

פרויקט ביג דאטה

שאלת מחקר - איזה סטטיסטיקות מנצחות משחק כדורסל

מקורות מידע:

המידע יילקח ישירות מהאתר הרישמי של ה-NBA. www.stats.nba.com

ניתוחי מידע:

שלב הראשון:

נבצע איסוף מידע בעזרת web scraping.

שלב שני:

בניית טבלאות csv, לעבודה עם R.

בטבלאות:

שורה - משחק.

עמודות - נתונים סטטיסטיים:

- FTPCT - free throw percentage
- 2PT% - 2 point throw percentage.
- 3PT% - 3 point throw percentage
- DREB - defensive rebound
- OREB - offensive rebound
- AST - assists
- STL - steals
- TOV - turnovers
- BLK - BLOCKS

שלב שלישי:

ננסה להגיע למסקנות מהם הנתונים המשמעותיים ביותר במשחק.

שיטות:

נעזר ב PCA כדאי לחפש נתונים בעלי קורלציה גבוהה לניצחון משחק.

נבחן את המסקנות אם הם מובהקים סטטיסטית.

נבצע קלסטרינג לחפש דמיון בין קבוצות.