

✓ Congratulations! You passed!

Grade received 60% To pass 1% or higher

[Go to next item](#)

Week 2 Quiz - Bias and Classification

Latest Submission Grade 60%

1. Scenario A

1 / 1 point

(Answer the following 5 questions based on the information given below and the material in this course.)

A study by ProPublica entitled "The Tiger Mom Tax" determined that the Princeton Review offered different prices for the same tutoring package. "Some ZIP codes entered into the company's website are offered The Princeton Review's Premier course for \$8,400, but other zips get the same course for as little as \$6,600. The tutoring service says the gap is because of income; the zip codes with higher prices tend to have an outsize number of financial services headquarters and other businesses that drive up The Princeton Review's prices. At the same time, the data show that a disproportionate number of those zip codes have a high percentage of Asian residents. The pricing gap for Princeton Review courses remains even for Asians in lower-income neighborhoods." This report led to allegations that Princeton Review was employing unethical racial discrimination in pricing.

Complete the following sentence:

"The most salient ethical issue about using ZIP codes to assign prices is that ZIP codes ____."

- A) ...are one of the US Federal Protected Classes.
- B) ...might act as a proxy for groups that it is not desirable to target.
- C) ...are private.
- D) ...are too specific for targeting the wealthy.



Correct

Correct answer: ZIP codes might act as a proxy for groups that it is not desirable to target.

ZIP codes are **not** one of the US Federal Protected Classes. It is a lesson of the class that it is useful to think outside and beyond legal discrimination rules. Indeed, in this case the problem might be that income and Asian demographics are correlated when analyzed by ZIP code, which is best described by option B. ZIP codes are not private and it is not clear what "too specific" means here, so the other choices don't make sense.

2. Assume that Princeton Review **is** price discriminating by race and that this is **legal** price discrimination. Are there still potential ethical problems raised by this practice?

1 / 1 point

- A) Yes, it reinforces a negative stereotype about Asians ("Tiger Mom/Dad").
- B) Yes, it promotes unequal access to education.
- C) Yes, it treats some people differently than others for reasons that are out of their control and may not be relevant to the decision at hand.
- D) Yes, customers may have been led to believe that they are getting the best price and therefore they will feel that the Princeton Review has lied to them.
- E) Yes, all of the above are true.



Correct

Correct answer: Yes, all of the above are true.

Although you might judge some of these to be more serious than others, all of these answers are potential ethical problems in this scenario. Many were raised in the examples of discrimination covered in lecture. Indeed, the lecture included the example of price discrimination and the example of educational testing. This question also illustrates a theme of the course: in general there are a large number of potential ethical problems for any given situation.

3. Someone born into a wealthy family with the means to pay for the Princeton Review may accrue other benefits throughout their lifespan that reinforce their initial superior position. For example, wealthy parents might be more likely to pay for tutoring and enroll their children in expensive, excellent private schools. The reverse might be the case for someone born poor. This effect is important because it might make later standardized educational testing an indication of birth rather than merit. This effect is called:

1 / 1 point

- A) The Dunning-Kruger Effect

B) Cumulative Disadvantage (or Cumulative Advantage)

C) Social Privacy

D) All of the above

E) None of the above

Correct

Correct answer: Cumulative Disadvantage (or Cumulative Advantage)

The key idea of the concept "cumulative disadvantage" is that a person's starting situation is out of their control (in this case, the wealth of your parents) and that this initial situation is not necessarily a good way to make subsequent decisions (your family tree shouldn't determine your ability to access an education), but that in fact the starting situation does influence the subsequent decisions (in this case, by wealthy parents paying for expensive test prep, and test scores being taken as a good indicator of ability). We believe in ideas like meritocracy or universality, but at the same time we recognize that cumulative disadvantage happens in almost every human system. This tension comes up a lot in ethical problems.

4. Imagine that we were able to determine that there was **no intent by any individual person** to discriminate by race in this example. Does this mean the problems discussed in this section are resolved? 0 / 1 point

A) Yes, without intent there is no longer a problem.

B) Yes, without intent there is no crime.

C) Yes, both A and B are true.

D) No, we still want to reduce harm no matter the intent.

E) No, in fact a lack of intent makes the ethical situation worse.

Incorrect

Correct answer: No, we still want to reduce harm no matter the intent.

Ethics is about identifying and reducing harms done to patients (or "stakeholders") regardless of the intent. Intent might be important as well in determining responsibility, but as we learned in the week 1 introduction there are many situations where well-meaning people can act to unintentionally produce broken, even dangerous analyses and systems. We still need to fix things that are broken and dangerous, even if we can't identify a villain who intentionally produced the danger.

This scenario is real. If you are curious about it, more information can be found at the ProPublica report: <https://www.propublica.org/article/asians-nearly-twice-as-likely-to-get-higher-price-from-princeton-review> and press coverage: <http://www.takepart.com/article/2015/09/04/tiger-mom-tax-test-prep-companies-charge-asian-americans-more>

5. Scenario B 0 / 1 point

(Answer the next 5 questions *based on the information given below*)

Amazon built an automated analysis tool to score job applicants. The gender of job applicants was not used as a variable, but the system was found to exhibit bias against women. "The team had been building computer programs since 2014 to review job applicants' resumes with the aim of mechanizing the search for top talent..... But by 2015, the company realized its new system was not rating candidates for software developer jobs and other technical posts in a gender-neutral way." This program was shut down by Amazon.

Question:

One explanation advanced for this problem is that there is a longstanding gender gap in the tech industry. Which statement from this week best speaks to this explanation?

A) Legally, in the US you must omit protected classes from analyses.

B) Within a data set, there is always proportionally less data available about minority groups.

