

Data Analytics

CS301

Introduction to Data Analytics

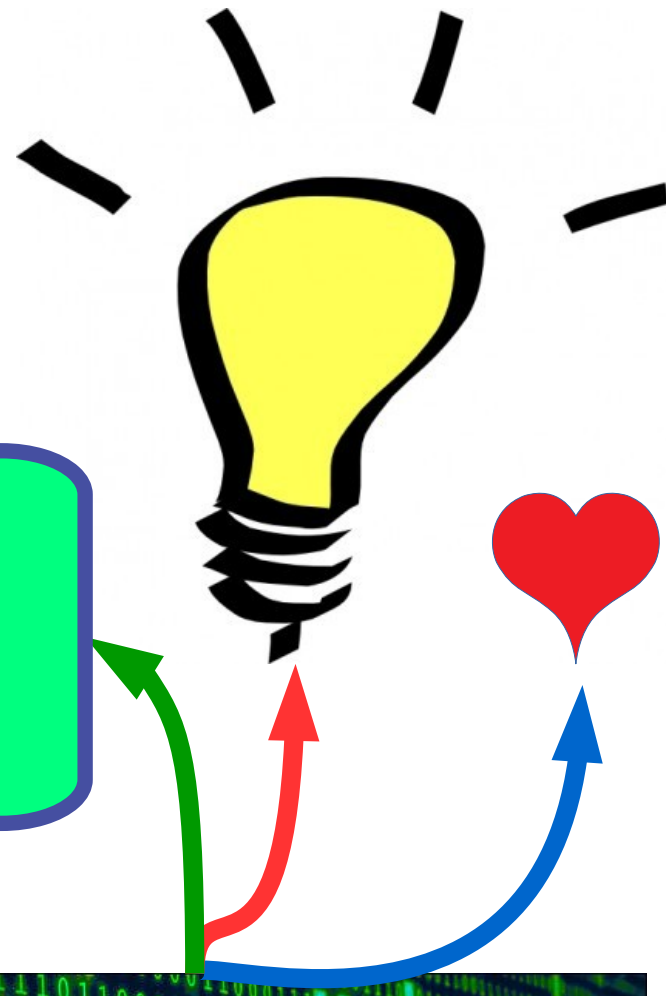
Week 1

Spring 2023

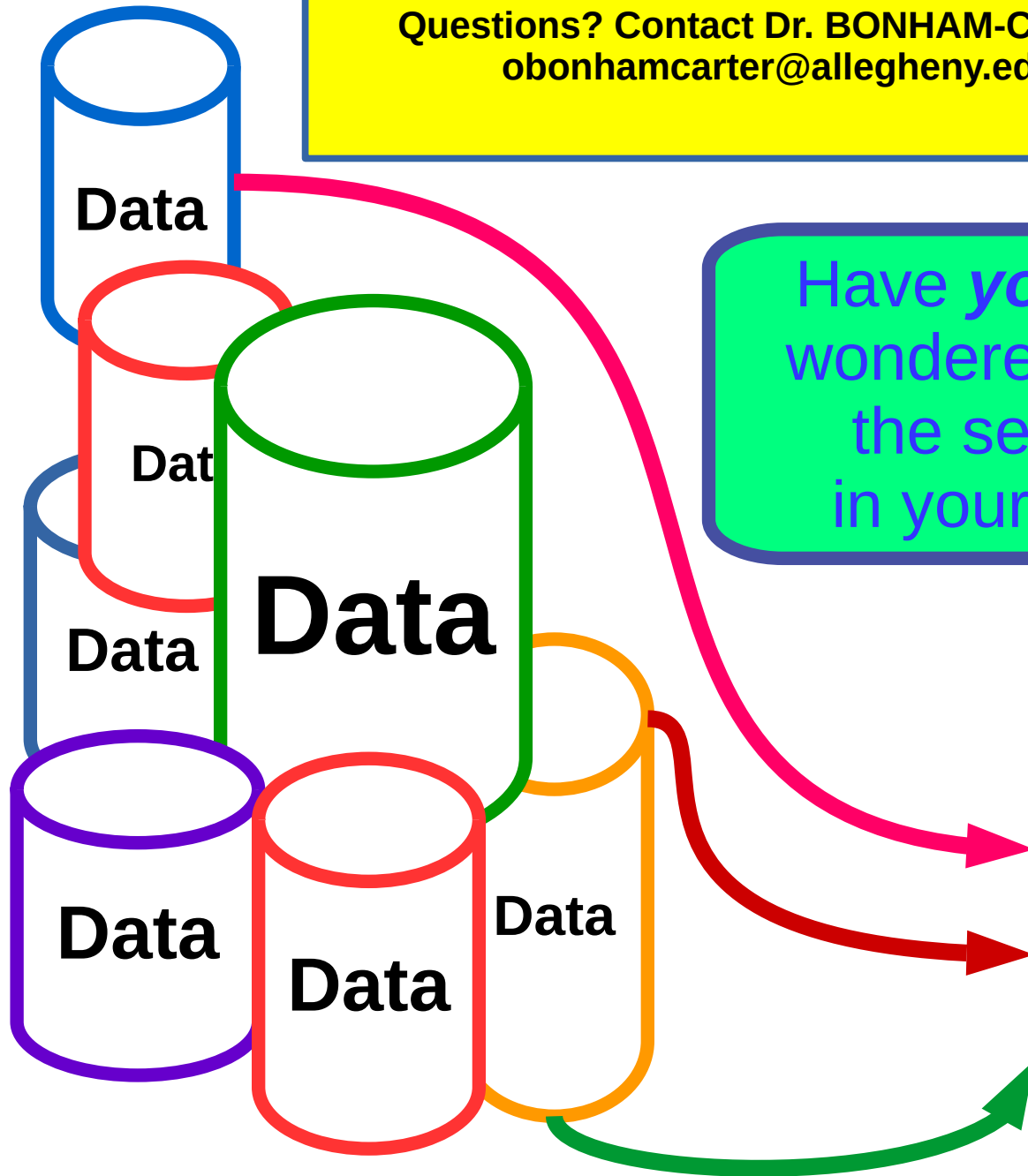
Oliver BONHAM-CARTER

Data Analytics CMPSC*301

Questions? Contact Dr. BONHAM-CARTER
obonhamcarter@allegheny.edu

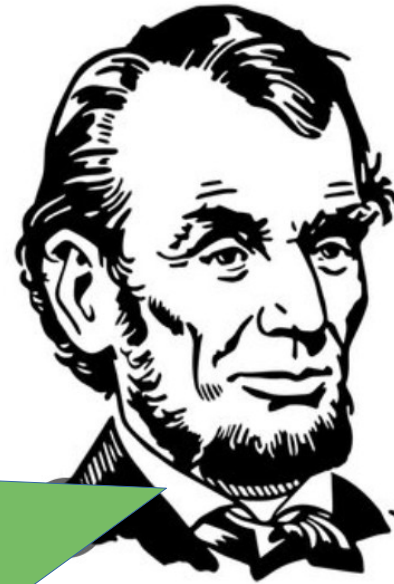


Have *you* ever
wondered about
the secrets
in your data?



