December 1st 2025

Does Increased AI Adoption Truly Impact the Job Market as Much as People Claim?

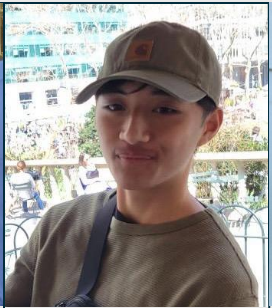
Presented By: James Apuyan, Luis Palacios, Louis Cheung, Sudipti Dantuluri, Jayden Lee, Alex Toohey (TA), and Sarah Zuo (Tutor)

Table of Contents

Meet the Team!

Team Members

Sudipti Dantuluri
Member



James Apuyan
Member

Louis Cheung
Member

Luis Palacios
Member

Jayden Lee
Member

Our Awesome Tutor & TA!

Alex Toohey
TA

Sarah Zuo
Tutor

Dataset & Google Colab Link:

Final Dataset

- Spreadsheet

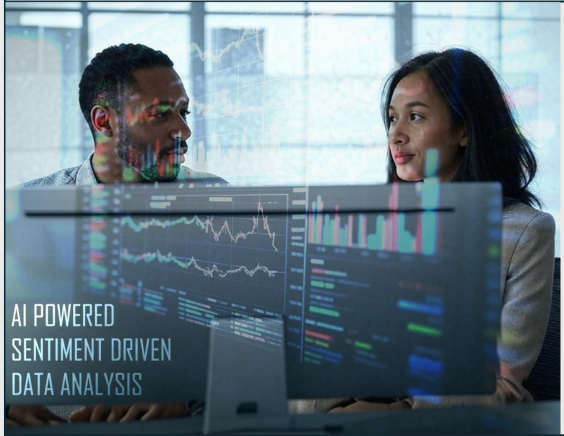
Datasets

- Global AI Tool Adoption Across Industries
- Industry Layoffs 2020 - 2023
- Layoffs Dataset 2024

Google Colab

- DSS Decal Fall '25 Colab

Background



- We wanted to explore trends related to how using AI impacted the job market i.e. displacement
- Used multiple datasets that contained industry, lay-off and AI adoption rates
- After cleaning and exploring, we visualized and standardized the data
  - Focused on layoffs; cleaned & combined
- Then aggregated the data into a final set

- Exploration (importing sets, filtered nulls, data types) → Cleaning (checking duplicates while putting data together) → explored industry, standardized them, applied their visualizations (industry and country mapping) → Aggregation to final dataset

