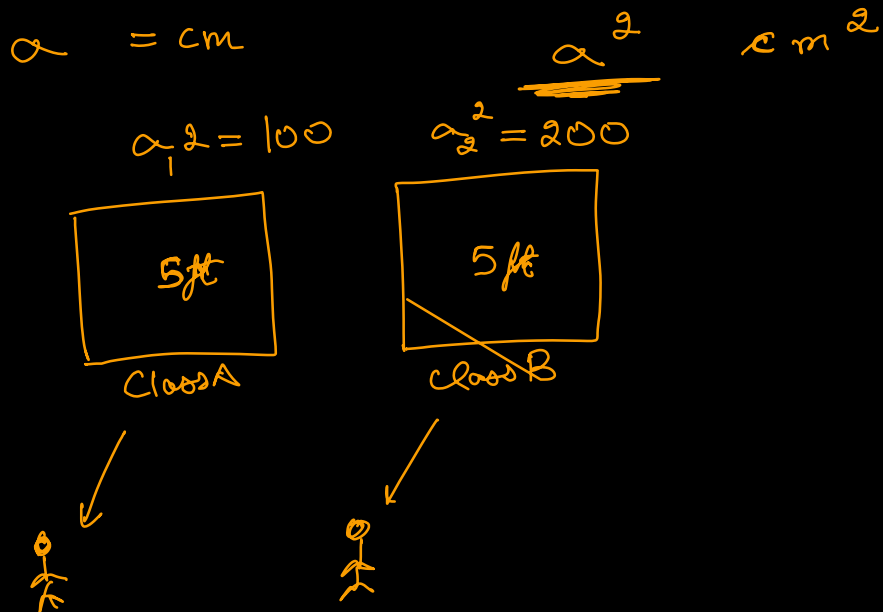


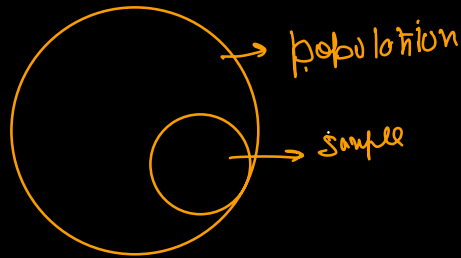
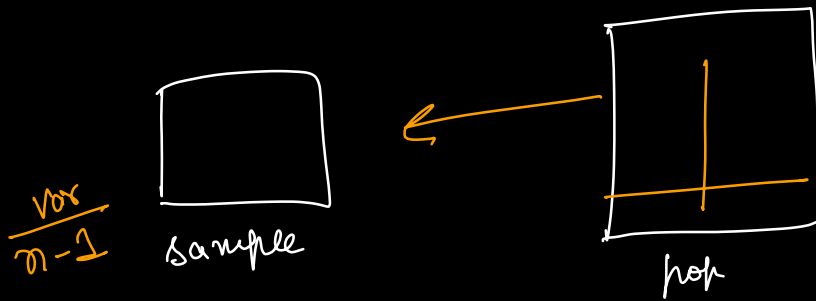
ML Bootcamp \Rightarrow Extra Class

MY DOUBTS :

- 1) why we need both std. deviation and variance \longrightarrow both tell how much values are away from mean
- 2) variance error
- 3) why $n-1$ in variance & std deviation formulae
- 4) Random variable
- 5) statistic(singular) & statistic(plural)
- 6) coefficient of variation



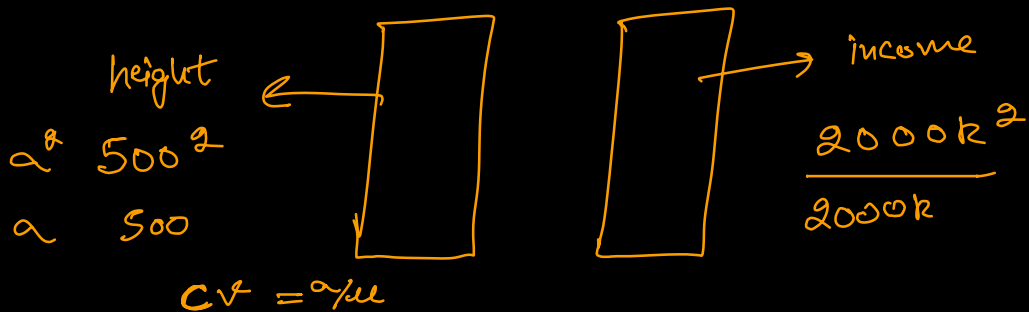
$n-1$



Statistic vs parameter

\bar{x} $\bar{\mu}$
 4.9 5.2 5

$$CV = \frac{\sigma}{\mu} \quad \frac{s}{x}$$



Pearson Coefficient r^2



Data Science Libraries

Agenda :-

1. Why libraries are required
2. What all DS libraries
3. How to Install them
4. Import & use
5. ML algo \Rightarrow code + library.

