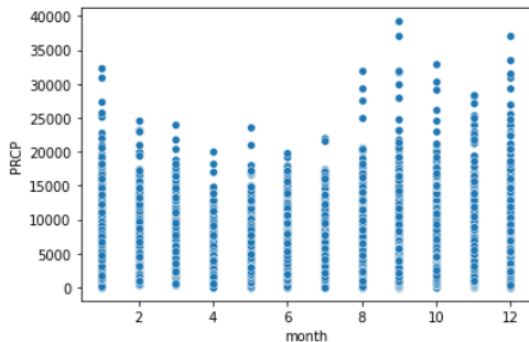


1. ניתוח מבוסס זמן:

מטרתנו הייתה למצוא קשר בין חודש מסויים לבין המשקעים באותו חודש (האם יש הבדל בין המשקעים בינואר ופברואר?). הקשר שמצאנו הוא קשר של טווחים שונים ושונות הנתונים. כפי שניתן לראות בתרשימים הבאים, ישנם חודשים בהם יש טווח גדול יותר ושונות גבוהה במשקעים. קשר זה מרמז על כך שהכללת במודל הלמידה יכולה לתרום לדיוקו.
תרשימים:

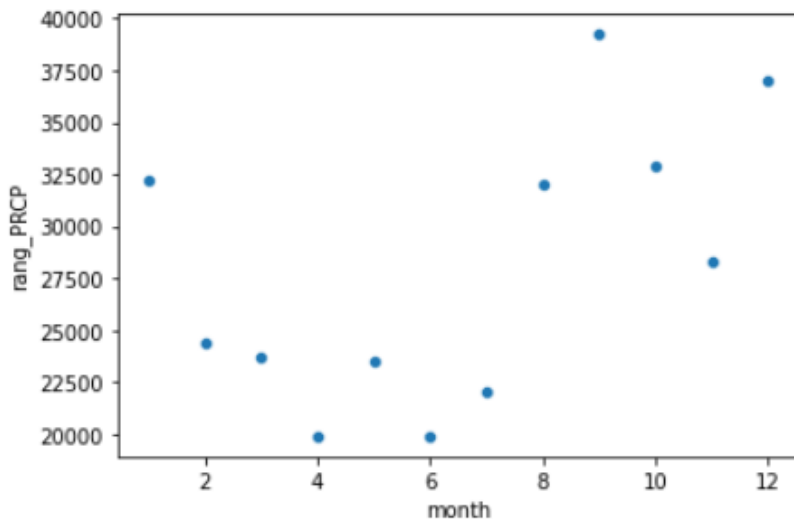
מופעי המשקעים בחודשים השונים:

אורך טווחי כמות המשקעים בחודשים השונים:



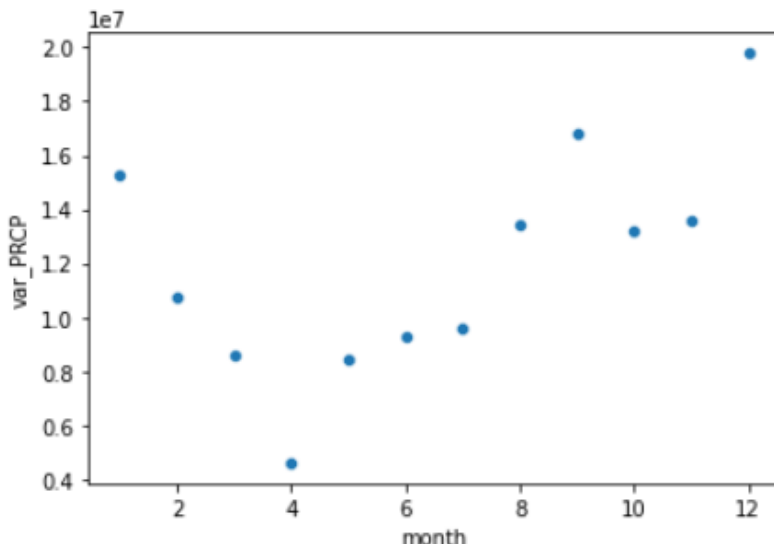
ניתן לראות כי קיימת שונות התפלגות בין החודשים השונים ולכן ישנה הצדקה לבחון את הימים על בסיס החודש שבו הוא נמצא שכן הדבר מסייע להבחנה מאיזה התפלגות הוא נלקח.