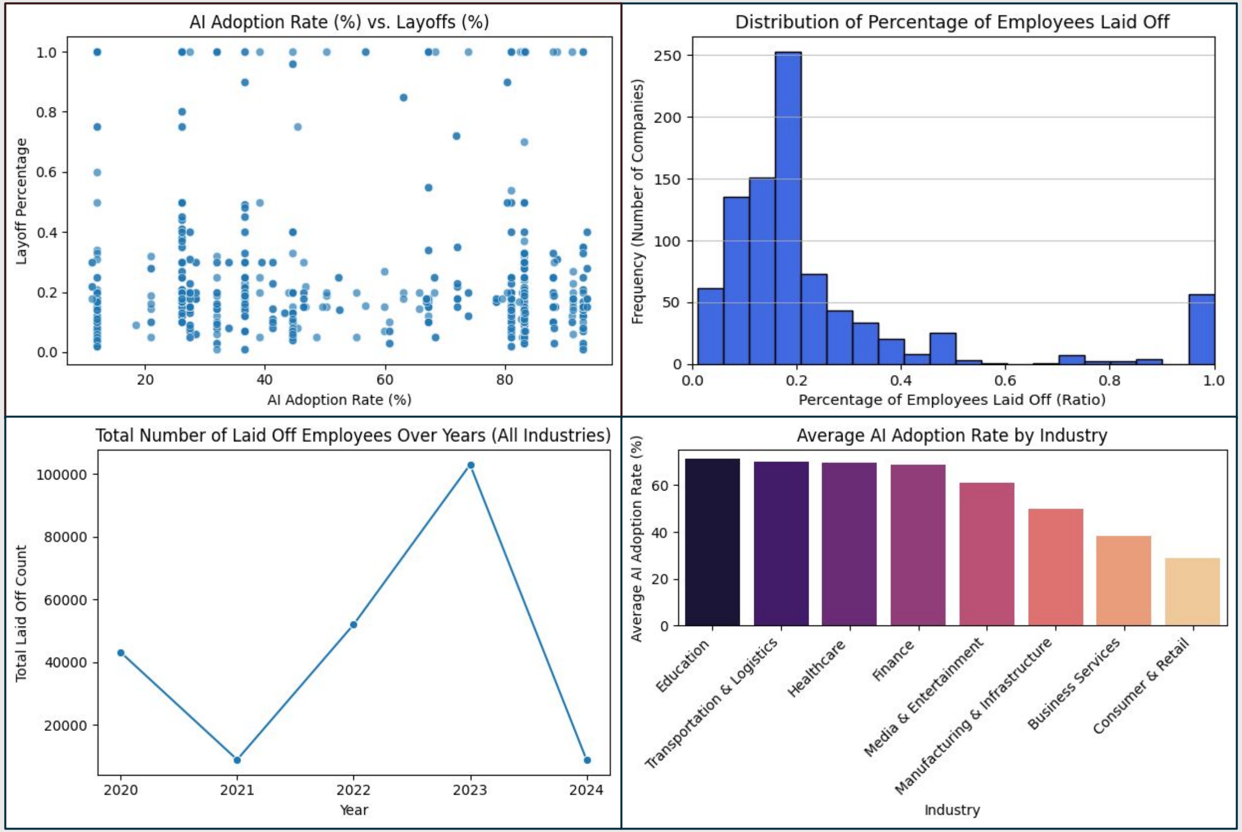
Individual Dataset Cleaning

	Previously	Currently	
Data Types	• Inconsistent data-type formatting across datasets • Numeric fields stored as text • Year fields inconsistent or unstandardized	• All datasets use consistent and correct data types • Numerical values properly cast as integers or floats • Categorical fields standardized across datasets • Years formatted consistently for comparison and merging	We verified and standardized data types across each dataset to ensure all numerical, categorical, and date fields were properly formatted before analysis.
Null Values	• Layoff datasets contained multiple missing numeric values • Some columns were almost entirely null • No imputation or null-handling applied yet • Risk of inaccurate analysis if left uncleaned	• Missing values imputed using industry-wise medians • High-null or unusable columns were removed • Layoff datasets now contain complete or meaningfully filled fields • No remaining nulls in critical analytical columns	We identified missing values, especially in the layoff datasets, and applied industry-specific median imputation and removed unusable columns with excessively high null rates.
Duplicates	• Potential duplicate rows due to multi-source data	• Zero duplicates confirmed across all datasets • Clean, de-duplicated records ready for merging • Dataset integrity validated before integration	We checked all datasets for duplicate rows using .duplicated() and confirmed that none were present, ensuring clean, consistent records.

Data Merging Process

	Previously	Currently	
Mapping	• Industry names did not match across datasets • Countries labeled inconsistently • Overlapping or ambiguous category labels • No unified grouping system for merging	• Industry names standardized into unified categories • Country names mapped to a consistent format • All datasets aligned on *Industry*, *Country*, and *Year* • Mapping enables accurate grouping and merging across datasets.	We standardized industries and countries across all datasets using custom mapping dictionaries. This ensured consistent labeling so that merging on key fields like Industry, Country, and Year would be accurate.
Merging	• Layoff data separated across two files (2020–23, 2024) • AI datasets and layoff datasets used different column naming conventions • Nulls and type mismatches prevented proper joining • Merging at this stage would have resulted in missing or misaligned records	• Layoff datasets successfully combined into a single dataset • AI industry, tools adoption, and layoff data merged into one final table • Fully aligned columns with consistent labels and formats • Complete unified dataset ready for analysis and modeling	After cleaning, imputing nulls, and standardizing labels, we merged the layoff datasets into one combined file, then integrated it with the AI adoption and industry datasets to form a unified analytical dataset.

