0:08

okay

0:27

so today in the second class

0:35

i'm going to talk about the importance of data

0:47

let's wait few more minutes

0:55

okay so [Music]

1:01

thomas edison said uh remember from the previous

1:07

semester thomas edison is an important guy and worked in very many different fields

1:15

and he says doctor of the future will give no medication but will interest his patients in the

1:22

care of human frame diet and the cause of prevention of disease so the modern

1:28

medicine will deal with prevention and well-being of the human rather than cure

1:36

that's what he proposes and even today we are not that point yet

1:42

because medical [Music] industry

1:48

is acting like an industry so [Music] there are patients and there are

1:54

diseases and there are cure and that

2:02

that is considered as a sector so without any patience

2:09

there will be no sector involved and the thomas edison says

2:16

[Music] a good medical advance

2:22

instead of curing people we can [Music] prevent disease and still

2:29

do some medical work for those prevention

2:35

operations that's what he said so today i am going to talk about data

2:42

and trust and trust in data and calculation and the

2:47

one of the guys in states he is in

2:53

he's working in a planetarium he's a science famous scientist popular scientist not a famous but popular

3:00

scientist he says the good thing about science is that it is true whether or not whether or not you believe in it

3:08

some other people also says i don't believe in science i trust science so you don't have to believe it because

3:14

regardless of your belief it's true [Music]

3:20

so it is an interesting fact let me come back to this this is an early

3:27

calculator from late 1970s hp calculator with the led displays and

3:34

the crystal displays we used to have some of these

3:40

i have some of these old calculators as well so there are problems like some

3:47

multiplication and trigonometric functions so well once we do this with

3:52

the calculator we trust the results so we don't question

3:59

the resulting number and we write them down as a result so how do we

4:05

how do we trust that's an interesting question and you should think about it

4:10

so since it is repeatable all the time and it is validated by many people

4:17

it is accepted as it is so it is accepted as the result is correct so

4:24

it's today for today some people believe that whatever computer says

4:31

is correct regardless of the algorithm correctness so if the computer says

4:38

certain record is not found people believe it that it is not theory so

4:44

it's an interesting phenomenon so we believe in machines because machines have

4:49

repeatable similar results every time you multiply these

4:56

two numbers you get the same result and if you do it manually you get the same result then if you multiply different

5:03

numbers you don't you don't think that the result will be

5:08

wrong you believe attracts to ashamed

5:16

in terms of data we are facing with that today's topic

5:23

we have a lot of data going around one of the data we are

5:28

having attention is the colonialist cases and the [Music]

5:34

statistics and the numbers collected variables

5:39

through applications and through medical systems so that's a lot of data going on

5:47

even then even the science is improving our lives are not that

5:53

advanced and today russia is invading another country

6:00

and as you see on the left top

6:05

children are eating unhealthy foods in schools

6:10

in restaurants bottom in furnitures there is a formaldehyde so

6:18

in baby room furnitures even has that matter has that chemical in it

6:25

children's do smoke etc so a lot of [Music]

6:32

a lot of things are happening against science still

6:37

uh we cannot eliminate those things from our lives therefore what i'm trying to

6:44

say here is that um even we trust the science

6:51

and calculation we don't follow the results

6:57

if we make a calculation it's evident that people who smoke will die earlier

7:04

but yet people still smoke

7:11

and there one thing is scientific research must be done properly and on the right side

7:19

what you see is [Music]

7:30

science [Music] the cigarette they say

7:37

is healthy and it keeps you awake etc so it has

7:42

some benefits scientific evidence evidence on effects of smoking which is

7:49

which is good what they say so it is interesting

7:54

therefore science is only helpful if done properly maybe that is the reason of disbelief

8:02

in some cases when doing science

8:07

and computations i'm i'm coming back to data a lot of variables

8:14

are [Music] indeed around those are

8:21

body variables such as height weight blood pressure

8:26

to a person status variables like address

8:31

of a person position of the person geographical position if it is running if he is or she is

8:37

running or driving speed of the person etc and in our life so this is daily

8:43

life variables input variables like things we read question

8:48

yeah uh you said that we should trust science but

8:54

as you know science changes sometimes and a lot of things

9:00

that we think is true turns out to be wrong and if there is the case how we

9:06

could trust signs actually

9:11

that doesn't happen for a good for a good result

9:18

something that is considered to be true doesn't get invalidated

9:24

it is extended so or

9:30

it is it is marketed in that way for certain reason

9:36

obviously for example in this research they know that

9:41

smoking is bad but they found a way to tweak

9:46

the results so that smoking is

9:51

good or good for your health smoking chesterfield cigarette is good for your

9:57

health actually it is not but they all they also know that but they tweak it so

10:02

if they change if then if you realize that signed smoking is bad

10:08

it is not a contradiction it is just extension and therefore

10:13

and [Music] and you have no choice other than believing in science

10:20

that is the best option [Music] having the possibility of

10:25

changing things in future shouldn't let you [Music]

