Berkeley Engineering | BerkeleyHaas

PROFESSIONAL CERTIFICATE IN MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

Office Hour with Vikesh Koul

Sep 03rd, 2024 at 8pm UTC

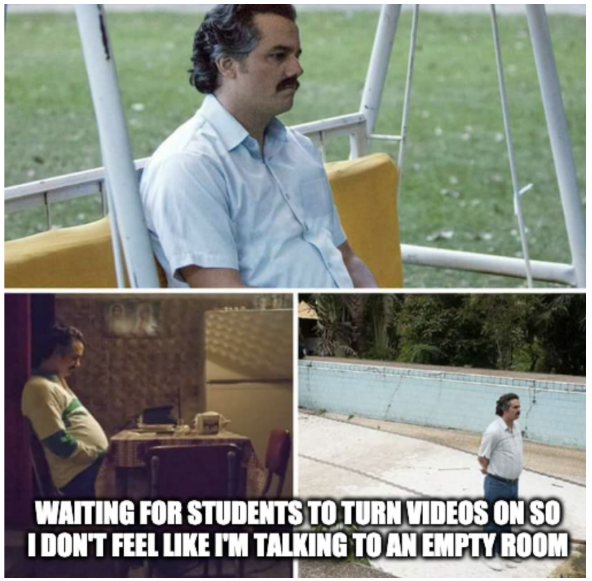
Note on my Office Hours

- 🤗Safe space to ask questions/doubts. It is your place to ask the simplest questions/doubts without hesitation. The session is about you.

- 🗒️Mostly focused on Case Studies. I usually like to take up a problem and solve it together with the learners.
  - Unresolved business situations, and would require your active engagement
  - Be a decision maker and develop decision orientation
  - More sticky, factual knowledge decays rapidly
- 🏃 Pace of the OH can vary, since you all have different backgrounds, it may feel slow or fast sometime.

- 💬 Please provide regular feedback; during sessions and surveys
  - Critical or good feedback, both are appreciated
  - I will incorporate the feedback as soon as possible

- 👩‍💻We are partners in learning for next 6 months. I will appreciate if you keep your videos on
  - Visual feedback for me
  - Better learning environment and fun too
  - Better sense of your peers!



You don’t need more time, you need more focus!

Studies shows that even having a phone in the room impacts your focus and concentration level

🔗📚Additional Resources

UC Berkeley: PC-MLAI

[Draft version 0.01]

Additional Resources

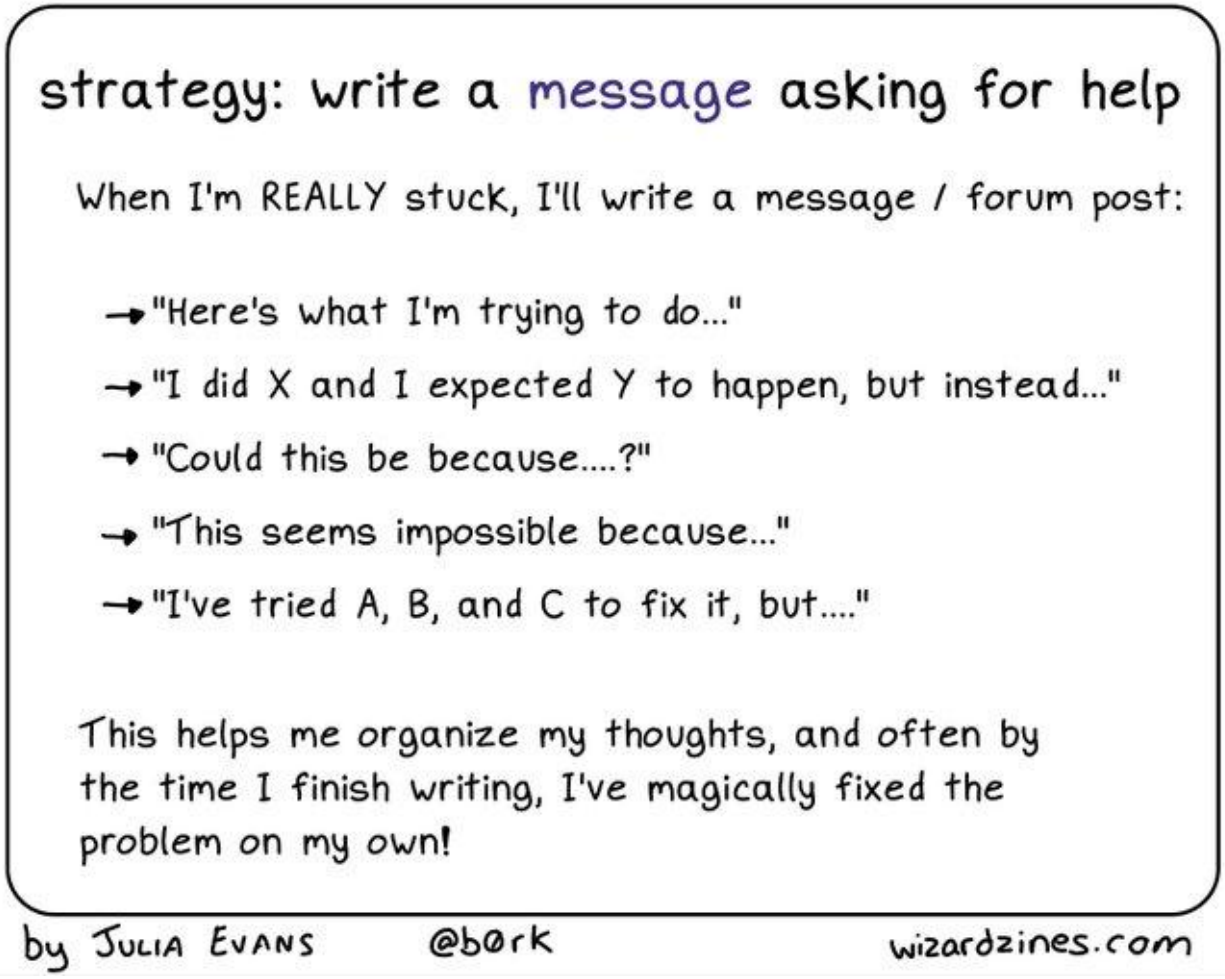
📢📢This is additional material if you want to go deep/explore into the topics covered in the course. It is optional and not required to complete the course.

