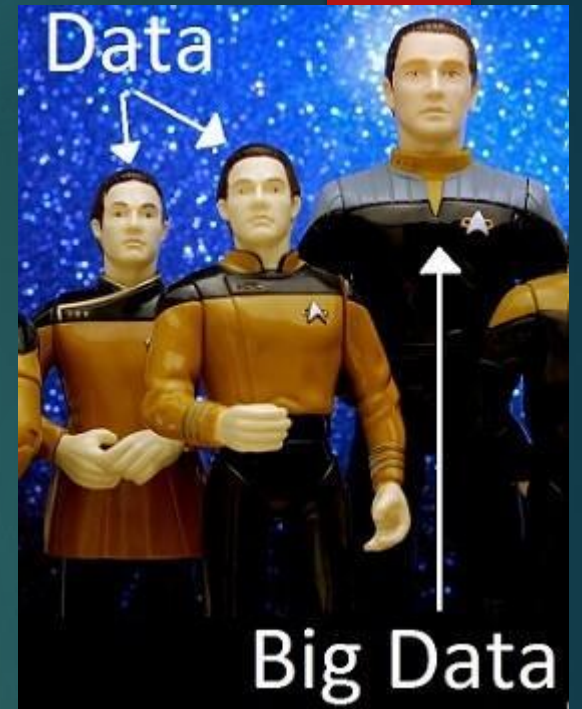


Big Data

SUKHJIT (BOB) SINGH

SINGHSUKHJIT@DEANZA.EDU



Post your questions here

- ▶ <https://todaysmeet.com/ACMClubBigData>



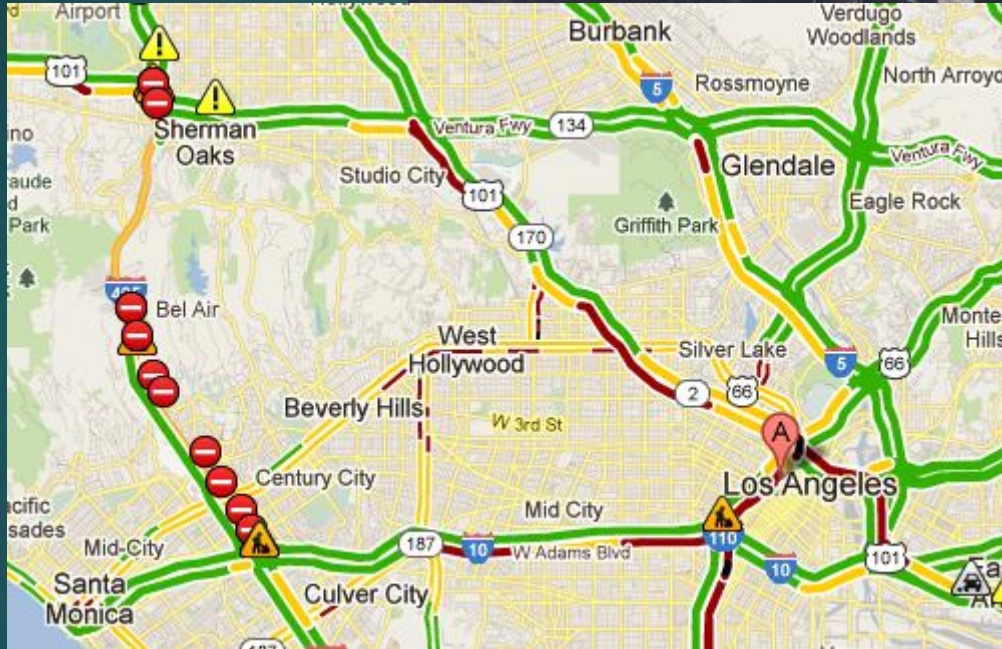
Agenda

- ▶ What is it?
- ▶ Why we need it?
- ▶ How it works?
- ▶ What can you learn at De Anza?
- ▶ How can you use it in the world?

What is Big Data

“extremely large data sets that may be analyzed computationally to reveal patterns, trends, and associations, especially relating to human behavior and interactions.”