Get Link-IN!

- Course web site:
 - <https://www.oliverbonhamcarter.com/classes/dataanalytics/dataanalytics/>
- Syllabus, Planning and ClassDocs:
 - <https://github.com/CMPSC-301-Allegheny-College-Spring-2023/classDocs>



I, Oliver
Bonham-
Carter 🖐️

Home / Classes / Data Analytics

Data Analytics

Computers and Information



ALLEGHENY
COLLEGE



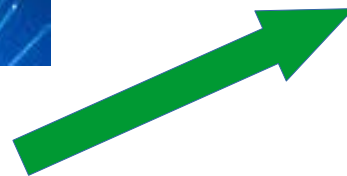


Computers and Information

- In this class, you will learn how to use machines to understand *trends* in data.
- (Making decisions by data)



Raw
Data



Meaningful
Information



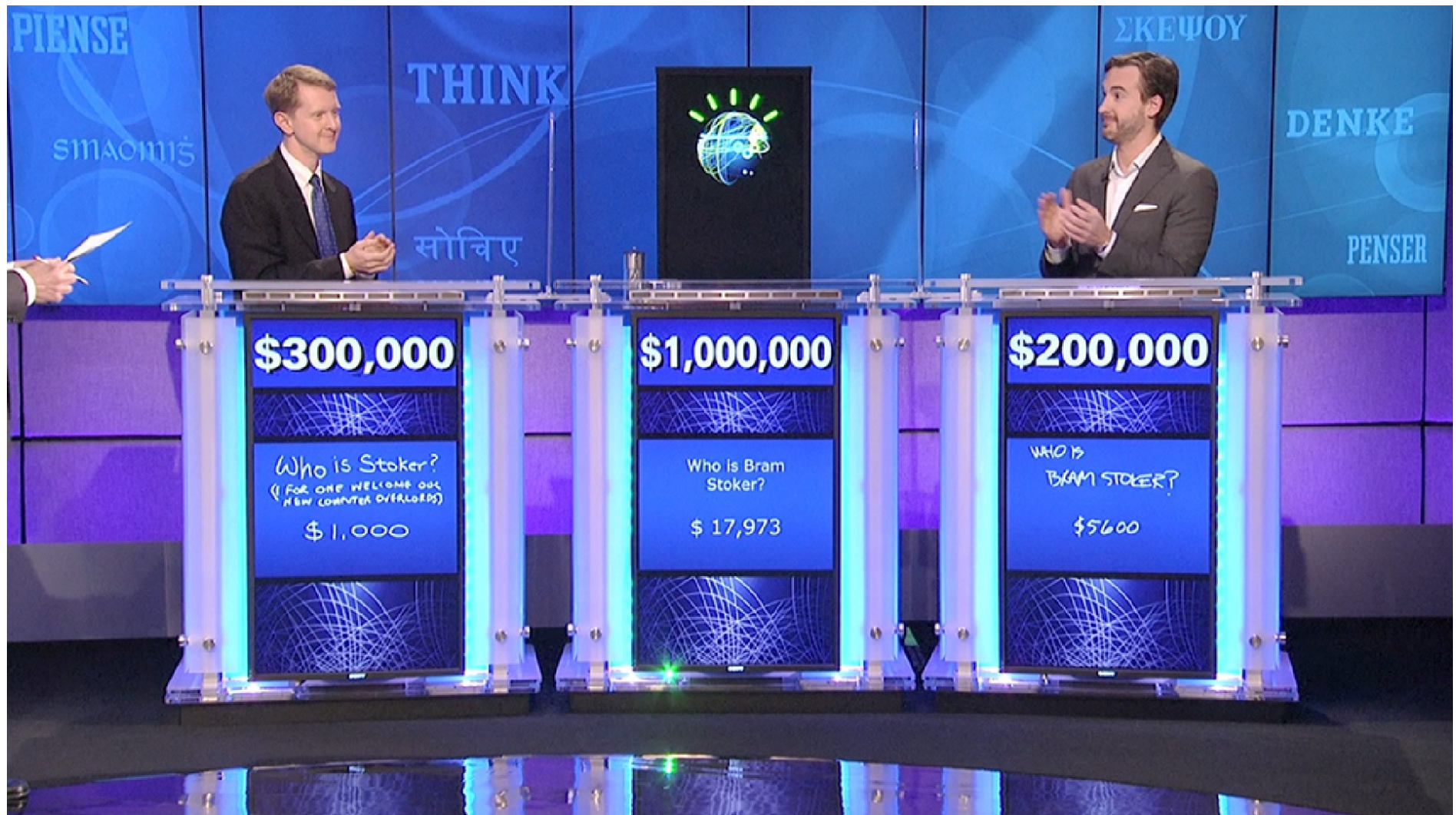


Analytics in Action

- The Jeopardy! Challenge of February 2011
- IBM's Watson beat the show's greatest champions: Ken Jennings and Brad Rutter.



