

Ethics:
Choices, Values, & Burritos

Learning Goals

1. Choices embed values
2. Define Bias & Fairness
3. Normative vs. Descriptive
4. Problem Formulation, Representation & Long Tail



Learning Goals

1. Choices embed values
2. Define Bias, Fairness, Problem Formulation, Long Tail ...
3. **Keep Mehran, Chris, and Juliette Busy in the Chat**



Mid-Quarter Evaluations

CS 106A Mid-Quarter Evaluation

* Required

Instructor Feedback

These first few questions come from Mehran Sahami, Chris Piech, and Juliette Woodrow. Your anonymous feedback will be reviewed by them, so please be specific and give suggestions for improvement.

What has Mehran been doing so far that has worked well for you? What should he continue doing? *

Your answer

What has Mehran been doing so far that has not worked well for you? Is there anything that he should stop or start doing?

Your answer

What has Chris been doing so far that has worked well for you? What should he



Assignment 5: Sneak Preview



Assignment 5: Ratings

QUALITY

5.0

106A

👍 AWESOME

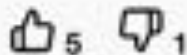
Feb 10th, 2015

For Credit: Yes Textbook: Yes

Mehran Sahami, what Stanford is all about, great teaching, new ideas, helping you ask big questions and build awesome stuff.

DIFFICULTY

4.0



Assignment 5: Ratings

QUALITY

5.0

CS106A

😄 AWESOME

Mar 25th, 2017

For Credit: **Yes** Attendance: **Mandatory** Would Take Again: **Yes** Grade: **A** Textbook: **Yes**

Chris Piech is everything! A natural teacher who loves his material and gets students to love it too. I want to be all that, do what he does, live that amazing life of being a great teacher at the world's greatest CS department.

DIFFICULTY

3.0

RESPECTED

INSPIRATIONAL

AMAZING LECTURES



Are These Ratings Biased?

QUALITY

5.0

DIFFICULTY

3.0

CS106A

😄 AWESOME

Mar 25th, 2017

For Credit: **Yes** Attendance: **Mandatory** Would Take Again: **Yes** Grade: **A** Textbook: **Yes**

Chris Piech is everything! A natural teacher who loves his material and gets students to love it too. I want to be all that, do what he does, live that amazing life of being a great teacher at the world's greatest CS department.

RESPECTED

INSPIRATIONAL

AMAZING LECTURES



What is in the Dataset?



You Have The Power to Find Out:



CS Skills

You Have The Power to Find Out:



Ethics Skills

CS Skills

You Have The Power to Find Out!

Ethics Skills

Values &
Habits

CS Skills



What is Bias?

Bias: difference between measured results and "true" value



What is Bias?

Bias: difference between measurement results and "true" value



What is Bias?