What is Big Data



What is Big Data

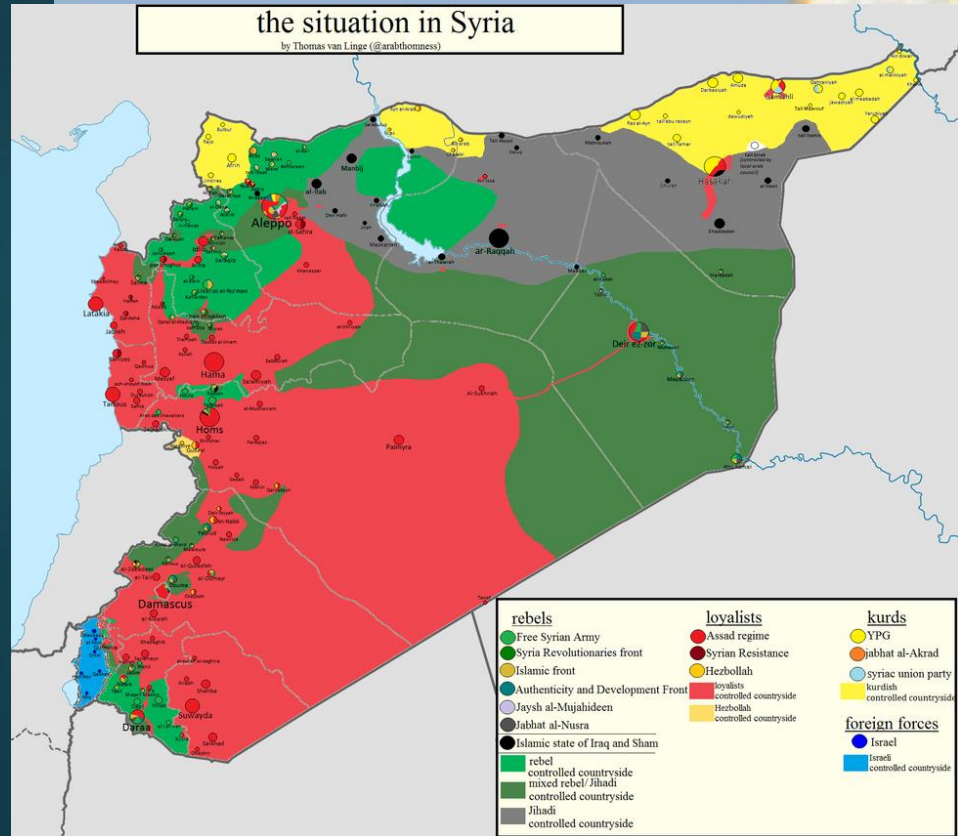


We collected 30,261,818 geocoded tweets between Oct. 1 and Nov. 1, 2012. Map by Mark Graham. Assistance by Nathan Carter and Ming Wang, Oxford Internet Institute. i@ox.ac.uk | www.ox.ac.uk/geography | www.ox.ac.uk

all geocoded tweets in October 2012

What is Big Data?

