



# Data Literacy for All: Data Science is for Everyone

Kirk Borne



@KirkDBorne

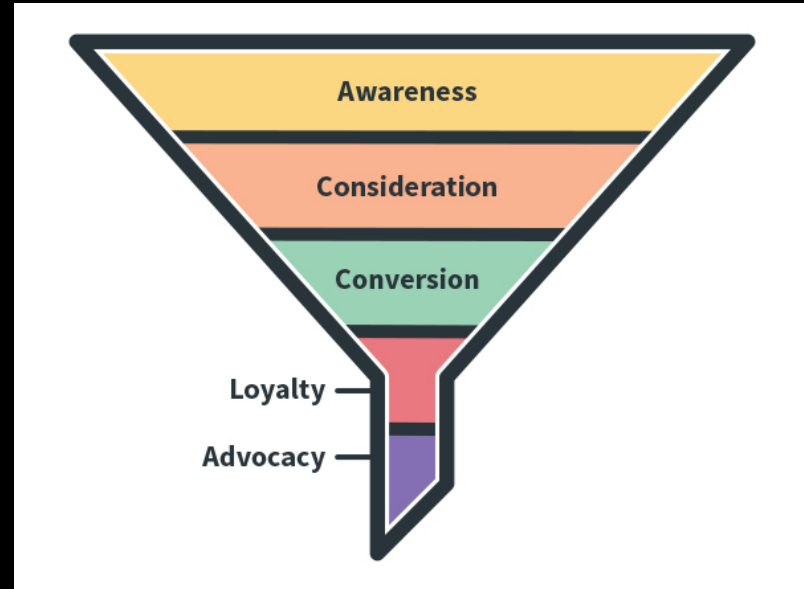


**Principal Data Scientist, Booz Allen Hamilton**

<http://www.boozallen.com/datascience>

# OUTLINE

- Awareness
- Consideration
- Conversion
- Loyalty
- Advocacy

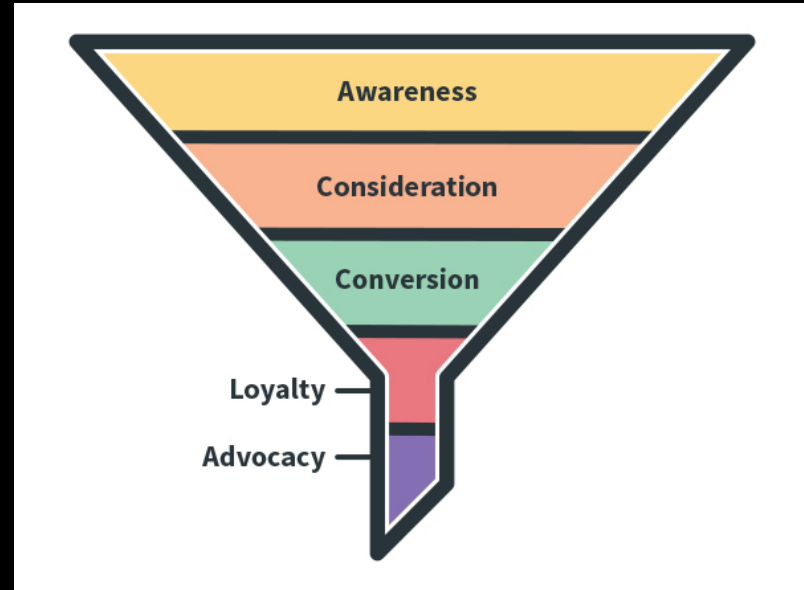


Source for graphic: <https://blog.aweber.com/email-marketing/understanding-the-marketing-funnel-5-strategies-to-improve-your-email-marketing.htm>

# Setting the Stage

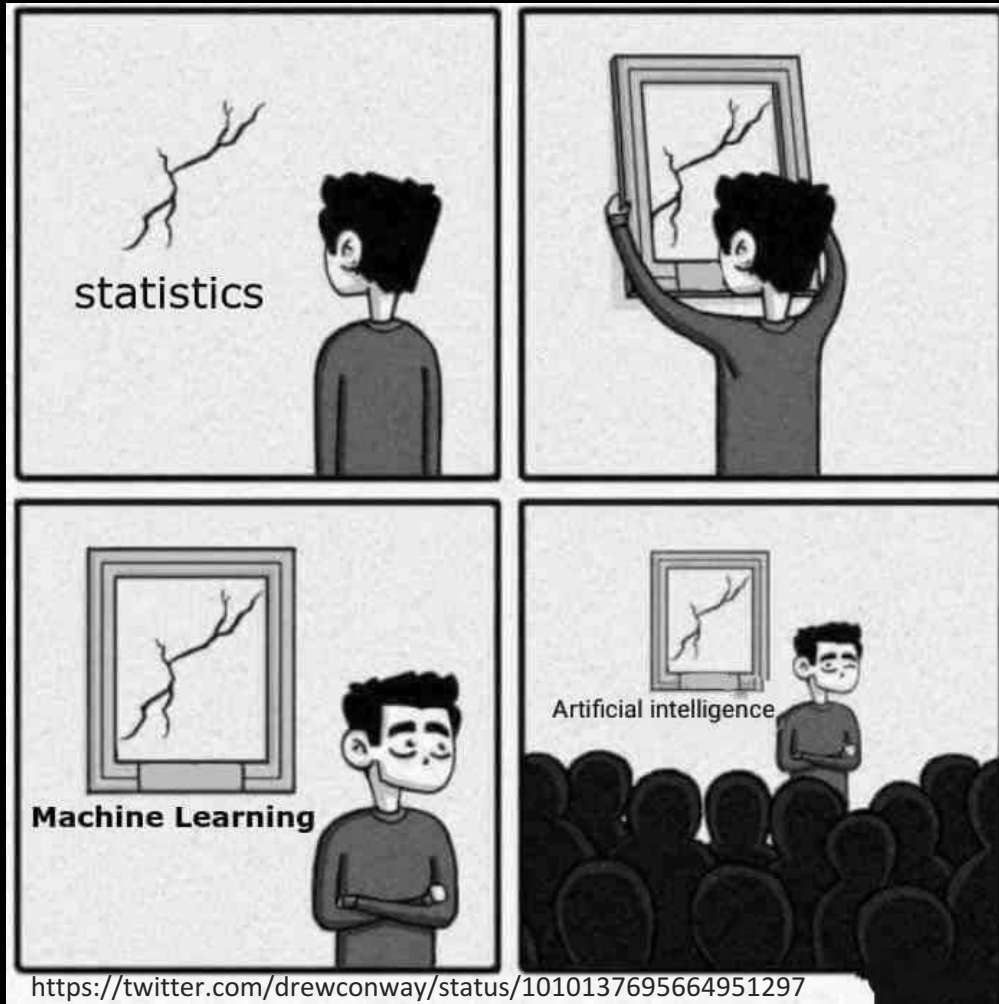


- Awareness
- Consideration
- Conversion
- Loyalty
- Advocacy



Source for graphic: <https://blog.aweber.com/email-marketing/understanding-the-marketing-funnel-5-strategies-to-improve-your-email-marketing.htm>

# How we frame the discussion really matters!



I wrote  
this...  
in 2015...  
=====

# Data Science Declaration

— @KirkDBorne (January 12, 2015)

*“Now is the time to begin thinking of Data Science as a profession not a job, as a corporate culture not a corporate agenda, as a strategy not a stratagem, as a core competency not a course, and as a way of doing things not a thing to do.”*

I wrote  
this...  
in 2015...

=====

Today I would add:  
*Data Competency*  
*is a way of thinking*  
*(Data Literacy),*  
*not a thing to*  
*think about.*

# Data Science Declaration

— @KirkDBorne (January 12, 2015)

*“Now is the time to begin thinking of  
Data Science as a profession not a job,  
as a corporate culture not a corporate  
agenda, as a strategy not a stratagem, as  
a core competency not a course, and as  
a way of doing things not a thing to do.”*

# From a talk that I gave in 2014...

(after 5 years  
teaching Data  
Mining at UMUC  
Graduate School  
and 7 years into our  
Data Science B.S.  
program at GMU)

## Data Science for the Masses: start small, Think Big

*"Big Data" is different!*

Visualize This:

A sea of Data (sea of CDs)



This is the CD Sea in Kilmington, England  
(600,000 CDs ~ 300 TB).

*"Big Data" are different!*

*We need more Data Scientists in order  
to discover the unknown unknowns in  
BIG DATA collections more efficiently  
and more effectively.*

### 1) Big Data in Education

Work with data in all learning settings:

- Data Science enables use and analysis of data in inquiry-based classroom learning.
- Learning is enhanced when students work with real data and information (especially online data) that are related to the topic (any topic) being studied.
- <http://serc.carleton.edu/usingdata/>  
("Using Data in the Classroom")

### 2) An Education in Big Data

Students are specifically trained to:

- access & query big data repositories;
- conduct meaningful inquiries into data;
- mine, visualize, and analyze the data;
- make objective data-driven inferences, discoveries, and decisions; and
- communicate "stories" through data.

### 3) Big Data for Education = **Learning Analytics**

Still very  
relevant  
in 2019...

# Data Science for the Masses: start small, Think Big

*"Big Data" is different!*



## 1) Big Data in Education

Work with data in the classroom settings:

- Data Science for the analysis of data in the classroom learning.
- Learning is enhanced when students work with real data and information (especially online data) that are related to the topic (any topic) being studied.
- <http://serc.carleton.edu/usingdata/> ("Using Data in the Classroom")

## 2) An Education in Big Data

Students are specifically:

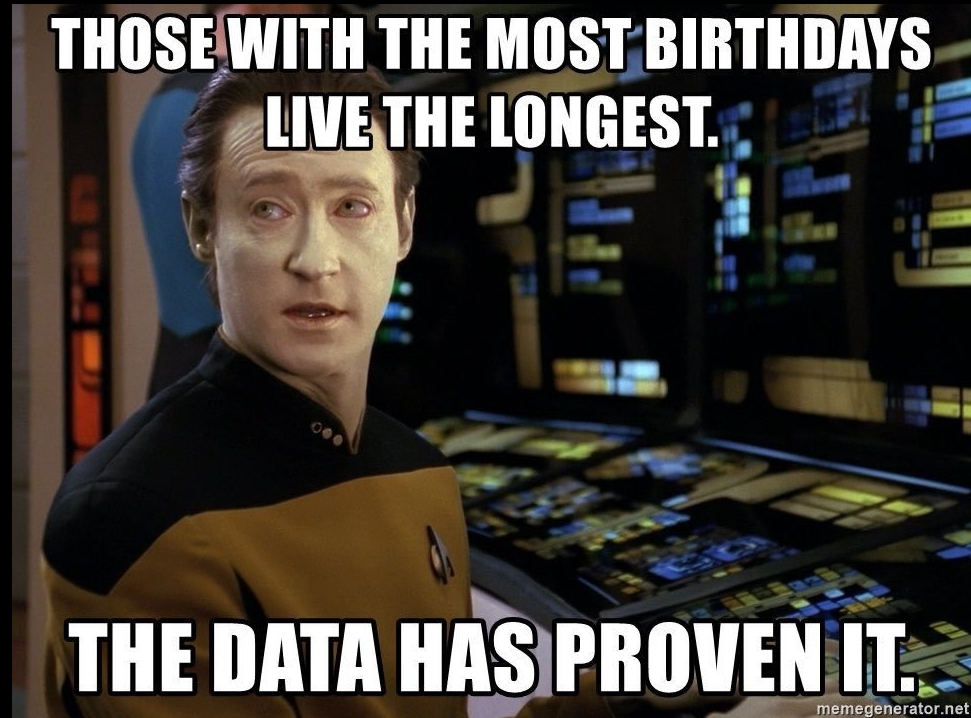
- access & manage data repositories;
- conduct inquiries into data;
- mine, visualize, and analyze the data;
- make objective data-driven inferences, discoveries, and decisions; and
- communicate "stories" through data.

