DATA 601 –Ethics, Privacy, and More

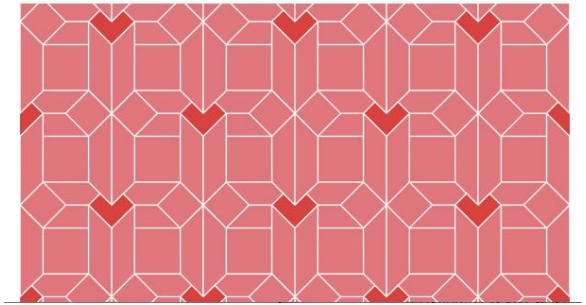
Felix Gonzalez, P.E.
Adjunct Instructor
Division of Professional Studies
Computer Science and Electrical Engineering
University of Maryland Baltimore County

Data Ethics is not a joke!

WIRED

MICHAEL ZIMMER SECURITY 05.14.16 7:00 AM

OKCUPID STUDY REVEALS THE PERILS OF BIG-DATA SCIENCE



theguardian

Facebook fiasco: was Cornell's study of 'emotional contagion' an ethics breach?

A covert experiment to influence the emotions of more than 600,000 people. A major scientific journal behaving like a rabbit in the headlights. A university in a PR tailspin

Facebook have recently come under fire for a controversial psychological study. Photograph: Dave Thompson/PA

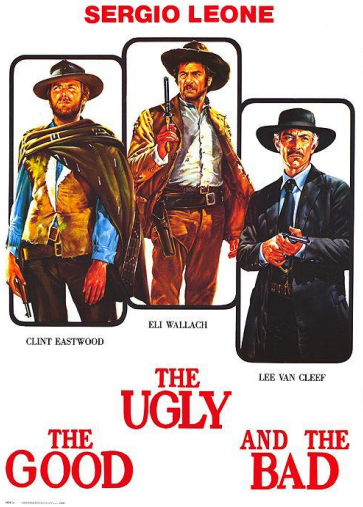
FORTUNE

Coming Soon: Ethics Training for Data Scientists

by Barb Darrow @gigabarb DECEMBER 4, 2015, 1:00 PM EST

The Good, the Bad and the Ugly

- Increasing efficiency, effectiveness, and competitiveness.
  - Data driven companies are more productive
- Use of data science make the societies more secure
  - Spam emails, fraud detection, terrorist attacks, shootings etc.
- Using AI for Sign Language Translation
- Tom Cruise (Deep Fakes)
- Barack Obama (Deep Fakes)



Ethics, Privacy, Legal Issues, and DS

Based on the nature of a project, you might see different issues at different stages,

- Initiating the project
- Creating the dataset (e.g. surveying)
- Cleaning the dataset
- Building the model
- Presenting the findings
- Sharing with others
- …

AI Enabled Surveillance

Goal: Determining terrorists

What if someone is having a heart attack?

Goal: Catching terrorists, criminals, drug traffickers

https://www.nytimes.com/2018/07/08/business/china-surveillance-technology.html

Current Use:

- Monitoring political activists and minorities
- Scoring people
- What about misuse?

Targeted Marketing - Customized Pricing

- Use of personal data increases the effectiveness of the marketing 3-5 times compared to traditional marketing approaches?
- Is it win-win or invasion of privacy?
- NYT-Target Case
- Different Pricing for the same product?

Forbes

Kashmir Hill, Forbes Staff
Welcome to The Not-So Private Parts where technology & privacy collide

TECH | 2/16/2012 @ 11:02AM | 1,958,223 views

How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did

Every time you go shopping, you share intimate details about your consumption patterns with retailers. And many of those retailers are studying those details to figure out what you like, what you need, and which coupons are most likely to make you happy. Target, for example, has figured out how to data-mine its way into your womb, to figure out whether you have a baby on the way long before you need to start buying diapers.

Charles Duhigg outlines in the New York Times how Target tries to hook parents-to-be

TARGET

Privacy (Anonymization is the ultimate solution)

Dr. Latanya Arvette Sweeney

In 1997, she identified then MA governor, William Weld, to his medical records using publicly accessible records.

An algorithm can identify 99.98 percent of Americans from almost any available data set with as few as 15 attributes, such as gender, ZIP code or marital status.

Rocher, L., Hendrickx, J.M. & de Montjoye, YA. Estimating the success of re-identifications in incomplete datasets using generative models. *Nat Commun* **10,** 3069 (2019)

https://www.nytimes.com/2019/07/23/health/data-privacy-protection.html

Bias and Fairness

- Credit approvals

AI Racial Bias Caused Mortgage Applicants To Be Denied (Link)

- Facial recognition

Facial recognition technologies accuracy may vary significantly across various demographic groups (Link)

- Criminal justice: recidivism

A computer program spat out a score predicting the likelihood of each committing a future crime. The algorithm was found to significant racial disparities (Link)

- Human resources: resumé screening

AI recruiting tool that showed gender bias (Link)

- Healthcare services

Racial bias found in health care algorithm (Link)

https://aiethics.princeton.edu/case-studies/case-study-pdfs/

AI's Invisibility Problem

- Diversity is missing from the data or not considered during development of algorithms resulting in inaccurate predictions for AI involving facial/image recognition

- Diversity challenges in data science related industries
  - https://money.cnn.com/2018/06/15/technology/google-diversity-report/index.html
  - https://www.npr.org/2020/07/02/886544638/we-have-a-black-people-problem-facebook-worker-claims-racial-discrimination
  - https://www.forbes.com/sites/carmenniethammer/2020/03/02/ai-bias-could-put-womens-lives-at-riska-challenge-for-regulators

What to do?

