Designing with DATA

@paminthelab
change sciences
I <3 DATA

But I’m still learning
First Half 9-10:30
Background & case studies
What is the point?

Second 11-12:30
Step-by-step
How to?
3 million emails
2 million Google searches
684,000 Facebook shares
$272,000 spent online shopping
100,000 tweets
3,600 photos to Instagram
2 million blog posts
It seems like TOO MUCH
Big data is like teenage *sex* - everybody talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it.

Dan Ariely
Complicated  Lack of Customization
Inconsistent  Not Actionable  Can’t Access it w/o IT  Ambiguous
Not User-Friendly  Inaccurate  Disparate Data Sources
No Real-time Access  Dirty Data  Lack of Buy-in
Takes too long to manage  People can’t use it well...
Disorganized  Sheer Volume
It's easy to come to the **WRONG CONCLUSIONS**
APP STORE

TORNADO GUARD

FROM DROLLO CADER 2187

PLAYS A LOUD ALERT SOUND WHEN THERE IS A TORNADO WARNING FOR YOUR AREA.

RATING: ★★★★★

BASED ON 4 REVIEWS

USER REVIEWS:

★★★★★ GOOD UI!

★★★★★ MANY ALERT CHOICES.

★★★★★ RUNNING GREAT, NO CRASHES

★★★★★ I LIKE HOW YOU CAN SET MULTIPLE LOCATIONS

★★★★☆ APP DID NOT WARN ME ABOUT TORNADO.

THE PROBLEM WITH AVERAGING STAR RATINGS
It isn’t just about TARGETS
poverty  rainfall  global happiness  copyright value  culture fit
attitudes  employee performance  DESIGN  automobile safety
emotion  economic growth  engagement  size of an atom  corruption
healthcare outcomes  learning  intelligence  emotional intelligence
potential output  social media ROI  risk  dolphin population  love
innovation  national security  reputation  public influence
customer satisfaction  team productivity  violence in households
cooperation  air pollution  asset value of advertising  trust  pH
length of Saturn’s day  online readership  stress level  accountability
Think of it like ARCHEOLOGY
It's incomplete.

TRACES
It’s actually about PEOPLE
3 people
3 things in common
9 high fives
1. A small amount of people can produce a lot of data
2. If you don’t decide to collect data, you won’t get much that is useful
3. If you aren’t consistent, you will get a mess of data
4. But it can be easier to see patterns with more data
5. Even with a mess, you can classify and interpret
Big Data

• What, where, when, how
• Multi-structured
• Collected by machines
• Broad
• Behaviors & actions of many people
• Collected as people do what they do
• People are not highly aware of data being collected
• Analysis uses statistical methods
✓ Transactional data
✓ Customer service logs
✓ Analytics
✓ A/B tests
✓ Social media posts
✓ Sensor data
Little Data

- It’s big data for one
- Focus on personal goals
- Individuals grant access, rather than companies
Thick Data

- How and why
- Description
- Collected by people
- In-depth
- Behaviors, actions, emotions, intentions, motivations of a few
- Collected as part of a study
- People are highly aware of data being collected
- Analysis includes developing codes, summaries, and themes
✓ Usability tests
✓ Contextual research
✓ Interviews
✓ Diaries
✓ Any study
The key is **PAIRING**
Public dataset & interviews
Social media analytics & lean
Emotion analytics & contextual
Customer data & interviews
<table>
<thead>
<tr>
<th>Public datasets</th>
<th>Interviews</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavioral analytics</td>
<td>Observations</td>
</tr>
<tr>
<td>Social media data</td>
<td>Lean research</td>
</tr>
<tr>
<td>Competitive data</td>
<td>Ideation</td>
</tr>
<tr>
<td>Emotional analytics</td>
<td>Diaries</td>
</tr>
</tbody>
</table>
Data & IMPROVING
A/B & analytics
Can you get Carer's Allowance?

What benefit does the person you care for get?

Don't include any benefits they've applied for if they haven't got a decision yet.

