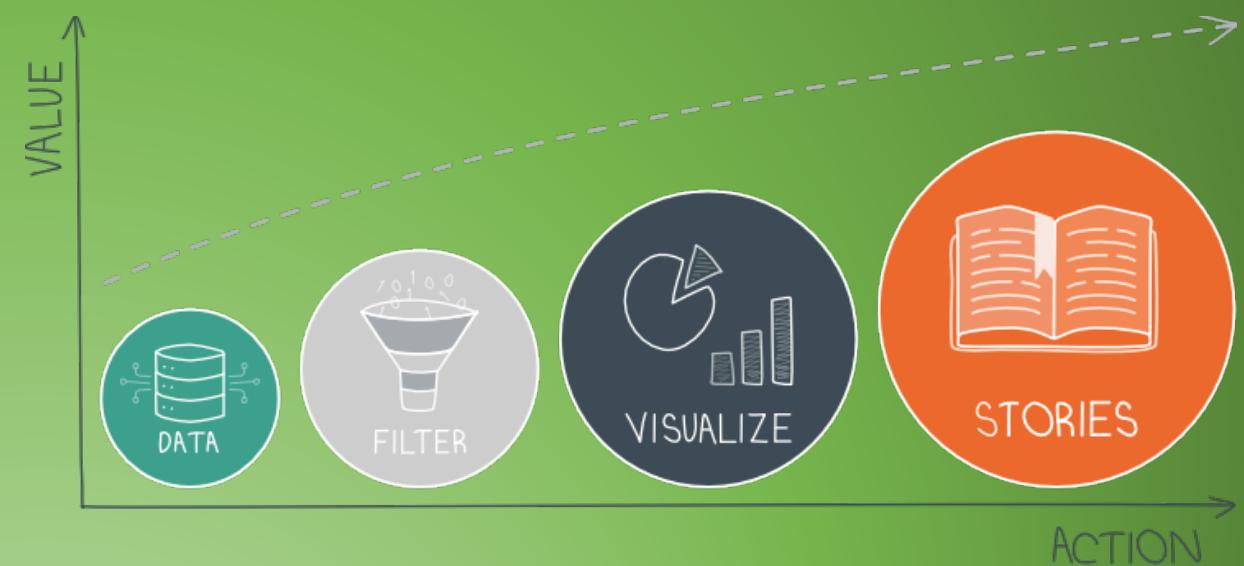


Effective Data Storytelling Using Data Science



Bastin Robins .J
Chief Data Scientist
CleverInsight



With great power comes great responsibility.



Tell data stories &
explain why

Relate experience &
formulate understanding

Act on insight &
drive change

Where are running towards

- Scope
- Data Source
- Prepare Your Data
- Discovering Insight
- Share your story



Questions to ask before you start on your story

1. Who are you writing for?
2. What are you trying to communicate?
3. Why is this important?

What are symptoms?
How much was your last BP
and Cholesterol?
What medicine you took?



Past
Descriptive Phase

Your Blood pressure is now x.
Cholesterol level today is Y.



Present
Diagnostic Phase

Based on past and present records
Chances of BP and Cholesterol hikes
are possible in Upcoming days.

A B C
A B C D
A B C D I
A B C D I A B C D I
A B C D I A B C D I



Future
Predictive Phase

Your Blood pressure is now x.
Cholesterol level today is Y.



Action
Prescriptive Phase

ID	Name	Age	Gender	Survived
1	Jack	19	M	No
2	Rose	18	F	Yes
3	Mike	40	M	No
4	Mary	32	F	Yes
5	Linda	28	F	Yes
6	George	40	M	?
7	Kim	32	F	?

Predicted Outcome: 70% Accuracy

SUPERVISED LEARNING

ID	Name	Age	Gender	Kids	Survived
1	Jack	19	M	0	No
2	Rose	18	F	1	Yes
3	Mike	40	M	0	No
4	Mary	32	F	2	Yes
5	Linda	28	F	2	Yes
6	George	40	M	0	?
7	Kim	32	F	3	?

