



Department of Computer Science  
UNIVERSITY OF COLORADO **BOULDER**



# What is Data Science

Introduction to Data Science Algorithms

Jordan Boyd-Graber and Michael Paul

AUGUST 25, 2016

## Roadmap

---

- What data science is
- What data science can do
- What the course is about

Data are everywhere.

## User ratings

---

<a href="#"><u>Ikiru</u></a> (1952)	UR	Foreign	
<a href="#"><u>Junebug</u></a> (2005)	R	Independent	
<a href="#"><u>La Cage aux Folles</u></a> (1979)	R	Comedy	
<a href="#"><u>The Life Aquatic with Steve Zissou</u></a> (2004)	R	Comedy	
<a href="#"><u>Lock, Stock and Two Smoking Barrels</u></a> (1998)	R	Action & Adventure	
<a href="#"><u>Lost in Translation</u></a> (2003)	R	Drama	
<a href="#"><u>Love and Death</u></a> (1975)	PG	Comedy	
<a href="#"><u>The Manchurian Candidate</u></a> (1962)	PG-13	Classics	
<a href="#"><u>Memento</u></a> (2000)	R	Thrillers	
<a href="#"><u>Midnight Cowboy</u></a> (1969)	R	Classics	

# Purchase histories

---

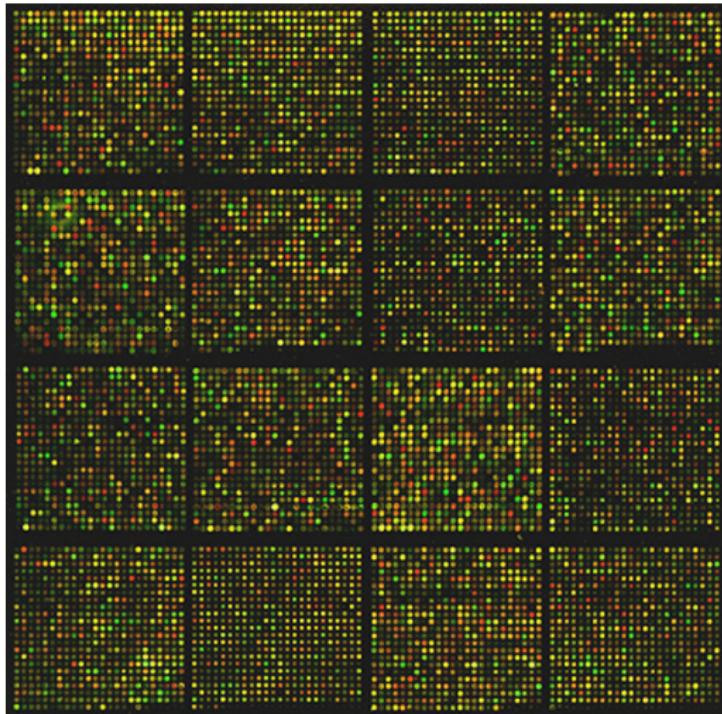
0.5/0.51 lb	<b>Cheese</b> <b>Cabot Vermont Cheddar</b>	0.51 lb	\$7.99/lb	<b>\$4.07</b>
<b>Dairy</b>				
1/1	<b>Friendship Lowfat Cottage Cheese (16oz)</b>		\$2.89/ea	<b>\$2.89</b>
1/1	<b>Nature's Yoke Grade A Jumbo Brown Eggs (1 dozen)</b>		\$1.49/ea	<b>\$1.49</b>
1/1	<b>Santa Barbara Hot Salsa, Fresh (16oz)</b>		\$2.69/ea	<b>\$2.69</b>
1/1	<b>Stonyfield Farm Organic Lowfat Plain Yogurt (32oz)</b>		\$3.59/ea	<b>\$3.59</b>
<b>Fruit</b>				
3/3	<b>Anjou Pears (Farm Fresh, Med)</b>	1.76 lb	\$2.49/lb	<b>\$4.38</b>
2/2	<b>Cantaloupe (Farm Fresh, Med)</b>		\$2.00/ea	<b>\$4.00 S</b>
<b>Grocery</b>				
1/1	<b>Fantastic World Foods Organic Whole Wheat Couscous (12oz)</b>		\$1.99/ea	<b>\$1.99</b>
1/1	<b>Garden of Eatin' Blue Corn Chips (9oz)</b>		\$2.49/ea	<b>\$2.49</b>
1/1	<b>Goya Low Sodium Chickpeas (15.5oz)</b>		\$0.89/ea	<b>\$0.89</b>
2/2	<b>Marcal 2-Ply Paper Towels, 90ct (1ea)</b>		\$1.09/ea	<b>\$2.18 T</b>
1/1	<b>Muir Glen Organic Tomato Paste (6oz)</b>		\$0.99/ea	<b>\$0.99</b>
1/1	<b>Starkist Solid White Albacore Tuna in Spring Water (6oz)</b>		\$1.89/ea	<b>\$1.89</b>