- Personal Independence Payment (PIP) daily living component
- Disability Living Allowance (DLA) - middle or highest care rate
- Attendance Allowance (AA)
- Constant Attendance Allowance (CAA)
- Armed Forces Independence Payment (AFIP)
- None of these benefits

Next
## IMPROVEMENT PAIRINGS

<table>
<thead>
<tr>
<th>A/B tests</th>
<th>Usability tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surveys</td>
<td>Intercept interviews</td>
</tr>
<tr>
<td>Customer service data</td>
<td>Interviews</td>
</tr>
<tr>
<td>Analytics</td>
<td>Usability tests</td>
</tr>
</tbody>
</table>
Data & Measurement

- Spring Temple Buddha: 153 m.
- Ushiku Daibutsu: 120 m.
- Statue of Liberty: 93 m.
- The Motherland Calls: 85 m.
- Christ the Redeemer: 39.6 m.
- 182 m. tallest in the world.
5 people in a group
1 measure per post-it

GO!
conversion rate
total revenue
profit margin
closed sales
cost per conversion
micro-conversions
viral coefficient
share of search
visits to purchase
net promoter score
growth of channel
referral traffic
number of visitors
user-generated content
social media sentiment
brand awareness
attention minutes
engaged time
page depth
total time reading
shares per campaign
perceived success
time on task
ease of use rating
cue recognition
system usability score
bounce rate
happiness rating
satisfaction rating
churn rate
task success
popular content
time on site
desired actions
navigation use
pageviews
back button presses
percent retained
uploads
RATIOS OVER COUNTS
“Friends don’t let friends measure page views, ever.

Avinash Kaushik
\[
\text{conversion rate} = \frac{\text{# of conversions}}{\text{# of visits}}
\]
\[
\frac{\text{# of conversions}}{\text{# of visits}} = \text{conversion rate} \uparrow 5\%
\]
TUNE IN TO THE

RIGHT SIGNALS
Single-signal
IMMERSE YOURSELF IN YOUR STAR WARS™ BATTLE FANTASIES

Available November 17, 2015 for PlayStation 4, Xbox One and on Origin for PC

WATCH THE LATEST TRAILER ▶
Multi-signal
D.R.E.A.M.
Dresses rule everything around me.
SHOP THE FINEST FROCKS

Big Day Bliss
SHOP STYLES FOR BRIDES & BEYOND

New Arrivals
Event stream
These photos from around the world should scare the pants off of McDonald's and Walmart.

From a small scientific network to the way your kids talk to Grandma, the web has come a long way.

April 15, 2015, was an incredible day for workers around the world.

A music video takes on a cultural obsession you may not know much about.

Everyone thought his ideas for making ice sounded crazy, but now…
ACCOUNT FOR BIAS
SIGNAL BIAS
(big data)

HEISENBERG UNCERTAINTY
PRINCIPLE (small data)
what is your bias
what is your own bias
what is your personal bias
what is your kpop bias
what does who is your bias mean
TIE IT TOGETHER WITH A FRAMEWORK
<table>
<thead>
<tr>
<th>METRICS</th>
<th></th>
</tr>
</thead>
</table>
| **ACQUISITION** | ▪ Subscribed to private Beta waiting list  
▪ Signed up for email newsletter  
▪ Engagement (# of interactions) |
| **ACTIVATION** | ▪ Opened the app and used it once  
▪ Created first project |
| **RETENTION** | ▪ Used the app two times in 30 days  
▪ Played a game to the next level  
▪ Ran a custom report |
| **REFERRAL** | ▪ Invited a friend  
▪ Shared to social media  
▪ Use of promo code |
| **REVENUE** | ▪ Monthly recurring revenue  
▪ Life time value |
<table>
<thead>
<tr>
<th>GOAL</th>
<th>SIGNALS</th>
<th>METRICS</th>
</tr>
</thead>
</table>
| **HAPPINESS** For users to feel like the site is unique | A satisfaction rating from a survey | ▪ Satisfaction  
▪ Perceived ease of use  
▪ Net promoter score |
| **ENGAGEMENT** For users to keep discovering more content | The amount of time people spend viewing from analytics | ▪ Number of videos watched per user per week  
▪ Number of uploads  
▪ Number of shares |
| **ADOPTION** For users to use the site or app | The number of signups per day from analytics | ▪ Upgrades to latest version  
▪ New subscriptions  
▪ Purchases by new users |
| **RETENTION** For users to continue to use the site or app | The number of return customers from customer data | ▪ Number of active users  
▪ Renewal rate  
▪ Repeat purchases |
| **TASK SUCCESS** For users to accomplish their goal | The number of successful completes from a study | ▪ Search result success  
▪ Time to upload  
▪ Profile creation completion |
# SEE-THINK-DO FRAMEWORK

