Sandia Data Challenge

UIUC FALL ‘25

Haley Kharvari, Julius White, Brisa Jasso, Chris Boukalis

Brisa Jasso

Statistics, minors in Math and Data Science

Junior

Chris Boukalis

Finance and Data Science, minor in Statistics

Junior

Haley Kharvari

Statistics and Economics

Junior

Julius White

Computer Science and Economics

Senior

About us

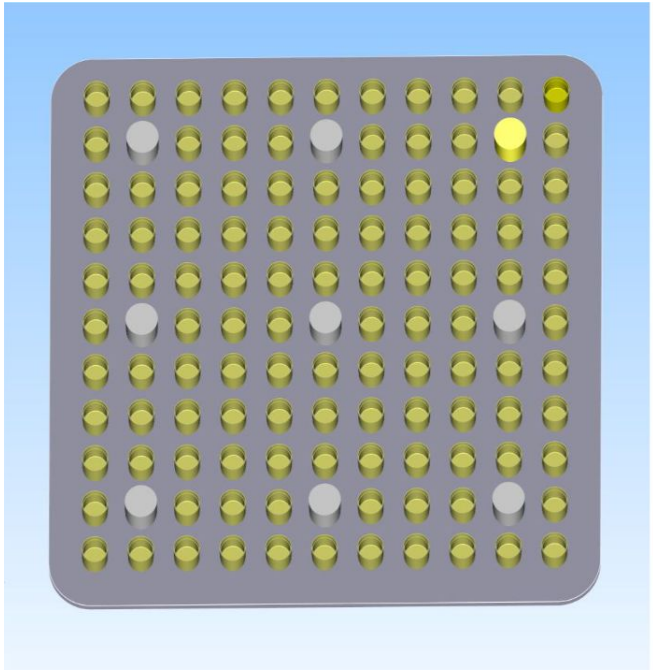
1. Data Problem
2. Preliminary Analysis
3. Method
4. Results
5. Conclusion

Data Problem

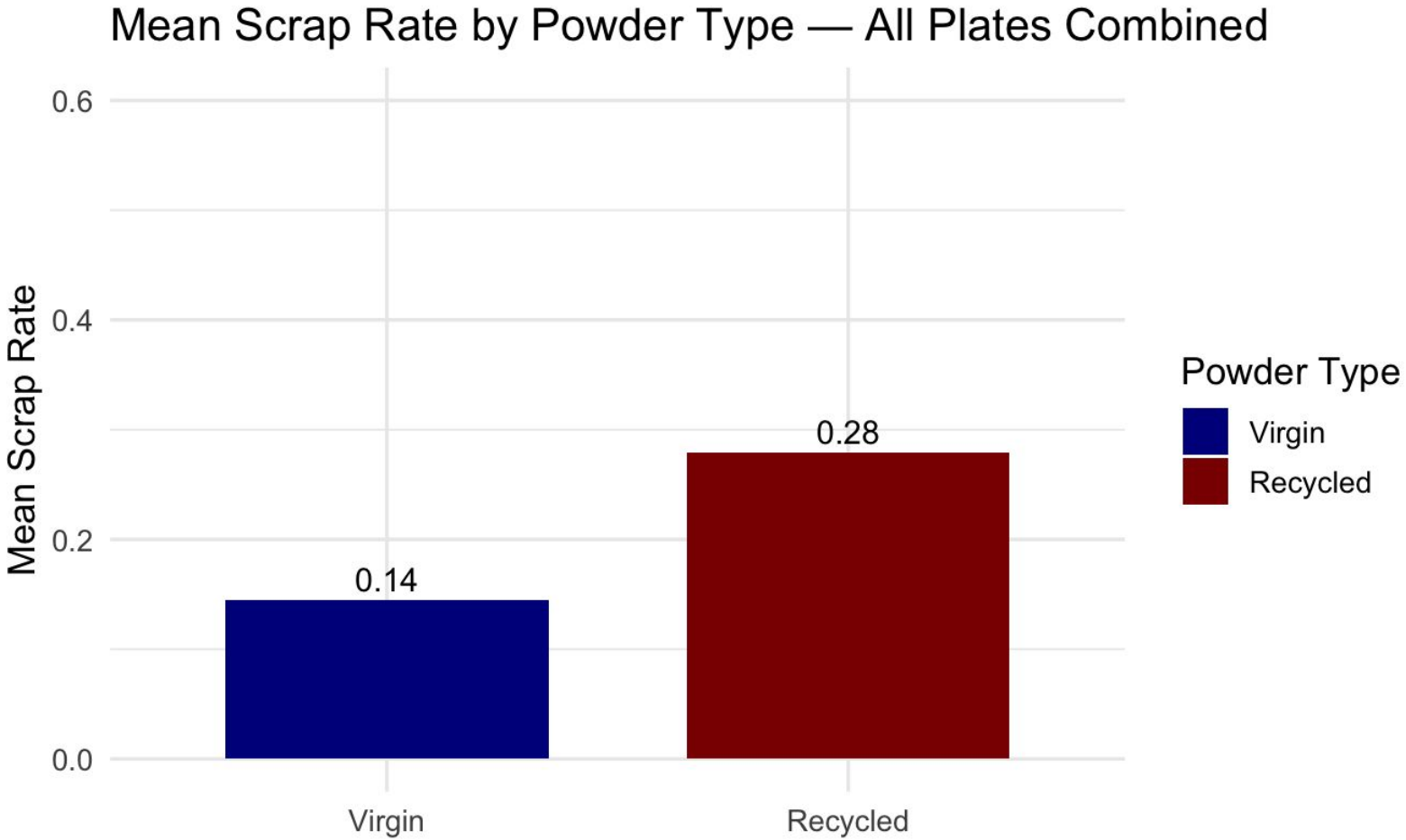
What combinations of build parameters minimizes the potential scrap rate/reduces the probability of scrapping?