Filter Reports By




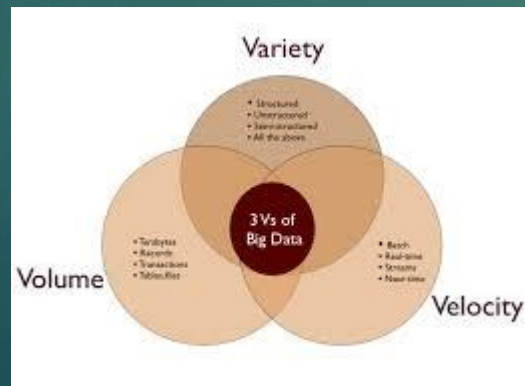
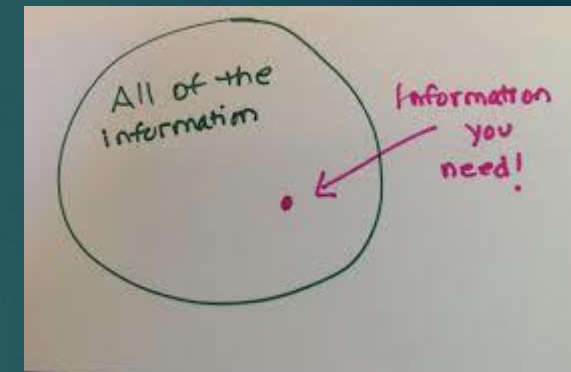
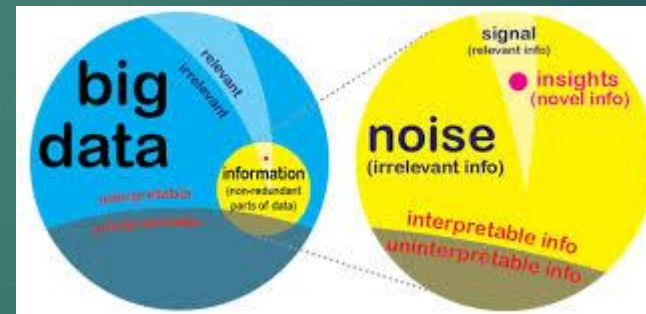
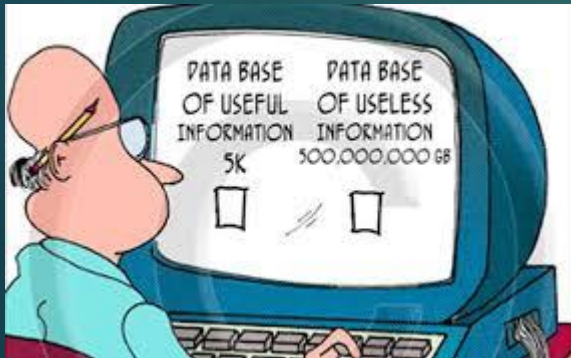
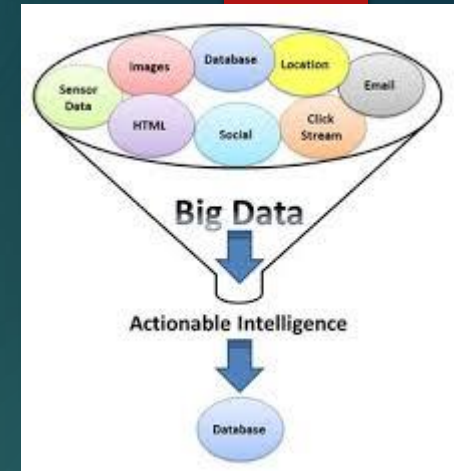
What is Big Data?

<http://tweetping.net/>

Why do we need it?

- ▶ Because it's here! Dude
- ▶ Data Driven (Communism) versus Data Informed (Democratic)
- ▶ Usability
- ▶Smarter Human Race

