



Machine Learning Vs. MLB



**“Never allow the fear of striking out
keep you from playing the game” -
Babe Ruth**

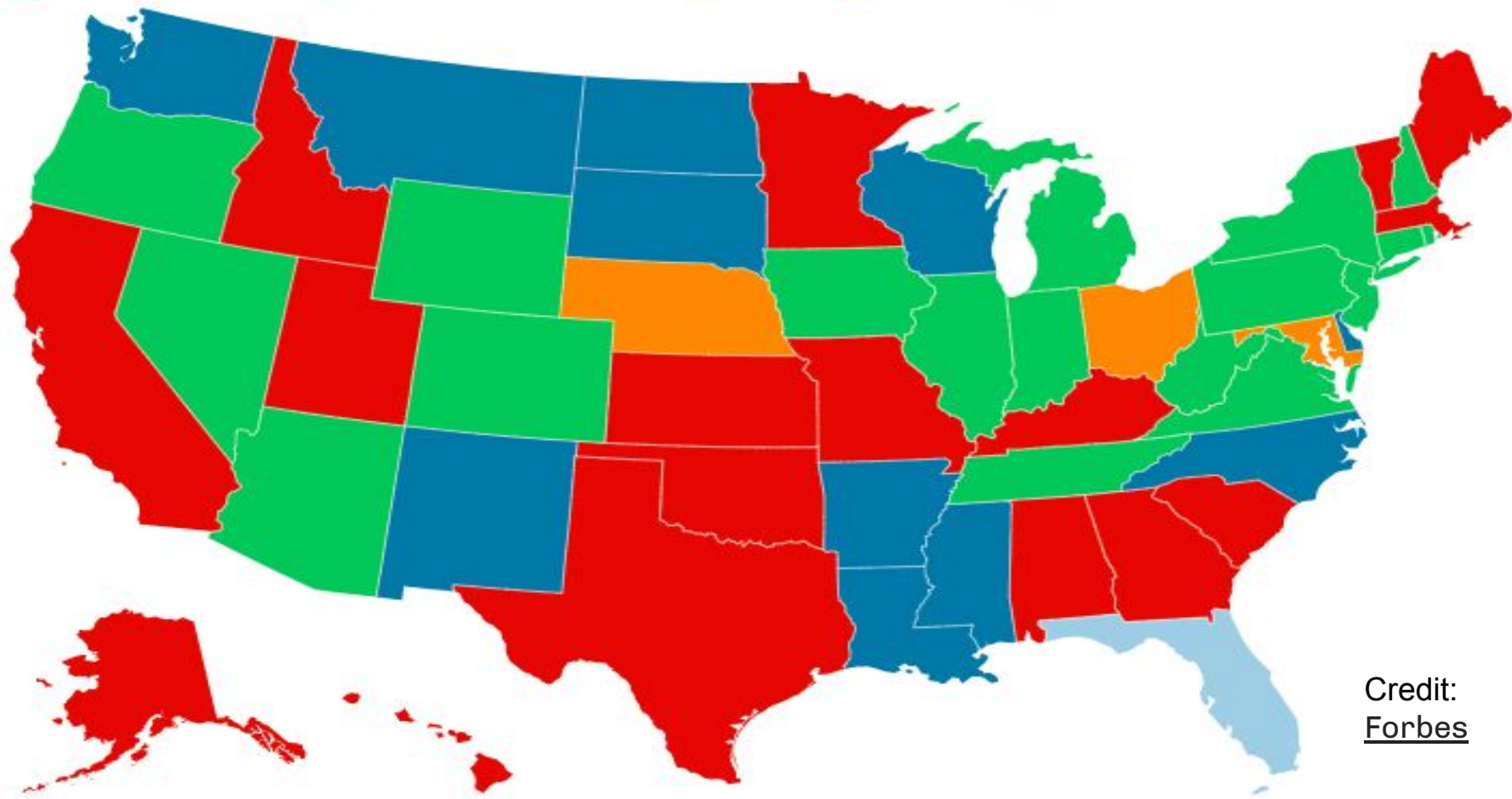


Process

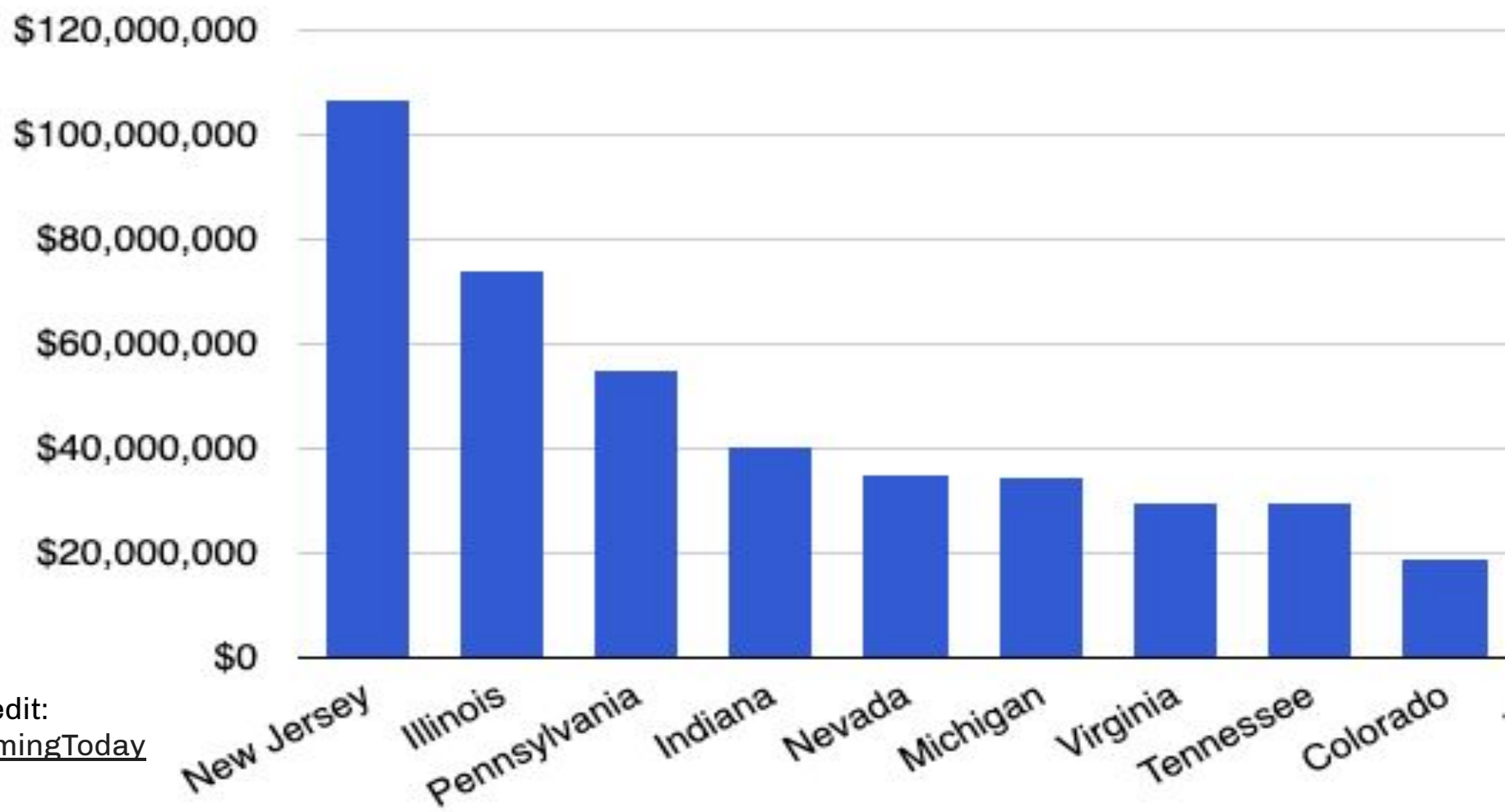


Legal Sports Betting in the U.S.

Legal online wagering Legal physical sportsbooks Pending Not legal In limbo



Credit:
Forbes



Credit:
[GamingToday](#)

Business Problem



"While I love betting on America's pastime, it's known to be the most demanding sport to win money. Baseball, unlike other major sports, is by far the most unpredictable sport to gamble on." - **GamblingSites**





“Of the four major professional sports, MLB home teams have the lowest win percentage — regular season or postseason.” - Action Network

