עבודת סיום -ביג דאטה

דניאל גולדשטיין – 203563028

רון אוזריצקי - 308385681

רקע:

התחום שלנו הוא משחק הכדורסל שהוא משחק כדור שמשחקים בו שתי קבוצות שבכל קבוצה חמישה שחקנים. מטרת המשחק היא לקלוע לסל של היריב כמה שיותר נקודות כאשר מס' הנקודות האפשרי שאפשר להשיג בהתקפה היא 1/2/3 . נקודה אחת היא מזריקת עונשין ,שתי נקודות סל בהתקפה מתוך קשת השלוש ושלוש נקודות קליעת זריקה מעבר לקשת השלוש. הקבוצה שקולעת יותר בסיום 40 דקות המשחק היא הקבוצה שמנצחת את המשחק.

השחקנים מחולקים לחמש עמדות עיקריות אף שאין איזה חוק כתוב במשחק שמגדיר אותם. עמדות אלו הן: רכז, קלע ,סמול פורוורד, פאוור פורוורד, סנטר.

בקצרה נתאר את עמדות אלו: רכז- לרוב השחקנים הנמוכים בקבוצה, הם אלו שבד"כ מובילים את ההתקפה, המומחיות העיקרית שלהם הא בלמסור.

קלע- גם בדרך כלל הם שחקנים נמוכים והתפקיד העיקרי שלהם הוא לקלוע.

סמול- פורוורד- השחקן שנחשב למגוון ביותר בחמישייה על המגרש כי הוא שולט בדר"כ ברוב היכולות של שחקן כדורסל הוא בדרך כלל שחקן שגם קולע מרחוק וגם יודע לקלוע מקרוב.גובה שחקן זה נמצא בדר"כ באמצע של ההירכיות בין העמדות

פאוור-פורוורד – נמצא בעיקר באזור הקרוב לסל לרוב וקולע יותר מתחת לסל. עמדה זו אמורה גם לקחת יותר כדורים חוזרים מהשחקנים היותר הנמוכים. שחקנים אלו הם גבוהים יותר מהעמדות שציינו עד כה.

סנטר- העמדה ששולטת בתוך הצבע של המגרש(האזור בין קו העונשין לסל) ולכן בהתקפה הוא קולע בעיר בתחת לסל ובהגנה הוא אמור להגן על הטבעת,וגם הוא כמו בעמדה הקודמת אמור להתמחות בלקחת כדורים חוזרים. בדר"כ שחקן בעמדה הזאת הוא השחקן הגבוה ביותר מבין השחקנים שעל המגרש.

מושגים בסיסים על המשחק שעליהם בדר"כ אוספים נתונים עליהם מעבר לנקודות כמובן הם:

אסיסטים- אסיסט מתרחש כאשר נקלע סל כתוצאה ממסירה של שחקן לשחקן אחר בקבוצה שקולע את הסל.

ריבאונד – כאשר יש החטאה של זריקה והשחקן הראשון שאוסף את הכדור זוקף לזכותו את הכדור החוזר. סעיף זה מתחלק לריבאונד התקפה וריבואנד הגנה אף שבפועל מדובר על הסך הכולל של שתיהם שהוא לקח במשחק

חטיפות – כאשר שחקן חוטף את הכדור לשחקן יריב וההתקפה עוברת אל קבוצתו.

חסימה- כאשר שחקן נמצא בהגנה והוא חוסם נסיון זריקה לסל של שחקן ההתקפה

הנושא שלנו הוא השפעת הגובה על פרמטרים ועמדות במשחק. ממה שאנו מכירים נהוג להכיר ששחקני כדורסל הם אנשים גבוהים בדר"כ. לכן אפשר להבין כי הגובה הוא גורם מאוד משמעותי שמשפיע על כל תחומי המשחק, בין אם בייצור נקודות ובין אם בדברים המשלימים שהם מסירות וריבואנדים וכו'. ממה שאנו מכירים בתור אוהדים של המשחק אז השחקנים ה"נמוכים" יותר מתמחים במסירות וחטיפות וקליעת שלשות, והשחקנים היותר גבוהים מתמחים בקליעה ל2 נקודות ריבאונדים וחסימות.

שאלת המחקר שלנו באה לבדוק האם באמת יש קשר בין גובהה של שחקן להצלחה בעמדה שלו? כאשר נגדיר הצלחה ע"י כך שנבדוק האם באמת השחקנים בעמדות שלהם שולטים בפרמטרים שמאופיינים בדרך כלל לכל עמדה.

המידע שיש לנו כדי לחקור את הנושא הוא מאגר מידע של סטטיסטיקות של שחקנים משנת 1950 עד שנת 2015 שבו יש עמודות כמעט מכל סוג של נתון שיש במשחק כמו אסיסטים ,ריבאונדים,חטיפות, נקודות , אחוזים של הדברים האלה ועוד הרבה דברים. מקור המידע נלקח מאתר basketball-reference והוא מכיל מידע של כ4000 שחקנים. כתובת המאגר שלנו הוא <https://www.kaggle.com/drgilermo/nba-players-stats#Players.csv>