retrieving the real-world data and cleaning them
 - analyzing and visualizing the data with tools

Software Tools to Use

- Tableau Desktop 10.3 (A free education license available.)
- Microsoft Excel
- Piktochart (We will use a free version.)
- A basic-level of skills in Excel is desired.

Tableau Desktop 10.3

- You can download and install Tableau from <https://www.tableau.com/products/desktop/download>
- You should request a free one-year license at <https://www.tableau.com/academic/students>
 - Be sure to use your @temple.edu email when filling out the form.

Schedule (1/4, subject to change)

Day	Date		Topic
1	Aug-28	M	Introduction
2	Aug-30	W	Data and Science
3	Sep-01	F	Data, Information, Knowledge
	Sep-04	M	<i>No Class (Labor Day)</i>
4	Sep-06	W	Data and Metadata
5	Sep-08	F	Prep for Assignment #1 (Create a Data Analysis Plan)
6	Sep-11	M	Introduction to Tableau
7	Sep-13	W	Introduction to Tableau
8	Sep-15	F	Introduction to Tableau
9	Sep-18	M	Prep for Assignment #2 (Analyze a Dataset Using Tableau)
10	Sep-20	W	Storytelling with Infographics
11	Sep-22	F	Exam #1 Review
12	Sep-25	M	Exam #1

Schedule (2/4, subject to change)

Day	Date		Topic
13	Sep-27	W	Data Visualization
14	Sep-29	F	Data Distortion
15	Oct-02	M	Integrating Data
16	Oct-04	W	Creating Interactive Dashboards
17	Oct-06	F	Dirty Data
18	Oct-09	M	Data Cleansing
19	Oct-11	W	Prep for Assignment #3 (Cleaning a Dataset)
	Oct-13	F	<i>No Class (Instructor travels for a conference)</i>
20	Oct-16	M	Visualizing Maps
21	Oct-18	W	Visualizing Maps
22	Oct-20	F	Visualizing Statistics
23	Oct-23	M	Visualizing Statistics
24	Oct-25	W	Prep for Assignment #4 (Temple Analytics Challenge)

Schedule (3/4, subject to change)

Day	Date		Topic
25	Oct-27	F	Key Performance Indicators
26	Oct-30	M	Aggregating Data
27	Nov-01	W	Visualizing Network
28	Nov-03	F	Visualizing Network
29	Nov-06	M	Exam #2 Review
30	Nov-08	W	Exam #2
31	Nov-10	F	Predictive Analytics
32	Nov-13	M	Predictive Analytics
33	Nov-15	W	Assignment #5 Plan Presentations
34	Nov-17	F	Assignment #5 Plan Presentations
	Nov-20	M	<i>No Class (Thanksgiving Break)</i>
	Nov-22	W	<i>No Class (Thanksgiving Break)</i>
	Nov-24	F	<i>No Class (Thanksgiving Break)</i>

Schedule (4/4, subject to change)

Day	Date		Topic
35	Nov-27	M	Data Science and Your Career
36	Nov-29	W	Data Science and Your Career
37	Dec-01	F	Exam #3 Review
38	Dec-04	M	Exam #3
39	Dec-06	W	TBA
	Dec-08	F	<i>No Class (Instructor travels for a conference)</i>
	Dec-11	M	<i>No Class (Instructor travels for a conference)</i>

Evaluation

Assignments #1 - #4	40% (10% each)
Assignment #5 (Term Project)	15%
Exams #1 - #2	20% (10% each)
Exam #3	15%
In-Class Exercise	10%

94% - 100%	A	73% - 76.99%	C
90% - 93.99%	A-	70% - 72.99%	C-
87% - 89.99%	B+	67% - 69.99%	D+
83% - 86.99%	B	63% - 66.99%	D
80% - 82.99%	B-	60% - 62.99%	D-
77% - 79.99%	C+	< 60%	F

Assignments

#	Assignment	Due Date (tentative)
1	Create a Data Analysis Plan	Fri, Sep 15
2	Analyze a Dataset Using Tableau	Mon, Oct 2
3	Cleaning a Dataset	Mon, Oct 16
4	Analytics Challenge	Tue, Oct 31
5	Original Data Analysis (Term Project)	Plan – Tue, Nov 14 Final – Mon, Dec 11

Assignment #5 – Term Project (15%)

- An original data analysis of your own with a dataset that you collect
- can be done individually or by a group of no more than four classmates
- Use any data source you'd like – government, sports, economics, politics, healthcare...
- You'll visualize your analyses with Tableau.
- By November 14, you're required to submit a plan.
- On November 15 or 17, you'll present your plan in class.
- The final term project is due by December 11.

Assignment Late Submission Policy

- I understand things happen sometimes. So, late submission is allowed.
- But there will be 10% penalty per each 12 hours.
- For example, if an assignment is submitted 30 hours after the deadline and it is graded 80, a 30% penalty is imposed, and the assignment will get $80 \times (100-30)/100 = 56$.
- I will not take any extraneous circumstance into consideration such as a PC malfunction or network outages.
 - A submission of one minute late will receive a 10% deduction.

