CTR & Other Data Myths

Become a data skeptic

Britney Muller
Founder, Data Sci 101
@BritneyMuller

MozCon
Virtual

Wag Dodge

MozCon
Virtual

400+ households evaluated a new TV pilot:

"No segment of the audience was eager to watch the show again."

MozCon
Virtual

400+ households evaluated a new TV pilot:

"You can't get too excited about going to the Laundromat"

Seinfeld

MozCon
Virtual

Superforecasters

77%
Average percentage of days on which our forecasts placed the highest probability on the "correct" outcome

74%
Average probability that Superforecasters assigned to the "correct" outcome over all forecasting days

SUPERFORECASTING
THE ART AND SCIENCE OF PREDICTION
PHILIP E. TETLOCK
DAN GARDNER

Guess what the most important driver of Superforecasting success is?

MozCon Virtual

It's easy to be led astray by data

Storks Deliver Babies ($p = 0.008$)

Robert Matthews
Aston University, Birmingham, England.
e-mail: rajm@compuserve.com

Summary
This article shows that a highly statistically significant correlation exists between stork populations and human birth rates across Europe. While storks may not deliver babies, unthinking interpretation of correlation and p-values can certainly deliver unreliable conclusions.

◆ INTRODUCTION ◆

Introductory statistics textbooks routinely warn of the dangers of confusing correlation with causation, pointing out that while a high correlation coefficient is indicative of (linear) association, association between storks and the concept of women as bringers of life, and also in the bird's feeding habits, which were once regarded as a search for embryonic life in water (Cooper 1992). The legend lives on to this day, with neonate-bearing storks being a regular feature of greetings cards celebrating births.

Spurious relationship: when two or more variables or events are associated but not causally related.
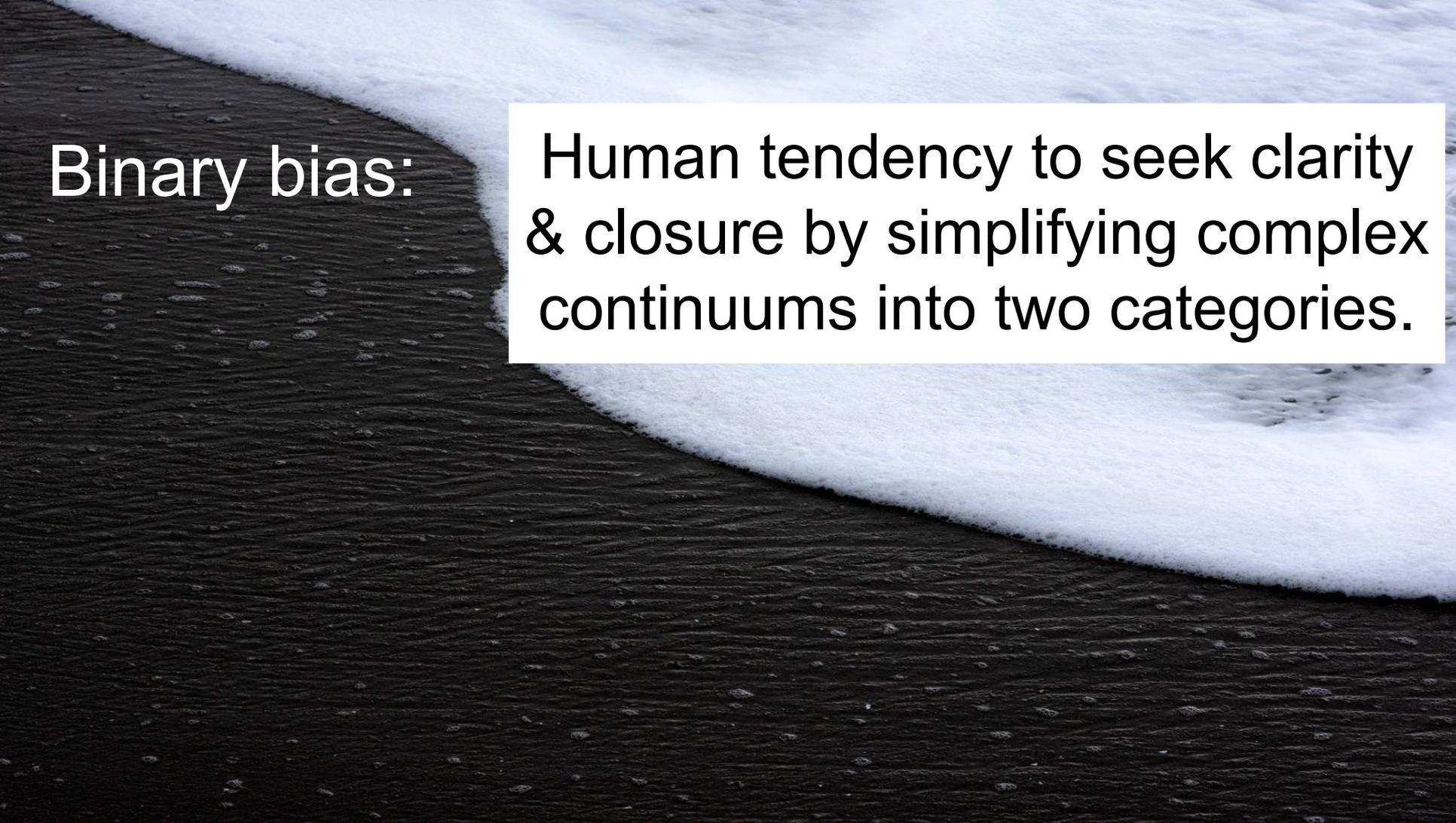
"But, I'm smart & not as likely to fall for these errors…"

MozCon
Virtual

The smarter you are ≠ more mental dexterity

In fact, the higher you score on an IQ test, the more likely you are to fall for stereotypes because you're faster at recognizing patterns.

Binary bias: Human tendency to seek clarity & closure by simplifying complex continuums into two categories.

Deep forces are behind our resistance to rethinking:

1. Makes the world more unpredictable

2. Requires us to admit facts may have changed

3. Threatens our identities

MozCon
Virtual

What data should we be rethinking?

Client Example:

Client: Why did X page's CTR go from 5.5% to 1.9%?

Me: Let me look into it [sure enough, the page's CTR tanked]

MozCon
Virtual

Client Example:

Client: Why did X page's CTR go from 5.5% to 1.9%?

Me: Let me look into it [sure enough, the page's CTR tanked]

Meanwhile:

- Clicks were up 5,824%!
- Impressions were up 18,144%!
- Rank was up 85%!

The page started to rank for a large number of new keywords, which brought the previous low impression:low clicks CTR (5.5%) way down.

MozCon Virtual

*example modified to protect an NDA

CTRs can be misleading

Beware of low numbers that can easily deceive:

Case A: 2 Impressions / 2 Clicks = **100% CTR**

Case B: 1,000 Impressions / 300 Clicks = **30% CTR**

Which would you prefer?

