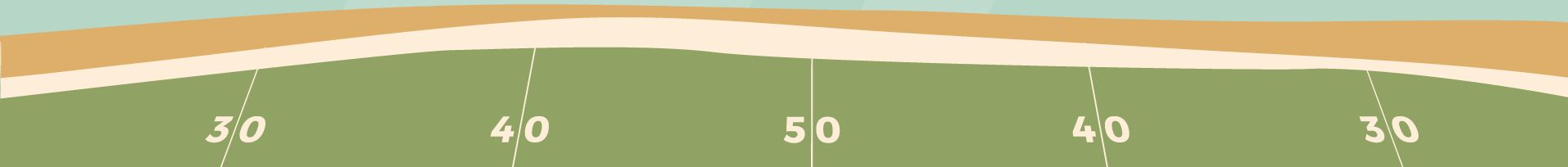
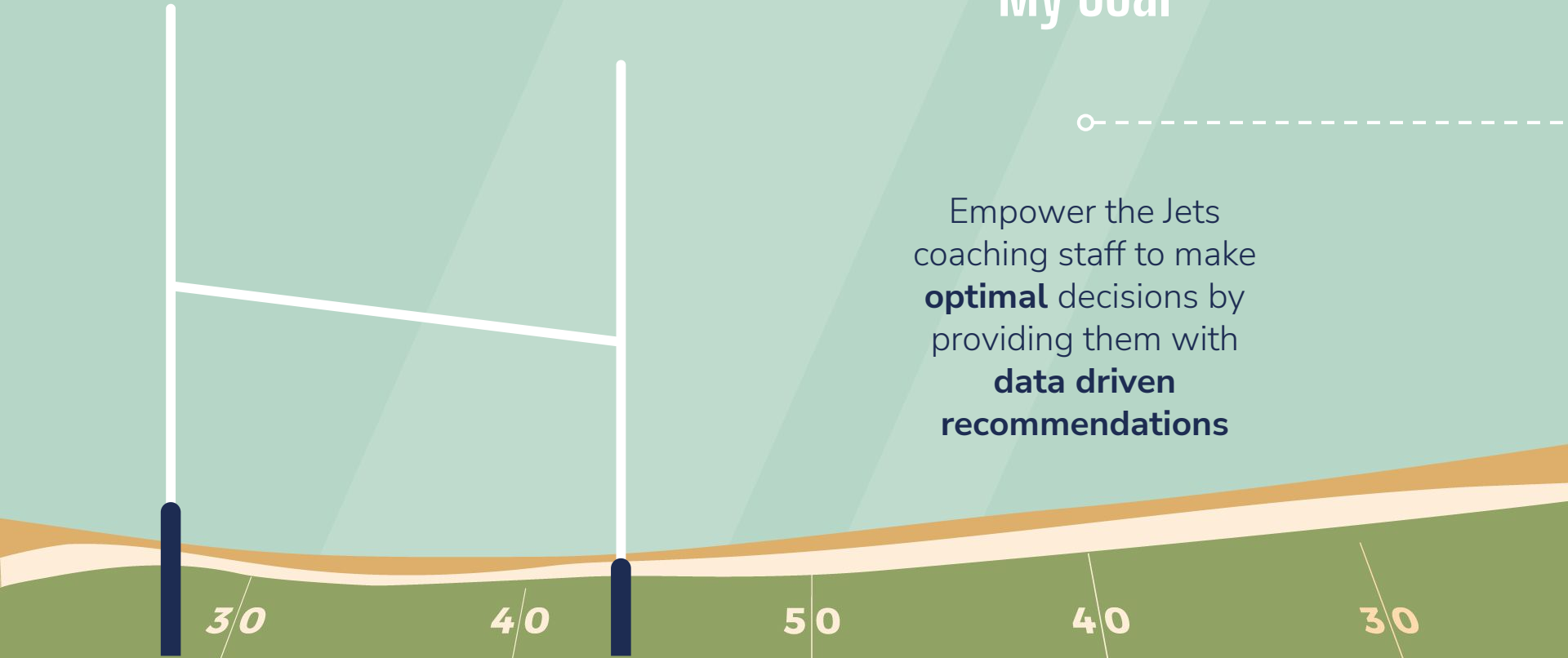
New York Jets
Linear Regression Analysis

