

DS2000

3/28- Tues.

Admin

- Hw7 due Fri 9pm
- Hw8 (re ad) opens Fri 3/31 due 4/7
 - upload Hw8
 - new score on original Hw
- no class on Fri! 3/31

→ no late submissions

Agenda

1. Recommendation Systems
2. Design choices
3. Python

1. Recommendation Systems

↳ Data Science: the more data, the better
the recommendation

ex: Netflix - what show to watch next *

Instagram - accounts to follow

Amazon - products to buy

Indeed - jobs to apply for

Spotify - playlists to listen to

All use different algorithms to make recs

All use lots of data

All try to mimic real life

DS2000 Rec. System

Context: TV Shows

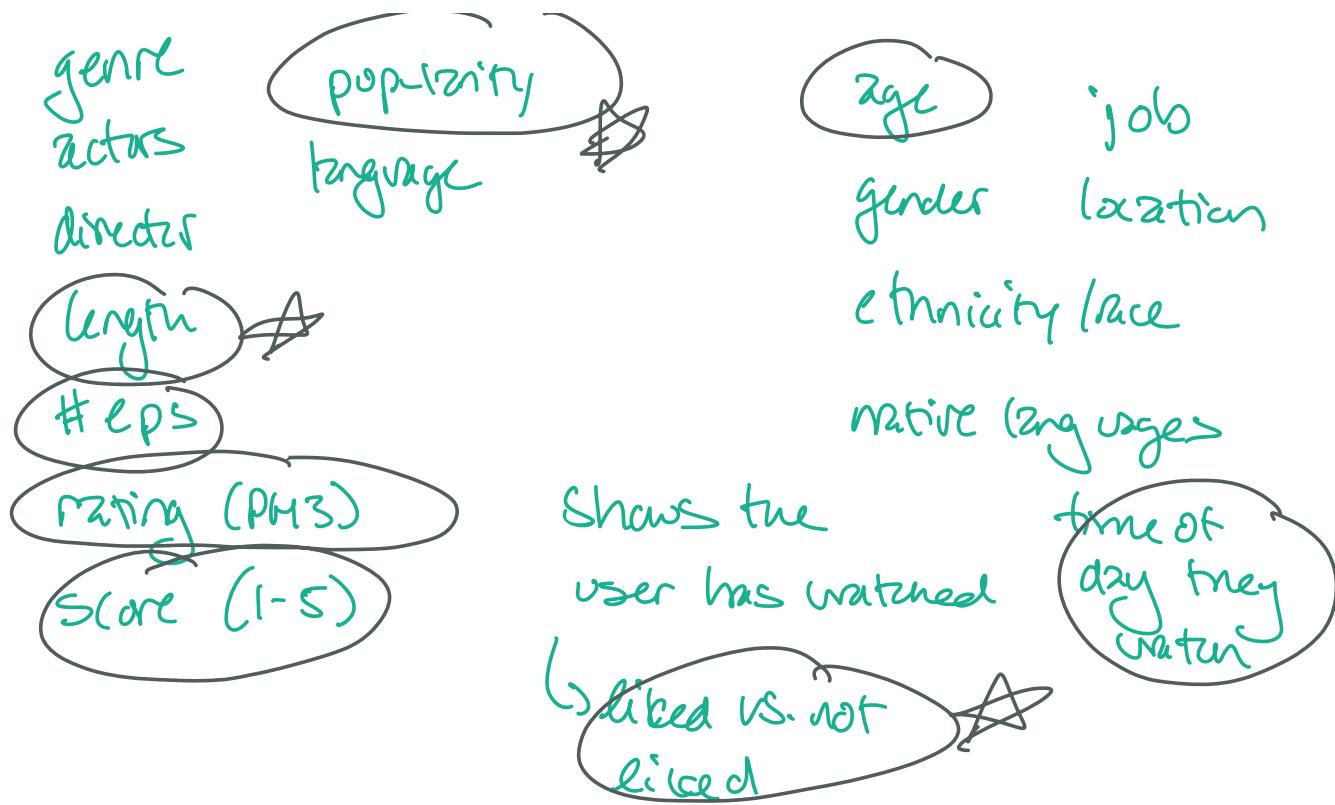
Rec what Laney should watch next

Improvement over real systems → any platform!

Start: what data do we need?

• About shows

• About users



numerical

↳ did you click off?

↳ ease of implementation

Info we have about a show:

- popularity (1-10)
- length (# mins of one ep)
- indv rating (1-10) → one person

Algorithm

- collect data on shows/people
- Ask Lany about a show she recently
watched

- get pop, length, rating of that show

- recommended shows with similar length, pop and good rating

Content-based filtering