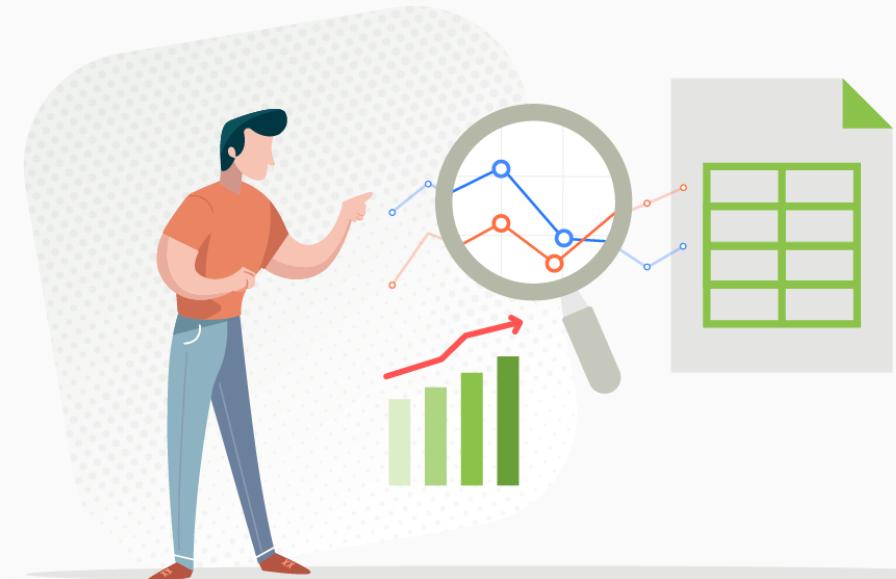


Hands-On Foundational Data Science Bootcamp

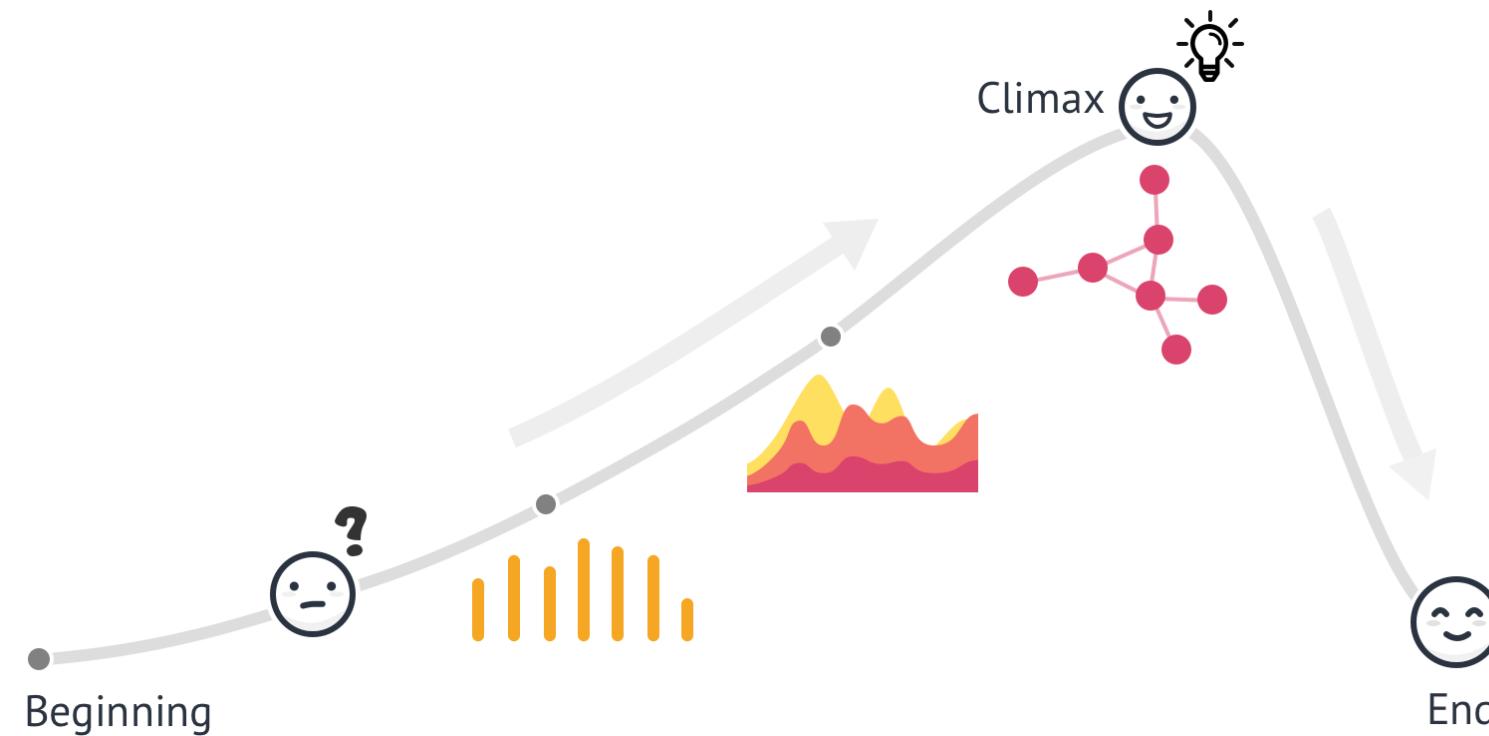


Instructor
Bastin Robins .J

Chief Data Scientist



- Intro to Data Science
- Analogy – Learn to Love Data
- How to define problem statement
- Break -
- Basic of Python Programming
- Understanding Essential Packages

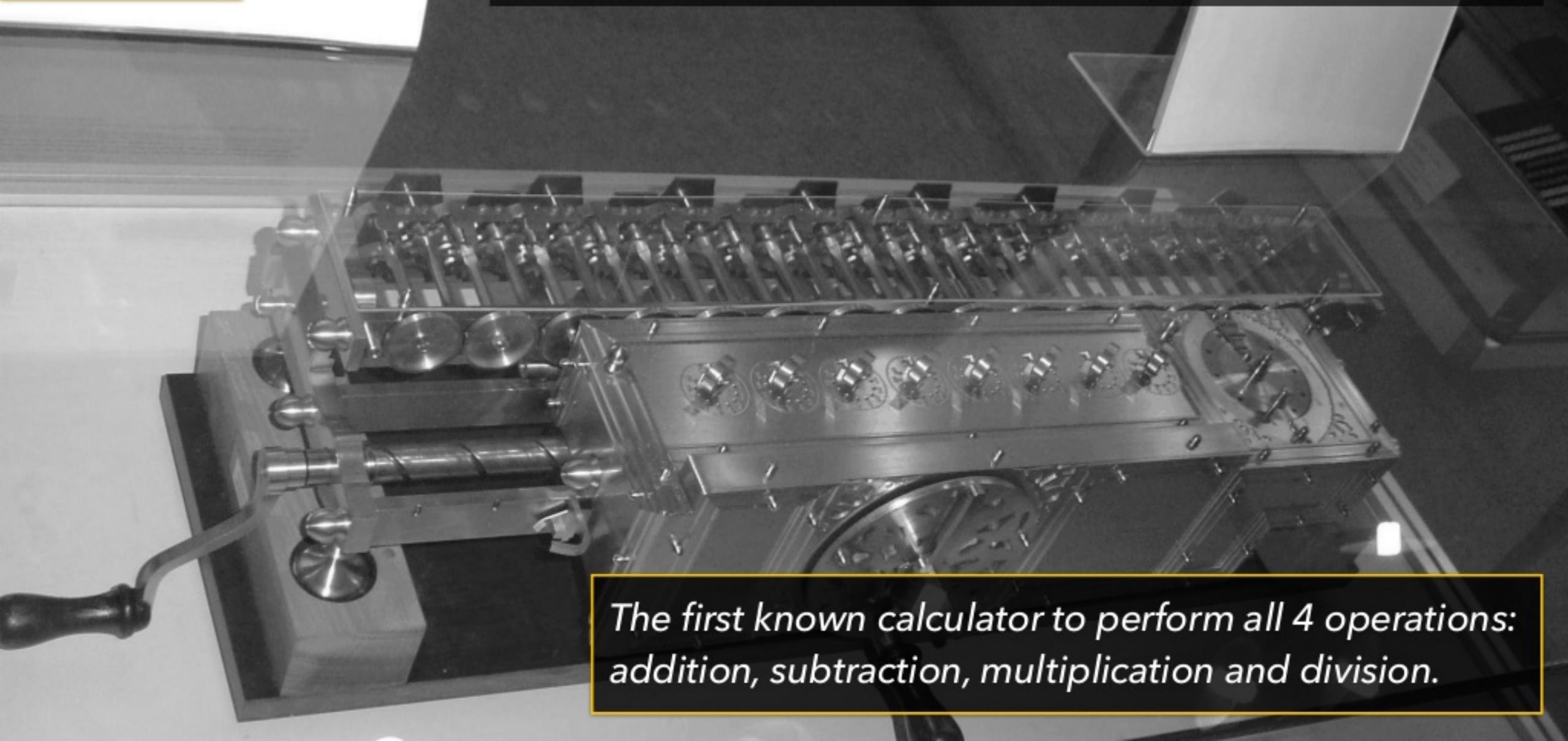


Data will talk to you
if you're willing to listen.

What?