Our team focused on four main parameters:

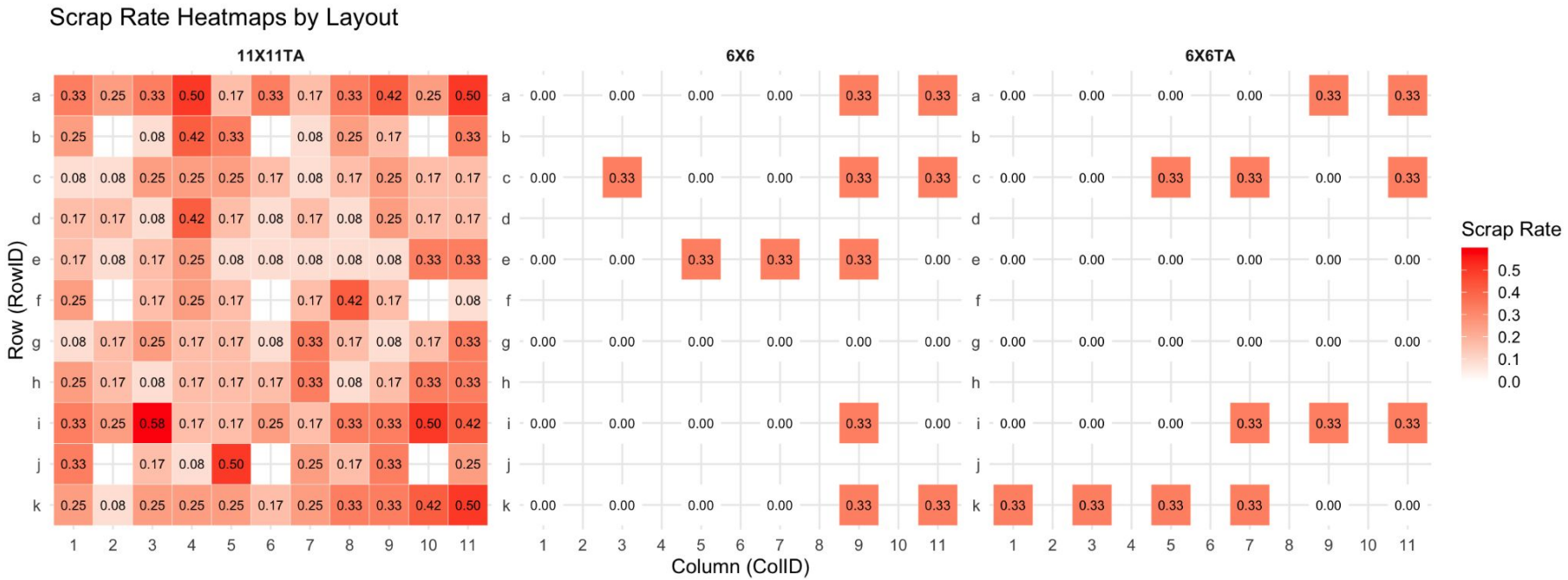
- Build Plate Layout
- AM Powder Type
- Location on the build plate
- Existence of test artifacts