גודלי טווחי המשקעים בחודשים השונים:

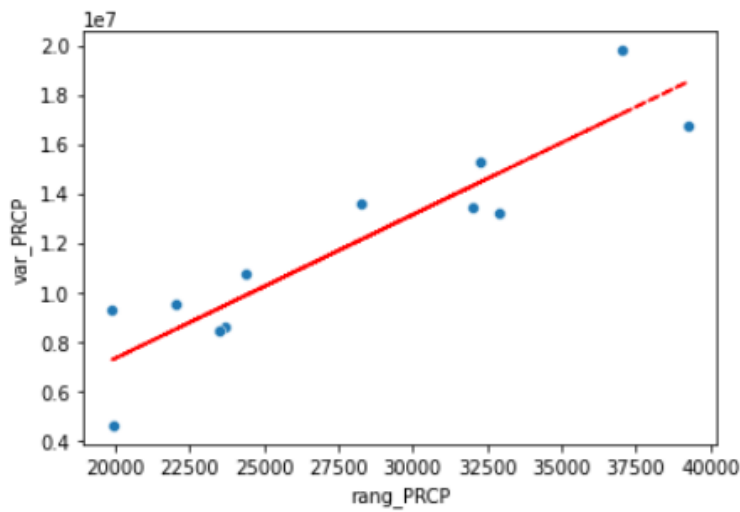


נשים לב כי קיימת שונות גדולה בין גודלי טווחי המשקעים. הטווח אינו גודל אינפורמטיבי במיוחד מפני שהוא רגיש מאוד לתצפיות קצה. לדוגמה עבור חודש יבש עם מספר מאורעות משקעים חזקים, הטווח יהיה גדול אך הדבר לא מעיד על דבר מעבר לכך שהיה יום בעל כמות משקעים השווה לערך הטווח.

שונות המשקעים בחודשים השונים:



ניתן לראות כי למעשה יש הבדל ניכר יחסית בשונות המשקעים בחודשים השונים. הדבר מעניק אינפורמציה על "רמת האמינות" של כל חודש. כלומר אם נתונה התפלגות של חודש מסוים ככל שהשונות של חודש זה נמוכה יותר כך ההתפלגות מעניקה אינדיקציה גדולה יותר לערך האמיתי של המשקעים באותו חודש

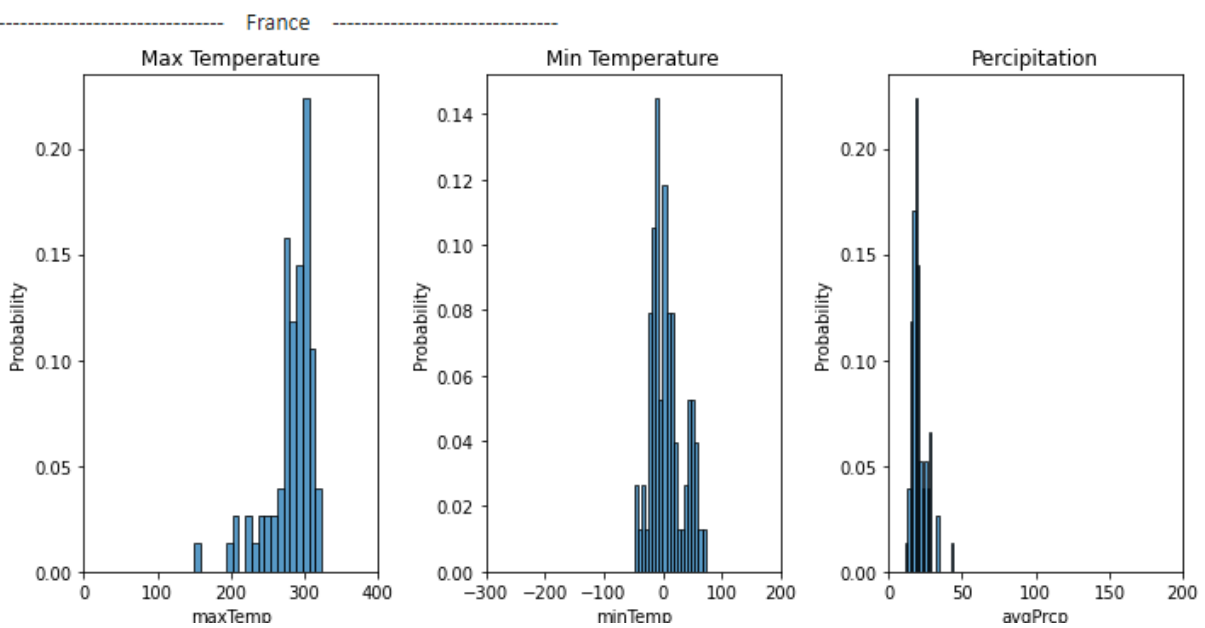
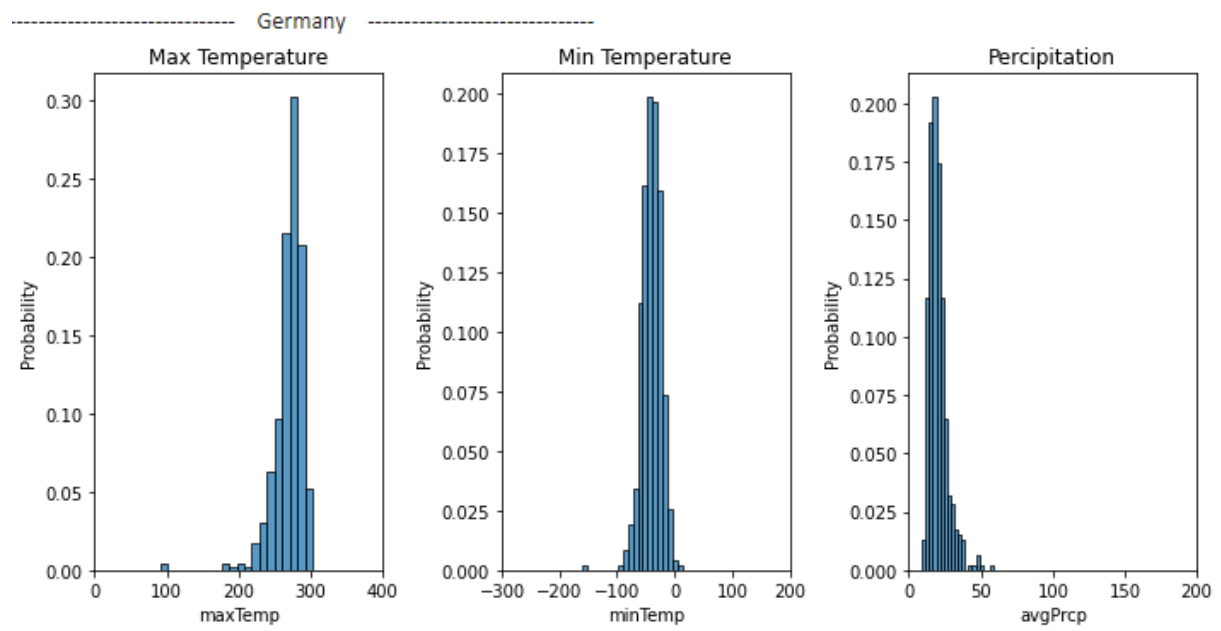