## 3) Big Data for Education

Learning

...for Education

*Data Literacy is a way of thinking about numbers and measurements of things ...*



# Let us kickstart with a Statistics Quiz ...

- Suppose that a survey finds that 10% of people believe that product X is bad for you.
- After a national advertising campaign to inform society of the dangers of product X, another survey is taken.
- The national media report the survey result:
  - *Following a national advertising campaign, the number of people who now believe that product X is bad for you has increased by 90%.*
- Your question: What percentage of people now believe that product X is bad for you?

What answer did you get ?



# The right answer is ... not obvious!

- If the news was unbiased, then the answer is probably 19%.
- If the news was trying to scare you away from Product X, then they may be misleading their listeners into thinking that the correct answer is 100%, though the correct answer is still 19%.
- Maybe the correct answer really is 100% (=10%+90%), but that fact was stated *very poorly* in the news release.
- Or maybe we should **re-read the original statement...**
  - *Following a national advertising campaign, the number of people who now believe that product X is bad for you has increased by 90%.*

# The right answer is ... not obvious!

- If the news was unbiased, then the answer is probably 19%.
- If the news was trying to scare you away from Product X, then they may be misleading their listeners into thinking that the correct answer is 100%, though the correct answer is still 19%.
- Maybe the correct answer really is 100% ( $=10\%+90\%$ ), but that fact was stated very poorly in the news release.
- So maybe only 10 people (out of only 100 that were asked) thought that the product was bad the first time, and then only 19 people (out of many thousands) thought that the product was bad the second time ... therefore, it is true that 90% more people now think that the product is bad, but the percentage of the total population is actually extremely small ... **maybe much less than 1%** (19 out of thousands)!!!!!!!

# The right answer is ... not obvious!

- If the news was unbiased, then the answer is probably 19%.
- If the news was trying to scare you away from X, then they may be misleading the public. The correct answer is still 19%.
- No, the news is not misleading the public. It is true that 19% of the people who were asked thought that the product was bad the first time, and then only 19 people (out of many thousands) thought that the product was bad the second time ... therefore, it is true that 90% more people now think that the product is bad, but the percentage of the total population is actually extremely small ... maybe much less than 1% (19 out of thousands)!!!!!!!

**Statistical and Data  
Literacy Matters!**

## Quote from H.G. Wells (1903; writer) ...

*“Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write.”*

Well, that day is here now!

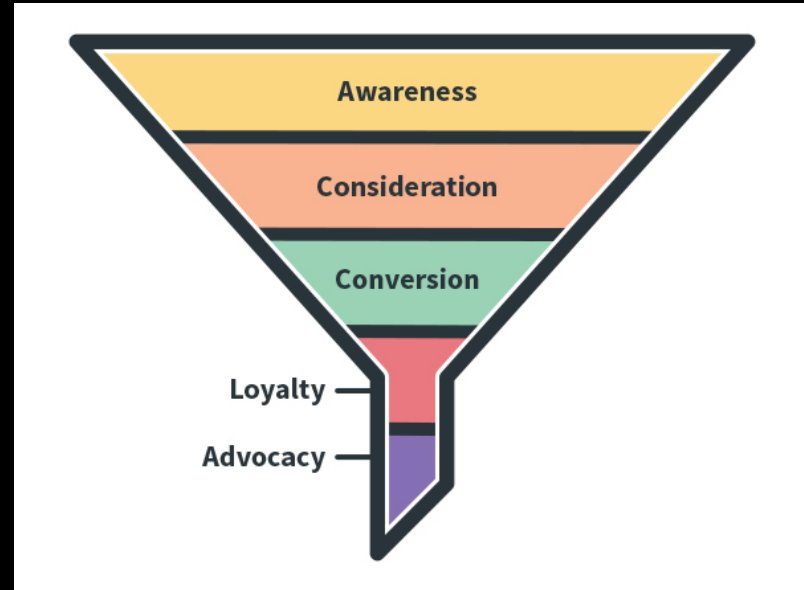
**Statistical & Data Literacy Matters!**

## Quote from somebody (?) ...

*“It is now beyond any doubt that  
cigarettes are the biggest cause  
of statistics”*

# OUTLINE

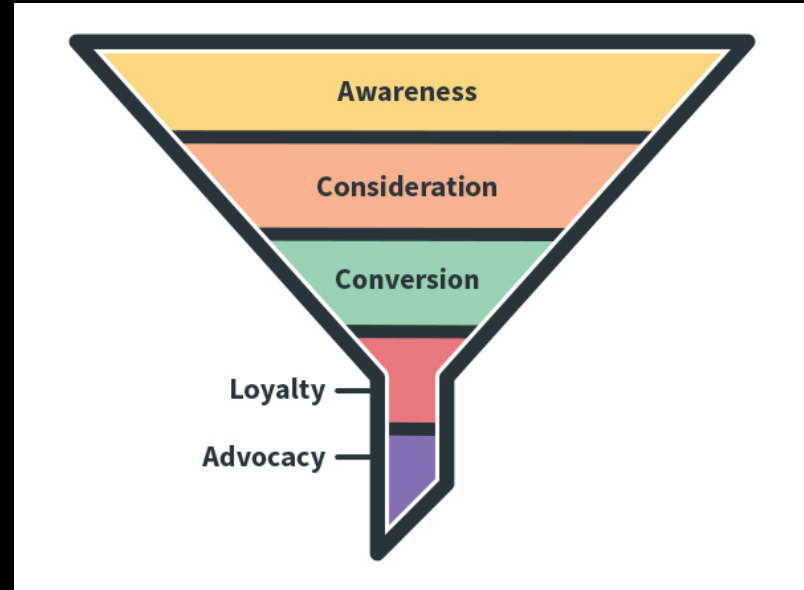
- Awareness
- **Consideration**
- Conversion
- Loyalty
- Advocacy



Source for graphic: <https://blog.aweber.com/email-marketing/understanding-the-marketing-funnel-5-strategies-to-improve-your-email-marketing.htm>

# Steps toward Data Literacy

- Awareness
- **Consideration**
- Conversion
- Loyalty
- Advocacy



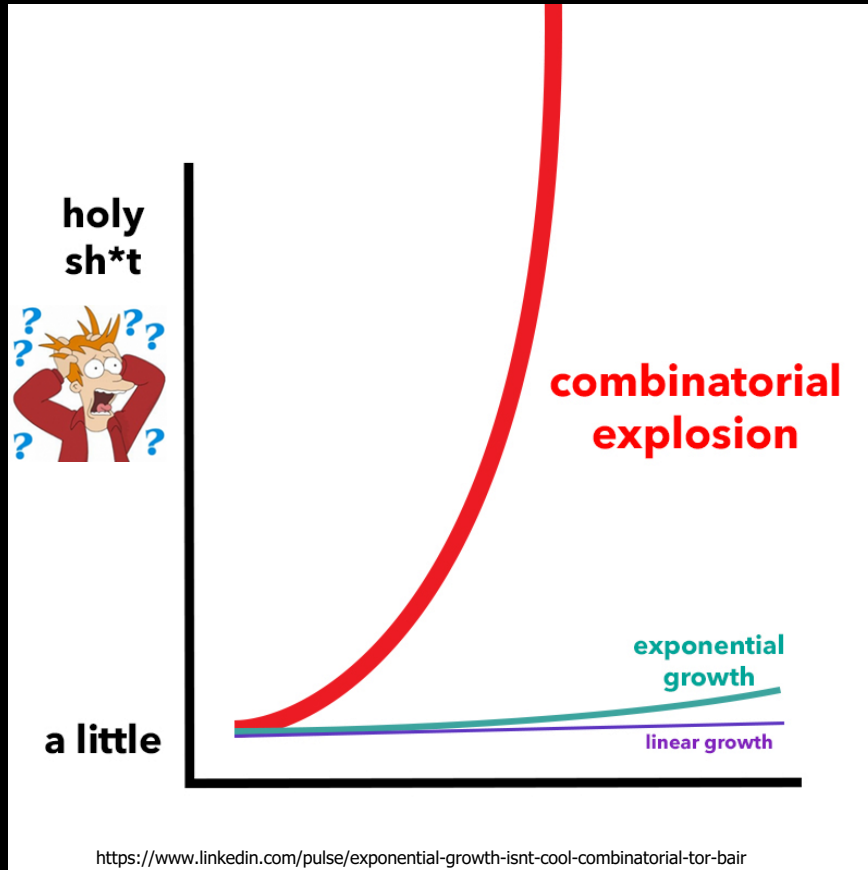
Source for graphic: <https://blog.aweber.com/email-marketing/understanding-the-marketing-funnel-5-strategies-to-improve-your-email-marketing.htm>

# Humans are very curious creatures..

Ever since we first explored our world...  
...we have asked questions about everything around us.



So, we have collected evidence (data) to answer our questions, which leads to more questions, which leads to more data collection, which leads to more questions, ..., which leads to **BIG DATA!**



**Knowledge is about connecting the dots.**

@KirkDBorne

$$y \sim x! \approx x^x$$

→ Combinatorial Growth!  
(all possible interconnections, linkages, and interactions)

$$y \sim 2^x \text{ (exponential growth)}$$

$$y \sim 2 * x \text{ (linear growth)}$$

# We are not talking **only** about this...

## The unicorns of the new data world...

### Which makes them hard to find...

#### Applied Science

- Statistics, applied math
- Machine Learning
- Tools: Python, R, SAS

#### Business Analysis

- Data Analysis, BI
- Business/domain expertise
- Tools: SQL, Excel, EDW

#### Data engineering

- Database technologies
- Computer science
- Tools: Java, Scala, Python, C++

#### Engineering

- Big data pipeline engineering
- Statistics and machine learning over large datasets
- Tools: Hadoop, PIG, HIVE, Cascading, SOLR, etc

figure eight

The Data Scientist Report 2018



<https://www.figure-eight.com/figure-eight-2018-data-scientist-report/>

## MODERN DATA SCIENTIST

Data Scientist, the sexiest job of 21st century requires a mixture of multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business. Finding a data scientist is hard. Finding people who understand who a data scientist is, is equally hard. So here is a little cheat sheet on who the modern data scientist really is.

