REGIS UNIVERSITY

MSDS640 Week 2:

Ethics, Privacy, and Social Justice Issues

Ethics, Privacy, and the Data Science Lifecycle

Ksenia Polson, Ph.D

Week 1 recap

- Data collection and Privacy
- Social Justice
- ACM code of ethics and ethical guidelines (7 general ethical principles)

1. Contribute to society and to human well being
2. Avoid harm
3. Be honest and trustworthy
4. Don’t discriminate
5. Respect others work to produce new ideas/inventions
6. Respect privacy
7. Respect Confidentiality

Week 2 Learning objectives

- Data science lifecycle – ethics, privacy concerns

- Risks/consequences of ethical misconduct
  - Fired, prison, reputation

Methodologies for data science projects

- Cross Industry Standard Process (CRISP-DM)
- KDD
- Domino Labs*
- Microsoft's TDSP*
- SEMMA

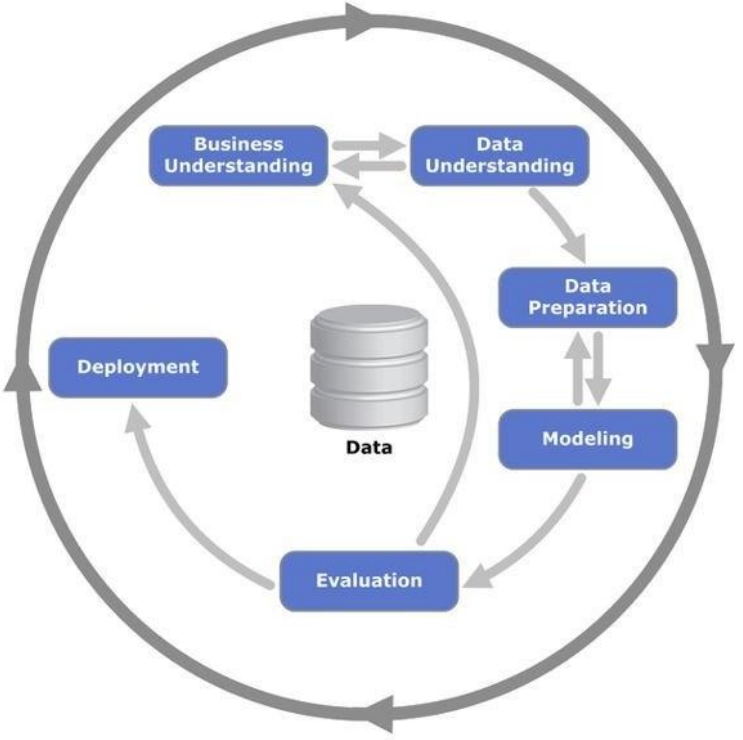
* approaches mixing conventional PM with agile

http://www.datascience-pm.com/

CRISP-DM Methodology

The CRISP-DM process has 6 steps

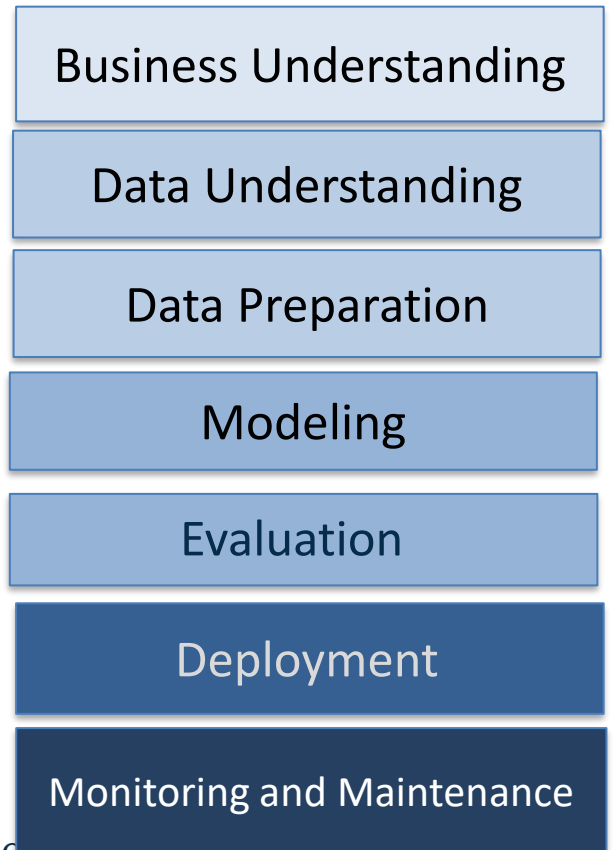
1. Business Understanding
2. Data Understanding
3. Data Preparation
4. Modeling
5. Evaluation
6. Deployment