TABLE OF CONTENTS



My Goal

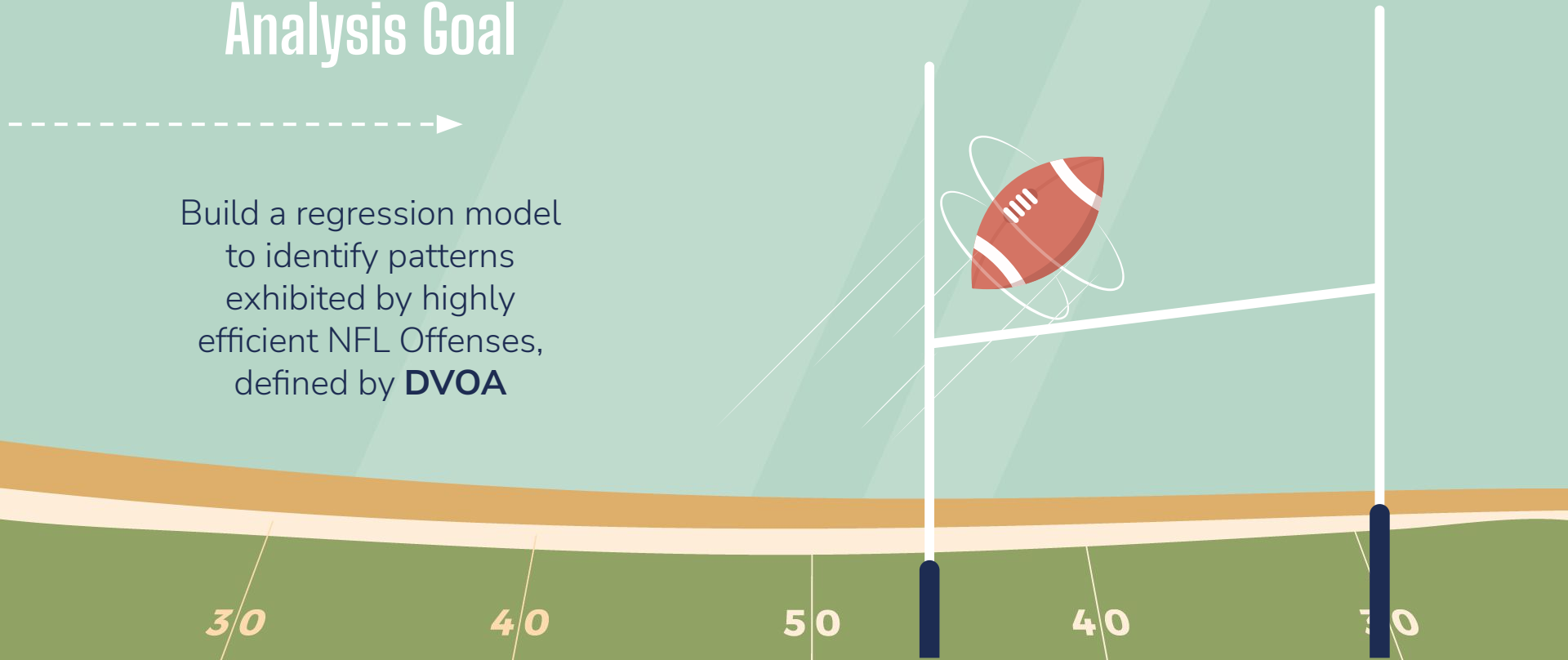
Empower the Jets coaching staff to make **optimal** decisions by providing them with **data driven recommendations**




81

Analysis Goal

Build a regression model to identify patterns exhibited by highly efficient NFL Offenses, defined by **DVOA**



