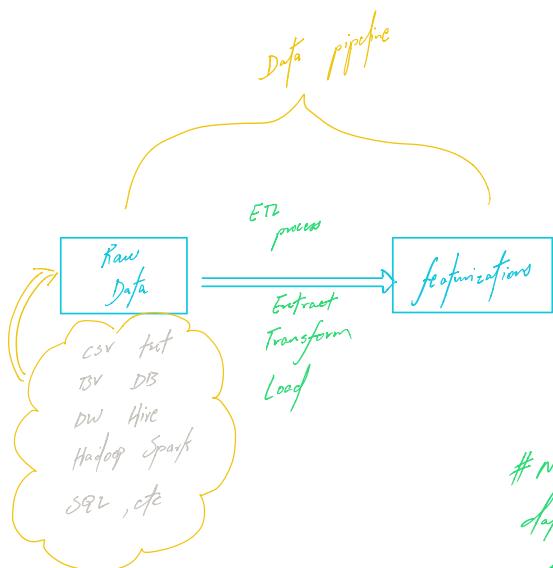


# stuff we need:  
 ML models  
 in production      # challenges  
 env.                # design choices  
 # more often than the current

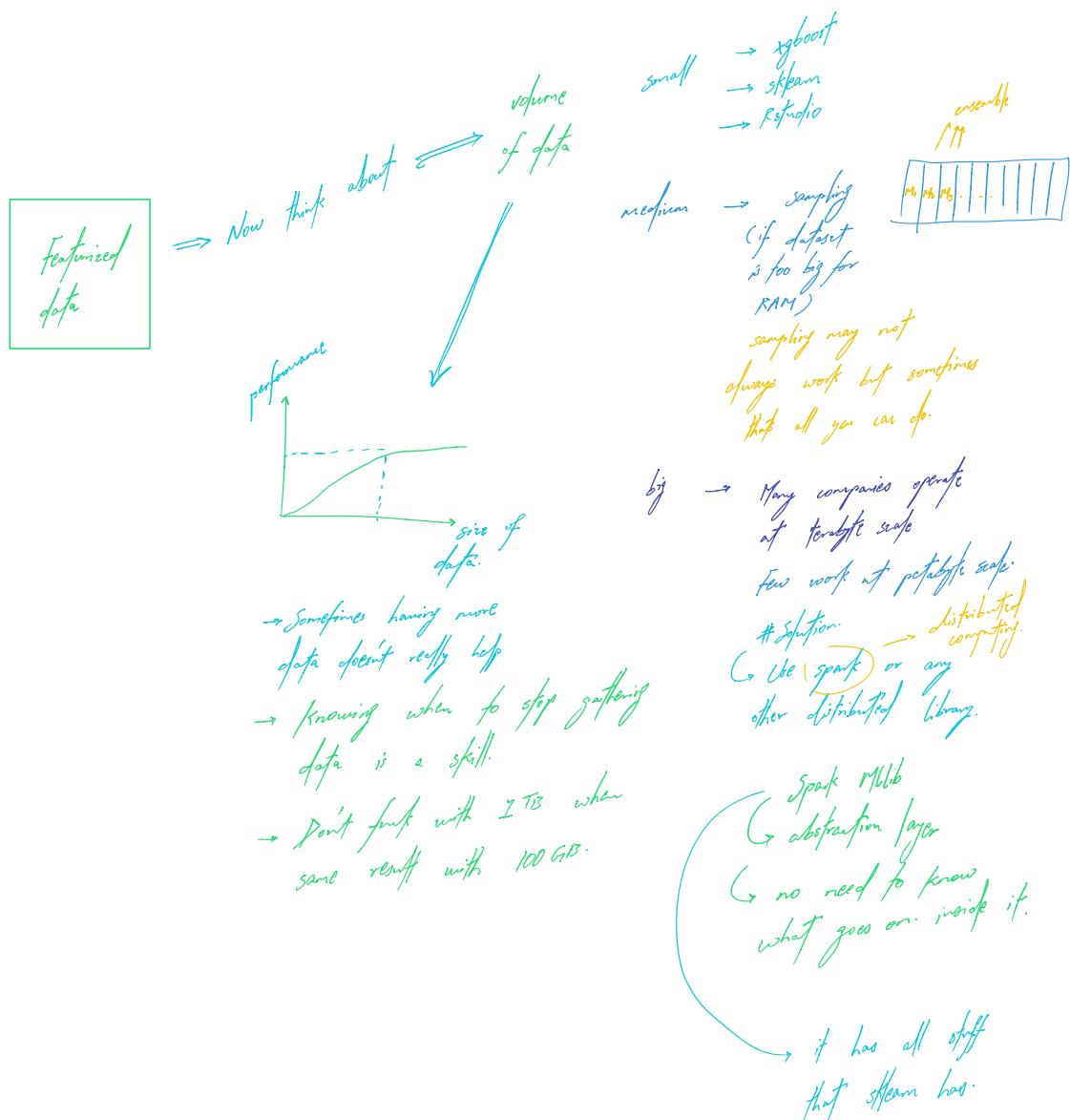
# let's assume we already  
 have a model.  
 ↳ data collection  
 ↳ feature engineering & EDA  
 ↳ training  
 ↳ testing  
 ↓  
 [M] → ML model  
 Lots assume everything is in place, we just need to release this. → green light from manager & client, ok staff.