📍This will be continuously updated

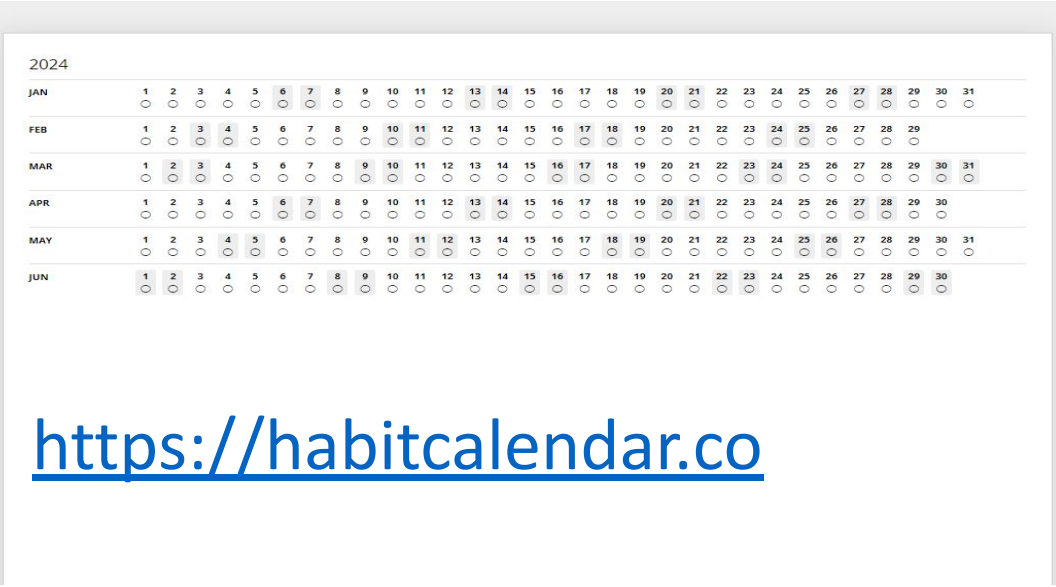
💬Please use the comment feature to suggest additional resource or any feedback

WK Num	Week Name	Article(s)	Book Chapter	Video
0	Program Orientation	• The Blank Sheet: The Note-Taking System You Never Learned • Intro to Google Colab • Markdown Guide	• Python Data Types • Python Data Structure • Python Reference Guide	Python Programming Python for Beginners - Learn Python in 1... Why Python is huge in finance? by Dani...
1	Introduction to Machine Learning	• Common Probability Distributions • Conditional probability explained visually • https://datavizcatalogue.co	• Pandas Overview • Pandas inbuilt Plotting	Problem on Baye's Theorem | Problem ... The Incredible Story Of Randomness Central limit theorem But what is the Central Limit Theorem?

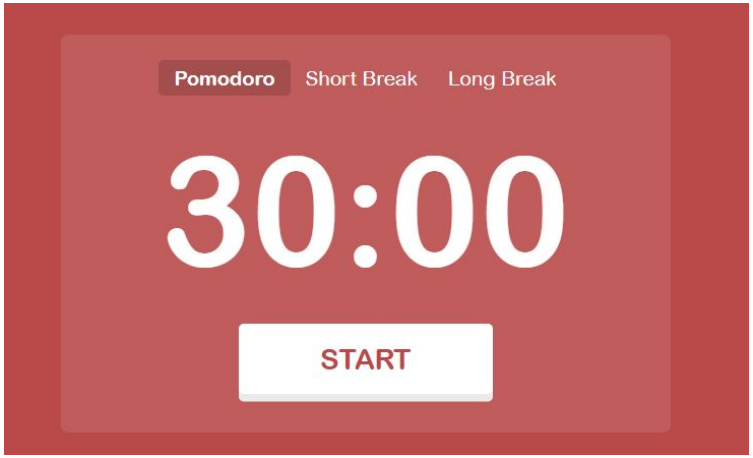
Asking for help in Tickets



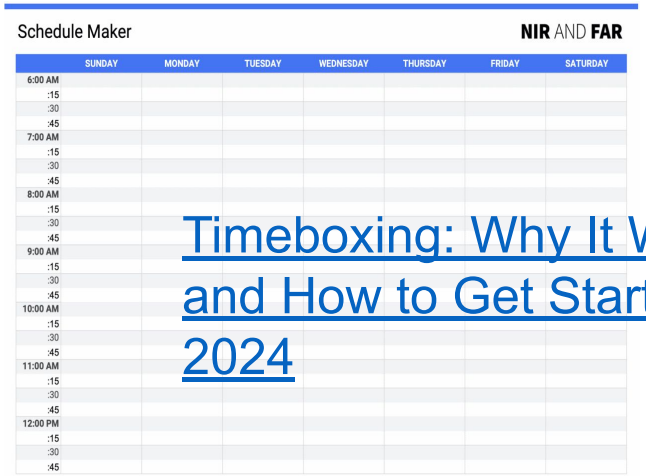
Track your time and improve your note taking