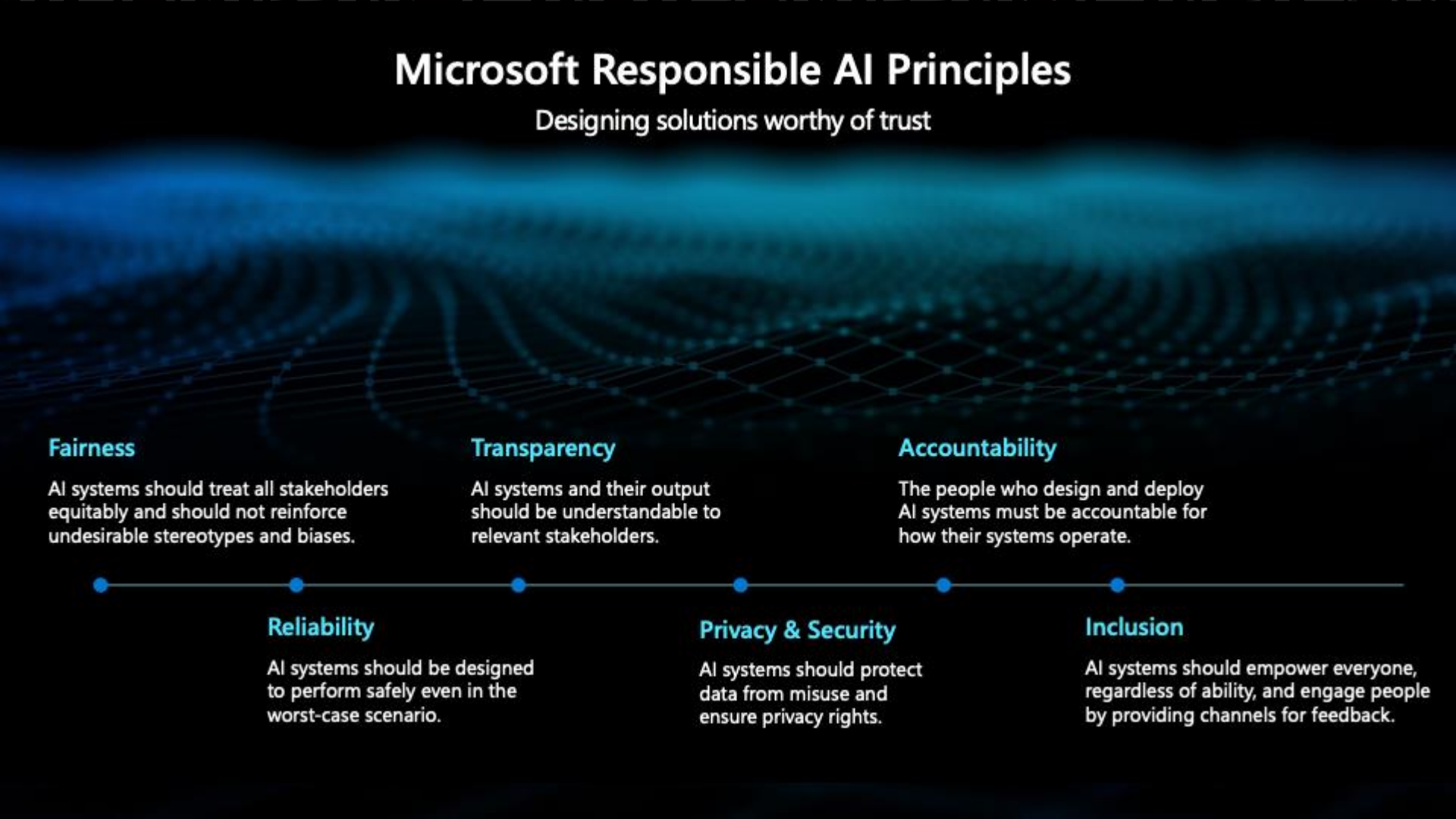
- Diagnosis by evaluating correlation between explanatory variables (Multicollinearity)
- Add more data
- Develop curated “gold standard” training data set
- Adjust the algorithm
- Unbalanced/Skewed Data ➔ Transform
- Rare events ➔ Oversample the data (Example)
- High frequency events ➔ Under-sample the data (Example)

Various Frameworks and Principles

- Examples:
  - Microsoft, Responsible AI Principles
  - O'Reilly, Radar; The five Cs of Data Products – O'Reilly, Radar
  - National Institute of Standards and Technology (NIST), Artificial Intelligence Risk Management Framework
  - U.S. Government Accountability Office (GAO), Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities

Microsoft Responsible AI Principles

Designing solutions worthy of trust



Fairness

AI systems should treat all stakeholders equitably and should not reinforce undesirable stereotypes and biases.

Transparency

AI systems and their output should be understandable to relevant stakeholders.

Accountability

The people who design and deploy AI systems must be accountable for how their systems operate.

Reliability

AI systems should be designed to perform safely even in the worst-case scenario.

Privacy & Security

AI systems should protect data from misuse and ensure privacy rights.

Inclusion

AI systems should empower everyone, regardless of ability, and engage people by providing channels for feedback.