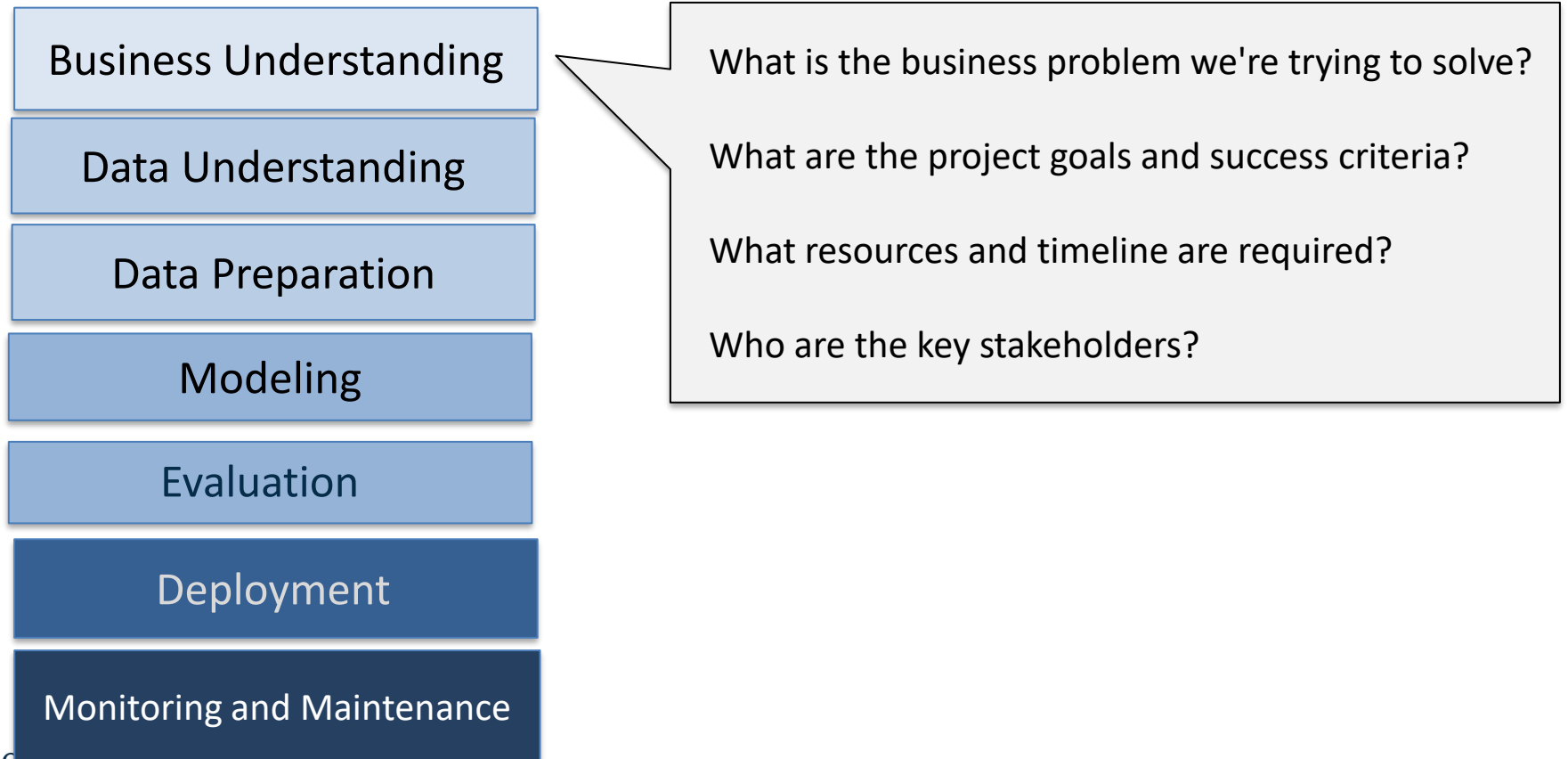
Ethical issues can arise at every stage of the process (will be discussed)

