

Probability and Statistics with PYTHON lecture 7

- QUANTIFYING CHANCE
- Lecture starts with a problem whose solution is hidden and after few lectures it will be revealed
- Then there is a problem displayed

- A disease is prevalent in 0.2% of a population.
- We have a test that, given to a sick person, gives a +ve result 85% of the time.
- Of all the people ever tested, 8% were positive.

Q: If Nazo is tested and test comes back positive, what are the chances that she actually has the disease?

- ☐ 85% ☐ 77% ☐ 21% ☐ 2%

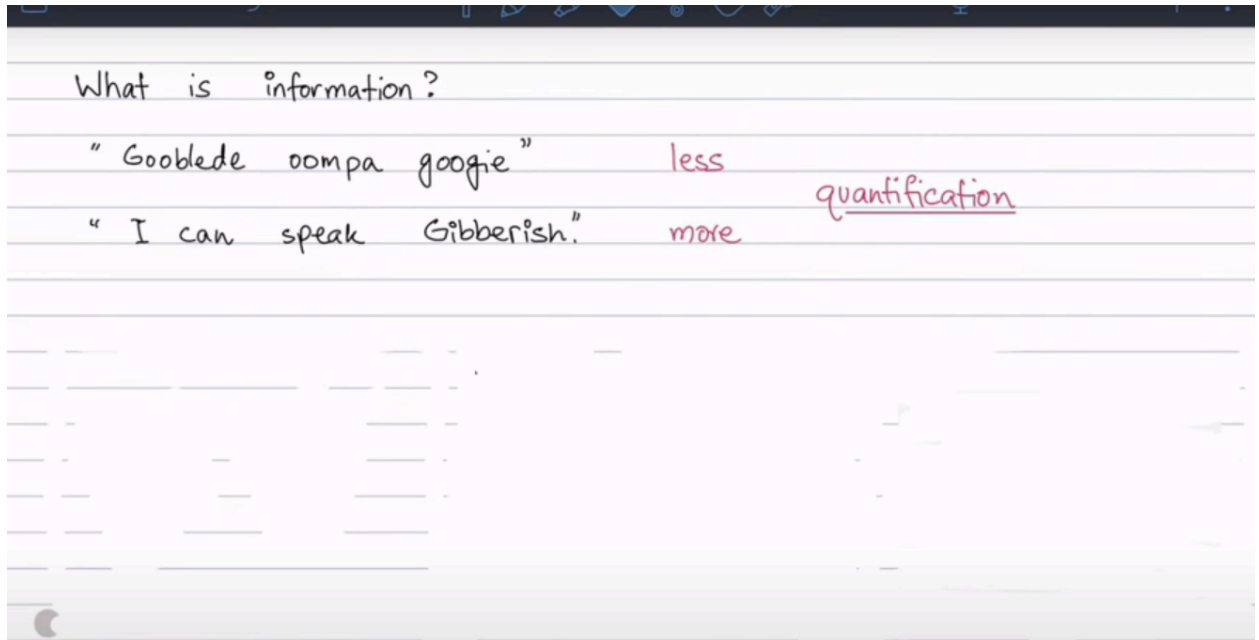
- The answer :

- A disease is prevalent in 0.2% of a population.
- We have a test that, given to a sick person, gives a +ve result 85% of the time.
- Of all the people ever tested, 8% were positive.

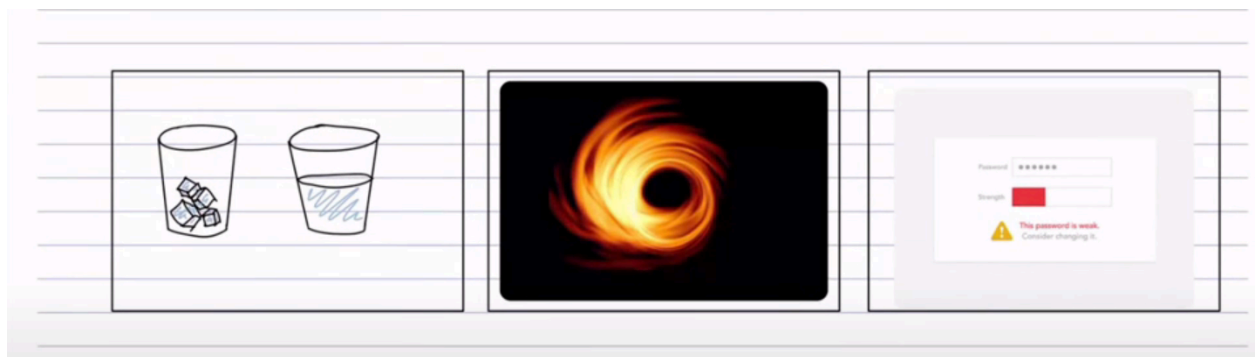
Q: If Nazo is tested and test comes back positive, what are the chances that she actually has the disease?

- ☐ 85% ☐ 77% ☐ 21% ☒ 2%

- One should be able to ask questions about success !
- Probability is not about computer science course it is about real life decisions



- We need to quantify things in order to understand them
 - As you can see in one we got less information in the other statement we got more information



- What do these things tell us ? , what information do we gather

Let's say you're doing cryptanalysis

Key 1 \rightarrow "Gooblede oompa googie" \rightarrow 29
✓ Key 2 \rightarrow "I can speak Gibberish." \rightarrow 1058

} Entropy

We can't eyeball 2.4 billion decryptions!

- In the above example :
- Let say the statement one has 29 information and statement two has 1058 information now we quantify it in such a way that we say ignore everything below 1000 and show me everything with above 1000 information , in that way we then select the useful information we can extract out from it
- This is called ENTROPY

— Video streaming — file size

— Password strength

— Huffman's codes — compression

— Computer vision

— Machine learning — throw away data/information

— Bits — foundations of all computation

- Higher entropy of video streaming means more data
- Password strength with lower entropy means weaker password
- Huffman codes , entropy is based on it
- Also used in computer vision
- Also in machine learning
- Bits
 - This is the foundation of entropy , bits is the unit of entropy actually
 - With the help of bits we can decide the entropy of anything
 - Basically tell how much information is stored in there

- The above we will cover in this course in coming lectures