

Introduction to Event Data

Maggie @jandwiches
Dustin @larimer

Keen IO

Analytics is the discovery and communication of meaningful patterns in data.

AGENDA

- Some context on big data & analytics
- What is the goal of your app?
- Event data
- Common analytics methods
- Analyze some data

SOME CONTEXT ON BIG DATA AND ANALYTICS

Every company is becoming a **software company**.
Every **software company** is becoming a **data company**.



Big Data and Analytics are kind of a thing right now.

THE WALL STREET JOURNAL.

U.S. EDITION Sunday, April 29, 2012 As of 9:44 AM EDT

Home World U.S. New York Business Tech Markets Market Data Opinion

TOP STORIES IN Technology  Apple, Samsung Back in Court  Kodak Gets Bid of More Than \$500 Million for Patents

April 29, 2012, 9:44 a.m. ET

Big Data's Big Problem: Little Talent

Article Video Stock Quotes Comments (46)

It seems that the markets are as much in love with "Big Data"—the ability to acquire, process and sort vast quantities of data in real time—as the technology industry.



Hilary Mason, chief scientist for the URL shortening service Bitly, outlines the key skills that data scientists must have.

The first Big Data initial public offering hit the market last week to roaring approval. Splunk Inc., **[SPLK +0.31%]** which helps businesses organize and make sense of all the information they gather, soared 109% on its first day of trading. Big Data, big price.

And this week, in cities in the U.S. and the U.K., Big Data Week events are being held to proselytize the unbelievers.

Big Data refers to the idea that an enterprise can mine all the data it collects right across its operations to unlock golden nuggets of business intelligence. And whereas companies in the past have had to rely on sampling, Big Data, or so the promise

Harvard Business Review

THE MAGAZINE BLOGS AUDIO & VIDEO BOOKS WEBINARS COURSES

Guest | limited access Register today and save 20%* off your first order! [Details](#)

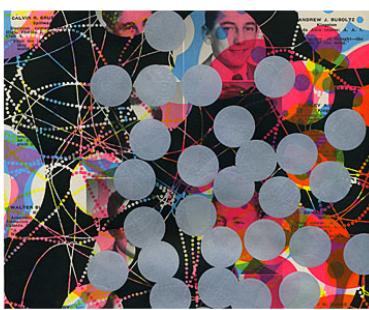
THE MAGAZINE
October 2012

ARTICLE PREVIEW To read the full article: [Sign in](#) or [Register](#) for free. HBR Subscribers [activate your free archive access »](#)

Data Scientist: The Sexiest Job of the 21st Century

by Thomas H. Davenport and D.J. Patil

Comments (6)



RELATED
[Executive Summary](#)
ALSO AVAILABLE
[Buy PDF](#)

Artwork: **Tamar Cohen, Andrew J Buboltz**, 2011, silk screen on a page from a high school yearbook, 8.5" x 12"

When Jonathan Goldman arrived for work in June 2006 at LinkedIn, the business networking site, the place still felt like a start-up. The company had just under 8 million accounts, and the number was growing quickly as existing members invited their friends and colleagues to join. But users weren't seeking out connections with the people who were already on the site at the rate executives had expected. Something was apparently missing in the social experience. As one LinkedIn

THE WALL STREET JOURNAL.

U.S. EDITION Wednesday, December 3, 2014 As of 3:57 PM PST

Home World U.S. New York Business Tech Markets Market Data Your Money

CIO Journal.

CIO Report | Consumerization | Big Data | Cloud | Talent & Management | Security

November 10, 2014, 6:45 PM ET

Universities Can't Train Data Scientists Fast Enough for CIOs

Article Comments (4)

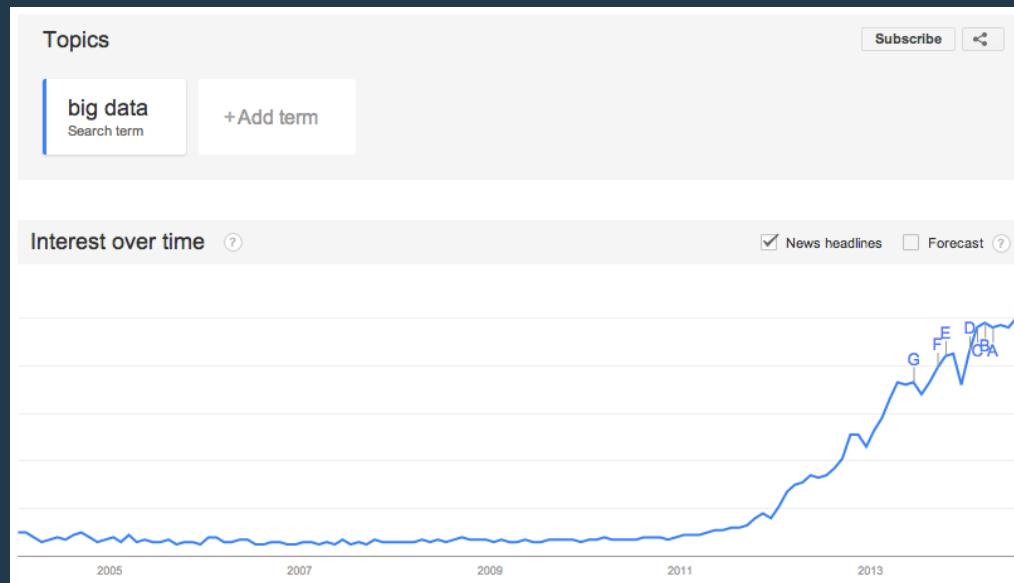
By CLINT BOULTON Reporter



Chief information officers are struggling to find data scientists, those individuals with engineering and business skills, as well as the statistical savviness required to analyze and derive value from Big Data companies generate. Academic leaders are endeavoring to produce more data scientists by offering school programs that span several disciplines, and encourage students to think more broadly about the technical and business levers they must pull to achieve desired outcomes.

