NBA Fantasy Analysis for Dynasty Points Leagues

By: Kofi Darfour

Understanding Fantasy Dynasty

- Long-term Commitment: Unlike traditional fantasy basketball, a dynasty league involves managing a team over multiple years. You'll keep most or all of your players from year to year, focusing on long-term strategy.
- Player Selection: Initially, you'll draft a team of NBA players. This roster is typically your foundation for several seasons, making the initial draft critically important.
- Scoring System: Points are awarded based on players' real-life performances in various statistical categories like points, rebounds, assists, steals, etc. Different leagues may have different scoring systems, so it's important to understand your league's specific rules.
- Roster Management: You'll manage your roster by starting players, benching others, and making trades. Since it's a long-term game, you'll also need to consider player development and aging.
- Prospects and Rookies: In dynasty leagues, young prospects and rookies are more valuable than in standard leagues. You'll need to balance the immediate contribution of veterans with the potential of younger players.
- Annual Drafts: After the initial draft, there will be yearly drafts to select rookies or unowned players. This is a chance to improve your team and adjust your strategy.
- Starter Level Player: A player who averages 35+ fantasy points
- Winning: Your goal is to accumulate the most points over the season. Strategy involves not just who is performing well now, but who will be valuable in future seasons.

Introduction

Research Questions:

- Which Position is the best to target in a draft?
- Which Position is worst?
- What's a good draft strategy based off the two questions above?

Rationale: Due to the nature of Dynasty Fantasy Basketball, having the best value players is key because you keep them until you trade them

Data Manipulation and Analysis

- Library Import: The script imports tidyverse and ggplot2, which are used for data manipulation and visualization.
- Data Loading: It loads NBA player statistics from the 2021-2022 season from a CSV file named "NBA_2021_2022_stats.csv".
- Renaming Columns: For better readability, the column names are renamed to more descriptive titles (e.g., 'Tm' to 'Team', 'G' to 'Games Played').
- Calculating Fantasy Points: The script introduces a new column, Average Fantasy Points, calculated using a custom formula based on various player statistics like rebounds, points, steals, blocks, and assists.
- Total Fantasy Points: Another new column, Total Fantasy Points, is computed as the product of 'Average Fantasy Points' and 'Games Played'.
- Splitting Position Data: The 'Position' column is split into two new columns, 'Position 1' and 'Position 2', to accommodate players who play in multiple positions.
- Column Selection: The original 'Position' column is removed after the split.
- Player Categorization: The script includes code for identifying players by age groups and roles, such as starter-level, young, and prime players.
- Data Visualization: Several ggplot2 visualizations are created, including scatter plots, box plots, density plots, and bar charts to analyze various aspects like fantasy points distribution by position and age.

