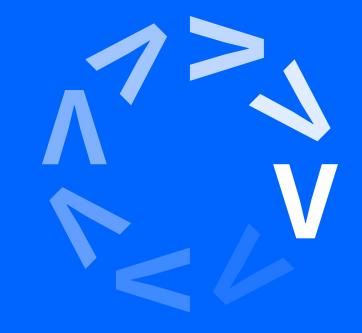


## Welcome Students!

- Please do the following:
  - Sign in on the attendance sheet
  - We will NOT be on computers today, so no need to turn them on

## **Mentor Introductions**



Presented by IBM Volunteers

## Putting the Science in Data

### The Basics of Data Science



Presented by IBM Volunteers



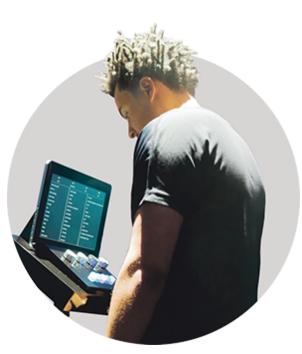
### What is Data Science



### Where does Data Science Lead?

- Data Scientist / Engineer / Analyst
- Marketing
- Consultant
- Risk Analyst
- Researcher
- Creative Arts





### It all starts with data...

ID	Gender	Status	Children	EstIncome	CarOwner	Age	LongDistance
1	F	S	1	38000	N	24.393333	23.56
6	М	М	2	29616	Ν	49.426667	29.78
8	М	М	0	19732.8	N	50.673333	24.81
11	М	S	2	96.33	Ν	56.473333	26.13
14	F	М	2	52004.8	N	25.14	5.03
17	М	М	2	53010.8	Ν	18.84	12.45
18	М	М	1	75004.5	Ν	64.8	26.52
21	М	М	0	19749.3	N	60.366667	20.22
22	М	S	1	57626.9	Y	43.906667	9.38
23	М	М	2	20078	N	32.846667	9.65
24	F	М	2	47902	N	26.033333	17.44
29	М	М	1	7545.96	Y	16.753333	22.39
35	F	S	0	78851.3	N	48.373333	0.37
36	F	S	1	17540.7	Y	62.786667	22.17
37	F	М	0	83891.9	Y	61.02	28.92
38	F	М	2	28220.8	N	38.766667	26.49
40	F	S	0	28589.1	N	15.6	13.19

Structured data such as Excel files and relational database tables

Fast and massive amounts of data. IoT, email, network traffic, sensors, etc.



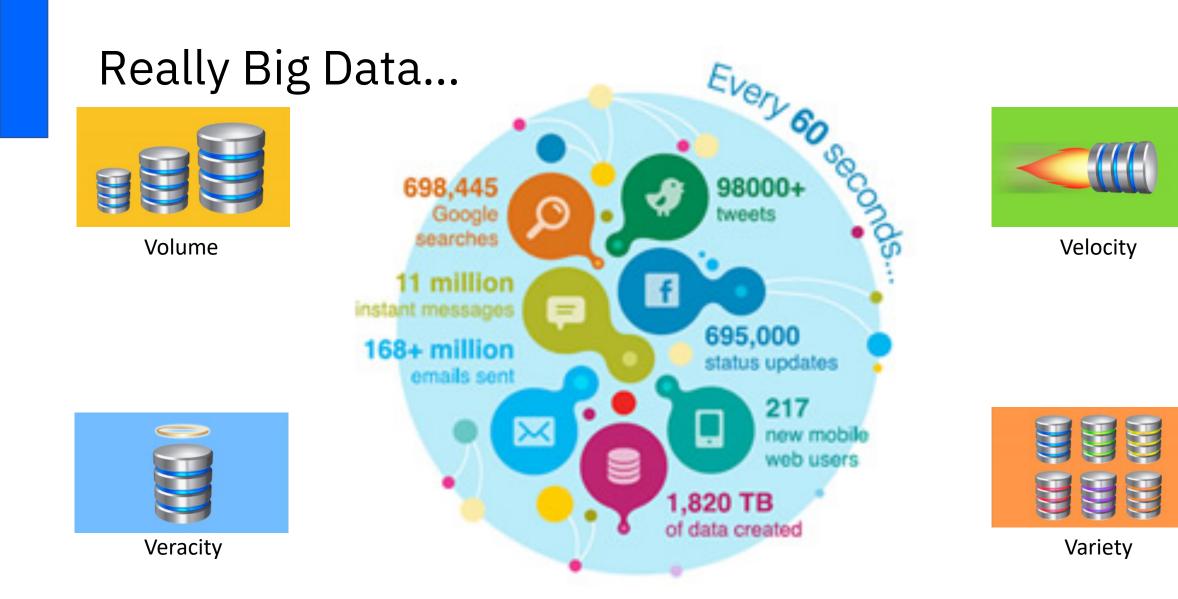
IBM 🤣 @IBM · 12h

Operating in more than 100 schools worldwide, **#PTECH** is accelerating students towards success in some of the world's fastest growing industries. Learn more: bitly.com/2Hw4YTj