But one data scientist asks why companies should bother. In time software, no

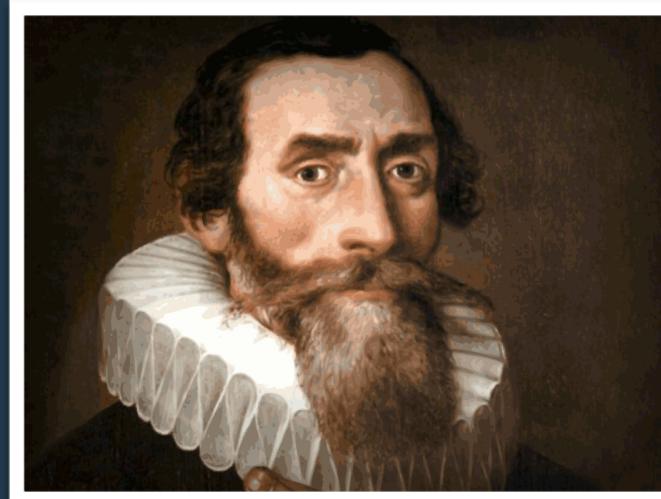
Trendz



COOL DATA STORIES



Tycho Brahe



Johannes Kepler



LinkedIn Account Type: Basic | Upgrade 23 1

Michelle Wetzer Add Connections

Home Profile Contacts Groups Jobs Inbox Companies News More People Search... Advanced

Want \$ for a Trust Deed - Get cash for real property secured loans, performing or non-performing.

Share an update Attach a link Share

All Updates ▾

LinkedIn Today recommends this news for you

£50m new investment in London's Tech City David C.

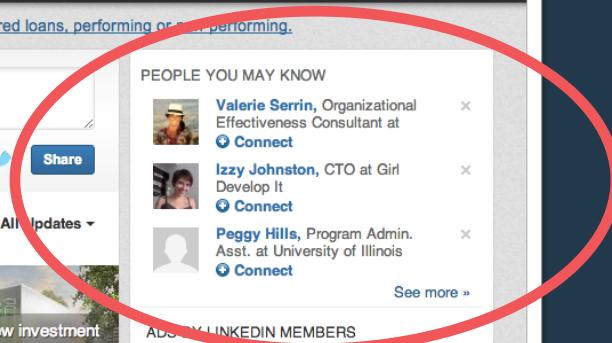
PEOPLE YOU MAY KNOW

- Valerie Serrin, Organizational Effectiveness Consultant at [+ Connect](#)
- Izzy Johnston, CTO at Girl Develop It [+ Connect](#)
- Peggy Hills, Program Admin. Asst. at University of Illinois [+ Connect](#)

See more »

ADS BY LINKEDIN MEMBERS

Build Huge Email List? Social media can increase your leads & sales within few weeks. Check it

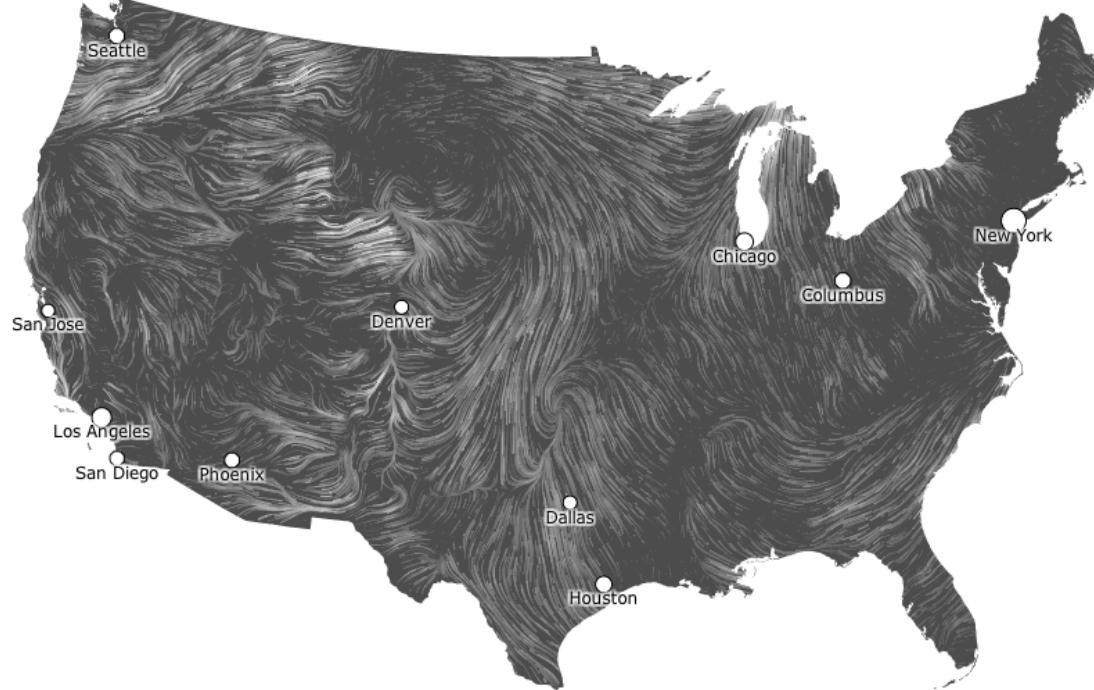
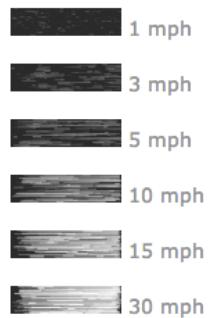


Dec. 6, 2012

5:59 pm EST

(time of forecast download)

top speed: **30.2 mph**
average: **6.2 mph**



<http://hint.fm/wind>

<http://infosthetics.com>

Le Petit Journal

ADMINISTRATION
61, RUE LAFAYETTE, 61

5 CENT. SUPPLÉMENT ILLUSTRÉ 5 CENT.

ABONNEMENTS

Les manuscrits ne sont pas rendus
On envoie sans frais
Une fois le bureau de poste