### MATH & STATISTICS

- ☆ Machine learning
- ☆ Statistical modeling
- ☆ Experiment design
- ☆ Bayesian inference
- ☆ Supervised learning: decision trees, random forests, logistic regression
- ☆ Unsupervised learning: clustering, dimensionality reduction
- ☆ Optimization: gradient descent and variants

### PROGRAMMING & DATABASE

- ☆ Computer science fundamentals
- ☆ Scripting language e.g. Python
- ☆ Statistical computing package e.g. R
- ☆ Databases SQL and NoSQL
- ☆ Relational algebra
- ☆ Parallel databases and parallel query processing
- ☆ MapReduce concepts
- ☆ Hadoop and Hive/Pig
- ☆ Custom reducers
- ☆ Experience with xaaS like AWS

### DOMAIN KNOWLEDGE & SOFT SKILLS

- ☆ Passionate about the business
- ☆ Curious about data
- ☆ Influence without authority
- ☆ Hacker mindset
- ☆ Problem solver
- ☆ Strategic, proactive, creative, innovative and collaborative

### COMMUNICATION & VISUALIZATION

- ☆ Able to engage with senior management
- ☆ Story telling skills
- ☆ Translate data-driven insights into decisions and actions
- ☆ Visual art design
- ☆ R packages like ggplot or lattice
- ☆ Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau

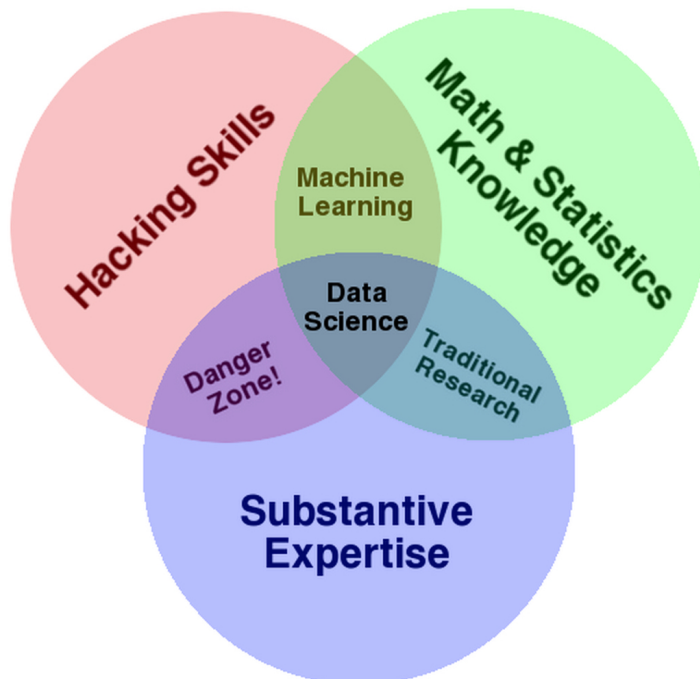
MarketingDistillery.com is a group of practitioners in the area of e-commerce marketing. Our fields of expertise include: marketing strategy and optimization; customer tracking and on-site analytics; predictive analytics and econometrics; data warehousing and big data systems; marketing channel insights in Paid Search, SEO, Social, CRM and brand.

Marketing  
DISTILLERY

<http://www.marketingdistillery.com/2014/11/29/its-data-science-a-buzzword-modern-data-scientist-defined/>

# Data Literacy is most relevant...

...at the intersection of Domain Experts  
and ML-trained Data Scientists!



<http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram>

Let's get real  
about this...

... says K.Borne



# Data Literacy includes ...

*“Data Literacy includes the ability to read, work with, analyze, and argue with data.”*

(Jordan Morrow, Qlik)

<http://www.dataliteracynetwork.org/definitions.html>

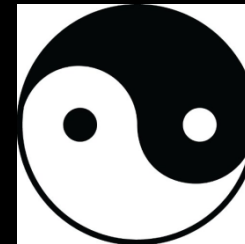


Source: <http://bit.ly/2mEzJsr>

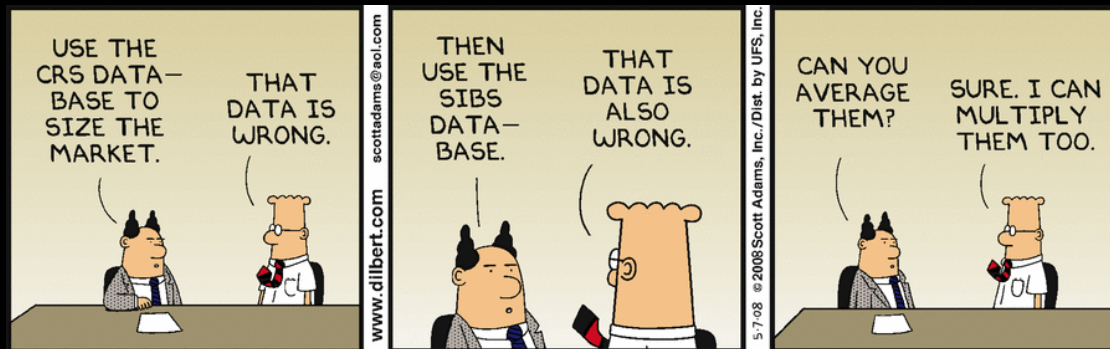


# Data Literacy in 2 parts: Data Science and Data Ethics

<http://www.kirkborne.net/cds151/>

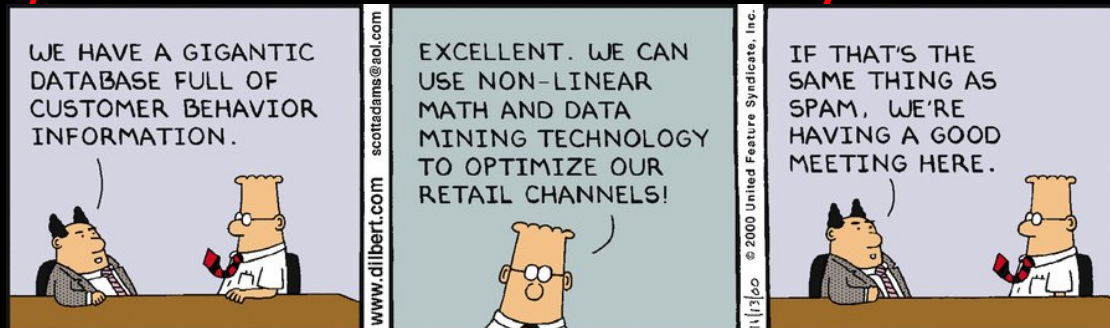


## 1) How to use data correctly



<http://dilbert.com/strips/2008-05-07>

## 2) How to use data ethically



<http://dilbert.com/strips/2000-11-13>

**Within any pile of big data, you can always find correlations and patterns – but are these random, or confessing some truth, or confirming some bias?**

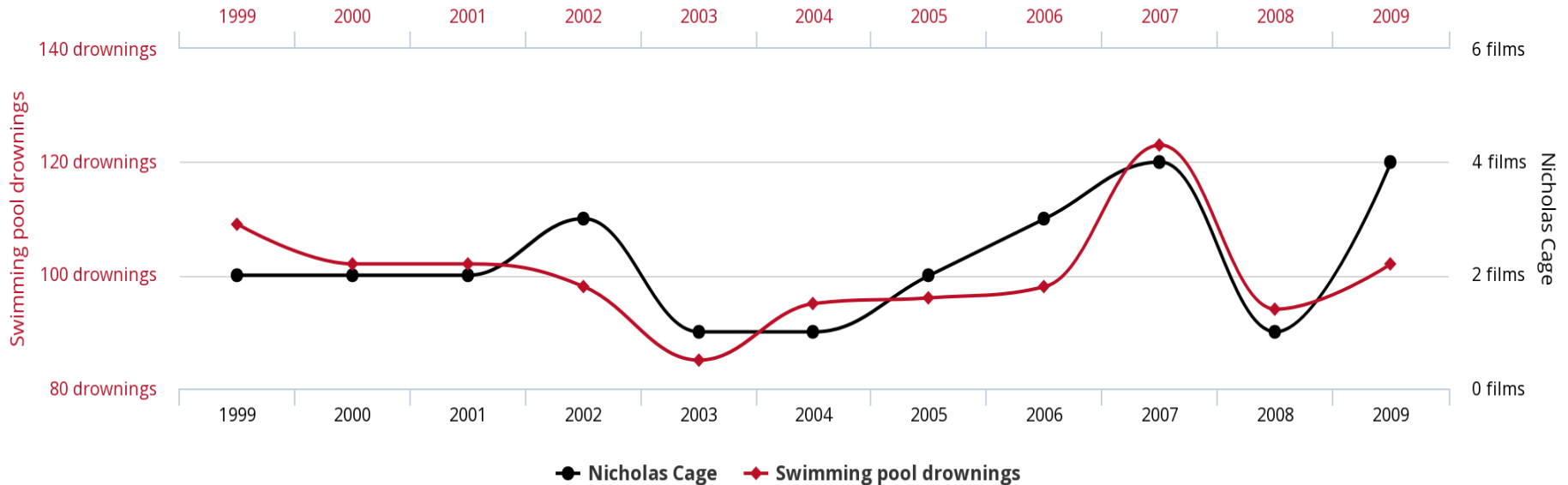
**“If you torture the data long enough,  
it will confess to anything.”**

**– Ronald Coase,  
Nobel Prize winning economist**

# Statistical Fallacies can still appear in Data Science in the era of Big Data

**Correlation  $\neq$  Causation**

**Number of people who drowned by falling into a pool**  
correlates with  
**Films Nicolas Cage appeared in**



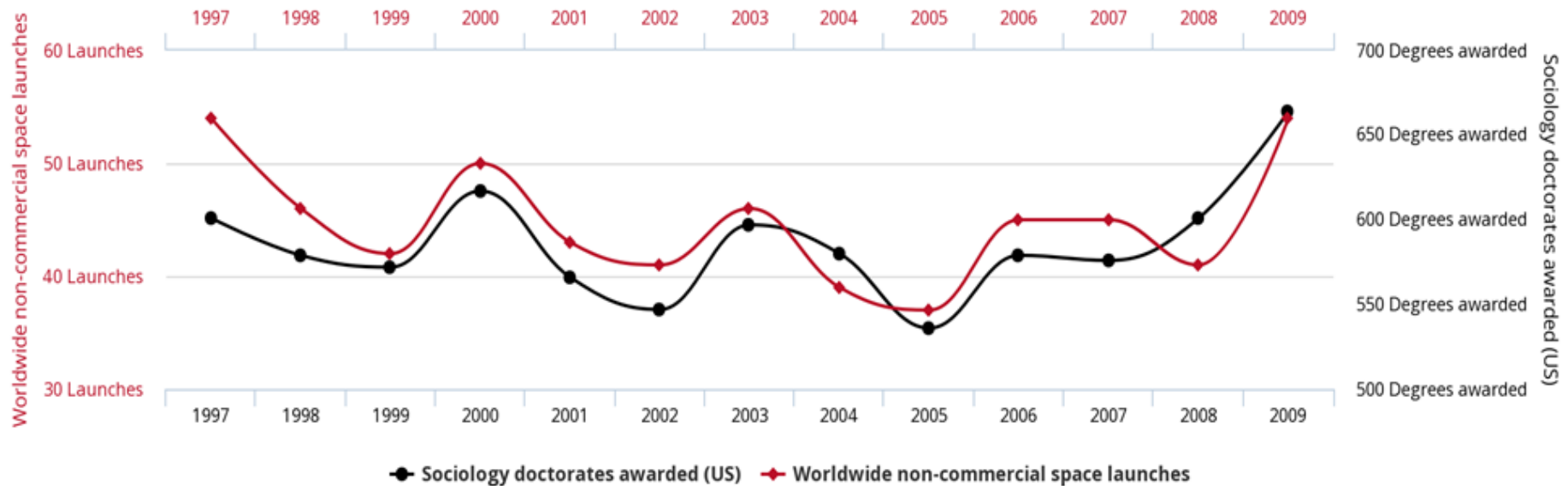
<http://tylervigen.com/spurious-correlations>

