

A photograph of a basketball on a wooden court floor. The basketball is in the lower-left foreground, resting on a black circular object. The court floor is made of light-colored wooden planks with black boundary lines. The background is a dark, solid color.

# **An Exploration of Predicting The Results of NBA Games**

Jerel Novick

August 27, 2019

# Agenda

- Objective
- Data
- Methodology
- Models & Results
- Conclusions
- Further Development
- Questions and Answers





## **Objective**

**Build A Model That Can Predict The Outcomes  
Of NBA Games  
(And Maybe Make A Little Money)**

# Data

- Sourced From Kaggle
  - Basketball Reference
- Regular Season Games For The 2012 to 2017 Seasons
  - 14,760 Games
  - 7,380 “Match Ups”
- Perfectly Balanced



# Methodology

- Four Factor Model
  - Shooting (Field Goals)
  - Turnovers
  - Rebounding
  - Free Throws
- Personal Fouls
- Game Location
- Time Off

# Methodology

- Training Data – 2012 to 2015 – 9,840 Games
- Test Data – 2016 – 2017 – 4,920 Games

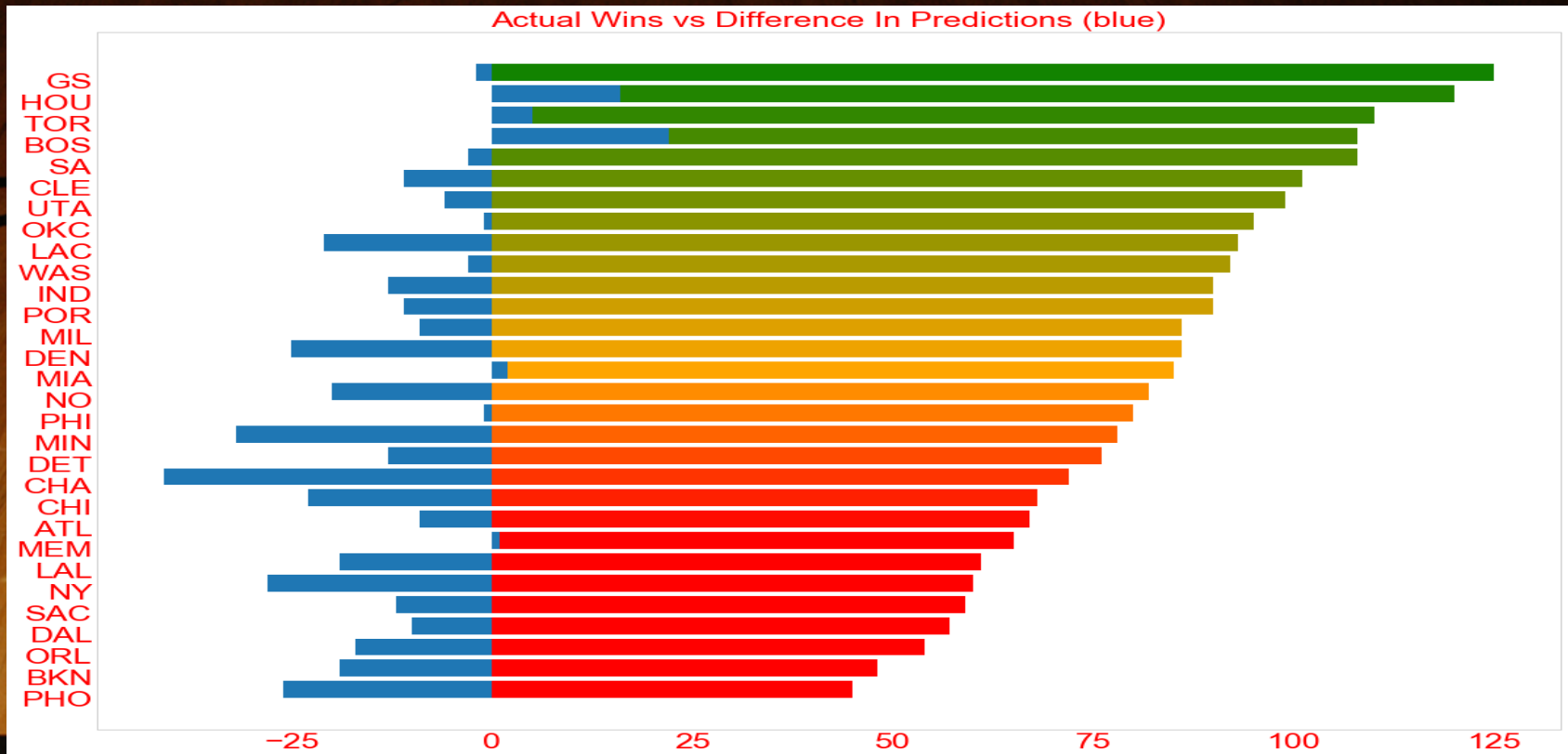


## First Model – Version One

- 80.77% Training Accuracy
- 79.13% Test Accuracy
- Coefficients

FGA	FGM	TO	ORB	DRB	FTA	FTM	Location	Days Off	PF
-0.25	0.41	-0.20	0.21	0.26	-0.09	0.18	0.44	-0.17	-0.08

# First Model – Version One





## **First Model – Version Two**

- Effective Field Goal Percentage
- Turnover Percentage
- Total Rebound Percentage
- Free Throw Percentage
- Personal Fouls
- Game Location