## Document collections

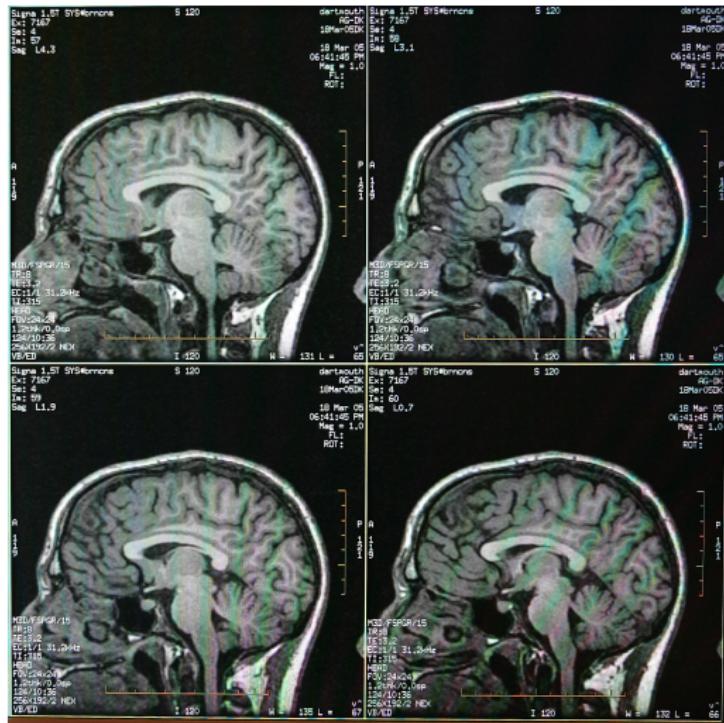


## Genomics

---

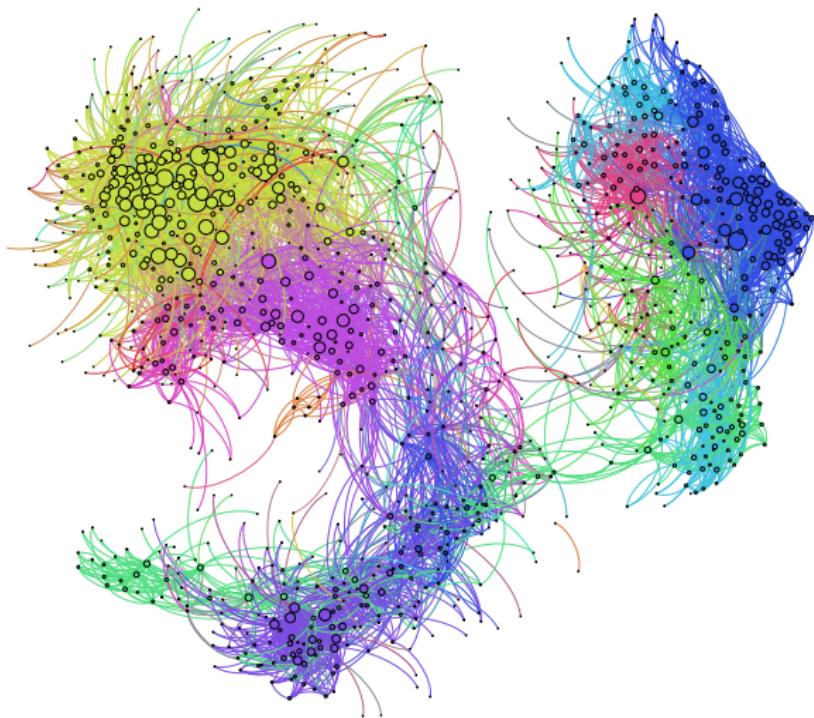


# Neuroscience



## Social networks

---



# Finance



Data can help us solve problems.

## Mathematical Foundations

---

Data

X

Answers

Y

Hidden Structure

Z