Beginners

\Downarrow

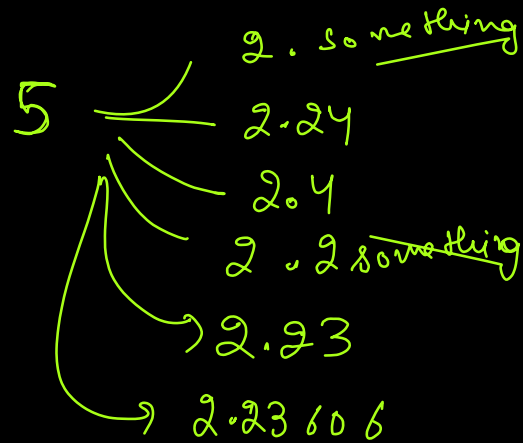
Intermediate

~~Complex~~

Linear Regression?

We will never be
using majorly
our own
code.

Sqrt of a number



π α

we always learn & then

implement most things.

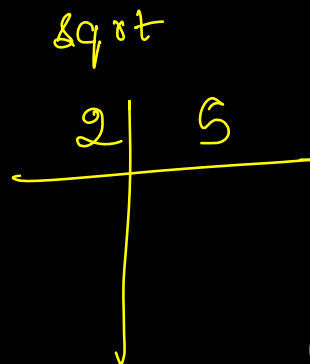
$$3335 \times 3387 =$$

$$\begin{array}{l} 4+5 \rightarrow 90 \\ 35 \times 35 = 1225 \\ \leftarrow 95 \times 45 = 2025 \end{array}$$

\Rightarrow very complex

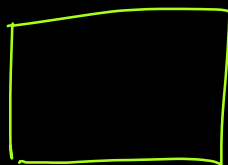
\Rightarrow implem. can be made much more efficient

