**תרגיל מעבדה 5**

יש לטעון את הקובץ lab.csv

1. יש לשמור את הטבלה בתור משתנה חדש שלא מכיל את המשתנה ה-Class.

2. יש להשתמש ב-kmeans (k=2) על-מנת לחזות לאיזה קלסטר תשתייך כל רשומה.

1. מהם הגדלים של הקלסטרים שהתקבלו?
2. כמה רשומות שנכללו בקלסטר הלא נכון? (בהנחה שאנחנו יודעים את הסיווג הנכון).
3. צרפו את וקטור ההשתייכות לקלסטר לטבלה המקורית (שלא מכילה את טור ה-Class) ושנו את שם העמודה שהתווספה ל-Cluster (היעזרו בגוגל כדי לגלות כיצד עושים זאת...).
4. יש לצייר גרף (הספריה ggplot2):
   1. ציר X הוא המשתנה X1, ציר Y הוא המשתנה X2
   2. צבע הנקודות נקבע ע"פ החלוקה לקלסטרים.
   3. כותרת הגרף תהיה Cluster.
   4. כותרת ציר X תהיה "day" , כותרת ציר Y תהיה "n of cases"
   5. \*\* שימו לב ש-Class הוא נומרי. כדי שהצבעים יוצגו יש לבקש מהפונק' Plot להתייחס לשדה זה כ-factor (אין צורך לשנות את סוג השדה ב-dataframe, אלא רק עבור יצירת התרשים).

**הוראות הגשה**

* הגשה בזוגות בלבד.
* עליכם להגיש קובץ word ובו תשובות לשאלות הנ"ל, הכוללות את הפקודות בהם השתמשתם ואת התוצאות שנתקבלו בכל סעיף.
* שם קובץ ה- word אמור לכלול את מספרי תעודות הזהות של שתי המגישים וביניהם קו תחתון.
* רק אחד מבני הזוג יגיש את העבודה באתר הקורס.
* ניתן להגיש את התרגיל עד שבוע לאחר המעבדה בה פורסם לתיבת ההגשה הרלוונטית באתר הקורס.

***בהצלחה!***