http://www.datascience-pm.com/crisp-dm-2/

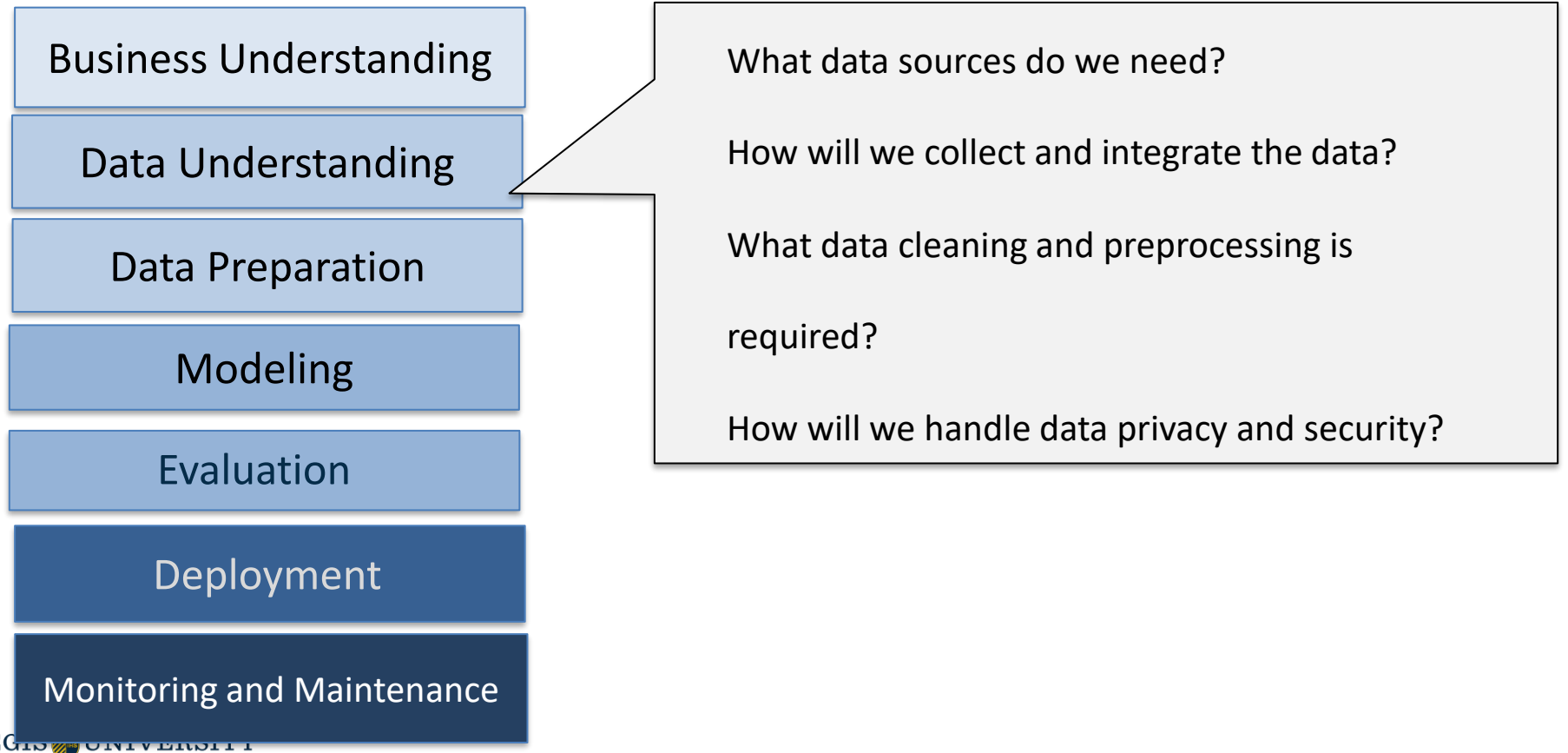
Data Science Project Lifecycle based on CRISP-DM



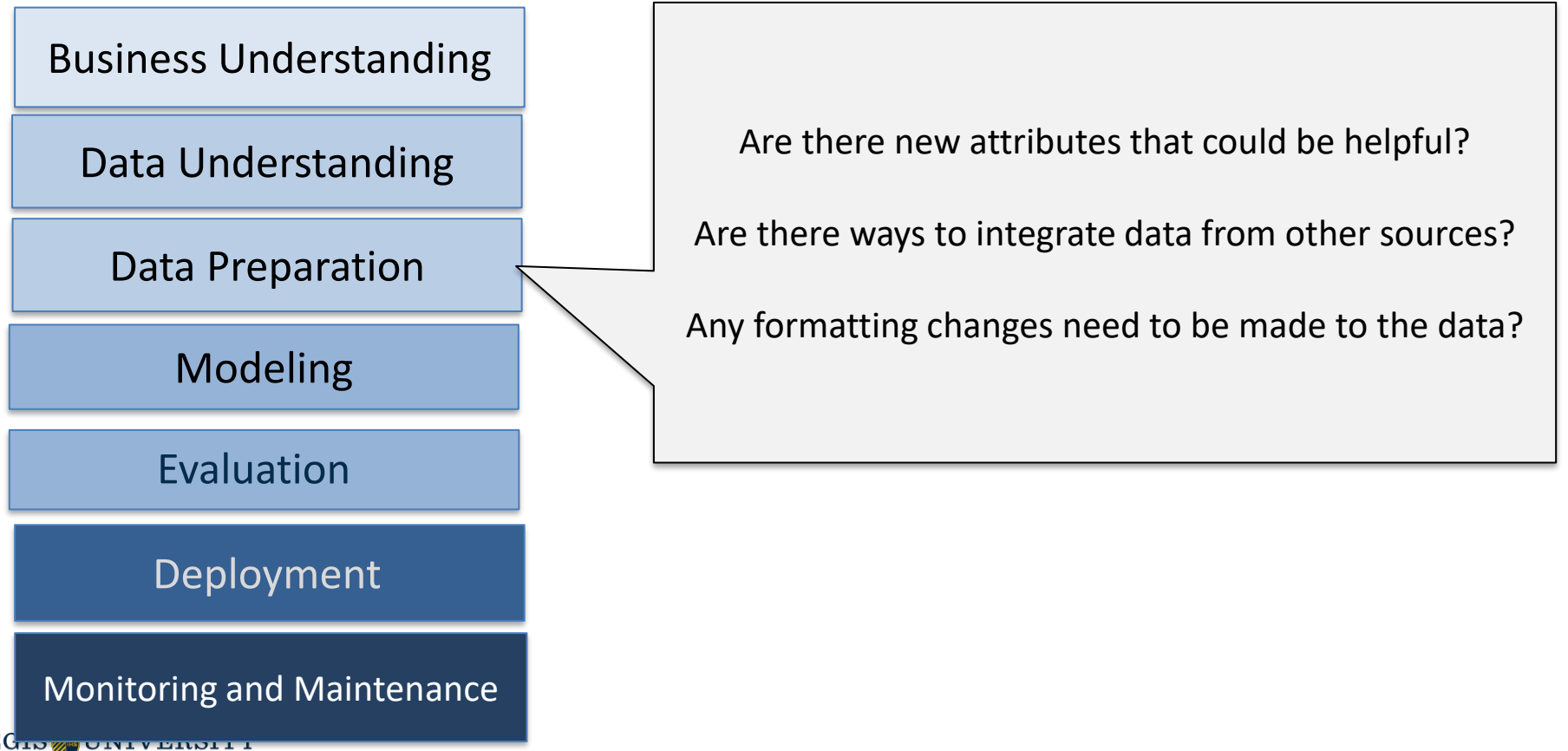
Data Science Project Lifecycle based on CRISP-DM