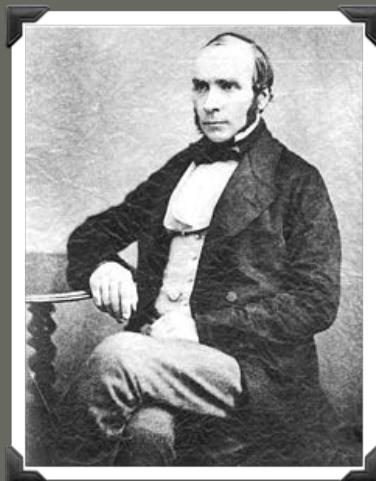
23^e Année 1150

DIMANCHE 1^{er} DÉCEMBRE 1912

PRISE MOIS UN AN
SOCIÉTÉ GÉNÉRALE... 2 fr. 3 fr. 60
DÉPARTEMENTS... 2 fr. 4 fr. *
ÉTRANGER... 2 fr. 6 fr. *

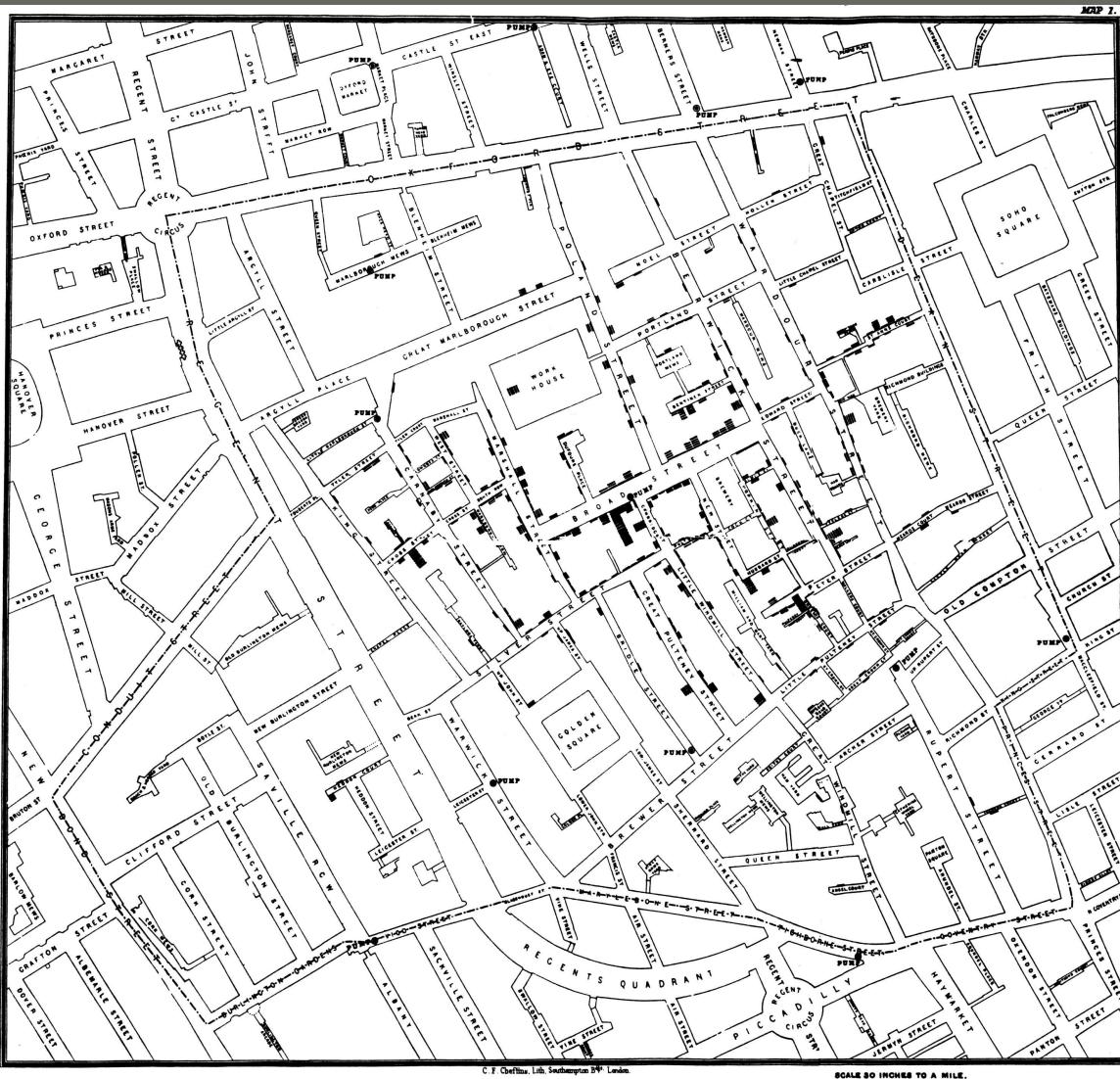


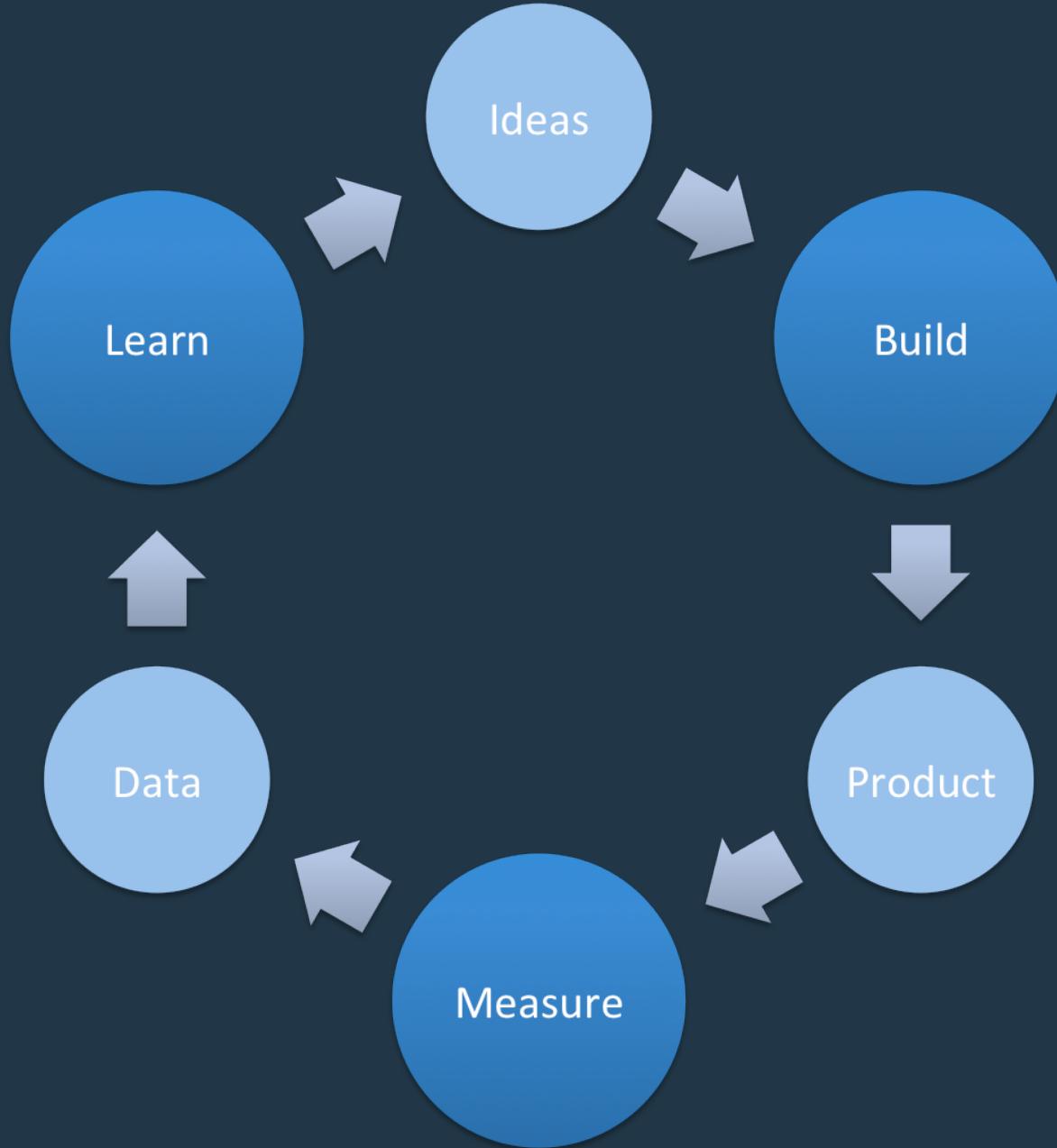
LE CHOLÉRA

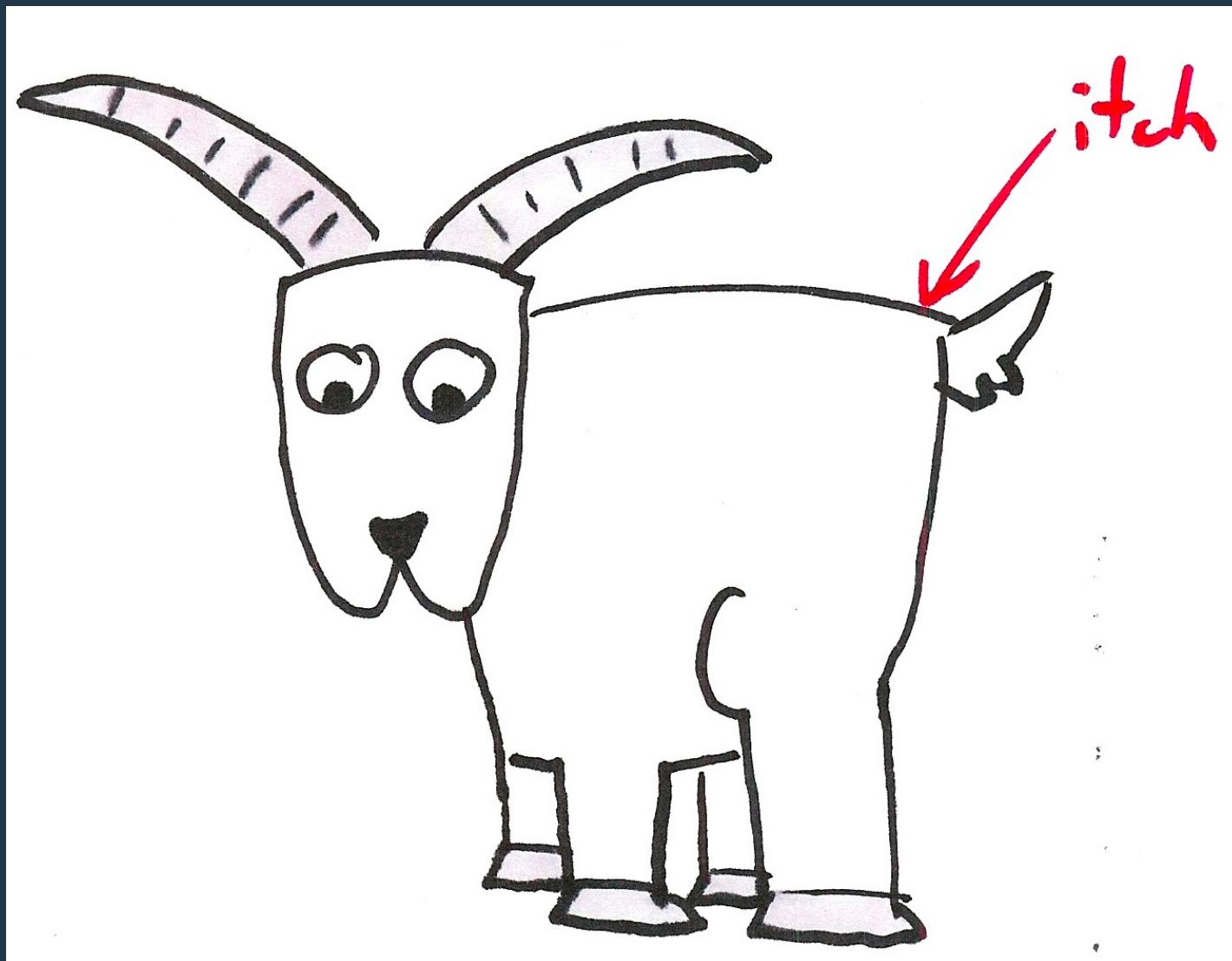


John Snow figured out Cholera spreads through water.