10:31

keep away from the science rather

10:38

should direct you for more research so

10:43

[Music] if you don't continue

10:50

the pursuit of the science and pursuit of the topic

10:57

then it will mean that you take it as granted

11:03

you accept it as it is and it becomes like it turns into beliefs

11:09

rather than truth so try to

11:15

try to keep looking at it if the topic is not clear so existing existing

11:22

uh status of the science existing status of the knowledge about the certain subject

11:28

for example earthquakes tells us certain amount of information

11:36

but new information usually does not validate does not invalidate the old

11:43

rather it extends it and sometimes extension

11:49

has exceptions those exceptions may may be

11:54

regarded as invalidation i think that's

12:01

that explains it so it never fails but it is extended and those extensions sometimes

12:08

looks like contradiction

12:20

the control variables that the things that we manipulate are things that we select pushed about said

12:27

wrote so that's that is the variables that we do [Music]

12:34

in our daily life and external variables like temperature weather pressure family economy salary

12:41

and so on so those are external variables other variables do exist in our life

12:47

and in the in the old days there were no data no recording now we recorded

12:54

these and we can process we can process what we read what we see what we watch what we

13:00

do what we click what we what we are where we go how fast we go etc how

13:09

much weight do we have what is our blood sugar etc blood pressure

13:15

so we can process those heartbeat etc

13:22

this is a good excerpt from the

13:30

from this document this document is a document generated in

13:37

i think 1954 the library entrance is 1964 but the machine is

13:43

made earlier the name of the machine becomes one of the first computers is bendix

13:49

bendix 15 computer it has a forward

13:56

preface of the user manual and i am going to share it with you in

14:01

the slides because it is very very good it says one man one of the man's first

14:07

problem was to find the shelter in the forest he solved it by walking over the land

14:14

until he found a cave so so that he can sleep in the cave rather

14:19

than the forest that was simple direct solution to a simple problem as centuries merged into ages he faced

14:26

more problems and learned to remember more facts concurrently his brain became more

14:32

active thinking power increased both in rate and complexity so that

14:37

he faced more problems because he was more intelligent he created more problems and learned more facts and at

14:44

an ever-increasing rate so that we that's what we call intelligence therefore it caught a milestone of

14:51

mental activity jumping the mental activity and man the human kind

14:57

came to the present day because of the increasing complexity of the problems

15:04

men were forced to become specialists division of labor that we have seen

15:11

complex complexity of problems men were forced to become specialists dividing the problems into sub problems then

15:16

dividing sub problems into other subproblems and

15:23

oops and one of the categories in which some men specialized was

15:29

mathematics so one category was mathematics category of the sub problems

15:39

other men came to mathematicians for solutions to those problems whose factors could be

15:45

expressed numerically so like in second world war soldiers

15:50

army came to mathematicians saying that let's make an messaging system so that enemy cannot

15:57

understand so cryptology came out and then other side tried to solve it so

16:05

that means decrypting the code problem is

16:11

uh came out as a problem to solve but the mathematicians then found two

16:18

types of problems were arising more and more frequently one problems of complex

16:23

such complexity that their solutions are impossible by passing on paper methods

16:28

so it requires numerical integration etc so problem is very difficult to solve by

16:35

hand it requires numerical solution second

16:41

problems are easy possible to solve by pen and paper

16:47

but it has to be repeated for each case so it will be very time consuming and it

16:53

will be taking months or years to make therefore a machine is required

16:59

therefore people implemented machine and they try to

17:04

make they try to compute things by machine so mathematicians go back to engineers

17:12

some of them are mathematicians and engineers together at the same time and asked for

17:17

development of machines to help man in the generation of solutions by increasing the speed

17:23

at which mathematician could work so speed increments

17:28

the engineers responded with various combinations of keys wheels gears cams etc that we have seen last week

17:36

which were referred to as calculating machines a calculated machine was placed on the mathematician desk and called as

17:43

desk calculator this machine system is generated solutions at a highly increased rate

