# F1 Grand Prix Race Win Predictor – Sprint 2

# Reintroduction to the Problem Space

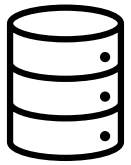
**Predicting F1 winner**: complex problem

**Binary classification problem** : Win (1) or Not Win (0)

**Models for Sprint 2**

- Logistic Regression
- Decision Tree
- Random Forest

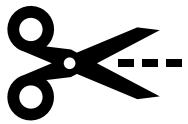
# Data Overview and Changes



**Dataset V2:** focuses on pre-race factors.

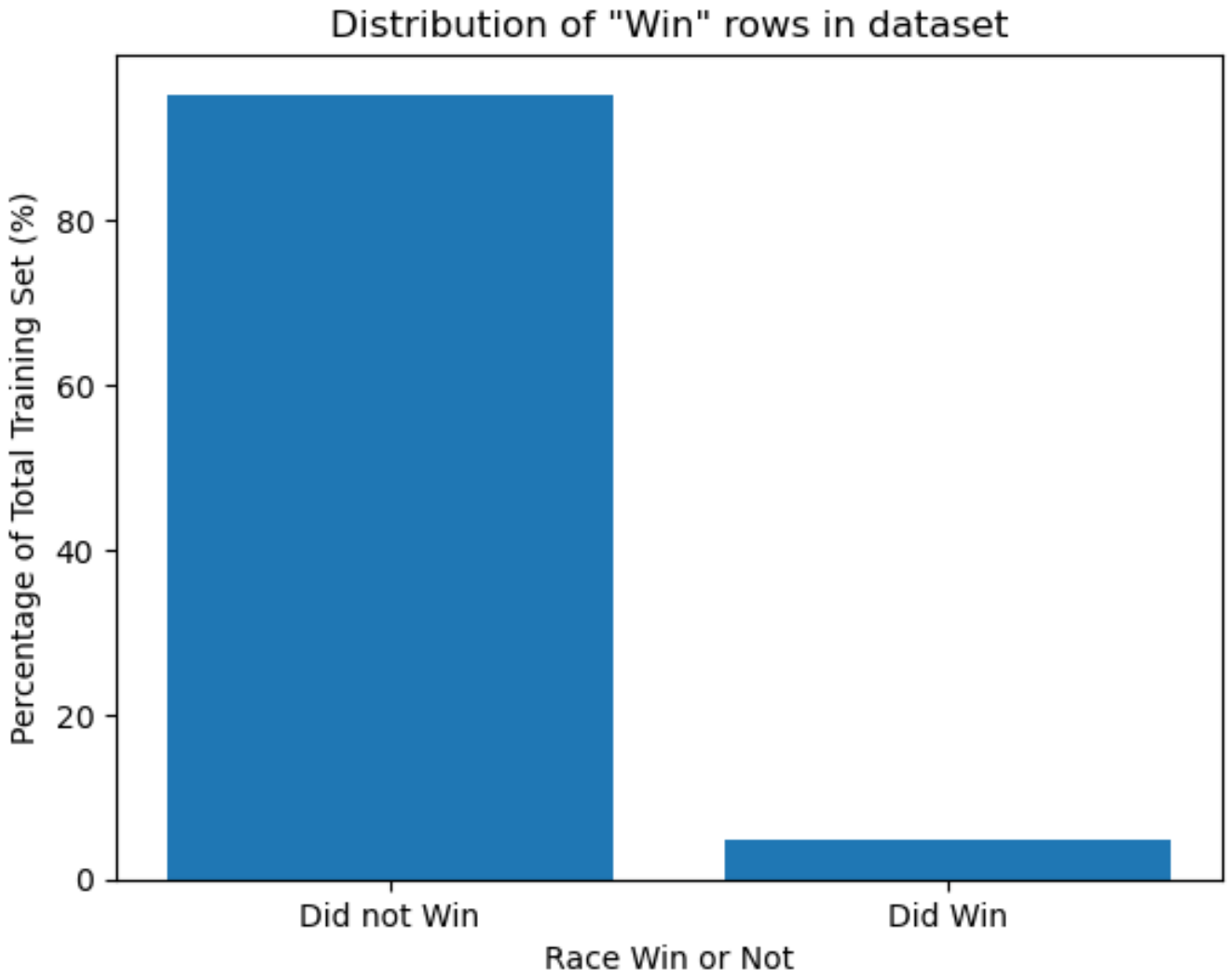
**Excludes in-race statistics:** e.g. pit stops.