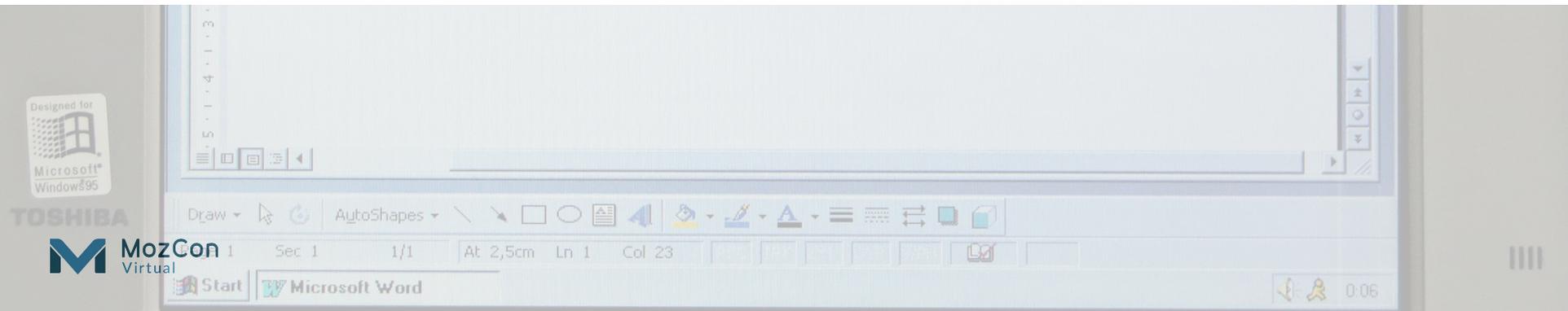MozCon
Virtual

example modified to protect an NDA

Website traffic falling can be a good thing
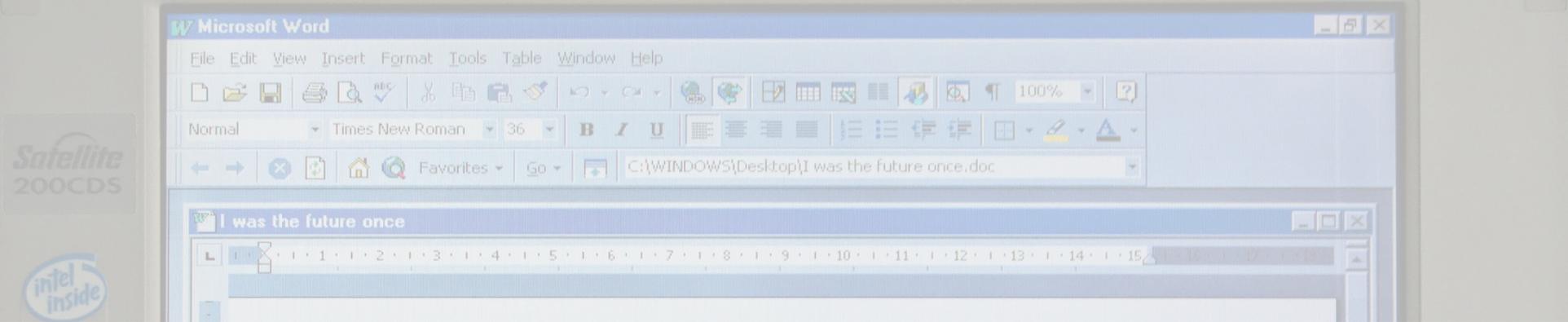
"We laugh at people who still use Windows 95, yet we still cling to opinions that we formed in 1995."
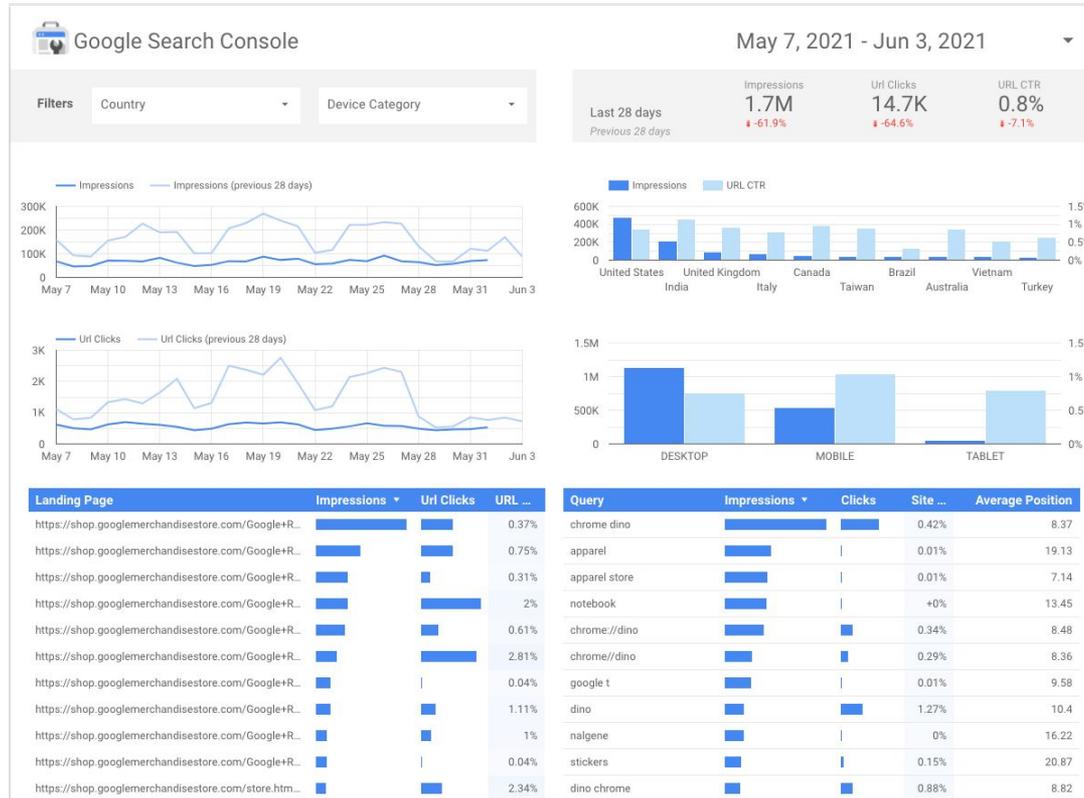—Adam Grant

Bounce rate going up can be a good thing!