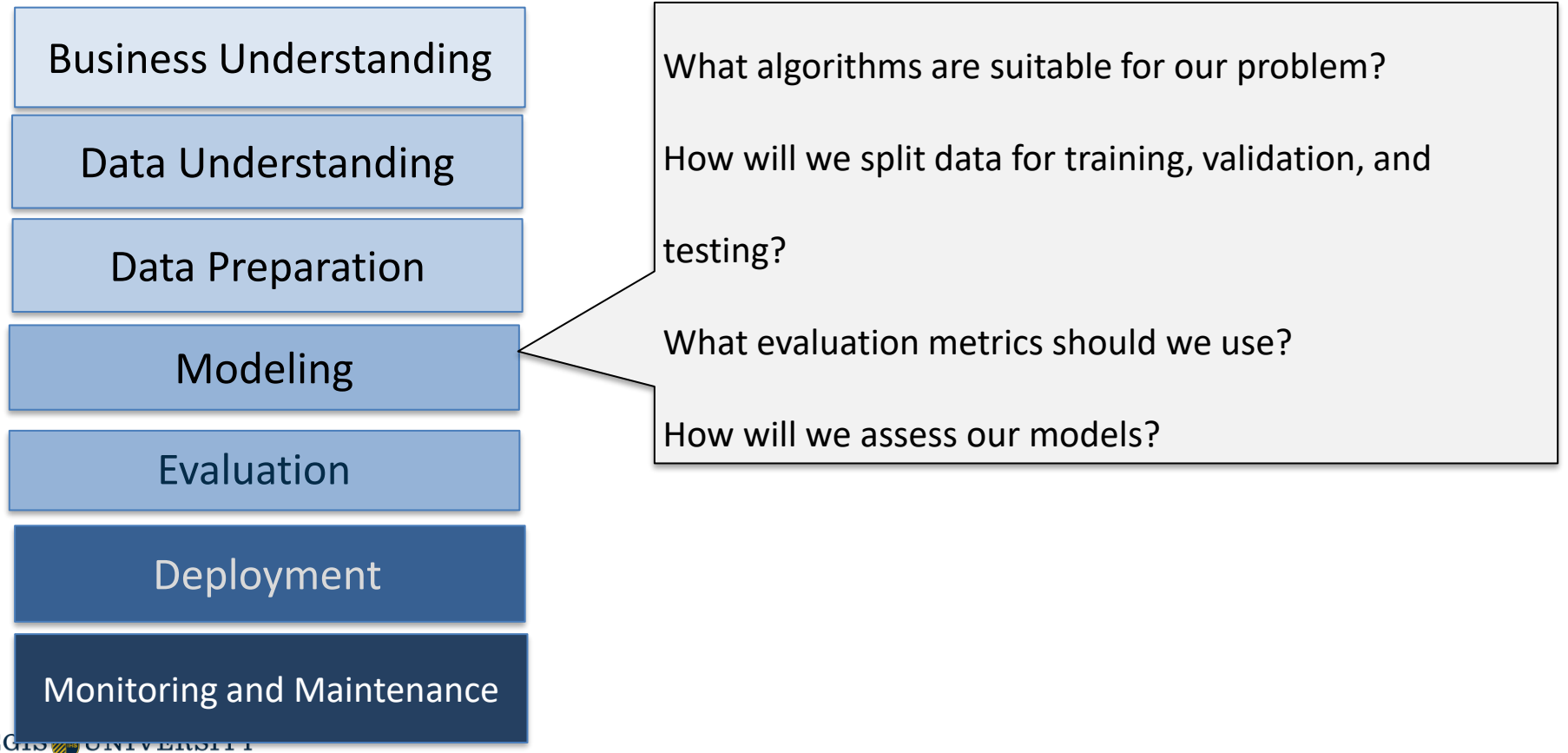
Data Science Project Lifecycle based on CRISP-DM