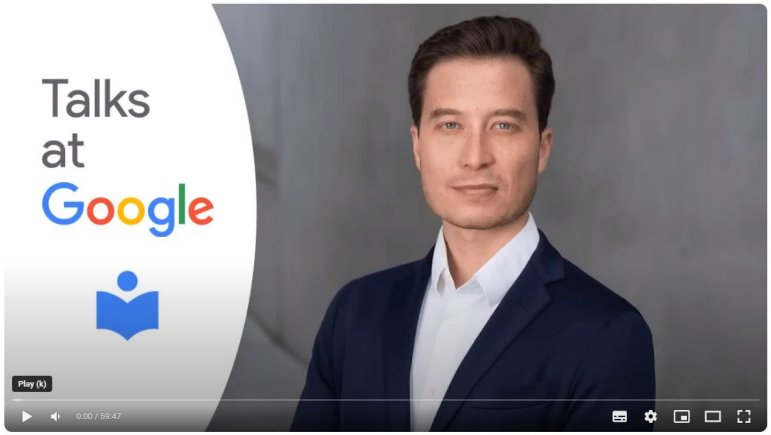
https://habitcalendar.co



https://pomofocus.io/



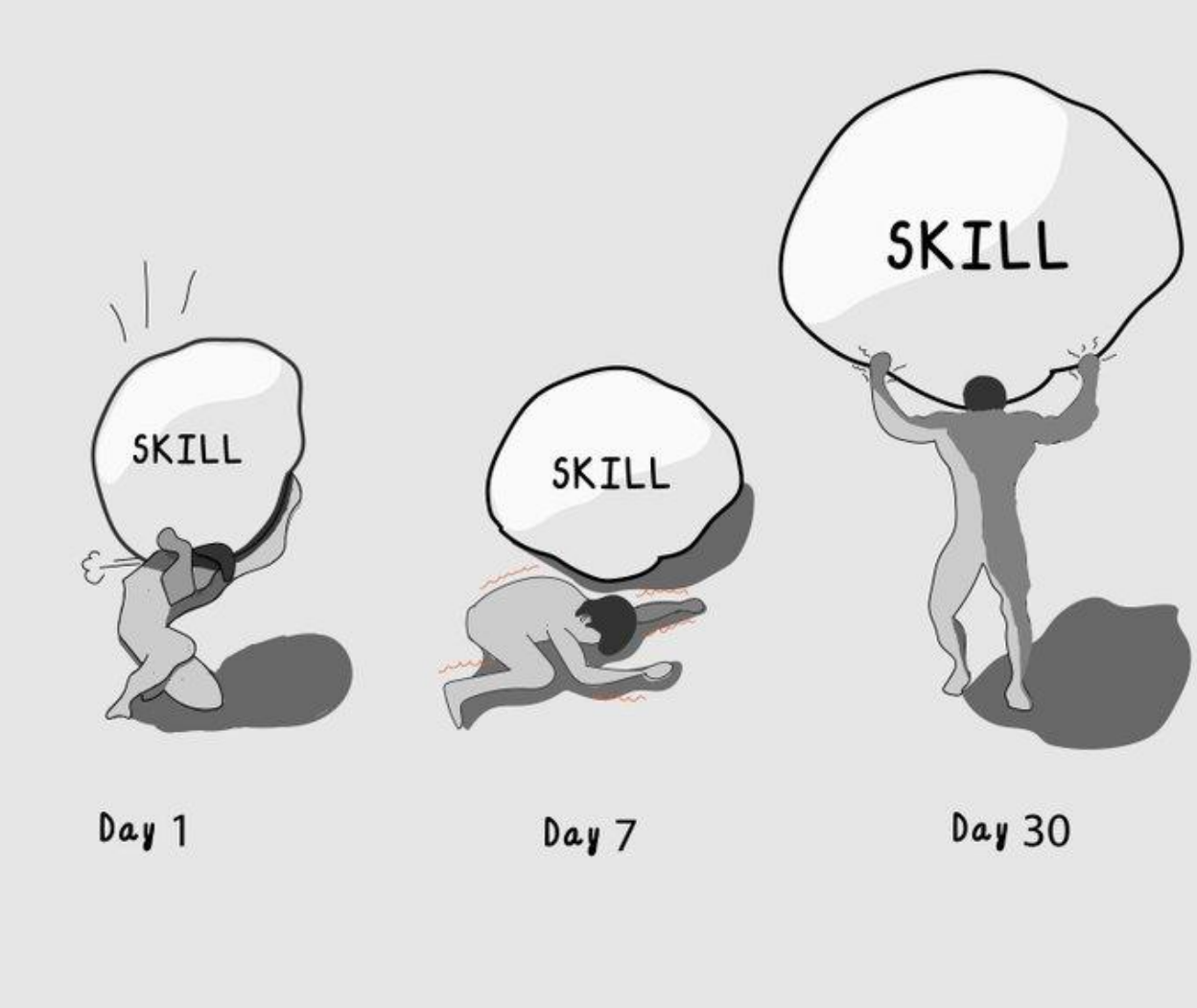
Timeboxing: Why It Works and How to Get Started in 2024