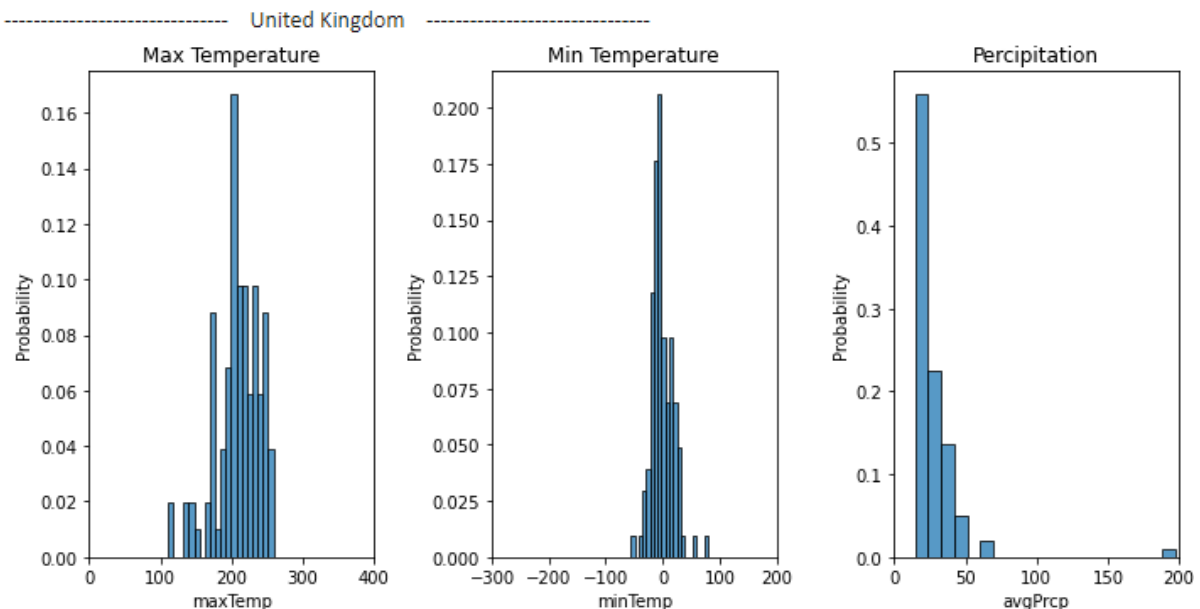
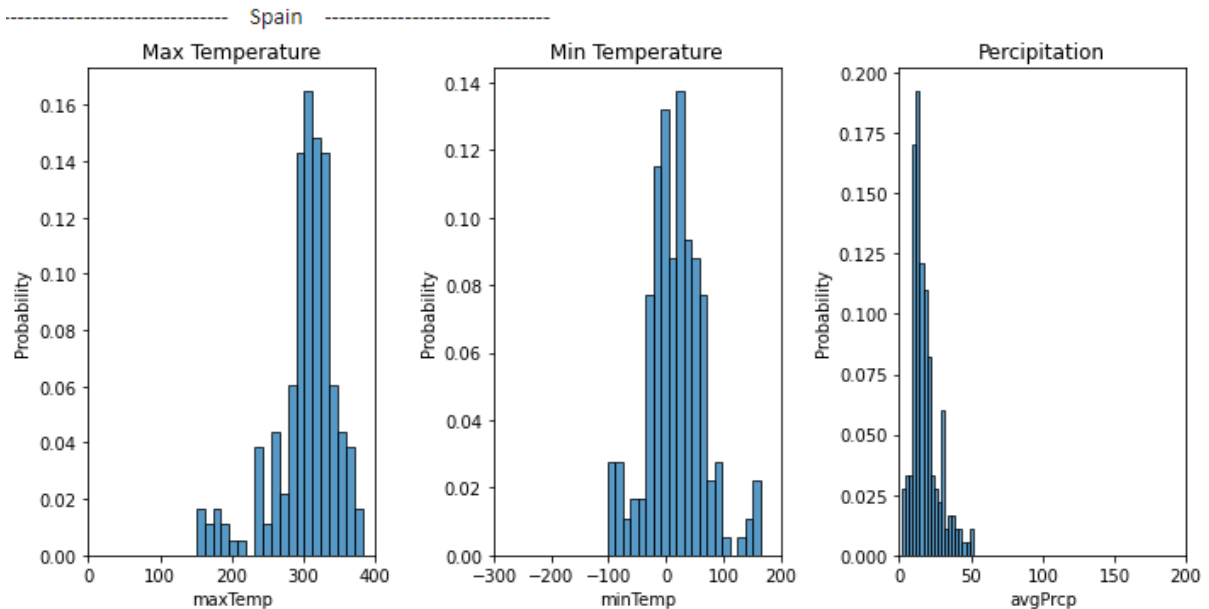
