



D.E.A.D Sales Analysis

James Rounthwaite, Jamie Luna, Lana Huynh,
Rachel Roggenkemper, Sophia Chung

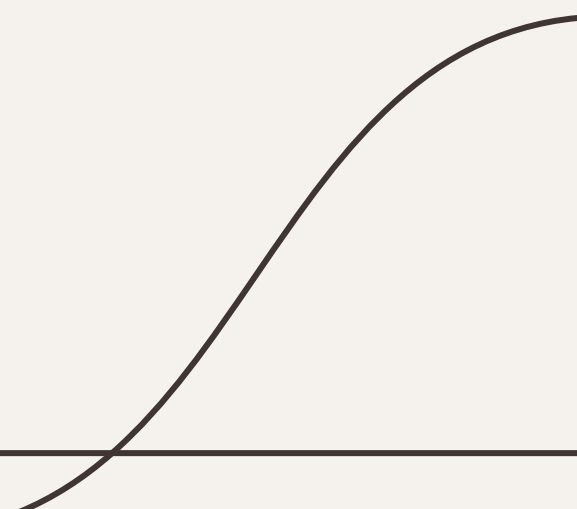


Overview

Provide pattern and factor analysis into alcohol sales

Goal: Model with clear interpretable variables

- Priority: Matching the trend



Data Cleaning & Methodology

- Focused on 1 category at a time
 - Highlight locational and time-based trends
- Models on a per year basis
 - Starting with most recent
- Standardized columns
 - Consistent model interpretation (especially for Ridge and Log-Cosh)

Model Selection Process

- Baseline: SKLearn
- Three separate model innards
 - OLS: Match data as close as possible
 - Ridge: Match data, but punish complexity
 - Patterns to be extrapolatable
 - Log-Cosh: Matching data, but resilient to outliers

Model Performance

Sales = Year + Month + CountyCode_01 + CountyCode_02 + ... + CountyCode_99 + Other

- OLS Regression
- R-squared = 0.775
 - Year, month, and county codes predicts 77.5% of the variability in sales.

Overall Takeaways and Conclusions

- Counties and months are the most impactful in influencing sales

Year	County 1	County 2
2021	Adair	Polk
2020	Story	Polk
2019	Polk	Warren

Thank
you!
