Universitat Politécnica de Catalunya

Multivariate Analysis

Final Project

Group members:

William Alejandro Cueva

Nashly Erielis González

Group 11 and 12

Interfaz de usuario gráfica, Texto

Descripción generada automáticamenteTexto

Descripción generada automáticamenteDecember 23, 2024

1. Introduction

The present work aims to perform a multivariate analysis of the performance of outstanding NBA players. This analysis explores general patterns, identifies outliers, and segments players into groups based on their performance and physical characteristics. Additionally, the association between different metrics is evaluated and performances are compared between teams and universities. The variables of special interest are the performance metrics in the game of basketball. For the present work, these variables correspond to the number of games played throughout the season (gp), average number of points scored per game (pts), average number of rebounds per game (reb), average number of assists per game (ast), percentage of available offensive rebounds that the player captured while on the court (oreb\_pct), percentage of available defensive rebounds that the player captured while on the court (dreb\_pct), a measure of the player's shooting efficiency that takes into account free throws, 2- and 3-point shots (ts\_pct), and percentage of field goals from teammates assisted by the player while on the court (ast\_pct). On the other hand, physical variables such as age, height, and weight are available. Finally, categorical variables such as university, school, year of contract, round in which the player plays, and jersey number are also available.

1. Exploratory Data Analysis
   1. Data Description
   2. Distributions
   3. Missing Values and Imputation

Objective 1: Analyze general patterns in player performance in relation to age and physical variables.

* 1. Multivariate Outliers

Objective 2: Identify the traits of those players with exceptional metrics

1. Hotelling T 2 Test

Objective 3: Compare performance metrics between the two universities with the highest number of outstanding players.

1. MANOVA

Objective 4: Determine whether performance metrics vary significantly between the 4 teams with the highest number of featured players.

1. Principal Component Analysis

Objective 5: Find components that reduce the dimensionality of the data.

1. Factor Analysis
2. Multidimensional Scaling
3. ~~Correspondence Analysis~~
4. Multiple Correspondence Analysis

Objective 7: Analyze the relationship between the levels of defensive rebounds, level of offensive rebounds, level of assists and the levels of points obtained.

1. Cluster Analysis

Objective 6: Establish segments that group players according to performance metrics and their physical characteristics.

1. Discriminant Analysis

Objective 8: Determine the characteristics about player performance metrics that influence the team's overall playing impact.