NBA PLAYER STATISTICAL ANALYSIS AND PREDICTION PROJECT

COLLABORATORS

JACOB EVANS
KARAN ANAND
MERT OZTOP
PRATIK PUROHIT
PRIYA MARINGANTI

UNIVERSITY OF TORONTO

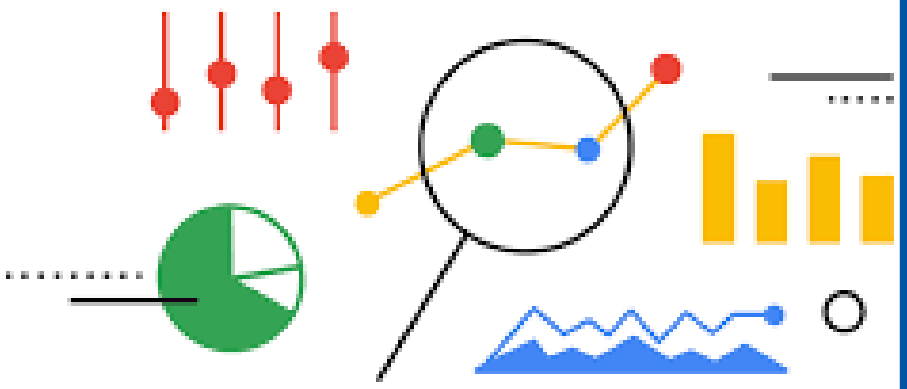
OVERVIEW

Our goal in this project was to attempt to predict the points for basketball player for the upcoming season using machine learning techniques.

METHODOLOGY

- Data Collection
- Data Preprocessing
- Exploratory Data Analysis
- Model Training and Evaluation

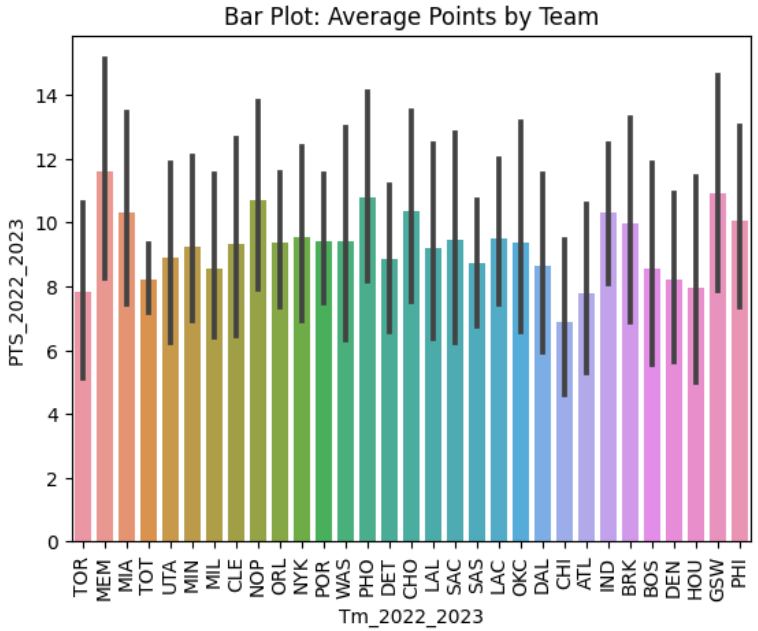
DATA SOURCES

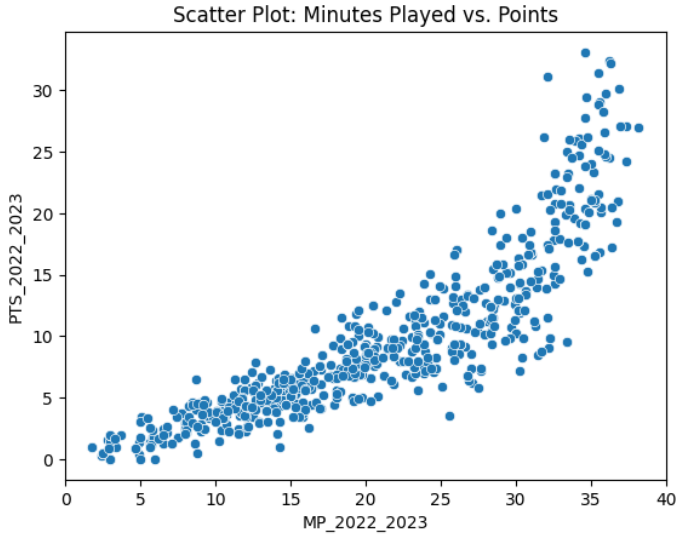


NBA Stats : Kaggle (2021/22 & 2022/23)

Athlete Head Shot: WebScraping (*NBA Website* and *Loodibees Logos*)

Exploratory Data Analysis





- Barplot of Average Points by Team for 2022/2023
- ScatterPoints for Minutes Played vs. Points