Exploratory Data Analysis (EDA)

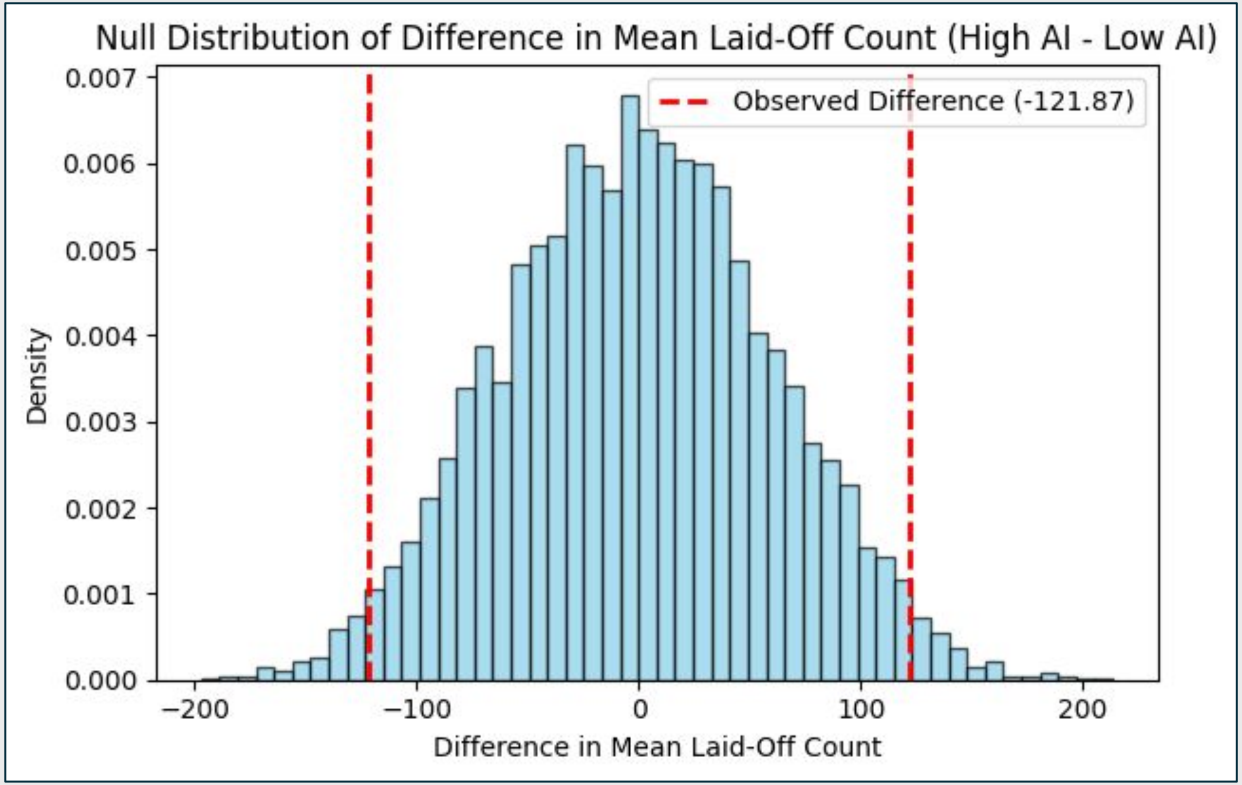


The EDA stage was crucial for understanding the **foundational properties** of our dataset. After cleaning and finalizing the data, our primary goal was to inspect the s**tatistical distribution** of individual variables and uncover **initial patterns** and **simple associations** between them.

Pictured to the left are some of our attempts to **visually inspect** the data.

This phase allowed us to formulate hypotheses and identify areas for **deeper modeling**.