O'Reilly Radar Five C's of Data Products

- **Consent:** A clear appreciation and understanding of the facts, implications, and consequences of an action
- **Clarity:** Who is going to use the data, when your data will be used, how long it will be kept, whether it will be shared with 3rd parties?)
- **Consistency and Trust:** How will be data stored? What will happen if stolen? Will company choose values when it comes to chose between values and money?
- **Control and Transparency:** Who has the control over the data? Will the participants be able to change/remove their data?
- **Consequences:** Unknown unknowns: whether a company invest on thinking possible consequences?

Source: The five Cs – O'Reilly (oreilly.com)

National Institute of Standards and Technology

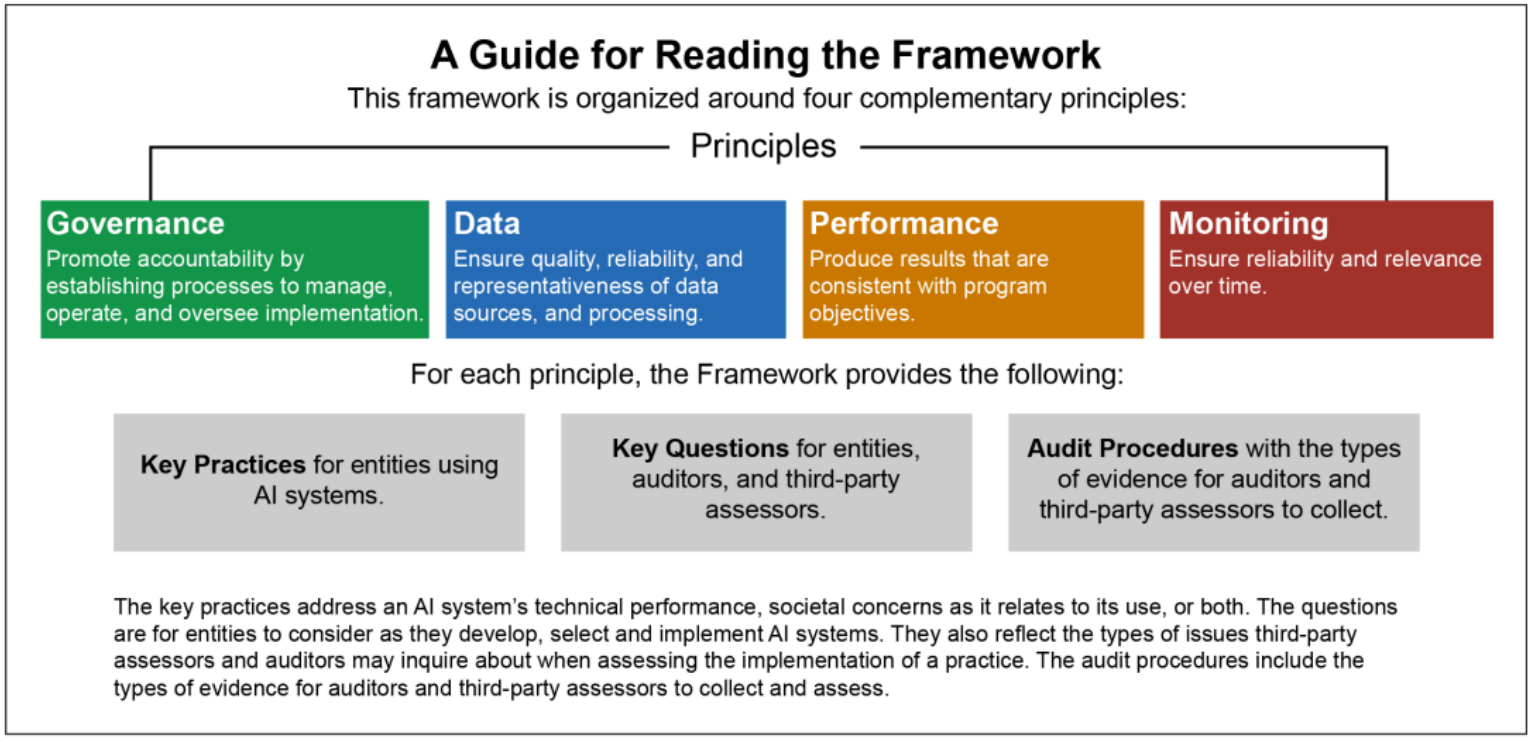
- NIST, Artificial Intelligence Risk Management Framework.

Lifecycle	Activities	Representative Actors
Plan & design	Articulate and document the system's concept and objectives, underlying assumptions, context and requirements.	System operators, end-users, domain experts, AI designers, impact assessors, TEVV experts, product managers, compliance experts, auditors, governance experts, organizational management, end-users, affected individuals/communities, evaluators.
Collect & process data	Data collection & Processing: gather, validate, and clean data and document the metadata and characteristics of the dataset.	Data scientists, domain experts, socio-cultural analysts, human factors experts, data engineers, data providers, TEVV experts.
Build & use model	Create or select, train models or algorithms.	Modelers, model engineers, data scientists, developers, and domain experts. With consultation of socio-cultural analysts familiar with the application context, TEVV experts.
Verify & validate	Verify & validate, calibrate, and interpret model output.	
Deploy	Pilot, check compatibility with legacy systems, verify regulatory compliance, manage organizational change, and evaluate user experience.	System integrators, developers, systems/software engineers, domain experts, procurement experts, third-party suppliers with consultation of human factors experts, socio-cultural analysts, and governance experts, TEVV experts, end-users.
Operate & monitor	Operate the AI system and continuously assess its recommendations and impacts (both intended and unintended) in light of objectives and ethical considerations.	System operators, end-users, domain experts, AI designers, impact assessors, TEVV experts, product managers, compliance experts, auditors, governance experts, organizational management, end-users, affected individuals/communities, evaluators.
Use or impacted by	Use system/technology; monitor & assess impacts; seek mitigation of impacts, advocate for rights.	End-users, affected individuals/communities, general public; policy makers, standards organizations, trade associations, advocacy groups, environmental groups, civil society organizations, researchers.