Key Variables

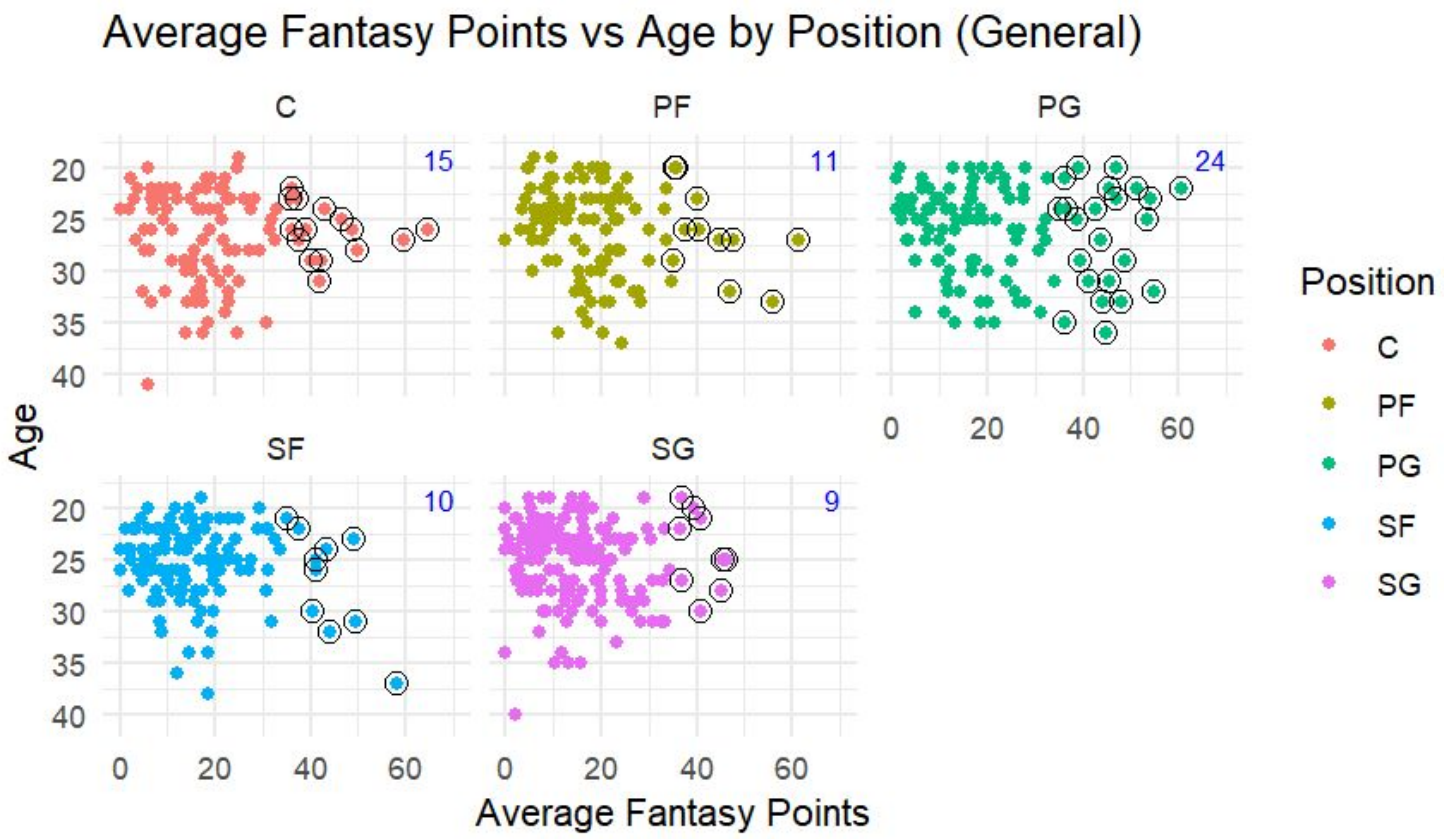
Key Variables Glossary:

- Games Played (G)
- Offensive (ORB) and Defensive Rebounds (DRB) Per Game
- Assists (AST) Per Game
- 3 Point (3P%) and 2 Point (2P%) Percentages
- Free Throw (FT%) and Field Goal (FG%) Percentages
- Steals (STL) Per Game
- Blocks (BLK) Per Game
- Points (PTS) Per Game
- Player Positions (Pos)
- Average Fantasy Points
- Total Fantasy Points

Variable Descriptions:

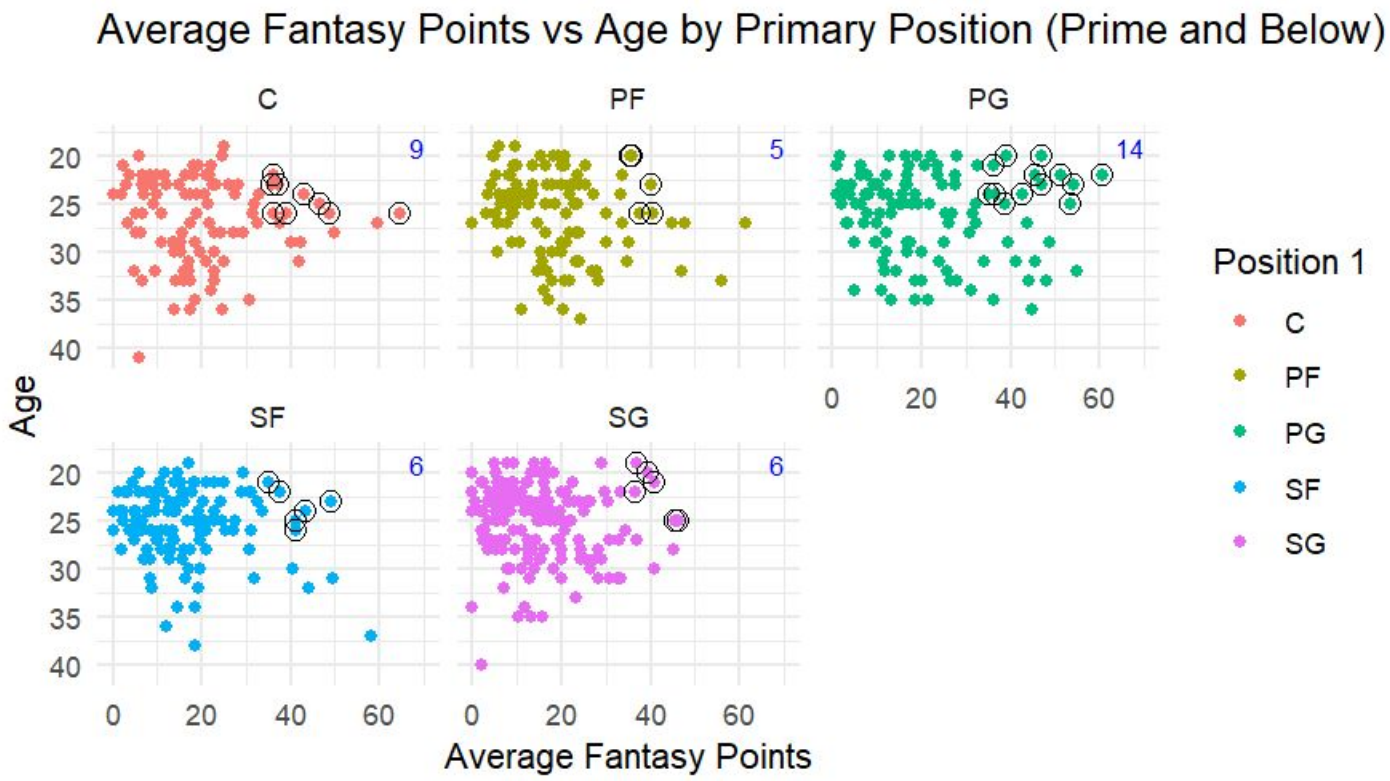
- Team (Team): The NBA team a player is part of.
- Games Played: Number of games a player participated in.
- Offensive and Defensive Rebounds Per Game: Average rebounds per game, both offensive and defensive.
- Assists Per Game: Average assists made per game.
- Percentage Metrics (3P%, 2P%, FT%, FG%): Player's shooting efficiency metrics.
- Steals and Blocks Per Game: Average number of steals and blocks made per game.
- Points Per Game: Average points scored per game.
- Player Positions: Playing positions of the players, split into primary and secondary positions.
- Average Fantasy Points: A calculated metric based on player performance across different categories.
- Total Fantasy Points: Total fantasy points accumulated over the season.

Graph: Scatterplot (General)