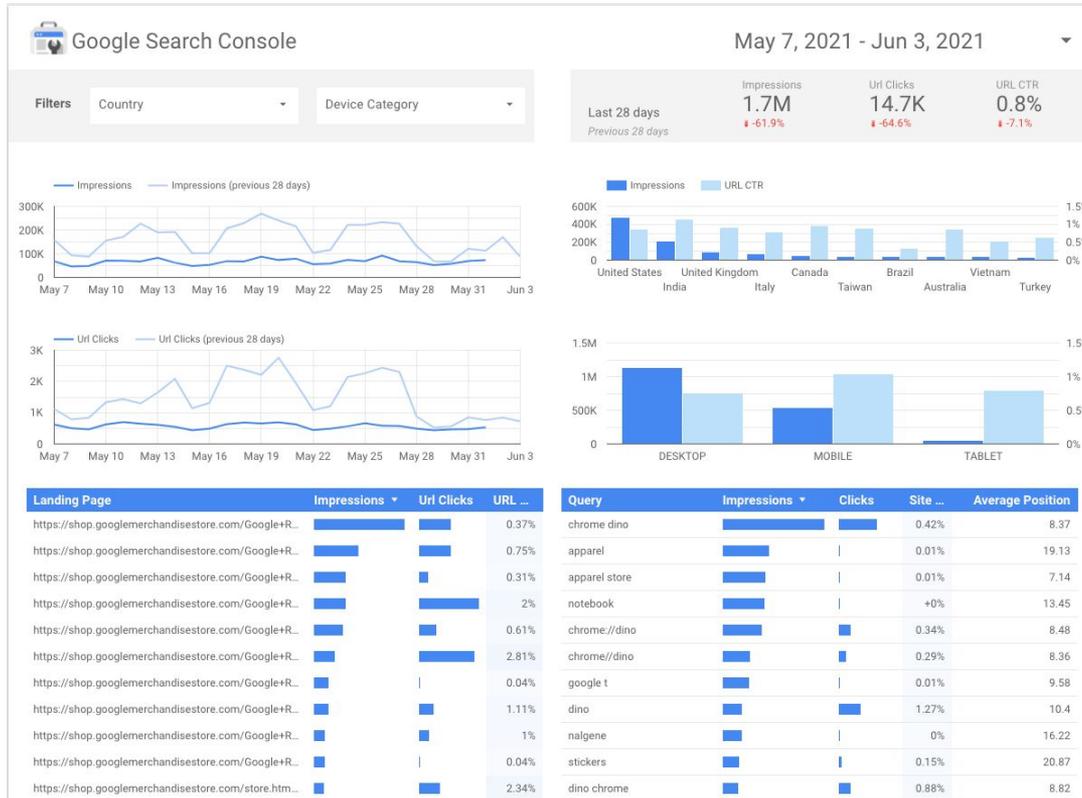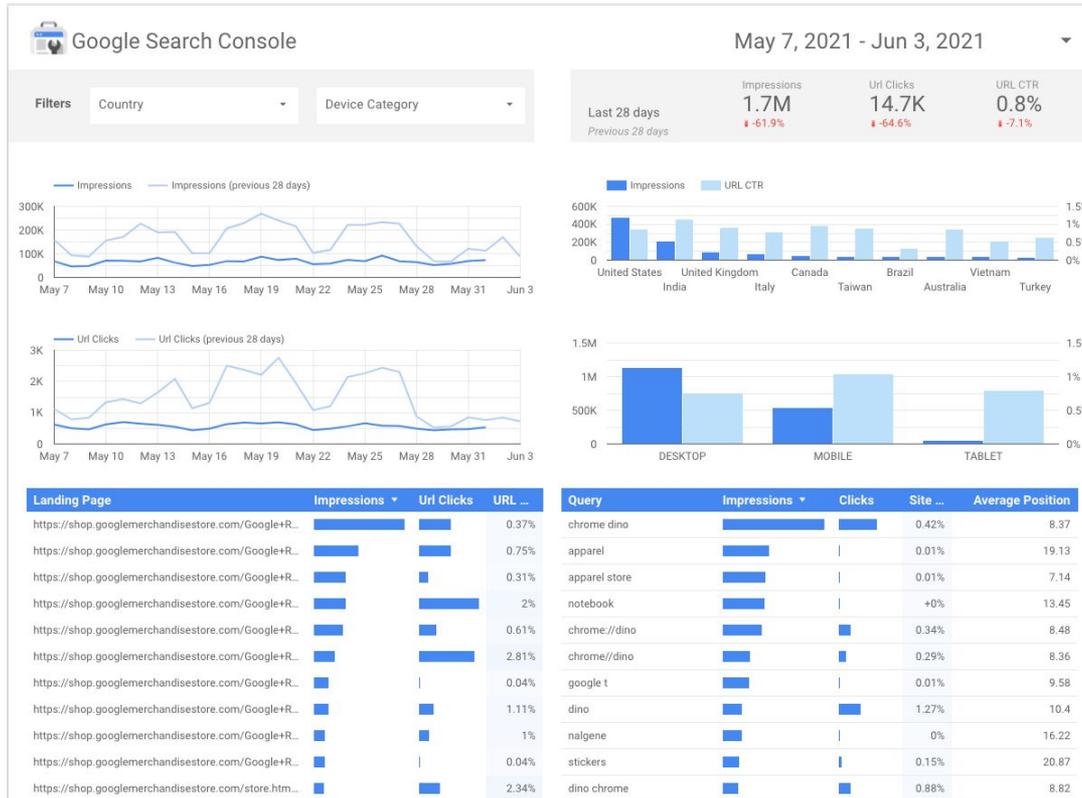# How do we determine what metrics are good vs. bad?

What does your sitewide 'CTR' really tell you?

What does 'Average Position' really tell you?

What does the # of ranking keywords really tell you?

We often favor <u>feeling</u> right over <u>being</u> right

Let's try to view CTR & other metrics like Wag

MANN GULCH FAME

WAG. DODGE 1944

What aren't we considering...

- Has the # (and quality) of ranking keywords changed?

- What sections of your site get the most clicks?

- What does your CTR curve look like with branded & non-brand queries?

- What are your keyword opportunities

Power Up Your Data Skeptic Skills!!!

✍️ **Keywords:**

- Has the # of ranking keywords gone up or down?
  - *How might this affect other metrics?*
- What is the ranking distribution of keywords
  - *Has this improved from previous period?*
- What keywords are within "striking distance"?
  - *Keywords with solid CTR but on page two 10-20.*
- What are the new vs lost keywords since previous period?
- How are branded vs nonbrand keywords doing?

🖱️ **Clicks:**

- Are total clicks up from previous period?
- How have top clicked pages changed from previous period?

📄 **Pages:**

- What pages are ranking for the same keywords (cannibalization)
- What pages have dropped in performance?
- What pages have improved?
- What are the top page categories on your site?

bit.ly/mozcon-21
bit.ly/mozcon-21-2

MozCon
Virtual

Colab allows you to write and execute Python in your browser

🔺 MozCon - CTR Data Skeptic ☆

File   Edit   View   Insert   Runtime   Tools   Help

💬 Comment   👥 Share   ⚙

+ Code   + Text   RAM / Disk   ✏ Editing

Import Libraries

```
[1]    import csv
       import json
       import requests
       import pandas as pd
       import numpy as np
       import re
       import seaborn as sns
       from matplotlib import pyplot as plt

       from IPython.display import display
```

Upload all CSVs

```
[9]    from google.colab import files
       upload = files.upload()
```

Choose Files   3 files

- **GSC - Data Skeptic Dashboard for SEOs_Page 1_Table (3).csv**(text/csv) - 846529 bytes, last modified: 6/16/2021 - 100% done
- **GSC - Data Skeptic Dashboard for SEOs_Page 1_Table (4).csv**(text/csv) - 799989 bytes, last modified: 6/16/2021 - 100% done
- **GSC - Data Skeptic Dashboard for SEOs_Page 1_Table (5).csv**(text/csv) - 4469323 bytes, last modified: 6/16/2021 - 100% done