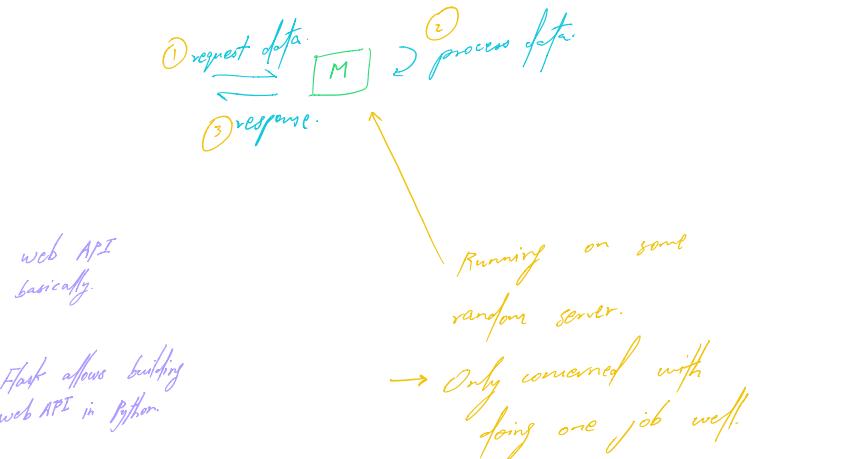


# normally we have  
 data engineers that  
 make the ETL pipeline  
 # but if company can't  
 hire them, DS and us

doing this process

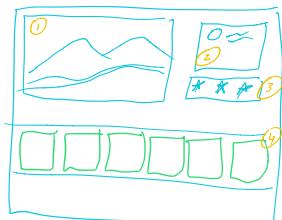
- \* What's the frequency of running this ETL pipeline?
- \* Big sales company might run this every hour.
- \* Some might want data every day.





- # Benefit of API
  - ↳ The client and server can be running on completely different languages and it would not matter
  - ↳ They have a common language to talk with each other → <sup>JSON</sup><sub>most popular</sub>

ML use case → { "datapoint":  
 { "fr": 95,  
 "fc": 48  
 } } }



①, ②, ③, ④ all can be different APIs. That's what happens in real life.

People know MLP CNN RNN by just easier when asked to build a recommender or chatbot or afae, etc.

THAT'S WHAT FUCKING HAPPENS WITH ME !!

We have to go through lots of case studies to develop this skill

↓  
breaking down a  
big problem into  
small palpable chunks

# Extracting text from image:

① If we know where the text  
will be. for eg → like a form, etc.  
then it's easy. using RCNN

② If it's free form, it's hard.

# If it's a PDF, there already tons of  
techniques out there. No need to re-invent the wheel.

# Use Amazon Sagemaker for deployment when  
you've already invested in EC2 and S3

core components

# For cloud, basic building blocks  
remain the same, the framework  
keeps on changing

# some storage  
# some compute  
# some other stuff

# use profile library in Python  
to sacrifice objects → doesn't matter what  
objects they are.

- # Haiku & graph
- are not important to learn for interviews.
- # People with good foundations can always  
learn and catch up.
- \* It doesn't hurt to learn though, IN CASE  
WE HAVE TIME.

# sepium → gets the help for us  
& then we can reason the  
fact out of it.  
# websites point to where people  
swipe data, they look us have  
to be careful.

# They kinda don't give a fuck about  
the entire bits of knowledge we gather.

ETL pipelines can be  
read from → "streaming off the pipeline"  
↳ only do it if needed

If's not about GDP or Awd or  
Acme, it's about how we use them.

[  
↳ if some tool makes  
your life simple, fucking use it...