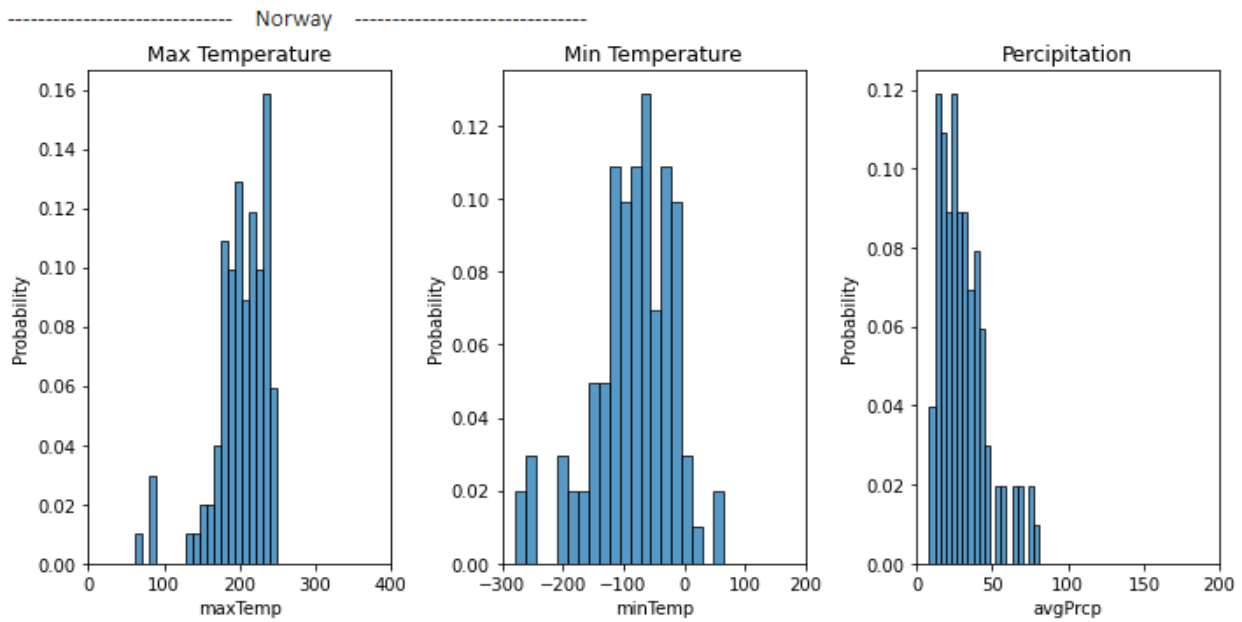
קשר בין טווחי המשקעים לבין השונות:

ניתן לראות כי קיים קשר לינארי בין טווחי המשקעים לשונות במשקעים, או במילים אחרות, ככל שטווח כמות המשקעים בחודש גדול יותר כל השונות יותר גדולה. משמעות הדבר היא שבחודש מסויים טווח כמות המשקעים המתקבל אינו תוצאה של תצפית קצה (outliers) אלא הוא אכן הטווח האפשרי של משקעים שמתקבל בחודש.

2. ניתוח מבוסס מקום:

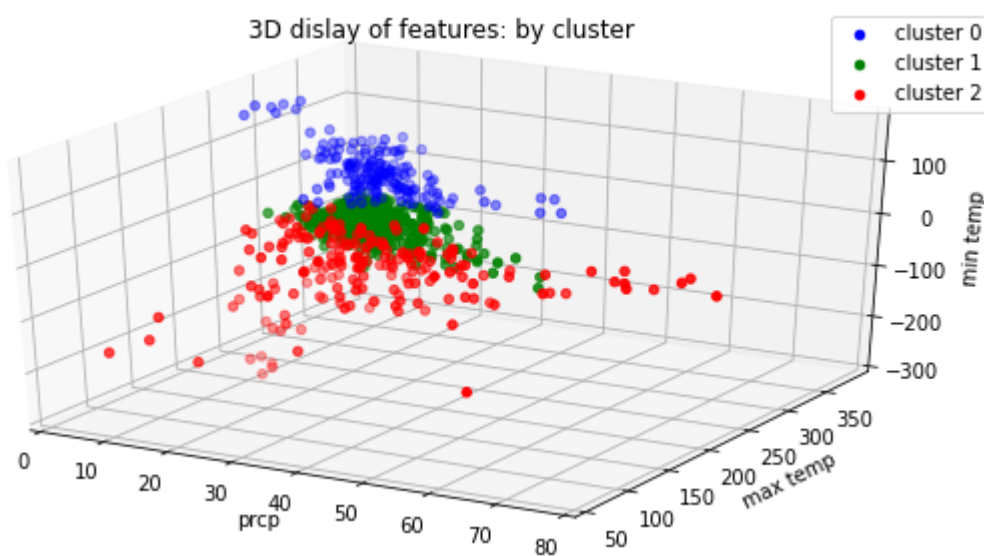
המטרה שלנו היא למצוא קשר בין מיקום (מדינה או אזור אקלים, כלומר קלסטר). תחילה נדון בניתוח מבוסס מדינות. לאחר בידוד כל נתון לפי המדינה, כפי שניתן לראות מהגרפים קיים קשר בין המדינה לטווח המשתנים והשונות בקרב הטווחים הללו (לדוגמא עבור אנגליה גודל הטווחים והשונות קטנים משמעותית). הדבר סביר כי לכל מקום יש טווח שונה של טמפרטורות ומשקעים והדבר מתבטא בקשר שמצאנו.





אך כפי שניתן לראות ההבדל בין רוב המדינות שונות לא משמעותי ולכן נסתכל על ניתוח לפי קלסטרים כאשר כל קלסטר מייצג אזור אקלים שונה על מנת לקבל. את הקלסטרים השונים יצרנו על ידי שימוש באלגוריתם kmeans על גבי טמפרטורה מקסימלית, מינימלית ומשקעים שכן לדעתנו משתנים אלו מגדירים אזור אקלים. לאחר ניסיון עם מספר שונה של קלסטרים החלטנו על 4 אזורי אקלים שונים ועל ידי הצגה במרחב קיבלנו את דרך ההצגה הבאה:

כאשר נוכל לראות שניתן לראות את השינוי בין הקלסטרים השונים, יש אזורי אקלים בעלי שונות גבוה יותר (ב3 יש שונות גבוהה של משקעים, ב1 יש שונות גבוהה של טמפרטורה מקסימלית) ויש אזורי אקלים בעלי ערכי שונות נמוכה יותר (0 ו2). בנוסף לכך ניתן לראות לדוגמא כי אזור אקלים 0 קר יותר מאשר אזור 2 מפני שבאופן כמעט מוחץ ערך הטמפרטורה המינימלית באזור 0 נמוך יותר מהנקודות באזור 1.