Colab allows you to write and execute Python in your browser

bit.ly/colab-yt

🧰 Google Search Console

Filters

🔲 Default Data
Click to select your data ▾

Search Console Query Report - May

Query	Impressions	Clicks ▾	Site CTR	Average Position
google kickball	1,150	750	65.22%	1
chrome dino	125,918	612	0.49%	8.27
google merchandise store	14,183	514	3.62%	1.03
youtube merchandise	2,635	437	16.58%	1.23
google t shirt	1,884	353	18.74%	1
dino game	67,776	349	0.51%	8.21
dino	26,171	311	1.19%	10.46
youtube merch	2,531	283	11.18%	1.19
dinosaur game	99,671	229	0.23%	11.26

1 - 20 / 24916 ‹ ›

Search Console Query Report - April

Query	Impressions	Clicks ▾	Site CTR	Average Position
dinosaur game	1,829,407	11,858	0.65%	8.1
dino game	315,945	5,945	1.88%	5.71
chrome dino	162,588	2,150	1.32%	7
google dinosaur game	118,681	1,601	1.35%	6.15
dino	57,069	819	1.44%	9.82

MozCon
Virtual

Export all tables (CSV)

Search Console Query Report - May

Query	Impressions	Clicks ▾	Site CTR	Average Position
google kickball	1,150	750	65.22%	
chrome dino	125,918	612	0.49%	8.2
google merchandise store	14,183	514	3.62%	1.0
youtube merchandise	2,635	437	16.58%	1.2
google t shirt	1,884	353	18.74%	
dino game	67,776	349	0.51%	8.2
dino	26,171	311	1.19%	10.4
youtube merch	2,531	283	11.18%	1.1
dinosaur game	99,671	229	0.23%	11.26

1 - 20 / 24916 ‹ ›

Sort by ▶
Reset
Download CSV
Download CSV (Excel)
Export to Sheets
Explore

Search Console Query Report - April

Query	Impressions	Clicks ▾	Site CTR	Average Position
dinosaur game	1,829,407	11,858	0.65%	8.1
dino game	315,945	5,945	1.88%	5.71
chrome dino	162,588	2,150	1.32%	7
google dinosaur game	118,681	1,601	1.35%	6.15
dino	57,069	819	1.44%	9.82

Do you rank for more or fewer keywords?

```
#May keywords minus April keywords:
len(df)-len(dfp)
```

1983

```
#You currently rank for 1983 more keywords than you did in April
#Note: This does NOT mean you rank for 1983 NEW KEYWORDS! **we'll get to that :)**
#Is this good or bad? --Let's try to uncover more info around the quality of keywords
```

What keywords have you gained?

```
df_outer[df_outer['April Rank'].isna()]
```

	Query	May Impressions	May Clicks	May CTR	May Rank	May	April Impressions	April Clicks
61	google kick ball	57.0	32.0	0.561404	1.000000	May	NaN	NaN
117	グーグルマーチャンダイズストア	18.0	15.0	0.833333	1.111111	May	NaN	NaN
139	/google+redesign/apparel/mens/mens+t+shirts	15.0	12.0	0.800000	1.000000	May	NaN	NaN
168	google camp charleston	77.0	9.0	0.116883	2.259740	May	NaN	NaN
198	googleマーチャンダイズストア	13.0	8.0	0.615385	1.000000	May	NaN	NaN
...	...	...	...	...	...	...	...	...
24900	bicycle turquoise	1.0	0.0	0.000000	1.000000	May	NaN	NaN
24901	google canada shop	13.0	0.0	0.000000	6.384615	May	NaN	NaN
24904	it's bags	1.0	0.0	0.000000	11.000000	May	NaN	NaN
24906	google get on youtube	1.0	0.0	0.000000	58.000000	May	NaN	NaN
24907	bags like timbuk2	4.0	0.0	0.000000	74.500000	May	NaN	NaN

9881 rows × 11 columns

What keywords have you lost?

```
df_outer[df_outer['May Rank'].isna()]
```

	Query	May Impressions	May Clicks	May CTR	May Rank	May	April Impressions	April Clicks
24916	google+redesign/shop+by+brand/youtube	NaN	NaN	NaN	NaN	NaN	13.0	8.0
24917	qwiklabs perk store	NaN	NaN	NaN	NaN	NaN	8.0	7.0
24918	google baby onesie	NaN	NaN	NaN	NaN	NaN	11.0	6.0
24919	google store clothing	NaN	NaN	NaN	NaN	NaN	10.0	5.0
24920	google bag price	NaN	NaN	NaN	NaN	NaN	15.0	5.0
...	...	...	...	...	...	...	...	...
32809	will tennyson merch	NaN	NaN	NaN	NaN	NaN	1.0	0.0
32810	phone stand for desk wood	NaN	NaN	NaN	NaN	NaN	1.0	0.0
32811	google or	NaN	NaN	NaN	NaN	NaN	2.0	0.0
32812	sticker cell	NaN	NaN	NaN	NaN	NaN	1.0	0.0
32813	branded shirt price	NaN	NaN	NaN	NaN	NaN	1.0	0.0

7898 rows × 11 columns

Sort lost keywords by clicks

Magically adds a sortable/filterable table

```
%load_ext google.colab.data_table
df_outer[df_outer['May Rank'].isna()]
```

Query	May Impressions	May Clicks	May CTR	May Rank	May	April Impressions	April Clicks ▼
google+redesign/shop+by+brand/youtube	NaN	NaN	NaN	NaN	NaN	13.0	8.0
qwiklabs perk store	NaN	NaN	NaN	NaN	NaN	8.0	7.0
google baby onesie	NaN	NaN	NaN	NaN	NaN	11.0	6.0
google cloud certification goodies	NaN	NaN	NaN	NaN	NaN	18.0	5.0
google bag price	NaN	NaN	NaN	NaN	NaN	15.0	5.0
google store clothing	NaN	NaN	NaN	NaN	NaN	10.0	5.0
google brand clothing	NaN	NaN	NaN	NaN	NaN	14.0	4.0
google merchandise store singapore	NaN	NaN	NaN	NaN	NaN	11.0	4.0
dinosauar game	NaN	NaN	NaN	NaN	NaN	93.0	3.0

Search for specific keywords inside table:

index: [____] to [____]   Query: [_____]   May Impressions: [____] to [____]   May Clicks: [____] to [____]   May CTR: [____] to [____]   May Rank: [____] to [____]

May: [____] to [____]   April Impressions: [____] to [____]   April Clicks: [____] to [____]   April CTR: [____] to [____]   April Rank: [____] to [____]   April: [_____]

Search by all fields:
[trex]

index	Query	May Impressions	May Clicks	May CTR	May Rank	May	April Impressions	April Clicks ▼	April CTR	April Rank
25023	trexx game	NaN	NaN	NaN	NaN	NaN	1.0	0.0	0.0	
25048	game trex google	NaN	NaN	NaN	NaN	NaN	2.0	0.0	0.0	
25092	trex running game	NaN	NaN	NaN	NaN	NaN	2.0	0.0	0.0	
25134	chrome trex game code	NaN	NaN	NaN	NaN	NaN	4.0	0.0	0.0	
25329	trex shop	NaN	NaN	NaN	NaN	NaN	4.0	0.0	0.0	
25375	jumping trex	NaN	NaN	NaN	NaN	NaN	1.0	0.0	0.0	
25377	dino trex google game	NaN	NaN	NaN	NaN	NaN	6.0	0.0	0.0	11.66666666666
25617	dino game trex	NaN	NaN	NaN	NaN	NaN	1.0	0.0	0.0	