IF GENDER = 'F' AND AGEBETWEEN = 18 to 25:
KIDS = 1

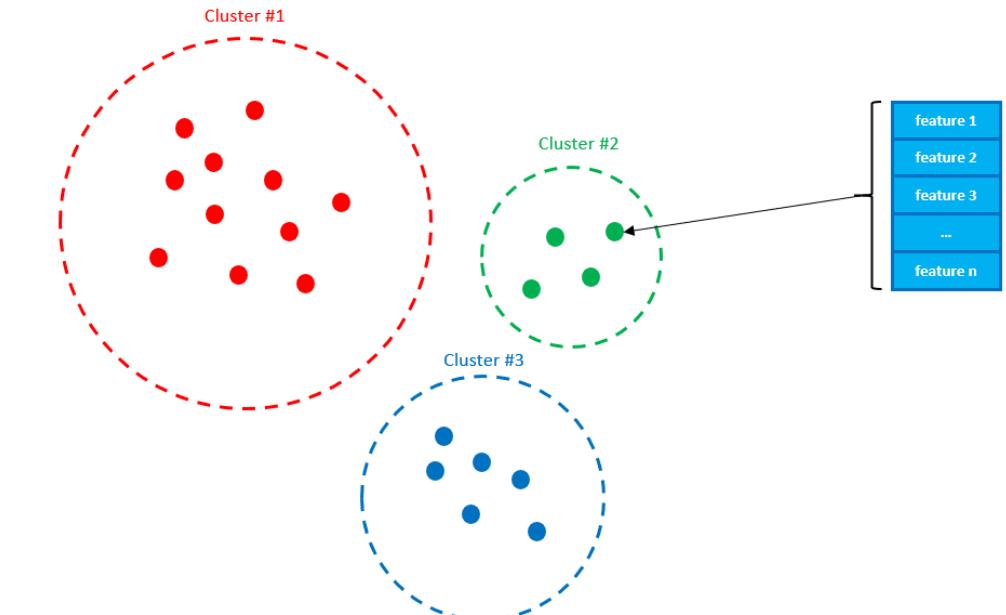
IF GENDER = 'F' AND AGEBETWEEN = 25 to 35:
KIDS = 2

IF GENDER = 'F' AND AGEBETWEEN = 35 to 45+:
KIDS = 3+

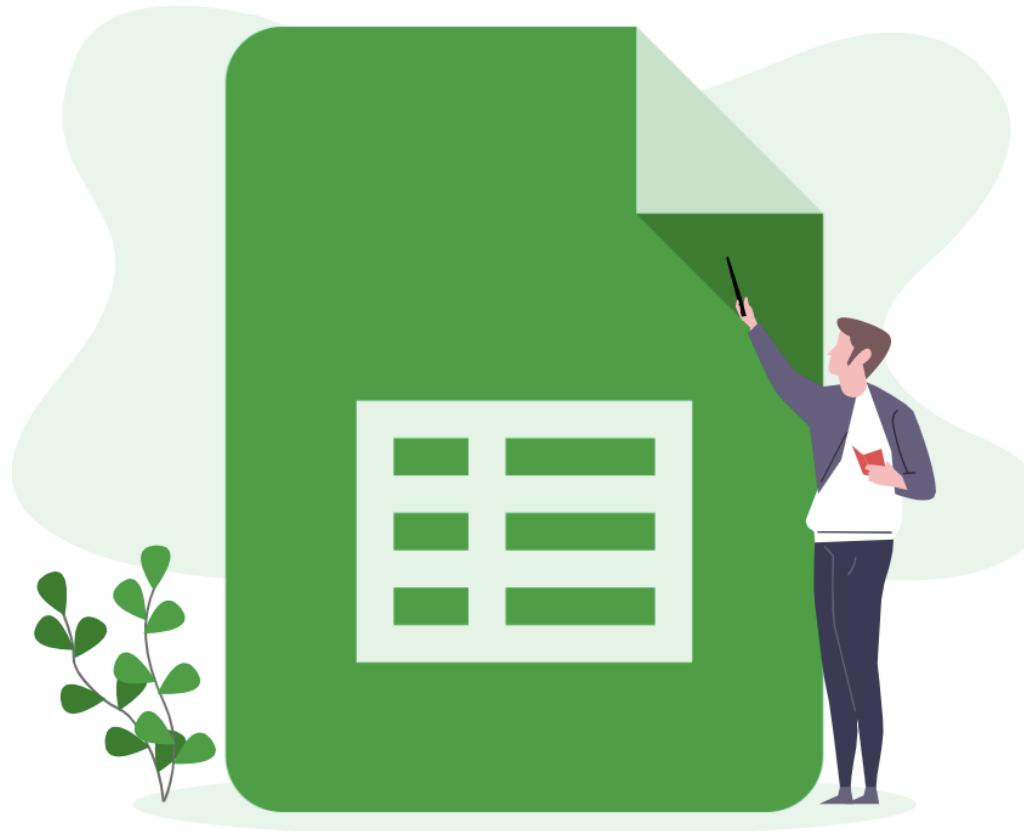
Predicted Outcome: 90% Accuracy

UNSUPERVISED LEARNING

1	Jack	19	M	0	No
2	Rose	18	F	1	Yes
3	Mike	40	M	0	No
4	Mary	32	F	2	Yes
5	Linda	28	F	2	Yes
6	George	40	M	0	?
7	Kim	32	F	3	?

