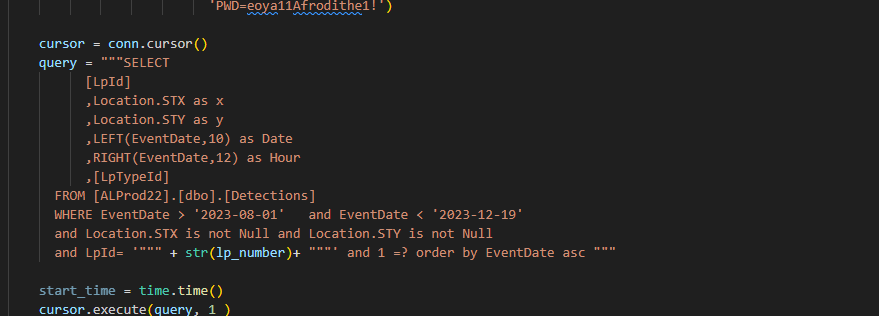
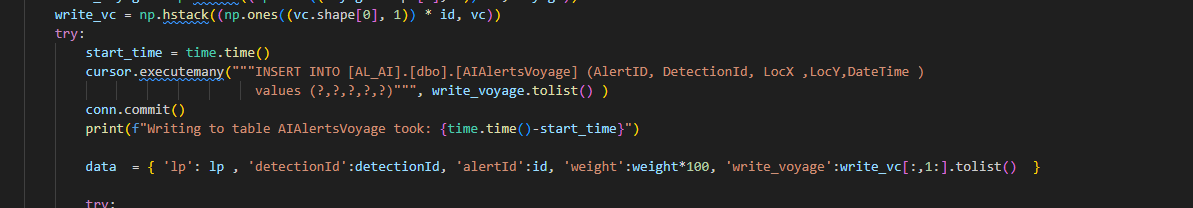
Mini documentation:



זה השאילתה ששימשה למשיכת המידע לאימון



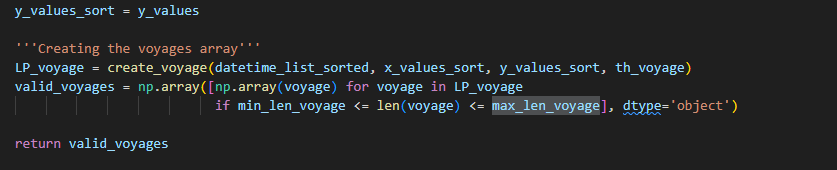
הטבלה שאליה מכניסים את החריגה משיגרה:.

נתוני קונפיגרציה :

אורך מינמיל של מסלול הוא 3 , מקסימלי 5 .