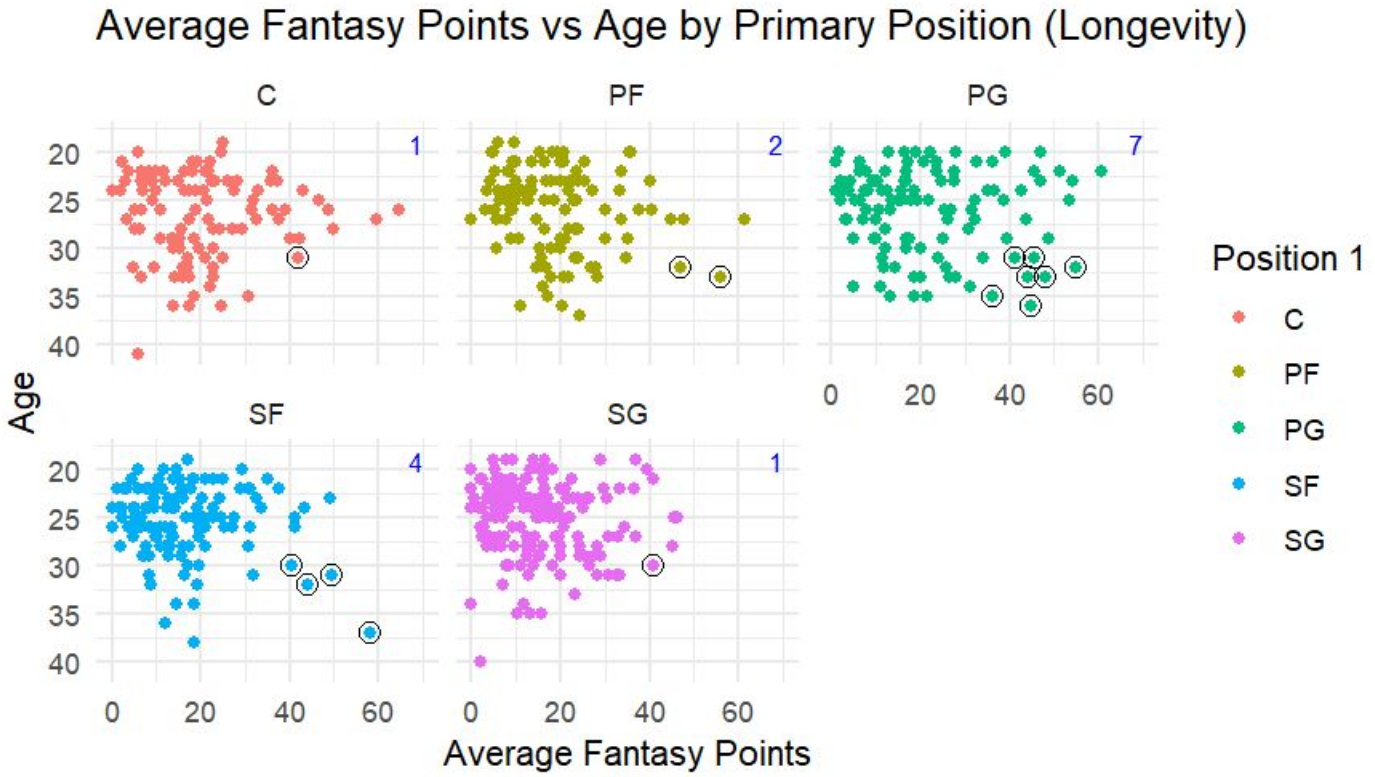
Highlighted All Players who are Starter Level

Graph: Scatterplot (Prime and Below)