Figure 2: AI actors across the AI lifecycle.

U.S. Government Accountability Office

- U.S. Government Accountability Office (GAO), Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities.



Source: GAO. | GAO-21-519SP

Minimum Practices for Data Science Projects

1. Documentation: promote transparency and trust
   a. Have we evaluated and documented our main goal, objective, roles and responsibilities, design, development, capabilities, limitations, data sources, methods for assessment, model specifications and parameters, performance evaluation, risks, and controls and barriers that are in place to prevent or mitigate risks?
   b. Have we conducted an AI impact or risk assessment?

2. Risks:
   a. Have we studied and understood possible sources of bias in our data? (Note the effect of data bias in supervised and unsupervised methods is different)
   b. Have we tested our training data to ensure it is fair and representative of use cases?
   c. Have we identified and have understanding of potential bias, inequities, societal concerns, and other risks?
   d. Have we tested for fairness with respect to different user groups?
   e. Have we tested for disparate error rates among different user groups?
   f. Have we listed how this technology can be attacked or abused?
   g. Is our system reliable, reproducible, traceable, and with transparent and explainable outputs (as permitted by the system)?

Minimum Practices, Cont.

7. Stakeholders:
   a. Multi-disciplinary team: Does our team reflect diversity of opinions, backgrounds, and kinds of thought?
   b. Do we have a process to capture feedback from stakeholders (e.g., domain experts, industry, public, etc.)?

8. Controls and Barriers:
   a. Do we have a mechanism for redress if people are harmed by the results?
   b. Do we test and monitor for model drift to ensure our software remains fair over time?
   c. Do we monitor and provide oversight to the system?
   d. Do we identify who is accountable for the system?

9. What kind of user consent do we need to collect to use the data and how to collect such consent? Have we explained clearly what users are consenting to?

10. Security and Privacy
    a. Can we shut down this software in production if it is behaving badly?
    b. Do we have a plan to protect and secure user data?

Other Examples and References

- Top 9 Ethical Dilemmas of AI and How to Navigate Them in 2022 (https://research.aimultiple.com/ai-ethics/)

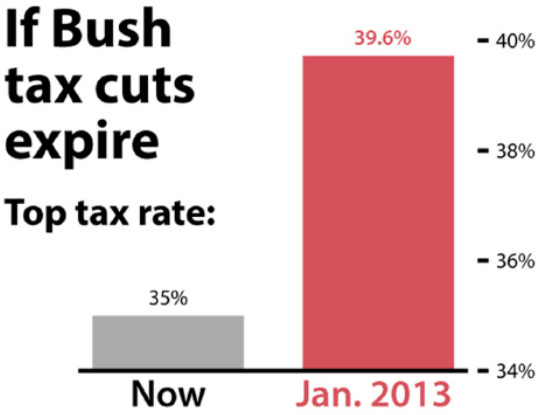
- Jumpstart Article on "AI Gone Wrong 5 Biggest AI Failures of All Time" https://www.jumpstartmag.com/ai-gone-wrong-5-biggest-ai-failures-of-all-time/

- Rogan grills Zuckerberg on how Facebook moderates' controversial content (https://www.youtube.com/watch?v=irCFLEVUuIo)

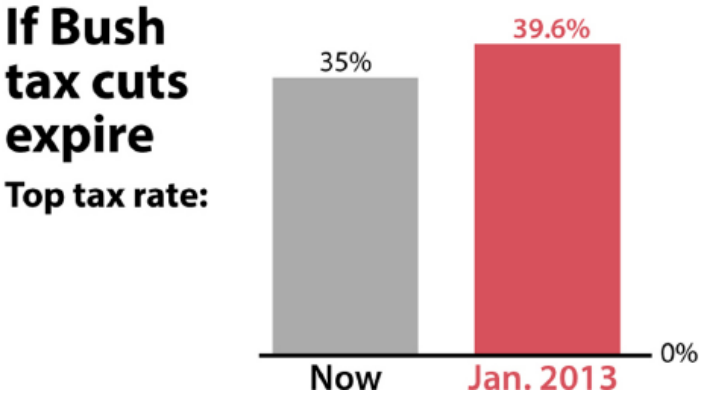
Ethical Principles of Data Visualization

- Well-designed charts are empowering.
- They can also deceive viewers regardless of whether they are designed with ill-intent or not.
- Data visualization has ethical consequences
  - Ethical thinking is not only about intentions, but also about consequences.
- Need to explicitly consider ethical principles when creating and distributing visualizations.
  - Transparency
  - Integrity
  - Accountability