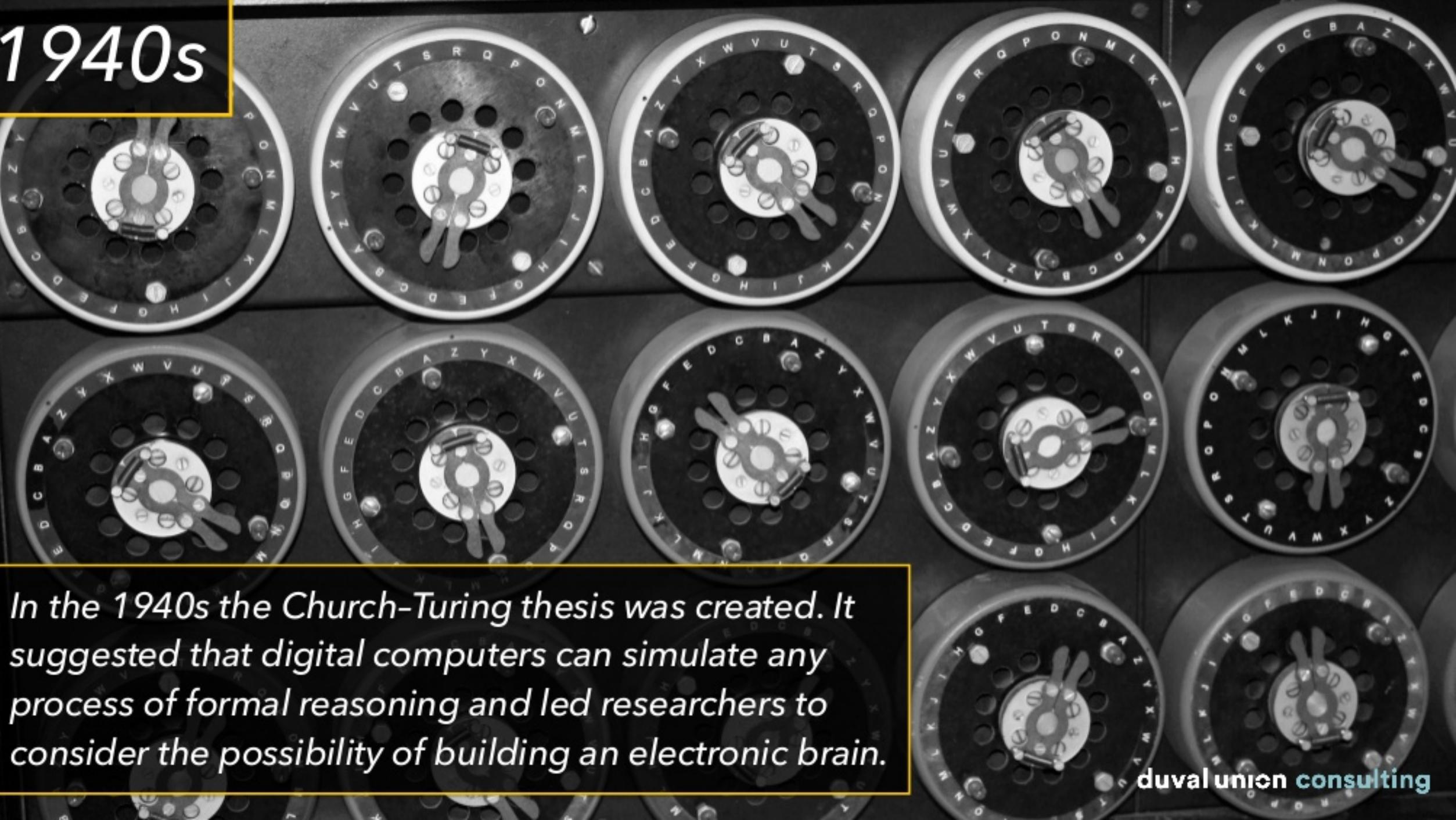
1672

First AI: THE STAFFELWALZE



*The first known calculator to perform all 4 operations:
addition, subtraction, multiplication and division.*

1940s



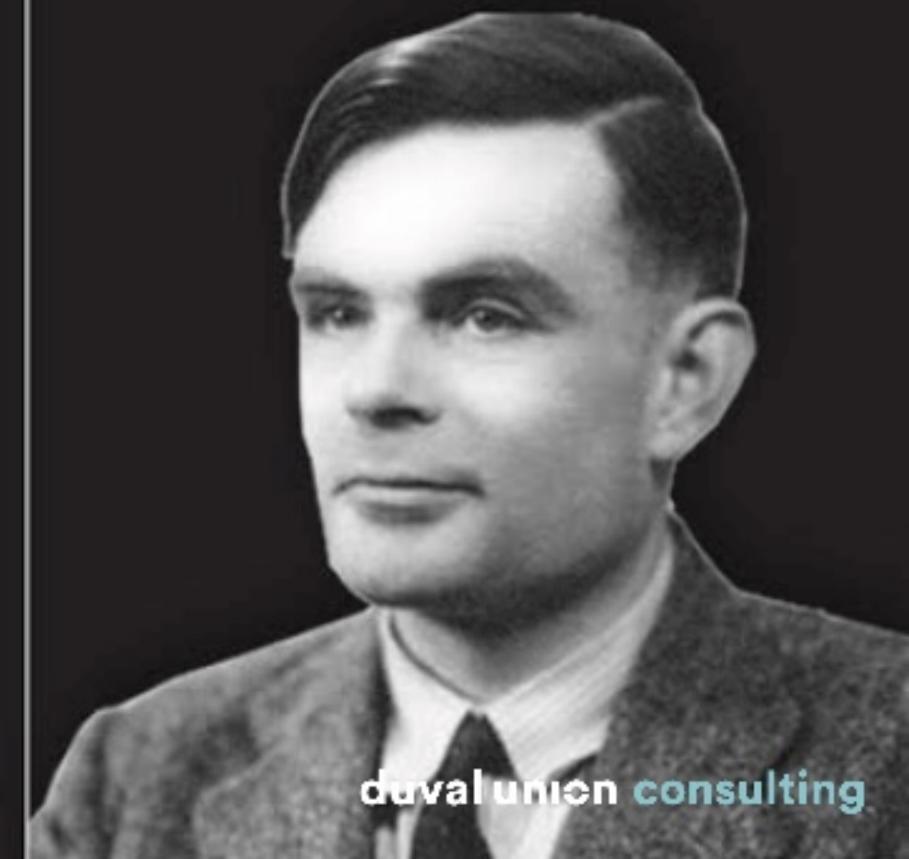
In the 1940s the Church-Turing thesis was created. It suggested that digital computers can simulate any process of formal reasoning and led researchers to consider the possibility of building an electronic brain.

1950s



AI started getting developed by philosophers and mathematicians in the 19th century. Alan Turing is one of the best known founding fathers of modern AI.

Turing test



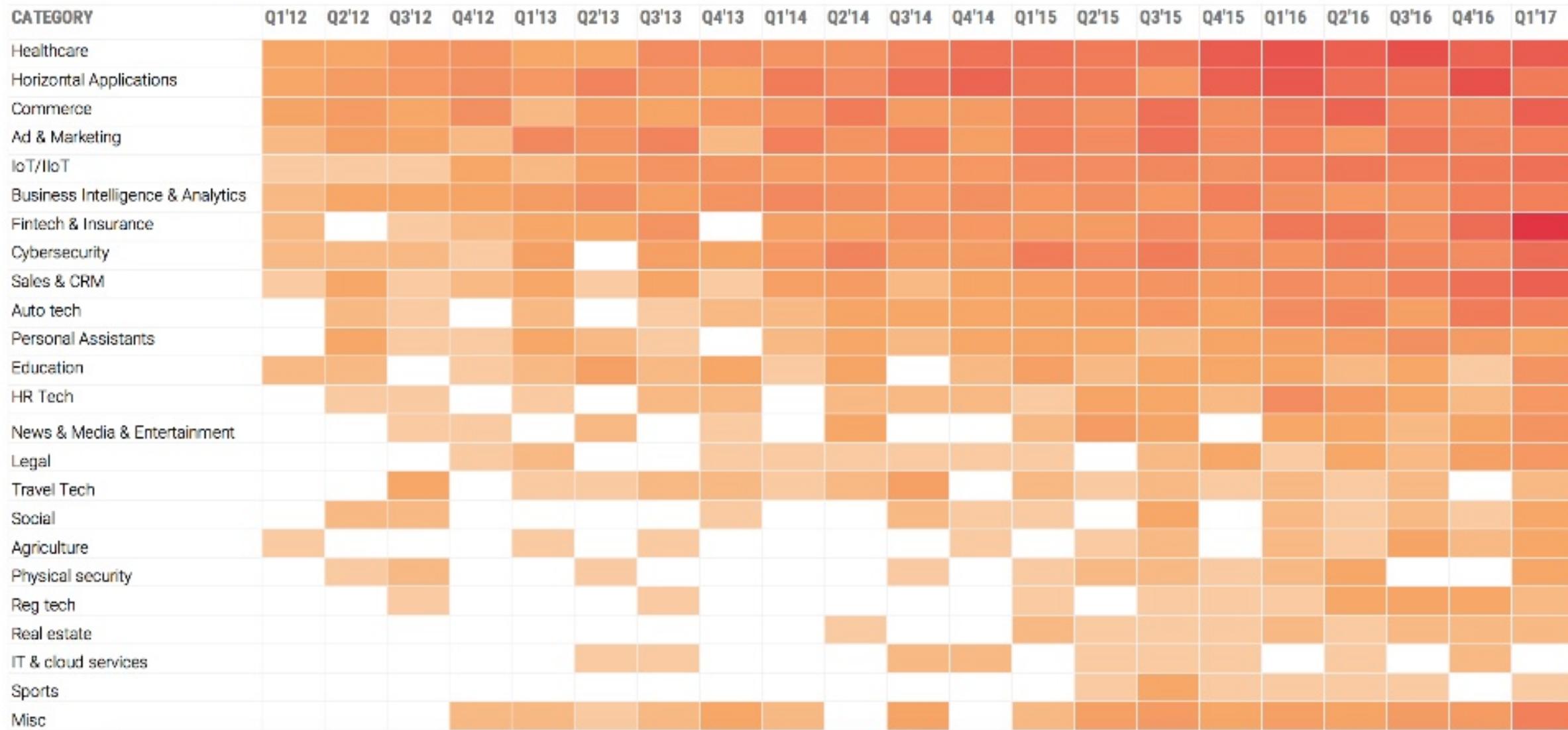
Why?



AI HEATMAP: DEALS DISTRIBUTION BY CATEGORY

Q1'12-Q1'17 (as of 3/23/17)

Across **ALL SECTORS**



How?

ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



MACHINE LEARNING

Machine learning begins to flourish.



DEEP LEARNING

Deep learning breakthroughs drive AI boom.



1950's

1960's

1970's

1980's

1990's

2000's

2010's

Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

Learn to ❤️ Data

Because the real hero is DATA

Do these four cities looks identical to you ?

Take a look at the sales report alongside. A company has branches in 4 cities, and each branch changes the product price every month. This leads to a corresponding change in the sales.

Here is the performance of the **4 branches** with their monthly price and sales for each month.

Looking at the average, the four branches have an **identical performance**.

DO YOU AGREE?

2010	Boston		Chicago		Detroit		New York	
Month	Price	Sales	Price	Sales	Price	Sales	Price	Sales
Jan	10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
Feb	8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
Mar	13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
Apr	9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
May	11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
Jun	14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
Jul	6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
Aug	4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
Sep	12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
Oct	7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
Nov	5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89
Average	9.0	7.50	9.0	7.50	9.0	7.50	9.0	7.50
Variance	10.0	3.75	10.0	3.75	10.0	3.75	10.0	3.75

Average price is the same. Average sales is the same too.
 Variance in price is the same. So is the variance in sales.

Are they really identical? Check Again..

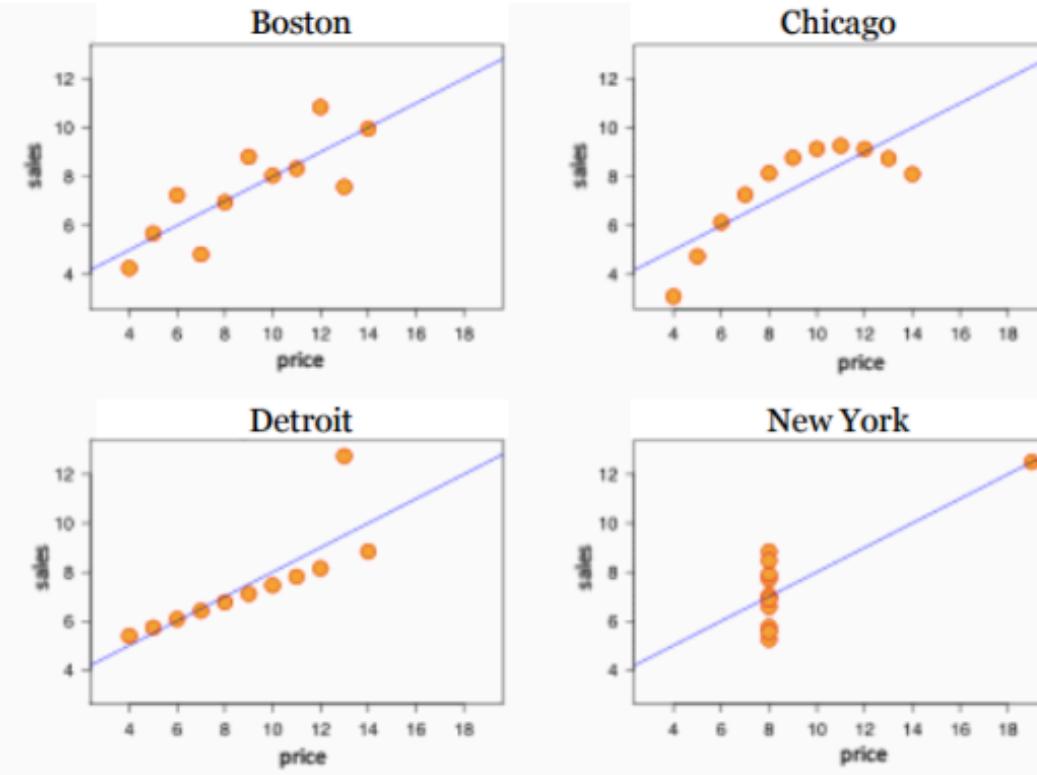
But in fact, the four cities are totally different in behaviour.