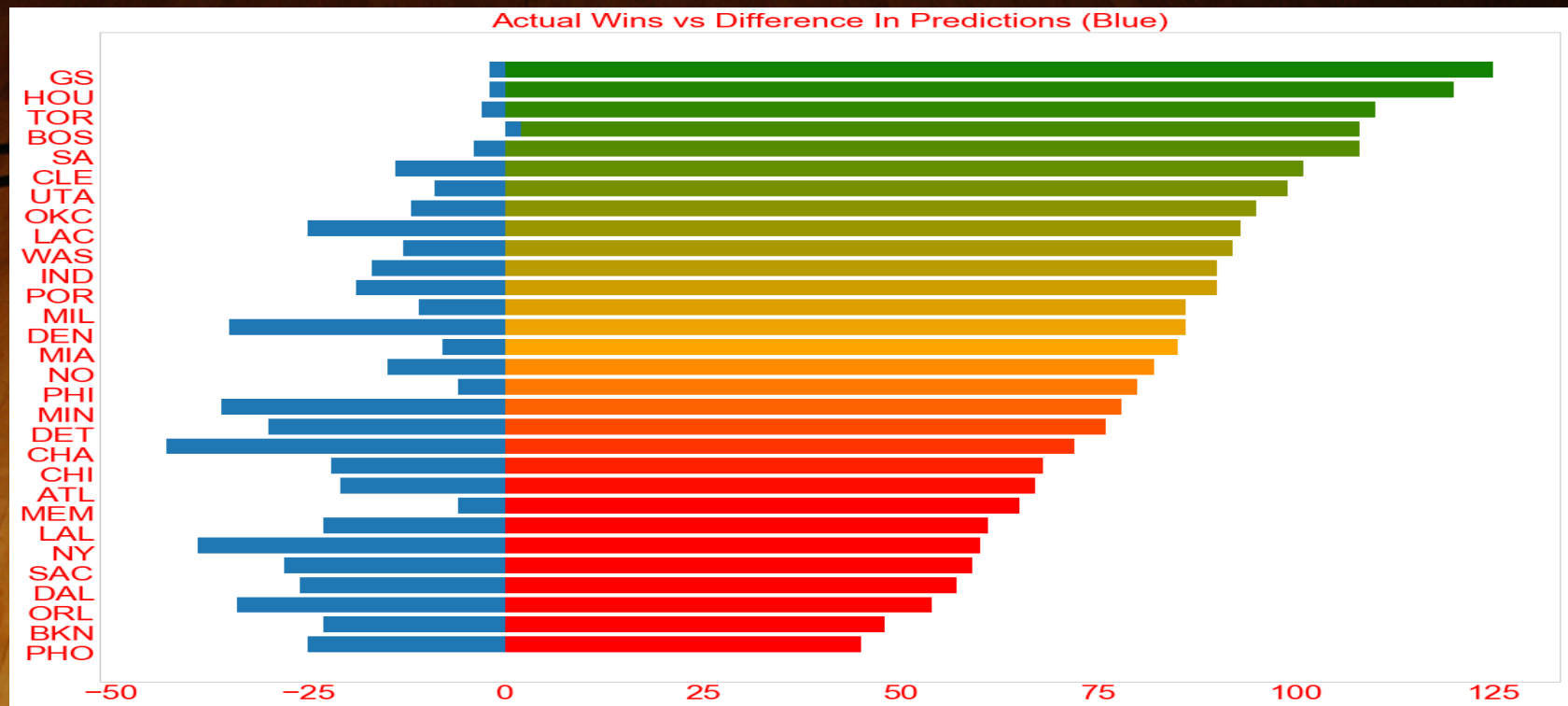
## First Model – Version Two

- 76.84% Training Accuracy
- 75.26% Test Accuracy
- Coefficients

EFG%	T0%	TREB%	FT%	Location	Days Off	PF
0.20	-0.18	0.16	0.02	0.50	0.004	-0.08



# First Model – Version Two



## **Second Model – Version One**

- Five Period Moving Averages