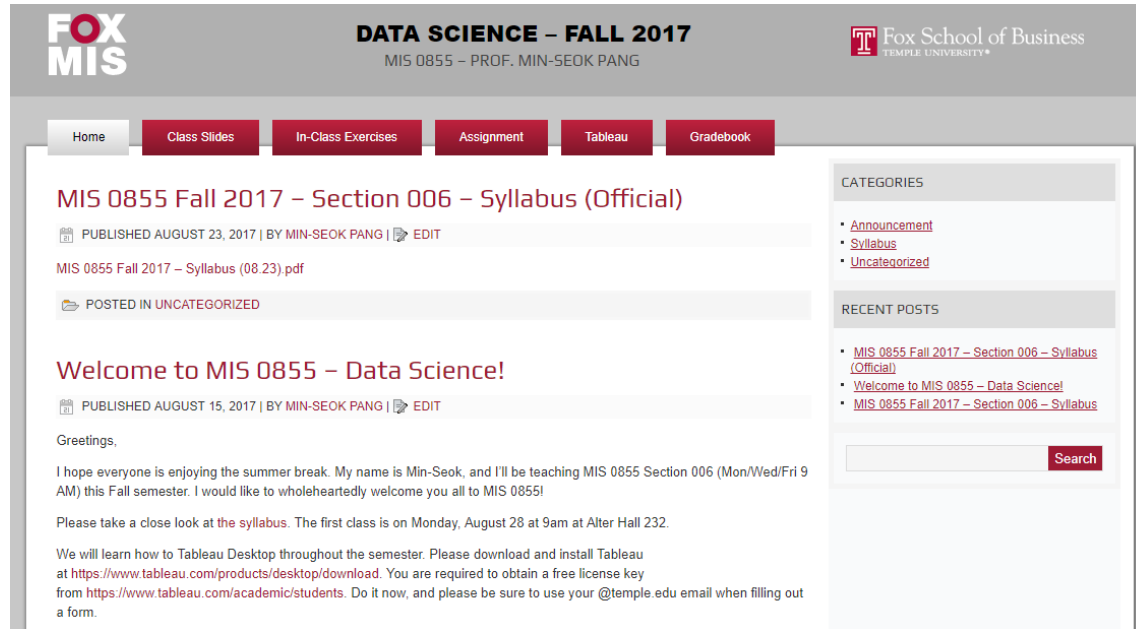
Three Exams (35%)

- September 25, November 8 and December 4 (tentative)
- Based on class slides, in-class exercises, and assignments
- Exam #1 (10%), Exam #2 (10%), and Exam #3 (15%)
- Closed-books-closed-notes
- Exam #2 is not cumulative. Exam #3 covers the whole course.
- Consist of multiple-choices
- Review sessions before the exams

In-Class Exercises (10%)

- There will be one in-class exercise (either group or individual) every session.
- All in-class exercises will be counted toward the grade.
- Deliverables from in-class exercises will be graded by success or fail.
- Up to four missing in-class exercises will be excused.
 - Use this opportunity only if it is necessary!

Class Site at MIS Community



The screenshot shows a web page for the MIS 0855 Fall 2017 class. The header includes the Fox School of Business logo and the course title "DATA SCIENCE – FALL 2017" by Prof. Min-Seok Pang. A navigation menu contains links for Home, Class Slides, In-Class Exercises, Assignment, Tableau, and Gradebook. The main content area features a post titled "MIS 0855 Fall 2017 – Section 006 – Syllabus (Official)" published on August 23, 2017, with a link to the syllabus PDF. Below this is a "Welcome to MIS 0855 – Data Science!" post from August 15, 2017, which includes a greeting and information about the first class on August 28 at 9am in Alter Hall 232. It also mentions that Tableau Desktop will be used throughout the semester and provides a link to download it. A search bar is located in the bottom right corner of the page.

- <http://community.mis.temple.edu/mis0855sec006fall17/>
- All class materials, announcements, and grades will be posted here.

Academic Integrity

- ZERO TOLERANCE
- Plagiarizing others' work without references will be reported to the University immediately.
- If deliverables submitted by multiple students are suspected to be work of one single person, the instructor will report to the University as a cheating.
- DO NOT SHARE ANY FILE THAT YOU CREATE WITH ANYONE.

Class Policies

- Keep monitoring announcements at the class site.
- Using laptops or tablets is allowed only for class-related activities.
 - Should close laptops while the instructor plays a video.
- Please use @temple.edu account for all correspondents with the instructor. Email messages sent from a non-Temple account may not be responded.
- Please turn off or mute cell phones. Do not send emails, texts, or tweets during class.
- In case of severe weather, we will meet as long as the University is open.