Boston's sales has generally increased with price.

Detroit has a nearly perfect increase in sales with price, except for one aberration.

Chicago shows a decline in sales beyond a price of 10.

New York's sales fluctuates despite a nearly constant price.





Can we see how Indian weather
Was Since 1902 till 2002?





4,650,000
pieces of
data



Which
variables
matter?



Organising
+
clustering



Analysis
of
data



What variables
contribute to a solution?

Results



Understand



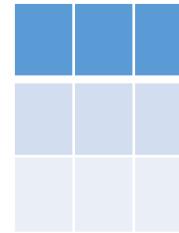
Improve



Learn

Data

Tabular



Text



Audio

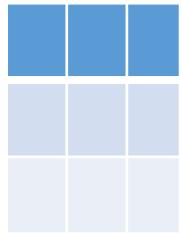


Video



Data > Information > Knowledge > Wisdom

Tabular



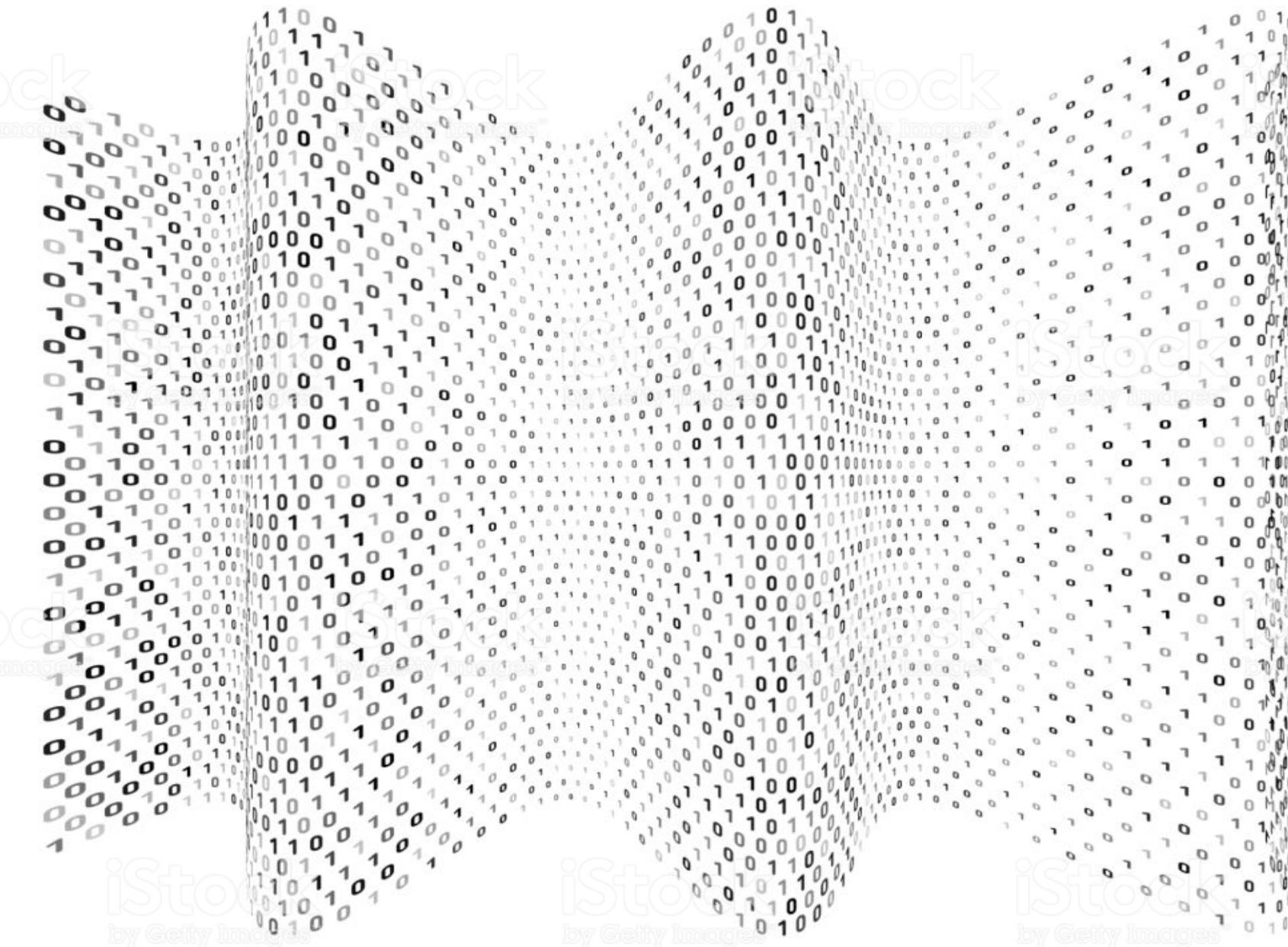
Text

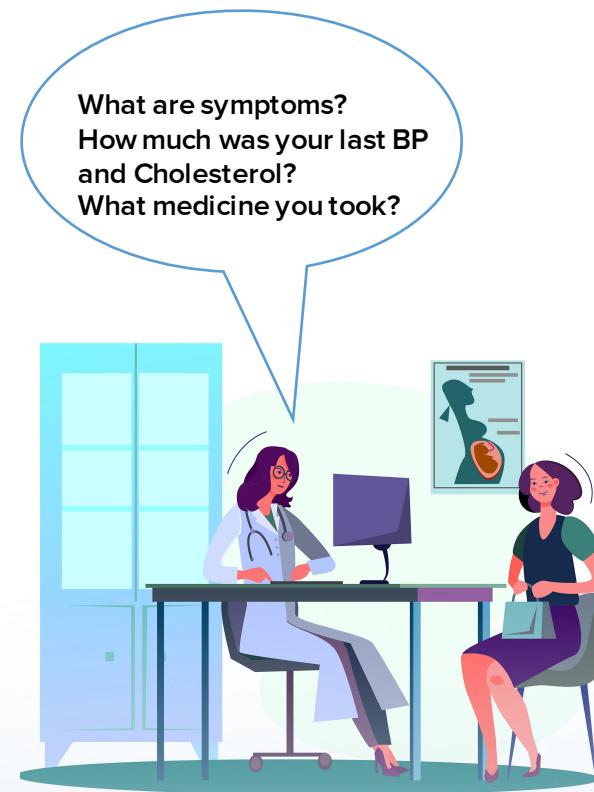


Audio

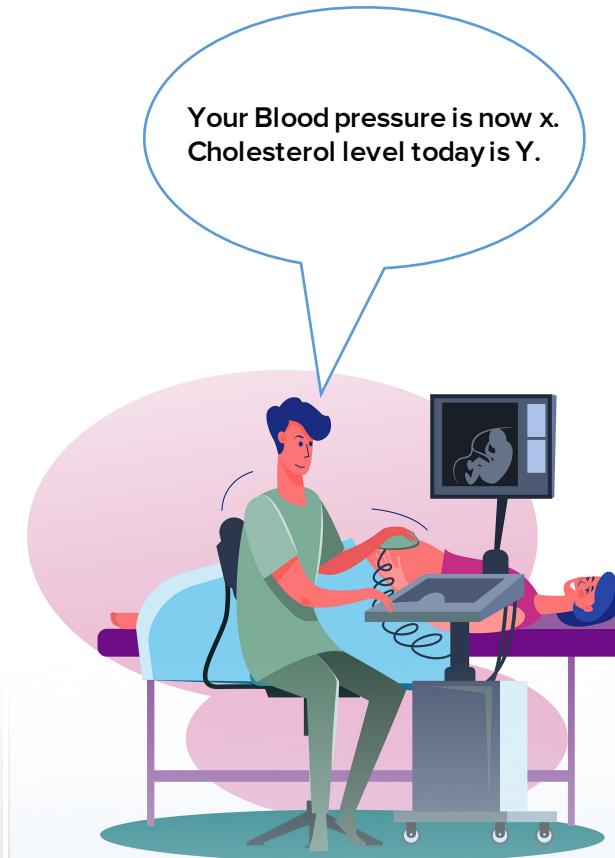


Video

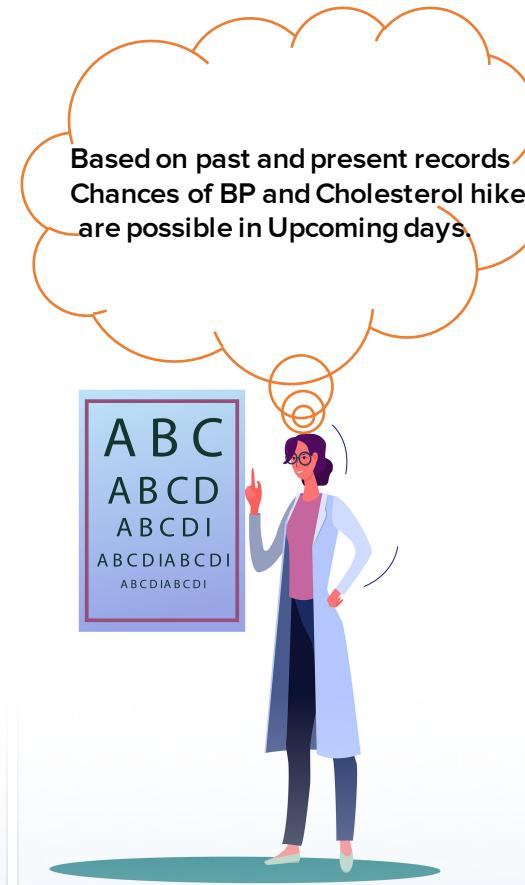




Past
Descriptive Phase
What happened



Present
Diagnostic Phase
Why it happened



Future
Predictive Phase
What will happen



Action
Prescriptive Phase
What action to take

Tom wants to play

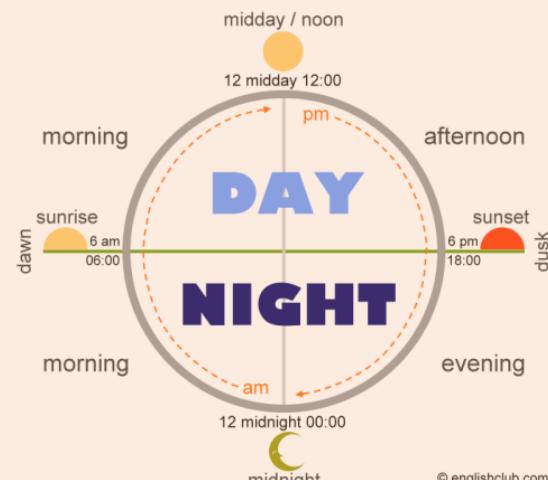
*But **Mumma** says **Tom**, outside its **rainy**,
see its also **very cold** and **wind** is **very
high**. So don't go to **play today**.*