EDA: Bootstrapping

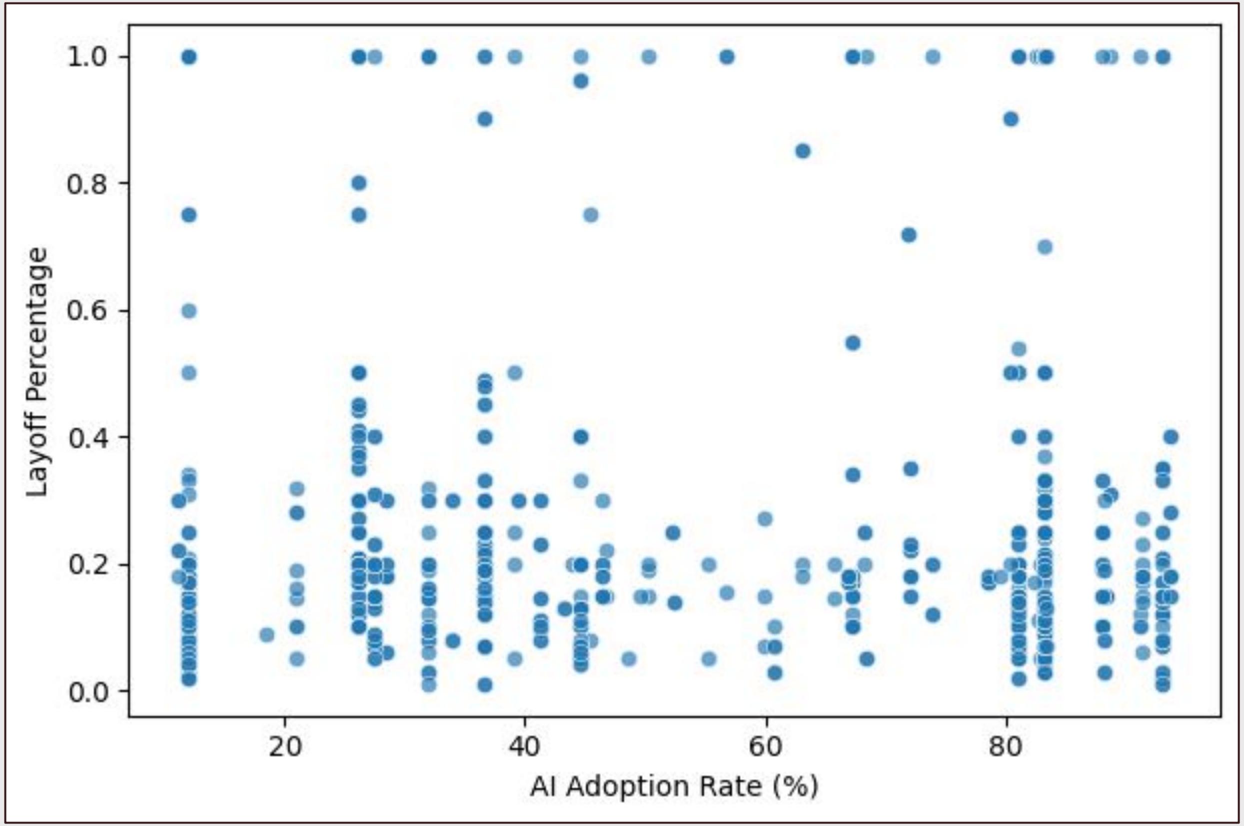


To test for a significant relationship between AI adoption and layoffs, we **bootstrapped** our data, comparing the mean layoff count of **High vs. Low AI Adoption** groups.

The observed difference in means resulted in a p-value of **0.0386**. Since this value is less than **5%**, we **rejected the null** and concluded there is a statistically significant difference...

Companies with Lower AI Adoption had a **higher average layoff** count than those with High AI Adoption.

AI Adoption Rate (%) vs. Layoffs (%)