Bias: difference between measurement results and "true" value



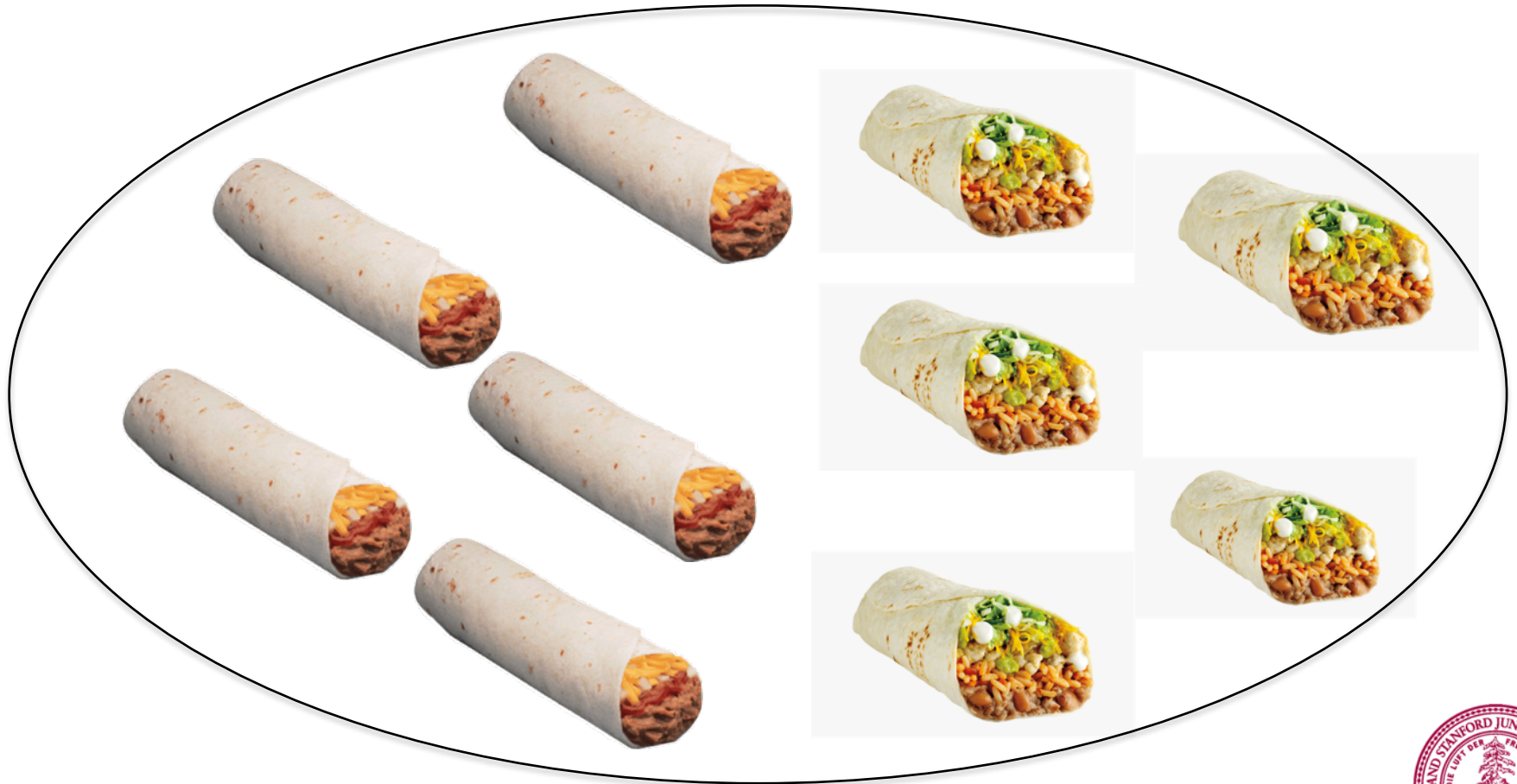
"Sampling Bias"

Statistical Bias: learn more in future computer science & statistics classes!



What is Bias?

Bias: difference between measurement results and "true" value



What is Bias?

Bias: difference between measurement results and "true" value



"Sampling Bias"

Statistical Bias: learn more in future computer science & statistics classes!



What is Bias?

Bias: difference between measurement results and "true" value

What kinds of bias raise ethical concerns?



Discriminatory Bias

Biased measurement or classification

+ use of that bias to compound existing injustice or fail to treat all as having equal moral worth

=> Unfair Bias



Unfair Bias in Ratings

How might this definition of work
in ratings?



Equality of Opportunity?

Equality of Opportunity: everyone has same opportunity to develop skills needed for the job, apply for the job, and get promoted.



Parity: Equality of Numbers

Parity: Everyone is equally likely to be a good teacher, so we should expect to end up with number of good teachers equal to population. Parity-fair ratings should reflect this.



Measures of Fairness

Simulating loan decisions for different groups

Drag the black threshold bars left or right to change the cut-offs for loans.
Click on different preset loan strategies.

Loan Strategy

Maximize profit with:

MAX PROFIT

No constraints

GROUP UNAWARE

Blue and orange thresholds are the same

DEMOGRAPHIC PARITY

Same fractions blue / orange loans

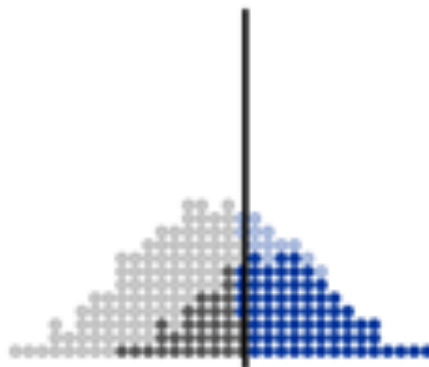
EQUAL OPPORTUNITY

Same fractions blue / orange loans to people who can pay them off

Blue Population

0 10 20 30 40 50 60 70 80 90 100

loan threshold: 59

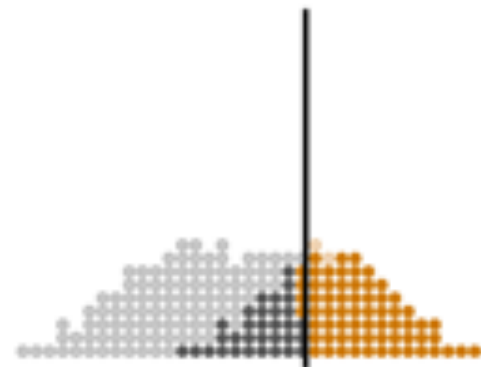


denied loan / would default: light grey
denied loan / would pay back: dark grey
granted loan / defaults: light blue
granted loan / pays back: dark blue

Orange Population

0 10 20 30 40 50 60 70 80 90 100

loan threshold: 53



denied loan / would default: light grey
denied loan / would pay back: dark grey
granted loan / defaults: light orange
granted loan / pays back: dark orange



Fairness: Beyond the Numbers

QUALITY

5.0

DIFFICULTY

3.0

CS106A

😄 AWESOME

Mar 25th, 2017

For Credit: **Yes** Attendance: **Mandatory** Would Take Again: **Yes** Grade: **A** Textbook: **Yes**

Chris Piech is everything! A natural teacher who loves his material and gets students to love it too. I want to be all that, do what he does, live that amazing life of being a great teacher at the world's greatest CS department.

RESPECTED

INSPIRATIONAL

AMAZING LECTURES



Evaluation Beyond the Numbers

**NATURAL
TEACHER**

QUALITY

5.0

CS106A

👍 AV
ME

Mar 25th, 2017

For Credit: **Yes** Attendance: **Mandatory** Would Take Again: **Yes** Grade: **A** Textbook: **Yes**

Chris Piech is everything! A natural teacher who loves his material and gets students to love it too. I want to be all that, do what he does, live that amazing life of being a great teacher at the world's greatest CS department.

DIFFICULTY

3.0

RESPECTE

INSPIRATIONAL

AMAZING LECTURES

GREAT

EVERYTHING!



Descriptive vs. Normative

Descriptive language

- is
- was
- what people did
- what happened

Normative language

- right
- wrong
- good
- bad
- should
- should not



Descriptive or Normative?

**NATURAL
TEACHER**

QUALITY

5.0

CS106A

👍 AV
ME

Mar 25th, 2017

For Credit: **Yes** Attendance: **Mandatory** Would Take Again: **Yes** Grade: **A** Textbook: **Yes**

Chris Piech is everything! A natural teacher who loves his material and gets students to love it too. I want to be all that, do what he does, live that amazing life of being a great teacher at the world's greatest CS department.

DIFFICULTY

3.0

RESPECTED

INSPIRATIONAL

AMAZING LECTURES





Descriptive



Normative

Thick Normative Terms

“Thick” normative terms are both descriptive AND normative:

- brave : foolhardy



Thick Normative Terms

“Thick” normative terms are both descriptive AND normative:

- brave : foolhardy
- Cowardly :



Thick Normative Terms

“Thick” normative terms are both descriptive AND normative:

- brave : foolhardy
- Cowardly : cautious
- Polite : (?)
- rude : (?)
- chill
- kind
- etc



Thick Normative Terms

Thick Normative Terms & Fairness

- We compare people in many ways, not just numerically
- Thick normative terms express “loaded” judgments



Are These Claims Descriptive or Normative?

Mark Zuckerberg on whether Facebook would fact-check false claims about election suppression:

- 1. *"We have a different policy, I think, than Twitter on this."*
- 2. *"You know, I just believe strongly that Facebook shouldn't be the arbiter of truth of everything that people say online."*
- 3. *"I think in general private companies probably shouldn't be—or, especially these platform companies—shouldn't be in the position of doing that."*



Are These Claims Descriptive or Normative?

Not surprising that statements setting the policy for platforms would be normative. What about the programs behind the platforms themselves?

Do programs like the ones you are writing contain normative claims or values?



How are values
embedded in design?

1. Problem Formulation Embeds Values

1. “Sandcastle” (warm-up) problem: Finding forest fires.