How Tom Understand
Basics



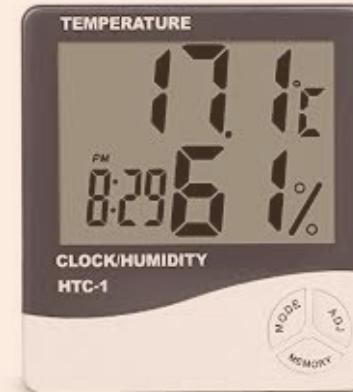
Climates



Time



Windy



Temperature



Play

Features

Target



Let see how
Tom Learns

Day	Climate	Temp	Wind	Time	Play
1	Sunny	40	2.5	8:00 AM	YES
2	Rainy	10	8.3	8:00 PM	NO
3	Sunny	35	4	8:00 PM	YES
4	Spring	25	4.6	8:00 PM	YES
5	Rainy	23.5	0.1	8:00 AM	YES
6	Rainy	13.3	6.6	8:00 PM	NO
7	Rainy	12.1	4.6	8:00 PM	NO
8	Rainy	11	8	8:00 PM	?

Identifier Categorical Continuous Continuous Categorical Categorical

Identifier

- Registration ID
- Name
- UID
- Aadhaar ID
- PAN ID

Discrete

- Age
- No of Students
- Total Subjects
- Total Questions
- Total Failed
- Total Passed

Categorical

- Gender
- Class
- Colors
- Types
- Age Group
- Educational Level

Continuous

- Height, Weight
- Temperature
- Humidity
- Cost
- Time
- Length

Example

- 10820830010
- 8008-80901-189
- Jack Smith

Example

- 20
- 30
- 22
- 10

Example

- Male / Female
- I,II,III, IV
- Red, Yellow, Green
- Rich, Middle, Poor
- 10-15, 15-20, 20-30
- Kid, Adult, Older
- Bsc, MSc, BE, B.tech

Example

- 118.7
- 20.4
- 5.5
- \$100
- 24 hrs
- 500 cm

Categorical

Aka Classification Problem

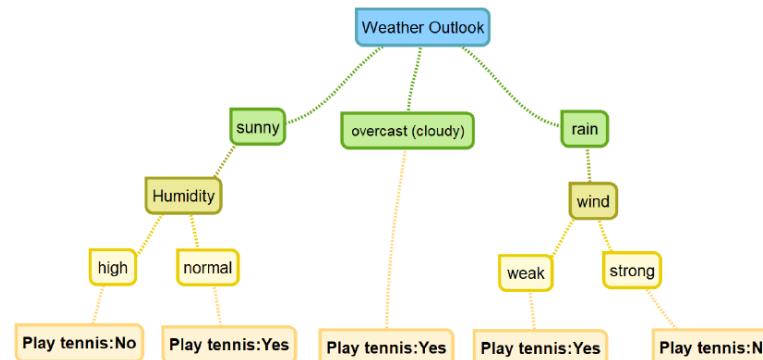


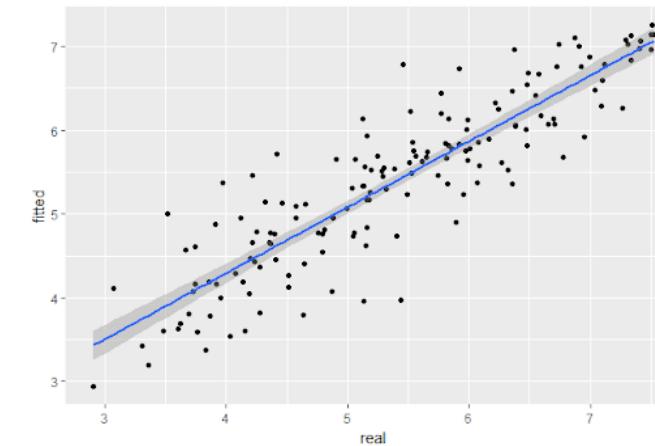
Figure 1: Play Tennis Decision tree

How to Measure Accuracy

- Sensitivity / Recall
- Specificity
- Precision
- F1 Score
- ROC
- Confusion Matrix

Continuous

Aka Regression Problem



How to Measure Accuracy

- Squared Error
- Mean Squared Error - MSE
- Root Mean Squared Error – RMSE
- R2
- Absolute Error
- Mean Absolute Error - MAE

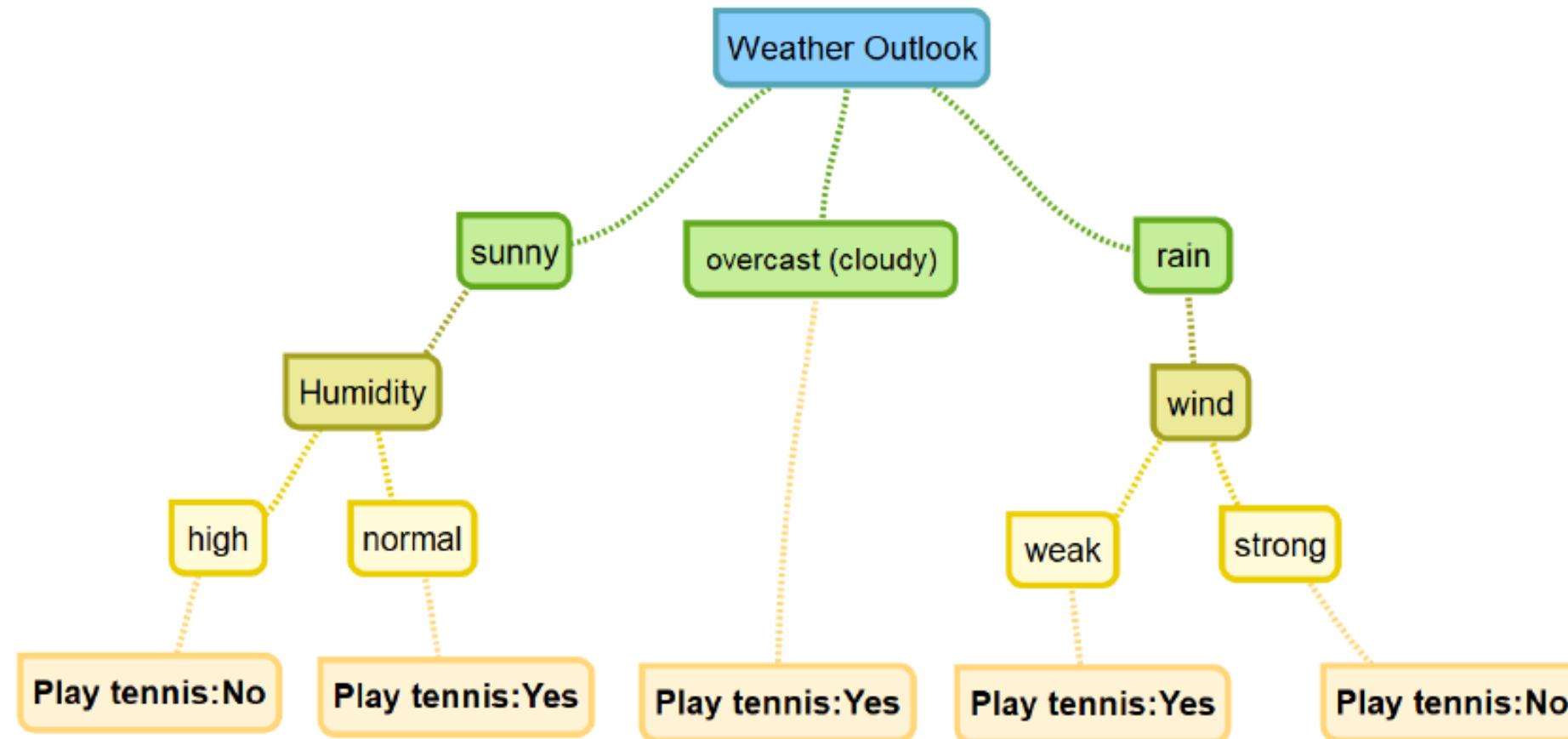
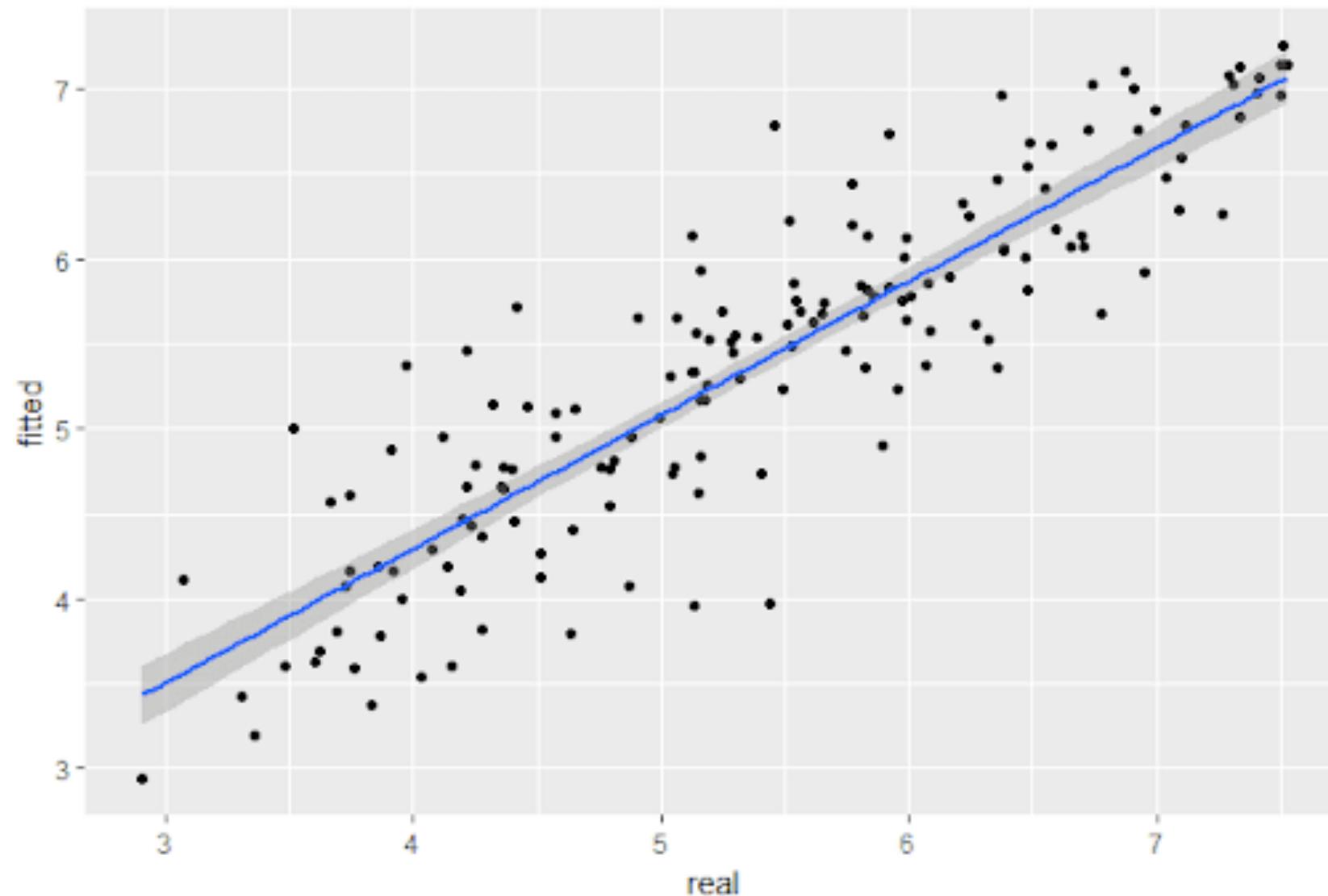