Tiago Forte | Building a Second Brain | Talks at Google

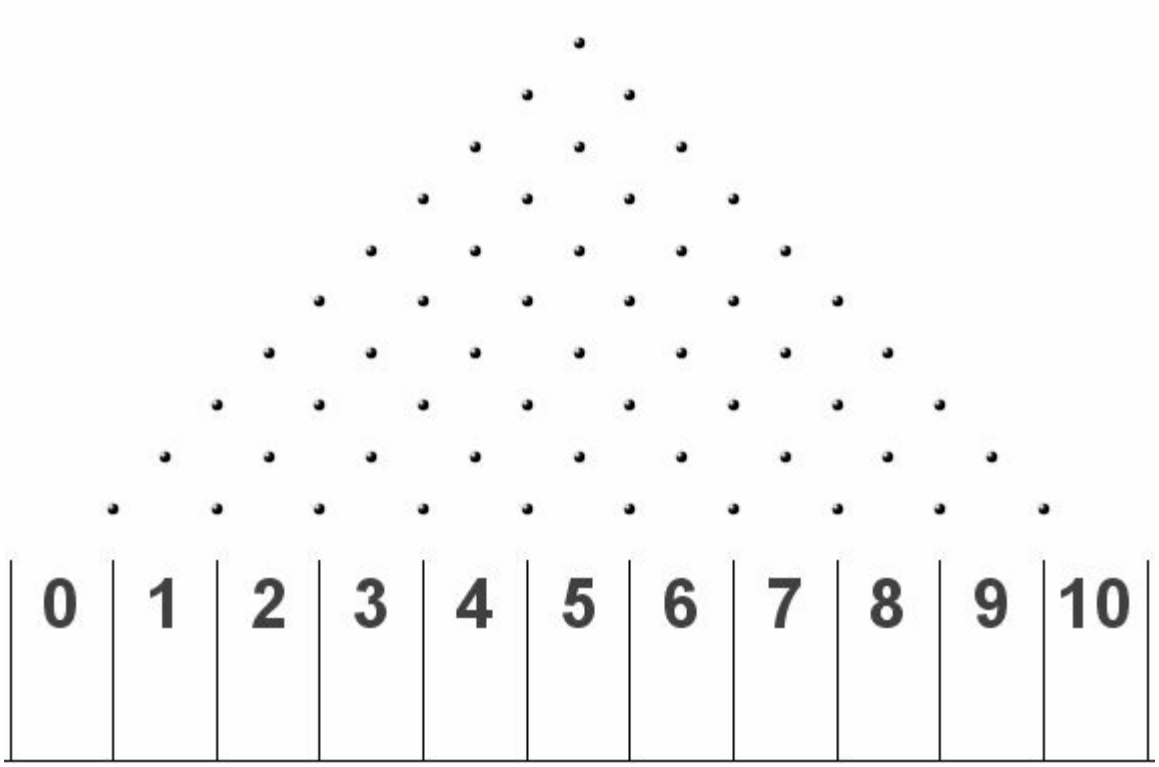
Building a Second brain

Remember it will become easy



Distributions

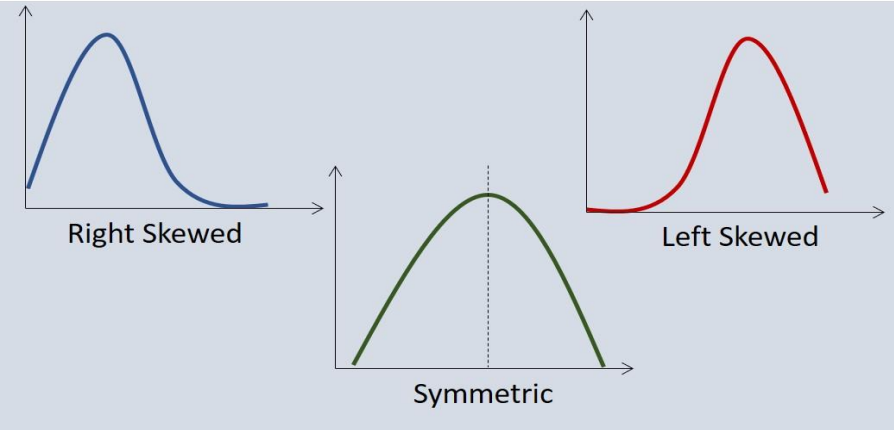
Random events give rise to a distribution



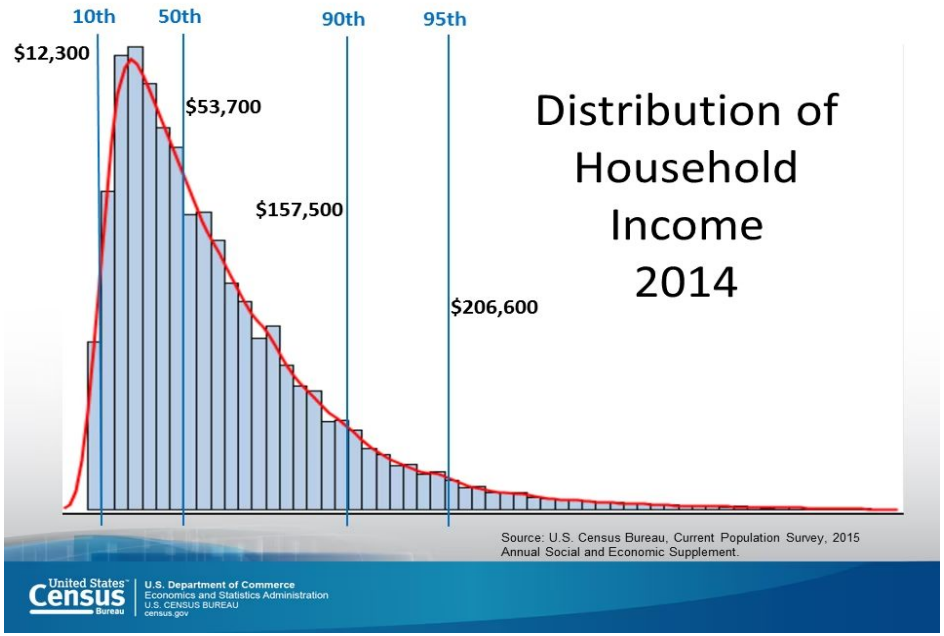
Distributions are relatively easy to tackle mathematically