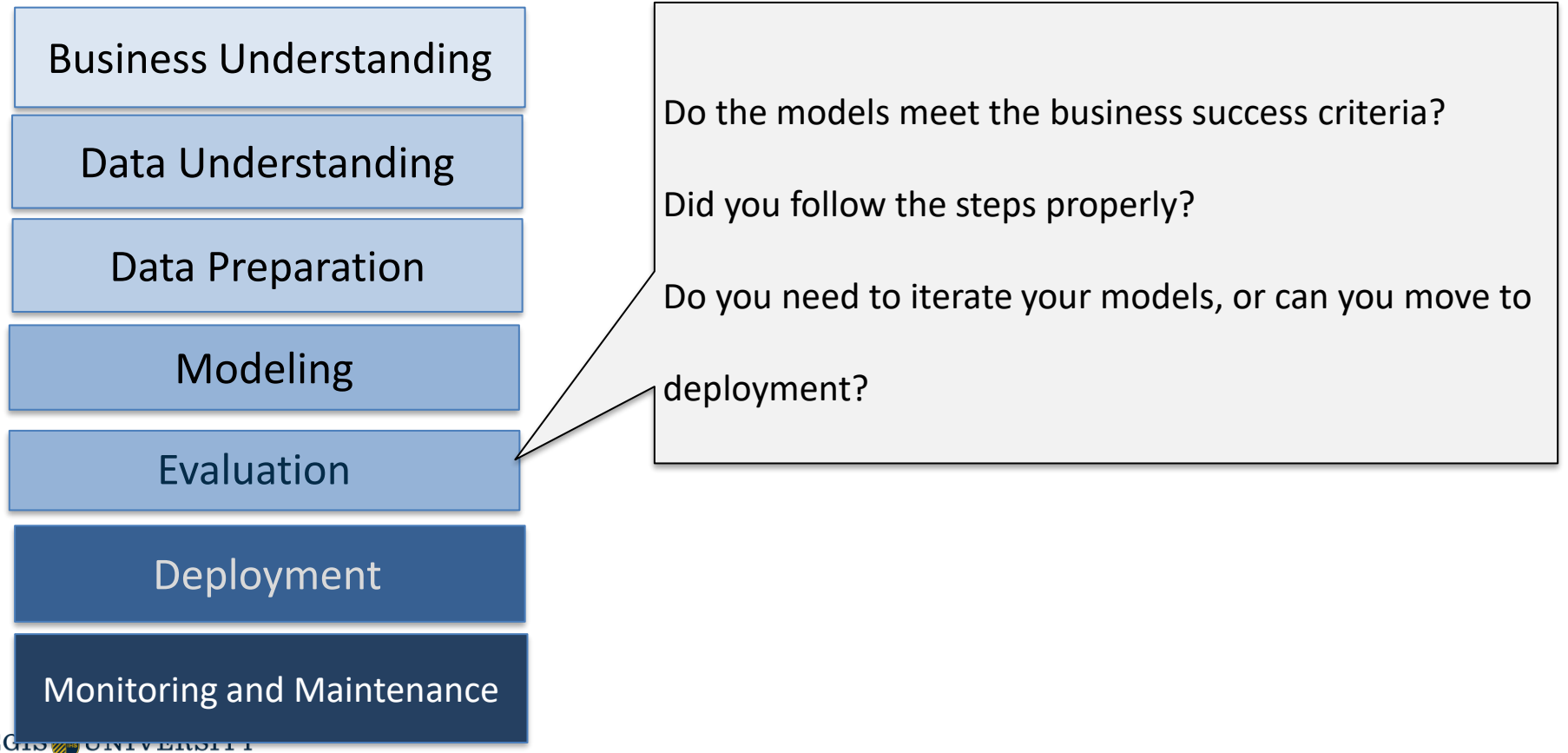
Data Science Project Lifecycle based on CRISP-DM



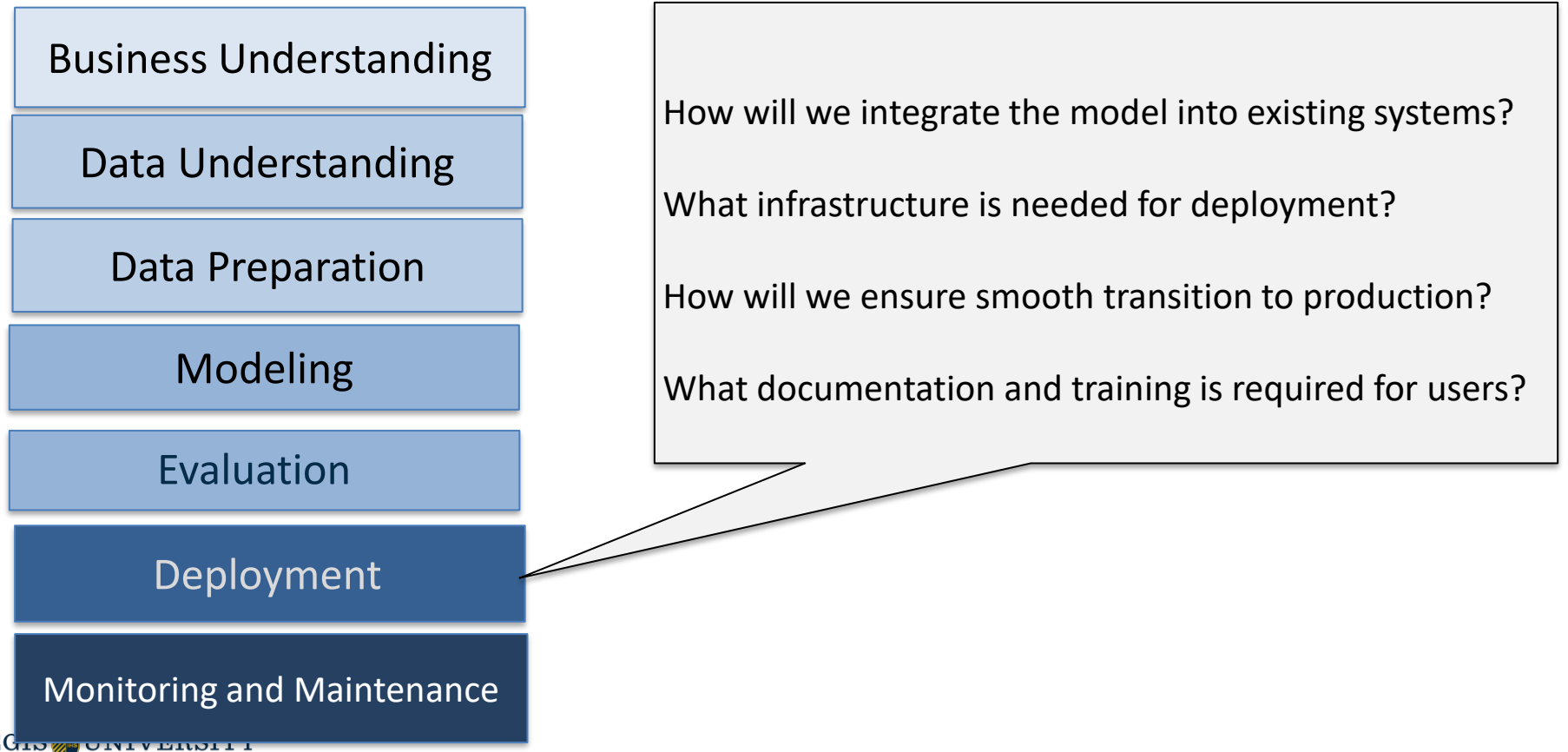
Data Science Project Lifecycle based on CRISP-DM



