CX.AI

Class #7:
Selection bias

Lesson #1: Selection bias

- **What's the selection bias?**
    - Customers who take part can be fundamentally different to those who don't
    - The motivation to participate can be linked to what matters
    - Typically, more loyal customers have higher tendency to participate

- **When is it harmful?**
    - It may whitewash your descriptive - ratings and feedback, but this alone does NOT influence harmful driver analysis outcomes.
    - It is harmful, when the reason to participate is logically link to BOTH:
        - The outcome metrics (NPS rating)
        - One (or more selected) topics
            - EXAMPLE: Talkative people love surveys and for them personal service is much more important.

Lesson #2: Strategies to reduce it.

- **Strategies to reduce it.**
  - **Increase reply rates**
    - **For email outreach: More touches**
    - **For popup/website: Gain more attention**
    - **Set incentives unrelated to your service**
    - **Phone or other in-person survey – social pressure makes people not to quite**
  - **Measure bias with modeling**

- **What's a proper reply rate?**
  - **Science asks for >70% to accept a sample to be "unbiased and representative"**

CX.AI

# Lesson #3: Controlling the bias with modeling

- **Option 1: Hard approach**
    - "Force" a subset of clients to answer, by e.g., calling them, sending survey email and following up again.
    - Whether or not a customer belongs to this subgroup is stored in a binary "Instrument variable"
    - Can be a one-time effort

- **Option 2: Soft approach**
    - While sending most customers just 1 invite emails
    - You "stalking" a subset of customers with more emails
    - The number of emails sent is stored in a so called "instrument" variable

CX.AI

Lesson #3: Controlling the bias with modeling

- **Items or topics**

- Context variables like demographics, segment, etc.

- **"Instrumental" variables** >

MODEL 1
(to be build by ML)

→ **NPS rating**

MODEL 2
(to be build by ML)

→ **Items or topics**
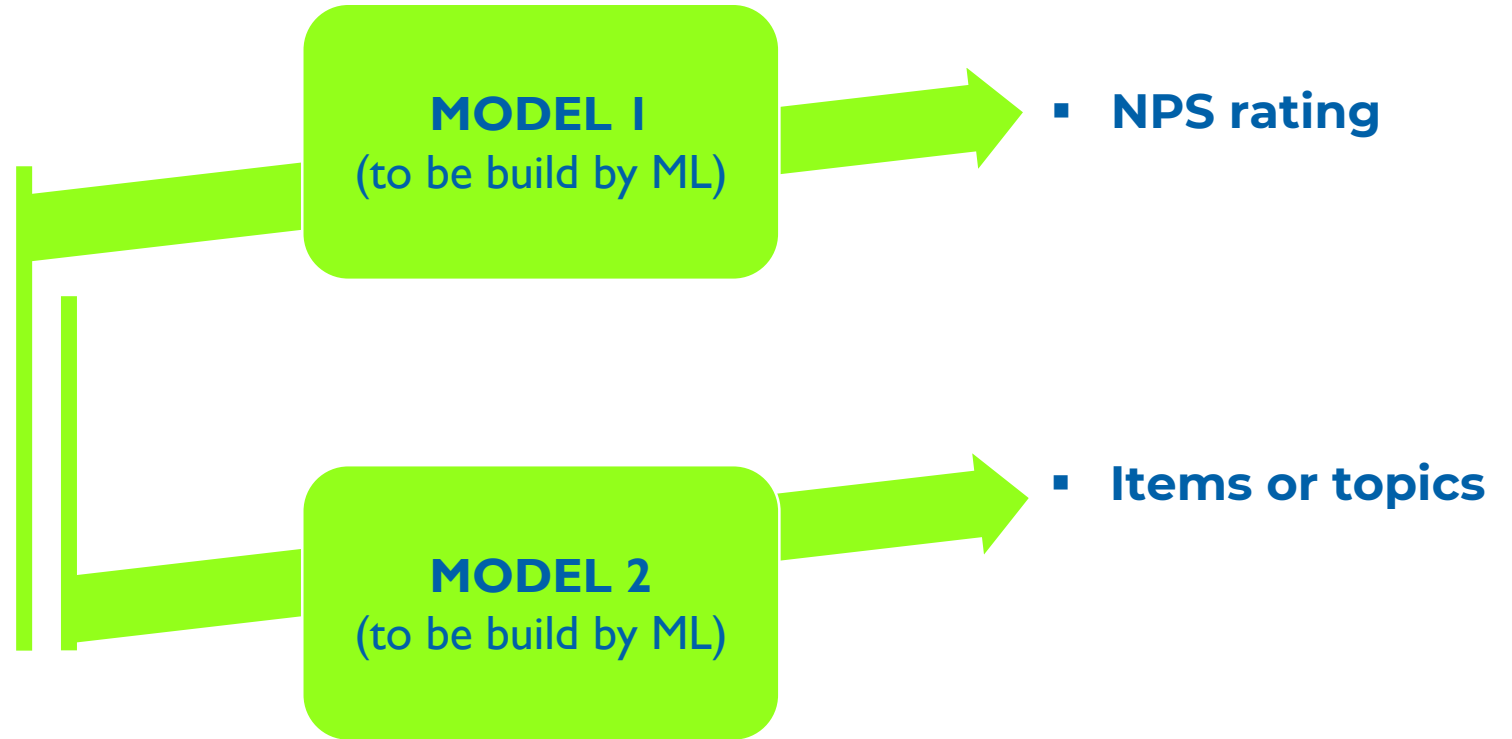
If an instrument variable influences BOTH, outcomes (NPS rating) AND items/topics, then we have biased results.

CX.AI

Summary Class #7

1. Selection bias is when the reason for participation is linked to the answers.

2. It is only harmful when biasing outcomes AND drivers.

3. Can be controlled with better responds rate and with modeling

4. Modeling can measure biases and clean results from them.

CX.AI