

חלק 2 - שאלות מחקר / שאלות עסקיות

1. שאלות המחקר:

- 1.1. האם לבשל באמצעות מסירה עדיף על בישול באמצעות שליחת כדור?
- 1.2. לאיזה קבוצות ריאל מדריד דומה על סמך עבירות, כרטיסים ושערים?

2. מדדי הצלחה - KPIs

- 2.1. אחוזי הגולים שנעשו באמצעות מסירה גדול מאחוזי הגולים שנעשו באמצעות שליחת כדור בין העונות 2011/2012 - 2016/2017.
ספציפית - מפורט מה ה-KPI צריך למדוד, את אחוז הניסיונות בהן הובקע גול באמצעות בישולים שונים.
מדיד - ניתן למדוד את אחוזי השערים שנעשו בעזרת הבישולים השונים.
בר השגה - הנתונים קיימים בסט הנתונים שבחרנו.
רלוונטיות - נותן תובנה עבור קבוצות כיצד הכי כדאי לנסות ולהבקיע שער.
תחום בזמן - 5 עונות שיש בדאטה סט.

- 2.2. נבדוק האם יש קרבה בין קבוצות על סמך התכונות עבירות, כרטיסים ושערים שקרו בין העונות 2011/2012 - 2016/2017.
ספציפית - מפורט אילו תכונות נמדוד על מנת להשיג את המטרה.
מדיד - האירועים מתועדים ולכן ניתן לספור אותם עבור כל קבוצה.
בר השגה - ניתן לבצע PCA כדי להוריד את מימדי הבעיה ולבצע kmeans.
רלוונטיות - נותן תובנה עבור קבוצות על אופי הקבוצות היריבות.
תחום בזמן - 5 עונות שיש בדאטה סט.

חלק 3 - הבנת הנתונים

5. עבור כל טבלה לקחנו 2 משתנים עם האנטרופיה הכי נמוכה.
6. בגלל שבטבלה הראשונה (של אירועי המשחק) לחלק מן הנתונים שערכם הוא מספרי מציינים נתונים קטגוריאליים ואין משמעות למספר עצמו, כך שאין משמעות לממוצע ולערכים הסטטיסטיים שמופקים שם.
בטבלה השנייה ניתן לראות שיש קורלציה בין משתנים שונים אך המשתנים הללו הם משתנים של יחסי ההימורים כך שקורלציה בהם היא בלתי נמנעת והגיונית.
כמות הגולים בממוצע שקבוצת הבית מבקיעה היא 1.548062 עם סטיית תקן של 1.309591
כמות הגולים בממוצע שקבוצת החוץ מבקיעה היא 1.154964 עם סטיית תקן של 1.142596