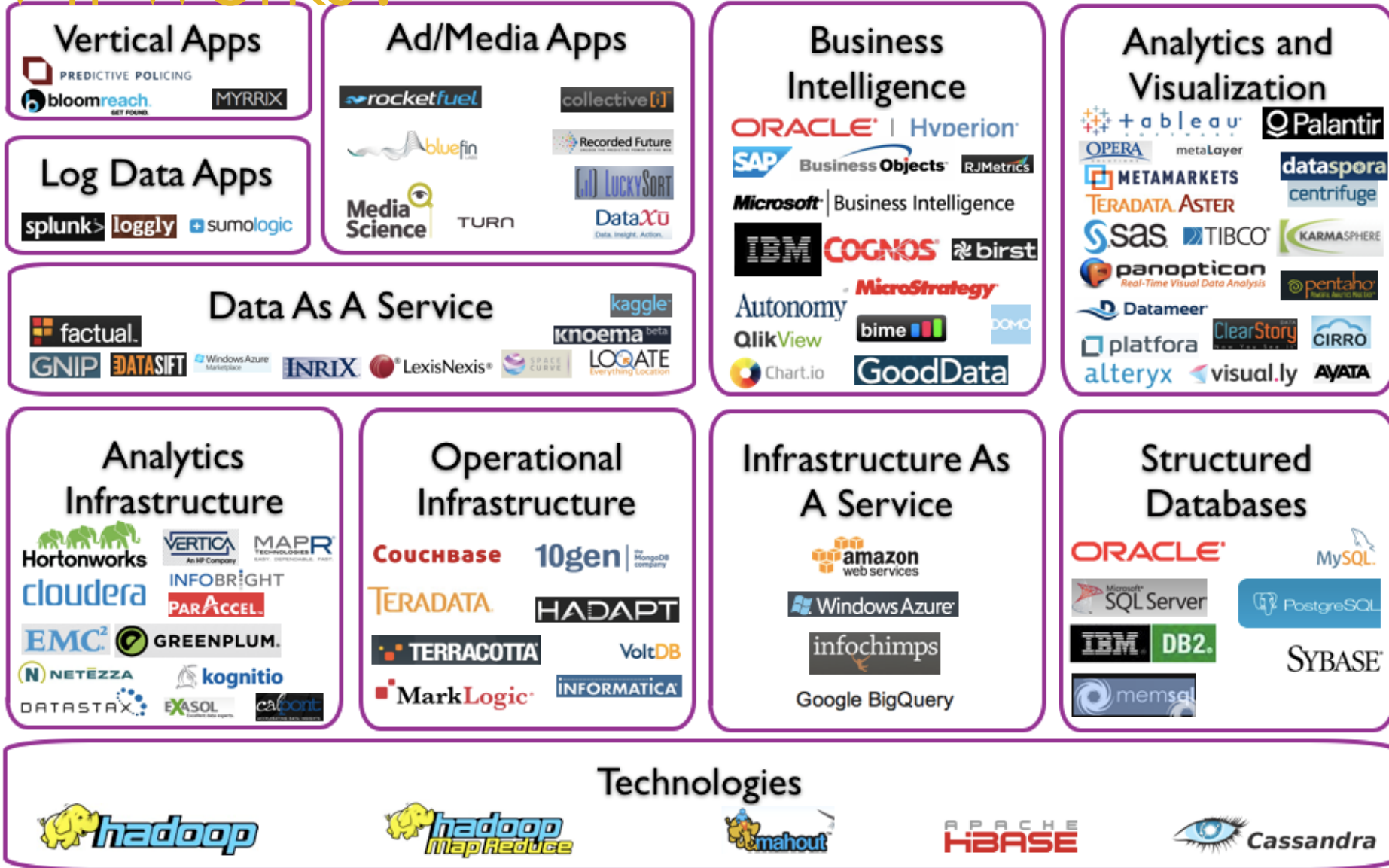
It is estimated that by 2020, there could be four times more digital data than all the grains of sand on Earth