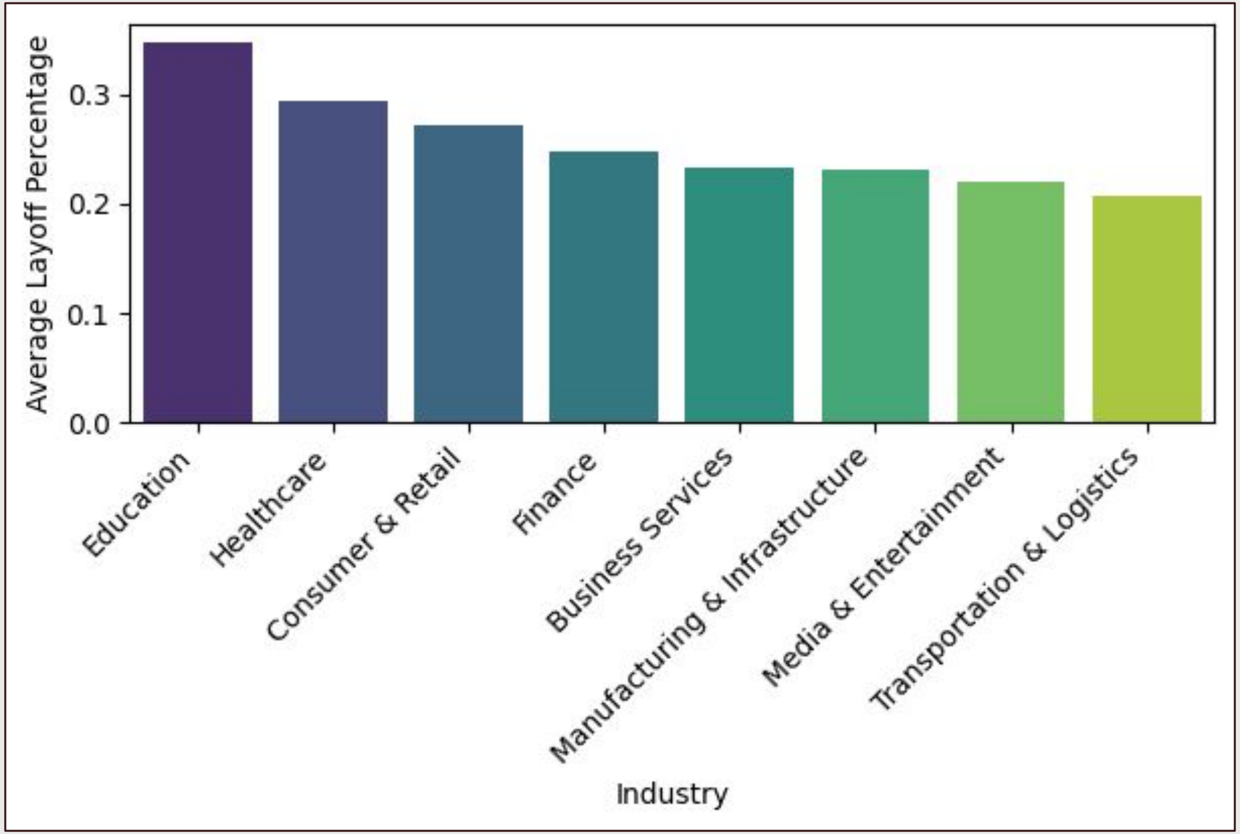


The Scatter Plot of AI Adoption Rate (%) versus Layoffs (%) suggests a **weak but potentially inverse** relationship.

The high density of data points shows that many companies, **regardless of their AI adoption level**, experienced low layoff percentages.

However, the most severe layoff events appear to occur more frequently in the **mid-to-low range of AI adoption**, supporting our hypothesis test conclusion that lower AI adoption may correlate with higher workforce instability.

Average Layoff (%) by Industry



The Bar Plot of the Average Layoffs (%) by Industry reveals **significant industry-specific risk**.

We calculated the mean layoff percentage for each sector to **isolate systematic differences.** Industries like Education and Healthcare exhibit notably **higher average layoff percentages** compared to others.

This suggests that sector-specific **economic pressures** or **industrial shifts** play a dominant role in workforce reduction, **independent** of the overall AI adoption rates within the sector.

Linear Regression Strategy

Why?

To formally evaluate whether AI adoption influences layoffs, we introduced linear regression as our first modeling approach.
This allowed us to move beyond visual patterns and quantify the relationship between these two variables in a clear and interpretable way.

Our goal was to examine the direction, magnitude, and overall stability of the effect.
By standardizing features and incorporating 5-fold cross-validation, we ensured the results were consistent and not tied to a single sample split.

Using linear regression helped us determine whether the hypothesized link between AI adoption and layoff percentage is supported statistically, or whether the relationship is weaker or more complex than expected.