17:50

compared to previous because pen and paper is replaced by the desk calculator

17:55

and making possible more solutions for mathematicians but it wasn't enough

18:03

increasing demands upon mathematicians soon exceeds this even the most efficient man

18:09

machine system with the desk calculator so far devised engineers were called again

18:15

and this time they switched to electronic equipment use of electronic circuits reduced the

18:21

time necessary to perform arithmetic operations to the point where a single operation such as addition

18:27

could be performed in a few microseconds this is 1954

18:33

one millionth of a second that's micro

18:39

and this new type of machine was called as digital computer that's what they say

18:44

so that is the definition of digital computer at the introduction of uh

18:51

bendix g15 computer and some engineers did not switch

18:57

entirely to electronic circuits they combined them with the mathematic mechanical features

19:02

of the old calculating machines developing a computing machine which used mechanical operations to represent

19:08

arithmetic and physical quantities such as degrees of rotation of the shaft how many degrees can contain the number

19:16

represent quantities or numbers such as weights on the right side this type of computer

19:22

is called as an analog computer the accuracy of an analog computer is limited by the accuracy

19:29

with which a physical quantity can be

19:34

measured um [Music]

19:42

let's wait i'm going to on the left side you are seeing a mechanical

19:48

calculator a mechanical device for the army it makes a calculation for hitting the

19:55

target so the calculations are made by mechanical [Music]

20:01

gears and levers rather than digitally computed

20:07

values so we call that type of computer as an analog computer

20:13

let me let me share the sound as well

20:25

there is a short video that we are going to watch now

20:46

so this is gonna be fun yeah

20:56

this is outsystems the most complete full stack application development platform for professional

21:03

so this is going to be a super short video on analog computers

21:08

and what you see right now is the output of a very simple analog computer that i put together

21:14

and here it is on the breadboard it's just made of a couple components just um

21:21

you have an op-amp there an analog switch and a couple of passive components

21:27

um and my lab power supplies powering at all and um

21:32

basically what i've programmed this computer to be doing is solving a differential equation which differential

21:39

equations are kind of what analog computers are really most suited for solving

21:45

that's that's really what they excel at and that's what people use them for so the differential equation that i'm

21:51

solving is this one right here you see i have x dot

21:57

equals x is the equation or that is the derivative of x

22:03

is equal to x the derivative of x with respect to time is equal to x where x is a function of

22:10

some kind and then you can also integrate both sides to get x equals the integral of x plus c

22:17

and the solution for this equation you can find analytically pretty easily is just

22:23

x equals e to the x or the trivial case where x equals zero

22:28

but x equals e to the x is what we're mostly interested in so the way you solve this

22:34

in the analog domain is you you set up the computer as shown in this little diagram here so

22:41

you have an integrator and a summer which is an inverting summer so

22:47

in analog computers the summing components are typically inverting as well as the integrators in

22:53

fact i think probably all the components except multipliers

22:59

are inverting and that's just kind of the the way it has to be because of

23:05

how how op amps work with negative feedback and stuff equations quite accurately and and very

23:12

quick as well so you can see this this chart on the oscilloscope here is

23:18

indeed the graph of e to the x we're pretty pretty close to it

23:25

and you can see i can i can scale it down as well so

23:30

the the way this works is the the integrator and the inverting sum block are both contained in this um

23:38

tllo72 op amp my my function generator which is

23:44

outputting a pulse like so um and the pulse basically resets the

23:51

integrator each cycle so that way you can it'll or that way it'll redraw itself on

23:58

the oscilloscope dx and it looks like the familiar

24:03

new level of insight i guess we should say that's it so this is an example of

24:09

analog computer that is used to compute a certain

24:16

output using an input if that is the value you are going to

24:23

that is the value you are going to use you can use it directly to control

24:29

an engine to control a lamp controller weapon system or control an

24:35

airplane or whatever that will be an analog control

24:41

through analog computing but uh with

24:48

the bandits that i was talking about uh

24:53

with the digital computers using electronic circuits um we represent numbers and quantities

25:02

with digits and any desired at any desired degree of accuracy can be obtained through adding more digits so

25:08

we can extend it we can make it more precise if you want to the precision of analog

25:16

computer is limited with the precision of the equipment precision of the resistors

25:22

precision of the capacitors and precision of the oscilloscope and the measuring devices so resistance of the