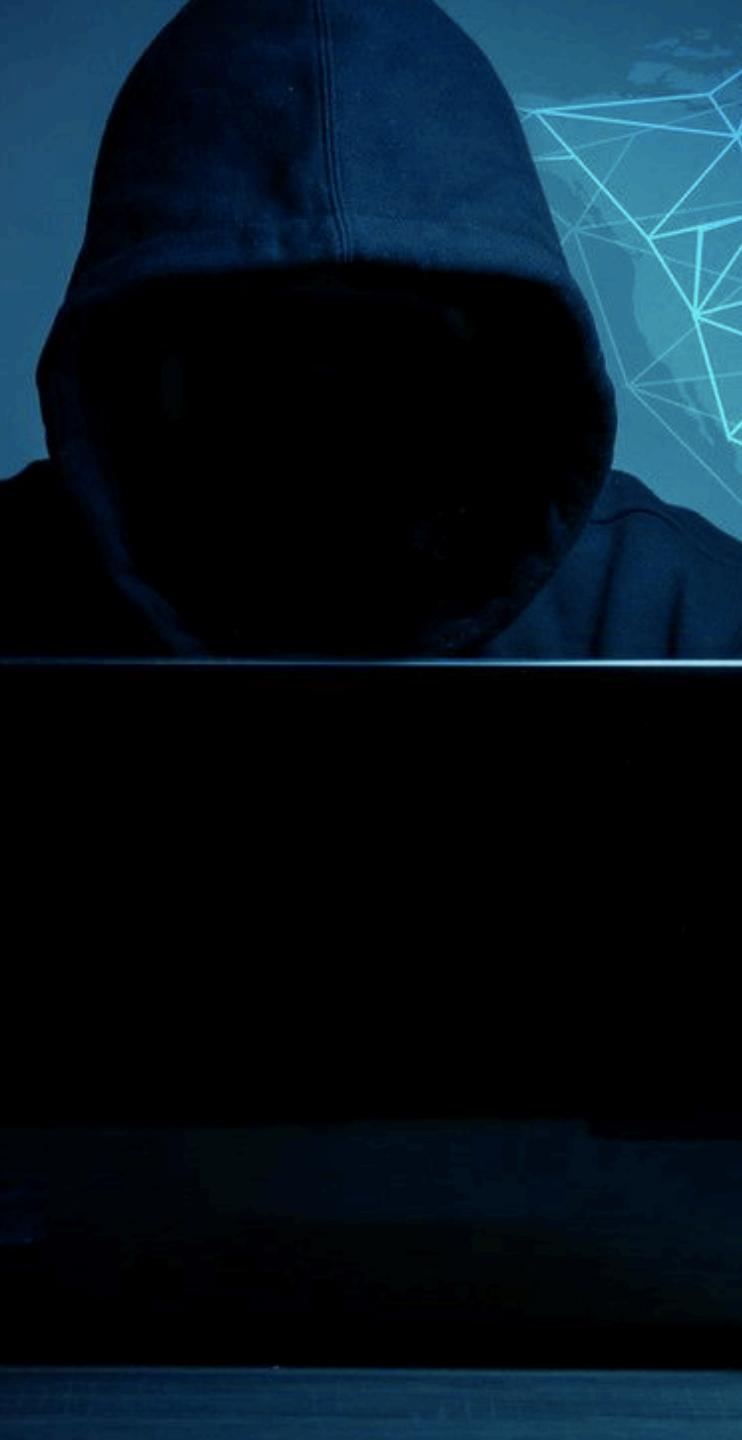






**Tools are important,
But Technique is Everything in
Data Science.**

Enough of Theory ;)