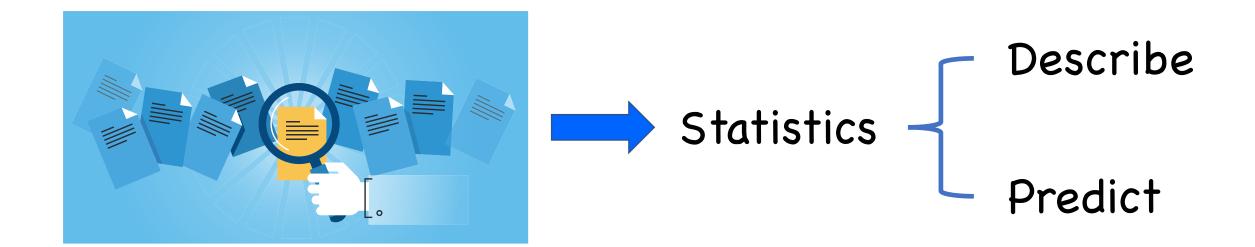


♀ 1↓ 14 ♥ 42 ♥

Unstructured data such as tweets, posts, pictures, music, video, etc.



### And add in science...



### Goals of Data Science

### Describe

Rain=no and Day=Saturday -> Play = Y

## Help to understand the relationship between the inputs and the output.

### Predict

Color	Texture	Shape	Avg Size (in)	Sweet	Object
Green	Smooth	Oblate	4	No	Tomato
Yellow	Smooth	Round	3.5	Yes	Apple
Green	Rough	Oblate	2	No	Tomatillo
Red	Lumpy	Cylindrical	2.5	No	Chili
Green	Rough	Oblong	11	No	Cabbage
Yellow	Smooth	Oblong	8	Yes	?

Make predictions for a new sample described by its attributes.

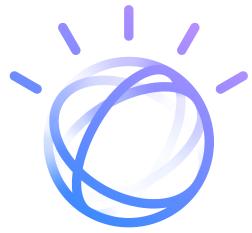
88% chance of being a banana

## What makes it possible?

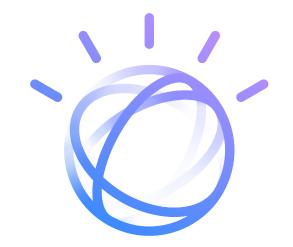


## Has Data Science impacted your life?

- It has if you
  - Use social networking websites like SnapChat, Instagram or Facebook
    - Sites collect data to present relevant ads/feeds
  - Go shopping
    - many stores rearrange products/change design based on data science
  - Receive recommendations
    - Watching movies, reading books, shopping
  - Ask Google, Siri, Alexa, Cortana, etc. a question

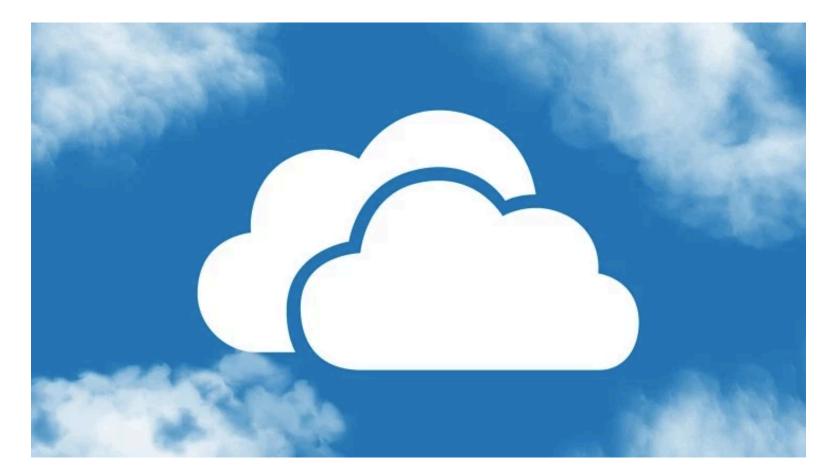


### Data Science Fun Facts



- It can apply to nearly any industry
- Nearly 90% of data scientists are contacted at least once a month for new job opportunities
- 88% of all data scientists are now happy or very happy in their job
- Less than 0.5% of all data we create is ever analyzed and used
- Presentation is key

## Any questions?



# Directions

- 1. Divide the class into four number off 1 to 4
- 2. Each number is a team
- 3. The teams will be given one challenge to work on as a group
- 4. After 3 minutes, the teams will need to be ready to discuss with the mentors
- 5. The team with the clearest explanation is the winner
- 6. The winning team gets bragging rights & a swag item per team member!

#### **IBM Volunteers**



What would a computer or data scientist see?