Figure 1: Play Tennis Decision tree



Day	Climate	Temp	Wind	Time	Play
1	Sunny	40	2.5	8:00 AM	YES
2	Rainy	10	8.3	8:00 PM	NO
3	Sunny	35	4	8:00 PM	YES
4	Spring	25	4.6	8:00 PM	YES
5	Rainy	23.5	0.1	8:00 AM	YES
6	Rainy	13.3	6.6	8:00 PM	NO
7	Rainy	12.1	4.6	8:00 PM	NO
8	Rainy	11	8	8:00 PM	YES
9	Sunny	40	2.5	8:00 AM	YES
10	Rainy	10	8.3	8:00 PM	NO
11	Sunny	35	4	8:00 PM	YES
12	Spring	25	4.6	8:00 PM	YES
13	Rainy	23.5	0.1	8:00 AM	NO
14	Rainy	12.1	4.6	8:00 PM	NO
15	Rainy	11	8	8:00 PM	NO
15	Rainy	11	8	8:00 PM	NO
15	Rainy	11	8	8:00 PM	NO

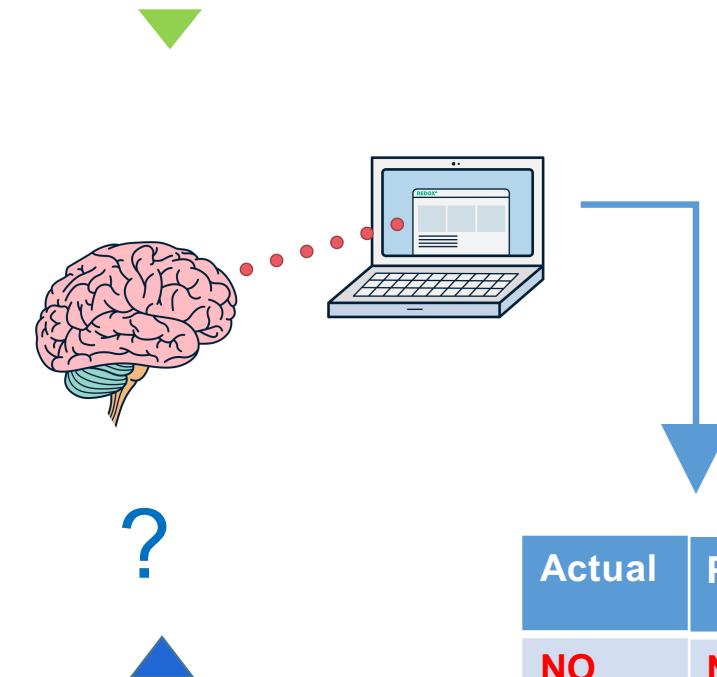
Total Yes: 8 | Total No: 9

70% Training

30% Testing

Day	Climate	Temp	Wind	Time	Play
1	Sunny	40	2.5	8:00 AM	YES
2	Rainy	10	8.3	8:00 PM	NO
6	Rainy	13.3	6.6	8:00 PM	NO
4	Spring	25	4.6	8:00 PM	YES
15	Rainy	11	8	8:00 PM	NO
3	Sunny	35	4	8:00 PM	YES
7	Rainy	12.1	4.6	8:00 PM	NO
8	Rainy	11	8	8:00 PM	YES
9	Sunny	40	2.5	8:00 AM	YES
10	Rainy	10	8.3	8:00 PM	NO
11	Sunny	35	4	8:00 PM	YES
14	Rainy	12.1	4.6	8:00 PM	
13	Rainy	23.5	0.1	8:00 AM	
12	Spring	25	4.6	8:00 PM	
15	Rainy	11	8	8:00 PM	
5	Rainy	23.5	0.1	8:00 AM	
15	Rainy	11	8	8:00 PM	

70% Training



30% Testing

95% Accurate Prediction

Actual	Predicted
NO	NO
NO	NO
YES	YES
NO	YES
YES	YES
NO	NO

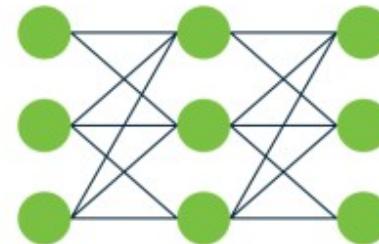
Machine Learning



Input



Feature extraction



Classification

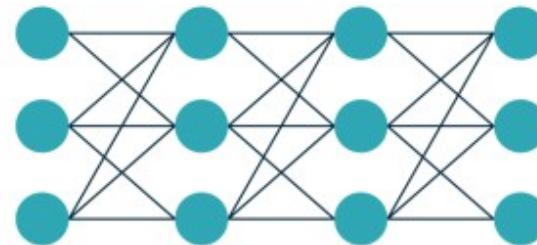
CAR
NOT CAR

Output

Deep Learning



Input



Feature extraction + Classification

CAR
NOT CAR

Output

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	
1	state	district	190101	190102	190103	190104	190105	190106	190107	190108	190109	190110	190111	190112	190201	190202	190203	190204	
2	Andaman and Nicobar Islands	North and Nicobar	2.661	2.729	2.654	2.693	2.674	2.594	2.538	2.545	2.462	2.585	2.603	2.559	2.605	2.646	2.754	2.84	2.67
3	Andaman and Nicobar Islands	Nicobar	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Andhra Pradesh	Anantapur	29.997	31.566	33.95	36.37	34.42	31.079	29.025	28.81	30.484	29.742	28.284	26.972	28.766	30.846	34.951	36.114	35.94
5	Andhra Pradesh	Chittoor	29.402	31.147	32.452	35.078	35.357	33.404	31.594	31.054	31.571	30.026	28.017	26.502	27.841	29.53	33.278	35.1	36.2
6	Andhra Pradesh	Adilabad	30.153	31.691	35.712	39.037	40.116	36.711	31.248	29.753	32.067	32.375	29.658	28.369	30.422	32.48	37.363	39.804	42.2
7	Andhra Pradesh	Karimnagar	29.783	31.112	34.74	38.018	38.625	35.538	30.704	29.503	31.638	31.588	29.175	27.865	29.561	31.465	36.268	38.452	41.0
8	Andhra Pradesh	Nizamabad	29.843	31.101	34.871	37.958	38.171	34.557	30.037	28.989	31.226	31.391	28.789	27.839	29.722	31.792	36.394	38.583	40.5
9	Andhra Pradesh	Hyderabad	30.208	31.485	34.939	37.985	37.131	34	30.654	29.931	31.831	31.308	29.185	28.209	29.586	31.462	36.239	37.985	40.1
10	Andhra Pradesh	Medak	29.796	30.982	34.663	37.698	37.293	33.906	30.024	29.203	31.328	31.112	28.836	27.865	29.368	31.332	36.144	37.886	40.1
11	Andhra Pradesh	Mahbubnagar	30.961	32.558	35.683	38.764	37.682	34.525	31.407	30.881	32.517	31.784	29.827	28.759	30.216	32.925	38.691	40.2	
12	Andhra Pradesh	Rangareddi	30.283	31.533	35.094	38.149	37.27	34.093	30.644	29.992	31.87	31.417	29.277	28.324	29.702	36.437	38.127	40.1	
13	Andhra Pradesh	Warangal	30.112	31.329	34.648	37.835	38.378	35.724	31.349	30.405	32.208	31.767	29.633	28.245	31.433	35.84	38.005	40.6	
14	Andhra Pradesh	Nalgonda	30.106	31.366	34.338	37.431	37.345	34.764	31.223	30.382	32.066	31.348	29.134	28.005	31.103	35.492	37.538	39.9	
15	Andhra Pradesh	Srikakulam	26.678	27.42	29.735	31.539	32.421	31.929	29.192	28.702	29.49	29.005	27.7	26.044	27.613	30.19	31.315	32.5	
16	Andhra Pradesh	Khammam	30.256	31.269	34.143	37.083	38.12	36.102	32.016	31.19	32.682	31.80	28.522	29.458	31.207	34.946	37.209	39.9	
17	Andhra Pradesh	Visakhapatnam	28.925	29.824	32.443	34.715	36.047	34.215	30.901	30.119	31.505	31.06	27.306	28.071	29.78	32.685	34.9	36.6	
18	Andhra Pradesh	Vizianagar	27.789	28.498	31.175	33.159	34.239	32.749	29.575	28.945	30.005	31.618	27.643	26.14	27.086	28.654	31.595	33.197	34.5
19	Andhra Pradesh	West Godavari	29.644	30.446	32.395	34.941	36.956	36	32.565	31	31.618	29.739	28.279	28.639	29.983	32.908	35.046	38	
20	Andhra Pradesh	East Godava	29.409	30.18	32.393	34.818	36.569	35.316	31.984	32.319	31.312	29.574	27.901	28.402	29.86	32.756	34.937	37	
21	Andhra Pradesh	Guntur	30.531	31.799	34.034	37.054	37.982	36.316	32.219	33.235	31.894	29.724	28.507	29.549	31.099	34.905	37.175	39.8	
22	Andhra Pradesh	Krishna	29.665	30.726	32.774	35.513	37.064	35.59	31.7	32.707	31.352	29.371	28.05	28.734	30.213	33.42	35.605	38.5	
23	Andhra Pradesh	Sri Potti Sriramula	30.393	32.177	33.676	36.694	37.0	33.907	33.086	33.449	31.561	28.942	27.792	28.989	30.661	34.471	36.915	39.1	
24	Andhra Pradesh	Prakasam	31.011	32.721	34.961	38.061	38.305	33.378	32.544	33.375	31.959	29.742	28.579	29.914	31.77	35.934	38.202	40.2	
25	Andhra Pradesh	Y.S.R.	29.873	31.706	33.818	35.59	33.42	31.219	30.627	31.534	30.248	28.226	26.941	28.5	30.556	34.807	36.633	37.5	
26	Andhra Pradesh	Kurnool	31.242	32.962	34.061	37.398	34.203	31.51	31.088	32.472	31.565	29.78	28.611	30.21	32.437	37.019	38.731	39.4	
27	Arunachal Pradesh	Anjaw	19.43	20.6	26.406	27.482	29.077	29.023	29.248	28.569	27.054	23.297	20.308	19.586	20.381	24.413	24.607	28.4	
28	Arunachal Pradesh	Dibang Valley	17.464	18.31	21.689	24.086	24.674	25.89	25.752	25.89	25.744	24.716	21.1	18.559	17.993	18.547	22.388	22.049	25.4
29	Arunachal Pradesh	Tirap	17.464	18.31	21.689	24.086	24.674	25.89	25.752	25.89	25.744	24.716	21.1	18.559	17.993	18.547	22.388	22.049	25.4
30	Arunachal Pradesh	Changlang	19.436	22.154	24.873	25.591	26.903	26.712	27.025	26.678	25.485	21.811	19.134	18.505	19.093	22.955	23.004	26.5	
31	Arunachal Pradesh	Tawang	14.524	17.143	20.65	23.278	23.958	26.097	25.651	25.728	25.02	23.71	19.45	16.081	16.103	17.258	21.555	21.278	24.6
32	Arunachal Pradesh	East Kameng	21.124	23.729	27.351	29.343	29.522	31.01	30.711	30.737	30.312	29.404	25.787	22.158	22.187	23.638	28.238	27.336	30.0
33	Arunachal Pradesh	West Kamer	17.572	20.228	23.733	25.944	26.471	28.211	27.953	27.809	27.376	26.188	22.316	18.885	18.87	20.197	24.648	23.906	27.1
34	Arunachal Pradesh	Kurung Kumey	16.979	18.988	22.215	24.707	25.643	27.547	27.288	27.206	26.668	25.291	21.397	18.105	17.9	18.869	23.015	22.637	26.2
35	Arunachal Pradesh	West Siang	14.715	16.28	19.342	22.273	23.906	26.144	26.254	26.2	25.203	23.349	19.162	15.839	15.327	16.113	19.989	20.179	24.5
36	Arunachal Pradesh	Upper Subansiri	14.084	15.722	18.749	21.647	23.145	25.436	25.289	25.165	24.379	22.588	18.399	15.216	14.921	15.611	19.434	19.591	23.7
37	Arunachal Pradesh	East Siang	19.444	20.93	24.086	26.497	27.719	29.808	29.864	29.906	29.022	27.57	23.765	20.397	19.86	20.726	24.785	24.395	28.3
38	Assam	Sonitpur	22.749	25.4	29.156	30.932	30.725	31.699	31.558	31.655	31.265	30.564	27.149	23.691	23.577	25.289	30.168	28.833	31.1
39	Assam	Nagaon	18.855	21.096	24.584	26.674	26.603	27.296	27.186	27.444	27.25	26.56	23.138	20.179	19.604	20.935	25.646	24.586	26.8
40	Assam	Dhemaji	22.558	24.239	27.716	29.703	30.385	31.97	31.837	32.009	31.34	30.214	26.585	23.232	23.02	24.068	28.493	27.603	31.0
41	Assam	Lakhimpur	22.809	24.933	28.469	30.366	30.752	31.971	31.797	31.932	31.433	30.526	26.967	23.602	23.432	24.792	29.384	28.267	31.2