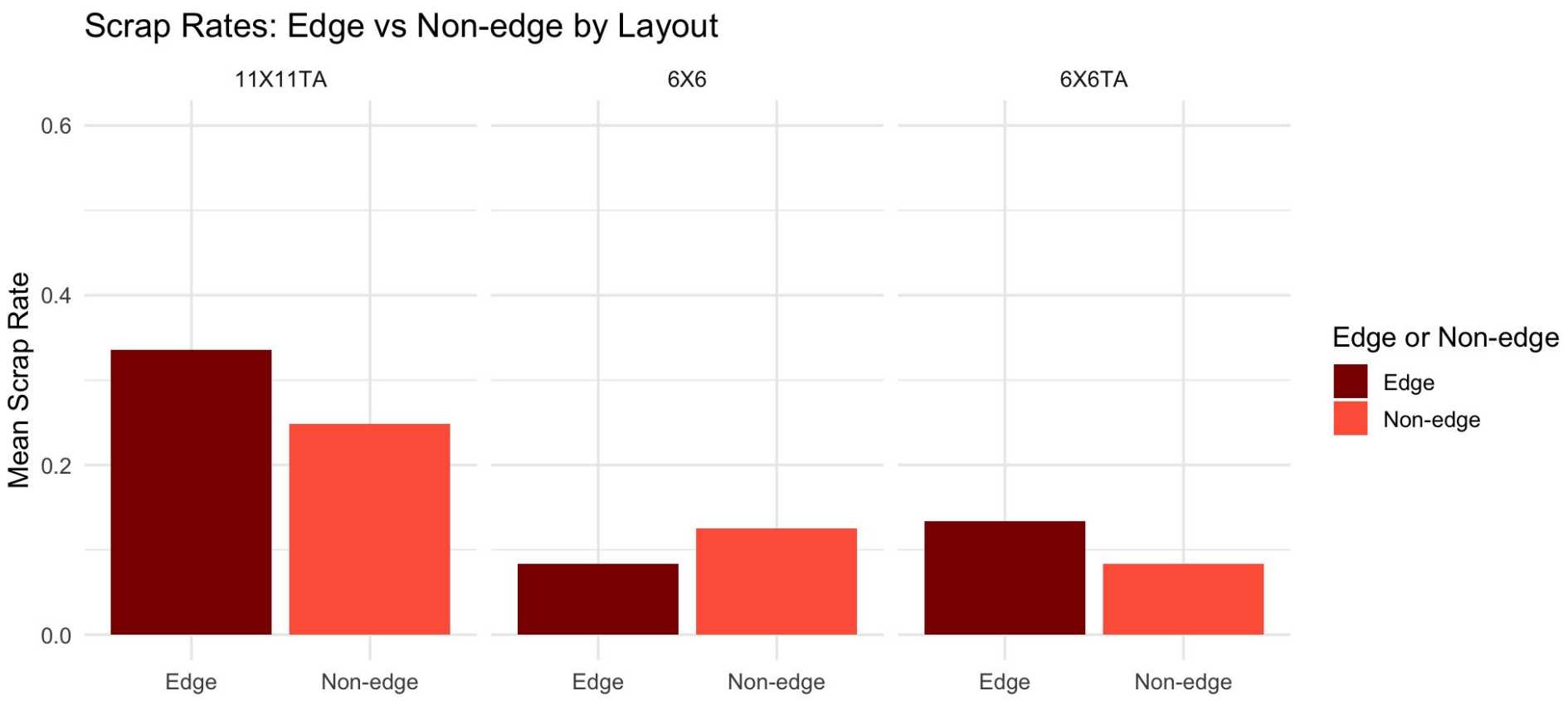
Preliminary Analysis



Preliminary Analysis Graphs



Preliminary Analysis Graphs



Preliminary Analysis Graphs

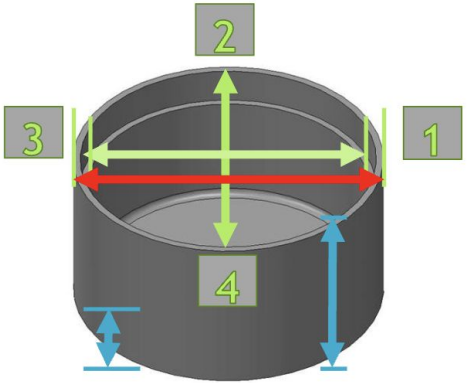
Method

Features:

- Layout
  - 6x6
  - 6x6 TA
  - 11x11 TA
- Powder
  - Virgin
  - Recycled
- Zone
  - Edge
  - Between
  - Center