- Exclude First Month of the Season
- 10,800 Games

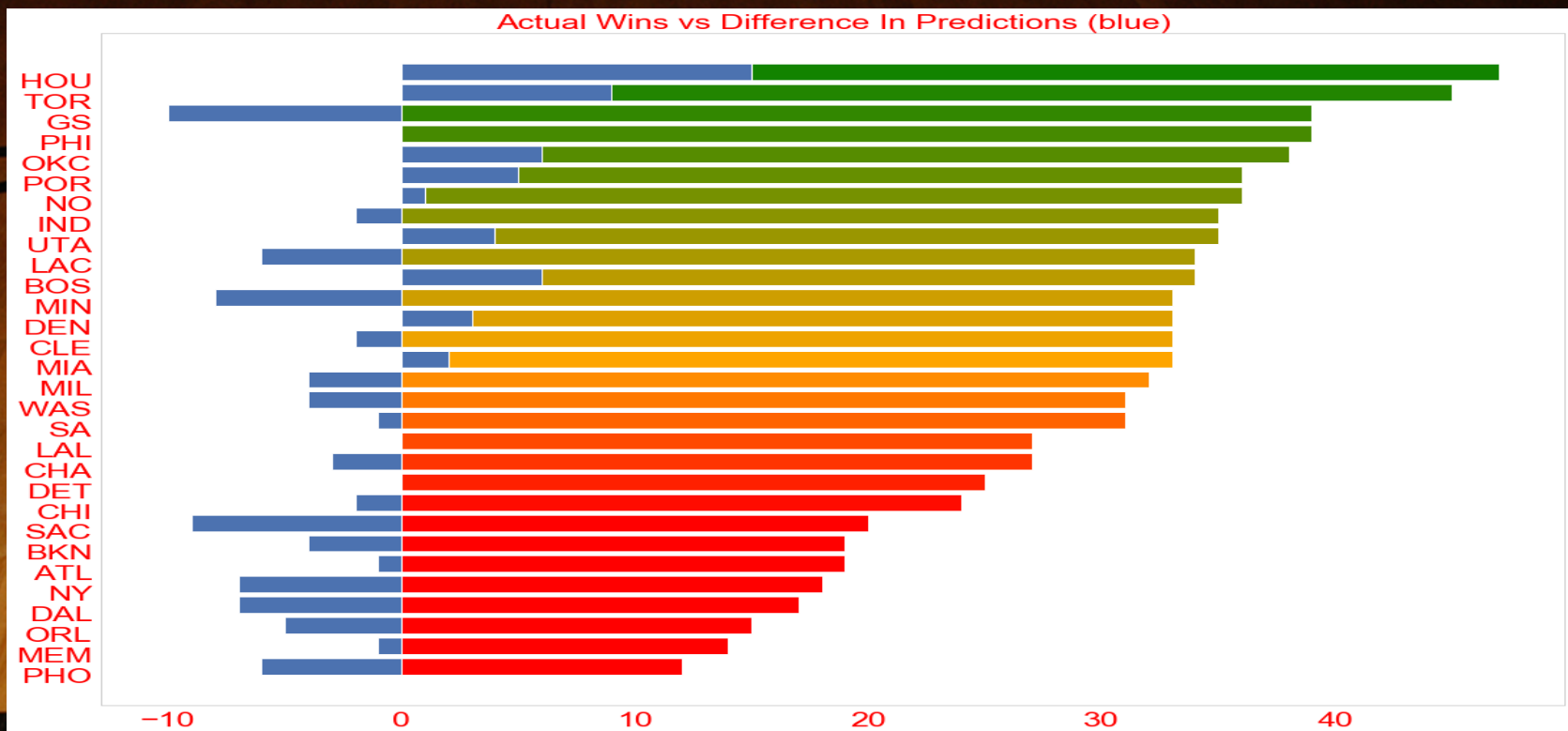


## Second Model – Version One

- 60.35% Training Accuracy
- 58.46% Test Accuracy
- Coefficients

FGA	FGM	TO	ORB	DRB	FTA	FTM	Location	Days Off	PF
-0.08	0.12	-0.05	0.06	0.04	-0.02	0.03	0.74	0.03	-0.004