Distributions help us explain real world events

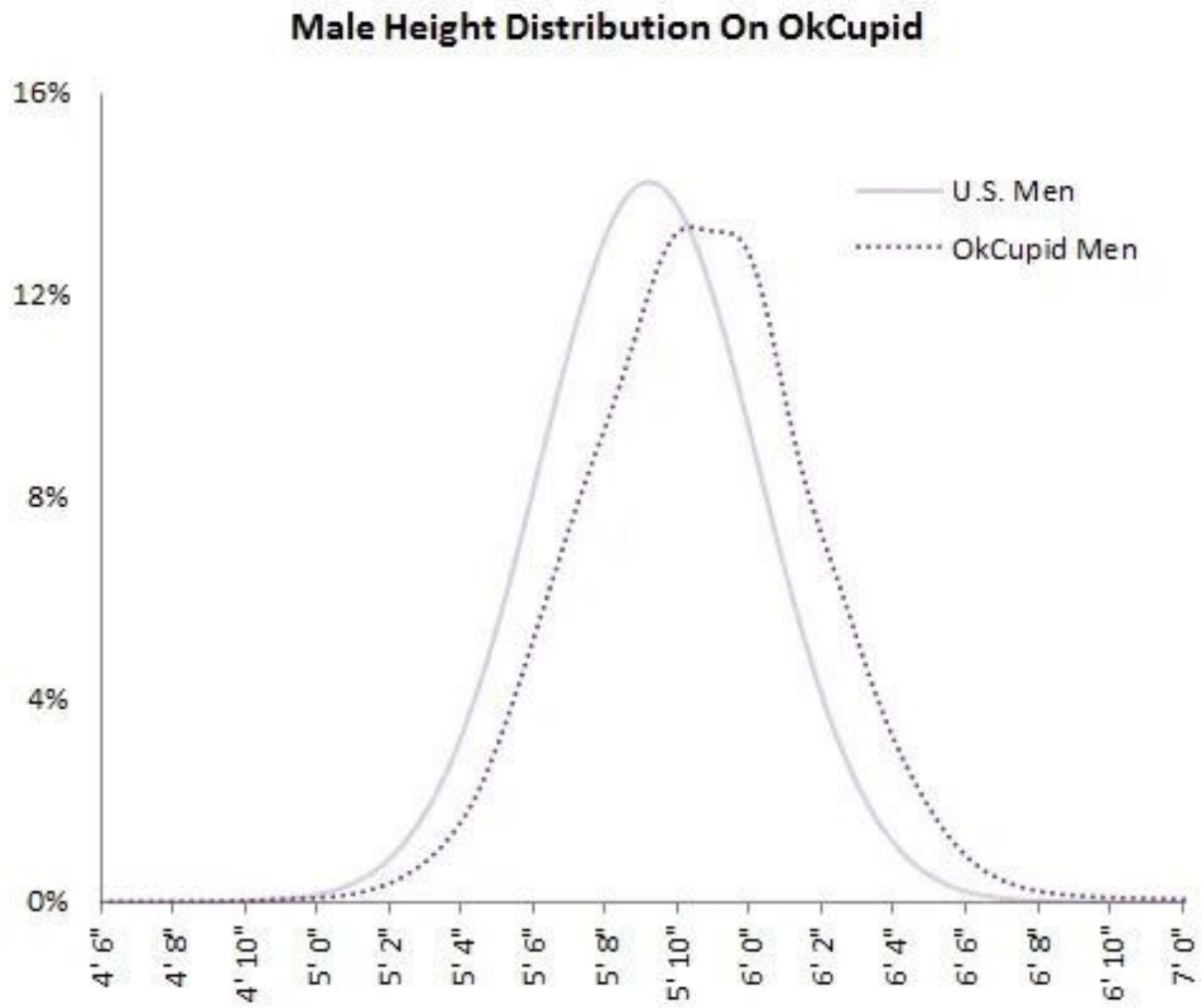
- Distributions help us to define events in the world
- Random events can be mapped to a distribution and that can be used to manage uncertainty
- Each distribution can be described using mean, median, and standard deviation, among other things
  - You can have skewness in your distribution which is also helpful to explain events



"Skew where there is few"

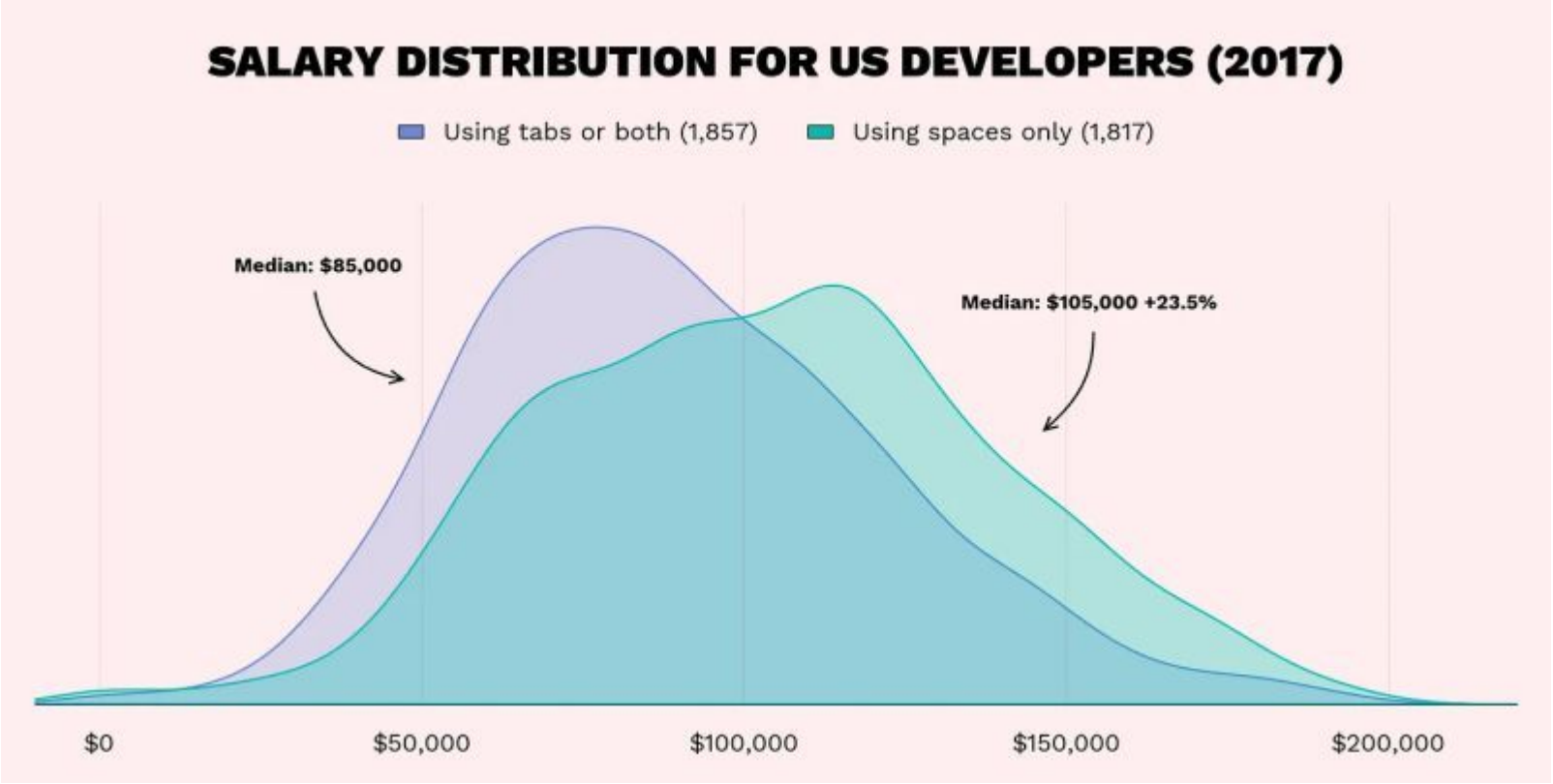


What's happening here?