tylervigen.com

# Statistical Fallacies can still appear in Data Science in the era of Big Data

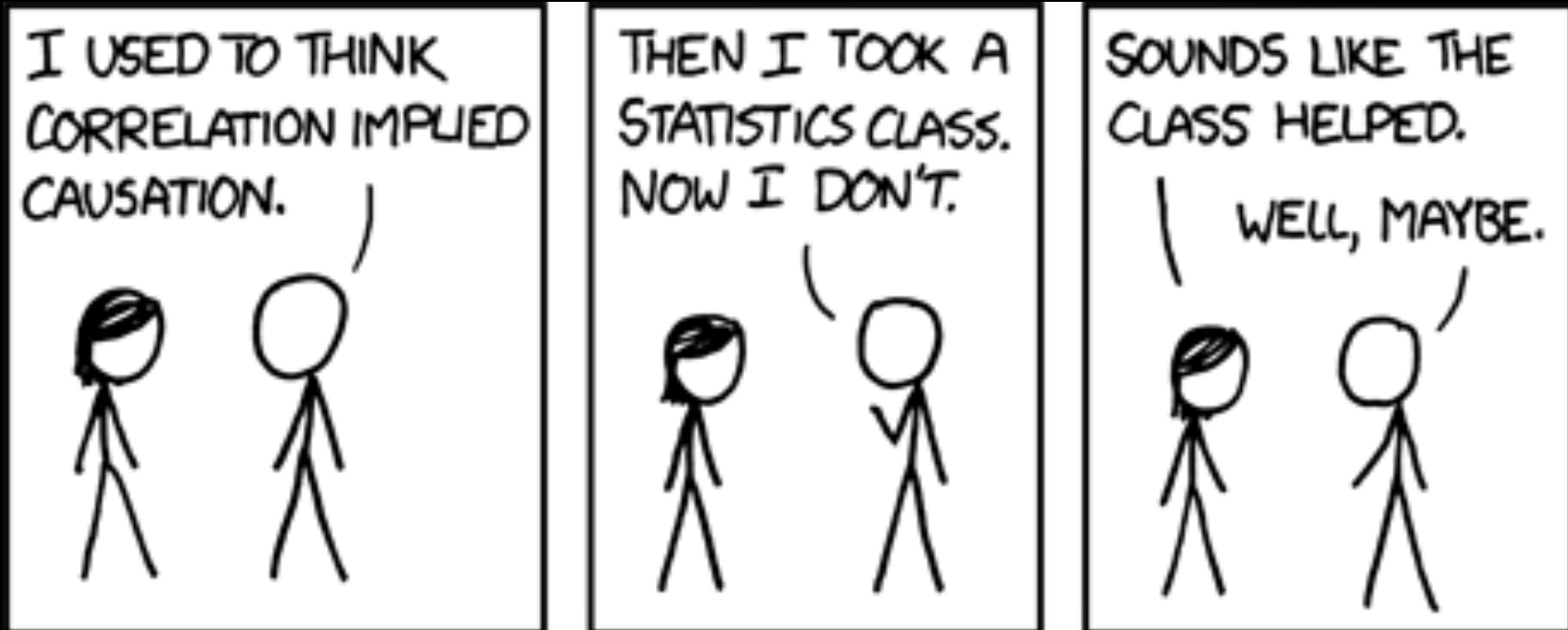
**Correlation  $\neq$  Causation**

**Worldwide non-commercial space launches**  
correlates with  
**Sociology doctorates awarded (US)**



# Statistical Fallacies can still appear in Data Science in the era of Big Data

**Correlation  $\neq$  Causation**



<https://xkcd.com/552/>

## DATA FALLACIES TO AVOID



### CHERRY PICKING

Selecting results that fit your claim and excluding those that don't.



### DATA DREDGING

Repeatedly testing new hypotheses against the same set of data, failing to acknowledge that most correlations will be the result of chance.



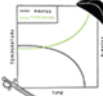
### SURVIVORSHIP BIAS

Drawing conclusions from an incomplete set of data, because that data has 'survived' some selection criteria.



### COBRA EFFECT

Setting an incentive that accidentally produces the opposite result to the one intended. Also known as a Perverse Incentive.



### FALSE CAUSALITY

Falsely assuming when two events appear related that one must have caused the other.



### GERRYMANDERING

Manipulating the geographical boundaries used to group data in order to change the result.



### SAMPLING BIAS

Drawing conclusions from a set of data that isn't representative of the population you're trying to understand.



### GAMBLER'S FALLACY

Mistakenly believing that because something has happened more frequently than usual, it's now less likely to happen in future (and vice versa).



### HAWTHORNE EFFECT

The act of monitoring someone can affect their behaviour, leading to spurious findings. Also known as the Observer Effect.



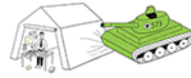
### REGRESSION FALLACY

When something happens that's unusually good or bad, it will revert back towards the average over time.

REPLICATION SUCCESS RATE			
GROUP	PHASE	PHASE	PHASE
GROUP A	100%	100%	100%
GROUP B	100%	100%	100%
GROUP C	100%	100%	100%
GROUP D	100%	100%	100%
GROUP E	100%	100%	100%
GROUP F	100%	100%	100%
GROUP G	100%	100%	100%
GROUP H	100%	100%	100%
GROUP I	100%	100%	100%
GROUP J	100%	100%	100%

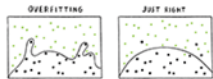
### SIMPSON'S PARADOX

When a trend appears in different subsets of data but disappears or reverses when the groups are combined.



### MCMANARA FALLACY

Relying solely on metrics in complex situations and losing sight of the bigger picture.



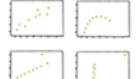
### OVERFITTING

Creating a model that's overly tailored to the data you have and not representative of the general trend.



### PUBLICATION BIAS

Interesting research findings are more likely to be published, distorting our impression of reality.



### DANGER OF SUMMARY METRICS

Only looking at summary metrics and missing big differences in the raw data.

In our rush to build and to promote our models, we are often too quick to overlook our own cognitive biases and other data fallacies:

<https://bit.ly/2pPnUSu>

# Feature Selection and Model Bias: choosing features in the dark

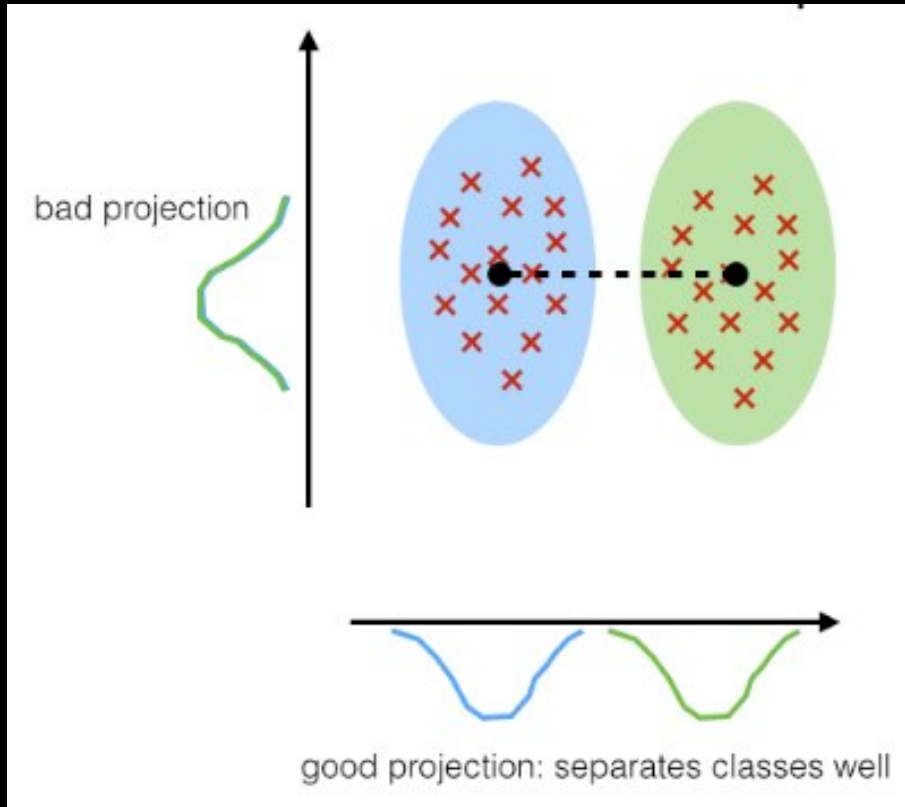


I picked out two socks from my sock drawer this morning!

It was still dark, but that shouldn't matter, right? After all, they are the **same size** ...  
**THE SAME ?!?**

The Era of Big Data represents the **END OF DEMOGRAPHICS** (*i.e.*, our models should no longer be based on and biased by a limited selection of attributes and features)

# Feature Selection and Projection



Source: <https://www.quora.com/How-was-classification-as-a-learning-machine-developed>

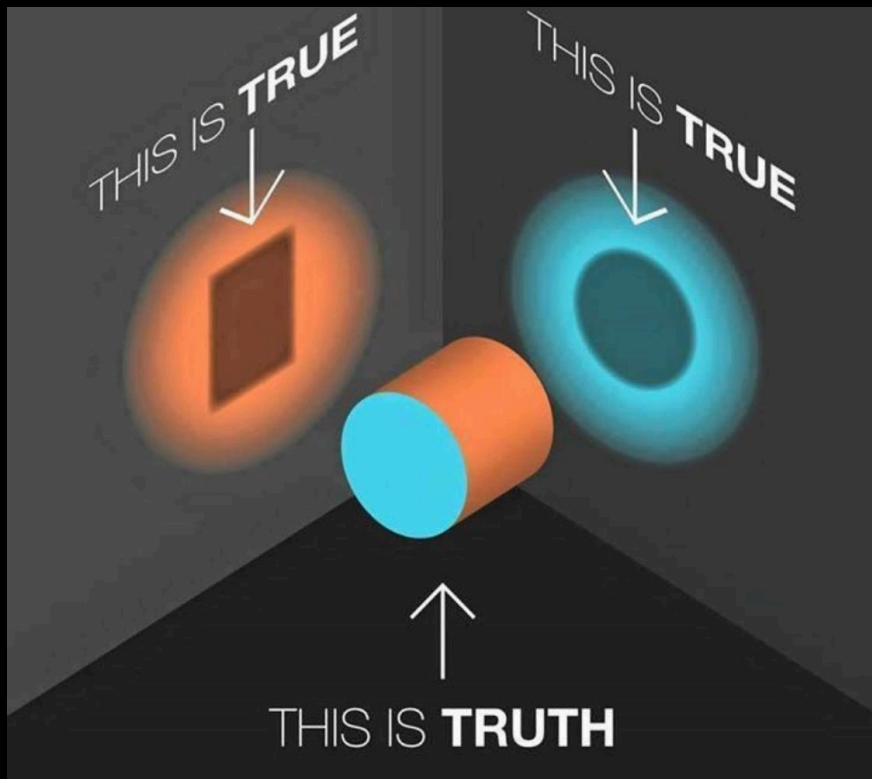
Feature Selection is important in order to disambiguate different classes.

More importantly, **Class Discovery** depends on choosing the right projection and selecting the right features!