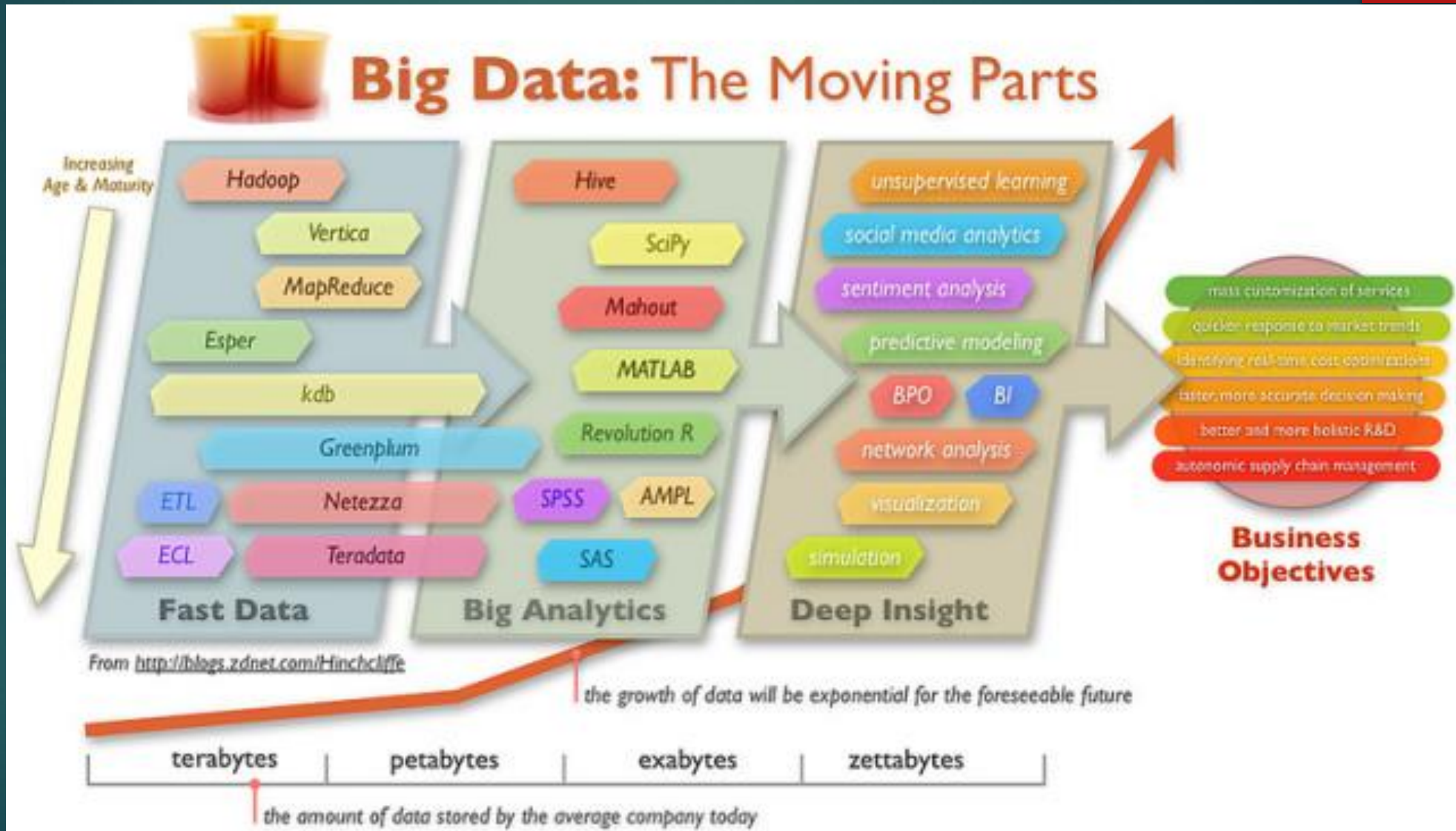
Why do we need it?



How it works? Big Data Landscape




How it works?



How it works?

Count the number of words in the following paragraph.
Make sure you do not count "Stop Words" (the, a, is)

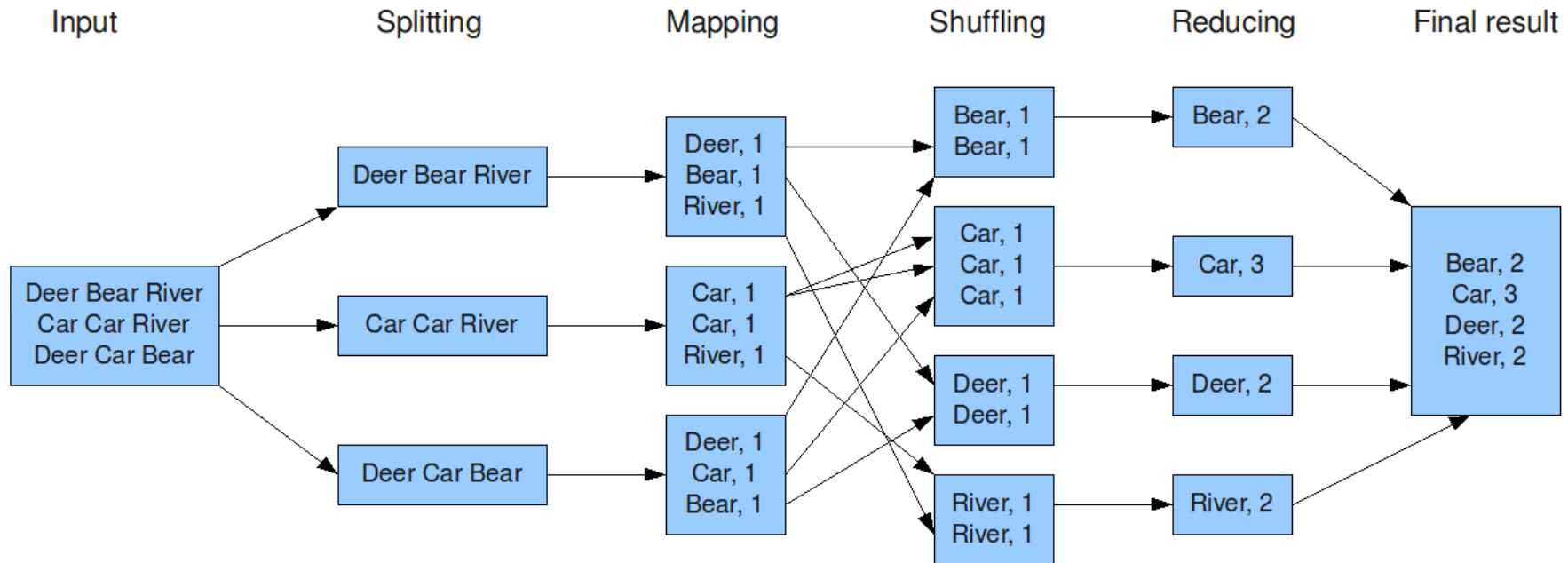
***The quick brown fox jumps over the lazy dog
A dog is a man's best friend***