MozCon
Virtual

Pull all questions from GSC

```
questions = cln[(cln['Query'].str.contains('(does|who|will|why|what|where|when|how|can|\?|was|were|being|been|versus|vs|vs\.|best)',
                                              regex=True)==True)]
questions.head(10)
```

Query	May Impressions	May Clicks	May CTR	May Rank	April Clicks	Brand vs Non
how to get google t shirt	37.0	12.0	0.324324	3.081081	9.0	Brand
how to get google developer t-shirt for free 2021	58.0	4.0	0.068966	10.896552	4.0	Brand
how to get google stickers	16.0	3.0	0.187500	2.187500	6.0	Brand
how much is a google bike worth	25.0	3.0	0.120000	3.040000	4.0	Brand
how to get google cloud goodies	26.0	2.0	0.076923	5.115385	1.0	Brand
google store cancel order	70.0	2.0	0.028571	6.828571	0.0	Brand
how to get google t shirt for free	25.0	1.0	0.040000	6.240000	0.0	Brand
how long does google fi take to ship	50.0	1.0	0.020000	4.560000	0.0	Brand
what is google merchandise store	79.0	1.0	0.012658	1.379747	3.0	Brand

MozCon Virtual

Create a brand vs. non-brand column:

```python
#The "google|chrome|youtube" below is regex and can be modified for your data:
cln["Brand vs Non"]=cln.Query.str.contains("google|chrome|youtube")
mask = cln.applymap(type) != bool
d = {True: 'Brand', False: 'Non-Brand'}
cln = cln.where(mask, cln.replace(d))
cln.head(10)
```

	Query	May Impressions	May Clicks	May CTR	May Rank	April Clicks	Brand vs Non	Rounded Rank
0	google kickball	1150.0	750.0	0.652174	1.003478	74.0	Brand	1
1	chrome dino	125918.0	612.0	0.004860	8.268627	2150.0	Brand	8
2	google merchandise store	14183.0	514.0	0.036241	1.031728	621.0	Brand	1
3	youtube merchandise	2635.0	437.0	0.165844	1.228463	404.0	Brand	1
4	google t shirt	1884.0	353.0	0.187367	1.002123	343.0	Brand	1
5	dino game	67776.0	349.0	0.005149	8.213025	5945.0	Non-Brand	8

MozCon
Virtual

Evaluate ranking distribution

```python
ax = brand_ranking_pivot.head(20).plot(kind='bar')
ax.set_ylabel('Count of Queries')
ax.set_title('Brand Ranking Distribution')
```

Evaluate brand vs. non-brand rank distribution

*Note the y-axis value changes
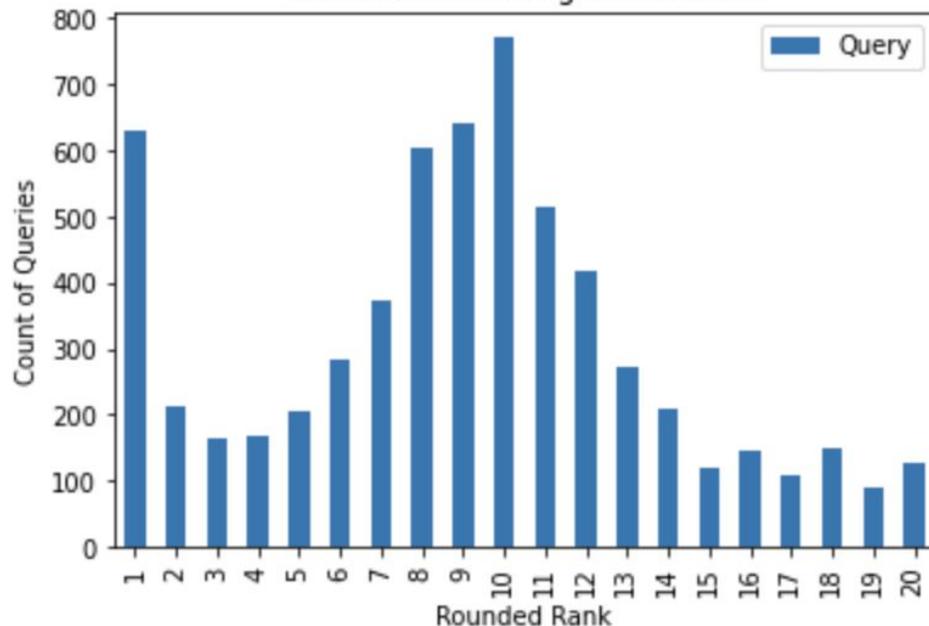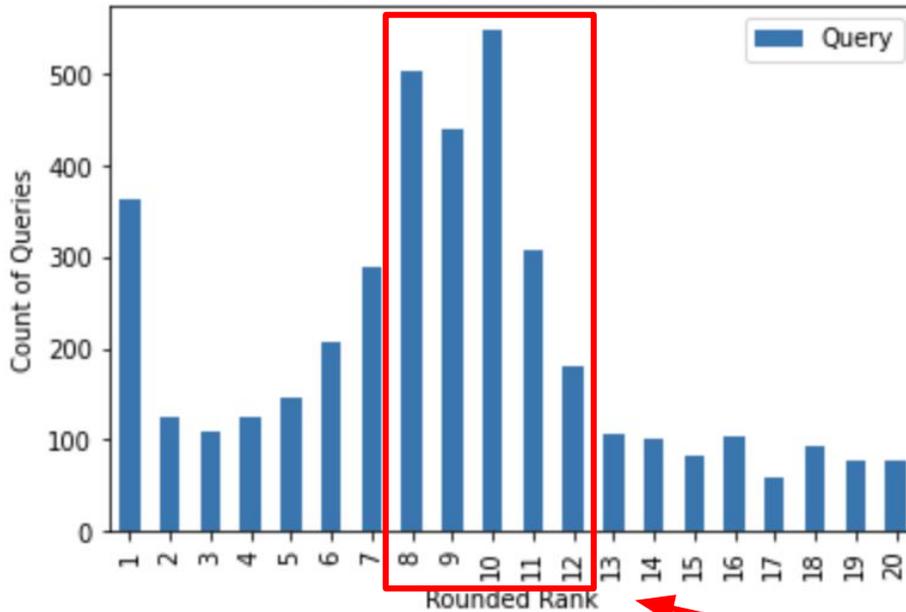
Evaluate brand vs. non-brand rank distribution
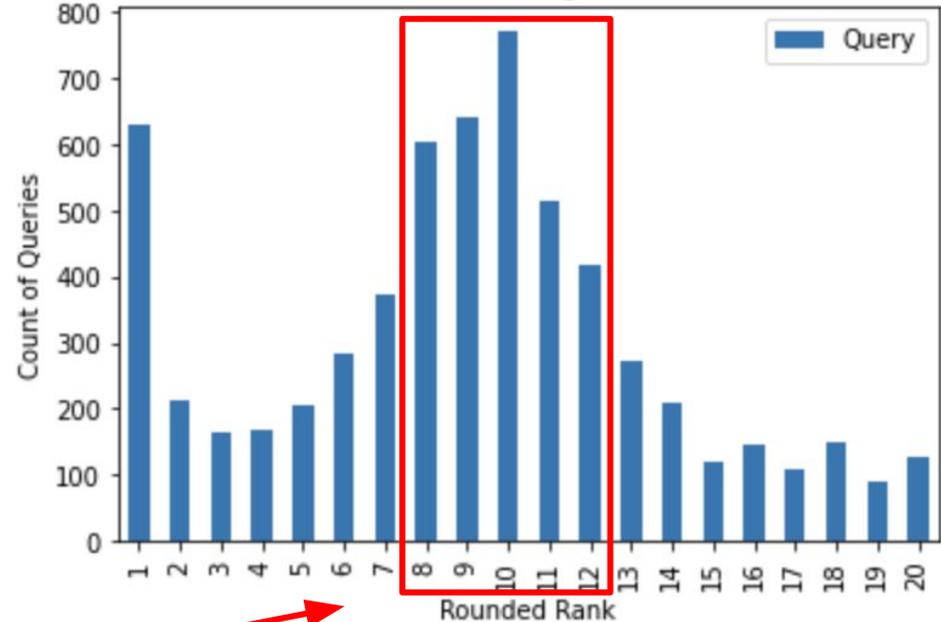
Potential keyword improvement opportunities