# High-Variety Data can be a Bias-Buster.

## Projection Matters!

→ <https://bit.ly/2CGHZjN> ←



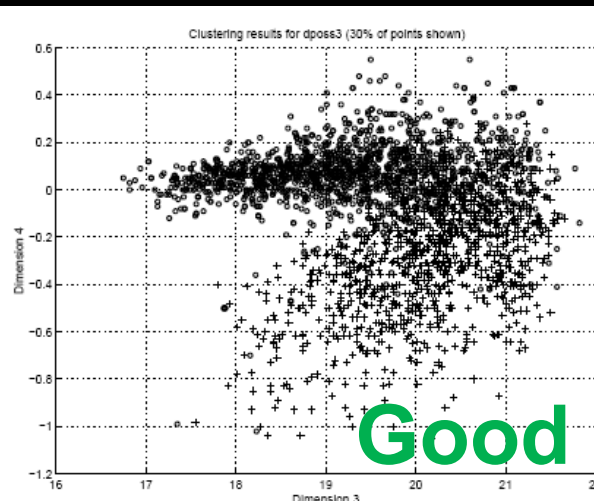
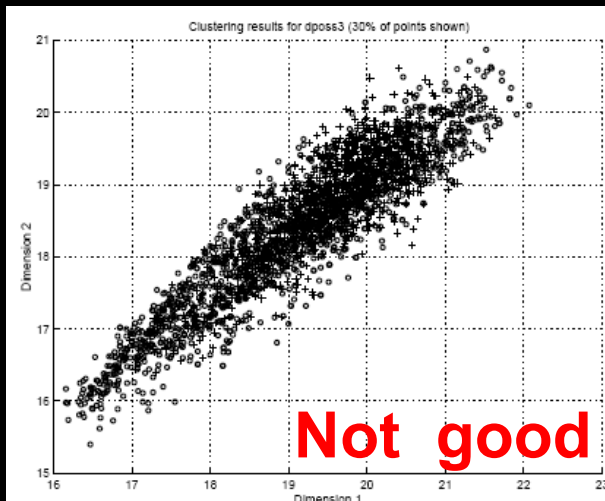
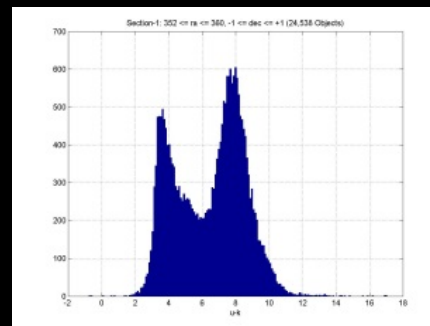
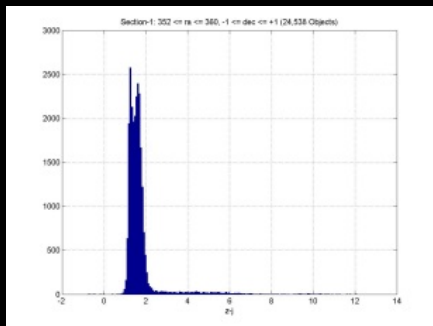
Source: <http://www.transformativeinsights.co.nz/blog/new-perspective-on-conflict>

Your chosen data attributes represent a low-dimension projection of the full truth – the feature space (dimensions) in which you explore your data is a form of cognitive bias – ... **it matters!**

# The 5 important **D**'s of Data **Variety**:

Entity **D**isambiguation, Entity **D**eduplication, **D**iscrimination between multiple classes, **D**iscovery of new classes, and **D**ecreased model bias (underfitting).

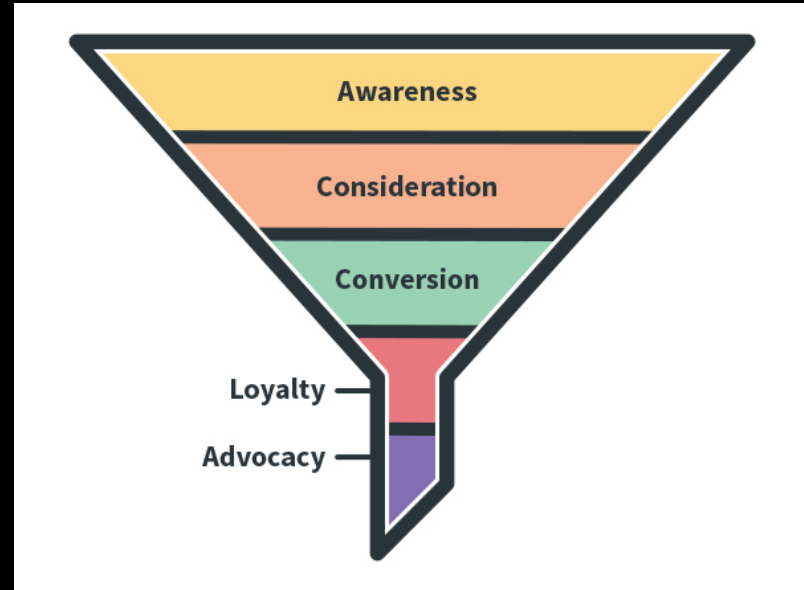
The separation and discovery of multiple classes improves when a sufficient number of “correct” features are available for exploration and testing.



Reference: <http://www.cs.princeton.edu/courses/archive/spr04/cos598B/bib/BrunnerDPS.pdf>

# OUTLINE

- Awareness
- Consideration
- **Conversion**
- Loyalty
- Advocacy

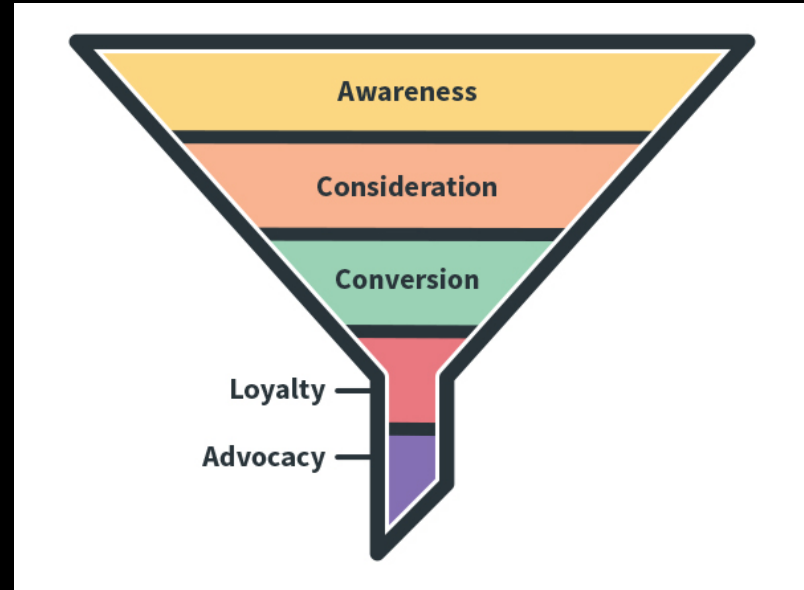


Source for graphic: <https://blog.aweber.com/email-marketing/understanding-the-marketing-funnel-5-strategies-to-improve-your-email-marketing.htm>



# Using Data Science to Create Value

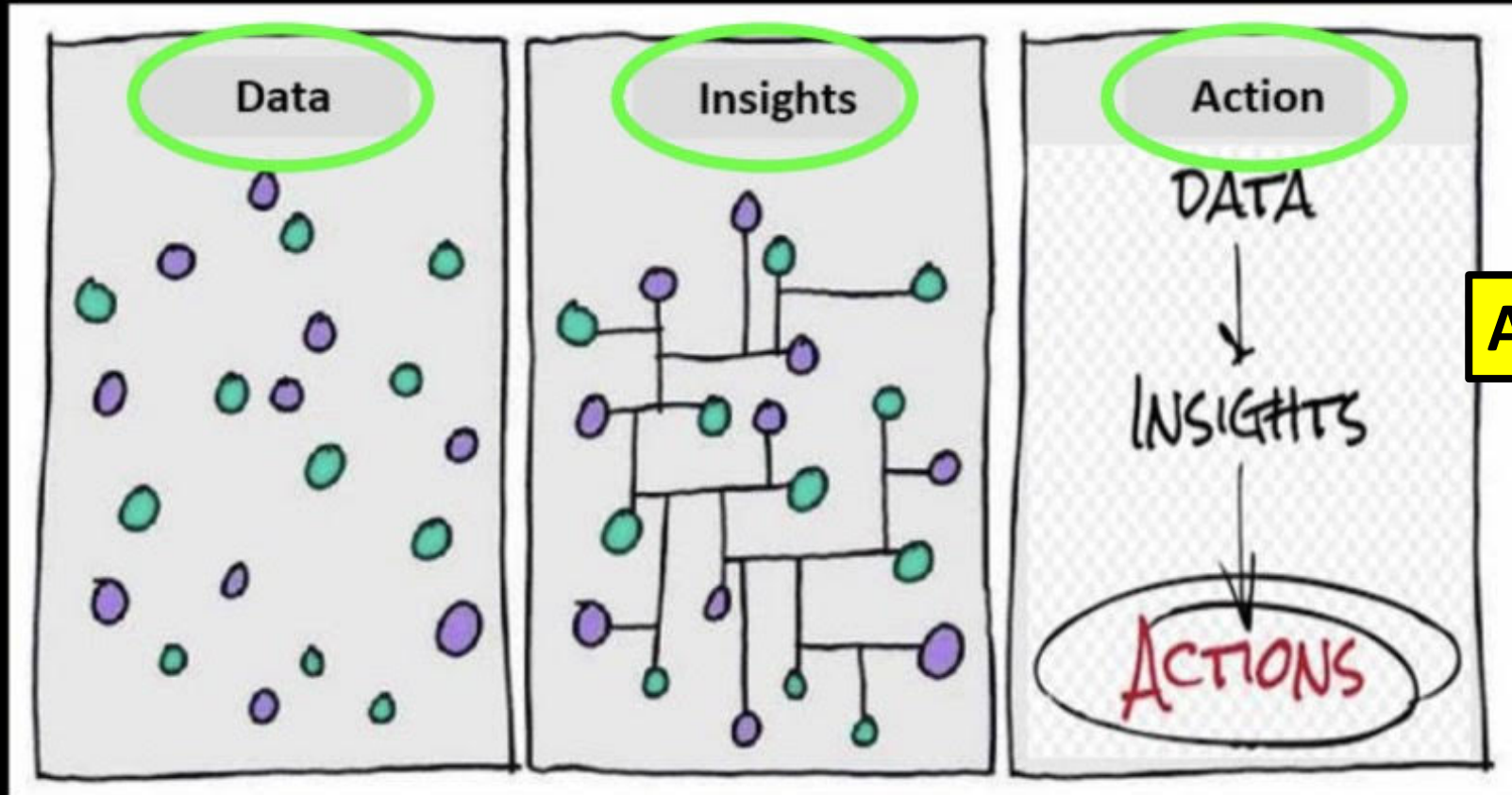
- Awareness
- Consideration
- **Conversion**
- Loyalty
- Advocacy



Source for graphic: <https://blog.aweber.com/email-marketing/understanding-the-marketing-funnel-5-strategies-to-improve-your-email-marketing.htm>

# Data are the fuel for insights that drive action...

## From Data to Insights to Actionable Intelligence



**Action!**

# The Real Power of A.I. – there is nothing “artificial” about it!

The New AI is better than Artificial Intelligence

**Accelerated**

**Applied**

**Actionable**

**Assisted**

**Intelligence**

**Adaptable**

**Augmented**

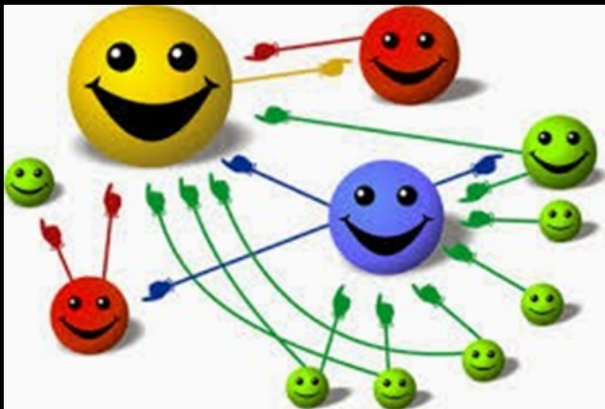
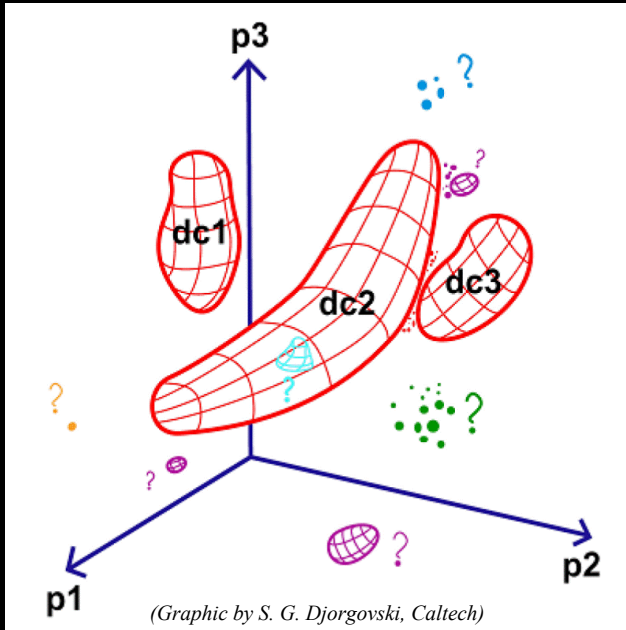
**Amplified**