נאפיין את אזורי האקלים:

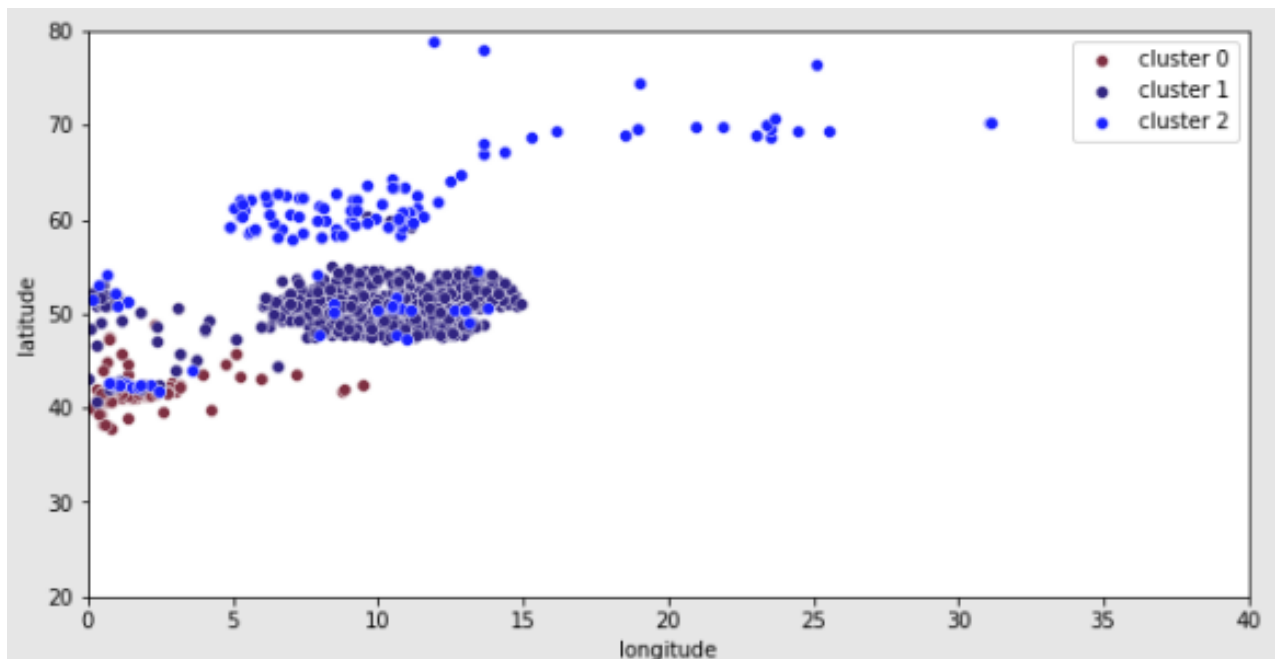
אזור 0 – מזג אוויר חם מאוד
בעל כמות משקעים נמוכה

אזור 1 – יש באזור זה
טמפרטורה מינימלית נמוכה
טמפרטורה מקסימלית גבוהה,
ובעל כמות משקעים נמוכה.

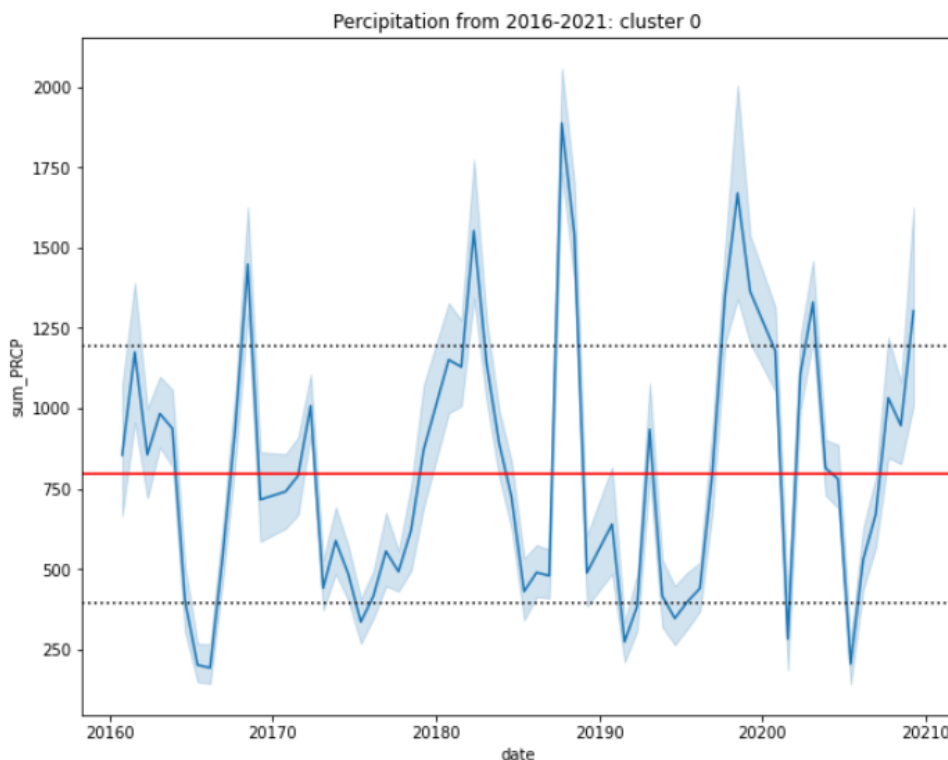
אזור 2 – כמות משקעים
וטמפרטורות בעלות שונות
גבוהה. כמות המשקעים גבוהה
באופן יחסי וטמפרטורה נמוכה (גם מינימלית וגם מקסימלית)