\Rightarrow our implem. might not be the best

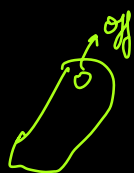


\Rightarrow implem. in code normally

\Rightarrow divide & conquer greedy algorithm

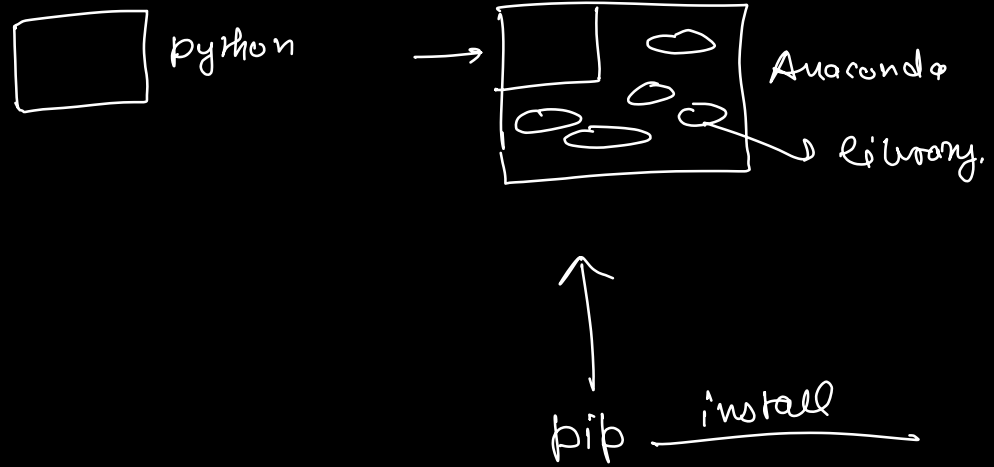


it works



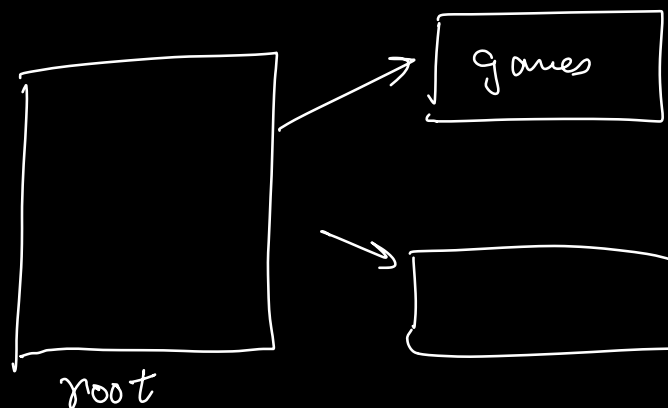
time plays a very critical role

$\log t ()$ 1000gms $\xrightarrow{\quad} 5min \rightarrow$
 \downarrow
 1 sec ✓



`pip install <library-name>`

multiple python installation on your
system via anaconda



Libraries
installation
Environment

Data Science Problem/Framework

1. Get data
2. Work to get statistics
3. Draw diagram to show the insights
4. ML model training
5. result.

hr play previous

--	--	--	--

Get your data



store our data



excel, csv/spreadsheet/JSON

csv

1, 2, 3
↑ ↑



numpy & pandas

to load data into python
to work on

70 65-70

not sure



we do use

them

& calculate

statistical

A/B by,

stats

⇒ Statistics

numpy

scipy

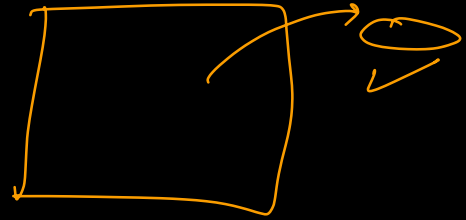
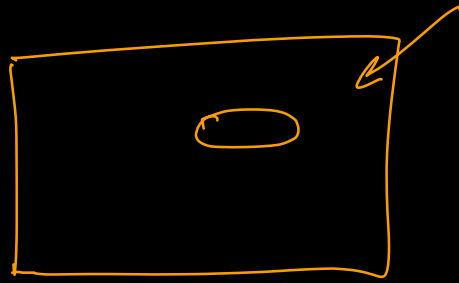
math