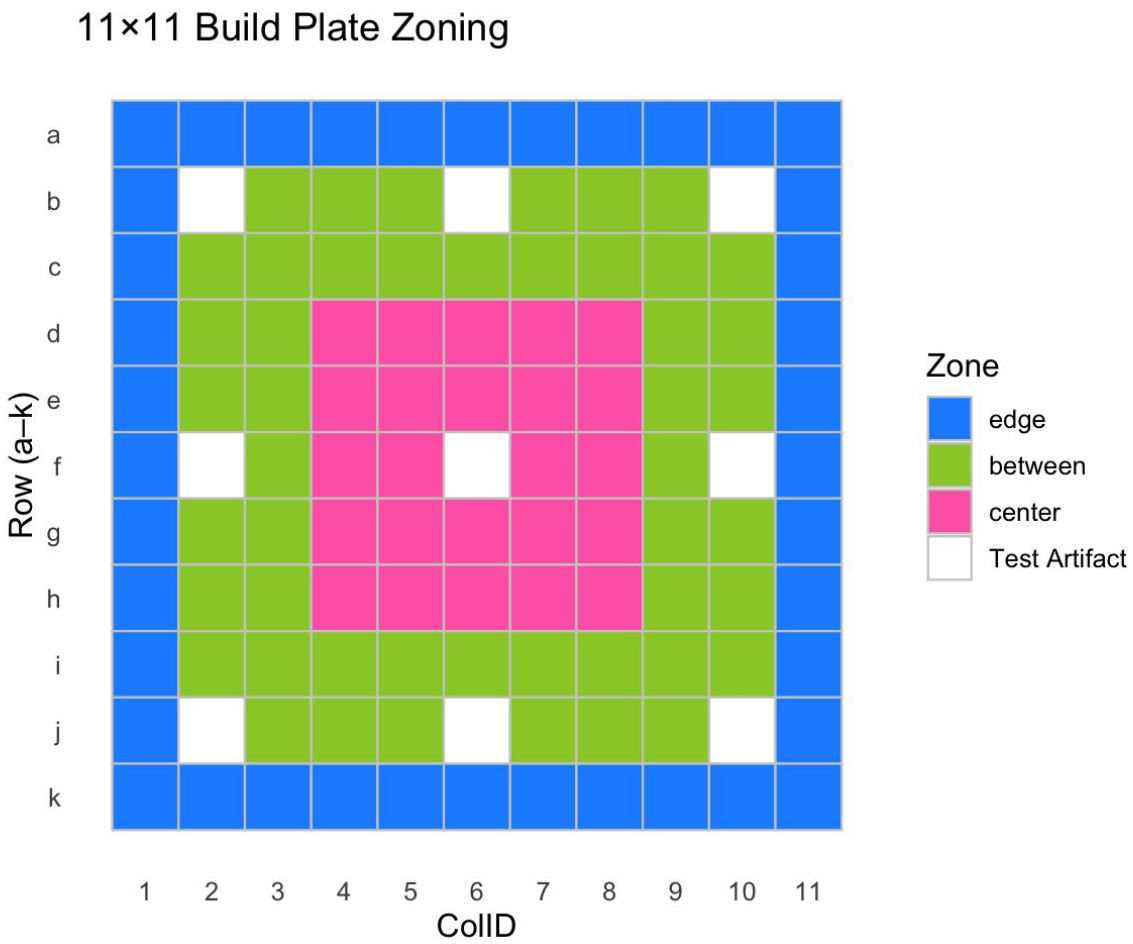
Response:

- Scrap
  - Each cup was determined to be a scrap (1) or not a scrap (0)



Responses (Form-Fit Geometry from CMM)	
Lip Interior Diameter	B3_DATUM_B_LOC [0.415”, 0.435”]
Lip Exterior Diameter	B3_REF_OD [0.445”, 0.469”]
Floor Height (with standoff material where applicable)	C1_LOC_INSIDE_PLN [0.049”, 0.069”]
Lip Height (with standoff material where applicable)	C4_LOC_TOP_PLN [0.261”, 0.281”]
Lip thickness (Measured at four locations around CUP lip)	B3_THICK1_WALL B3_THICK2_WALL B3_THICK3_WALL B3_THICK4_WALL [0.010”, 0.017”]

Method


11×11 Build Plate Zoning
Row (a–k)
a
b
c
d
e
f
g
h
i
j
k
1
2
3
4
5
6
7
8
9
10
11
ColID
Zone
edge
between
center
Test Artifact

