Data Appendix

We used two datasets from the data.world NBA Salaries data (https://data.world/datadavis/nba-salaries/workspace/intro). These two datasets are named players and salaries_1982-2018.

players.csv

The dataset RAW_interactions contains 5 columns. The columns are called user_id, recipe_id, date, # rating, and review. Each row/entry represents a review and rating from a user on a Food.com recipe.

salaries_1985-2018

The dataset RAW_recipes contains 12 columns. The columns are called name, id, # minutes, contributor_id, submitted, tags, nutrition, n_steps, steps, description, ingredients, and n_ingredients. Each row/entry represents a Food.com recipe.

Our Final Dataset

After joining based on the player ID, our final dataset that we used for our analysis contains 4 columns. The columns are called season, salary, position, and name

Variables

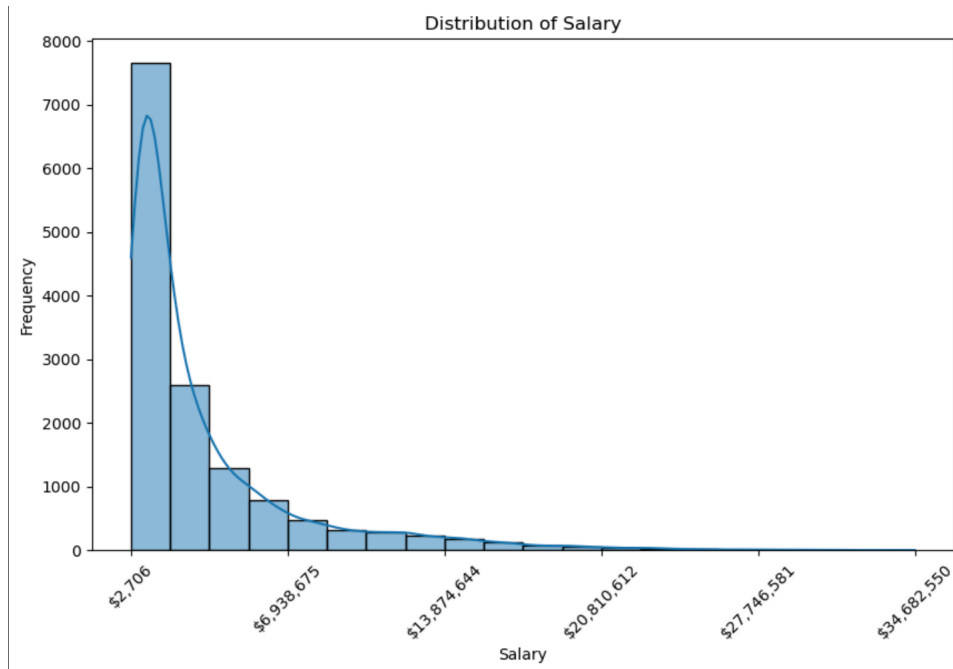*Season -* The NBA season the data pertains to.
Type: String (e.g., "2019-2020")
Notes: Indicates the specific season for which the salary and performance metrics are recorded.

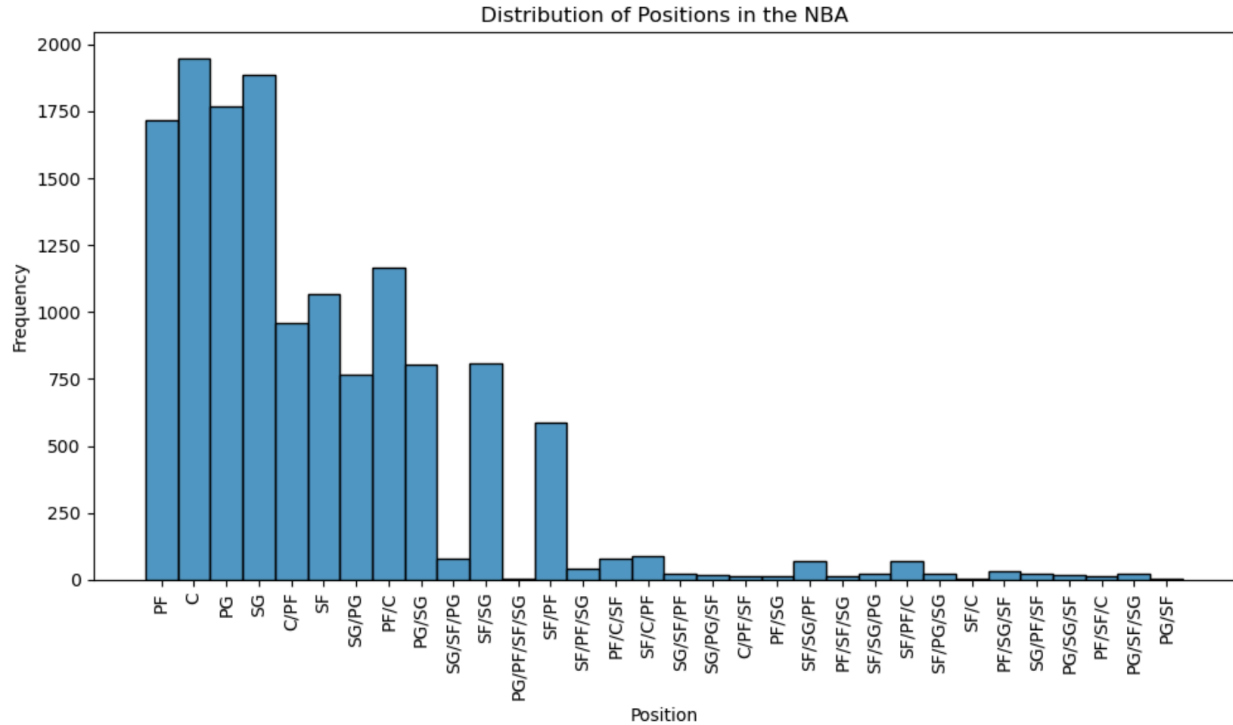*Salary -* The player's salary for the specified season.
Type: Numeric (USD)
Notes: Represents the total salary the player received for the season, excluding endorsements and other income sources.

Distribution of Salary

Position-the position the player primarily plays.

Type: String (e.g., "PG" for Point Guard)

Notes: Reflects the player's main role on the team. Common positions include PG (Point Guard), SG (Shooting Guard), SF (Small Forward), PF (Power Forward), and C (Center).

Distribution of Positions in the NBA

Name-The name of the NBA player.

Type: String

Notes: Full name of the player as recorded in official NBA rosters.