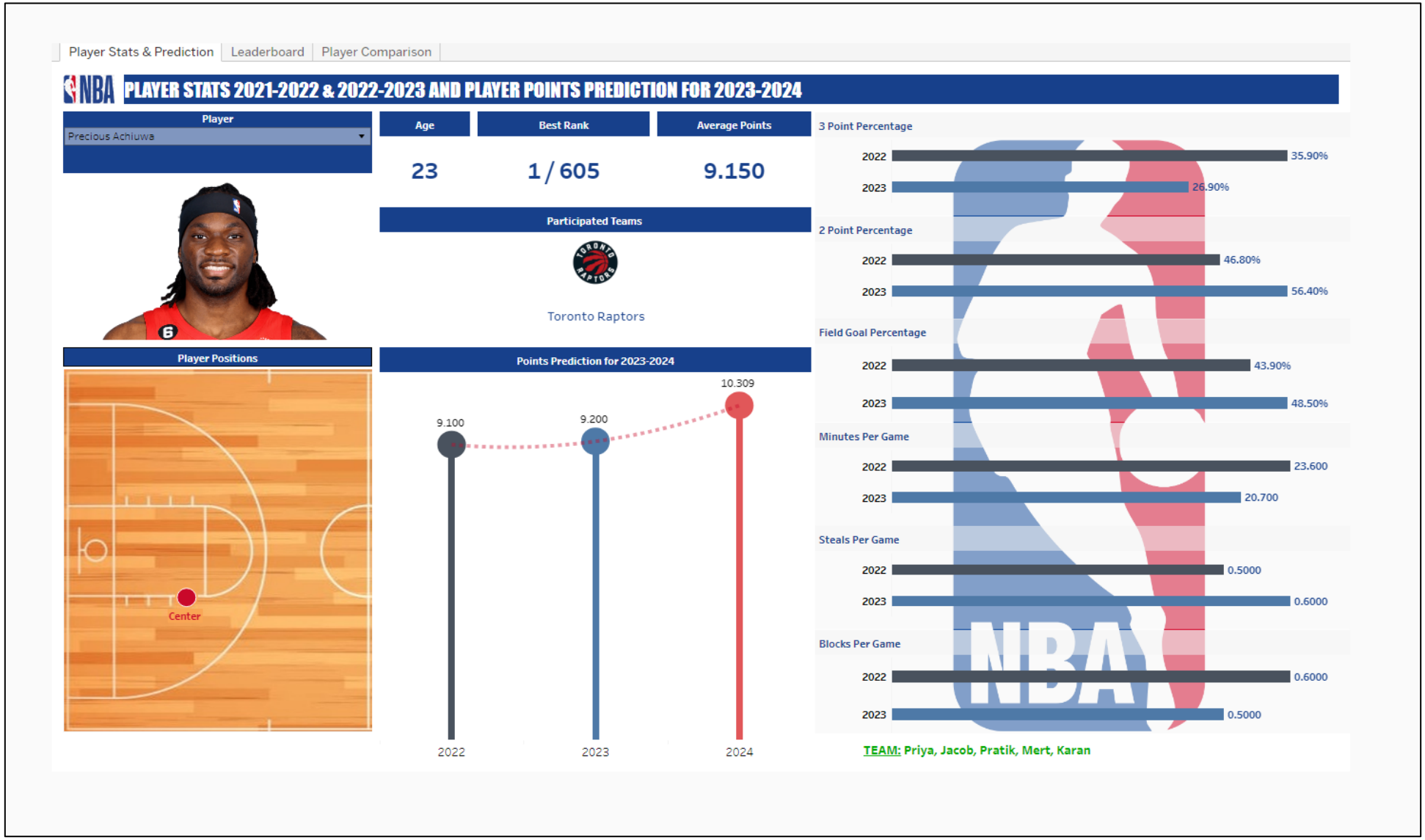
Model Training & Evaluation

Model	R-Squared	Mean Square Error
Linear Regression	0.9988	0.077
Random Forest Regression	0.9987	0.548
Random Decision Tree	1.0	1.041
Lasso Model Regression	0.9982	0.256

- Experimented with various regression: linear regression, decision tree regression, random forest regression, and lasso model.
- The Linear regression model was our best choice because it showed the least MSE.