**Features engineered:** e.g. cumulative season points

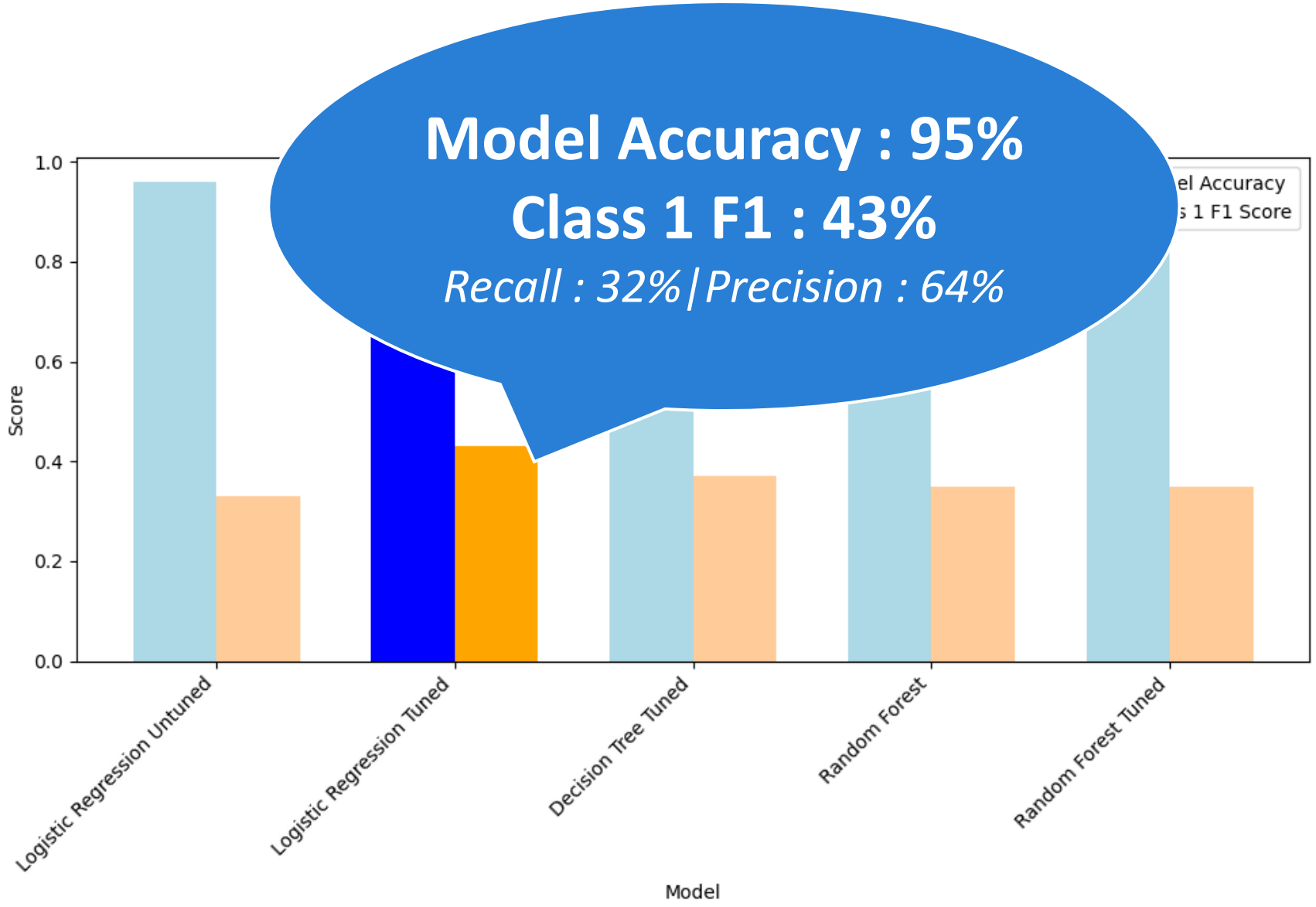


**Original data trimmed** for predictive relevance.

# Dataset is heavily Class Imbalanced…

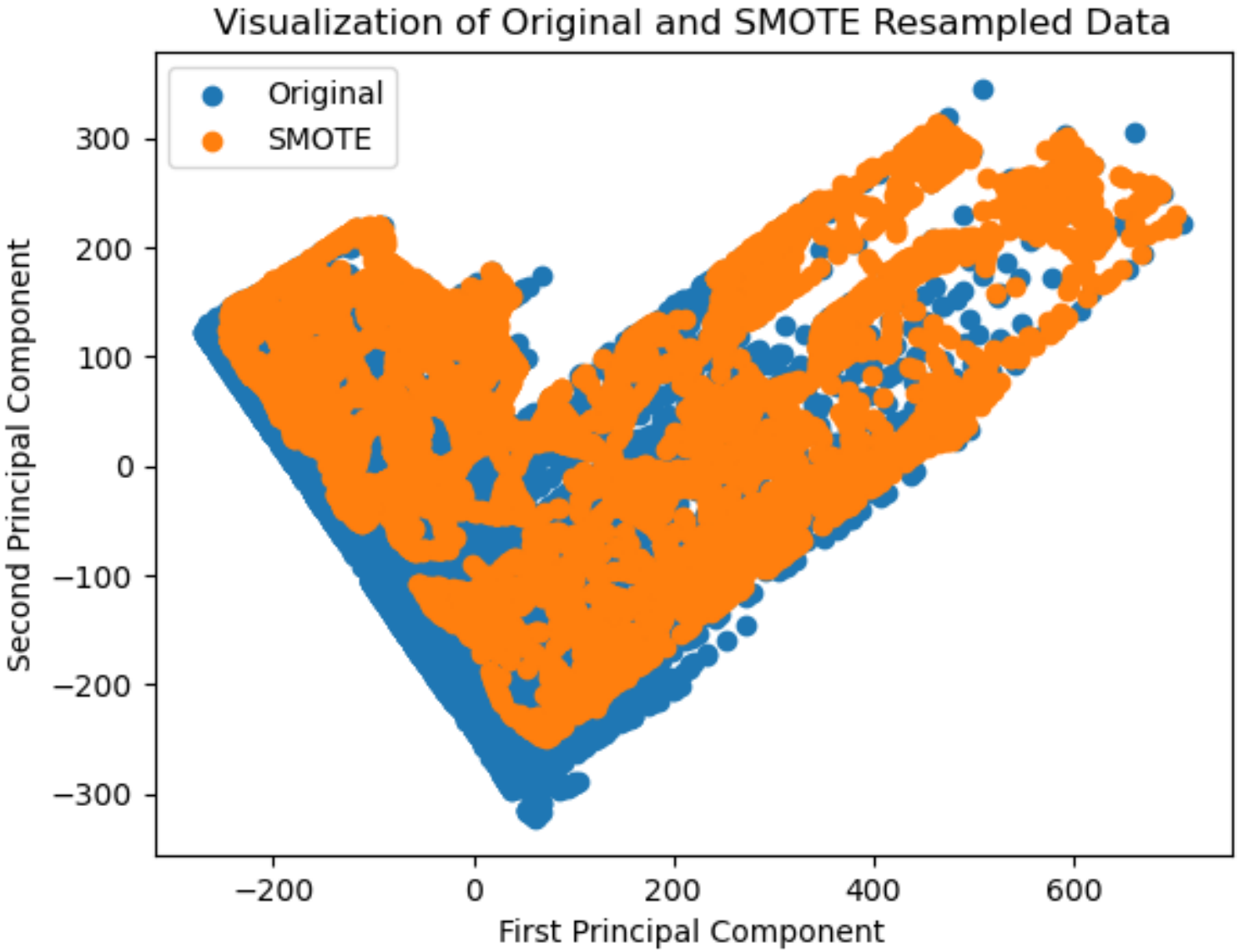


# Results of our first models



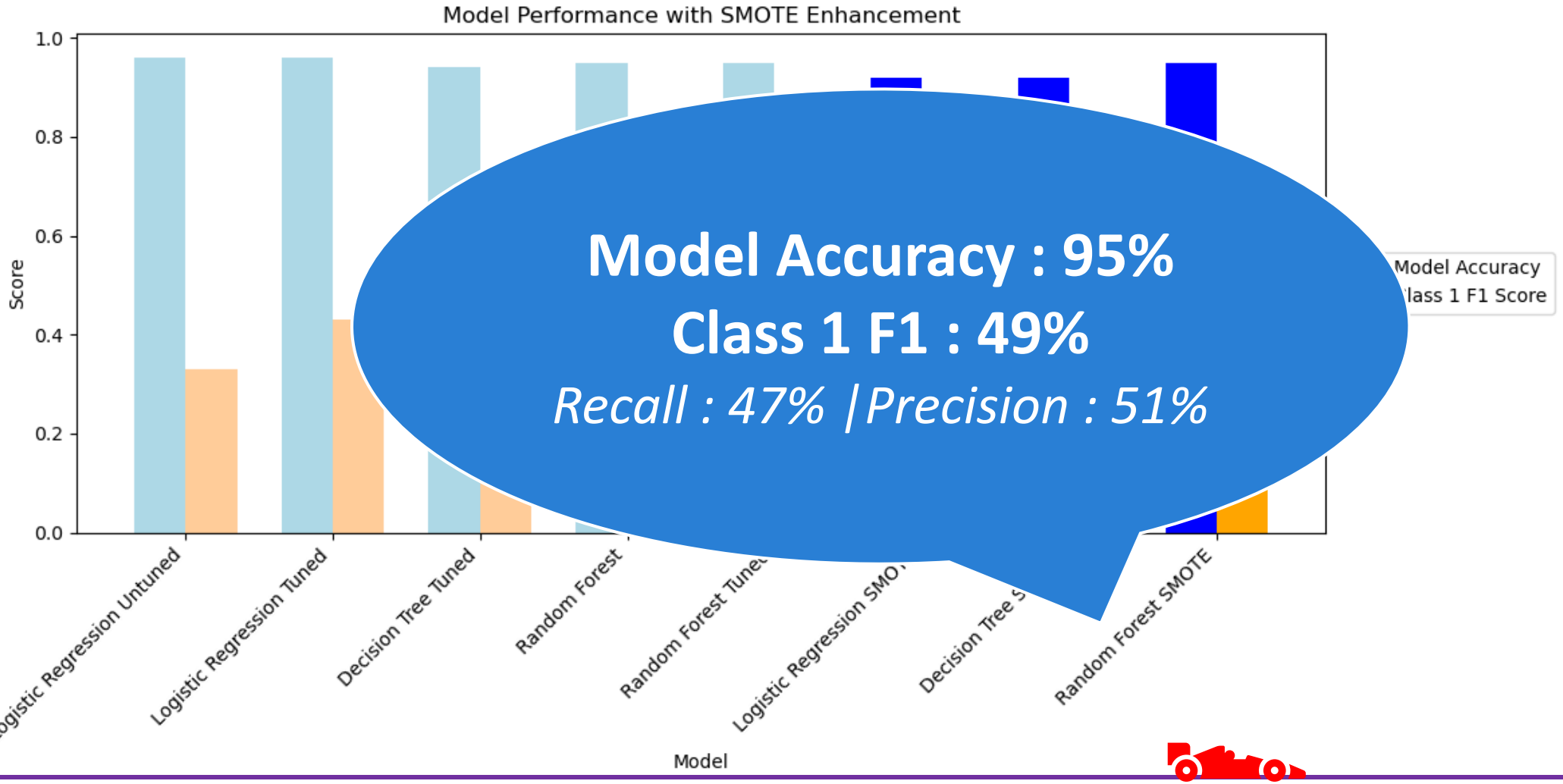

Benchmark : Predict each row will not win and you will be 95% accurate.

# Data Pit Stop: SMOTE's Role in Tuning Predictive Models

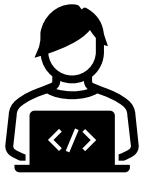


# SMOTE – Produced the best models so far…

Benchmark : Predict each row will not win and you will be 95% accurate.



# Next Steps for Sprint 3…



**Further Feature Engineering:** enhance predictive power

**Test Other Methods To Reduce Class Imbalance**

**Further Modelling Techniques:** e.g. Neural Networks

**Profit and Loss Analysis:** Make money?

**Model Interpretation** : What makes a winner?

# Thanks for your time