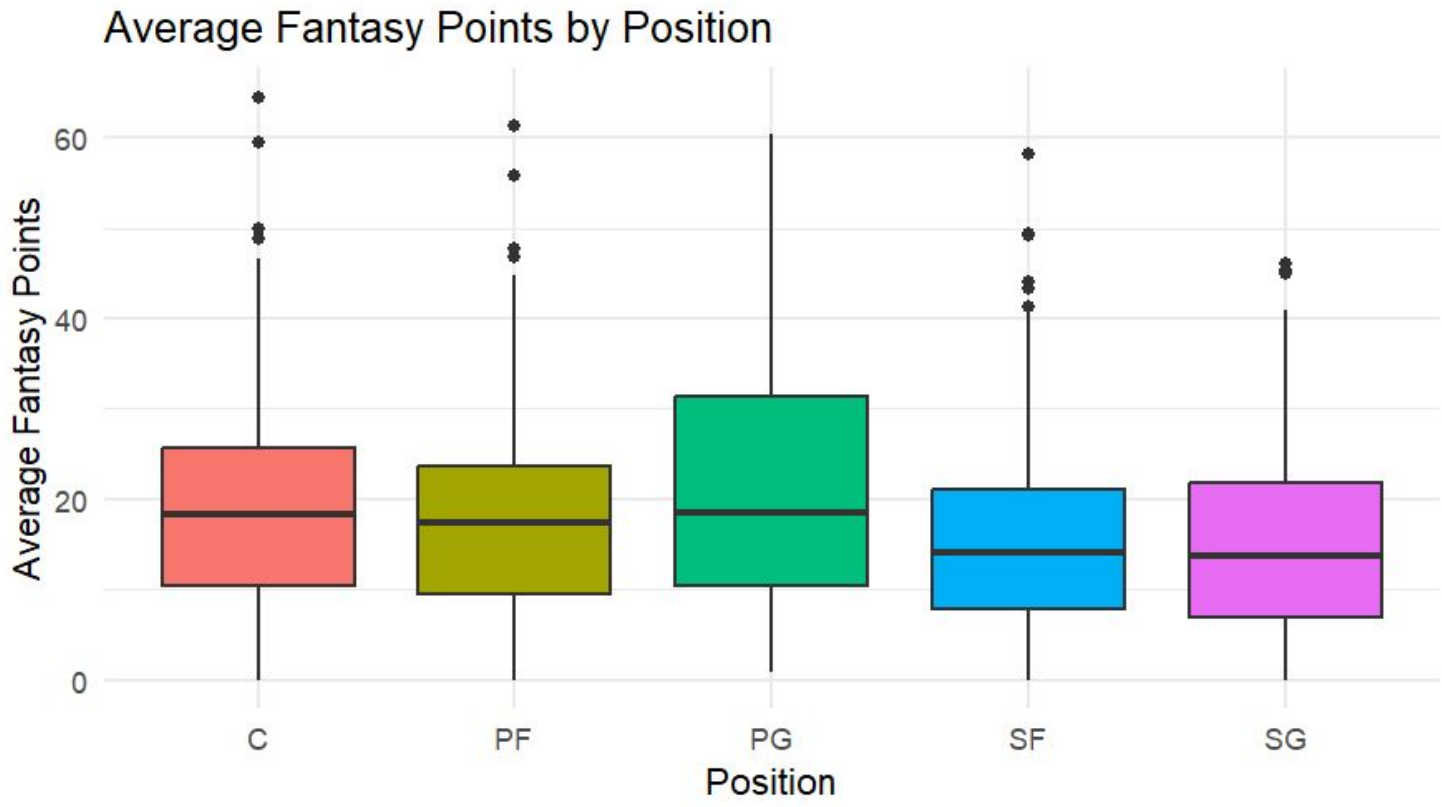
Highlighted Players 26 y/o and younger who are Starter Level

Graph: Scatterplot (Longevity)