**Awesome**

<https://datamakespossible.westerndigital.com/real-power-ai/>

## 4 Types of Discovery from Data:

- 1) **Class Discovery:** Find the categories of objects (population segments), events, and behaviors in your data. + Learn the rules that constrain the class boundaries (that uniquely distinguish them).
- 2) **Correlation (Predictive and Prescriptive Power) Discovery: (INSIGHT DISCOVERY)** – Find trends, patterns, and dependencies in data that reveal the governing principles or behavioral patterns (the object's "DNA").
- 3) **Outlier / Anomaly / Novelty / Surprise Discovery:** Find the new, surprising, unexpected one-in-a-[million / billion / trillion] object, event, or behavior.
- 4) **Association (or Link) Discovery:** (Graph and Network Analytics) – Find both the usual and the unusual (interesting) data associations / links / connections across the entities in your domain.



# Levels of Analytics Maturity in Data-Driven Applications

## 1) Descriptive Analytics

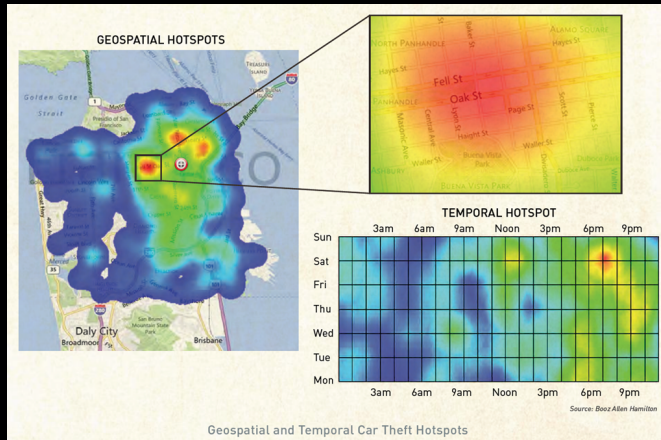
- **Hindsight** (What happened?)

## 2) Diagnostic Analytics

- **Oversight** (real-time / What is happening? Why did it happen?)

## 3) Predictive Analytics

- **Foresight** (What will happen?)



# 5 Levels of Analytics Maturity in Data-Driven Applications

## 1) Descriptive Analytics

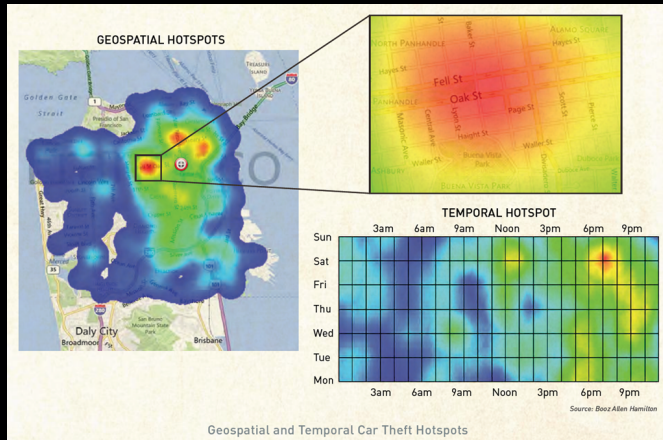
- **Hindsight** (What happened?)

## 2) Diagnostic Analytics

- **Oversight** (real-time / What is happening? Why did it happen?)

## 3) Predictive Analytics

- **Foresight** (What will happen?)



## 4) Prescriptive Analytics

- **Insight** (How can we optimize what happens?) (Follow the dots / connections in the graph!) **Insights Discovery**

## 5) Cognitive Analytics

- **Right Sight** (the 360 view , **what is the right question to ask for this set of data in this context** = Game of Jeopardy)
- Finds the right insight, the right action, the right decision,... right now!
- Moves beyond simply providing answers, to **generating new questions and hypotheses.**



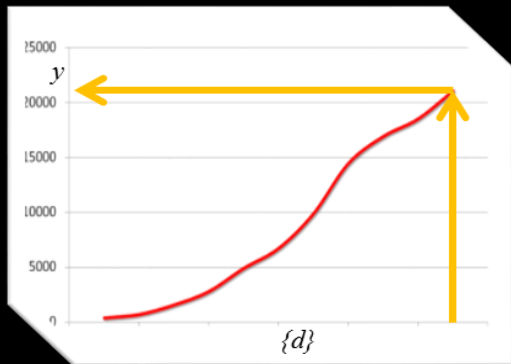
# Predictive vs Prescriptive: What's the Difference?

## PREDICTIVE

### *Analytics*

Find a function (i.e., the model)  $f(d,t)$  that predicts the value of some predictive variable  $y = f(d,t)$  at a future time  $t$ , given the set of conditions found in the training data  $\{d\}$ .

=> Given  $\{d\}$ , find  $y$ .

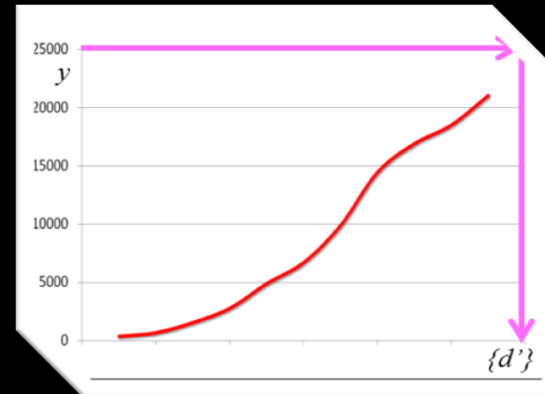


## PRESCRIPTIVE

### *Analytics*

Find the conditions  $\{d'\}$  that will produce a prescribed (desired, optimum) value  $y$  at a future time  $t$ , using the previously learned conditional dependencies among the variables in the predictive function  $f(d,t)$ .

=> Given  $y$ , find  $\{d'\}$ .



# Predictive vs Prescriptive: What's the Difference?

## PREDICTIVE

### *Analytics*

Find a function (i.e., the model)  $f(d,t)$  that predicts the value of some predictive variable  $y = f(d,t)$  at a future time  $t$ , given the set of conditions found in the training data  $\{d\}$ .

=> Given  $\{d\}$ , find  $y$ .

Confucius says...

“Study your past to know  
your future”

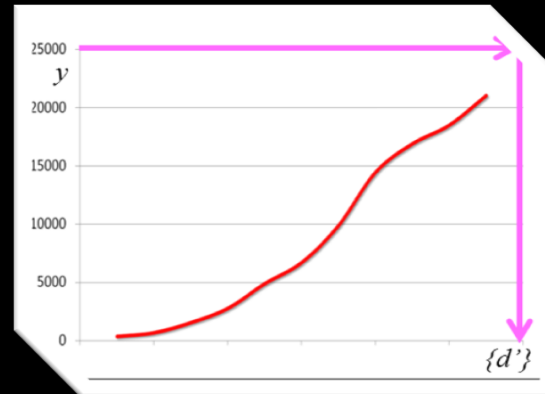
$\{d\}$

## PRESCRIPTIVE

### *Analytics*

Find the conditions  $\{d'\}$  that will produce a prescribed (desired, optimum) value  $y$  at a future time  $t$ , using the previously learned conditional dependencies among the variables in the predictive function  $f(d,t)$ .

=> Given  $y$ , find  $\{d'\}$ .



# Predictive vs Prescriptive: What's the Difference?

## PREDICTIVE

### *Analytics*

Find a function (i.e., the model)  $f(d,t)$  that predicts the value of some predictive variable  $y = f(d,t)$  at a future time  $t$ , given the set of conditions found in the training data  $\{d\}$ .

=> Given  $\{d\}$ , find  $y$ .

Confucius says...

“Study your past to know your future”

$\{d\}$

## PRESCRIPTIVE

### *Analytics*

Find the conditions  $\{d'\}$  that will produce a prescribed (desired, optimum) value  $y$  at a future time  $t$ , using the previously learned conditional dependencies among the variables in the predictive function  $f(d,t)$ .

=> Given  $y$ , find  $\{d'\}$ .

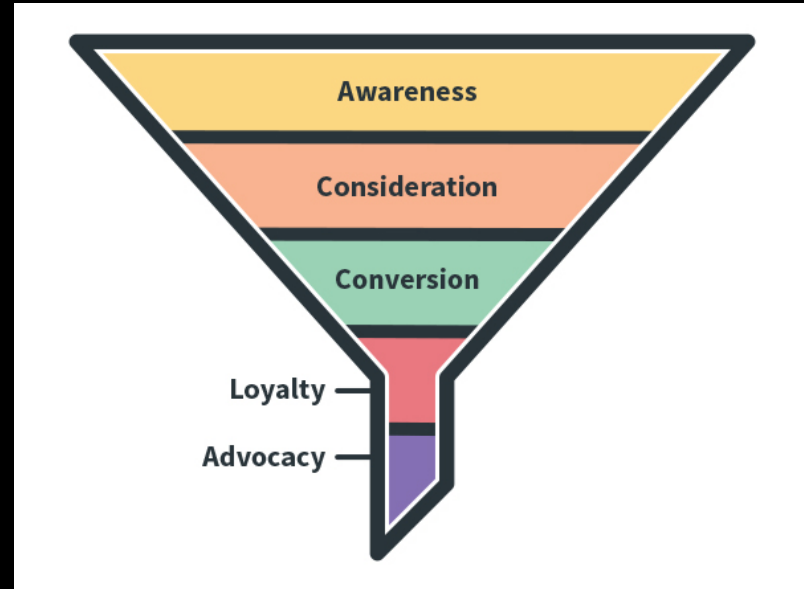
Baseball philosopher Yogi Berra says...

“The future ain't what it used to be.”