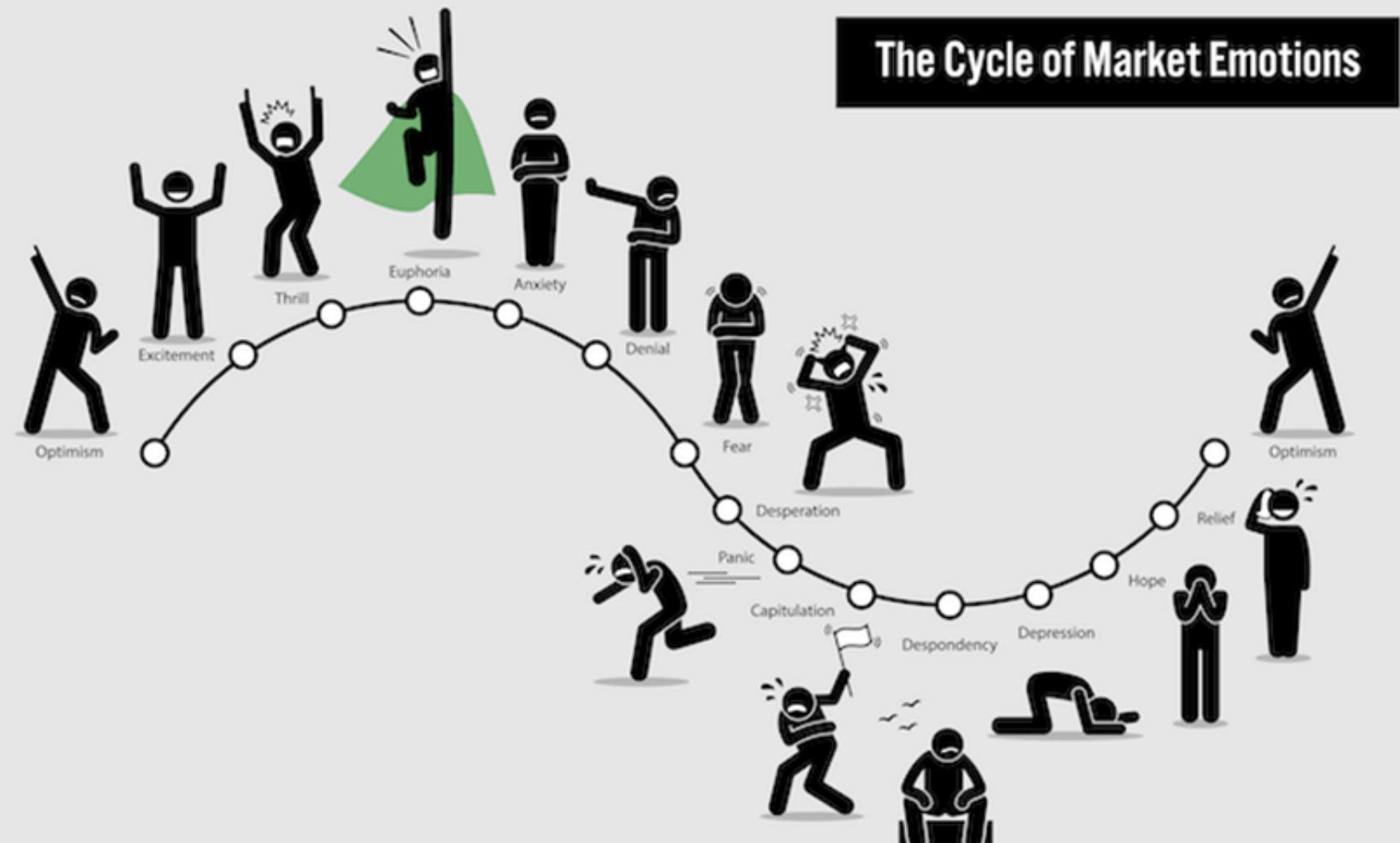


Can we see how people double-deal
Services ;)



Stock Signal

Oakmont is growing exponentially, we need to maintain multiple investment Portfolio, in present day scenario lots of sentiment's plays a major role in trend.



THE ECONOMIC TIMES

MONTDAY, 4 OCTOBER 2018

WWW.ECONOMICTIMES.COM

NBFCs Trying Everything Out is Risky; JM Fin MD
BRANDS & COMPANIES

Kavanaugh Sworn in as US Supreme Court Judge
AROUND THE WORLD

Rethinking Home-and-Away Format can Save Test Cricket
SPORTS: THE GREAT GAMES

Diwali Dhoom may Elude Mkts Amid Global Gloom



Rupee likely to slide further, decline in stocks to continue on rising crude and US rates

Satish Das & Suresh Marichandani

A Dull Diwali

How much more will the Rupee fall before Diwali? (Our final week)

ET POLL

ET Prime

ET NOW

Greenko Acquiring Orange for \$850m

Greenko is acquiring Orange Renewable from Singapore's AT Capital Group for seven times its pre-money value, report Arjit Barman. The company is also in talks to buy Shriram Renewable Energy, a private company held by Suresh Founder Tulsit Patel and his family, for \$25 billion. [++ 17](#)

ET NOW

Sentiment Analysis

Understanding the different types of customers and segmenting them into few categories based on their likelihood, helps me to recommend right products to similar user groups.