Source: OKcupid Blog

What's happening here?

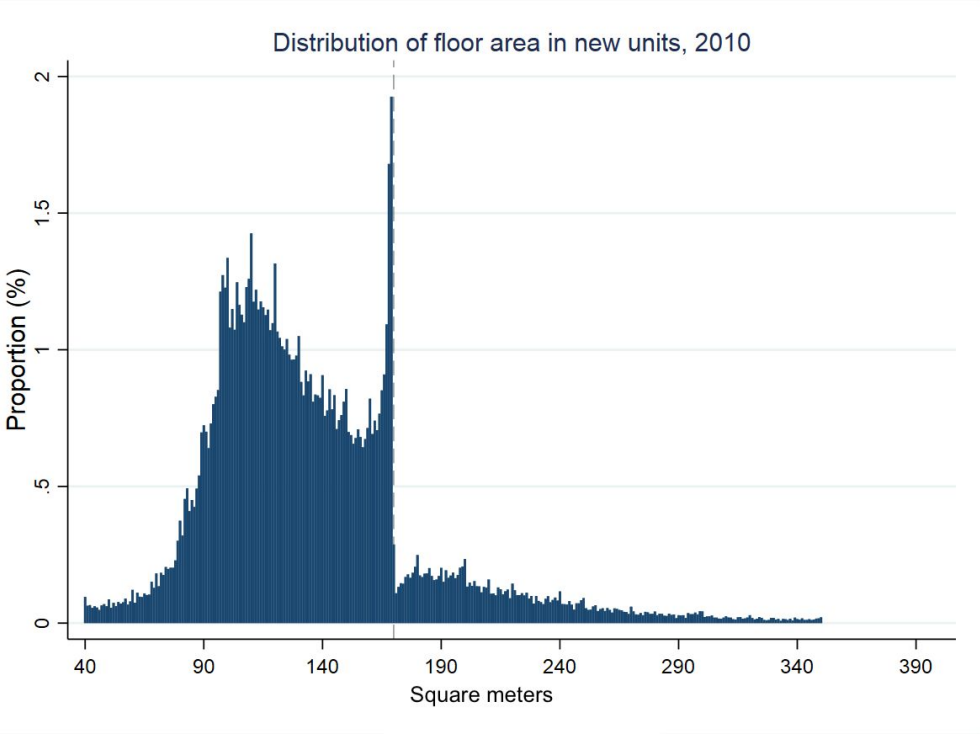


Spaces vs. Tabs: The Surprising Impact on Programmer Salaries |

Also see this! Tabs versus Spaces

Source: Levelup Blog

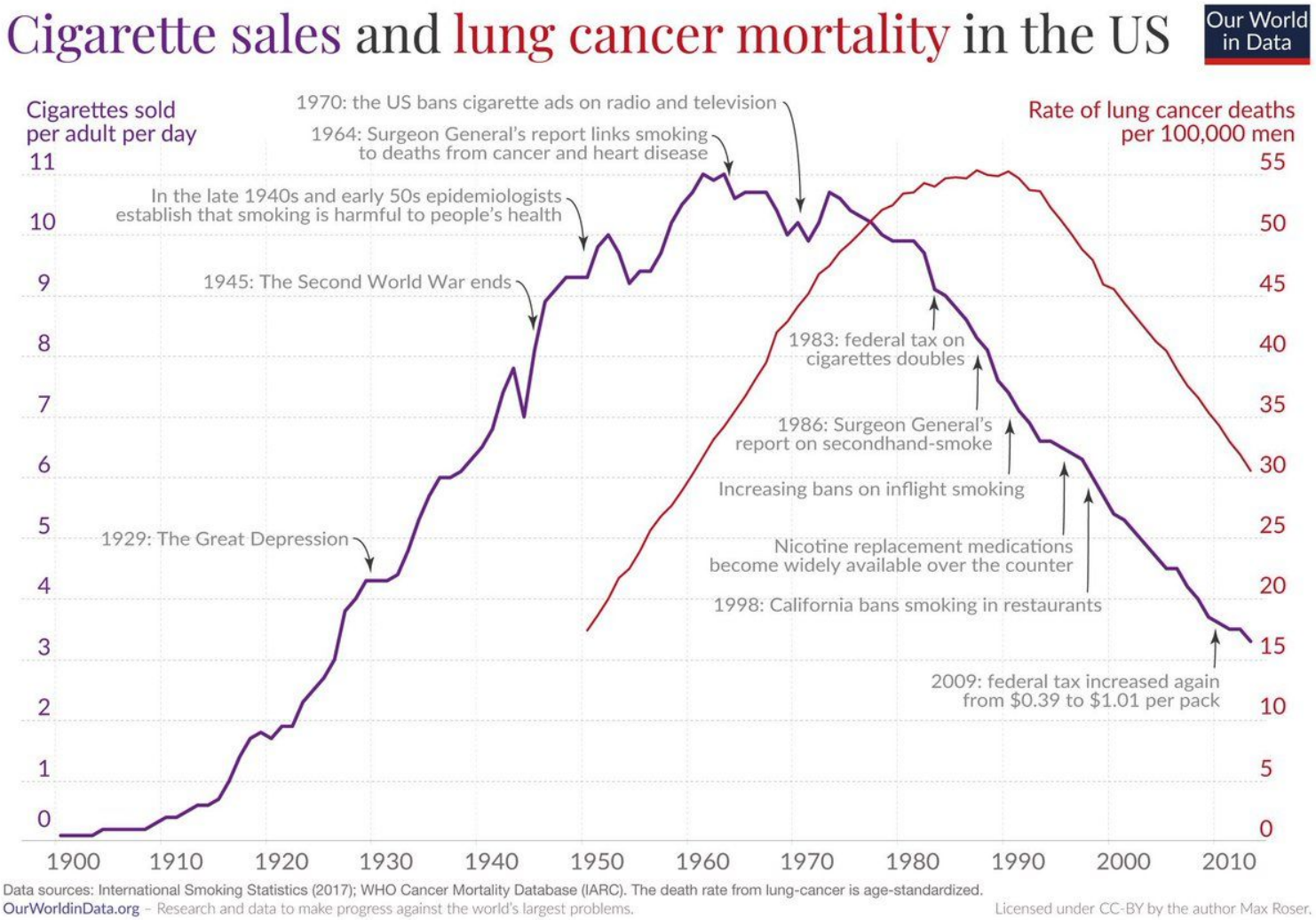
What's happening here?