Curse of dimensionality

Learn to ❤️ Data

Importance of Knowing Data

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

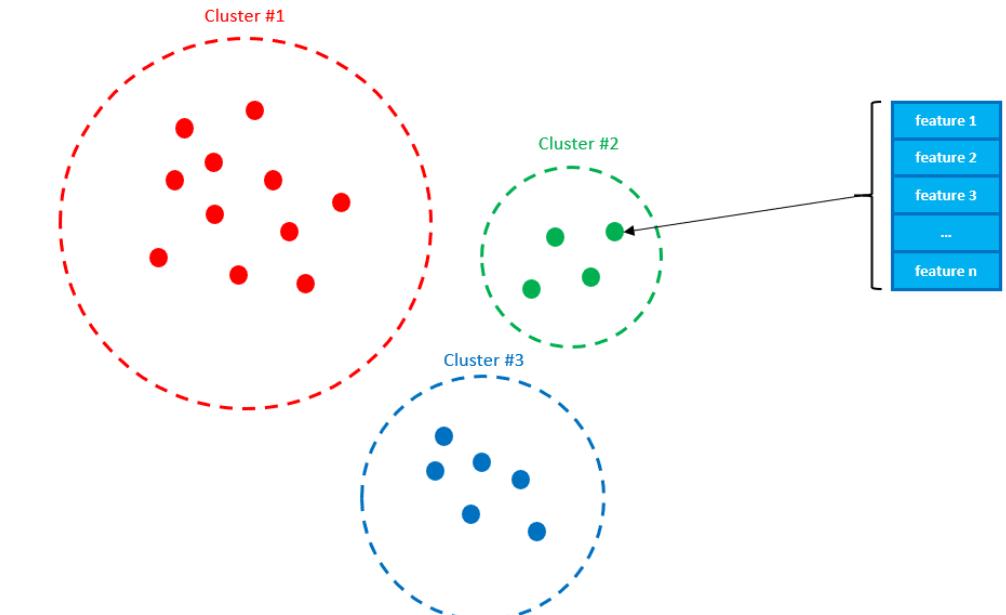
IF GENDER = 'F' AND AGEBETWEEN = 25 to 35:
KIDS = 2

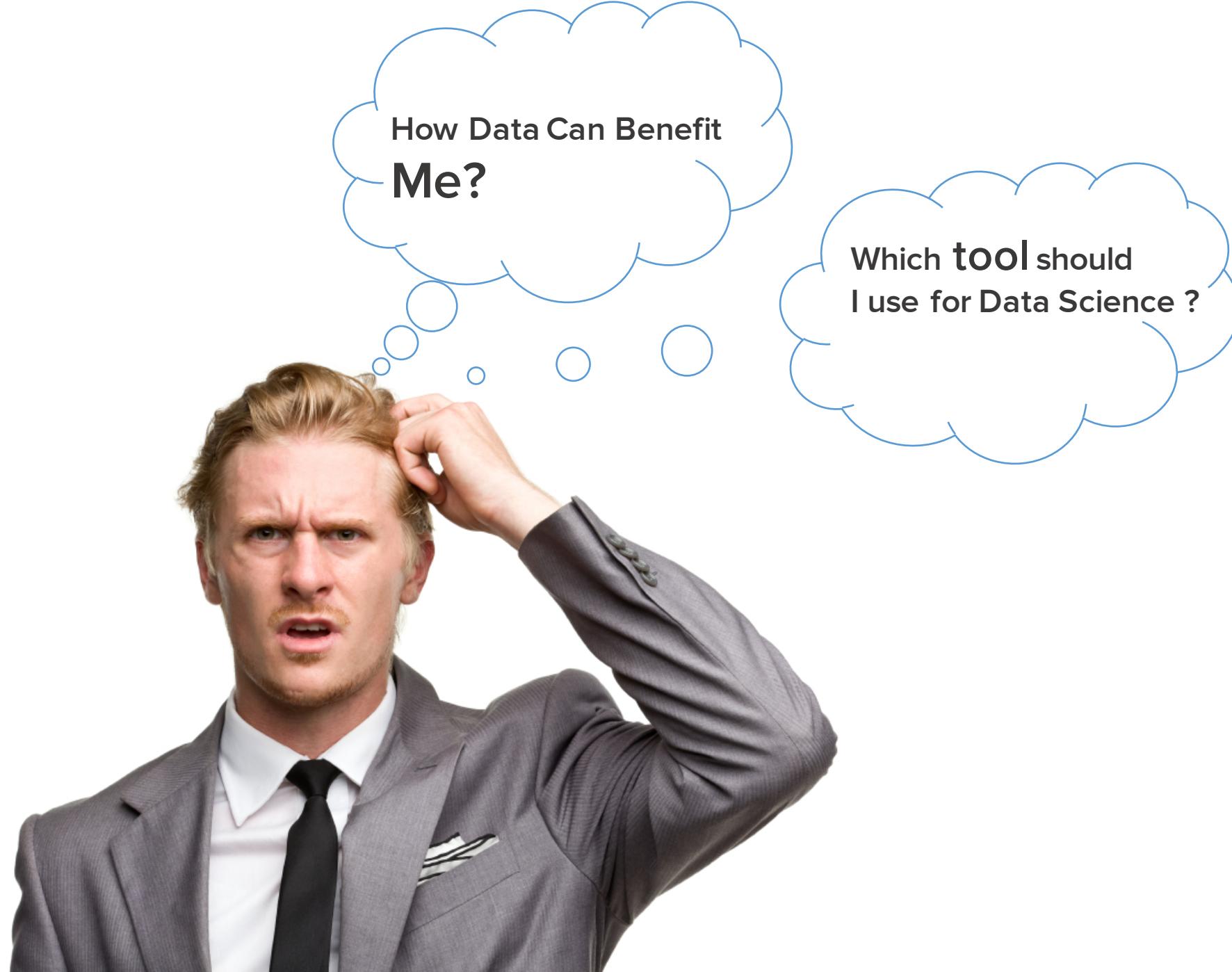
IF GENDER = 'F' AND AGEBETWEEN = 35 to 45+:
KIDS = 3+

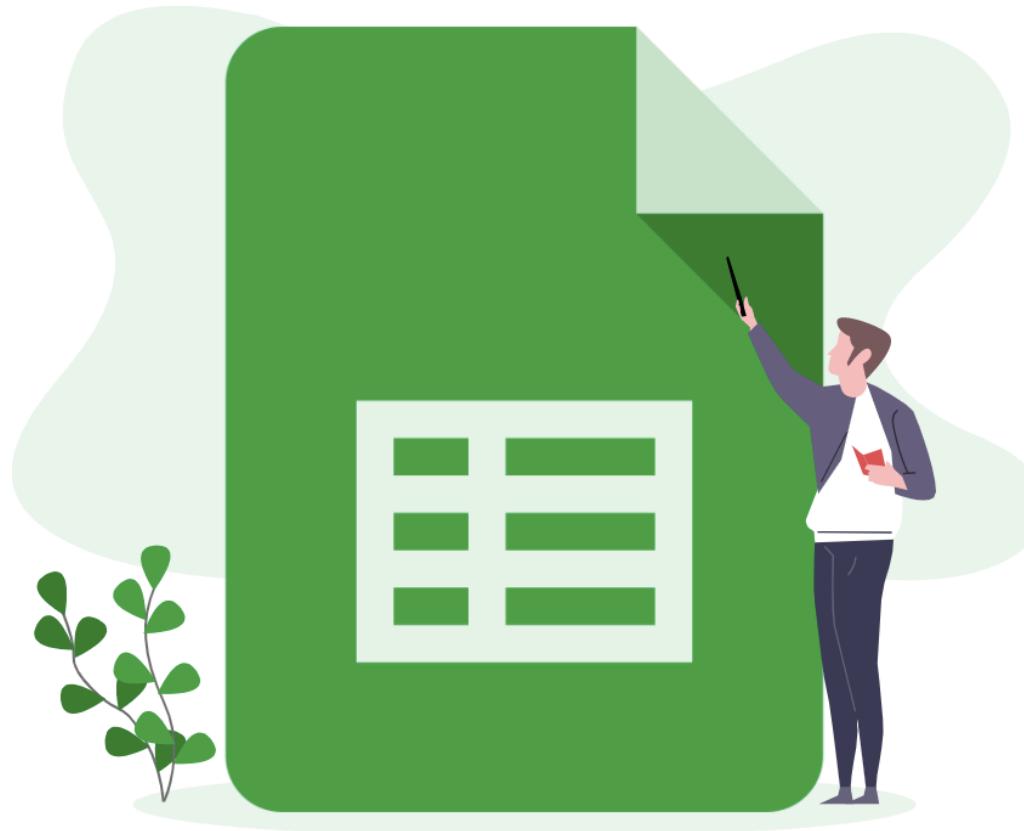
Predicted Outcome: 90% Accuracy

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.**

Where Can I Use this?