25:28

cables all play important role however when you are doing digital

25:34

that doesn't happen so you can go any any deep

25:43

any kind of accuracy as you want

25:50

therefore [Music] at that time bendix says

25:55

most digital computers are general purpose most digital computers computers are

26:01

general purpose designed for multitude of users and scientific and commercial

26:14

[Music] so

26:20

let's think about what data is as i said human is interacting with the

26:27

environment changes the environment and observe the result and makes decisions to continue interaction so that human

26:34

will live true five senses continuously it makes me he makes measurement and

26:41

evaluate using brain and stay in existence that is what human

26:48

mankind do modern human uses not only five senses

26:53

but also uses measurement techniques and tools to understand the surrounding environment and process the external

27:00

world interact with the external world these collection of measured and

27:05

obtained variables are called as data so if we measure variables

27:10

and write them down somewhere it could be paper that is called as data

27:19

this data also variables uh can be in different forms it could be

27:26

qualitative or quantitative nitel and nigel

27:31

if it is qualitative data for example very beautiful day we understand it as human but

27:38

it's hard to measure it although it gives you a measurement it is kind of subjective

27:44

so we don't say that it is not a measurement it is a measurement but

27:50

the scale is difficult to tell the quantitative on the other hand

27:55

has multiple dimensions one is discrete the other one is a continuous

28:01

continuous data means real numbers the quantities that are represented by

28:08

real numbers such as height weight etc and discrete ones are

28:15

the ones that are countable four five six etcetera

28:23

about a dog for example uh data that is qualitative is brown long hair powerful

28:29

so those are qualitative data it gives a lot of information but it's qualitative

28:34

so it doesn't mean that it has low value it is just different

28:41

discrete data about the dog is it has four legs how many legs four legs usually

28:47

has two brothers or the weight is continuous variable so 25

28:54

kilograms something kilograms and something something something centimeters

29:01

on the other hand if we switch to if we switch to digital world nothing is fully continuous so there is

29:07

a limit of the number precision and iterability format

29:13

that you have seen in the previous 151 or 101

29:18

you have seen that in 101 the

29:23

science that the the traditional science that deals with data is called as statistics

29:30

uh what what it does is he collects it collects and analyzes numerical data in large quantities

29:37

especially for the purpose of inferring meaning extracting proportions in a

29:42

whole from those in a representative sample meaning that

29:49

you get you measure certain things and you you conclude that

29:54

the idea is this

30:00

most of the disease is coming from this direction or

30:05

those who study more can pass exams

30:11

better because of the data obtained so that is what statistics

30:16

do and on the picture it says it is the science of collecting

30:22

organizing presenting analyzing and interpreting numerical data to assist in the making more effective decisions at

30:29

the end you will make decision in management or engineering or

30:37

defense or whatever so that is what statistics is and it is the root of

30:45

actually data science in another

30:51

video it says it is the science of collecting classifying presenting and interpreting

30:57

data so collecting classifying presenting and interpreting data

31:02

examples are these uh how data is grouped is classifying how

31:08

data is presented how what kind of graphs are used graphs

31:13

that will be graph graph not grass um [Music]

31:20

and some of them is some part of the statistics is called

31:26

descriptive statistics it just explains the case

31:32

so it will say like 10 percent of the population is

31:38

disabled that is descriptive statistics inferential statistics says

31:45

by the year 2050 world temperature will increase

31:51

by 2 degrees that is what inferential statistics so it will

31:56

extract in information for future or existing cases so it will

32:04

extract additional information

32:09

and as i said in the old days there were no data and the data that we had

32:16

on the left were collected manually recording techniques

32:22

after these years emerged so we don't collect data manually

32:27

we use computer databases but

32:32

database means a file system that holds active data

32:39

to read and write so databases became traditional

32:44

relational databases became insufficient because we will

32:50

be talking about big data which i am going to explain in a few moments

32:57

due to the reason of number of data sources number of frequencies and number of variables that we are dealing with

33:04

since we can collect more data then people decided to collect it for certain reasons and but

33:12

you need to you need to process that too just collection is not

33:18

sufficient one definition is a data warehouse

33:23

in addition to database in the database there is an active information

33:30

but the data warehouse very amber in turkish

33:37

contains static data so historic data previous data to be

33:43

analyzed usually it is passive you don't change it anymore

33:51

okay it is known an enterprise data warehouse edw or dvh or dw is a system used for

33:59

reporting and data analysis and considered as a core component of business intelligence

34:04

it is repositories of integral data from one or more sources they store current and historical data

34:11

in one single place used for creating analytical reports for workers through the enterprise and companies etc and the