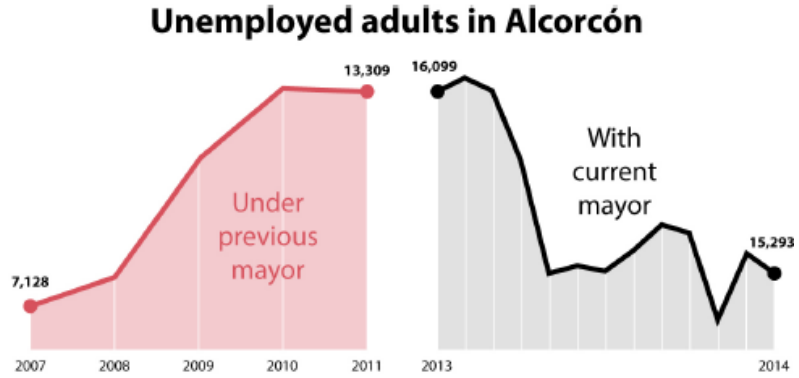
Don’t



It should have been

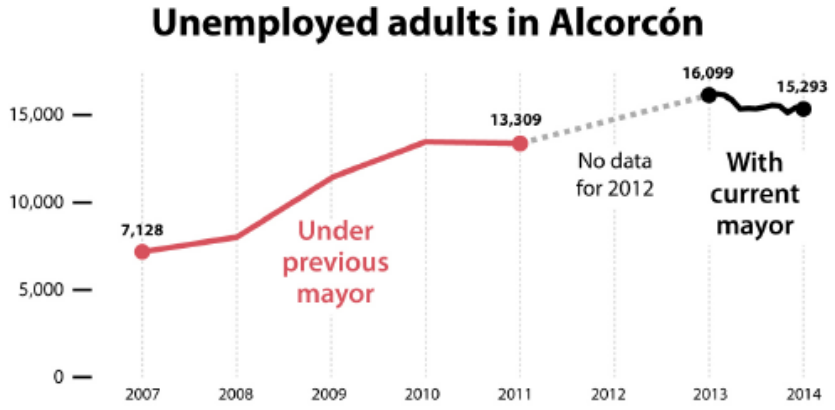


Don’t



Created by the Spanish city of Alcorcon in 2014 to celebrate the success of the current Mayor in increasing employment.

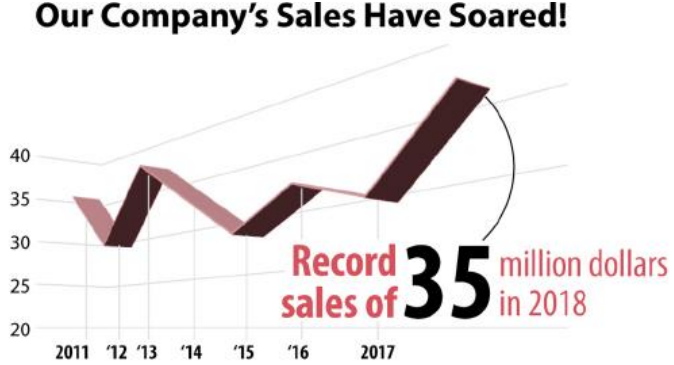
Distorted both vertical and horizontal scales.

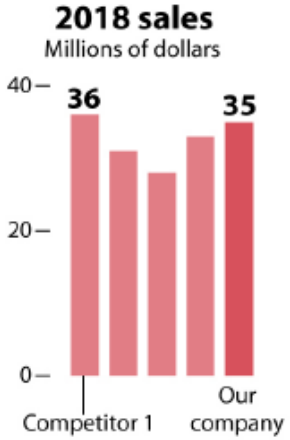


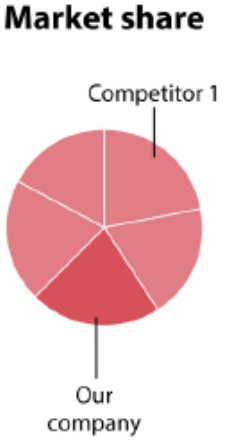
Don’t

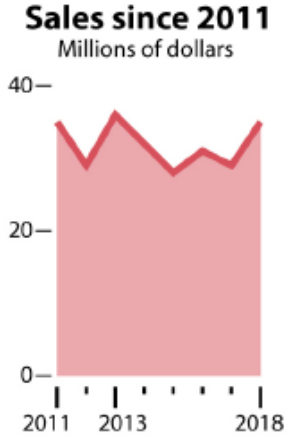

Record Sales!
35
million dollars in 2018
Competitor 1
Competitor 2
Competitor 3
Competitor 4
Our Company


Share of All Companies in Our Market
Competitor 4
Competitor 1
Competitor 3
Competitor 2
Record sales of 35 million dollars in 2018


Our Company’s Sales Have Soared!
40
35
30
25
20
2011
’12
’13
’14
’15
’16
2017
Record sales of 35 million dollars in 2018