We’re going to start by writing a function called `highlight_fires` (in the file `forestfire.py`) that highlights the areas where a forest fire is active. You’re given a satellite image of Greenland’s 2017 fires (photo credit: Stef Lhermitte, Delft University of Technology). Your job is to detect all of the “sufficiently red” pixels in the image, which are indicative of where fires are burning in the image. As we did in class with the “redscreening” example, we consider a pixel “sufficiently red” if its red value is greater than or equal to the average of the pixel’s three RGB values times some intensity threshold. In this case, we have provided you with an appropriate intensity threshold of 1.05 via a constant named `INTENSITY_THRESHOLD` in the file `forestfire.py`. Note that this is a different intensity threshold value than we used in class for the “redscreening” example, as different applications often require different intensity threshold.

When you detect a “sufficiently red” pixel in the original image, you set its red value to 255 and its green and blue values to 0. This will highlight the pixel by making it entirely

Formulating a problem means
describing the **desired solution as**
good or worthy of being done.



Problem formulation and goal statements are normative

Formulating a problem means describing the **desired solution as good** or worthy of being done.

- What is the **problem to be solved**?
- **For whom is this a problem**? Who would benefit from its solution?
- **Who can agree** that this is a problem worth solving?



What is the problem to be solved?

“Homeless people are sleeping here and we (who is we?) want them to stop”



What is the problem to be solved?

“Homeless people are sleeping here and we (who is we?) want them to stop”



What is the problem to be solved?

“Homeless people are sleeping here and we (who is we?) want them to stop”

“Some people in our community don’t have a place to sleep and we (who is we?) think they should”



What is the problem to be solved?

“Homeless people are sleeping here and we (who is we?) want them to stop”

“Some people in our community don’t have a place to sleep and we (who is we?) think they should”



What is the Problem to be Solved?

Search Engines

Ratings of Professors



Problem formulation embeds values

Piech + Sahami, CS106A, Stanford University



What are the Problem(s) to be Solved?

Ratings of Professors



What are the **problem(s)** to be solved?

For whom are these **problems**? Who would benefit from their solution(s)?

For each problem, who can agree that the problem is worth solving?



What are the Problem(s) to be Solved?

Search Engines



What are the **problem(s)** to be solved?

For whom are these **problems**? Who would benefit from their solution(s)?

For each **problem**, who can agree that the problem is worth solving?



2. Choice of Data Embeds Values

CS 106A Mid-Quarter Evaluation

* Required

Instructor Feedback

These first few questions come from Mehran Sahami, Chris Piech, and Juliette Woodrow. Your anonymous feedback will be reviewed by them, so please be specific and give suggestions for improvement.

What has Mehran been doing so far that has worked well for you? What should he continue doing? *

Your answer

What has Mehran been doing so far that has not worked well for you? Is there anything that he should stop or start doing?

Your answer

What has Chris been doing so far that has worked well for you? What should he

- Surveys are cheap to run
- They measure opinions
- What are other ways to measure quality of professors?



Underrepresentation & The Long Tail



Unicode & Representation

Covers over 1 million symbols,
Full coverage of 90 languages,
basic coverage for 200 languages



U+30ED



U+15D2



U+2666



U+043E



U+2690



U+2B07



U+FF0F



U+2B1C



U+1F490



U+2639



U+263B



U+13EA

Adopt a Character



U+1F643

Emoji



Basic Info



U+1F9E1

News

Events

Connect



U+0994

Membership



U+05E9



U+0E07



U+1D25



U+51F8



U+0B37

Press



U+203F



U+1F60A



U+03CE



U+127F



U+060F



U+1557

Everyone in the world should be able to use their own language on phones and computers.

[LEARN MORE ABOUT UNICODE](#)



Representation & Surveys



Representation & Survey Data



If I had a different experience, my ratings may not affect the numerical average - but they matter!



Representation & Search



I want to see
myself
represented in
search in my own
terms, but other
people are using
different terms
to describe me

Learning Goals

1. Choices embed values
2. Define Bias & Fairness
3. Normative vs. Descriptive
4. Problem Formulation, Representation & Long Tail



Questions?

Questions?



Discrimination

Laws and policies demean when they express that a person is of lesser moral worth in a context in which the actor has the kind of social power to lower the status of the individuals affected (Hellman, 2008, 7).

Is the same true for bias in data?