Taylor Winkel @twinkelnews

Seattle Native. Reddit junkie. Yoga. Ira Glass. Freelancer. News Desk Assistant @King5Seattle. Troller @Crosscut. Frmr @nwNewsNetwork. twinkel@king5.com



Alex Garland @AGarlandPhoto

Freelance photographer looking for stories of people, tall trees and fast moving rivers. Tweet me for photo work. More @ thedignityvirus.com



Dan Savage @fakedansavage

I write Savage Love and host the weekly Savage Lovecast. Opinions my own, RTs mandatory.



gene balk @genebalk

FYI Guy data-journalist & researcher @SeattleTimes. Librarian (MLS Rutgers), cycling advocate, sometimes soul DJ. RTs are not endorsements, opinions are my own.



Jim Brunner @Jim_Brunner

Seattle Times political reporter. Sloppy guitarist. Baseball geek. VP Pacific NW Newspaper Guild. jbrunner@seattletimes.com



Taylor Soper @Taylor_Soper

@UW '12 | @GeekWire reporter | Rip City | Golf | Sushi | Tom Hanks | Tweeting tech, journalism, athletics, music, food

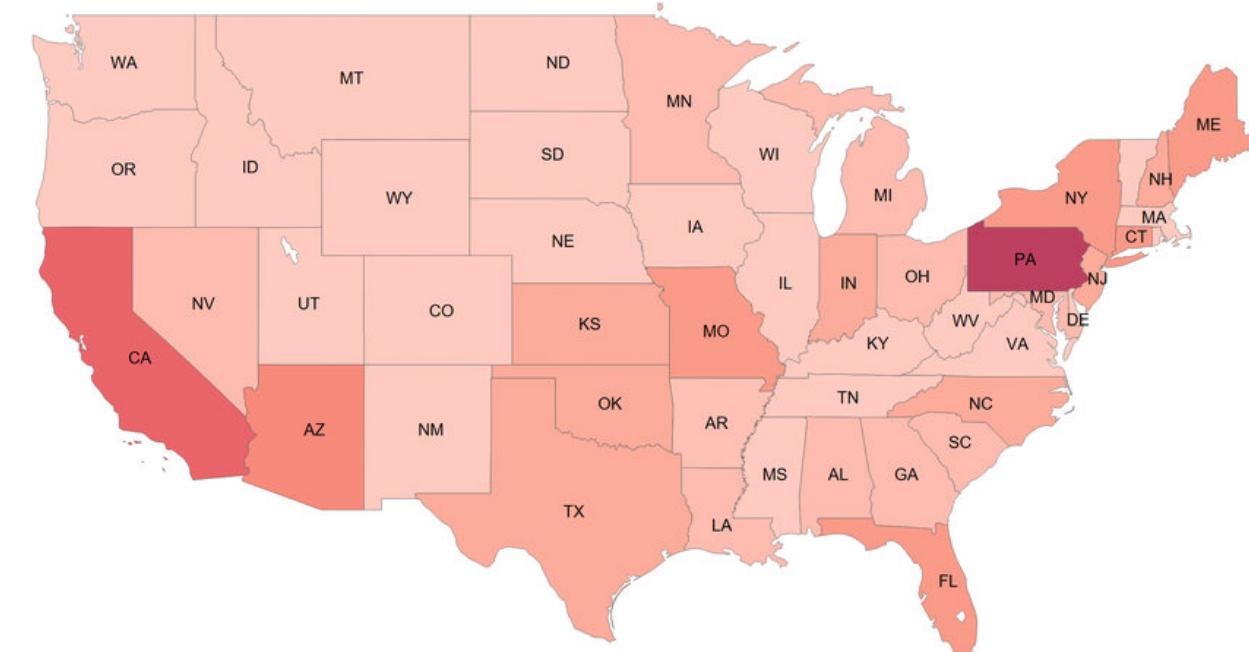


Matt Day @mattmday

Business reporter minding Microsoft and technology matters for the @SeattleTimes. Ex-@WSJ. mdav@seattletimes.com.