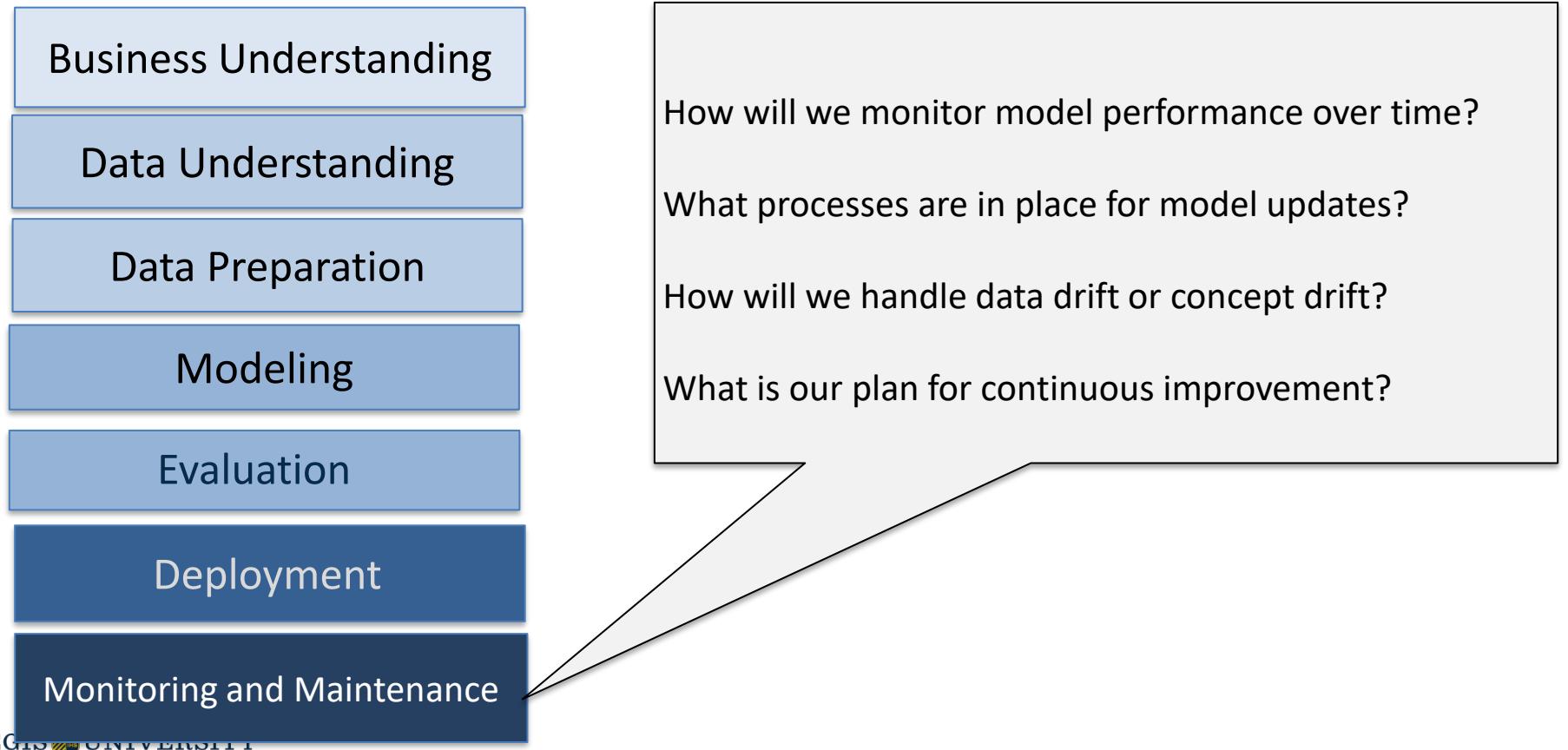
Data Science Project Lifecycle based on CRISP-DM