Machines, Data and Information



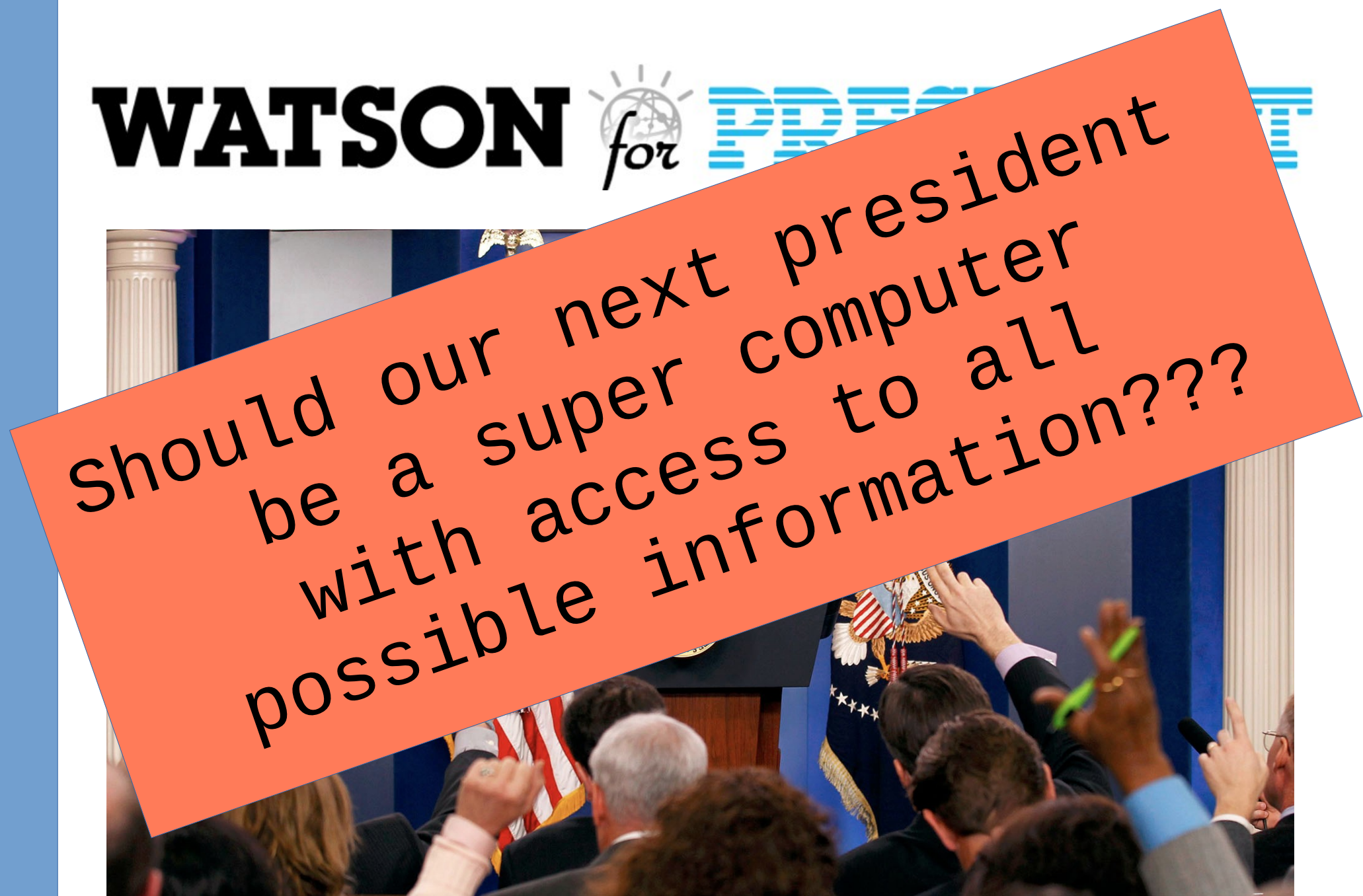
WATSON for PRESIDENT



Is Watson magic??

<http://watson2016.com/>
(The Electronic Frontier Foundation)

WATSON for PRESIDENT



Is Watson magic??

<http://watson2016.com/>
(The Electronic Frontier Foundation)

Surrounded by DATA!

- We live in the “Information age”
- Actually, we live in the “Data age” since there is more data available than information
- Data != Information



Surrounded by DATA!

- It is cheap (and free or even lucrative) for businesses to collect data concerning:
 - in e-commerce,
 - customer behaviors,
 - purchase interests,
 - health and medical data.



Meet the Fitbit Family



EVERYDAY FITNESS

ACTIVE FITNESS

PERFORMANCE FITNESS

SMART SCALE

zip

one

flex

alta

chargeHR

blaze

surge

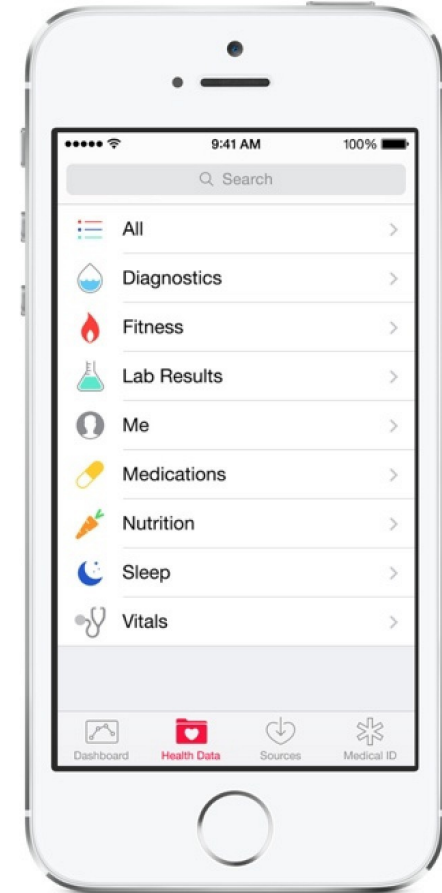
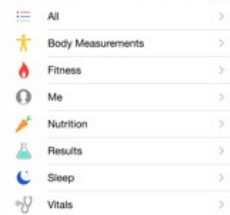
aria

We [Voluntarily] Contribute to this Data





Our Phones Create Data



- Smart phones constantly monitor us and keep data.
- **Q: How does the iPhone decide whether we are actually getting enough sleep?**
- **Who keeps this data?**