Visualization

import matplotlib

seaborn

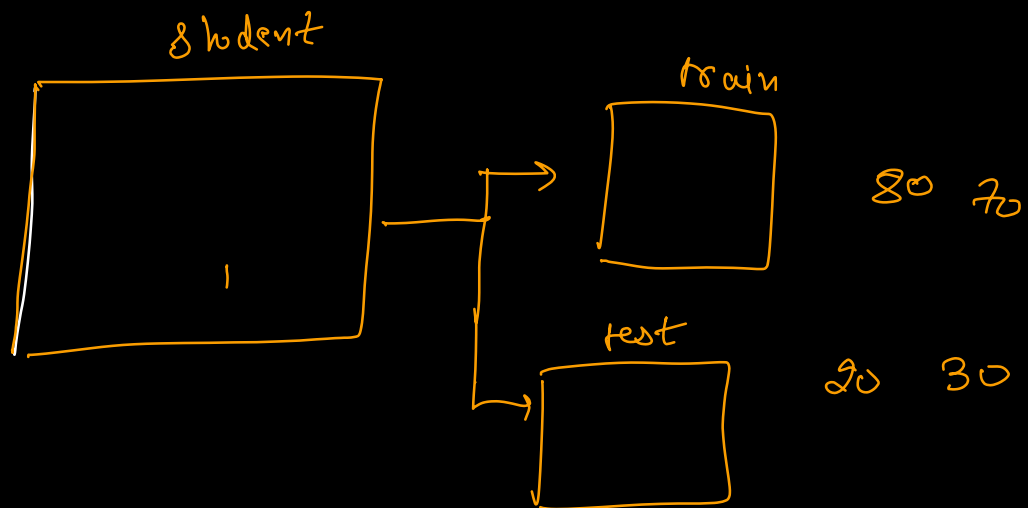
plotly



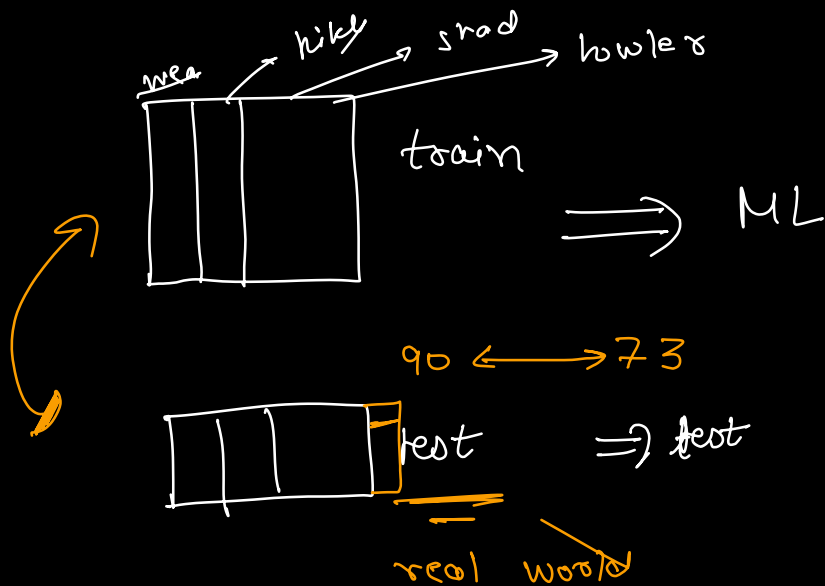
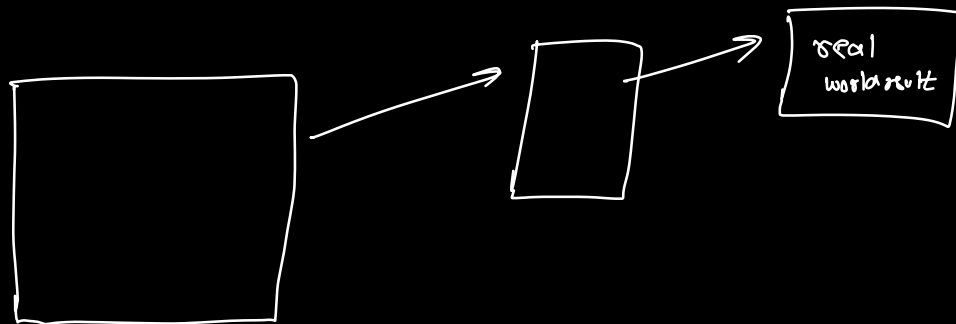
2 things

\Rightarrow we should be able to predict future
with a little good accuracy

\Rightarrow we cannot predict future as completely



Q \Rightarrow Someone ask me or you, to
make a team for T20 world
cup.



90 - 20
90 - 20
90 - 20

We have libraries to work on splitting
as well as for apply ML
models.

Scikit learn

Keras

Pytorch

Tensorflow

live

prediction

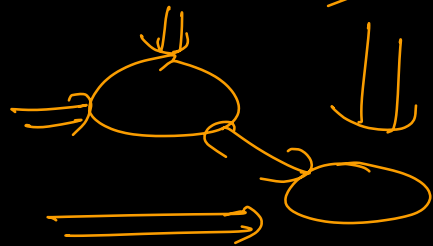
Live

Not

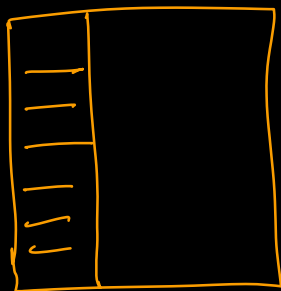
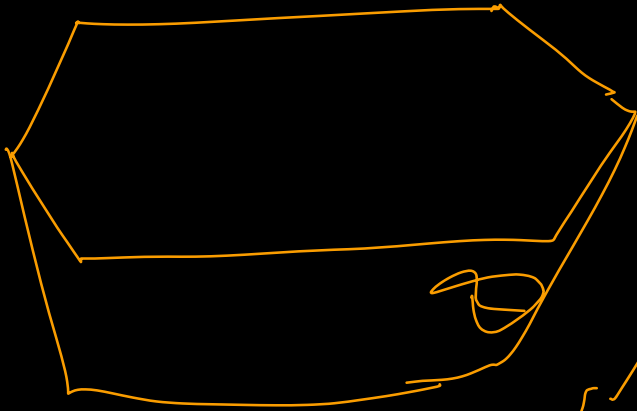
Actual

Live

Not live



london
↓
ny



100

Confusion matrix

Predicted \ Actual		<u>Survived</u>	Not Survived
Survived			5
Not Survived			