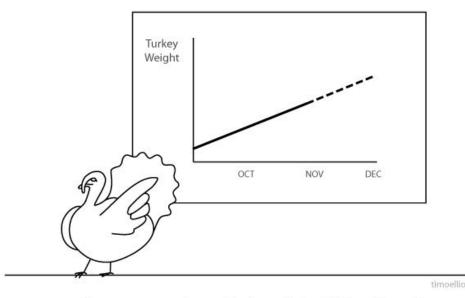
#### **IBM Volunteers**

15

## 2 - Predict

### What is wrong with the turkey's prediction?

### THANKSGIVING PREDICTIVE ANALYTICS



"I see no reason why excellent growth shouldn't continue ... "

### **IBM Volunteers**

## 3 - Predict

Can you determine a pattern to predict "Friend of Elvis?

Type of Animal	Caught a Rabbit	Friend of Elvis
Cat	No	No
Dog	No	No
Dog	Yes	Yes
Eagle	Yes	No

#### **IBM Volunteers**

17

## 4 - Describe

Determine a pattern for related purchases?

Item 1	ltem 2	ltem 3	ltem 4
Apple	Banana	Cherry	Durian
Apple	Durian		
Banana	Durian		
Durian	Banana	Cherry	
Banana	Durian		
Apple	Banana		
Apple	Cherry	Durian	

### **IBM Volunteers**



### **IBM Volunteers**

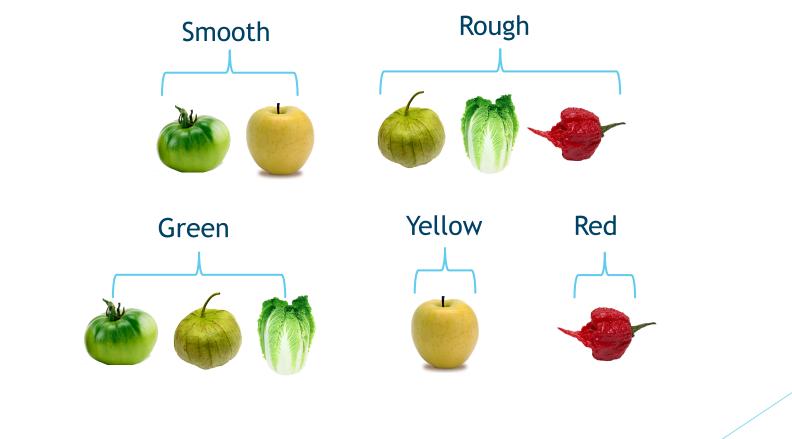
## **Present Answers**

**IBM Volunteers** 

20



What would a computer or data scientist see?



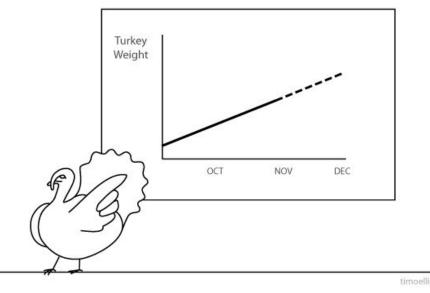
**IBM Volunteers** 

21

## 2 - Predict

### What is wrong with the turkey's prediction?

### THANKSGIVING PREDICTIVE ANALYTICS



"I see no reason why excellent growth shouldn't continue..."

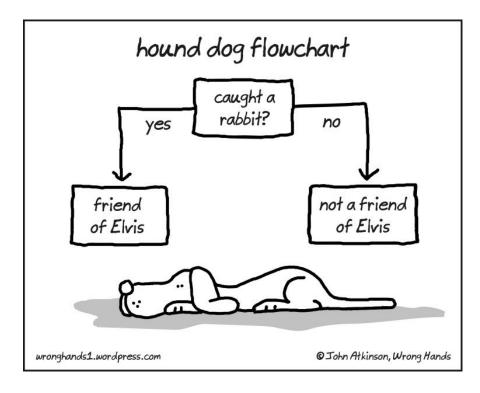
The main issue is the turkey's weight may not keep going up after Thanksgiving and weight cannot go up forever.

### **IBM Volunteers**

22

## 3 - Predict

Can you determine a pattern to predict "Friend of Elvis?



#### **IBM Volunteers**

## 4 - Describe

Determine a pattern for related purchases?



Banana/Durian 80% Cherry/Durian 100%

#### **IBM Volunteers**

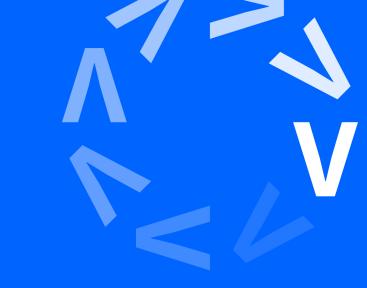
## **Discussion questions**

- What did you learn about data science through this activity?
- What lingering questions/thoughts about data science do you have related to being a web development college student and your future career?

Q&A



## Thank you.



Presented by IBM Volunteers