Data Gathering and Baseline

- Sportsipy
- Baseball Reference
- 52-54% accuracy



Models

1. Logistic Regression

2. KNN

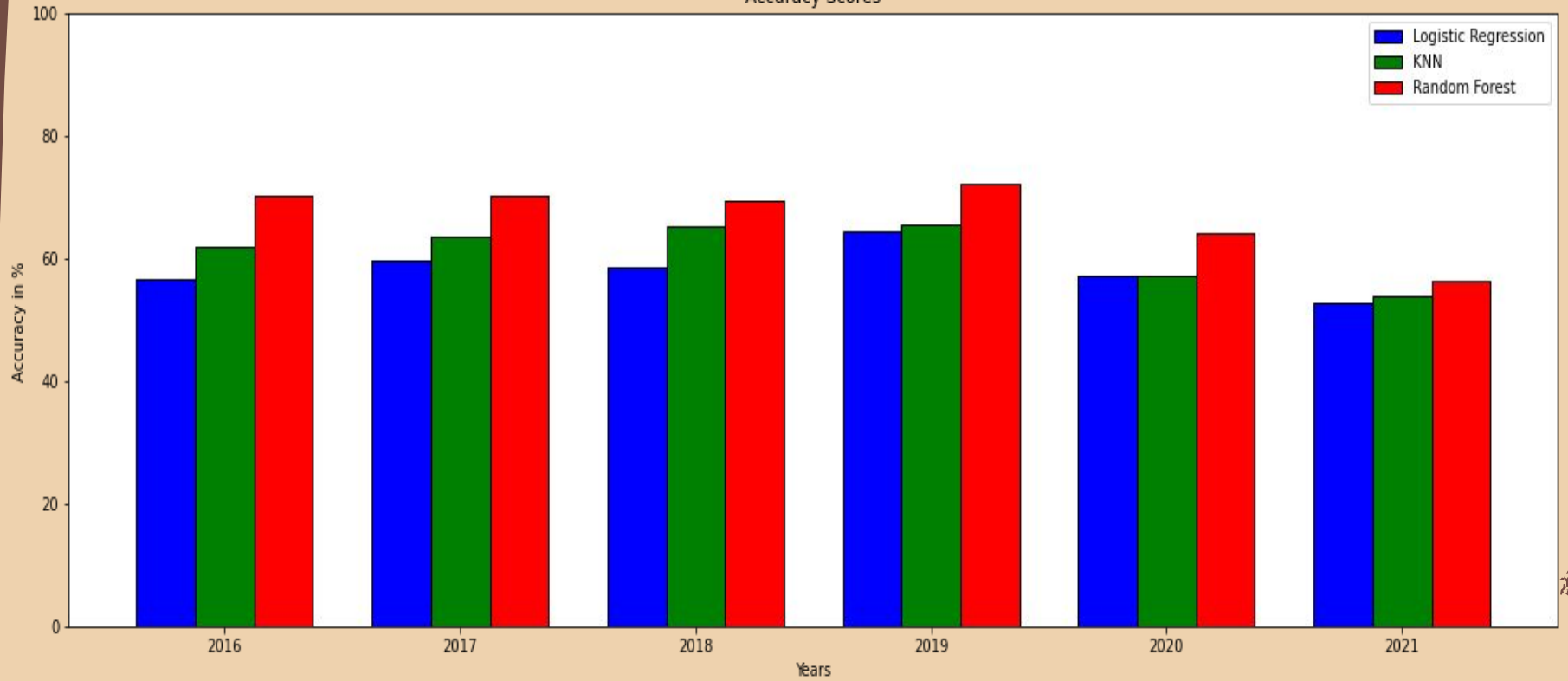
3. Decision Trees

4. Random Forest

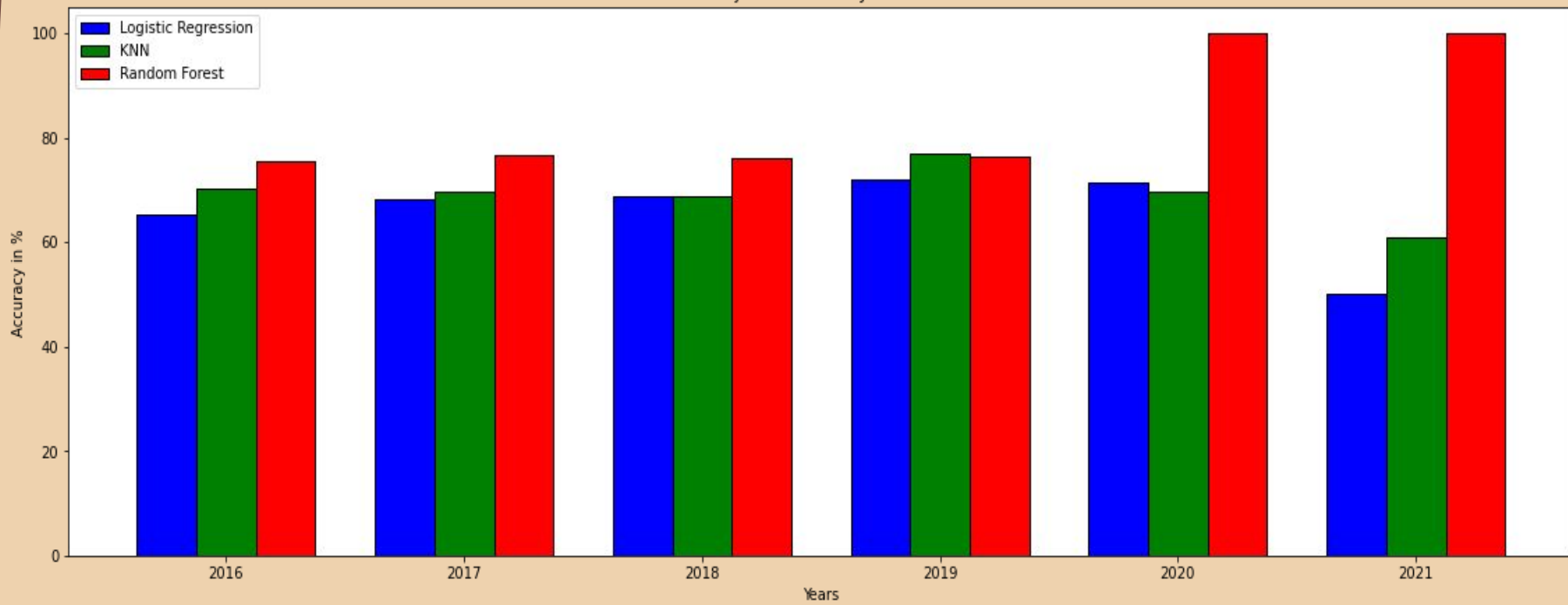