Data Science Project Lifecycle based on CRISP-DM



Data Science Project Lifecycle based on CRISP-DM

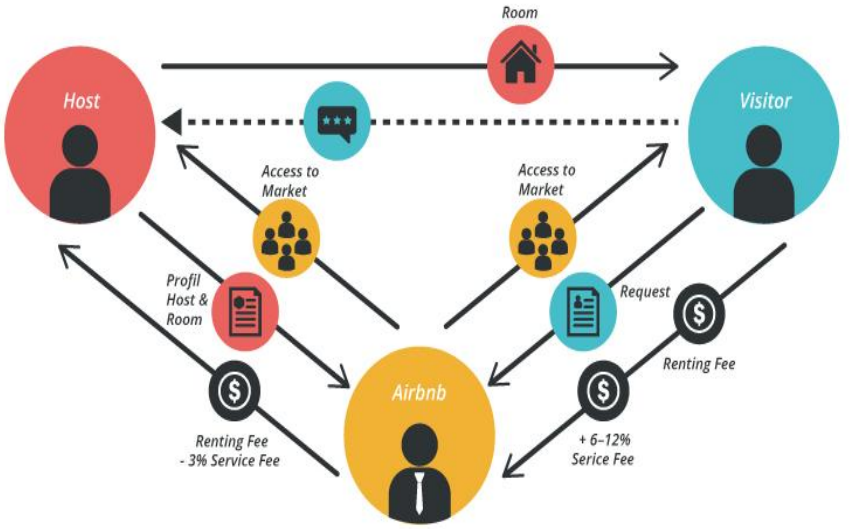


CRISP-DM examples

- Kaggle AirBNB

  TDS AirBNB

  MIMIC-III



https://medium.com/@mayrazrodriguez/sentiment-analysis-and-topic-modeling-using-airbnb-dataset-c65848b98a57

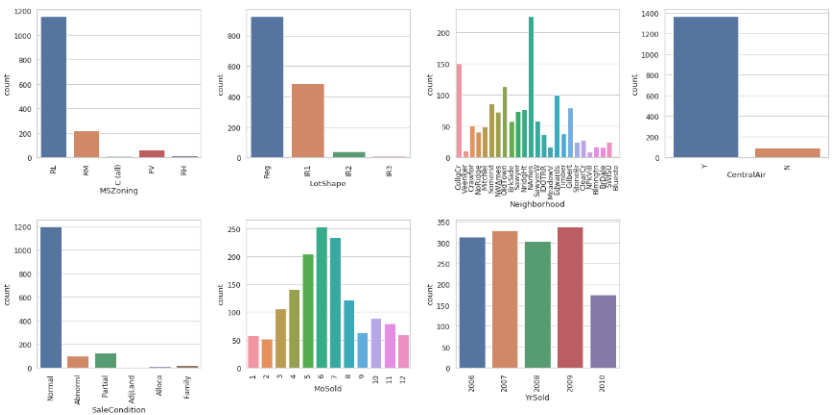
DS PM ethical considerations

- (1) Business understanding
  - Any ethically questionable assumptions or undertones?
  - Will data collection entail privacy concerns?
  - Will any ML models or actionable insights result in injustice?
  - AirBNB example – shouldn't we be looking for issues from reviews? E.g. potential safety issues?
  - MIMIC-III example – if we are predicting if people in ICUs decease, will this prediction be used ethically?

http://www.lionessesofafrica.com/blog/2017/3/1/understanding-the-difference-between-a-business-strategy-vs-a-marketing-strategy

(2/3) Data understanding/prep

- How balanced are classes in the data? Will this cause any biases in conclusions or models?
- AirBNB – How balanced is data across demographics (e.g. neighborhoods)? Might require data joining with census data (from BigQuery or AWS, etc)
- MIMIC-III – Are there privacy concerns? Could data be de-anonymized?
- Is data biased towards certain groups?
- Are we cherry-picking data?
- Are we throwing out or correcting errors in an ethical way?



https://towardsdatascience.com/how-to-perform-exploratory-data-analysis-with-seaborn-97e3413e841d

(4) Modeling

- Were proper techniques used for the situation?
  - e.g. if we only have a handful of datapoints (fewer than 50ish) most ML algorithms cannot be used
  - If data is highly unbalanced (e.g. many of one class/target and not many of another), need to consider this
  - Hyperparameters should be tuned; data leakage avoided
- AirBNB – didn't really try many models or optimize RF model at all