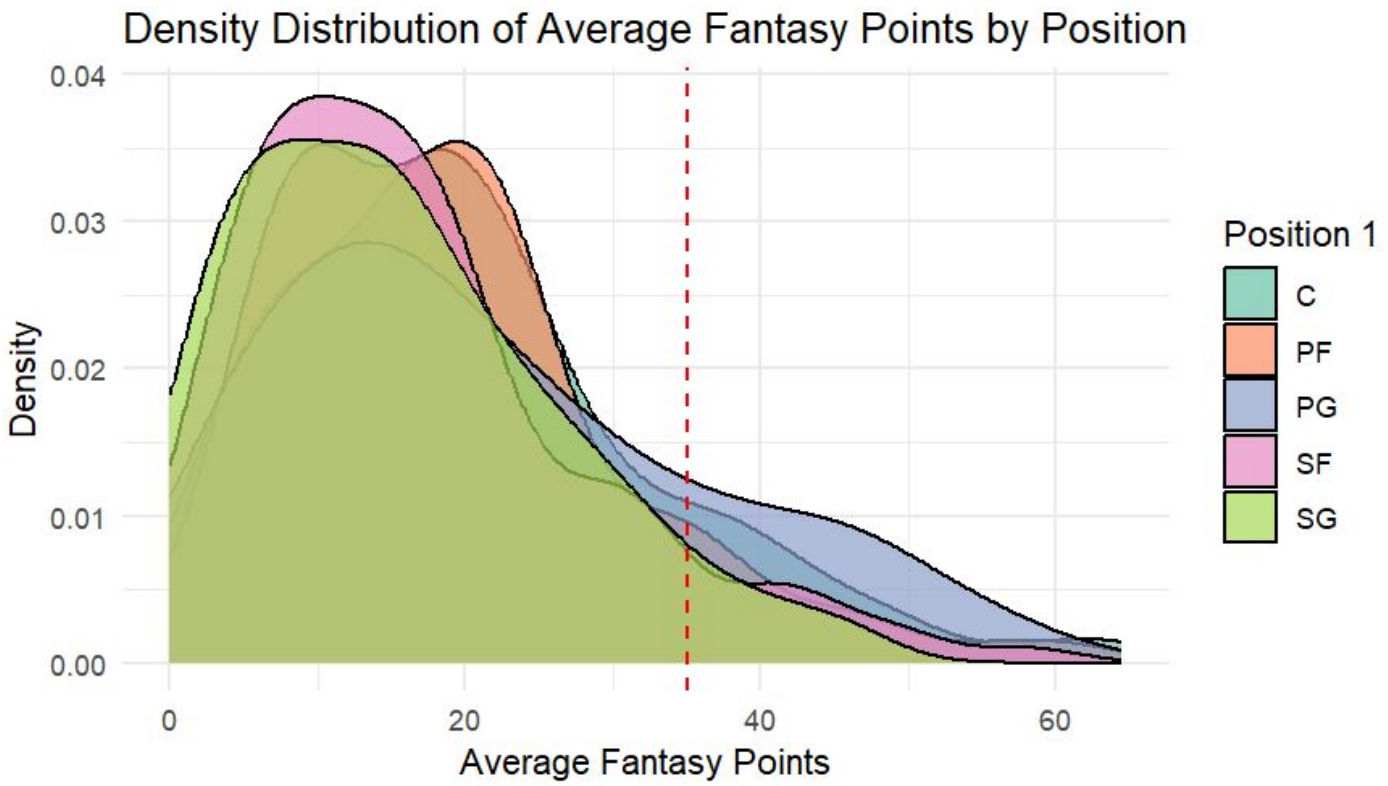
Highlighted Players 30 y/o and up who are Starter Level

