# Scoping Pair Discussion Points

Below are examples of discussion points that may be used to facilitate the pair

## Exercise 1: Admissions Team Help

Problem: Decrease the number of applicants who drop out of the application process

Definitions to get clear on:

- What do we mean by "drop out", i.e. how is it calculated?
- Focus on decreasing the **number** or **percentage** of drop outs? Need to account for the possibility of our actions affecting the number of initial applicants
- What is the current number and percentage of drop outs?
- What exactly is the desired impact resulting from "decreasing the number of applicants who drop out of the application process"?
    - More students enrolled? Graduated?
    - What about background/skill sets of enrolled/graduated students? Should it remain the same?
- Any constraints? e.g. No more than X more hours of Admissions team time can be spent implementing the recommended action(s)

What types of actions are available for the Admissions team to take? For example:

- More or different types of outreach to students identified as at risk of dropping out
- Altered application process
- Adjust our target applicant pool

Potential data sources:

- Applicant info from application form
- Sourcing info (how did they arrive at the application?)
- History of applicant interactions
- Applicant performance in each part of the process
- Student/applicant survey results

Types of analyses & desired findings:

- Exploratory analysis to describe the data we have:
    - What type of applicants are dropping out?
    - At what point in the application process?
    - Any commonalities between applicants who drop out?
- Classification model to predict whether an applicant will drop out at some point in the process; Admissions team can pay extra attention to and take additional action on those students predicted as likely to drop out
- Unsupervised clustering to better understand the types of Metis applicants and their respective drop-out rates; each cluster might warrant a different intervention

Validation & timeline:

- Any sort of intervention/action coming out of this analysis should be tested on a subset of students in the upcoming admissions cycle(s)

- Comparing the results of this experiment against typical drop-out figures will reassure us that the results are as expected, in the desired direction of lower drop-outs

Potential risks:

- Sample size
  - Need to make sure our results are statistically significant, especially in regards to the validation method described above
- Effects on quality of Metis experience for students and staff
  - Need to make sure these actions don't have adverse effects on parties involved

## Exercise 2: Shipping Company Safety

Problem: Decrease the number of employee accidents in warehouses

Definitions to get clear on:

- What is the current number of accidents overall? In each warehouse?
- Do we want to reduce the overall number of accidents? Or per-warehouse?
- Aside from the obvious goal of improved employee safety, are there other desired impacts from "decreasing the number of employee accidents in warehouses"? For example:
  - Improved employee morale
  - Lower insurance costs
  - Improved public opinion
- Any constraints? For example:
  - No more than $X can be spent implementing the recommendation action(s)
  - Order processing time and volume must remain the same across all warehouses

What types of actions are available for the executive team to take? For example:

- Implement new rules and regulations across all warehouses
- Adjust order processing time / volume, or allocation of orders between warehouses
- Provide incentives for accident-free warehouses

Potential data sources:

- Accident reports
- Characteristics of each warehouse: size, age, shift schedule,
- Order history: what's being processed, at what speed
- Employee survey results
- Rules and regulations in each warehouse
- Industry safety best practices

Types of analyses & desired findings:

- Exploratory analysis to describe the data we have:
  - When, where, what type, and to whom are most of the accidents happening?
  - What is the most-reported reason?
  - Any patterns over time?

    - Any patterns in relation to order volume/speed or warehouse characteristics/rules?
- Regression model or time series to predict the number of accidents that a warehouse will encounter in the next X days
    - Warehouse management team can take additional preventative measures when predictions are high
    - Interpret the coefficients of regression to understand the relationship between features and predicted number of accidents; causal relationships will need to be tested by setting up experiments
- Natural language processing on accident reports and employee surveys to understand the context around existing accidents
- Unsupervised clustering on accident reports and warehouse data to better understand the types of accidents; each cluster might warrant a different intervention

Validation & timeline:

- Any sort of intervention/action coming out of this analysis should be tested on a subset of warehouses over following few months
- Comparing the results of this experiment against typical employee accident figures will reassure us that the results are as expected, in the desired direction of fewer accidents

Potential risks:

- Perverse incentives
    - Need to avoid unwanted consequences of incentives, such as under-reporting of accidents to meet new rules and/or incentives
- Inconsistent roll-out:
    - Any recommended actions will eventually need to be rolled out to specific warehouses; executive team must take extra steps to ensure that roll-outs are implemented correctly and consistently, to achieve the desired results