<table>
<thead>
<tr>
<th>AUDIENCE</th>
<th>CONTENT</th>
<th>INTERACTIONS</th>
<th>METRICS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SEE</strong></td>
<td><strong>AUDIENCE</strong></td>
<td><strong>CONTENT</strong></td>
<td><strong>INTERACTIONS</strong></td>
</tr>
<tr>
<td>People who wear clothes</td>
<td>Blog</td>
<td>Search</td>
<td>Amplification</td>
</tr>
<tr>
<td>People who wear clothes &amp; think they need some</td>
<td>Style gallery</td>
<td></td>
<td>Applause</td>
</tr>
<tr>
<td>People who wear clothes &amp; think they need some &amp; are looking to buy right now</td>
<td>Pinterest</td>
<td></td>
<td>% new visits</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Brand awareness</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Page depth</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Engaged time</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>% assisted</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cart abandonment</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Conversion rate</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Revenue</td>
</tr>
</tbody>
</table>
Measure with Meaning
Marketplace
Personal well-being
Collective well-being
Water efficiency credits promote smarter use of water, inside and out, to reduce potable water consumption.

Energy and atmosphere credits promote better building energy performance through innovative strategies.

Sustainable sites credits encourage strategies that minimize the impact on ecosystems and water resources.

Indoor environmental quality credits promote better indoor air quality and access to daylight and views.

Innovation credits address sustainable building expertise as well as design measures not covered under the five LEED credit categories.

Regional priority credits address regional environmental priorities for buildings in different geographic regions.
7.8 Safety
2.9 Health
5.6 Housing
6.9 Access to services
3.4 Civic Engagement
8.1 Education
7.9 Jobs
5.4 Income
10.0 Environment
7 Steps to designing with data
1 Frame the question
What we think we know

- Competitors
- People
- Behaviors

What we know we don’t know

- Outliers
- Untapped sources
Be a bit Bayesian

- Updates beliefs
- Iterative
- Incorporates multiple sources
- Yes, and...
- Never 100%
2 Check your values
Data with a Soul

Don’t collect it without a purpose
Use only what is needed
Use data with empathy
Choose a partner
Ask your data points

GO!
Data unto others as you would have others data unto you.

Pamela P
Start with data on hand
Behavioral Analytics

- Piwik, Clicky
- Google
- Chartbeat, Crazy Egg, Mixpanel, KISS metrics
Get paid anywhere.

Accept credit cards and checks wherever your customers are – with PayPal Here<sup>™</sup>.

Watch how it works

Download the app to get the free card reader.

Already have PayPal Here? Quick Tips to Get Started

Works with iPhone and Android devices only.
Site analytics for **NAVIGATION & SEARCH**
Combo analytics for CONTENT
Understanding Expectations

- Acquisition > Referrals
- Acquisition > Social
- Channels > Organic Search
- Bounce Rates
- User Flows
Gauging Interest

- Audience > Demographics
- Audience > Interests
- Ready-made segments
- Behavior > New & Returning Visitors
- Behavior > Site Content
- Behavior > Site Search
- Behavior > In-page
Informing a Customer Journey

- Behavior > Behavior Flow
- Acquisition > User Flow
- Funnels
- Event Tracking
- Audience > Technology & Mobile
- ALSO: Alexa and Compete
Competitive analytics for JOURNEY
Over 26.77% of traffic is from Referrals

**Top Referring Sites:**
- Bookingbudd...
- Airfarewatch...
- Cheapflights.c...
- CNN.com
- Expedia.com

**Top Destination Sites:**
- Aa.com
- Airbnb.com
- Justfly.com
- AlaskaAir.com
- Facebook.com

See 1,308 More Referring Sites
See 89 More Destination Sites
Improve your Analytics skills with free online courses from Google.