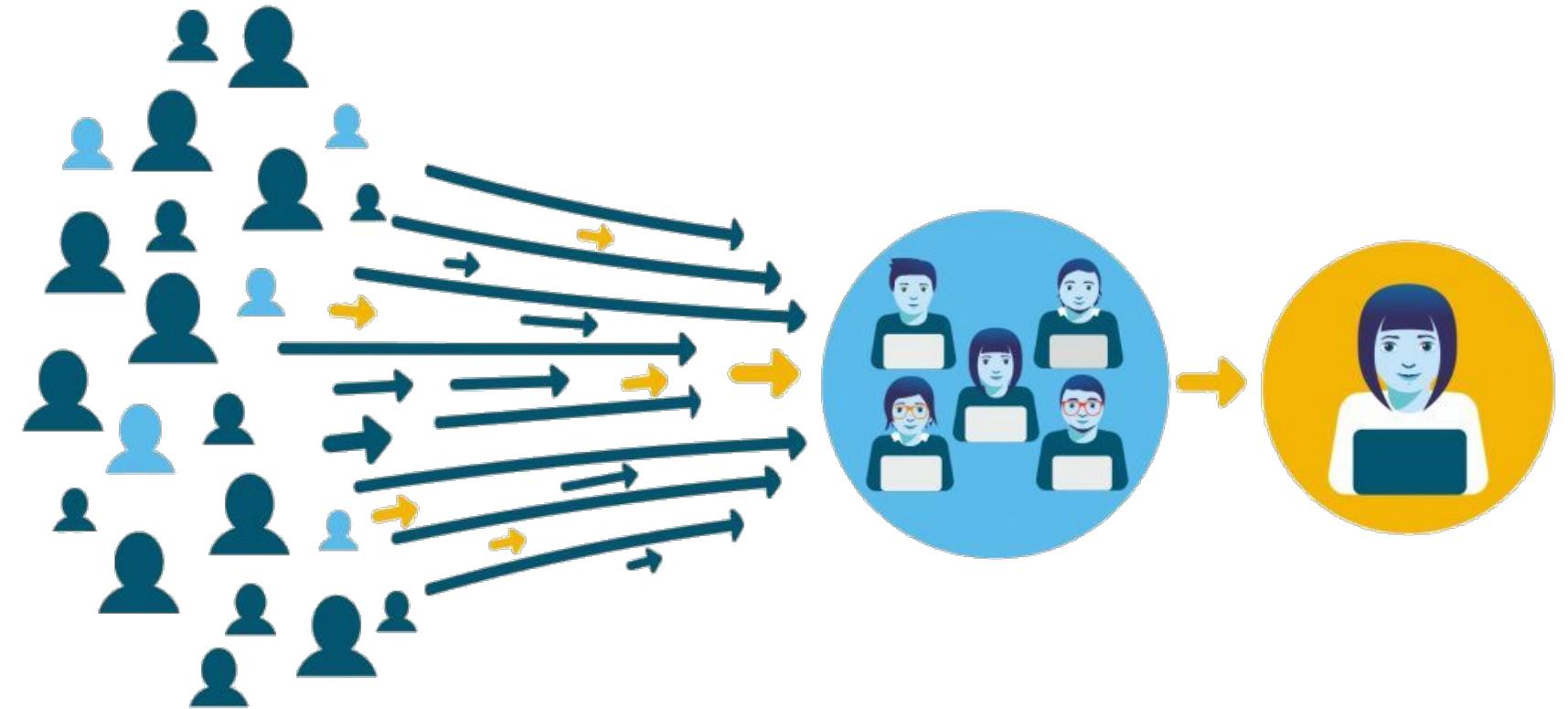


Sentiment Across Geographies



HR Resume Scoring

Oakmont is growing exponentially, we are in need of best team.
But the problem is we received more than 1000 application per day.
Its impossible for my HR team to call each and every person.





LOREM IPSUM
New York, New York lorem ipsum.com
EDUCATION
LOREM IPSUM 2010-2012
LOREM IPSUM 2010-2012
SIT AMET
LOREM ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. Ut wisi enim ad minim veniam, quis nostrud exercitatio nonnullam ut laoreet dolore magna aliquam erat volutpat.
EXPERIENCE
LOREM IPSUM DOLOR 2010-2013
PROFESSIONAL SKILLS
LOREM IPSUM DOLOR LOREM IPSUM DOLOR DOLOR DOLOR
PERSONAL SKILLS
LOREM IPSUM DOLOR SIT AMET
Languages
ENGLISH SPANISH GERMAN



Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. Sed ut perspiciatis unde omnis iste natus error sit voluptatem accusantium doloremque laudantium, totam rem aperiam, eaque ipsa quae ab illo inventore veritatis et quasi architecto beatae vitae



Frequency Count

Keyword	Score	Count
Python	5	4
C++	5	4
MS Excel	2	2
Pivot	3	1
Neural Network	3.5	1
CNN	5	1
TOTAL	23.5	



Keywords	Score
C++	5
Python	5
Programming	2
MS Excel	2
MS Powerpoint	2
English Proficiency	5
Communication Skill	5
Machine Learning	3
Neural Networks	3.5
Data Analytics	3
Regression Analysis	4
CNN	5
ANN	5
Pivot Table	3

User ID	Keyword	Score	Personality	Interests
Mike	Python, C++, CNN	23.5	Social Buff	Politics, Sports
Jack	Python, C++, CNN	22	Quiet	Art
Kevin	Python	18	Extrovert	Politics, Public

Now sort column of choice we get the Top candidates from a pile of resumes.