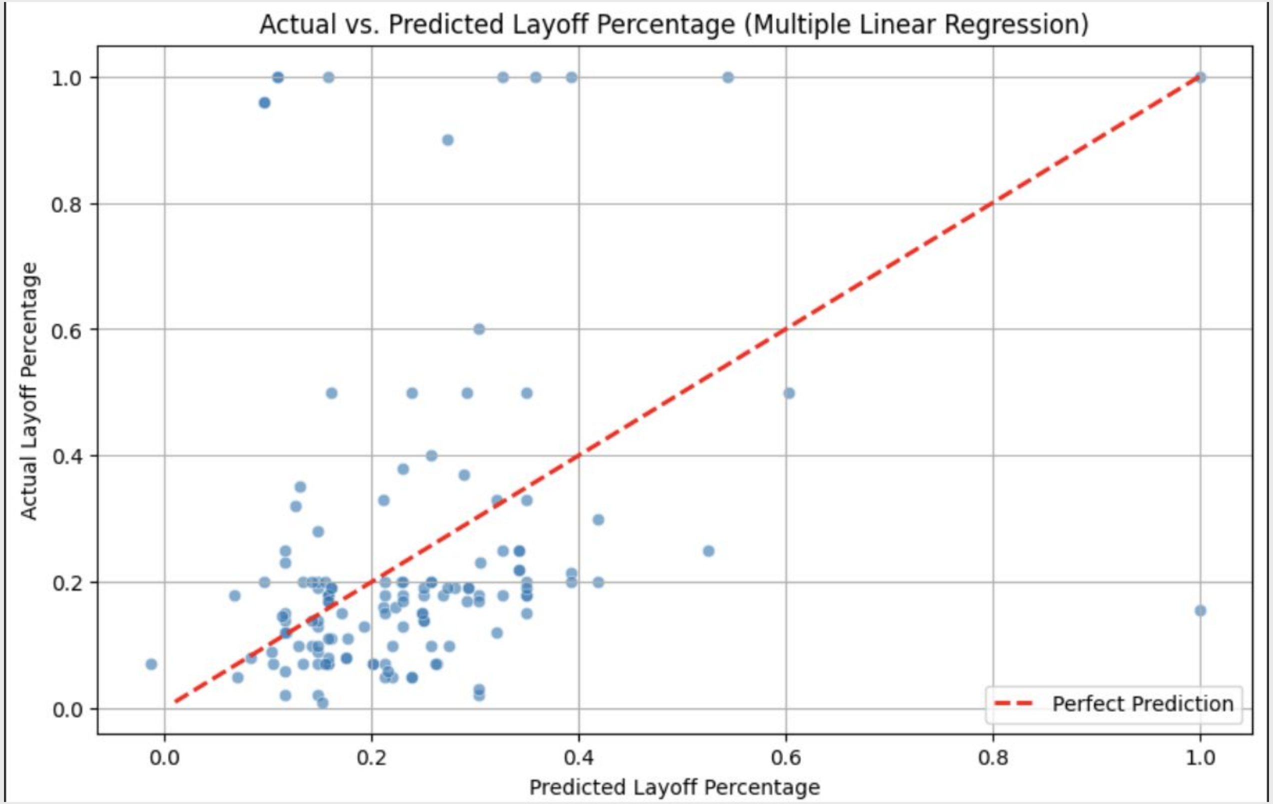
Our Approach

We Built a linear regression model to estimate how changes in AI adoption rate relate to changes in layoff percentage. This involves standardizing all numerical features to ensure the model treated each variable on the same scale and avoided undue influence from large-valued features.

We applied 5-fold cross validation, training the model on multiple subsets of the data and averaging the results. This process allowed us to access how consistently the relationship appeared across different segments of the datasets.

Our evaluation focused on coefficient size, statistical significance and predictive performance.

Linear Regression Results



The scatter plot comparing actual versus predicted layoff percentages highlights the limitations of our linear regression model. The majority of points fall far from the perfect prediction line, indicating that the model struggled to accurately estimate layoff outcomes based on the available features.
While adoption showed a statistically significant positive coefficient, the predictive power of the model is still very weak, with R^2 values close to zero across all validation folds. This means that layoffs are influenced by many complex factors that a simple linear model cannot capture.

Verdict

Does AI adoption truly impact the job market?

1

AI Adoption Rate showed a statistically significant positive relationship with layoffs

2

Across all 5 cross-validation folds, the AI adoption coefficient remained consistently positive

3

95% Confidence Interval: [0.0024, 0.0042]

4

Most models struggled to predict layoff percentages, with many R^2 scores near zero or negative

Why Did Random Forest Perform Best? Results?

The relationship between AI adoption and layoffs is weak and complex, not the simple cause-and-effect that headlines suggest.

Also, a good use of random forest is the implementation of hypertuning specific parameters for the model.

Random Forest Advantages:

- Ensemble of Decision Trees
  - Trains 100+ trees on random data samples (bootstrap aggregating)
- Random Feature Selection
  - At each split, considers only random subset of features
  - Prevents overfitting
- Handles Complexity
  - Captures non-linear relationships (unlike linear models)

Key Takeaways:

1

Statistical significance != Practical significance: Just because an effect is statistically significant doesn't mean it has major real-world impact. Our AI adoption coefficient was tiny (0.003), meaning a 10% increase in AI adoption correlates with only a 0.03% increase in layoff percentage

2

Iteration and experimentation: We tried multiple modeling approaches, added features, tested different algorithms, and had to adjust our expectations based on results

3

Negative results are still results: Poor model performance told us something important: the relationship between AI and layoffs is far more nuanced than a simple causal link

Thank you!