Analysis Breakdown – Target Selection

DVOA - Defense-adjusted Value Over Average

Widely used to compare teams

Compares every NFL play to a league-average baseline

Opponent's strength is considered

Redzone plays count for more

Analysis Breakdown – Feature Selection

Play Call

- 1st Down
- 2nd Down
- 3rd Down

Game Score

- Play Call while Leading
- Play Call while Tied
- Play Call while Trailing

Conversions

- Average yardage until 1st
- Conversion Rates per down

Analysis Breakdown – Tools and Data Used

Beautiful Soup

Se

Web Scraping

pandas

Data Organization

scikit learn

Machine Learning

matplotlib

seaborn

Data Visualization

PRO FOOTBALL REFERENCE

FOOTBALL OUTSIDERS™

Data

Analysis Results – Model Selection

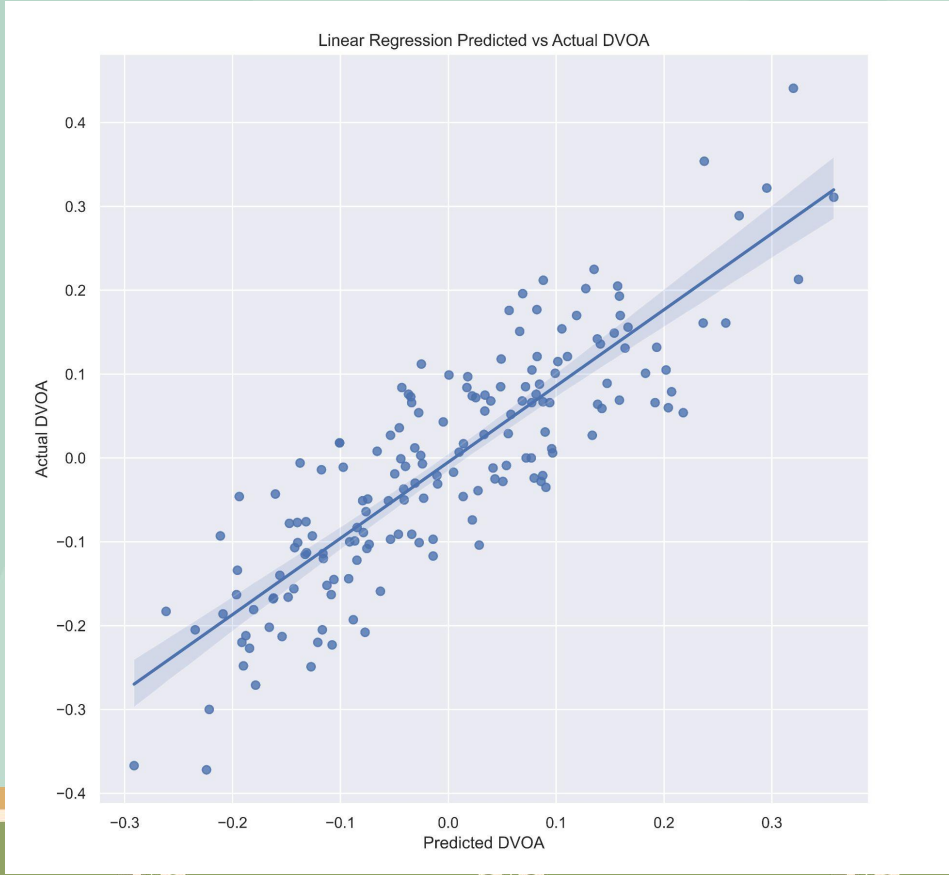
	Linear	Lasso	Ridge	Polynomial	Polynomial Lasso
Valid R^2	77.23%	77.8%	77.3%	70.25%	76.85%
Test R^2	73.59%	73.11%	73.1%	70.89%	73.44%

Analysis Results - Model Selection

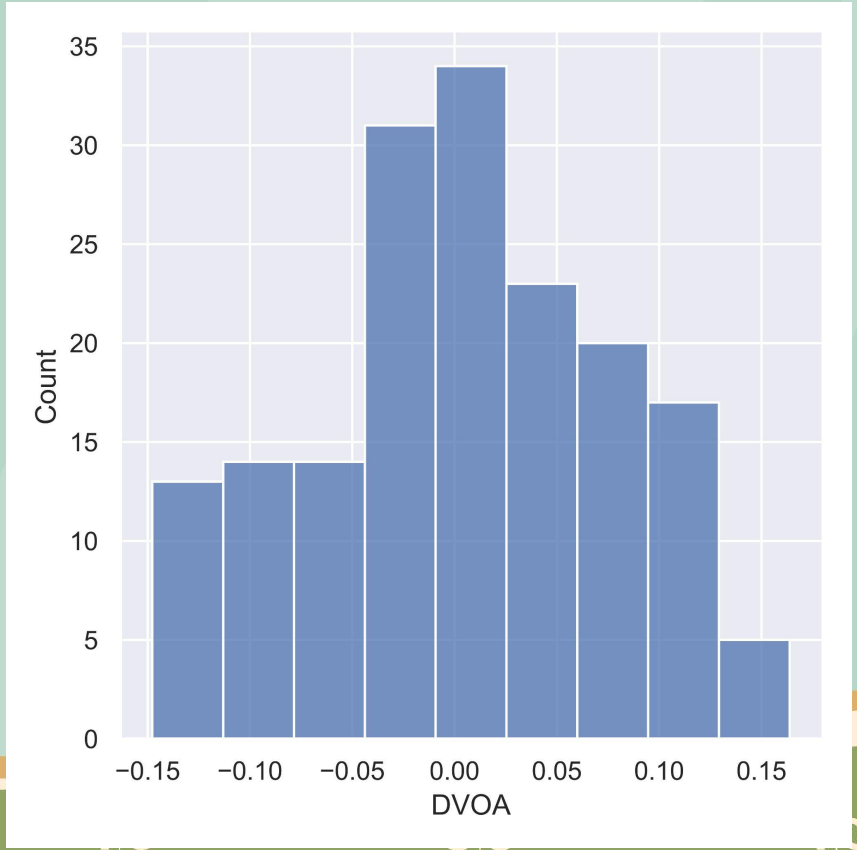
- **Complexity vs. Interpretability Tradeoff -** feature engineering and polynomial regressions weren't worth the lack of interpretability

- **Key Patterns -** Conversion rate, regardless of down, highly related to DVOA

Analysis Results – Linear Regression Prediction vs Actual



Analysis Results – Linear Regression Residuals



Future Work

- **Better Data -** Eliminate “garbage time” stats, focus more on neutral state of game
- **Explore Polynomial Relationships** - Multiplying conversions, passing rates, etc...
- **Explore Other Models** - Decision Trees, Random Forests

THANKS!

Appendix

Analysis Results – Linear Regression Residuals