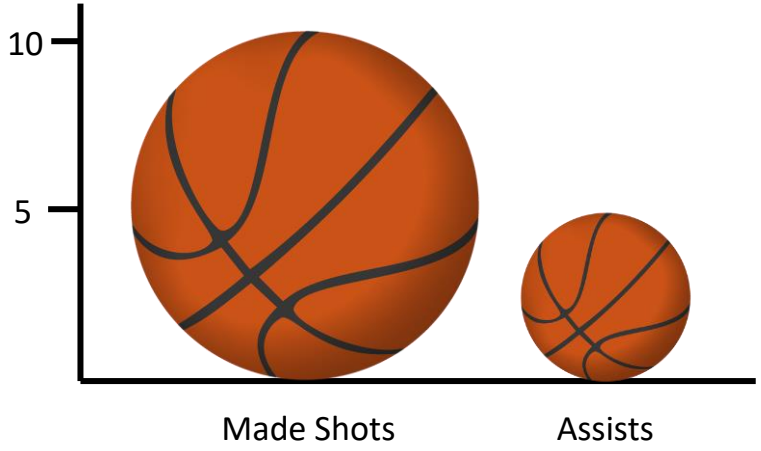

2018 sales
Millions of dollars
40
20
0
36
35
Competitor 1
Our company

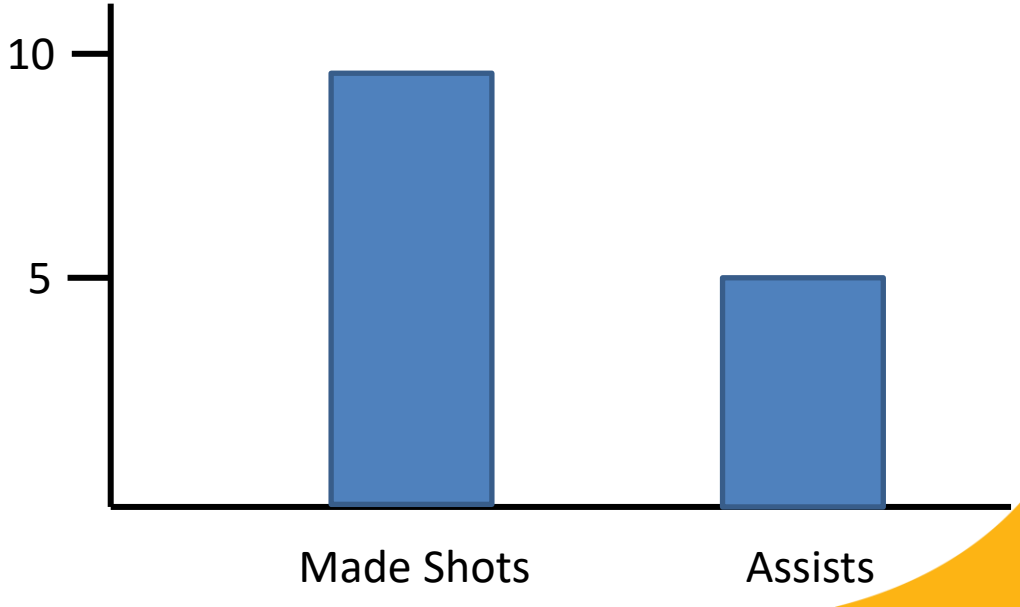

Market share
Competitor 1
Our company

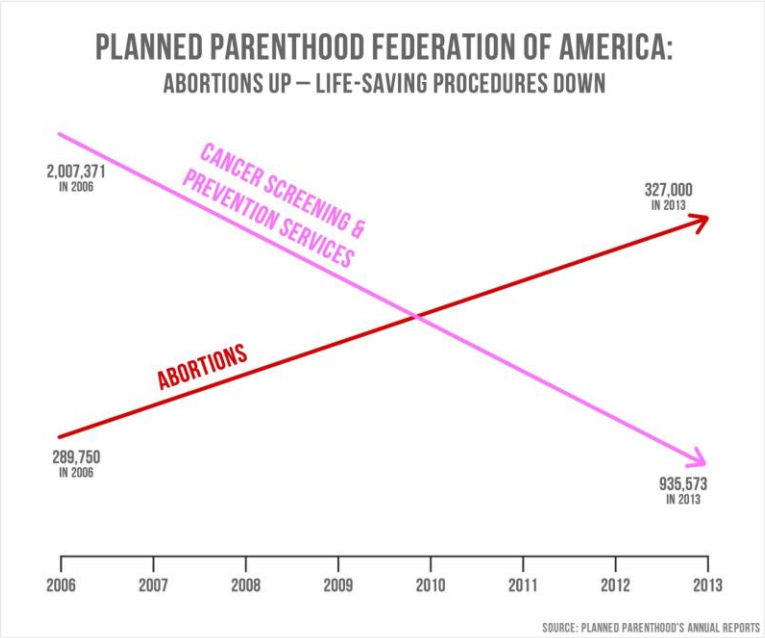

Sales since 2011
Millions of dollars
40
20
0
2011
2013
2018

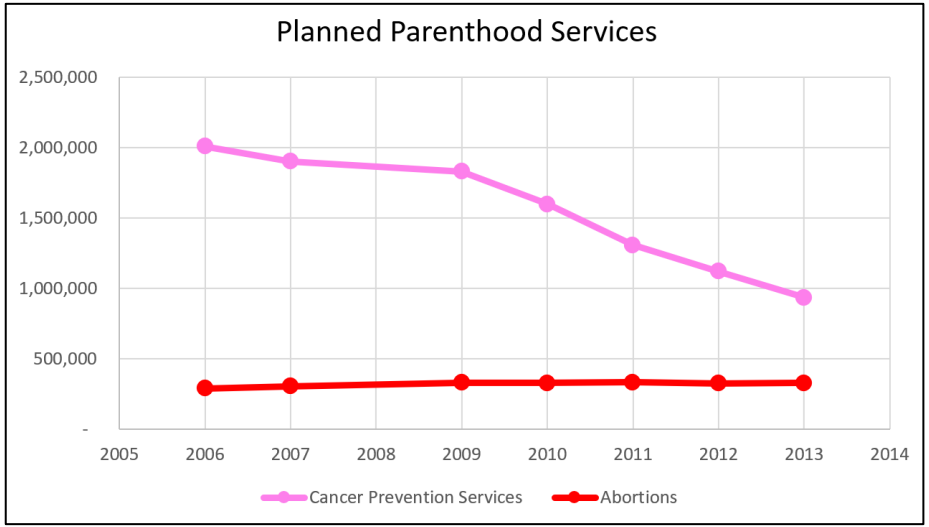
Was Kobe a ball hog?