MozCon
Virtual

*Note the y-axis value changes

Visualize (interactive) rank vs. clicks

Organic clicks vs. rank data
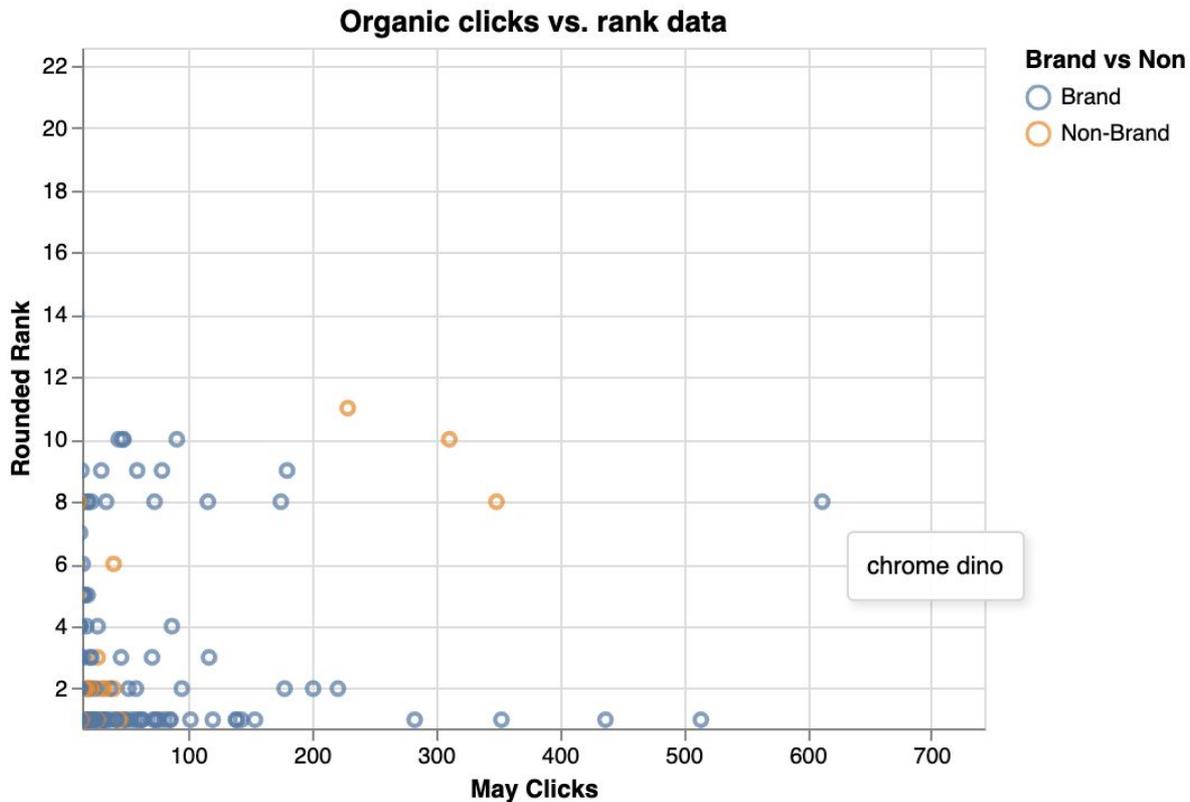
Identify keyword cannibalization

Number of competing landing pages for query

Query	Landing Page	URL Clicks
site:shop.googlemerchandisestore.com	119	0
google merchandise store	54	533
google merchandise	43	162
shop.googlemerchandisestore	41	10
if the google merchandise store	38	0
shop.googlemerchandisestore.com	33	0
google lanyard	30	0
google gift shop	29	2