House sizes in France

Source

Cigarette sales and death rate are same curve, shifted by 25 years

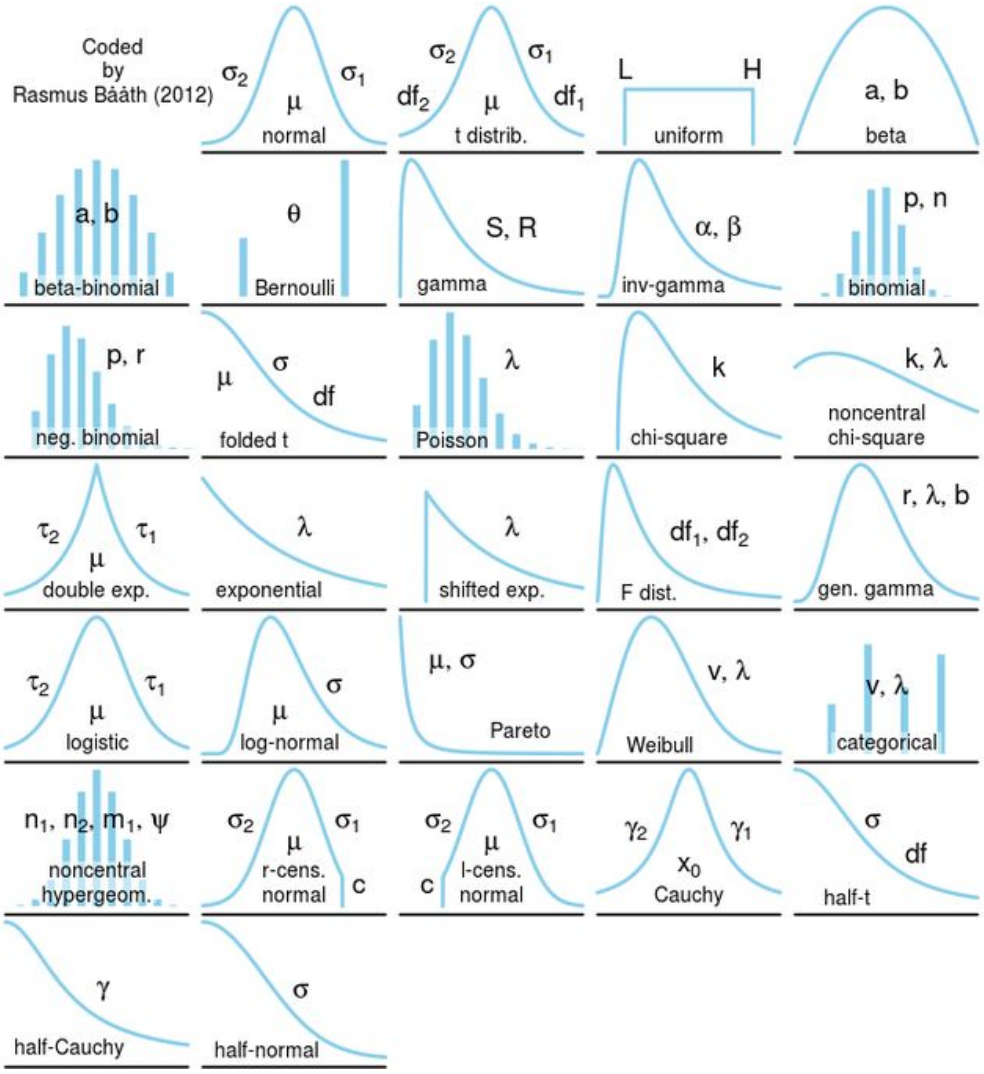


Which distribution to choose?

- Source of distributions:
  - **Empirical** - Based on past data
  - **Theoretical**- based on a mathematical form
- If you wish to model a phenomena in your business, which one should you use?
- Empirical distributions can get influenced by outliers
- **Theoretical distributions**:
  - allow for more robust modeling
  - already have quick shortcuts in how to characterise it
  - Look for distributions that fit the data : Like the customer service calls follow a Poisson distribution

Common Theoretical Distributions

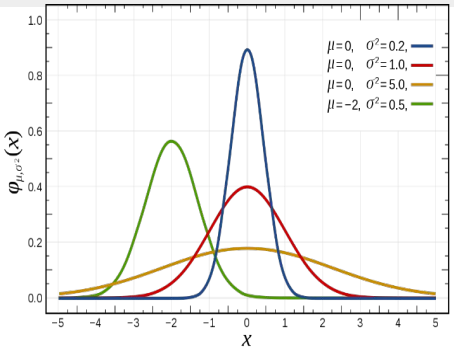
Probability distributions come in many shapes with different characteristics defined by the ***mean***, ***standard deviation***, ***degree of freedom***, ***skewness***, **and** ***kurtosis***.



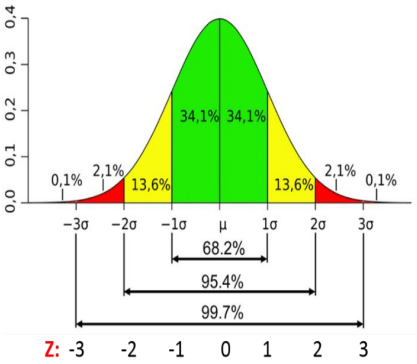
Know this guy?