In the 2009-2010 season, Kobe made 9.8 shots per game and had 5.0 assists per game, on average.

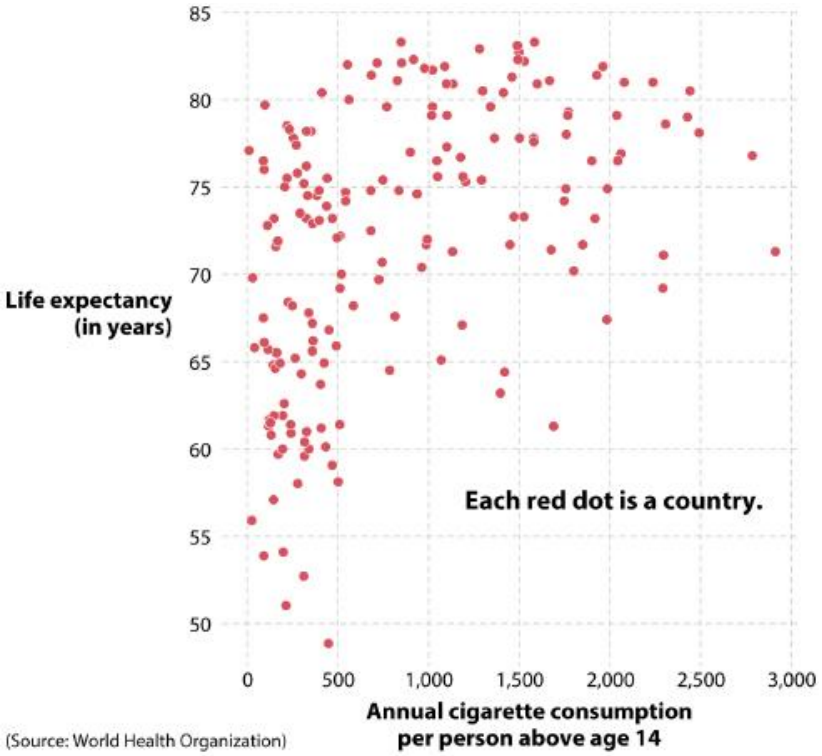





PLANNED PARENTHOOD FEDERATION OF AMERICA:
ABORTIONS UP – LIFE-SAVING PROCEDURES DOWN
2,007,371
IN 2006
CANCER SCREENING &
PREVENTION SERVICES
327,000
IN 2013
ABORTIONS
289,750
IN 2006
935,573
IN 2013
2006
2007
2008
2009
2010
2011
2012
2013
SOURCE: PLANNED PARENTHOOD'S ANNUAL REPORTS

