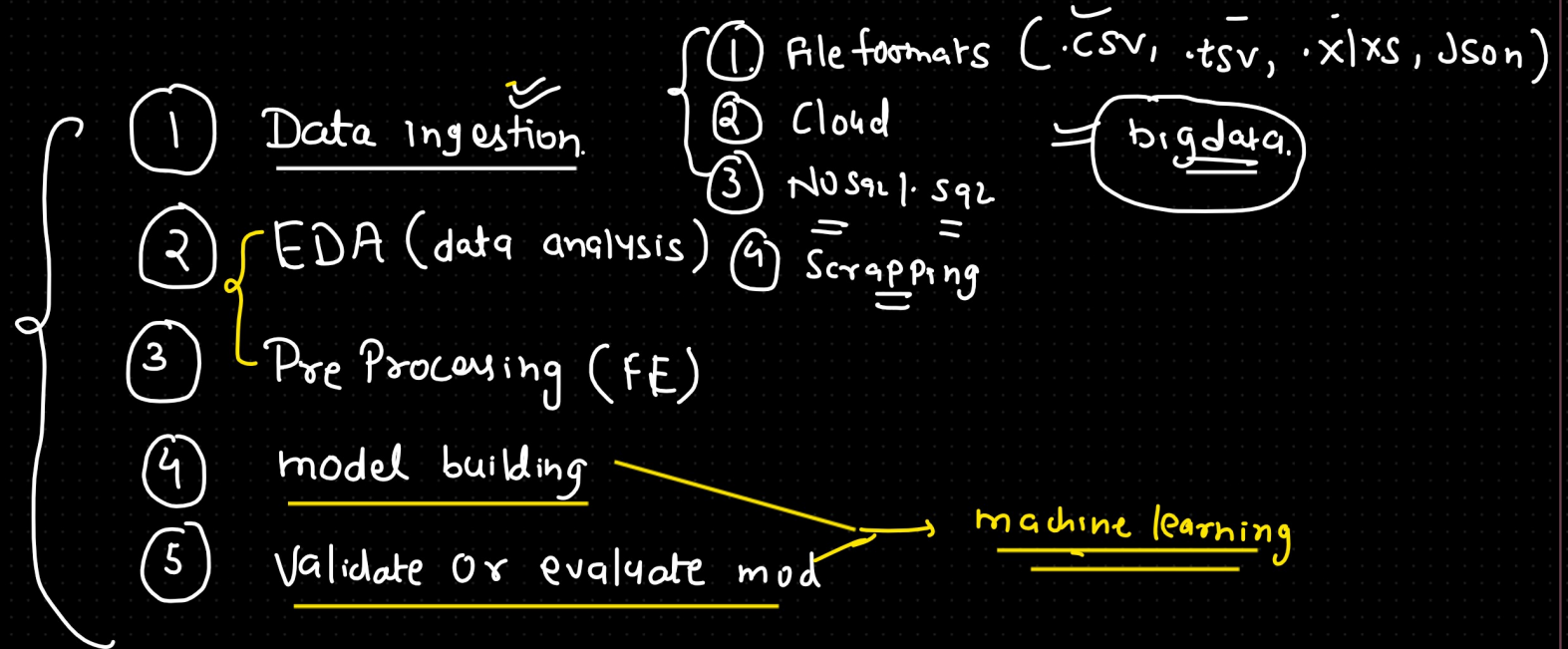


EDA

Core ML steps



EDA + PP(FE)

EDA
(analysis)

- ⇒
- ① Profile of the data.
 - ② Stats based analysis.
 - ③ Graph based analysis.

Collector \Rightarrow analysis (EDA)

Raw.



① Profile of the data.

maybe or it is not
correct in terms of
content

additional.

Data cleaning

Rows
cols.
missing
outlier
size
duplicate
dtype
num, cat-
RAM
...

② stats based analysis

Var

Std

Cov

Corr

Chi-square test

t, z, f - test

Skewness

mean, median, mode

Range, Percentile

③ Graph based analysis (you can make statistical conclusion)

- ① Box Plot-
- ② Scatter plot
- ③ KDE
- ④ Count Plot-
- ⑤ histogram
- ⑥ heatmap

① collected Raw data [playstore google playstore]

② EDA

↳ Profile of the data (Cleaning)

↳ stats based analysis (pymon)

↳ Graph based analysis



③ Feature eng