C) Classification should be as objective as possible.

D) As long as rating systems are numeric there is little ethical risk.

E) None of the above

Incorrect

Correct answer: Within a data set, there is always proportionally less data available about minority groups.

If there is little data about a minority group--in this example, women--it might be easy to wrongly conclude that a behavior closely associated with that group is a marker for failure. Sources inside Amazon reported that because few graduates from women's colleges hold technology jobs, the system used the negative association between women's colleges and tech jobs to conclude that candidates from these colleges should not be hired. In other words, when you build a system with biased data then the output will exhibit the same biases.

The other answers don't work because we highlighted the difficulty with objectivity in classification. In week 1 we emphasized that simply converting something into a number does nothing to eliminate ethical risk or make things objective. Although option "A" may be true in some circumstances, it doesn't explain this problem with the hiring system.

6. This is an example of an "automated scoring system" as described by the Citron et al. article. According to the article, which of the below regulatory responses would be effective against such a system?

0 / 1 point

- A) Applicants must have the right to inspect their own data that is used by the system.
- B) Applicants must have the right to demand corrections for errors.
- C) If there is a dispute about erroneous information, they should have the right to attach a note of clarification.
- D) All of the above.
- E) None of the above



Incorrect

Correct answer: None of the above.

Items A, B, and C are in fact the current rules in the US for credit scores. Citron et al. argue that the current system is a "failure." While these solutions might sound like a good idea in general, this case shows that they won't help much. For example, job applicants in the Amazon case already have the right to inspect their own data used by the system (i.e., "Applicants must have the right to inspect their own data that is used by the system.") -- the Amazon system scans the resume that the applicants uploaded themselves. Having this "right" does nothing to stop harmful discrimination by the system.

7. How would you analyze this situation to identify "harm" and spot potential ethical problems?

1 / 1 point

- A) Harm refers only to physical harm.
- B) Harm is individual.
- C) A lost job opportunity is a harm.
- D) Reinforcing or producing negative ideas about gender is a harm.
- E) Both "A lost job opportunity is a harm" and "Reinforcing or producing negative ideas about gender is a harm" are true.



Correct

Correct answer: Both "A lost job opportunity is a harm" and "Reinforcing or producing negative ideas about gender is a harm" are true.

In Johnny Cash's segment we highlighted the importance of thinking broadly about harms and addressing harms beyond the individual. Indeed, in the dictionary the definition of "harm" already goes beyond physical harm. The word refers to any adverse consequence.

8. What is the best statement of advice from this class that you could use to think through problems like this one in the future?

0 / 1 point

- A) Don't ever consider gender at all.
- B) Consider gender, but only after the system is built in order to test its performance.
- C) Consider gender, but only its prevalence in the data used to train/build the system.
- D) It may be impossible for a system like this to avoid gender because identity categories like gender may be fundamentally intertwined with human behavior.
- E) None of the above



Incorrect

Correct answer: It may be impossible for a system like this to avoid gender because identity categories like gender may drive human behavior.

It may be impossible for a system like this to avoid gender because identity categories like gender may drive human behavior.

Answers A, B, and C are unrealistic here because it doesn't seem likely that you could produce an unbiased, fair, and/or just system by, for example, never thinking the key factor that made this example system a problem (answer "A"). The word "only" makes B and C poor answers -- although perhaps they would be better than doing nothing (answer "A"). While answer B is reasonable it is too narrow -- in this class we also emphasized that considering the distribution of minority populations in your data was also important

(answer "C").

Indeed, in this real system the data scientists involved found something like answer "D". They found that machine learning was scanning resumes and using gendered verb choices on resumes to rank candidates, even though they did not intend for this to happen. For example, men were more likely to use verbs like "executed" and "captured" on resumes, and these words were then used as a signal of success by the system even though gender was not explicitly considered as a variable.

9. What does this Amazon scenario share with the COMPAS system for scoring future violent crime as described in the Angwin et al. article?

1 / 1 point

- A) It is difficult to construct a score that doesn't include items that can be correlated with unwanted variables (race, gender).
- B) The company does not publicly disclose the calculations used to arrive at scores.
- C) If implemented, the system may exacerbate unwarranted and unjust disparities.
- D) The systems can have serious negative consequences for the people who are inaccurately scored.
- E) All of the above



Correct

Correct answer: all of the above

The first three quotes are taken from the Angwin et al. article and slightly paraphrased. All four apply equally to both systems. Even though we are using a lot of "all of the above" answers in these quizzes, in part this is to make the point that the same pattern of ethical problems involving discrimination recurs in systems that are in different contexts. In this case, hiring vs. predicting future criminality. These are very challenging problems because of A-D: some say they are unsolvable.

If you are interested in the Amazon hiring tool scenario, it is a real project. It is described here:
<https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scaps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>

10. A job applicant is notified of being eliminated from the recruitment process due to the results of a risk algorithm. The system scrapes candidates' social media accounts for pictures, content of the posts and likes and makes inferences about which candidates are unsuitable for employment and provides a report to the employer.

1 / 1 point

What is the ethical issue in this case?

- The risk algorithm may result in bias, arbitrariness or unfair mischaracterizations due to the data or coding of risk
- Job applicant might not be in a position to refuse access to social media accounts due to power imbalances between the applicant and employer
- Employer may learn more private information about the applicant's lifestyle than the applicant would be normally willing to share
- All of the above
- None of the above



Correct

Correct answer: All of the above.

This question covers multiple course concepts and ethical problems such as privacy, problem with consent, power relationships and harm. The candidate may not only lose this job opportunity but might also start self-control or filter expressions or thoughts in order to avoid future harm.