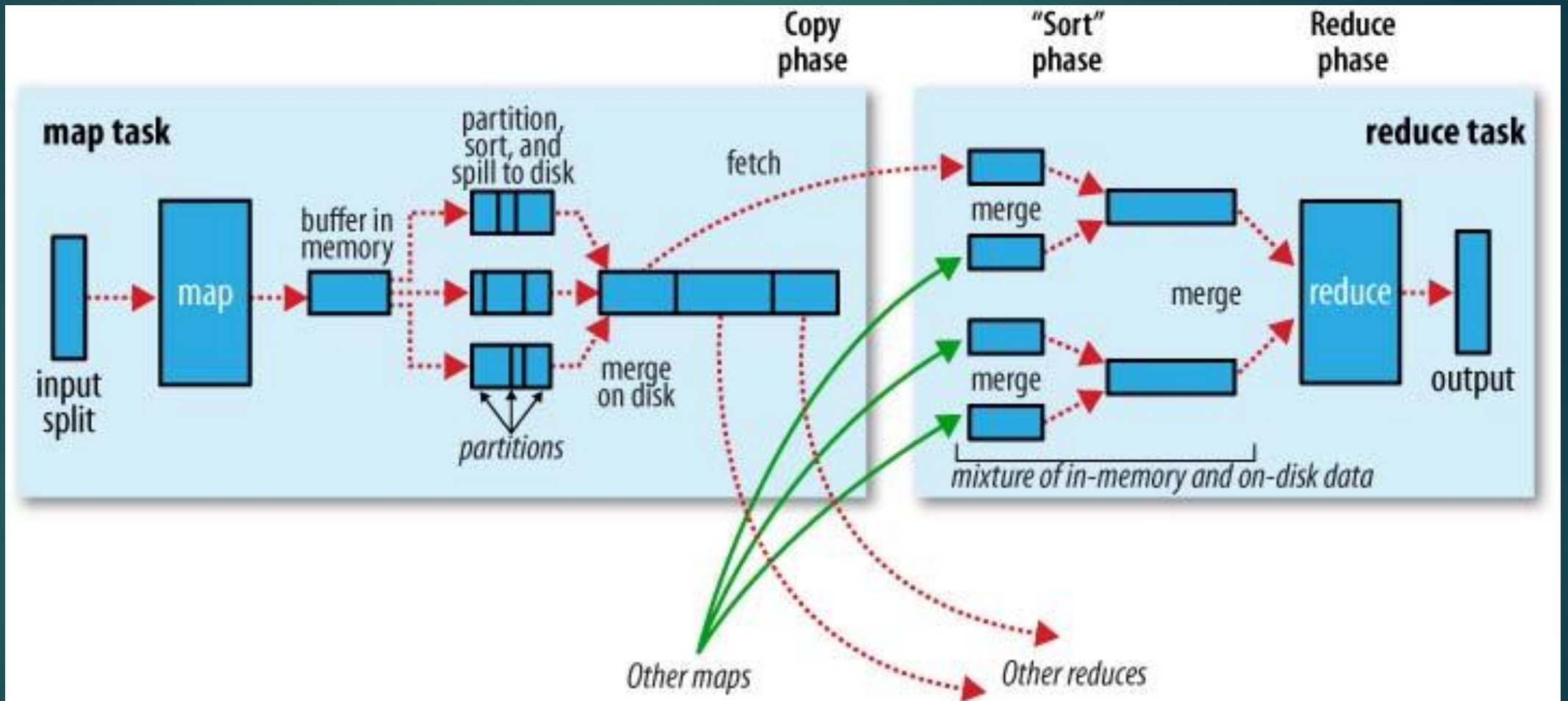
quick	=> 1
brown	=> 1
fox	=> 1
jumps	=> 1
over	=> 1
lazy	=> 1
dog	=> 2
mans	=> 1
best	=> 1
friend	=> 1