No one believed him :(







APPLYING ANALYTICS TO YOUR BUSINESS

Use analytics to measure progress toward a goal.

Use analytics to monitor health.

Use analytics to test new hypotheses.

Use analytics to explore.

Use analytics to provide value to your customers.

EXERCISE 1!

1 MINUTE

WHAT IS THE GOAL OF YOUR APP?

Examples:

Vine: reach 1M user-generated videos (increase usage)

G-School: increase efficacy of the program

A COMMON GOAL: ENGAGEMENT

- Account creations
- Deploys
- Purchases
- App Launches
- Donations
- Posts
- Shares/Tweets/Likes

INTRODUCING EVENT DATA



Actions & State + Time

Google  

Web News Images Shopping Videos More Search tools

About 230,000,000 results (0.43 seconds)

event.data | jQuery API Documentation
api.jquery.com › Events › Event Object › jQuery ›
event.dataReturns: Object. Description: An optional object of data passed to an event method when the current executing handler is bound.

Event data - Wikipedia, the free encyclopedia
en.wikipedia.org/wiki/Event_data › Wikipedia ›
Event data is a synonym to an audit trail. Modern computer software applications and IT infrastructure have adopted the term **event data** over audit trail. Events ...

Analytics For Hackers: How To Think About Event Data ...
https://keen.io/blog/.../analytics-for-hackers-how-to-think-about-event-d... ›
Jun 26, 2013 - There's this new and really powerful type of data: **event data**. Well, it's not really new. I bet you are already familiar with **event data**, even if you ...

jQuery event.data Property - W3Schools
www.w3schools.com/jquery/event_data.asp › W3Schools ›
Return the data passed with the on() method for each <p> element: ... The **event.data** property contains the optional data passed to an event method when the ...

Use the EVENTDATA Function - MSDN - Microsoft
https://msdn.microsoft.com/.../ms187909.a... › Microsoft Developer Network ›
Information about an event that fires a DDL trigger is captured by using the **EVENTDATA** function. This function returns an XML value. The XML schema includes ...

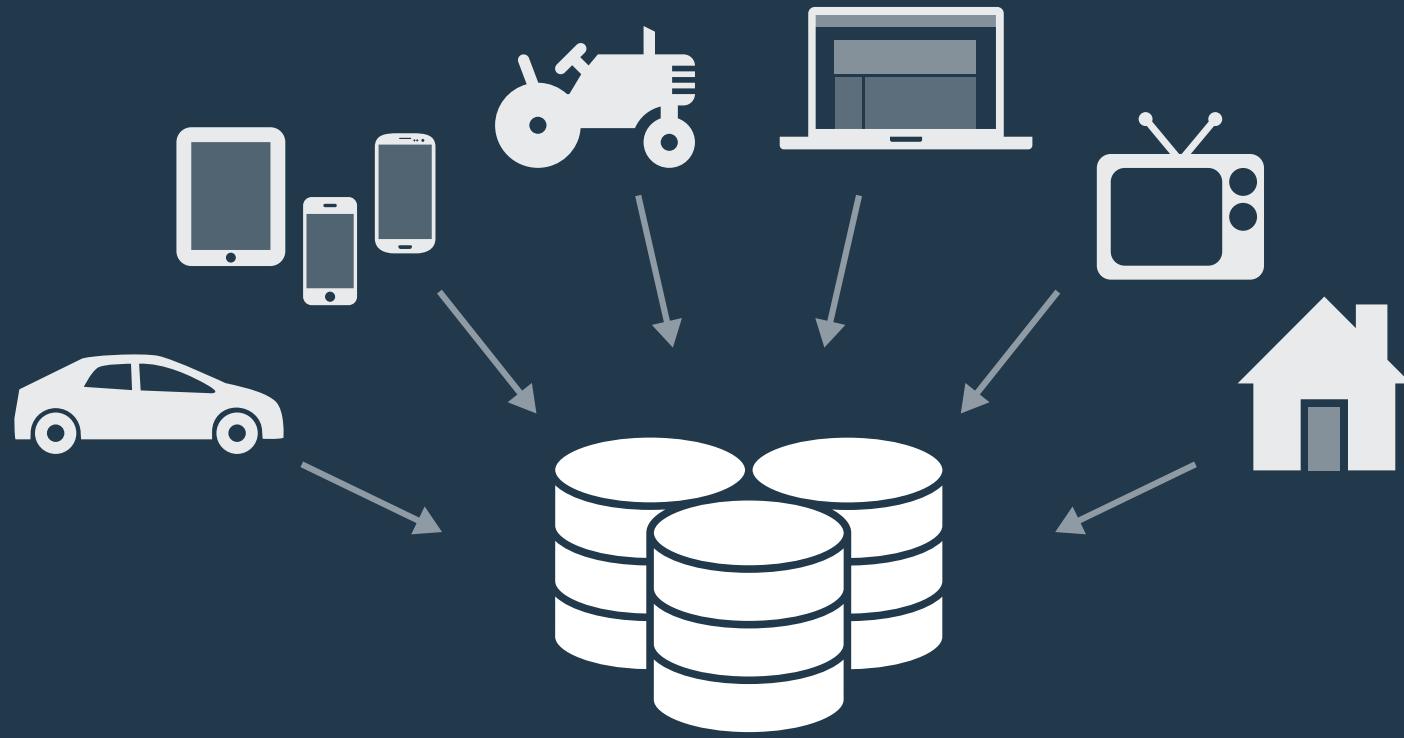
EVENTDATA (Transact-SQL) - MSDN - Microsoft
https://msdn.microsoft.com/.../ms173781.a... › Microsoft Developer Network ›