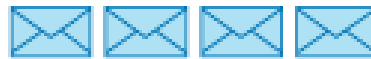
ALLEGHENY
COLLEGE

And So, Data is Increasing

 **65 billion**

Location-tagged payments
made in the U.S. annually

154 billion



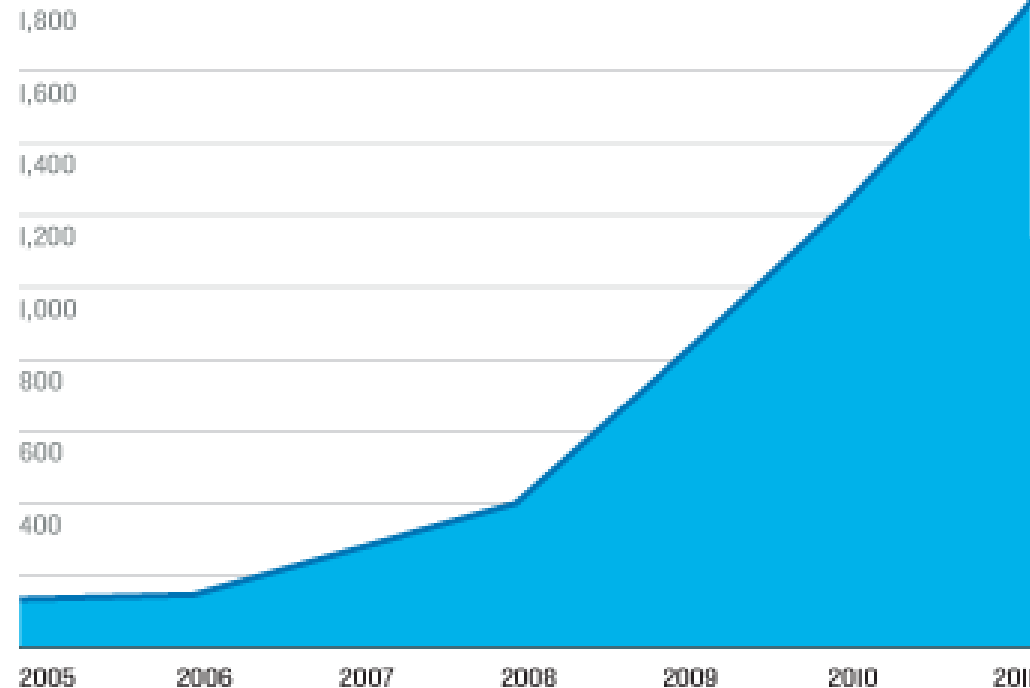
E-mails sent per day

 **87%**

U.S. adults whose location is
known via their mobile phone

Digital Information Created Each Year, Globally

2,000 BILLION GIGABYTES



2,000%

Expected increase in
global data by 2020

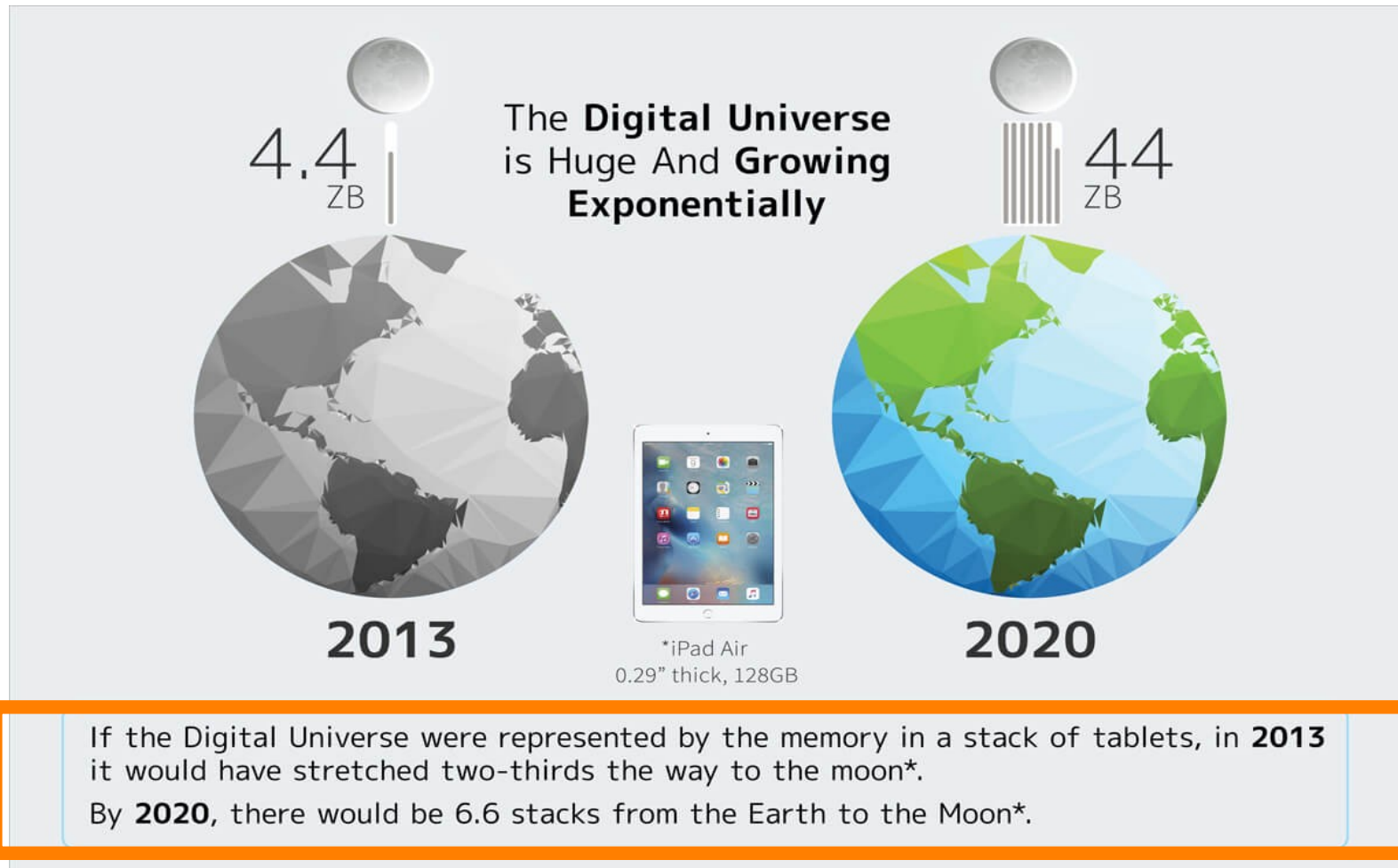
**III
Megabytes**

Video and photos stored
by Facebook, per user

75%

Percentage of all digital
data created by consumers

Data, Data, Data, Data!

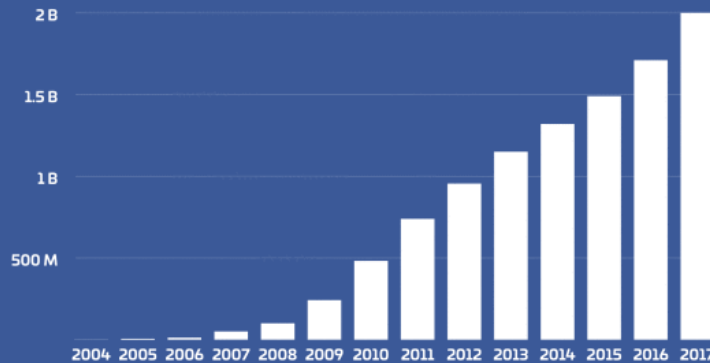


- How much data is there?
 - <https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-creat-e-every-day-the-mind-blowing-stats-everyone-should-read/#76dc5de060ba>
 - Data created each day: <https://youtu.be/VLAnBI2B4OY>