Google GOOGLE says its AI catches 99.9% of Gmail SPAM

Gmail ▾

in:spam

C

More ▾

1-34 of 34

<

>

N

Settings

COMPOSE

Delete all spam messages now (messages that have been in Spam more than 30 days will be automatically deleted)

Inbox

Starred

Important

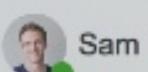
Sent Mail

Drafts

▶ Circles

▶ [Airmail]

More ▾



Sam ▾



<input type="checkbox"/> Nicole Gonzalez	Sam Wouters, Urgent Reminder! - Dear Sam Wouters, The 29Designx in association with	27 Feb
<input type="checkbox"/> Piotr Pawlak	Kinetise 2.0 - Source Code Generator - Hi! Thank you for signing up at www.kinetise.com	27 Feb
<input type="checkbox"/> Video Animation Agency	Sam Wouters, Urgent Reminder! - Hey Sam Wouters , Animated Video Department in ass	24 Feb
<input type="checkbox"/> Sohail	RE:Website Design & Development Proposal - Hi, Hope you are doing well. I recently se	24 Feb
<input type="checkbox"/> Elizabeth Gonzalez	Sam Wouters, Here's Logo for GBP19 Only - Hey Sam Wouters, Your domain needs a Lo	23 Feb
<input type="checkbox"/> suman kumari	RE:Website Design & Development Proposal - Hi, Hope you are doing well. I recently se	23 Feb
<input type="checkbox"/> Best Content Help	Complete Your Website! Win Your Customer With Great Content - Add Brandon@bestc	22 Feb
<input type="checkbox"/> Elizabeth	Sam Wouters, Weekend Special! 🎉 50% Off - Branding Bundle Deal 🎉 Extended Due	22 Feb
<input type="checkbox"/> priya singh	Google Help - Best Wishes! We are offering web services for our beloved customers only! '	22 Feb
<input type="checkbox"/> Nicole Gonzalez	Sam Wouters, Reminder! 💰 Get 80% Off Your Website Design & Development! - Dear	22 Feb
<input type="checkbox"/> Web Services	Revamp Your Website At Low Price !! - Hi, Hope you are doing well! We are an Indian co	22 Feb
<input type="checkbox"/> Sierra Alex	Custom Branding with Unique Sense - Do not Wish to receive these Newsletters.Reply u	22 Feb
<input type="checkbox"/> vineet	Website e-Commerce SEO Apps at affordable price - Hi, I am Vineet Now get 10 Pa	21 Feb
<input type="checkbox"/> ankit.mishra23	Web Design - Hi, Hope you are doing well! W	21 Feb
<input type="checkbox"/> Info	Important: Complete Search Engine Regis	21 Feb

No recent chats

Start a new one



CleverInsight Pvt Limited, Bangalore

House of Cards

★★★★★ 2013 TV-MA 1 Season HD 61

Sharks gliding ominously beneath the surface of the water? They're a lot less menacing than this Congressman.



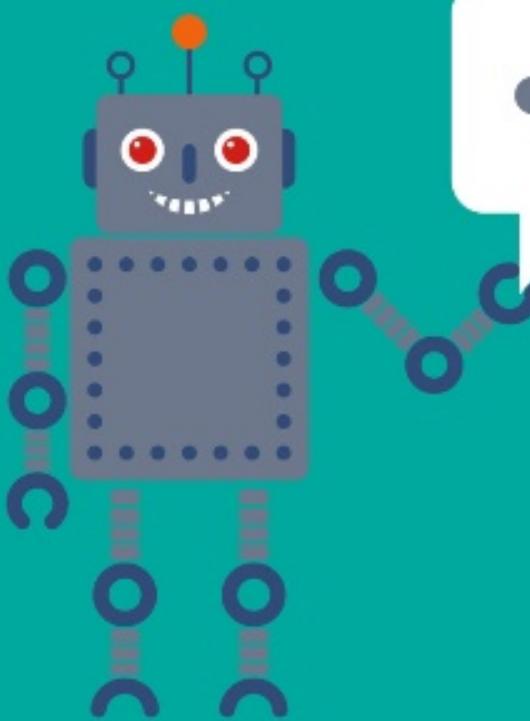
This winner of three Emmys, including Outstanding Directing for David Fincher, stars Kevin Spacey and Robin Wright.



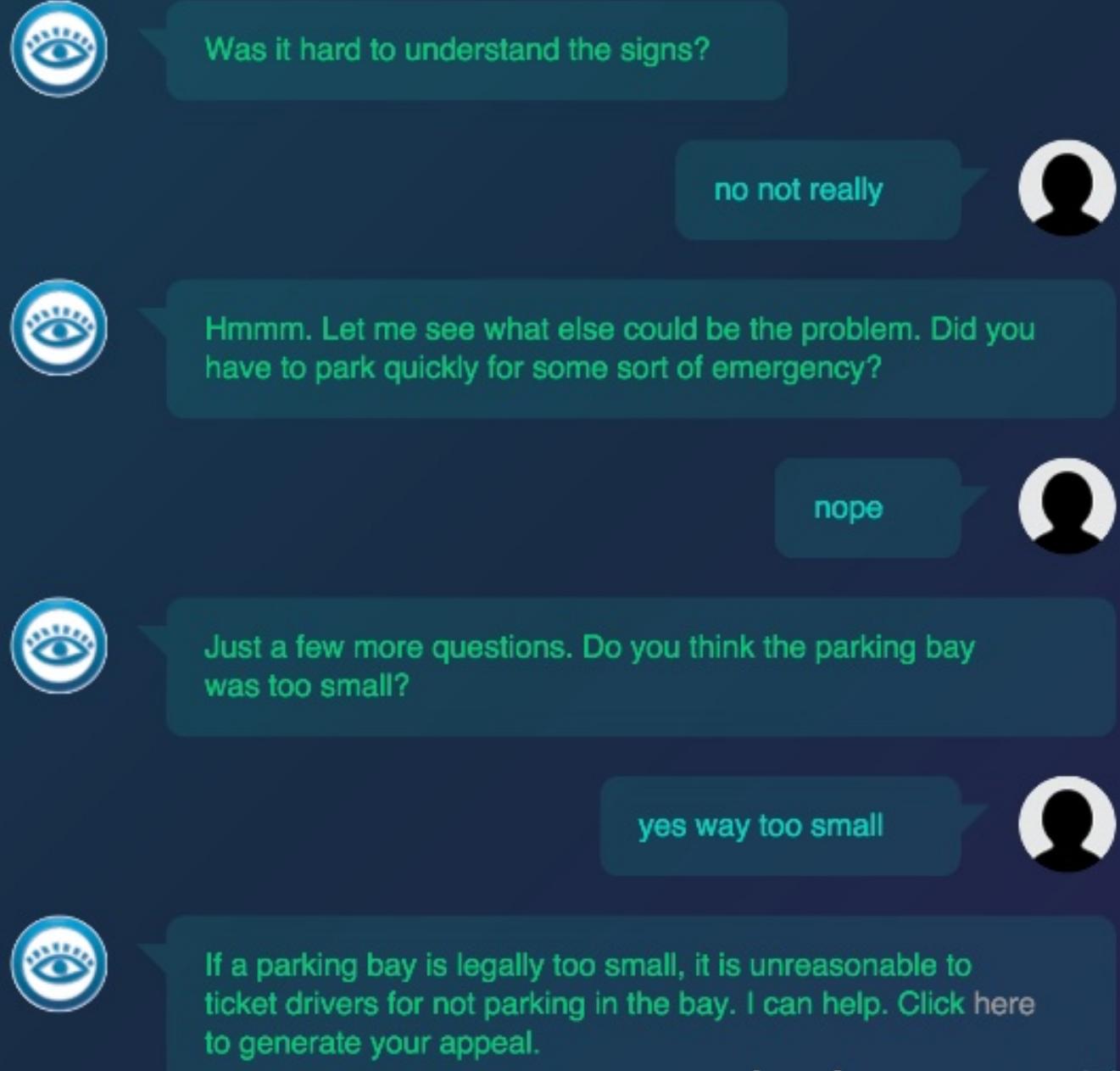
Because you watched Orange Is the New Black



NETFLIX uses AI to go from **RECOMMENDATIONS** based on what you've seen, to what you like



A 19-year-old made a free **CHATBOT LAWYER** that has appealed **\$3M IN PARKING TICKETS**