UID	twitter handle	age	Account ID
773345	@hipsterhacker	29	443556
773346	@TNG_S8	27	432354
773347	@modernseinfeld	28	336658
773348	@michellewetzler	28	2115789

```
{  
  "event": "death",  
  "timestamp": "2013-05-23T1:50:00-0600",  
  "cause": "creeper explosion",  
  "enemy": {  
    "type": "creeper",  
    "power": .887,  
    "distance_from_player": 3.43,  
    "age": .6677  
  },  
  "player": {  
    "UID": "99234890823",  
    "experience": 8873729,  
    "age": 338,  
    "inventory": ["diamond sword", "torches"]  
  }  
}
```



entity data	event data
strict schema	flexible schema
normalized	denormalized
shorter	wider
describes nouns	describes verbs
describes now	describes trends over time
updates	appends
big data	big big big data



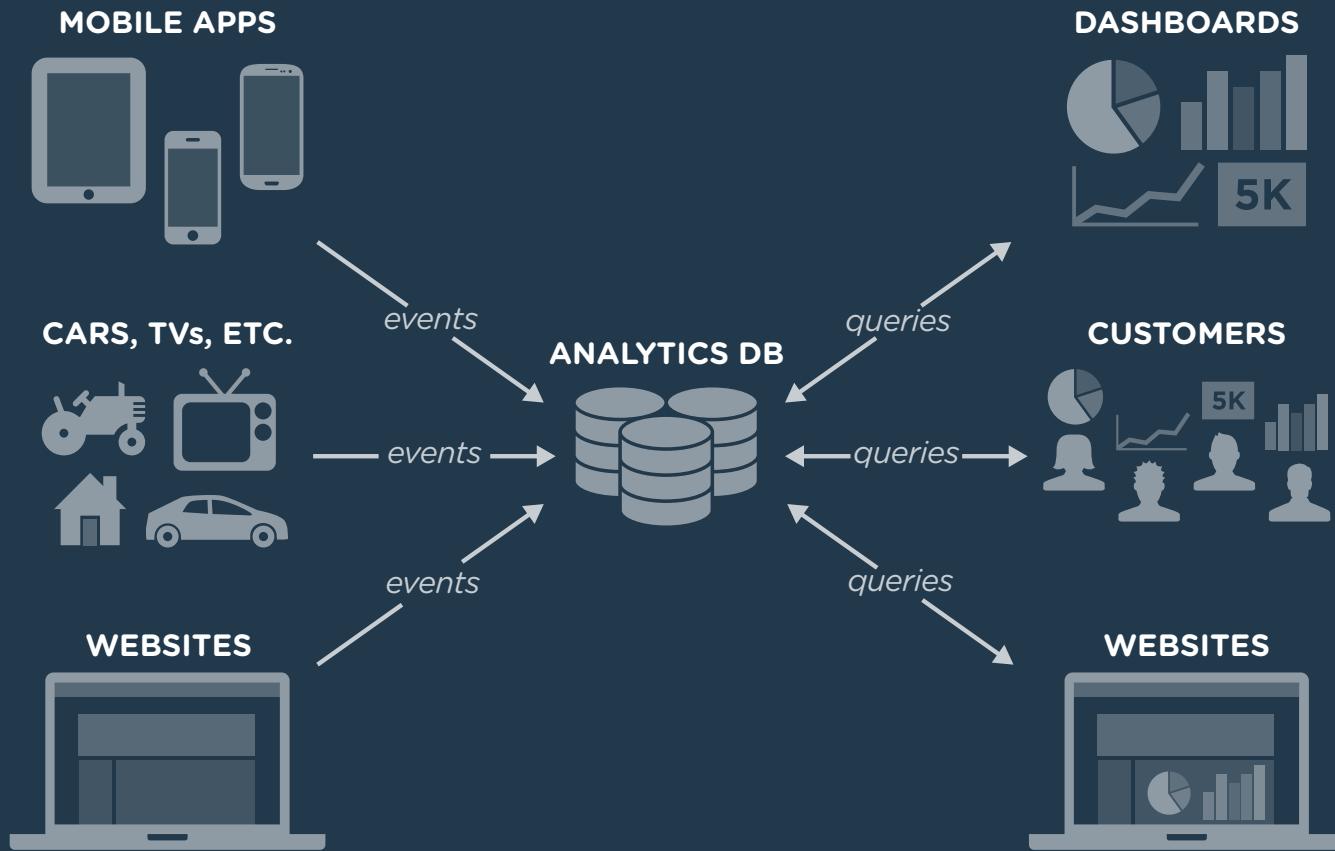
MORE EXAMPLES

ACTIONS: Signup, Login, Upgrade, Submit, Scroll, Send, Share, Search, Check-In, Vote, Update, Purchase, Level Up, Fail, Favorite, Vote, Crash, Rate, Start, Modify, Check, View, Capture

STATE INFO: User, Company, Organization, Team, Platform, Device, App, Level, Garden, Favorites, Interests, Inventory, Cart, Video, Location, Item, Record, Product, Account, Form, Picture, Story

ELEVATOR PITCH TECHNIQUE

- Describe your app to a stranger and listen to the words you use.
- Verbs are the actions you should record.
- Nouns are the important contextual information you should include in your data model.
- Most apps can be very robustly described by 5-10 key events and 5-10 key nouns.



EXERCISE 2!

2 MINUTES

Recall your goal from the previous exercise.

Write down at least one event you can track to measure your progress toward that goal.

EXERCISE 3!

2 MINUTES

Say you have a mobile app that allows users to purchase hats for their avatars. How would you model a purchase event?

COMMON ANALYTICS TECHNIQUES

99% of analytics work involves
what mathematical operation?

COUNTING!