View lessons from experts
View lessons from Google's digital analytics experts at your own pace.

Test your knowledge
Apply what you learn by completing short quizzes and practice exercises.

Join the learning community
Engage with other students and analytics experts in course forums and on Google+.

Courses

Google Tag Manager Fundamentals
My last post, perhaps provocatively, called for a reduction of data in executive dashboards (digital, online, offline). More English (IABI, specifically) would lead to a smarter understanding of performance, and of course glory for data practitioners.

Here's the post: Strategic & Tactical Dashboards: Best Practices, Examples.

In the post Adil commented that he's observed that attribution modeling is missing from most web analytics dashboards. My reply was that attribution modeling might not be present in a physical manner, but that it should be there below the surface to ensure our executives have a smarter understanding of a channel's true contribution.
R Is Not So Hard! A Tutorial, Part 18: Re-Coding Values

by David Lillis, Ph.D.

One data manipulation task that you need to do in pretty much any data analysis is recode data. It’s almost never the case that the data are set up exactly the way you need them for your analysis.

In R, you can re-code an entire vector or array at once. To illustrate, let’s set up a vector that has missing values.
NEW - SEE THIS INDEX LAID OUT IN A TEXT FORMAT
Social Analytics

Shoutlet
Reachli
Trackur

Tagboard
Social Mention
Open Status
Google Trends
Social media analytics for **EMOTION**
Mentions about "Amy Schumer"

Sort By: Date | Results: Anytime

- **Chris Hardwick, Amy Schumer stickers.**
  - Chris Hardwick, Amy Schumer stickers. - 20130418_214451_zps660060c0.jpg
  - 3 minutes ago - by krinting on photobucket

- **Amy Schumer pranks red carpet**
  - Amy Schumer pranks red carpet - AmySchumer_RedCarpetPrank-EleMagazine_zps660060c0.jpg
  - 3 minutes ago - by krinting on photobucket

- **Amy Schumer at MTV MA**
  - Amy Schumer at MTV MA - AmySchumer_MTVMA-EOntLine_zps7q3q56.jpg
  - 3 minutes ago - by krinting on photobucket

- **Inside Amy Schumer 02**
  - Inside Amy Schumer 02 - InsideAmySchumer_newseason_zpsem01thl.jpg
  - 2 minutes ago - by krinting on photobucket

- **Inside Amy Schumer 3**
  - Inside Amy Schumer 3 - InsideAmySchumer_newseason_zpsem01thl.jpg
  - 2 minutes ago - by krinting on photobucket

- **Amy Schumer cropped**
  - Amy Schumer cropped - Amy_Schumer-1.jpg
  - 2 minutes ago - by brocmen on photobucket

- **Inside Amy Schumer**
  - Inside Amy Schumer - InsideAmySchumer_Vulture_zps4q0.jpg
  - 3 minutes ago - by krinting on photobucket

- **Amy Schumer x**
  - Amy Schumer x - AmySchumer-THRmovies_zps1ym4pp2.jpg
  - 2 minutes ago - by krinting on photobucket

---

**Sentiment**
- positive: 12
- neutral: 62
- negative: 8

**Top Keywords**
- schumer: 124
- cosby: 28
- laik: 24
- submitted: 23
- style: 20
- laughing: 20
- schumeranthony: 20
- kinos: 20
- skull: 20

**Top Users**
- krinting: 12
- brigida: 10
- feek731: 4
- clarkkazo: 2
- brocmen: 2
- ilovepetersburg: 2
- Prasad: 1
- Tovis13: 1
- Meghan: 1
- lizzybuczak: 1
- kakachihalata: 1
What's your biggest challenge?

I'm frustrated with the lack of leadership from my team lately. It feels like nobody wants to take ownership for what we're building.

- STEPHEN BRENN
  - Software Engineer
    - Platform team
    - Dublin, Ireland
    - Disengaged
    - See more details
Primary Mood

DISAGREE
Talkativeness, extroversion.

AGREE

Secondary Mood

DISAGREE
Very charismatic. Ideological leadership. Verbal power.