Brand vs. non-brand clicks by category
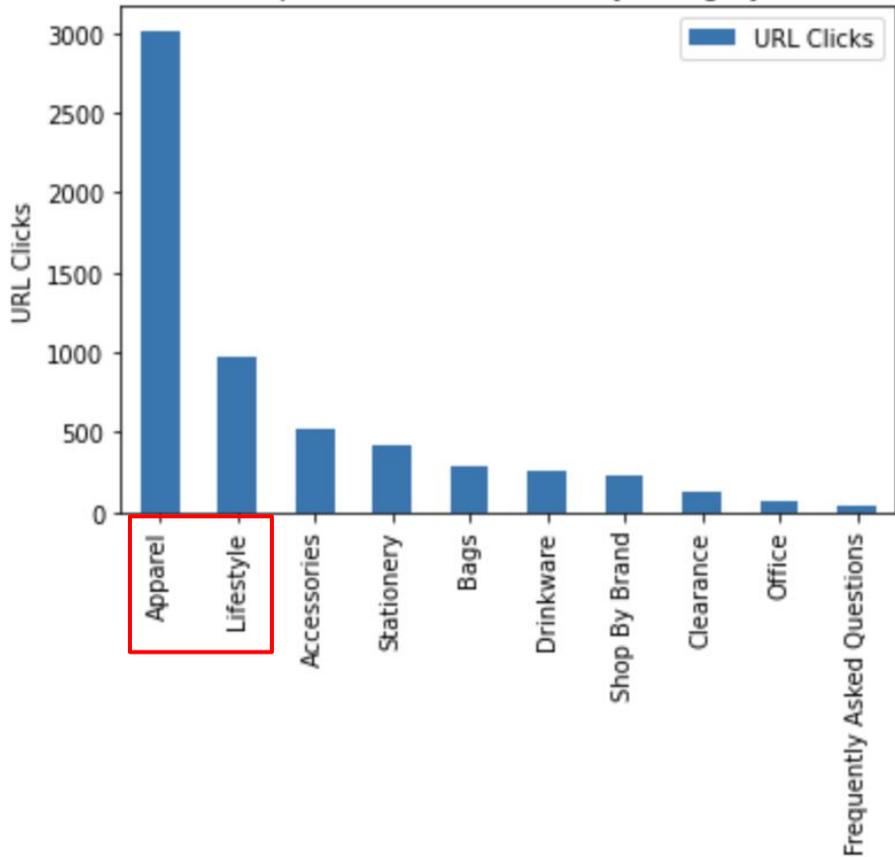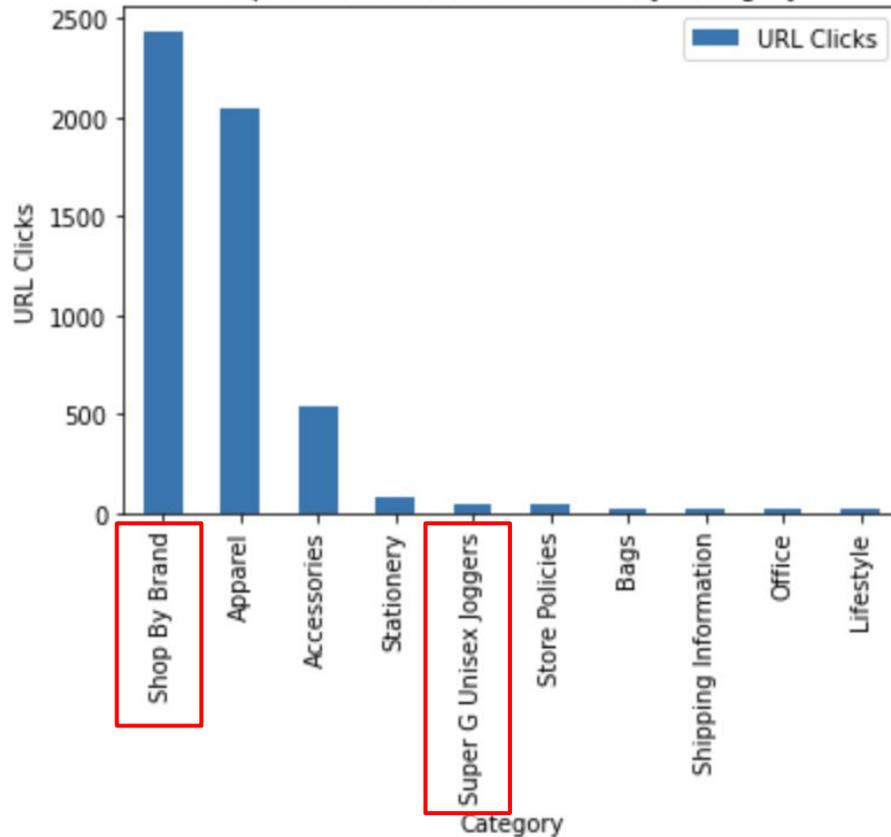
Learning cultures innovate more

Learning cultures thrive under a combo of **psychological safety** & **accountability**

How to develop the habit of thinking again?

1. Think like a scientist

2. Define your identity in terms of values, not opinions

3. Seek out information that goes against your views

How to cultivate a learning culture

Ask questions like:

- What leads you to that assumption?

- What might happen if it's wrong?

- What are the uncertainties in your analysis?

- What are the disadvantages?

- How do you know?

MozCon
Virtual

Guess what the most important driver of Superforecasting success is?

MozCon Virtual

How often the forecasters updated their beliefs

Typical predictors: Updated predictions **2x** per ?

Superforecasters: Updated predictions > **4x** per ?

MozCon
Virtual

What will you rethink?

Books that inspired this talk:

The Black Swan — Nassim Nicholas Taleb

Outliers — Malcolm Gladwell

Originals: How Non-Conformists Move the World Forward — Adam Grant

How to Lie with Statistics — Darrell Huff

A New Earth: Create a Better Life — Eckhart Tolle

Think Again: The Power of Knowing What You Don't Know — Adam Grant

The Rise of Theodore Roosevelt — Edmund Morris
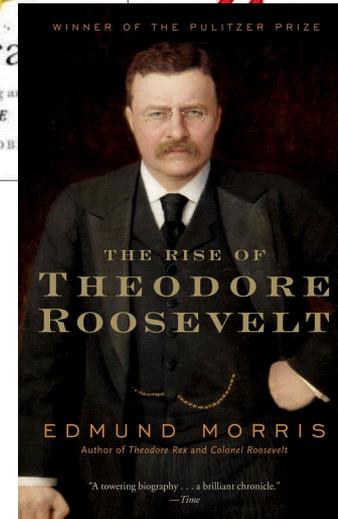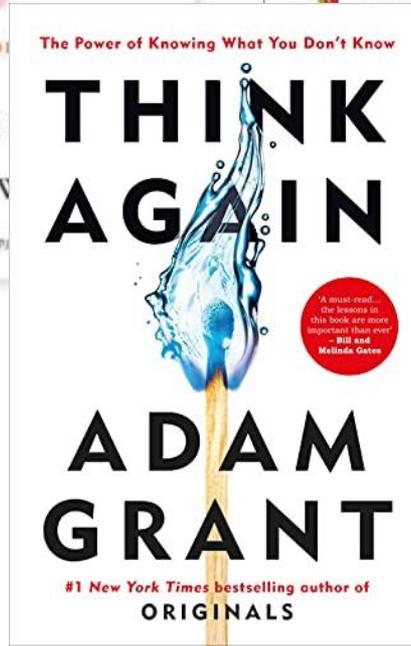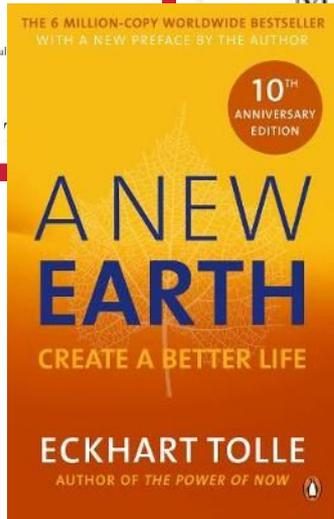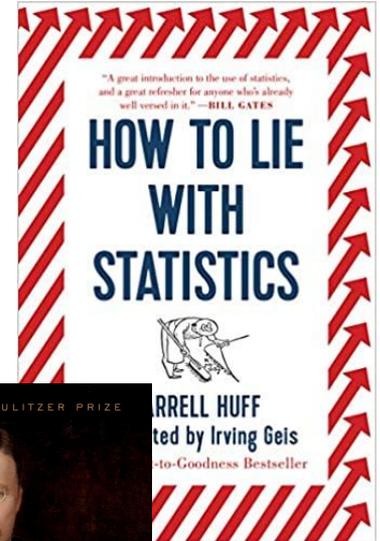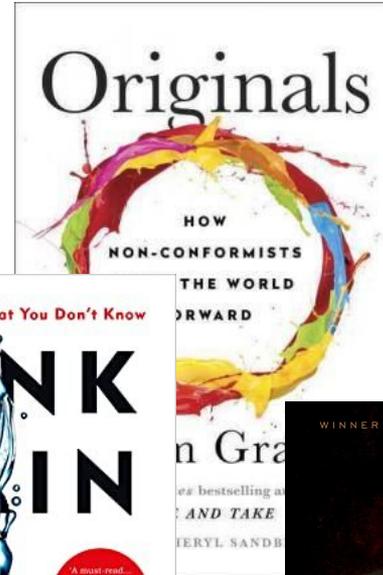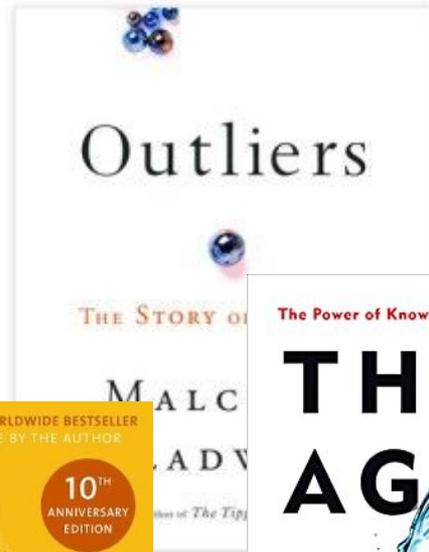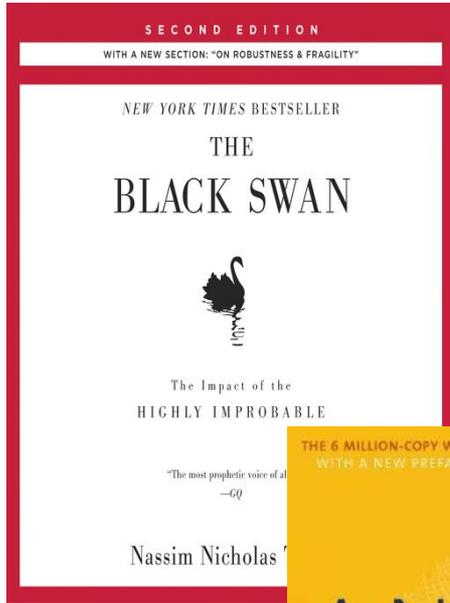
Brilliant people to follow:

@alexisksanders

@rvtheverett
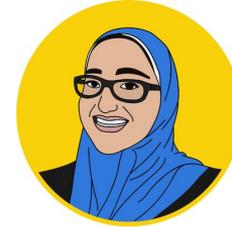
@areej_abuali

@rachellcostello

@kristintynski

@jessthebp

@ariannedonoghue

@nikrangerseo

@hulyacobans

@daisyquaker

@rebberbel

@aleyda

Brilliant people to follow:

@maxxeight

@defaced

@saksters

@fighto

@graysonparks

@datachaz

@suganthanmn

@tylerreardon
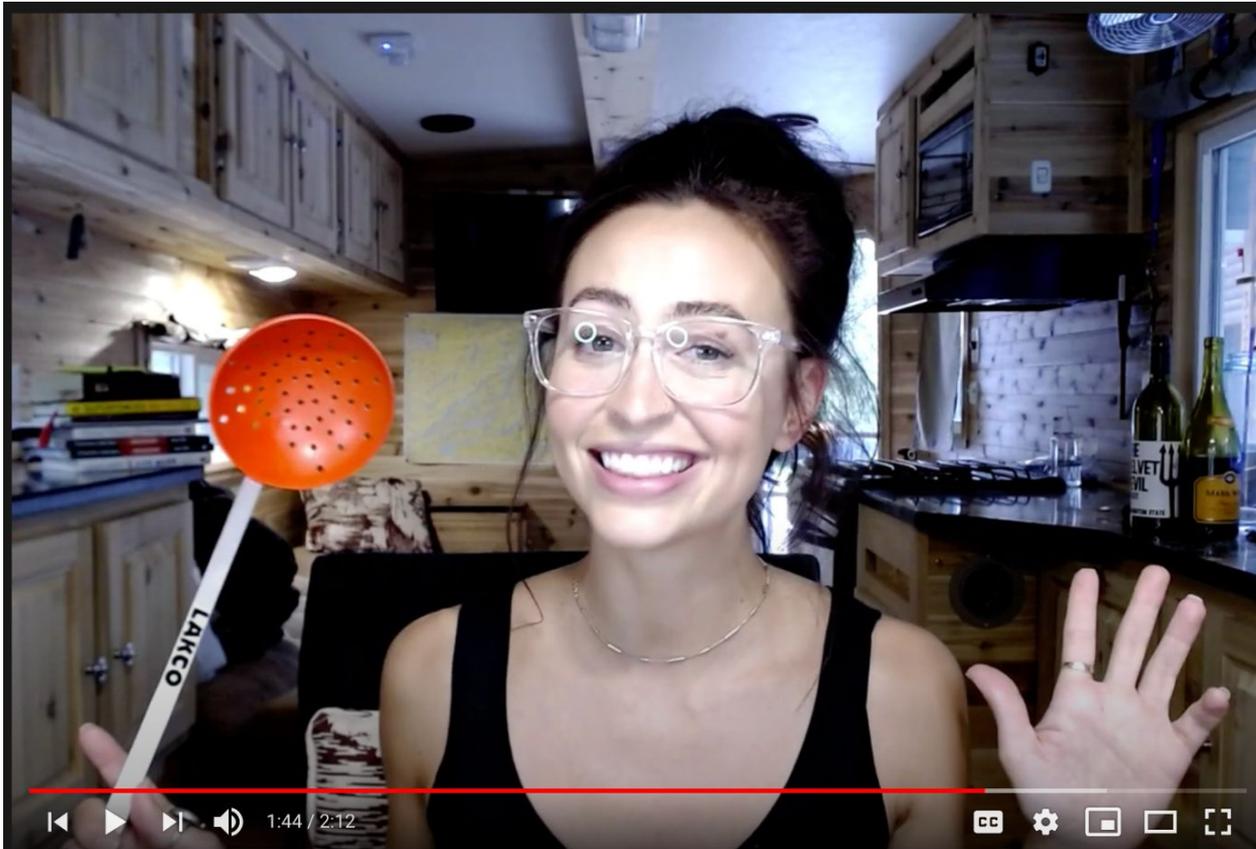
@jroakes

@dsottimano

@kevin_indig
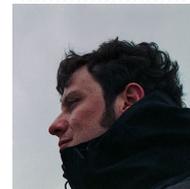
@cyberandy

@edfryed

---Thank you---

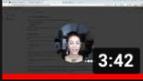**DataSci101.com**
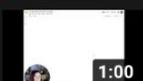@BritneyMuller | SEO Consultant + Data Science Student

bit.ly/colab-yt

Give these two a follow:

Masaki Okazawa
@saksters Follows you

🇯🇵🇻🇳🇺🇸 ▪ SEO ▪ I wear blue a lot

📍 Pittsburgh, PA   🎈 Born September 6   🗓 Joined August 2013

263 Following   **1,040** Followers

Followed by Lee Foot, Greg Bernhardt 🟢🔵, and 98 others you follo

Get SEO results faster
See results in weeks, not months using advanced automation!

Automate in Sheets + Experiment in CDN = Effective Changes

Following

RankSense
@RankSense

RankSense detects and removes traffic-killing SEO issues before it's too late. Speed up SEO from 6 months to 6 weeks using #Agile #SEO 🏃💨

📍 New Jersey, USA   🔗 ranksense.com   🗓 Joined December 2008

171 Following   **1,579** Followers

Followed by Gordon Carslaw, Lee Foot 🟢📈, and 57 others you follow