Normal distribution is the most important distribution

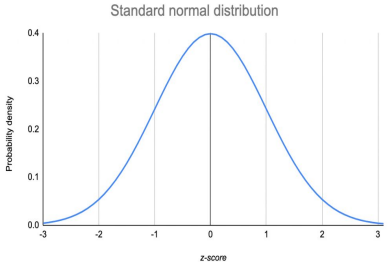
- A normal distribution can be expressed using two parameters; mean and std. deviation
- Mean determines the peak of the distribution and std. deviation determines the spread



- A useful feature is the 68-95-99 rule:
  - 68% of the values lie within 1sd
  - 95% of values lie within 2sd
  - 99% of values lie within 3sd
  - Useful in a business context



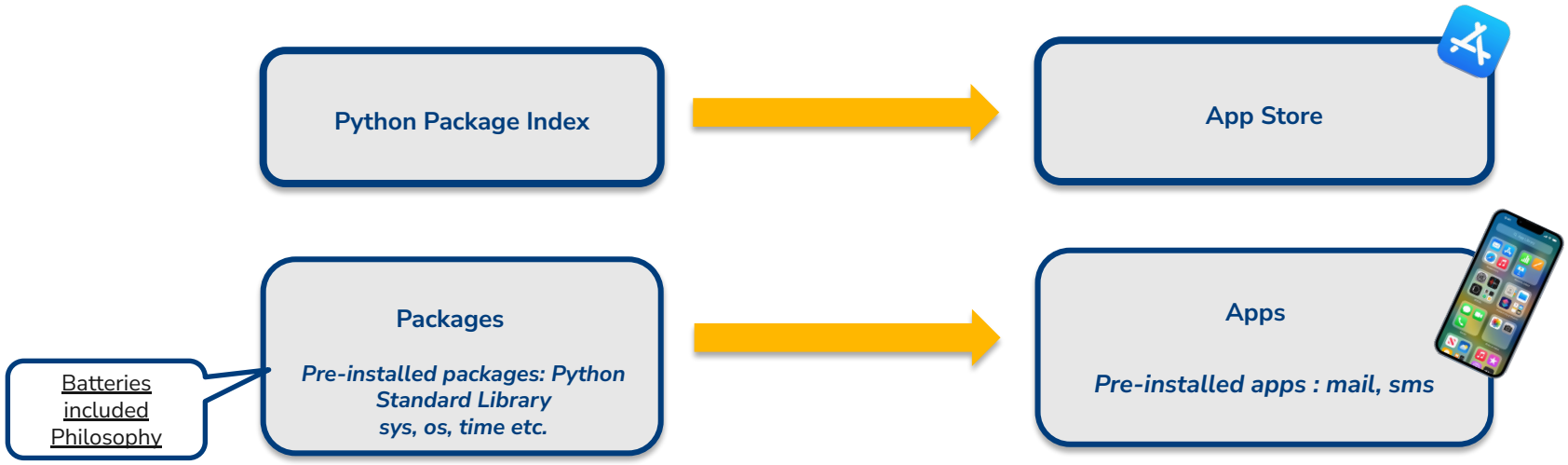
- Standard normal distribution is a special that has a
  - mean=0
  - standard deviation = 1
- Very useful for answering questions about probability



Python Overview

Understanding Python Ecosystem via a Smartphone analogy

Python is a platform for which packages can be built; similar to smart phones



Packages is what makes Python very powerful, like the apps makes your smartphone powerful and useful

Important Packages

Data Manipulation

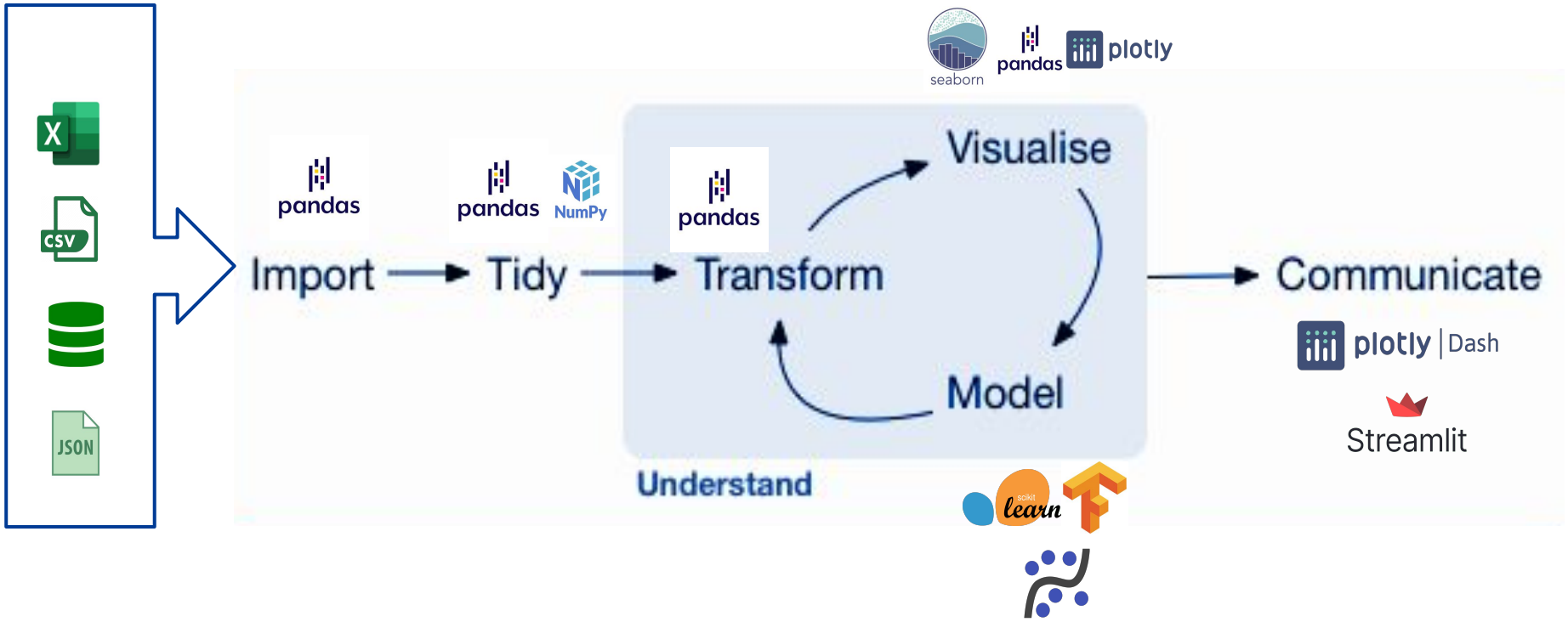
Data Viz

Machine Learning

TensorFlow

Dashboard

Data Science Workflow



Time to practice! ⚡👩‍💻