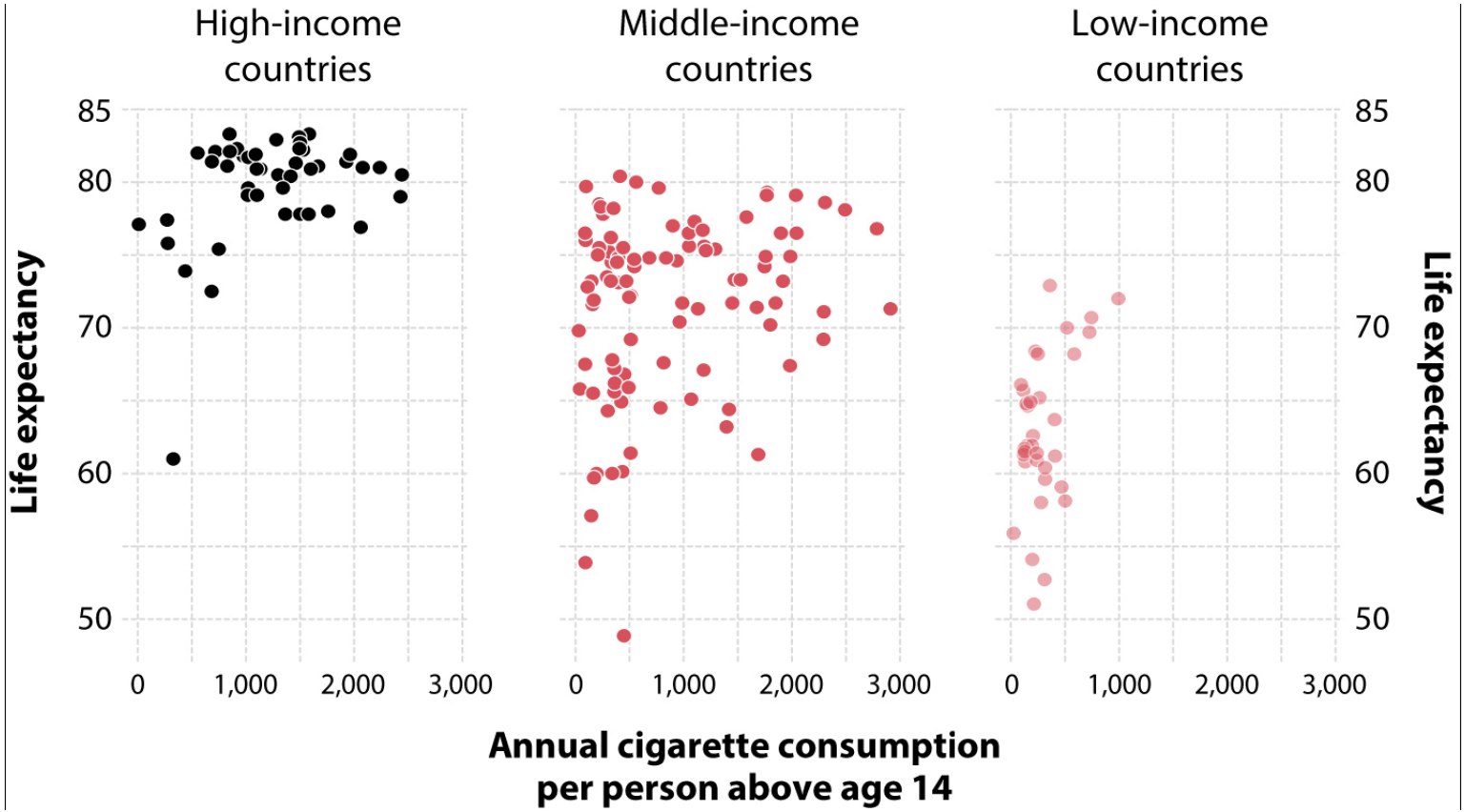

Planned Parenthood Services
2,500,000
2,000,000
1,500,000
1,000,000
500,000
-
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
Cancer Prevention Services
Abortions

Correlation is not equal to causation



The more cigarettes we consume, the longer we live!!!

Correlation is not equal to causation



Before You Create a Chart

Ask

- Who is my audience?
- What questions do they have?
- What answers am I finding for them?
- What am I trying to say?
- What other questions will my visualization inspire or what conversations may result?

Data Visualization Best Practices

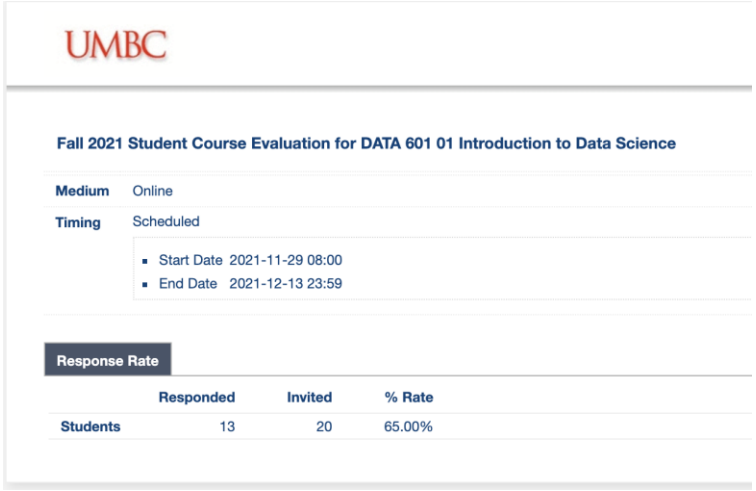
- Identify your target audience
- Make sure your data is clean
- Select the right chart
- Label your chart effectively
- Emphasize the important points
- Format your chart for accessibility
- Make use of color
- Ensure your data is readable in any format
- Ask for feedback

Questions?

Additional Slides

Survey Design is Almost an Art

- Population matters
  - Females more likely to respond than males
  - Younger generation less likely to join



UMBC

Fall 2021 Student Course Evaluation for DATA 601 01 Introduction to Data Science

Medium	Online
Timing	Scheduled

- Start Date 2021-11-29 08:00
- End Date 2021-12-13 23:59

Response Rate

	Responded	Invited	% Rate
Students	13	20	65.00%

- Questions matter ➔ Next Slides

Survey Design: Questions Matter

A seemingly straight forward question can lead to inaccurate data

The global economy is the most important issue in the world today.

strongly disagree	disagree	neither agree nor disagree	agree	strongly agree

Now look at the revised version

How important is the issue of the global economy in the world today?

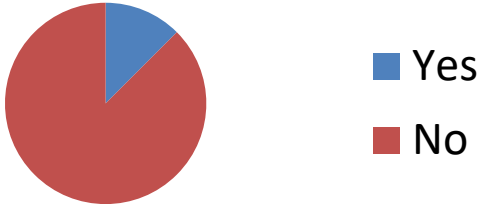
not at all Important	slightly important	moderately important	quite important	extremely important

Survey Design: Questions Matter

Providing a reason…

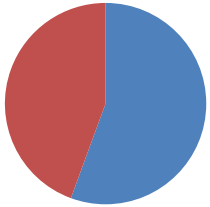
Q8) As some of you may know, the university is debating whether to move some parts of the university to a new section of campus in Rockville. Do you think the university should move to Rockville?

Yes = 12.5%



Q8) As some of you may know, the university is debating whether to move some parts of the university to a new section of campus in Rockville. Do you think the university should move to Rockville ***so that the school can have more space?***

Yes = 55.6%



Giving a reason will improve the extent to which people agree with you, and impact the accuracy of your survey

Survey Design: 7-Step Process

- Step 1: Literature Review
- Step 2: Interview Experts & Create Focus Groups from Target Population
- Step 3: Synthesize
- Step 4: Develop Items
- Step 5: Expert Validation
- Step 6: Cognitive Interviewing
- Step 7: Pilot Test