$\{d'\}$

# OUTLINE

- Awareness
- Consideration
- Conversion
- **Loyalty**
- Advocacy

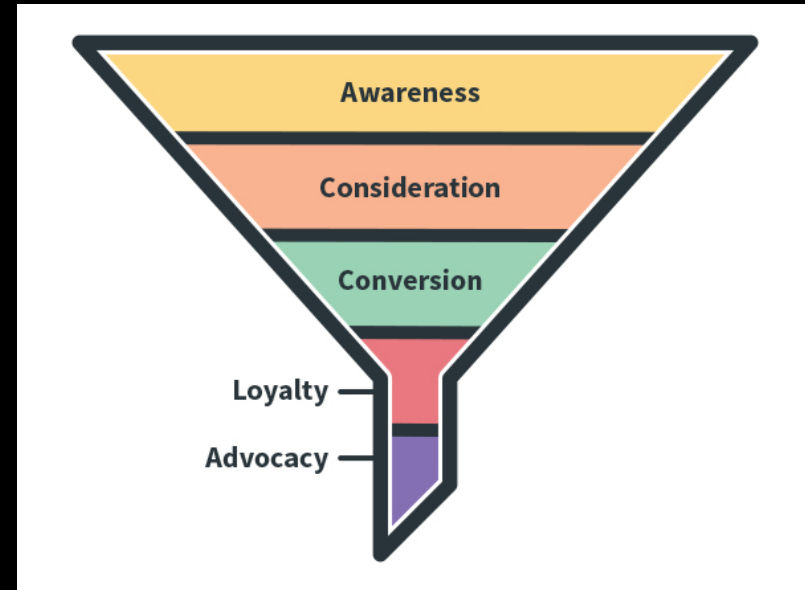


Source for graphic: <https://blog.aweber.com/email-marketing/understanding-the-marketing-funnel-5-strategies-to-improve-your-email-marketing.htm>

# Data Storytelling :

## tapping into the neuroscience of decision-making

- Awareness
- Consideration
- Conversion
- **Loyalty**
- Advocacy

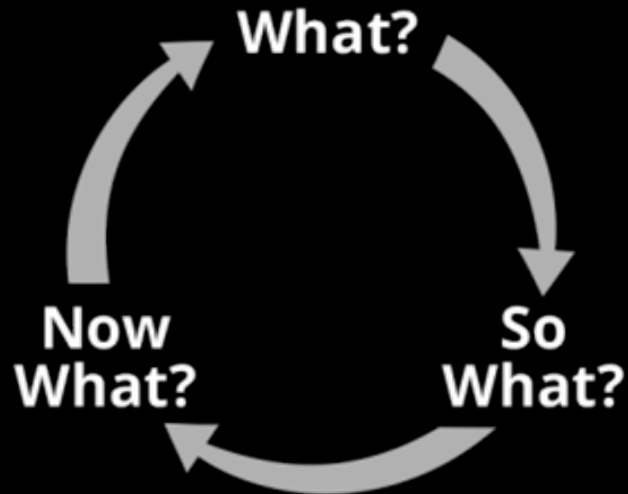


Source for graphic: <https://blog.aweber.com/email-marketing/understanding-the-marketing-funnel-5-strategies-to-improve-your-email-marketing.htm>

# Data Storytelling

"No one ever made a decision because of a number. They need a story."

*Daniel Kahneman, Quoted in Vanity Fair article  
"How Two Trailblazing Psychologists Turned the World of Decision Science Upside Down,"  
November 2016*



"People will forget what you said,  
people will forget what you did,  
but people will never forget  
how you made them feel"

Maya Angelou



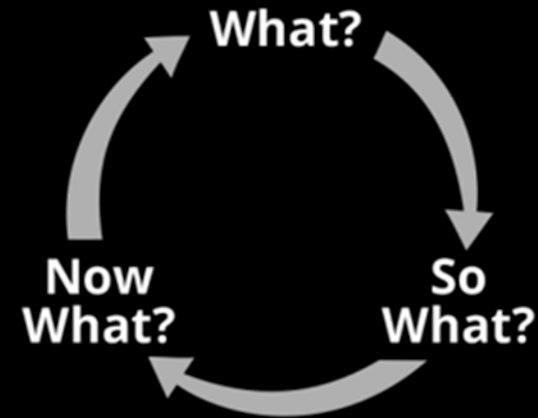
“Data, I think, is one of **the most powerful mechanisms for telling stories**. I take a huge pile of data and I try to get it to tell stories.

Steven Levitt  
Co-author of Freakonomics



# 3 Data Short Stories: the art & science of being data literate!

- 1) Personal
- 2) Civic
- 3) Commercial

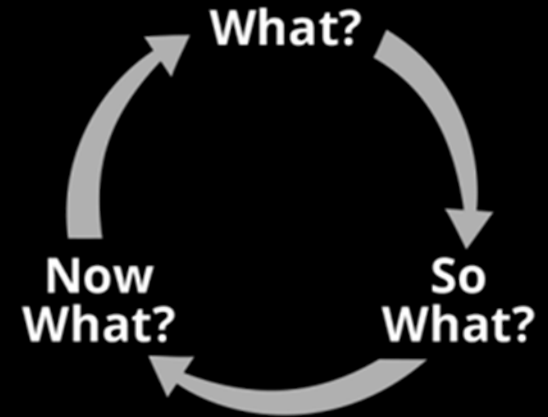


# 3 Data Short Stories: the art & science of being data literate!

**1) Personal**

2) Civic

3) Commercial



# We are ALL data generators!

## Shouldn't we all be value generators?

The most important “V”  
of Big Data is Value!



<https://knowyourmeme.com/photos/1119756-the-internet>

**It's a Digital World after all...**  
**Data helps our world go around!**



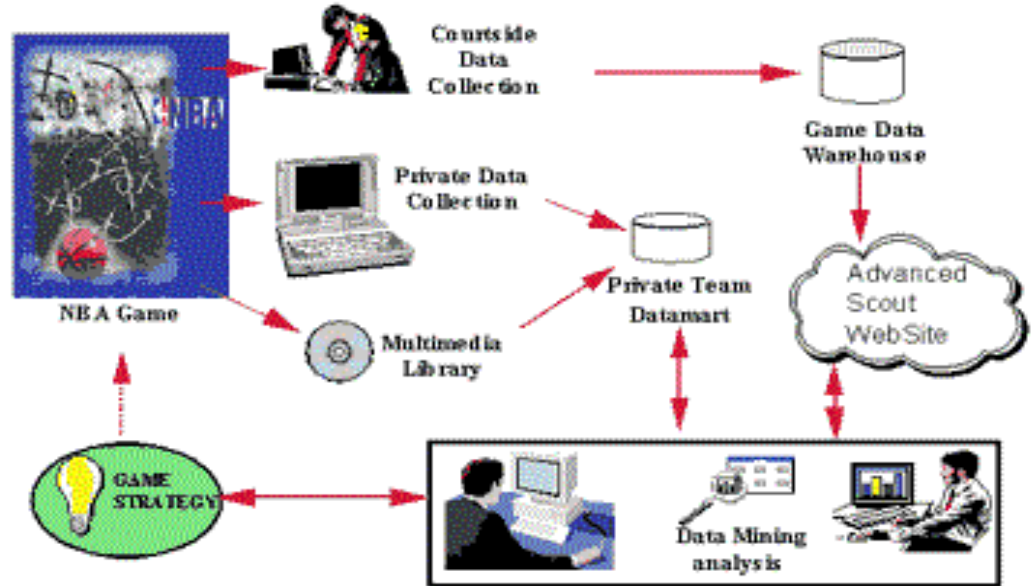
# Engage students at the intersection of data and their place of passion and personal interests!

Me in 1998 ...

Data Mining what?



## IBM Advanced Scout for NBA Coaches Information Flow



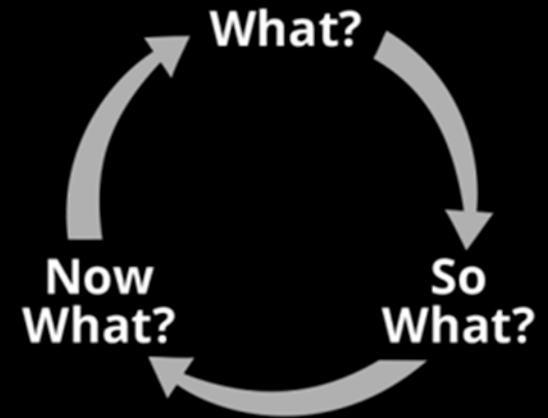
<http://dssresources.com/cases/DSScatalog.html#ADSCOUT>

# 3 Data Short Stories: the art & science of being data literate!

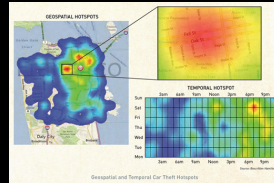
1) Personal

**2) Civic**

3) Commercial

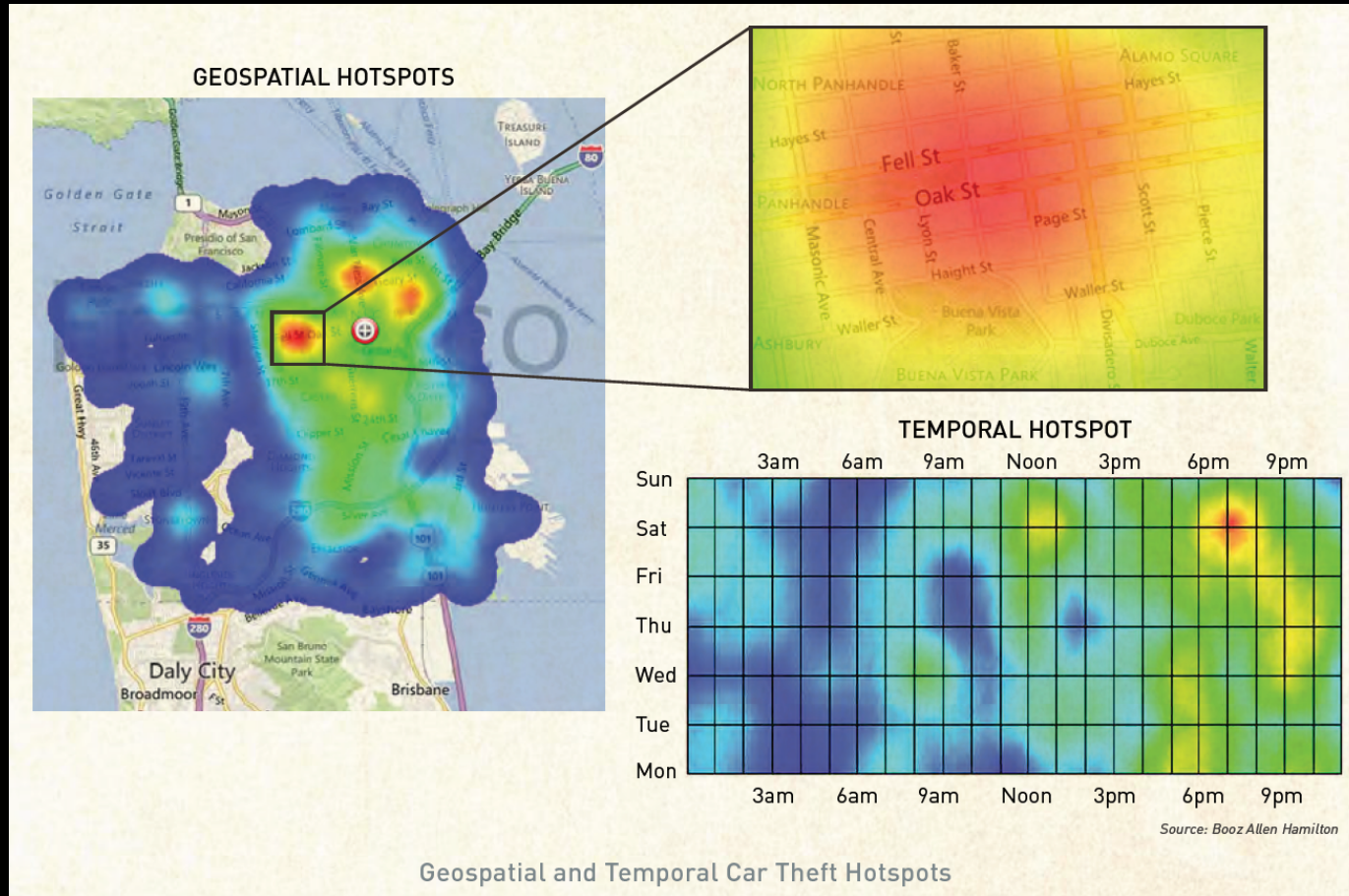


# Zoom deeper into your **Descriptive** Data for **Predictive** and **Prescriptive** Power Discovery... And for **Cognitive** Question Discovery!