How it works?

The overall MapReduce word count process



How it works?

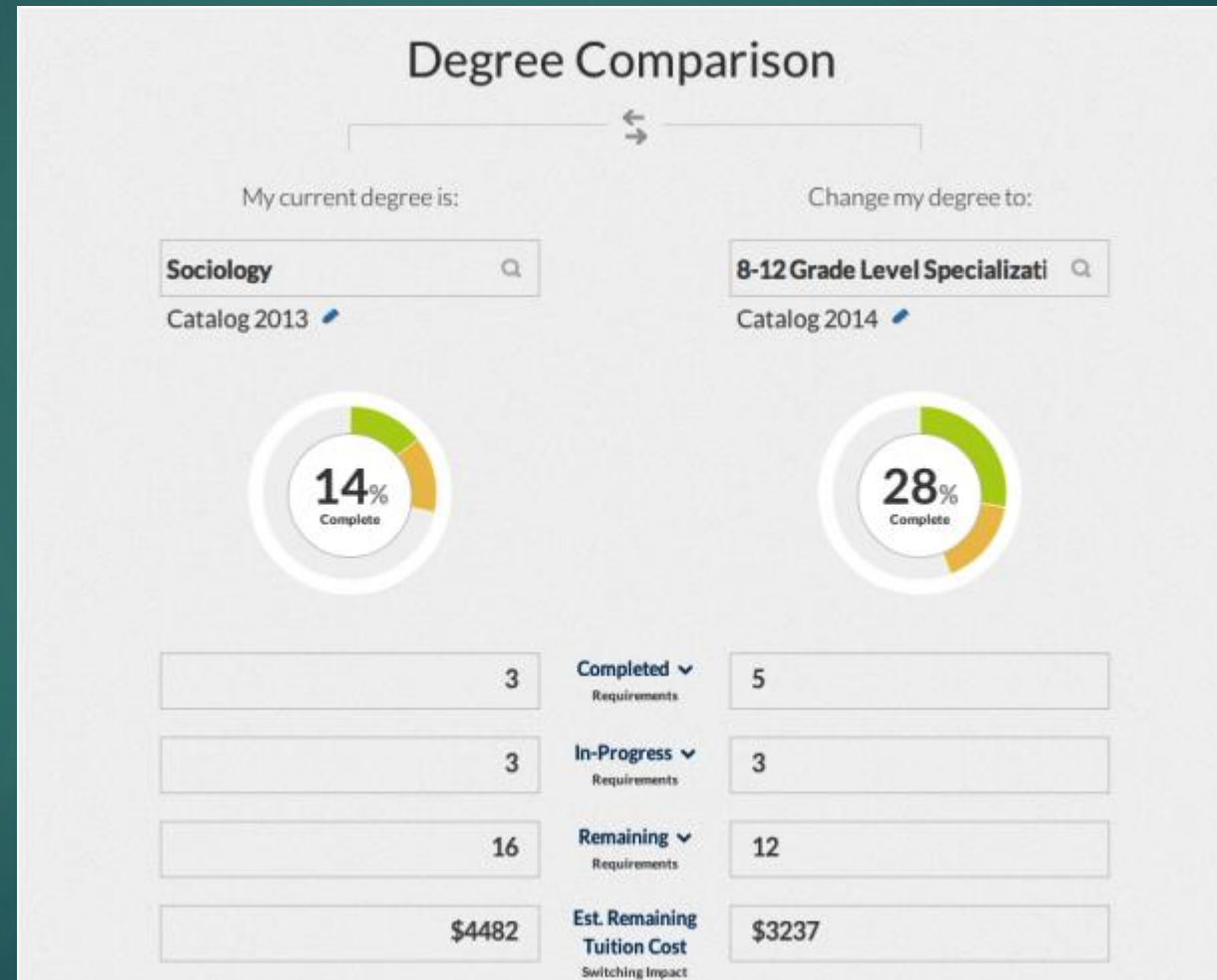


What can you learn at De Anza?

Google - De Anza Database Design

<http://www.deanza.edu/counseling/pdf/degrees/cis2-database.pdf>

How can we use it @ De Anza?



THE BIG DATA REVOLUTION IS HERE

As more and more companies use big data to influence their business decisions, the demand for skilled data analysts is seeing unprecedented growth—across all industries. Here's a snapshot of how big data creates even bigger job opportunities.

THE DATA ANALYTICS PROCESS

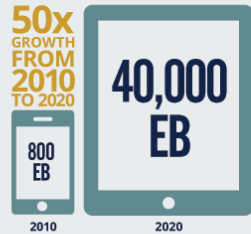


THE SOURCE OF IT ALL

BIG DATA

Big data is used to describe data sets that are too large and complex to be processed and analyzed using traditional technologies.

THE DIGITAL UNIVERSE



1 exabyte (EB) = 1,000,000,000,000,000 bytes
Source: IDC's Digital Universe Study, sponsored by EMC, December 2012

DATA ANALYTICS JOB OPPORTUNITIES

TOP 5 REGIONS



Source: Burning Glass International report of job postings for bachelor's and graduate degree holders in the data analytics field during 2012

TREND AMONG MAJOR COMPANIES

90% of Fortune 500 companies are predicted to have at least some big data initiatives under way.
Source: "Below and below: Big data becomes a big deal," Deloitte, Technology, Media & Telecommunications Predictions 2012

3/4 of executives at large companies in the U.S. and U.K. said they are working to increase their company's use of analytics.
Source: "Counting on Analytical Talent," Accenture, March 2010

1.5 MILLION

data-savvy managers and analysts will be needed in the next 5 years to drive decisions.

Source: "Big data: The next frontier for innovation, competition, and productivity," McKinsey Global Institute, May 2011