34:18

army or any kind of enterprise but the key part is it is combined

34:24

and it contains a lot of data that is static not changing

34:32

data stored in the data warehouse is uploaded from operation operational systems such as marketing or sales in a

34:39

company it may pass through an operational data store and may require data cleaning

34:45

for additional operations to ensure that data quality is good before it is used for reporting

34:52

for that purpose etl techniques are used etl means extract transform load

35:00

many of the tasks that you will be doing when you are working in the banks in

35:07

companies many places will be in the form of etl

35:12

meaning that you don't do anything new but you have data in some place

35:18

and the target structure is different so what your task will be extracting

35:25

data from its original sources processing it transforming it and loading it to

35:31

the target place so that other people can

35:37

understand it and can process it

35:42

now becomes big data uh

35:48

the concept of big data comes it refers to data that is so large

35:54

fast or complex that is difficult or impossible to process using traditional

35:59

methods traditional method is usually considered as relational databases like sql

36:08

etc then there are special tools to assess

36:13

and examine big data statistics the most popular

36:20

of these big data [Music] tools are nosql

36:26

structures like spark and hadoop which you will have a five minute video

36:34

coming the big data has certain properties a

36:41

usually it was three weeks but now it is extended to five weeks

36:48

the thing is the first three is

36:55

more important one is the volume

37:00

the volume of big data must be very big than the traditional ones second

37:10

the big data must be

37:15

structured or semi-structured or maybe unstructured

37:20

so the type of data must be filtered and processed so variety is

37:26

interesting third too much data coming very frequently so

37:35

it is coming all the time like location of vehicle

37:41

we say that velocity of data is very big

37:46

velocity means the degree that degree of trust

37:52

and value means business value of the data so that is also

37:59

that is also important for um

38:05

business

38:12

so here are some numbers about big data in the world for year 2000

38:21

big data industries worth is 270 billion dollars

38:28

this year how many google searches per day is 2

38:34

3.5 billion searches 100 billion messages

38:41

exchanged through whatsapp every day and stored

38:47

95 percent of businesses still today cannot manage unstructured data so those

38:53

data is lost and cannot be turned into value

38:59

eighty percent of data available today is unstructured just logs and

39:05

things like logs are unstructured not in tables

39:11

in the healthcare in the medical platforms

39:16

big data value is 67 billion it will be 67 billion

39:22

in 2025 so if you focus on big data and data science in future

39:30

in your career then you'll be making out of money

39:36

and one interesting further interesting numbers

39:42

79 zettabytes of data were generated in 2021 last year so zettabytes i have

39:51

i have an analogy in the next slide that is

39:56

one zetabyte is actually 250 billion dvds

40:04

billion so 79 zettabytes of data were generated in 2025

40:11

180 zetabyte of data will be generated in 2025.

40:19

more than 1.2 billion years have been spent online

40:24

by the people each day

40:30

internet users generate 2.5 quintillion bytes of data new data

40:38

in the year 2020 that was the two years ago statistics into each

40:43

internet user generated 1.7 megabytes of data per second

40:50

each each user

40:56

so through through the years uh

41:03

the data that that is generated is increasing in a very fast rate

41:09

and someone has to process it

41:14

global spending on big data as i said will be over 270 billion

41:22

and as you see the difference the

41:28

rate is increasing significantly they made a trick here they jumped

41:34

from 2019 to 2022 thinking that

41:40

it's a big jump but actually in the middle there will be 2020 and 2021.

41:47

even then the rate of increase is bigger

41:55

so data-driven approaches allow people and individuals yeah

42:07

do we have question no

42:12

data driven approaches allow individuals and businesses to watch past trends

42:18

to predict future so in order to predict future you need to

42:24

check the previous and existing data

42:30

in order to predict futures ellen k the famous scientist a computer scientist says best way to predict the future is

42:36

to invent it so if you want something goes in one direction

42:43

just do it so that the future will be in the shape of uh in the in the way

42:49

that you want and i have put a video of ellen k

42:56

uh to you so [Music]

43:03

you will have it for the

43:08

next for the next week

43:14

inventing the future is a talk by lnk

43:22

a real world problem uh classical house problem the pool problem

43:30

or work problems we trust it so

43:36

instead of guessing it instead of

43:42

estimating manually we compute the result

43:48

and we trust because the problem is analytically solved

43:54

so that's a real life problem for example so for for that purpose it is evident that