(from the Booz Allen “Field Guide to Data Science”)

“What is going on in that neighborhood on Saturday evenings between 6pm and 8pm?”



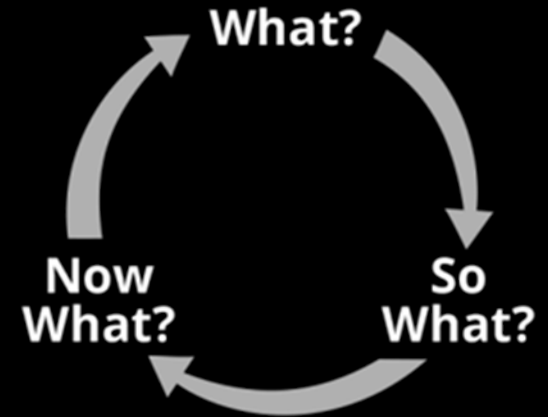
Source for graphic: <https://www.boozallen.com/s/insight/publication/field-guide-to-data-science.html>

# 3 Data Short Stories: the art & science of being data literate!

1) Personal

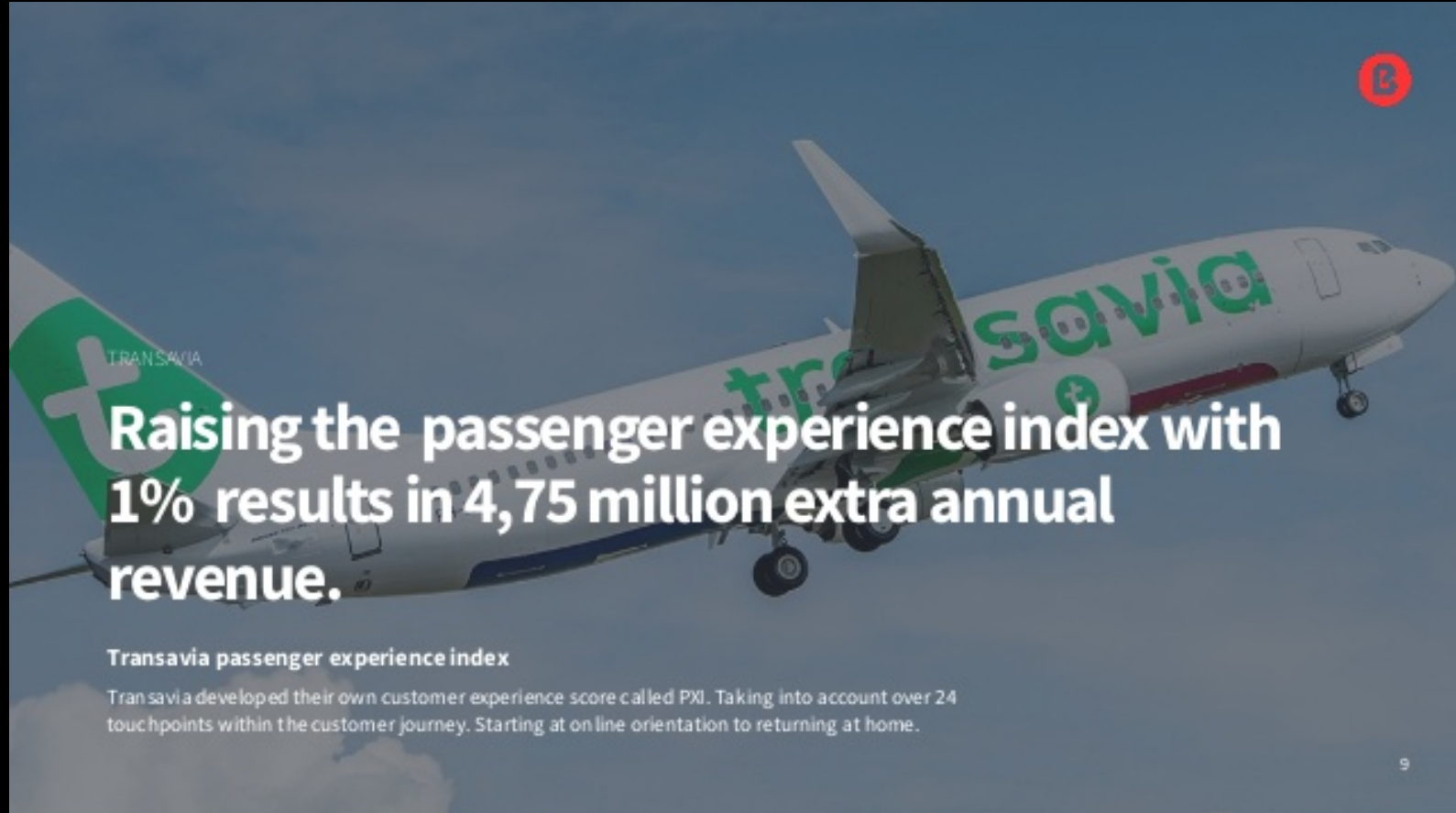
2) Civic

**3) Commercial**



**Data Democratization = Get on board with it, now!**

CEO declares, *“There is no more front office and back office.”*

A Transavia airplane is shown in flight against a blue sky with light clouds. The aircraft is white with green and red accents. The word 'transavia' is written in green on the side of the fuselage. A red circle with a white 'B' is visible in the top right corner of the image.

**Raising the passenger experience index with 1% results in 4,75 million extra annual revenue.**

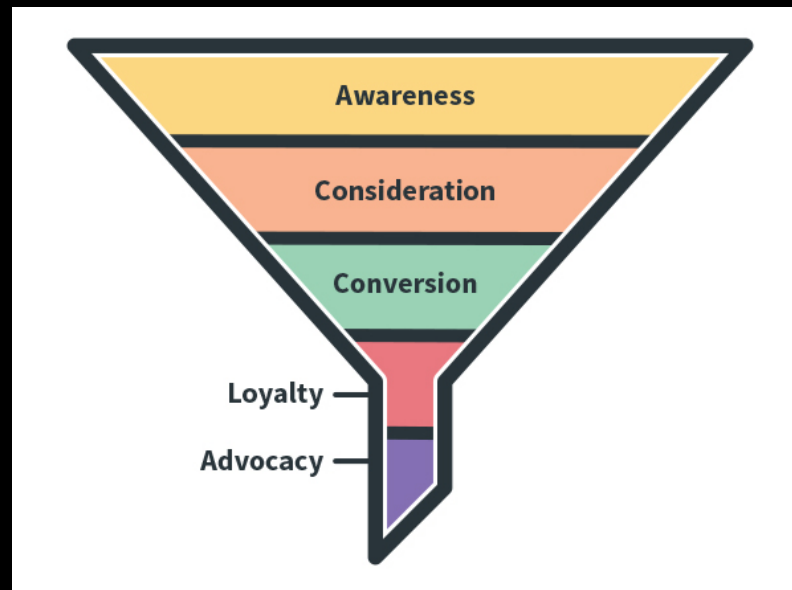
**Transavia passenger experience index**

Transavia developed their own customer experience score called PXI. Taking into account over 24 touchpoints within the customer journey. Starting at online orientation to returning at home.

9

# OUTLINE

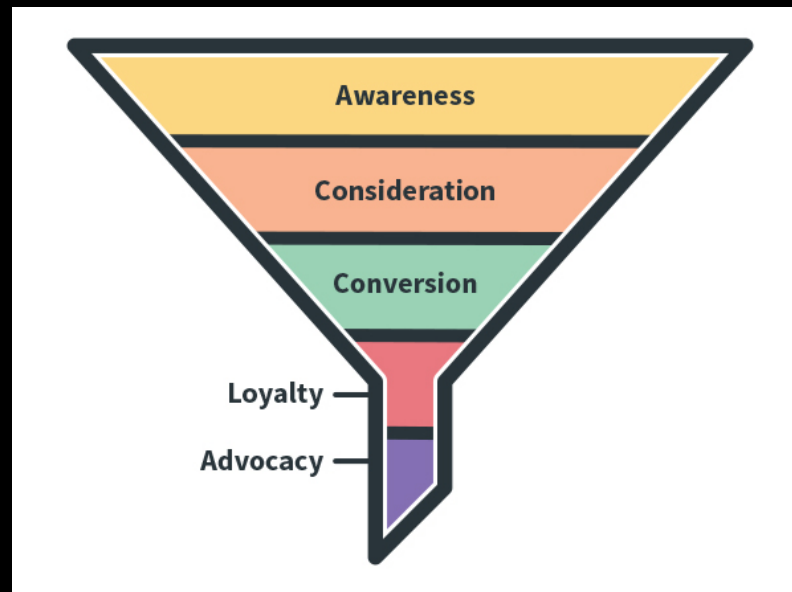
- Awareness
- Consideration
- Conversion
- Loyalty
- **Advocacy**



Source for graphic: <https://blog.aweber.com/email-marketing/understanding-the-marketing-funnel-5-strategies-to-improve-your-email-marketing.htm>

# The Data Literati – Passionate Advocates for the Use of Data

- Awareness
- Consideration
- Conversion
- Loyalty
- **Advocacy**



Source for graphic: <https://blog.aweber.com/email-marketing/understanding-the-marketing-funnel-5-strategies-to-improve-your-email-marketing.htm>

# Data Literacy as a Passion:

Building a Workforce of Explorers – to explore vast and endless seas of data!

*“If you want to build a ship,  
don’t drum up people to  
gather wood and don’t  
assign them tasks and work,  
but rather teach them to  
yearn for the vast and  
endless sea.”*

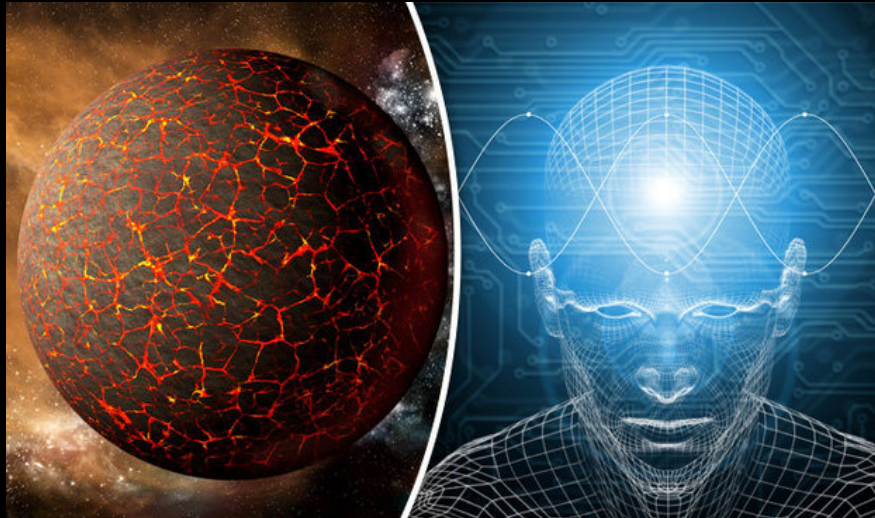
- Antoine de Saint-Exupery

<http://www.nytimes.com/2008/04/11/world/europe/11exupery.html>



<https://www.pinterest.com/pin/377106168772298092/>





*Come for the Data. Stay for the Science!*

**Thank you!**

Twitter: [@KirkDBorne](https://twitter.com/KirkDBorne) or Email: [kirk.borne@gmail.com](mailto:kirk.borne@gmail.com)

Get slides here: <http://www.kirkborne.net/UMUC2019>