MORE BASICS

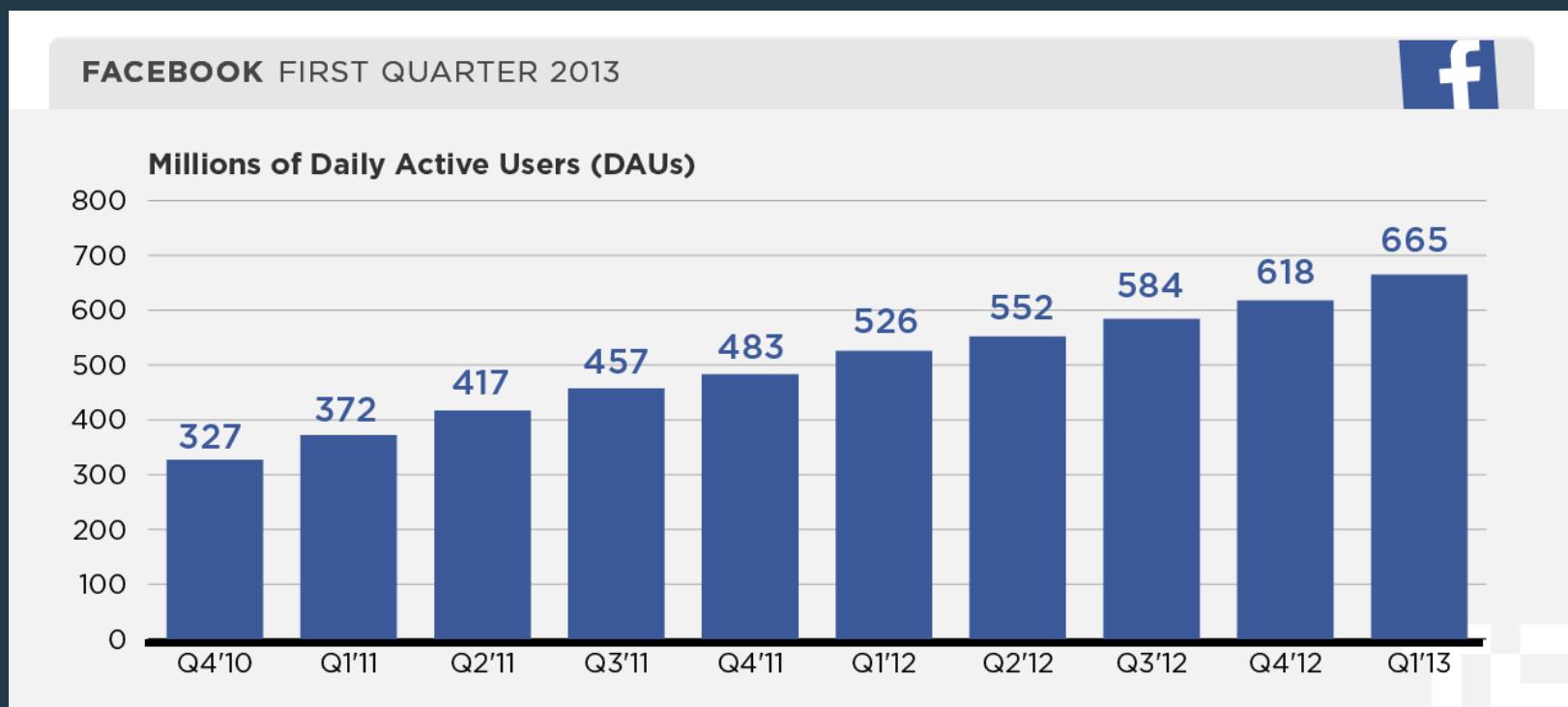
- Count Unique
- Sum
- Average
- Min
- Max

ADVANCED

- Statistical Analysis
- Correlation Analysis
- Predictive Analysis

Fancy Terms for Counting Stuff

DAU/MAU



EXAMPLE

What was the average revenue per active user last month?

1. Count the number of unique users who performed some action in June (2300)
2. Sum all of the purchases from June (\$5564)
3. Divide 2 by 1 (\$2.40)

SEGMENTATION

- Sorting data into buckets. Commonly used to sort users into groups.
- Examples: Gender, Age, Location, Department, Referrer, Version, Device

Country / Territory	Visits	% Visits
1. United States	33,399	59.17%
2. Canada	3,578	6.34%
3. United Kingdom	3,244	5.75%
4. Australia	1,588	2.81%
5. India	1,384	2.45%
6. Germany	1,183	2.10%
7. (not set)	1,153	2.04%
8. France	838	1.48%
9. Netherlands	697	1.23%
10. Russia	654	1.16%

FILTERING

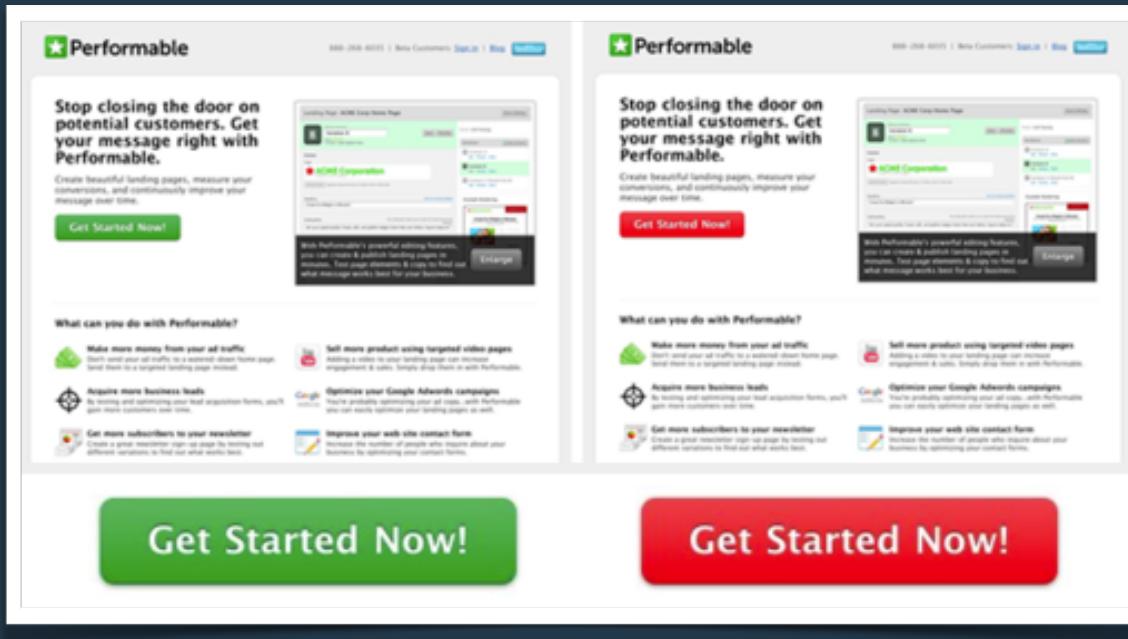
- Used to reduce the data set to which a query applies
- Use any of your event properties to do filtering.
- Example: Count the number of purchases events where item.category = “add-ons” and item.price > 100.