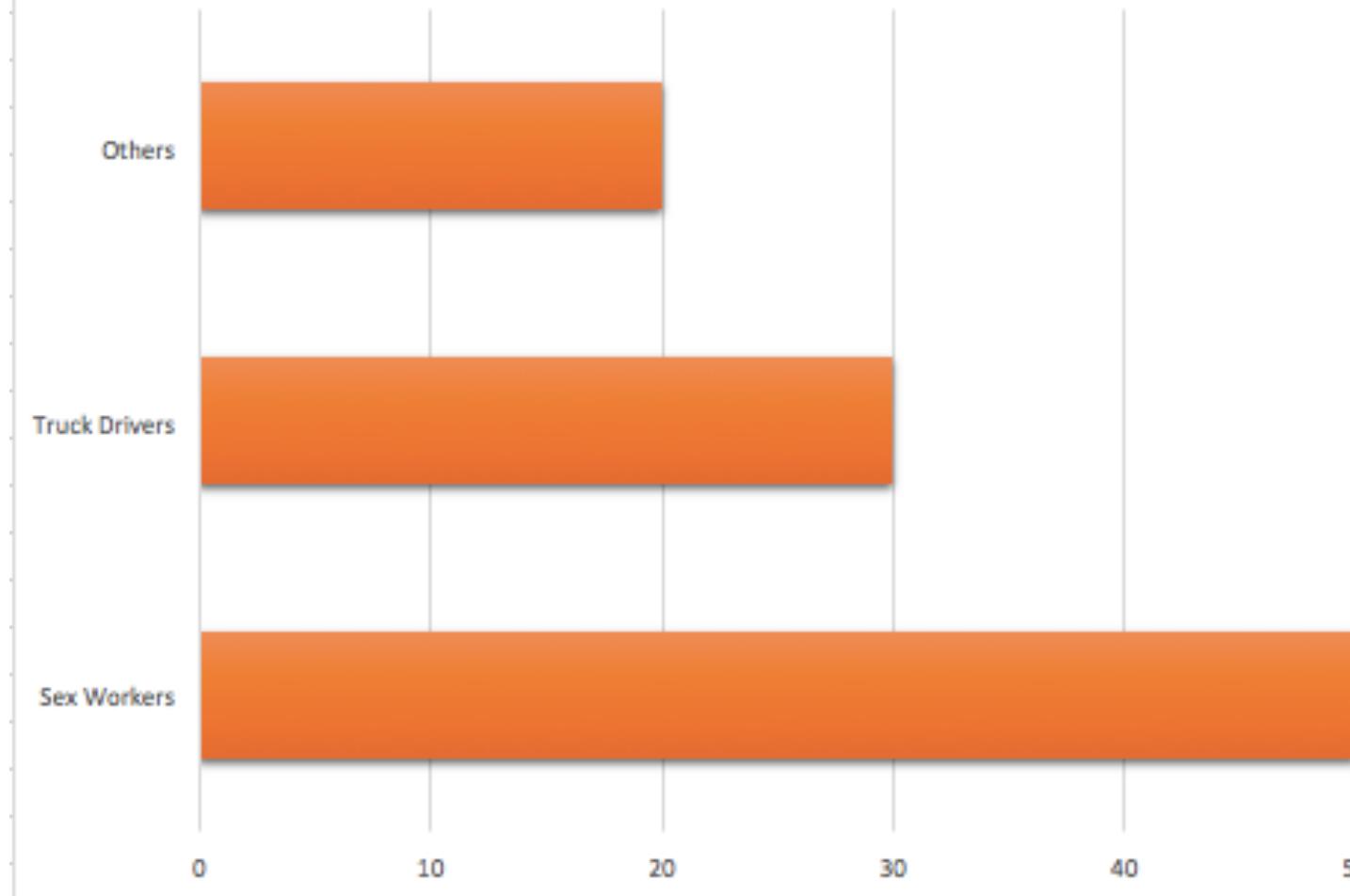


Data In Action

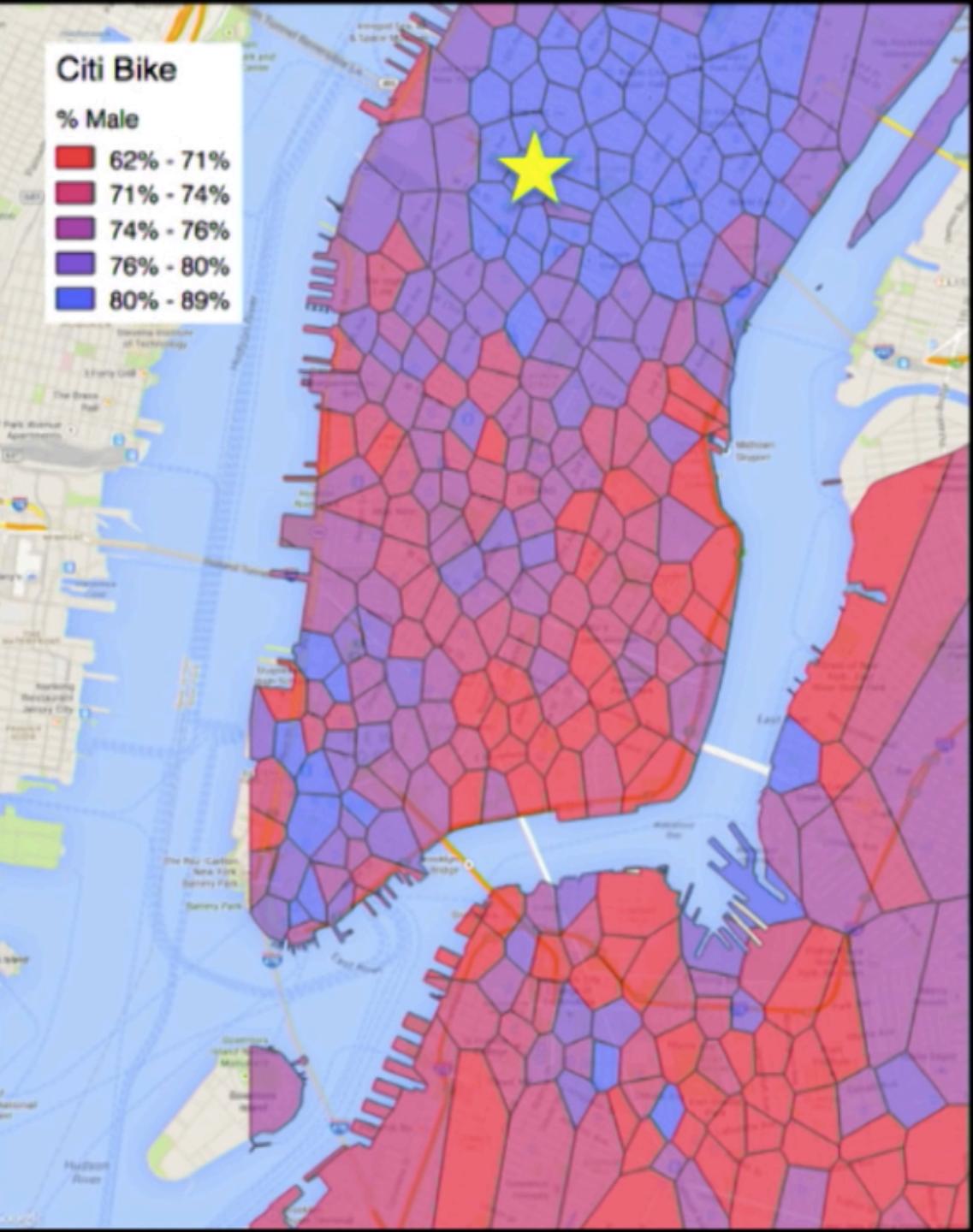
**After Implementation of GST,
HIV AIDS transmission started decline.**



AIDS TTransmission



Citi Bikes





Can we see what people search in Google.

What Google Search Tells About India.



Why is _____ so ?

Can we automate data narratives?



Deployment	pypl v0.0.1b3.dev0
Build Status	build passing
Metrics	coverage 28%
GitHub	contributors 8 Stars 6
License	License BSD
Community	chat on glitter

We are in beta

Cognito is an exclusive data preprocessing library that involves transforming raw data into a machine-understandable format. We at CleverInsight Open Ai Foundation took the initiative to build a better automated data preprocessing library and here it is.

Contributors



Collaborators



CHRIST (Deemed to be University) Lavasa provides developer & mentor contribution via CleverInsight Open AI Foundation Incubation Center at Christ University Lavasa since 2019.

<https://pypi.org/project/cognito/>

Become a contributor

<https://github.com/Cleverinsight/cognito/issues>

Join Cognito Community

<https://gitter.im/cognito-dev/community>

**Ask the right question, Data will
confess.**

In God We Trust; all others must bring DATA.

- *William Edwards Deming* -

Business Outcomes > ML,DL & NLP

Leave you Feedback

<https://bit.ly/37AXngH>



Time For Questions

Thank You

CleverInsight

CleverInsight
Open AI Foundation

149 Commonwealth Drive
Menlo Park, CA 94025
+1-408-888-4507



Bastin Robins .J

Chief Data Scientist
[@CleverInsight](https://twitter.com/CleverInsight)
robin@cleverinsight.co
<http://linkedin.com/in/bastinrobin>