

README – הוראות למפעיל – נושאים נבחרים בסטטיסטיקה

קובץ R - regression: תוך כדי הרצת הקוד, יש לבחור בקובץ ה-excel אשר נקרא regression.

קובץ R - logistic regression: תוך כדי הרצת הקוד, יש לבחור בקובץ ה-excel אשר נקרא logistic regression.

קובץ פייתון - main: מכיל קוד אשר מחולק לפונקציות, פירוט עבור כל אחת מהפונקציות, לאיזה חלק רלוונטית בעבודה:

t_test(): ביצוע מבחן t בין הסמסטרים לטובת בדיקה אם חל שיפור בציון התלמידים מסמסטר א' לסמסטר ב'.

outliers_samples(): תרשימי קופסא למשתנים גיל והיעדרויות לצורך בדיקה אם קיימים תלמידים חריגים.

graphs_of_explanatory_variables(): יצירת תרשימים עבור כל אחד מהמשתנים לצורך בדיקת התפלגויות וחלוקה של קטגוריות.

feature_selection(): ביצוע סינון של כמות המשתנים לטובת ביצוע רגרסיה לפנים/לאחור/בצעדים.

unification_of_categories(): איחוד קטגוריות לאחר בחירת המשתנים.

feature_selection_logistic_regression(): בחירת משתנים לרגרסיה לוגיסטית לאחר הפיכת המשתנה התלוי לבינארי.

unification_of_categories_logistic_regression(): איחוד קטגוריות פעם נוספת עבור הרגרסיה הלוגיסטית (בוצע רק עבור המשתנים שנבחרו גם ברגרסיה הרגילה).

ML_models(): ביצוע מודלים נוספים של למידת מכונה.