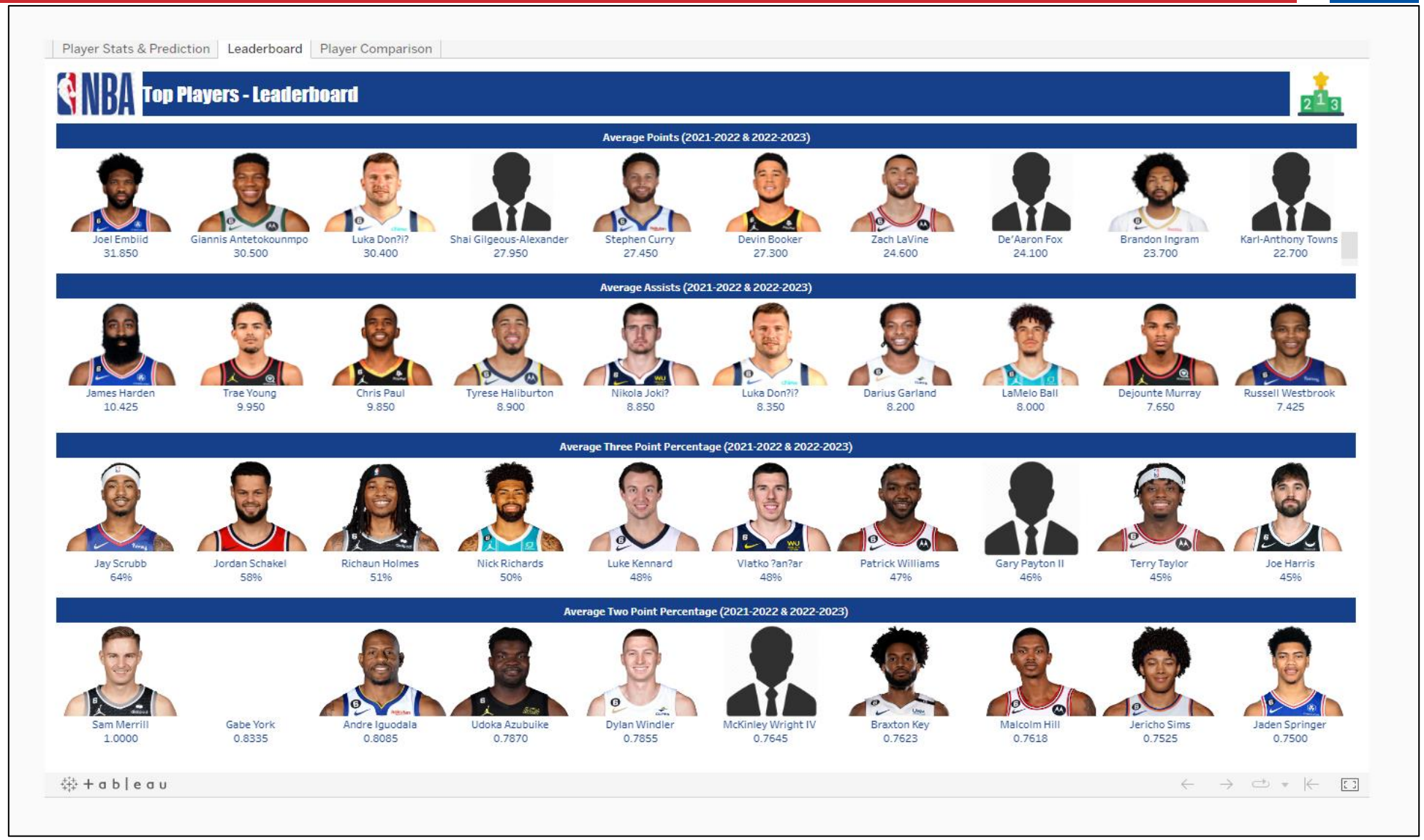
User Interface

Player Stats and Prediction



User Interface

Player Stats and Prediction



User Interface

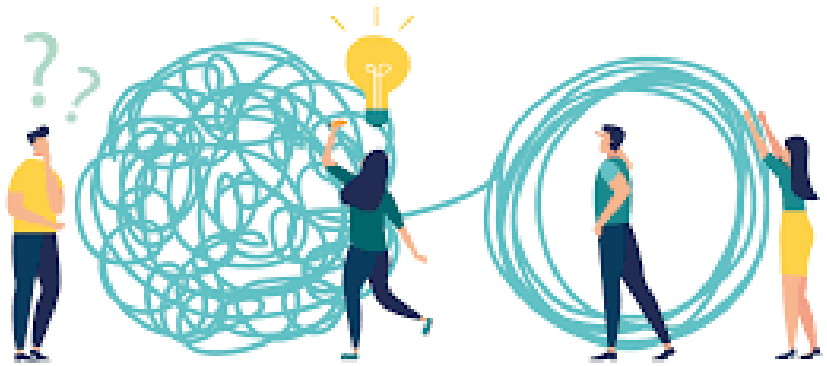
Player Comparison

RESULT

The prediction system achieved R-squared value of 0.9998, indicating a high level of accuracy in predicting player statistics based on the historical data.

- The system can assist with team selection, player scouting, and forecasting player statistics for the upcoming season.

CHALLENGES



- Bad Encoding
- External factors
- Outliers

THANK YOU

QUESTIONS ARE WELCOME