Graph: Box Plot



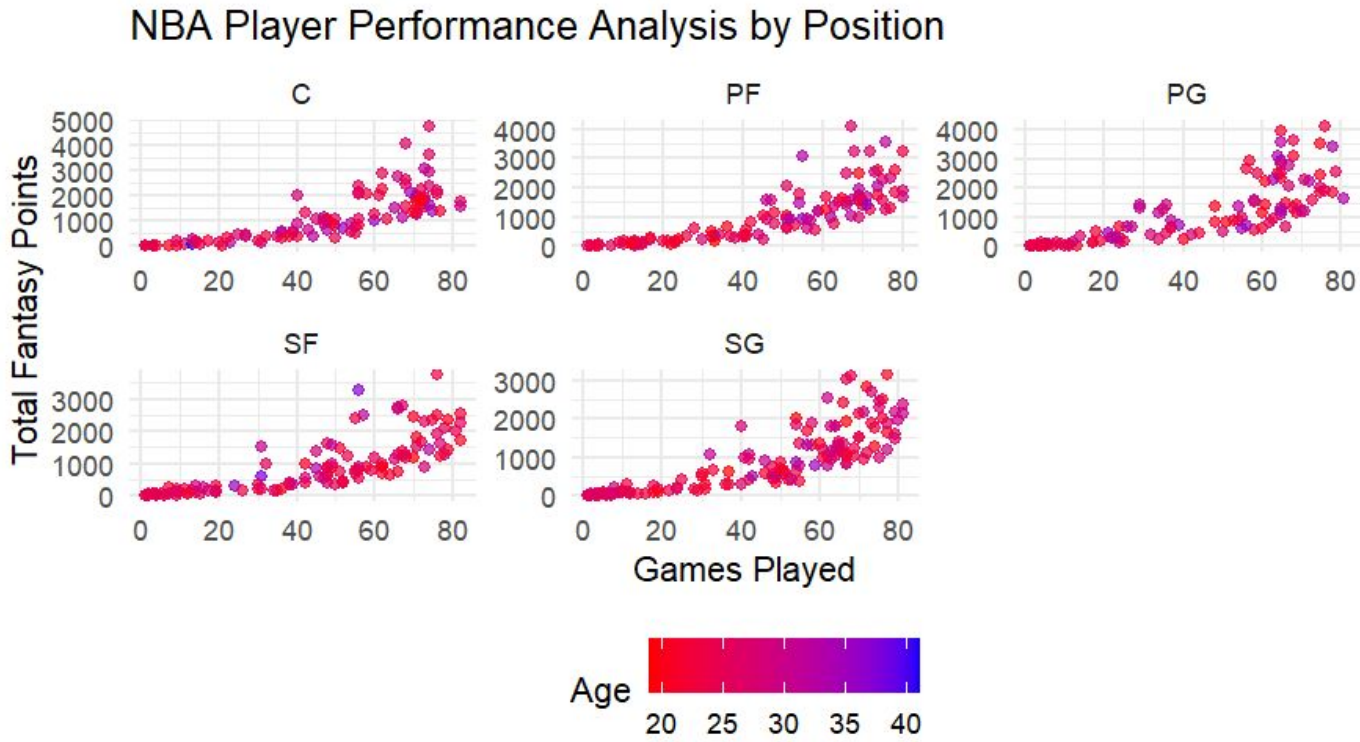
Shows that PGs are consistently high performing scorers compared to other positions.

Graph: Density



Red line and on indicates Starter Level Player, and here you can see PGs have the highest density in that case

Graph: Faceted Scatterplot



Demonstrates how of all positions PGs play the most and score the most within a season

Methods/ Analysis for Statistical Inference

- Test Chosen: ANOVA (Analysis of Variance)
- Appropriateness: ANOVA is appropriate for comparing means across three or more groups. In your case, it seems you have conducted an ANOVA to compare the means of a variable (possibly 'Total Fantasy Points') across different positions ('Position 1'). ANOVA helps determine whether there are statistically significant differences in means between the groups (positions).
- Hypotheses:
  - Null Hypothesis: There is no significant difference in the means of 'Total Fantasy Points' between different positions.
  - Alternative Hypothesis: There is a significant difference in the means of 'Total Fantasy Points' between at least two positions.
- Analysis Conclusion: Based on the ANOVA results, it appears that there is a statistically significant difference in the means of 'Total Fantasy Points' across different positions. The p-value (Pr(>F)) is very low (0.000175), indicating that there is strong evidence to reject the null hypothesis. The 'Position 1' variable is significant in explaining the variation in 'Total Fantasy Points'.

ANOVA Table Results:

```
             Df Sum Sq Mean Sq F value    Pr(>F)
`Position 1`  4   3401   850.4    5.67 0.000175 ***
Residuals    600  89988   150.0
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Conclusions

Summary of Findings: Based off the graphs, it looks like PGs are the best position to target and its between PF and SG as the worst. A good strategy would be to either target higher level players of weaker positions or stash high level PGs and trade for positions you need

General Conjectures: With a machine learning algorithm, you could make a model that tells you what players to target in a draft based off remaining available players and current roster using this analysis, but with more data.

Limitations and Future Research: Was not provided with more stats needed to calculate fantasy scoring such as turnovers, field goals made/missed per game, and fouls. This would have made fantasy point averages more accurate for analysis.