5. Bagging Classifier

6. XGBoost

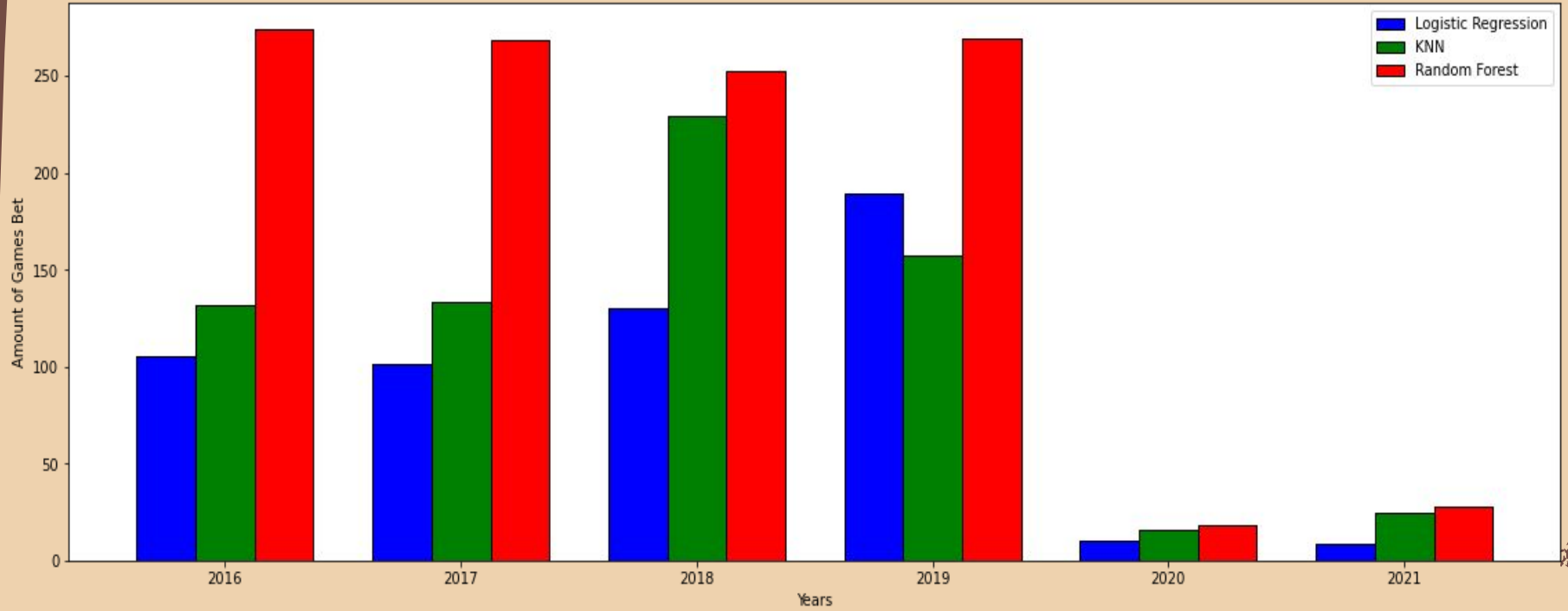
Accuracy Scores



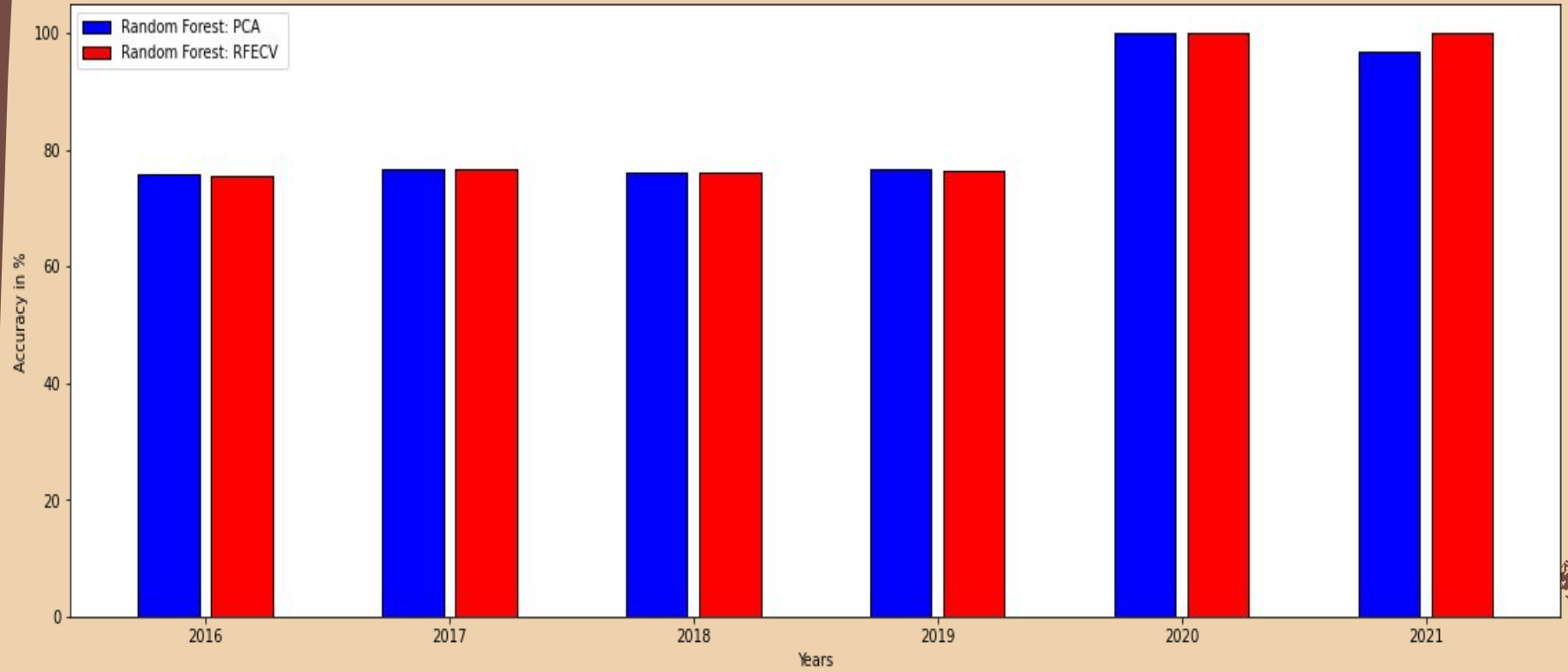
Adjusted Accuracy Scores



Games Bet on by Year



Random Forest: RFECV vs. PCA



Next Steps

01

*Moneyline
Data*

*Individual
Stats
Implementation*

02

03

Bet Sizes

Thank You!

Do you have any questions?



Matthew Reinhart

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon** and infographics & images by **Freepik**