https://www.appcoda.com/create-ml/

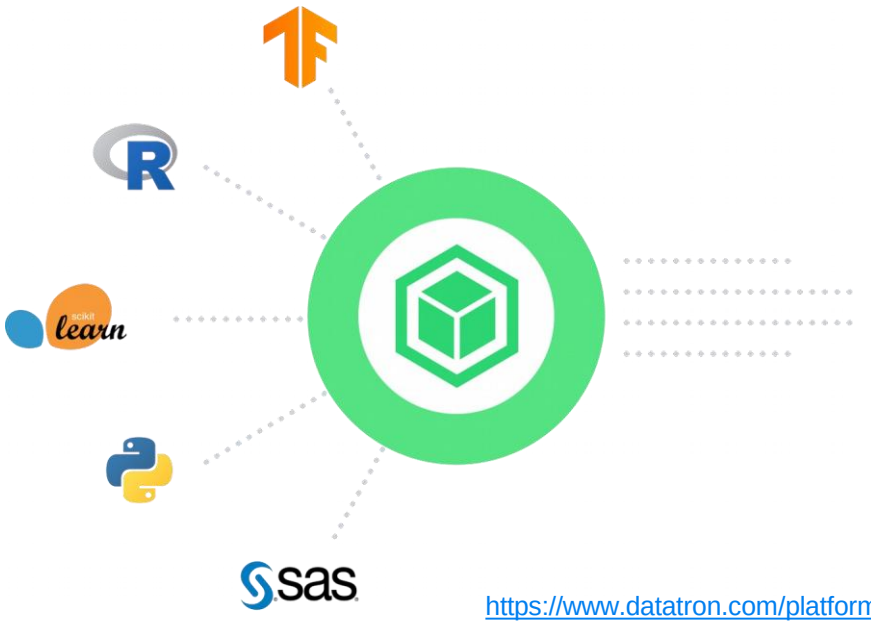
(5) Evaluation

- Be mindful of anything learned during data understanding
  - Are classes imbalanced? Accuracy vs Cohen's Kappa, no info rate
  - Have common mistakes been avoided (e.g. data leakage, overfitting)?

https://towardsdatascience.com/cross-validation-explained-evaluating-estimator-performance-e51e5430ff85

(6) Deployment

- Will model drift be accounted for?
- Are caveats clearly communicated?
- How will biases and model errors be monitored?



https://www.datatron.com/platform/deployment

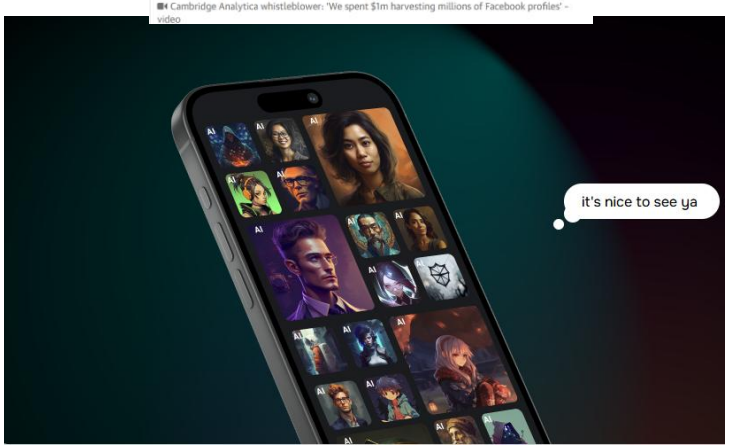
Example of Ethics Concerns

Cambridge Analytica scandal

- Whistleblower describes how firm linked to former Trump adviser Steve Bannon compiled user data to target American voters. More here

Character.ai lawsuit

- Meet Ais that feel alive. Chat with anyone, anywhere and in real time. For more, see here



Good Will Hunting (1997)

https://www.imdb.com/title/tt0119217/
https://www.youtube.com/watch?v=tH0bTpwQL7U

Week 2: AI Judicial Assistant Assignment

Design considerations and ethical concerns for the criminal justice system in Cook County.

Lots of info on COMPAS

We will perform EDA on the felony case data in Cool County and create a written report intended to go to the state attorney's office (including Kim Foxx and her data team). The report should describe the ethical issues found within the data, such as bias, ethical issues, and opportunities for improvement of the criminal justice system. Use at least 3 EDA plots and 3 or more of your professional-looking charts in your paper. The paper should be APA formatted and as short as possible while clearly conveying your messages. The most recent version of the data set is here.

Some questions that you may use to start could be:

• Are there racial, geographical, or socioeconomic biases present in the data or Cook County criminal justice system?

• Are there ways to reduce the populations of jails and prisons?

• Are there particular locations that might benefit from community services, like mental health or substance abuse treatment services?

Although a paper on this topic could be quite long, aim for fewer than 8 pages of double-spaced writing if possible**. Due next Thursday at 6pm.**

https://endtimeheadlines.org/2019/10/robot-judges-will-pass-sentence-with-no-human-bias-in-ai-courts/

AI Judicial Assistant Assignment Rubric

Criteria	Points
Grammar, formatting, organization, inclusion of at least 3 EDA plots and 3 professional looking charts, writing quality	5
Ethical, common good issues and opportunities for improvement in the Cook County criminal justice system and community as a whole	5
Total points:	**10**

Student Case Rankings

Please fill out the Case Ranking Spreadsheet by Sunday, **November 3 at 11:59pm**.

Conversation Theme

Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) is a software system used in U.S. courts. This system is based on an algorithm that assesses the likelihood of defendants awaiting trial to re-offend. This algorithm relies on data gathered from participants that includes demographics, previous arrests and employment history. COMPAS raises concerns regarding unfair discrimination, the reinforcement of human biases and a lack of transparency. Sources here and an article where the Cook County data is analyzed. Here is the latest article on it here.

1. What are some problems with a system like COMPAS?
2. If we were a data scientist (or similar role) on a team working on the COMPAS system, is there anything we can do to address these problems and make the system more equitable?
3. To open conversation for a wider AI discussion: What are the implications of AI-driven tools on ethics and privacy issues? Can we protect our information?