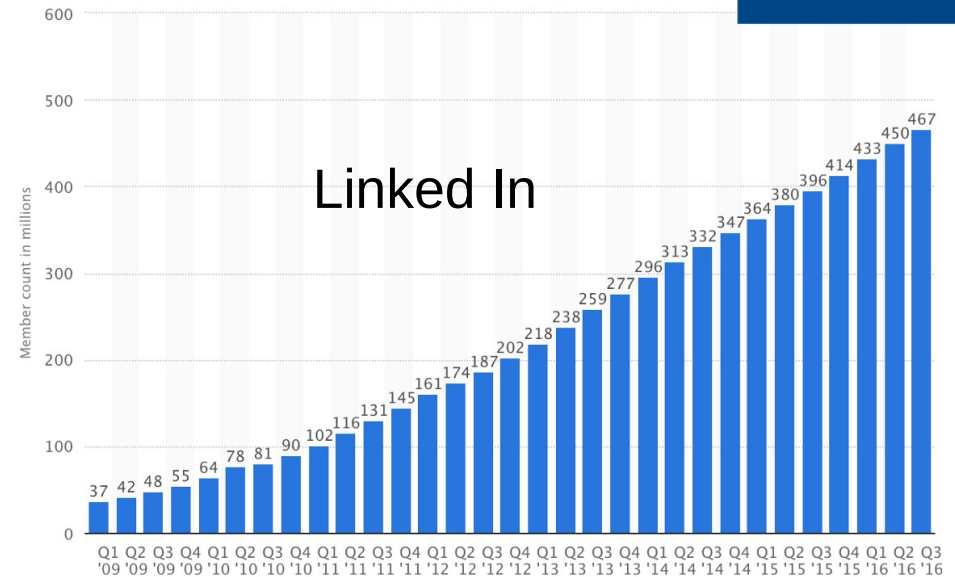


Using Data?!