EXERCISE 4!

Recall the purchase event we just modeled.

What property or properties might you use for filtering or segmentation?

A/B TESTING AKA SPLIT TESTING



21% more people clicked on the red button than on the green button!

EXAMPLE OF SPLIT TESTING DATA

user.id	user.name.first	user.name.last	form.version	form.fields
223655	zach	morris	A	[first name, middle name, last name, organization, gender, age, email, password]
223656	kelly	kapowski	B	[email, password]
223657	screech	powers	A	[first name, middle name, last name, organization, gender, age, email, password]
223658	lisa	turtle	A	[first name, middle name, last name, organization, gender, age, email, password]
223659	ac	slater	B	[email, password]
223660	jessie	spano	B	[email, password]
223661	mr.	belding	B	[email, password]
223662	mrs.	culpepper	B	[email, password]
223663	stacey	carosi	A	[first name, middle name, last name, organization, gender, age, email, password]
223664	allison	fox	B	[email, password]
223665	tori	scott	B	[email, password]
223666	mr.	dewey	B	[email, password]
223667	ollie	creeky	B	[email, password]
223668	violet	bickerstaff	B	[email, password]
223669	rhonda	robistelli	B	[email, password]

Which version of the form was more effective?

FUNNELS



COHORT ANALYSIS

A cohort is a group of people who share a common characteristic over a certain period of time.



RETENTION

How many customers remain customers?

How many users came back a second time?

Do my customers value my product?

Measure retention by counting how many users did an action X days after their first usage.

RETENTION ANALYSIS BY COHORT



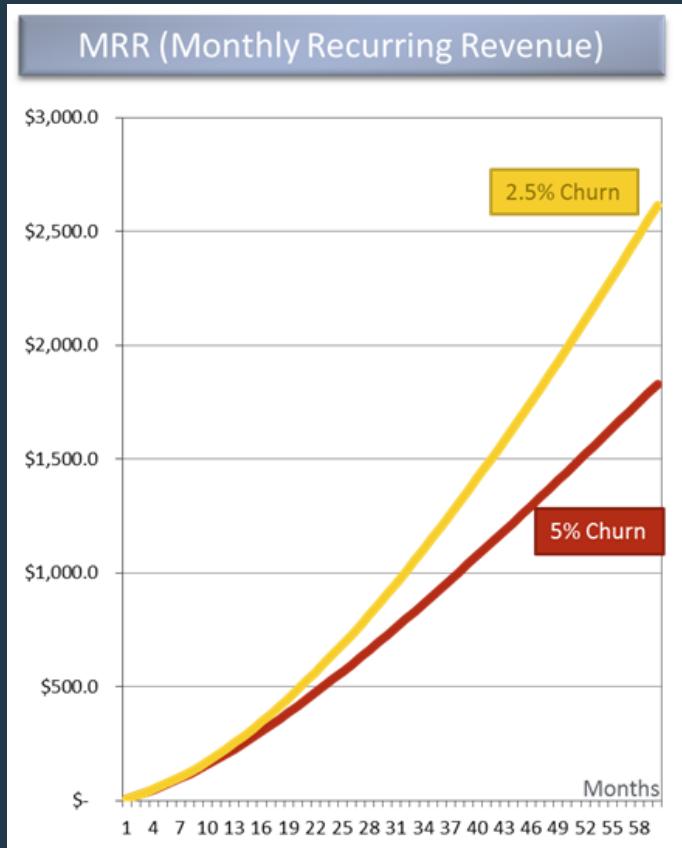
CHURN

How many users are we losing?

Churn is the total number of users you lost in a given timeframe, divided by the total number of users you had at the beginning of the timeframe.

CHURN

Churn impacts growth & profits significantly.



CUSTOMER LIFETIME VALUE (CLV, CLTV, LCV, LTV)

How much is a customer worth?

$$CLV = \text{Monthly Revenue} \times \text{Margin} \times \text{Number of Months}$$

$$CLV = \$100/\text{mo} \times 25\% \times 10 \text{ months} = \$250$$

Customer Acquisition Cost

How much did it cost to get that user?

CAC = \$ spent / number of users acquired

Include amount invested in marketing, advertising, and sales.

```
44
45 $steps = [
46   {
47     :event_collection => "create_organization",
48     :actor_property => "organization.id",
49     # :filters => [],
50     #       :property_name => "organization.name",
51     #       :operator => "ne",
52     #       :property_value => "TestOrg"
53     #       ],
54   },
55   {
56     :event_collection => "events_added_api_call",
57     :actor_property => "project.organization.id",
58   },
59 ]
60
61 # Protip: You may have more than two steps. Middle steps will further refine the number of candidates which make it to the
62 # Step 4 - Run this script!
63
64 x = 0
65 sums = {}
66 while x < (num_weeks + 1) do
67   sums[x] = {
68     "count" => 0,
69     "percent" => 0
70   }
71   x = x + 1
72 end
73 $row_sums = []
74 $row_sums << "totals"
75
76 #=====00=====
77 # These two nested loops run through all of the weeks since num_weeks ago, building funnels queries and running them.
78
79 SimpleXlsx::Serializer.new(Time.now.to_s[0..19]+".xlsx") do |doc| # This will create an excel file to output results
```

Querying Exercise

Use the event data in Keen IO (or google sheet) to answer the following questions.

1. What's the total revenue from all purchases, excluding the "High Rollers" campaign?
2. How many unique customers made purchases?
3. Which customers purchased the most products?
4. Which customers spent the most money?
5. What are the most popular products?
6. Are sales increasing or decreasing?

@jandwiches
maggie@keen.io