AGREE

Got it wrong?
Read these tips & try again
Answers: 1-b, 2-a, 3-c, 4-b, 5-c, 6-c, 7-b, 8-a, 9-c, 10-a, 11-c, 12-a, 13-b, 14-c. As for 15, Clinton was +227, Dubya was -25, Poppy Bush was +45, Reagan was +135, and Obama is +127 so far. Now obviously, a president can’t be fully credited or blamed for what a number happens to be the day he took and left office, but even so, it seems apparent that with two years to go, Obama’s Dow number will overtake Reagan’s.
4 Fill in the WHY
Research methods by **Data Source** vs. **Approach** vs. **Context of Product Use**

- **Behavioral**
  - Eyetracking
  - Usability Benchmarking (in lab)
  - Usability Lab Studies
  - Ethnographic Field Studies
  - Online User Experience Assessments ("Vividence-like" studies)

- **Attitudinal**
  - Participatory Design
  - Focus Groups
  - Phone Interviews
  - Diary/Camera Study
  - Message Board Mining
  - Customer feedback via email
  - Desirability studies
  - Cardsorting
  - Intercept Surveys
  - Email Surveys

---

**Key for Context of Product Use during data collection**

- **Natural use of product**
- **Scripted (often lab-based) use of product**
- **De-contextualized / not using product**
- **Combination / hybrid**

© 2008 Christian Rohrer
<table>
<thead>
<tr>
<th></th>
<th>January</th>
<th>April</th>
<th>July</th>
<th>October</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer Service Log Review</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Site and Search Log Analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>User Research/Field Study</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usability Testing (live/rem)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surveys/NPS Study</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Product Releases</td>
<td>⭐️</td>
<td>⭐️</td>
<td>⭐️</td>
<td>⭐️</td>
</tr>
<tr>
<td>NAVIGATION</td>
<td>P1</td>
<td>P2</td>
<td>P3</td>
<td>P4</td>
</tr>
<tr>
<td>--------------------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>noticed top nav</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>understood R &amp; S category</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>trouble with drop-down (slippery)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>missed athletes at first</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>missed tv listings at first</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>got icon for home</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>noticed travelling nav</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>used bottom nav</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| HOME PAGE           |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| too much going on   |    |    |    |    |    |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |
| noticed highlights & replays |    |    |    |    |    |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |
| scrolled feed       |    |    |    |    |    |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |
| noticed video to the side of main feature |    |    |    |    |    |    |    |    |    |     |     |     |     |     |     |     |     |     |     |
| scrolled past medals |    |    |    |    |    |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |
| noticed link to more (medals or news) |    |    |    |    |    |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |

| RESULTS & SCHEDULES |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| weren't sure how to get schedule |    |    |    |    |    |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |
| couldn't tell b/t tv and this schedule |    |    |    |    |    |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |
| missed results/medals |    |    |    |    |    |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |
| used pull-down      |    |    |    |    |    |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |

| LIVE VIDEO          |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
| didn't get that it was live |    |    |    |    |    |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |
| understood login    |    |    |    |    |    |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |
| had authenticated before |    |    |    |    |    |    |    |    |    |     |     |     |     |     |     |     |     |     |     |     |

**You can QUANTIFY QUALITATIVE**
5 Test alternatives
A/B Testing Tools

- Optimizely
- Unbounce
- Monetate

- Apptimize
- Tplytics
Vodafone – Product Page – 2015 Honorable Mention Winner

Web Testing Award Winner: 2015

How good is your gut?

Guess which version you think won... and then see real-life test results:

[VERSION A] OR [VERSION B]

Or click to see results without guessing
A/B testing for CHOICES

Version A  Version B (Winner)
"Keeping track of our business contacts was a pain when we all shared a single spreadsheet. Highrise is much better."

- Jocelyn, Senior Manager at One Design

"Searching through old email conversations was a pain. Highrise makes it easy to remember everything."

- Michael, Owner of NES, Accounting

**4.78%**
Ceci n’est pas un data viz.
End with questions
Data is a great input but it doesn’t always work as an output.
Designing with Data

1. Frame the question
2. Start from empathy
3. Explore data on hand
4. Add the why
5. Test alternatives
6. Do something
7. End with more questions
Get the free e-book
oreilly.com

Data-Informed Product Design
Pamela Pavliscak