  min\_length\_voyage = 3 # min length of a voyage

        max\_length\_voyage = 50 # max length of a voyage

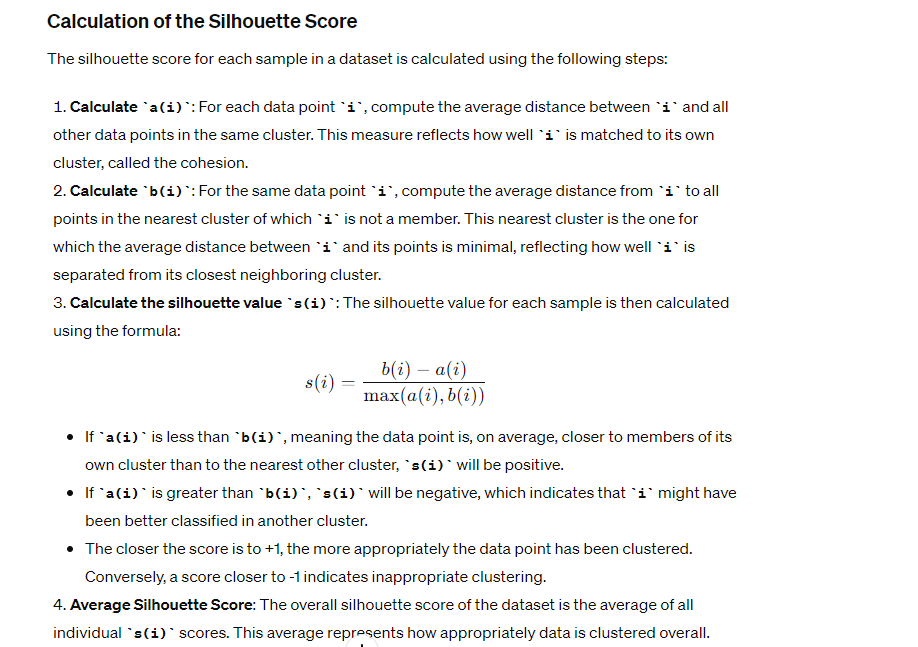


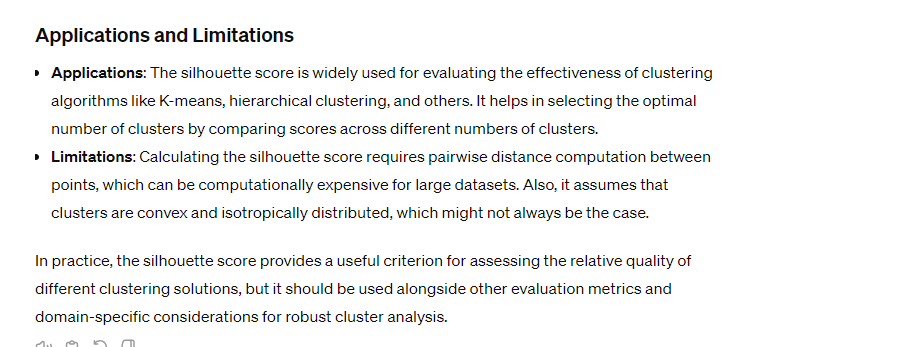
הclusttering נעשה על לפחות 30 מסלולים 

יש בין cluster 1 ל10 clustters , מספר הקלאטסטרים נקבע לפי silouette score.

הסבר על silouette score:

Higher silluate score indacte better matching





**בחירת מספר הקבוצות :** מבוסס לפי k-means כאשר מטריקת המרחק היא לפי dtw ,ומספר הקבוצות היא לפי sillhute score :

התנאי כדי לבחור את מספר הקבוצות מבוסס לפי silhouette score :

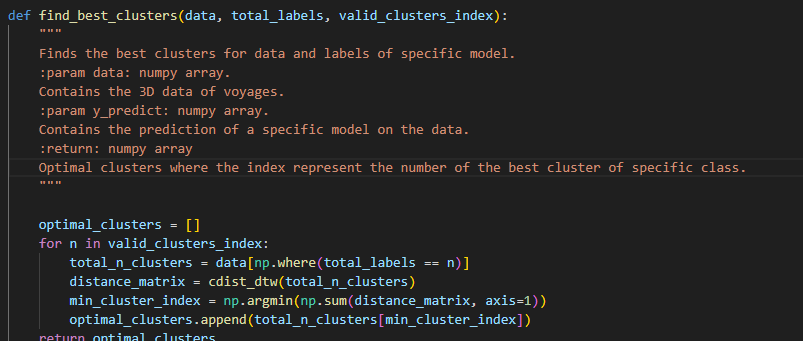
אם אין אפחד שגדול מה threshold לוקחים את האינדקס המקסימלי ואם יש אז לוקחים את האחרון שעומד בתנאי

(יכול להיות שלא מייצג ואי אפשר לקלסטר) ?

במקרה שזה כן עומד זה מחזיר את המספר המקסימלי של קלאסטרים שעומדים בתנאי של 0.85.

אחרכך יש עוד פלטור על אף שמצאנו את מספר הקלרטסרים שאם יש קלאסטר שמספר הvoyages שתויגו אליו הוא פחות מאחוז אז לא מתייחסים אליו .

המסלול שנחשב מייצג נבחר הבאה :



יש שדות נוספים של סוג רכב צבע רכב , סוג רכב פלסטני,פלסטני ישראלי צבאי במצלמות hts .