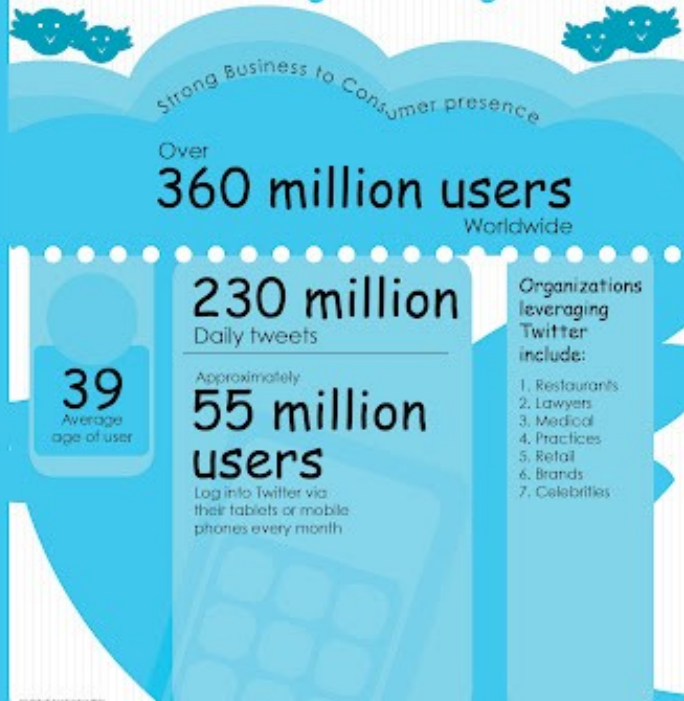
FACEBOOK MONTHLY ACTIVE USERS
JUNE 2017



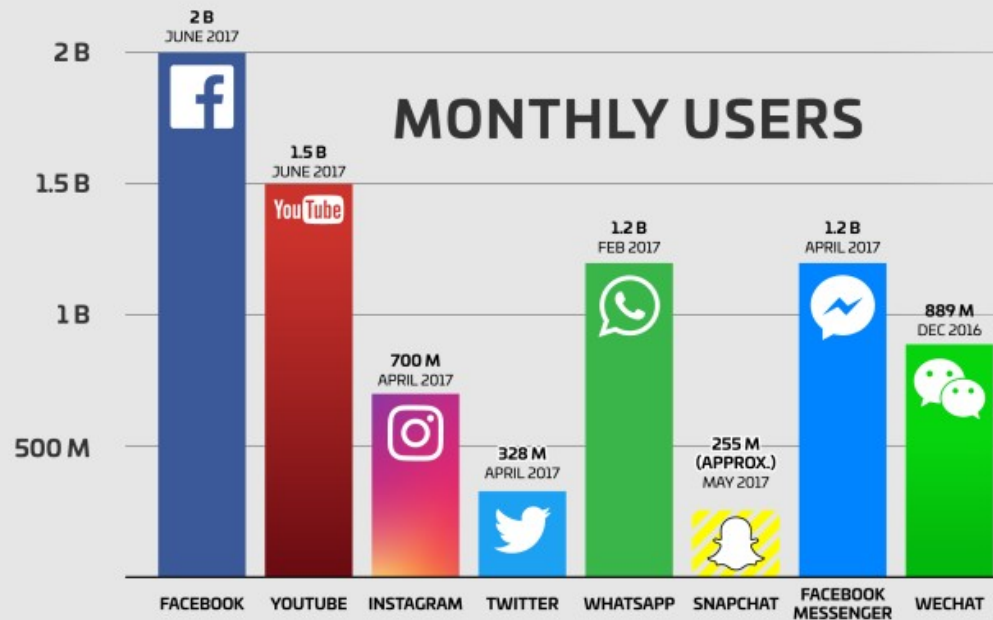
Linked In



twitter fast facts



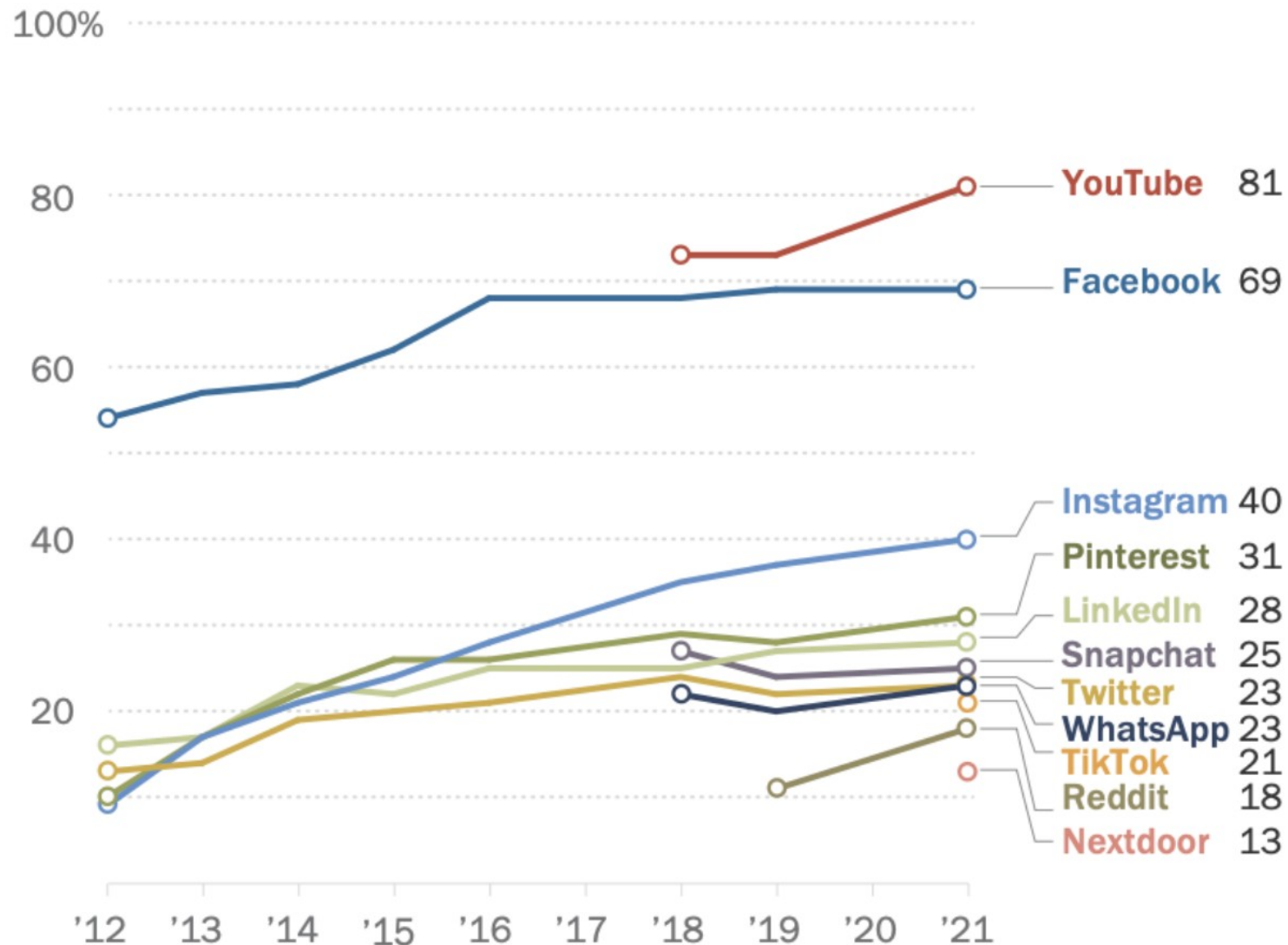
MONTHLY USERS





Growing share of Americans say they use YouTube; Facebook remains one of the most widely used online platforms among U.S. adults

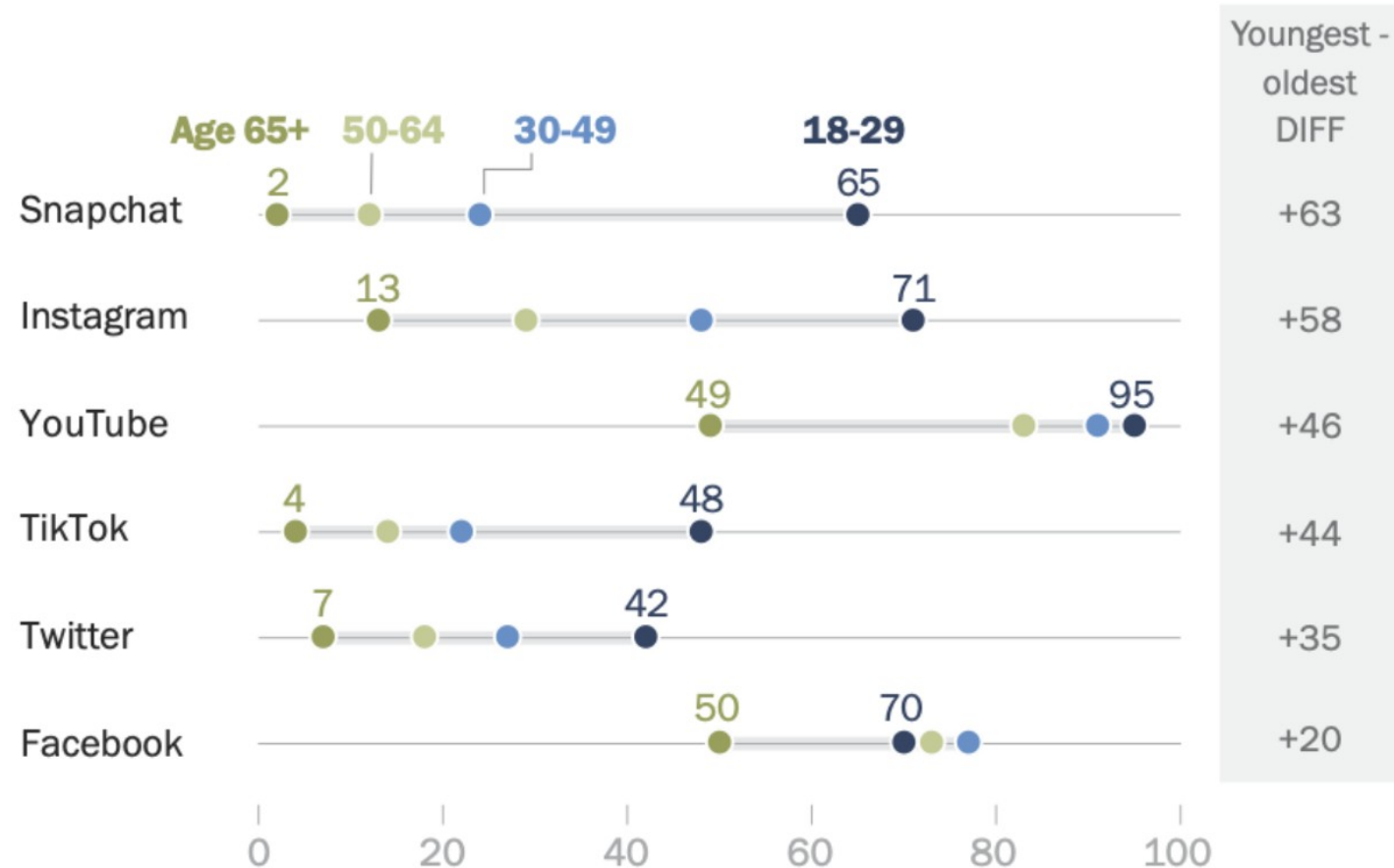
% of U.S. adults who say they ever use ...





Age gaps in Snapchat, Instagram use are particularly wide, less so for Facebook

% of U.S. adults in each age group who say they ever use ...



Note: All differences shown in DIFF column are statistically significant. The DIFF values shown are based on subtracting the rounded values in the chart. Respondents who did not give an answer are not shown.

Source: Survey of U.S. adults conducted Jan. 25-Feb. 8, 2021.