הגרף להלן מציג את אזורי האקלים השונים על גבי מפת אירופה:

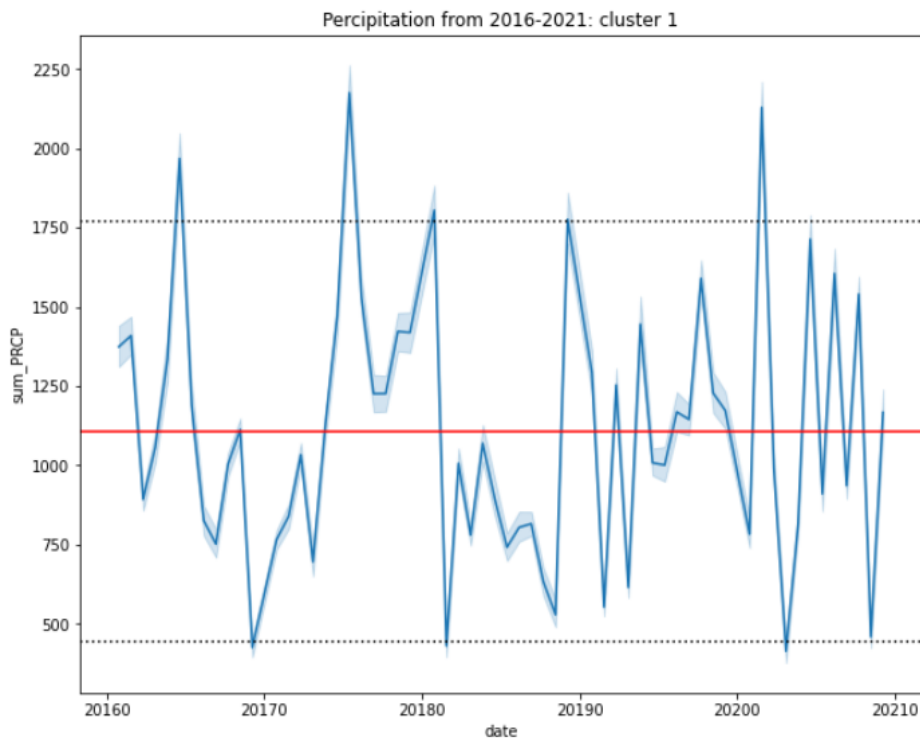
ואכן ניתן לראות כי הקלסטרים אכן מייצגים אזורי אקלים שונים על גבי יבשת אירופה, תופעה זו משתלבת עם ההנחה שכל קלסטר מציין אזור אקלים שונה. וכך החלוקה לקלסטרים שמצאנו מחלקת את החלק הנל באירופה לאזורי אקלים שונים כאשר לכל חלק מאפייני אקלים שונים



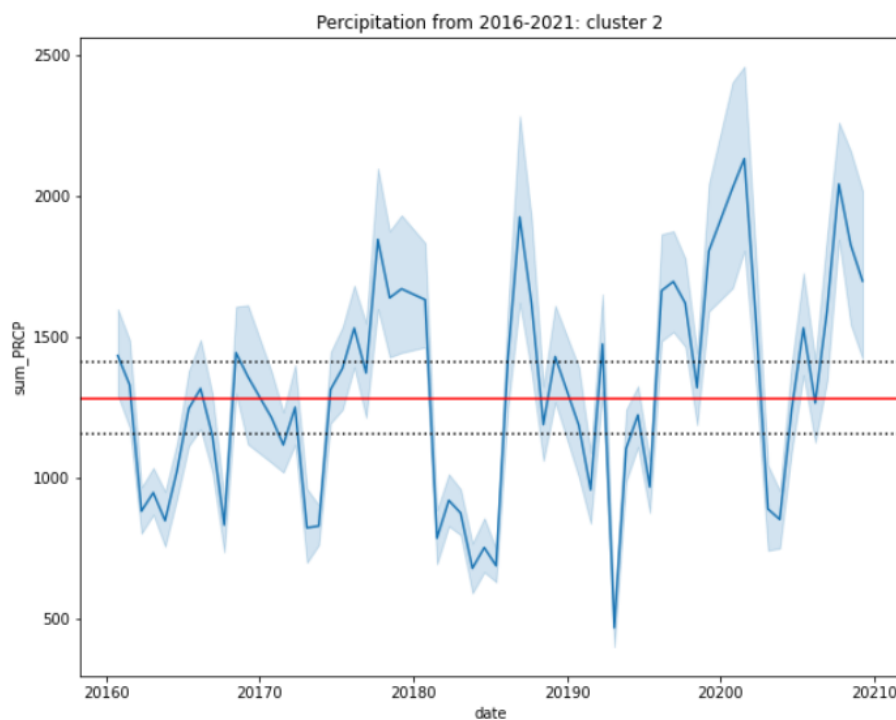
3. ניתוח מבוסס זמן מקום: המטרה שלנו היא לבחון את השוני בין אזורי אקלים שונים לאורך השנים, ואת השינויים במשקעים לאורך השנים בכל אזור אקלים. לצורך כך יצרנו גרף לכל קלסטר המתאר את כמות המשקעים במהלך השנים בנוסף לשלושה קווים המציינים את ממוצא המשקעים בקלסטר (הקו האדום) ומרחקי שונות המשקעים מהממוצע. על ידי שני הקווים המקוונים נוכל למעשה לסווג תצפיות קצה, כאשר הערך לא בין שני הקווים המקוונים. הגרפים:



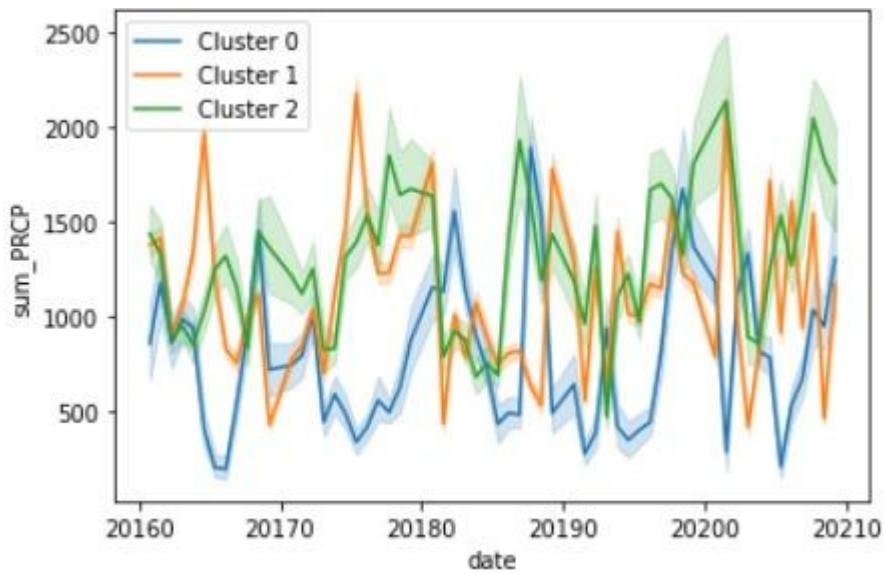
ניתן לראות כי בקלסטר זה ממוצע כמות המשקעים היא בערך 800, השונות היא לא גדולה במיוחד ואין יותר מידי תצפיות קצה. בנוסף טווח המשקעים נע בין 1800 ל-200. ניתן להסיק כי המשקעים הממוצעת באזור אקלים זה נמוכה יחסית ולא משתנה לרוב. כמו כן אין הרבה אירועים בעלי כמות משקעים יוצאת דופן.



ניתן לראות כי בקלסטר זה ממוצע כמות המשקעים היא בערך 1100, השונות היא גדולה יחסית ויש מעט מאוד נקודות קצה. בנוסף טווח המשקעים הוא בין 400 ל- 2250. ניתן להסיק כי המשקעים הממוצעת באזור אקלים זה מומצעת אך משתנה. יתרה מכך ישנה כמות נמוכה של ימים בהם כמות המשקעים יוצאת דופן



ניתן לראות כי בקלסטר זה ממוצע כמות המשקעים היא בערך 1300, השונות היא קטנה יחסית ויש הרבה נקודות קצה. בנוסף טווח המשקעים הוא בין 400 ל- 2500, אך בניקוי נקודות הקצה הטווח הוא בין 700 ל- 2000. ניתן להסיק כי המשקעים הממוצעת באזור אקלים זה גובהה ואינה משתנה. נשים לב כי ישנה כמות גדולה יחסית של ימים בהם כמות המשקעים היא יוצאת מגדר הרגיל.



נשים לב שלמעט תצפיות קצה
נשמרת המגמה שבה סכום הגשמים
עבור התכנות בקלסטר אפס קטן
מהסכום עבור תחנות בקלסטר 1
שקטן מאלו בקלסטר 2. ולכן לאורך
השנים המגמה נשארת זהה.

לסיכום, ניתן לראות כי אכן רמת המשקעים בכל קלסטר "מתנהגת שונה" גם כאשר מסתכלים על המשקעים
לאורך זמן. יתרה מכך, "יחס הסדר" בין הקלסטרים השונים נשמר לאורך השנים כלומר אזור גשום באופן יחסי
לאזור אחר נשאר כזה לאורך זמן.