Method

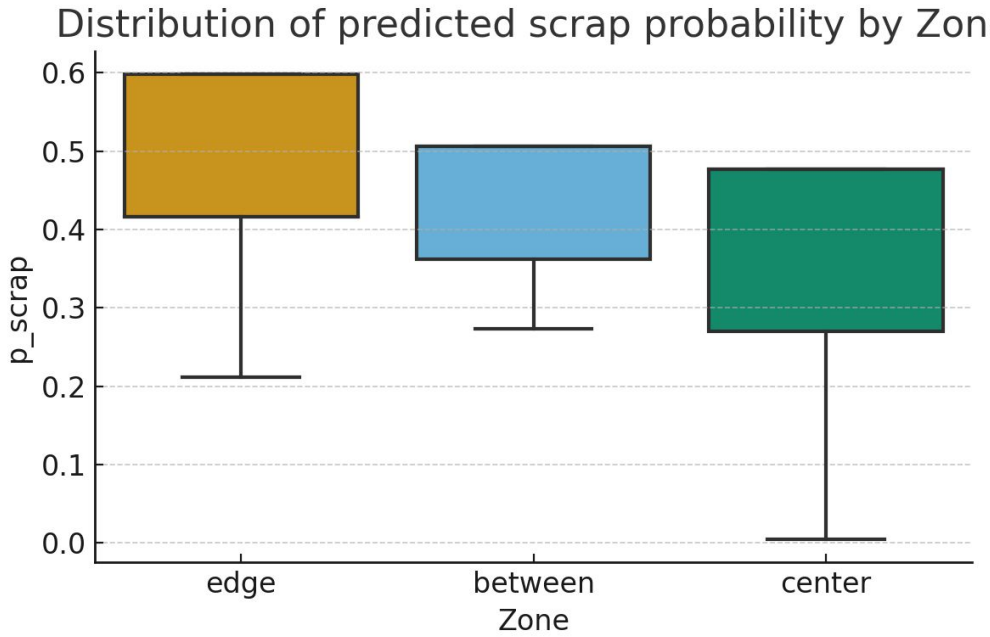
Model Goal: Learn which setups make scraps the least likely.

- Extreme Gradient Boosting (XGBoost)
  - Decision Tree based model
  - Feature importance insights, handles categorical data, imbalanced data

- Improvements made
  - Stratified and balanced class weights
  - Cross-Validation
  - Hyperparameter tuning

Results

Our model found:



	Layout	Powder	Zone	p_scrap
0	6X6TA	Virgin	center	0.004657
1	6X6	Virgin	edge	0.211493
2	11X11TA	Virgin	center	0.269927
3	6X6TA	Virgin	between	0.273675
4	6X6	Virgin	between	0.281289
5	6X6TA	Virgin	edge	0.319819
6	11X11TA	Virgin	between	0.362018
7	6X6	Virgin	center	0.385710
8	11X11TA	Virgin	edge	0.416335
9	11X11TA	Recycled	center	0.476952

Conclusion

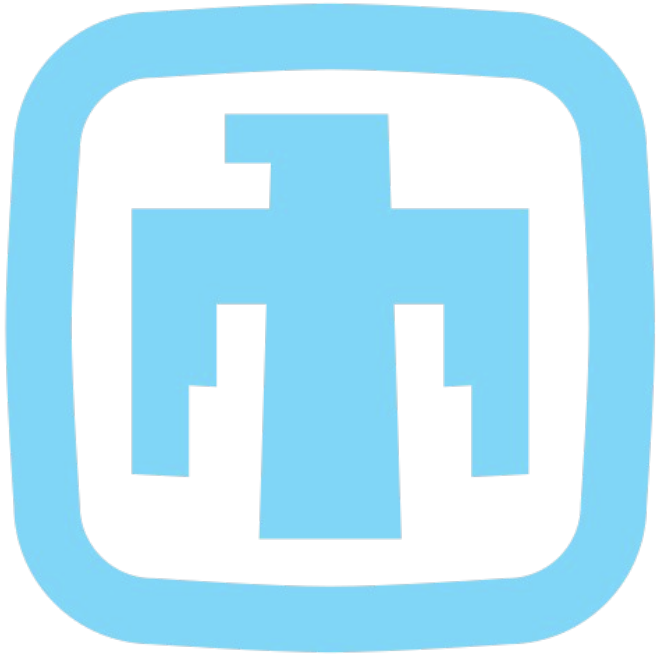
Our analysis highlights how our models can drive meaningful impact and address real-world challenges

Data Improvements for more impact:

- heat transfer analysis
- sample diversity
- increase in sample size

We aim to grow these skills further:

- model identification
- data organization



Thank You!