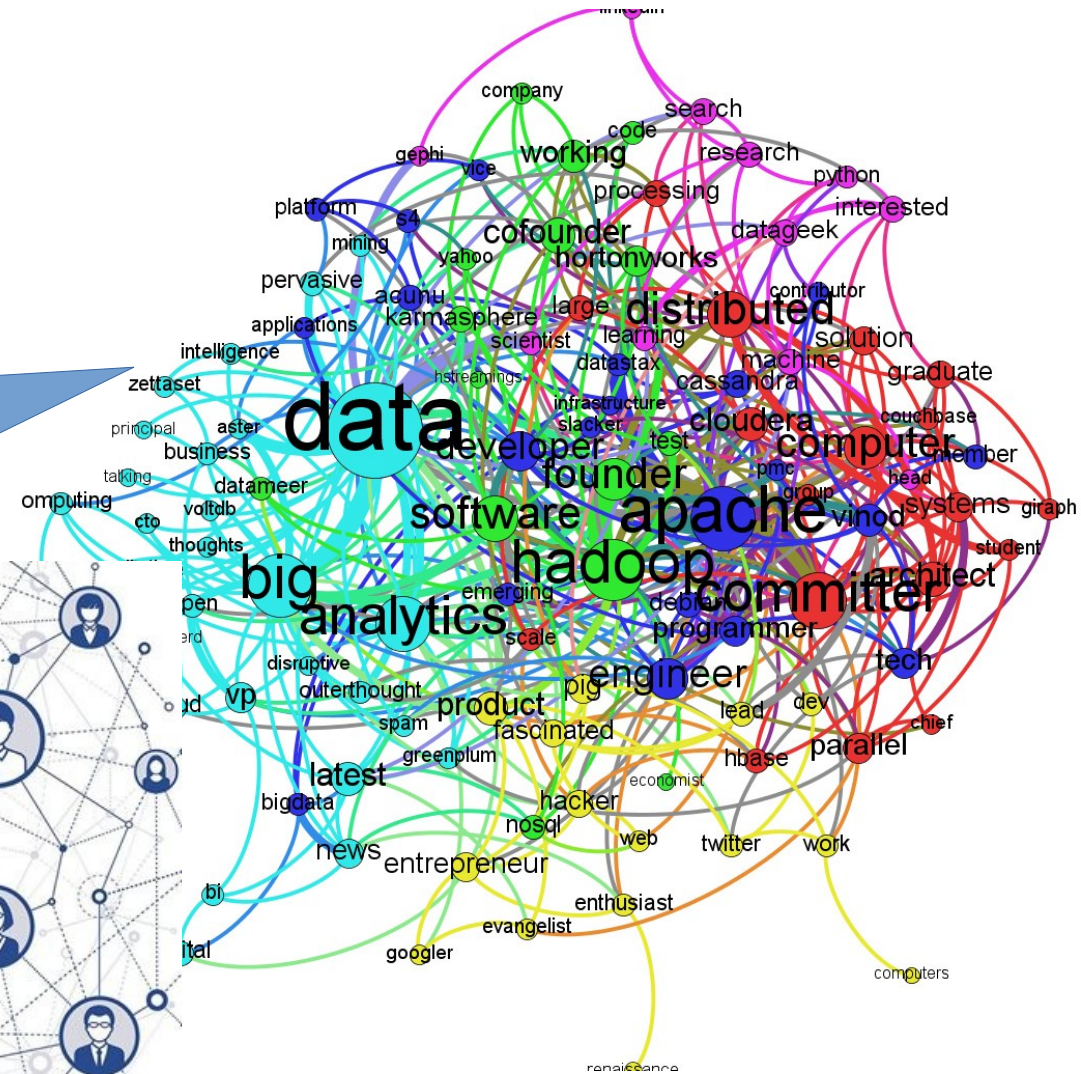
"Social Media Use in 2021"

*By the way: These last slides
visually describe trends ...*

- Graphics have informed us:
 - Which apps are popular
 - Number of people in age groups for social networking sites
 - How much data is created each year, in relation to other years
 - And much more!
- **How did we learn this information to make these previous visualizations?**

***Seriously, where did this
information come from???***

- Algorithms processed seemingly unconnected data to filter out unimportant material.





How Do We Know?

- The previous graphs came to us via raw Big Data from sites like Google, Facebook, Twitter and others.
- **Raw Data:** Seemingly meaningless clutter-like gibberish in which patterns are masked.

*Big data is high-**volume**, high-**velocity** and high-**variety** information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making.*

-- Gartner





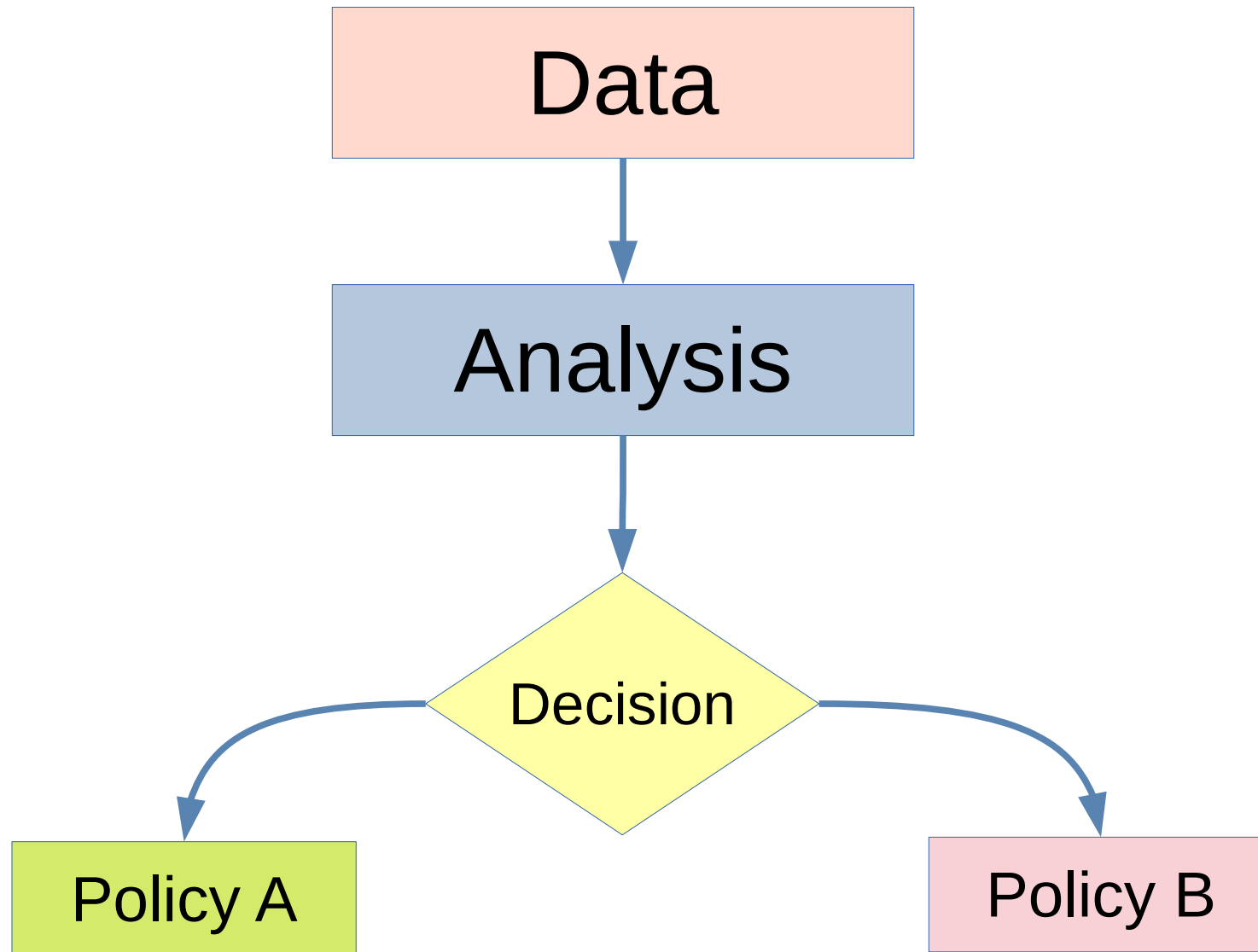
So, It Looks Like We Need **Data** to *Live Intelligently*