44:01

for the real life problems we can use math and calculations

44:06

for uh for better for better environment for better

44:14

for better living or real life problems not only for mathematics itself

44:23

so at the beginning i said we trust computing

44:28

once we compute the results of this problem for example we trust

44:37

and for a good society

44:42

new trend is to generate data-driven culture today

44:48

data-driven culture is about replacing the gut feeling to make decisions with

44:54

facts and assumptions this is a definition that passes in many places on the internet

45:00

what is data-driven culture data-driven culture data-driven company culture data-driven school culture data-driven

45:06

universities data-driven schools data-driven professors data-driven students data-driven cars etc so that means

45:14

decisions are made by facts with data rather than

45:25

rather than just personal decisions

45:33

at the beginning only people will on the left

45:39

only people will decide data-driven culture collects data and

45:44

man machine works together machine intelligence on the other hand

45:50

has more automation it has more automation

45:56

meaning that decisions are decisions are

46:01

made by the machines so we have the video and after the video i will be ending the class

46:08

we have a meeting with the school today at the university

46:19

if you have any questions we have the five minutes video if you have any questions let me know

46:30

we all use smartphones but have you ever wondered how much data it generates in the form of texts phone calls emails

46:38

photos videos searches and music approximately 40 gigabytes of data

46:43

generated every month by a single smartphone user now imagine this number multiplied by 5

46:50

billion smartphone users that's a lot for our mind even process isn't it in

46:55

fact this amount of data is quite a lot for traditional computing systems to handle and this massive amount of data

47:02

is what we term as big data let's have a look at the data generated per minute on

47:08

the internet 2.1 million snaps are shared on snapchat 3.8 million search queries are made on

47:15

google one million people log on to facebook 4.5 million videos are watched

47:20

on youtube 188 million emails are sent that's a lot of data so how do you classify any data

47:28

as big data this is possible with the concept of five v's volume

47:34

velocity variety veracity and value

47:39

let us understand this with an example from the healthcare industry hospitals and clinics across the world generate

47:46

massive volumes of data 2314 exabytes of data are collected annually

47:52

in the form of patient records and test results all this data is generated at a very

47:57

high speed which attributes to the velocity of big data variety refers to the various data types

48:04

such as structured semi-structured and unstructured data examples include excel

48:10

records log files and x-ray images accuracy and trustworthiness of the

48:15

generated data is termed as veracity analyzing all this data will benefit the

48:21

medical sector by enabling faster disease detection better treatment and reduced cost

48:27

this is known as the value of big data but how do we store and process this big

48:33

data to do this job we have various frameworks such as cassandra hadoop and

48:39

spark let us take hadoop as an example and see how hadoop stores and processes

48:45

big data hadoop uses a distributed file system known as hadoop distributed file system

48:52

to store big data if you have a huge file your file will be broken down into smaller chunks and stored in various

48:58

machines only that when you break the file you also make copies of it which goes into

49:04

different nodes this way you store your big data in a distributed way and make sure that even if one machine fails your

49:11

data is safe on another map mapreduce technique is used to

49:16

process big data a lengthy task a is broken into smaller tasks

49:22

b c and d now instead of one machine three machines take up each task and complete

49:29

it in a parallel fashion and assemble the results at the end thanks to this the processing becomes easy and fast

49:37

this is known as parallel processing now that we have stored and processed

49:42

our big data we can analyze this data for numerous applications in games like halo 3 and call of duty

49:50

designers analyze user data to understand at which stage most of the users pause restart or quit playing

49:58

this insight can help them rework on the storyline of the game and improve the user experience

50:04

which in turn reduces the customer churn rate similarly big data also helped with

50:09

disaster management during hurricane sandy in 2012 it was used to gain a better understanding of the storm's

50:16

effect on the east coast of the u.s and necessary measures were taken it could predict the hurricane's landfall five

50:22

days in advance which wasn't possible earlier these are some of the clear indications of how valuable big data can

50:29

be once it is accurately processed and analyzed so here's a question for you which of

50:35

the following statements is not correct about hadoop distributed file system hdfs

50:41

a hdfs is the storage layer of hadoop b data gets stored in a distributed manner

50:48

in hdfs c hdfs performs parallel processing of

50:53

data d smaller chunks of data are stored on multiple data nodes in hdfs

51:00

give it a thought and leave your answers in the comment section below three lucky winners will receive amazon gift val

51:06

that's it so

51:12

that is the big data that we were talking on

51:17

uh if you have any questions let me know

51:24

next week we will be focusing on a different topic

51:31

turning up to recording