# Second Model – Version One



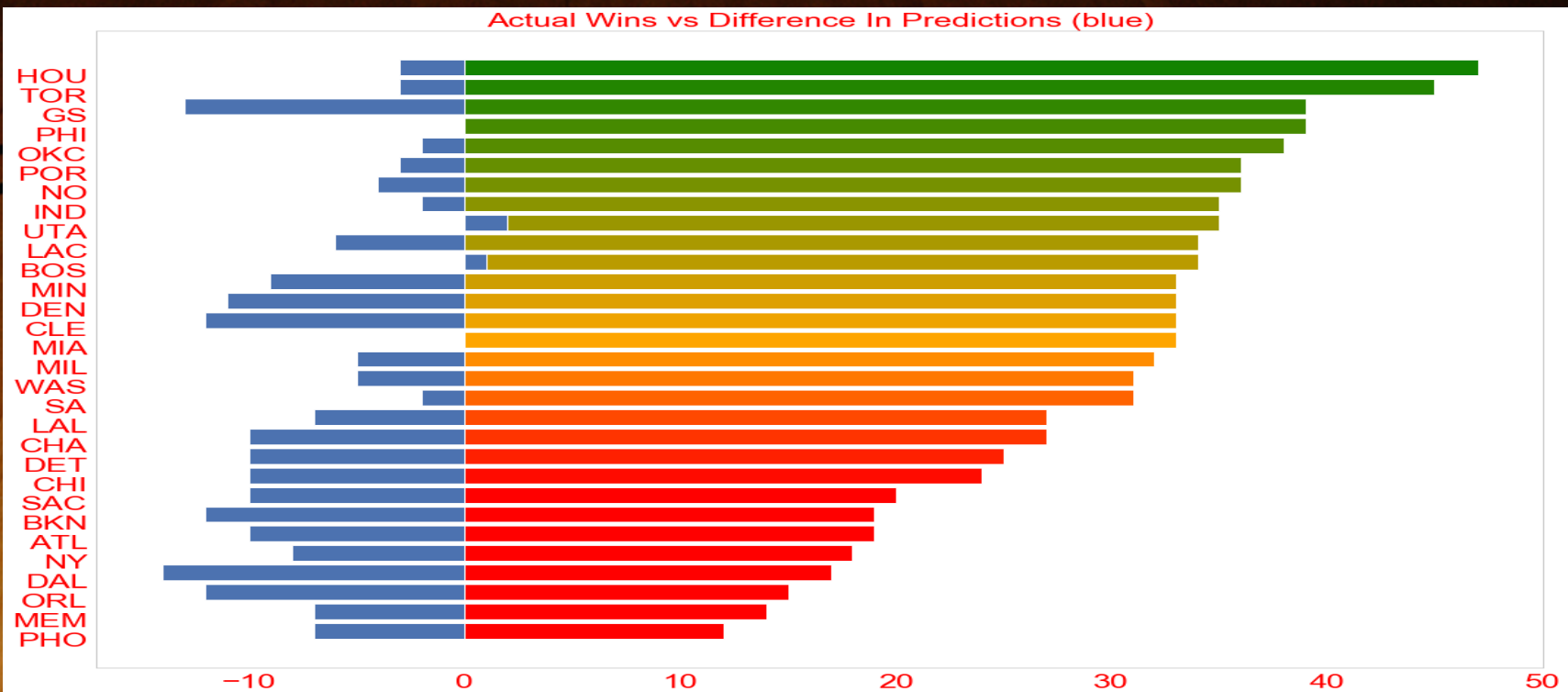


## Second Model – Version Two

- 59.85% Training Accuracy
- 57.89% Test Accuracy
- Coefficients

EFG%	T0%	TREB%	FT%	Location	Days Off	PF
0.07	0.08	0.03	0.002	0.73	0.04	-0.02

## Second Model – Version Two





# Conclusion

- First Model
  - Overall Accurate But Loses Predictive Power On Weaker Teams
  - Backward Looking
- Second Model
  - Moving Averages
  - Smoothing And Trends Less Predictive
- Predicting Favorites?

## **Further Work**

- Create And Employ Better Metrics
- Look For Interactions
- Calculate Win Probabilities
- Favorites vs. Dogs?
- NCAA
- Pythagorean Expectation And Log 5 Models
- Look For Value Against Vegas Moneylines
- Phone Interview Tomorrow



A photograph of a wooden basketball court floor. A thick black curved line, likely a three-point arc, is visible. The word "Questions?" is written in white text in the center of the image.

Questions?