- Making *smart* (?) decisions:
 - *Can we make reliable decisions without **data**?*
 - *Is the quality of our society diminished by bad or missing **data**?*
 - *How can we improve commerce, trade without knowledge from **data**?*
 - *How can we make better health decisions without knowledge from **data**?*
- You could give surveys to gather ideas from people but few are likely to respond...

***But, when was the last time
YOU took a survey?***



Policy Creation by Analytics



Thus, Much Interest in Data Analytics

- The present and future are information-driven
- Some of the decisions made after studying trends in a population
 - **Commerce:** what have customers already bought?
 - **Media:** What themes of films, music make money?
 - **Industry:** What products should we make to build, satisfy a market? Which market?
 - **Life Sciences and Medicine:** Reasons for sickness? Bad types of foods? Exposures to toxins?





Your Career Could Be Here!

- *“Big Data & Analytics Is The Most Wanted Expertise By 75% Of IoT (Internet of Things) Providers”*
 - <https://www.forbes.com/sites/louiscolumbus/2017/08/21/big-data-analytics-is-the-most-wanted-expertise-by-75-of-iot-providers/#52082a4e5188>
- *“75% of IoT providers are prioritizing big data and analytics expertise in their hiring decisions.”*
 - <http://www.forbes.com/sites/louiscolumbus/2017/08/21/big-data-analytics-is-the-most-wanted-expertise-by-75-of-iot-providers/>
- *“68% of vendors developing IoT solutions are struggling to find and recruit employees with development expertise.”*
 - <http://www.forbes.com/sites/louiscolumbus/2017/08/21/big-data-analytics-is-the-most-wanted-expertise-by-75-of-iot-providers/>




Glassdoor Informs of Careers

Prodigy Game 4.3★
Game Data Analyst

Job Company Rating

Bonus Points For:

- Degree in Engineering, Computer Science, Stats, Mathematics
- Demonstrated ability to solve hard mathematical, algorithmic, and statistical problems
- Expertise in advanced game analytics - user segmentation, player modelling
- Knowledge of Python/R languages
- Experience working with cloud platforms like AWS (Redshift, Athena, S3)
- Experience working with A/B testing and experimentation
- Significant accomplishments that required both technical and strategic capabilities, such as research projects, open source software contributions, and entrepreneurship




Prodigy Game
Game Data Analyst
Oakville

4.3★

ADURO 3.9★
Data Analyst

Job Company Rating Salary Reviews Why W

- Demonstrated ability to maintain absolute confidentiality
- Proficiency in BI Tools (Sisense, Tableau, PowerPivot, DOMO or other comparable tools)
- Aptitude with SQL, R, and other languages supporting data analysis
- Experience with C based object-oriented programming
- Experience working within a large reporting Data Warehouse
- Experience working with web and application analytics tools (Firebase, Google Analytics)
- Experience using analytics to support product development
- Familiar with source-control repositories and associated practices (Git, GitHub)
- Proven attention to detail and accuracy



ADURO
Data Analyst
Redmond, WA
\$46K-\$78K (Glassdoor Est.) ⓘ

3.9★

- **An Analytics Expert**
 - To apply data analysis skills to help development teams better understand users by applying analytics
 - Find and integrate data from multiple sources to provide analysis
 - Develop tools & methods to ensure data accuracy
 - Collaborate with Data & Analytics team members
 - R skills

In the News

TECH · EVERYDAY A.I.

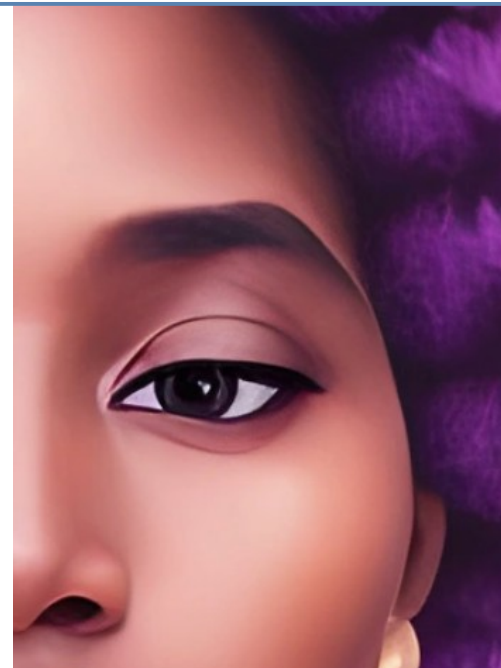
A closer look at the A.I. apps taking over social media

BY JOHN KELL

January 17, 2023 at 3:00 PM EST



A.I. image generators have proliferated across Silicon Valley and have gone viral across social media.



In Production: Nightcafe

The background of the landing page is a close-up, artistic rendering of a woman's face, looking slightly down and to the side. The lighting is soft and warm, highlighting her features. The text is overlaid on this image.

AI Art Generator

Create beautiful art in seconds with the help of
Artificial Intelligence

START CREATING

A family having a picnic on a green lawn near a boat yard



A family having a picnic on a green lawn near a boat yard