GAINING THE COMPETITIVE ADVANTAGE

THE TOP IN-DEMAND SKILLS

1. Structured Query Language (SQL)
2. Business intelligence
3. SAS business analytics software
4. Data analysis
5. Business analysis
6. Data warehousing
7. Business processes
8. Data management
9. Extract, Transform & Load (ETL) dev.
10. Data modeling

Source: Burning Glass International report of job postings for bachelor's and graduate degree holders in the data analytics field during 2012

"I keep saying that the sexy job in the next 10 years will be statisticians. And I'm not kidding."

— HAL VARIAN, chief economist at Google

Source: "Top Today's Gradients: Just One Word: Statistics," New York Times, August 3, 2009

TOP COMPANIES HIRING IN DATA ANALYTICS IN D.C., MARYLAND AND VIRGINIA

1. Deloitte
2. Capital One
3. IBM
4. Booz Allen Hamilton
5. Northrop Grumman
6. SAIC
7. CGI Group
8. General Dynamics
9. CACI
10. Freddie Mac

Source: Burning Glass International report of job postings for bachelor's and graduate degree holders in the data analytics field during 2012

GROWING JOB MARKET

\$97,000-\$108,000

The mean salaries for positions in the data analytics field.

Source: Burning Glass International report of job postings for bachelor's and graduate degree holders in the data analytics field during 2012

140,000-190,000

The predicted shortage of talent in the next five years with deep analytical skills to take advantage of big data.

Source: "Big data: The next frontier for innovation, competition, and productivity," McKinsey Global Institute, May 2011

UMUC
University of Maryland
University College

For more information on UMUC's data analytics program, visit umuc.edu/data

Questions
THANK YOU