IDENTIFYING DISEASES by comparing huge amounts of data



01000110
01110010
01100101
01100101

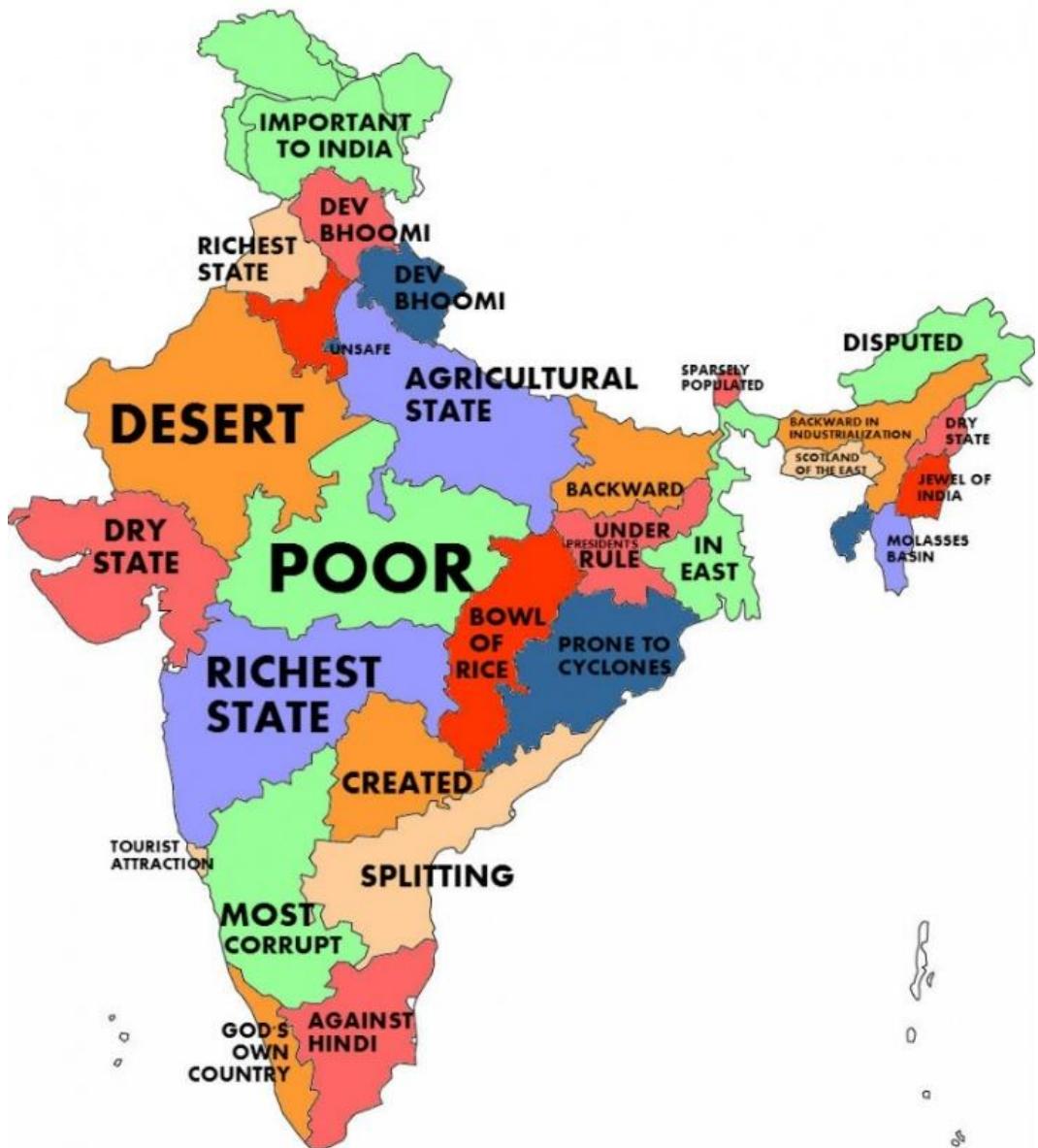


FREENOME, a startup focused on **DETECTING CANCER THROUGH AI**, set up by a 28-year old, landed a **\$65 MILLION INVESTMENT** by Andreessen Horowitz & Google Ventures

UBER is using AI for ROUTE OPTIMIZATION



What Google Search Tells About India.



Why is _____ so ?

Data Is Our Hero

In God We Trust; all others must bring DATA.

- William Edwards Deming -



Time For Questions

Time For Quiz

<https://bit.ly/2zuLJS7>

Break

CleverInsight

CleverInsight
Open AI Foundation

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



Bastin Robins .J

Chief Data Scientist
robin@cleverinsight.co

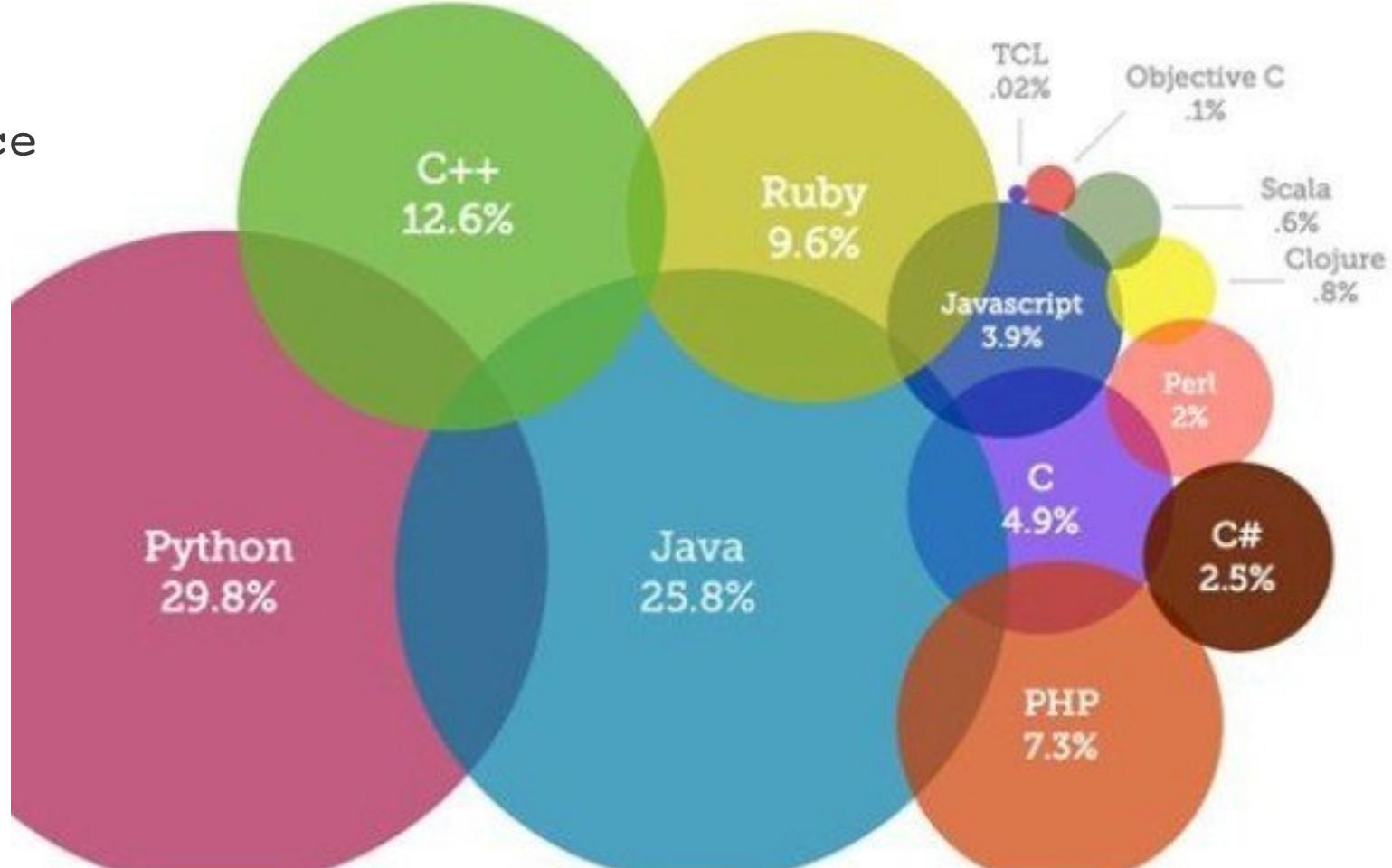
Python For Novice Beginners



Why Python?

- Simplicity
- Large Community
- Support for Data Science
- Expressive syntax
- Platform independent

Most Popular Coding Languages of 2013



See Python In Action

Tell me the given number `n` is odd or even

Simple Math

n = 10

n % 2 equal to 0 - EVEN

n % 2 not equal to 0 - ODD

Python code

```
n = 10  
if n % 2 == 0:  
    print("EVEN")  
else:  
    print("ODD")
```

Give me first `n` odd numbers

Simple Math

Given N = 10

1
3
5
7
9

Python code

```
n = 10  
for i in range(n):  
    if i%2!=0:  
        print(i)
```

Lets Dissect Syntax

```
n = 10
```

```
if n % 2 == 0:
```

```
    print("Even") ➤
```

```
else:
```

```
    print("Odd")
```

n is a variable

Variable Name : n

Value inside variable : milk



Lets Dissect Syntax

```
n = 10
```

```
if n % 2 == 0:  
    print("Even")
```

```
else:  
    print("Odd")
```

- Put value 10 inside container named n
- Now check if value inside container n is divisible by 2. if 0 then true it is divisible print even If not 0 false its not divisible.
- else print odd

What is a loop

Loops helps us to do a repetitive task

Example 1:

```
n = 10  
for item in range(n):  
    print(item)
```

Example 2:

```
n = 10  
for item in range(n):  
    if item % 2 == 0:  
        print("Odd")  
    else:  
        print("Even")
```

What are datatypes

```
a = 10      #integer  
b = 1.5     #float  
c = "R"      #character  
d = "Python" #string  
e = "Python is a basic programming language" #string  
f = True     #boolean  
g = [1, 2, 4] #list  
h = (1, 2, 3,) #Tuple  
i = {"name": "Robin", "class": "Data Science"} #Dictionary
```

What are list, tuple or dictionaries

```
name = "Bastin Robin" #variable
```

```
names = ["Bill gates", "George Cloney", "Kevin"] #list
```

```
names = ("Bill gates", "George Cloney", "Kevin") #tuple
```

```
#Dictionary
```

```
name = {  
    "name": "Bill gates",  
    "company": "Microsoft",  
    "age" : 64,  
    "kids": ["Adele", "Jennifer", "Rory"]  
}
```

What are list, tuple or dictionaries

```
name = "Bastin Robin" #variable
```

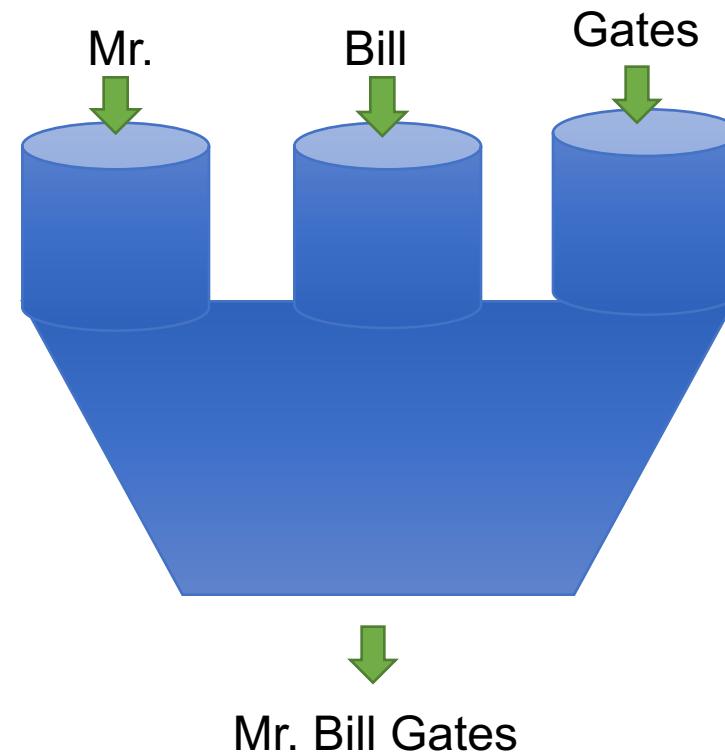
```
names = ["Bill gates", "George Cloney", "Kevin"] #list
```

```
names = ("Bill gates", "George Cloney", "Kevin") #tuple
```

```
#Dictionary
```

```
name = {  
    "name": "Bill gates",  
    "company": "Microsoft",  
    "age" : 64,  
    "kids": ["Adele", "Jennifer", "Rory"]  
}
```

Functions



What are functions

```
def sum(a, b):  
    print(a+b)
```

```
>>> sum(10, 20)  
>>> 30  
>>> a = sum(10, 20)  
>>> a  
>>>
```

Without return type

```
def sum(a, b):  
    return a + b
```

```
>>> sum(10, 20)  
>>> 30  
>>> a = sum(10, 20)  
>>> a  
>>> 30
```

With return type

Class & Methods

Quiz

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

Take away

"The power of machine learning comes from its ability to learn patterns from large amounts of data. Understanding your data is critical to building a powerful machine learning system."

Tools are nothing, Techniques are everything.

- Data Wrangling & Cleaning
- Data Visualization & Grammar of Graphics
- Exploratory Data Analysis
- Break -
- Machine Learning Models
- How to pick right model