

השוואת מערכות בחירות 2020-2021 בישראל

Oded Mousai

Goal

בבעבודה זו ננסה למצוא הבדלים בין מערכות הבחירות מרץ 2020 ו-מרץ 2021 בישראל בעזרת שימוש בכלים להורדת מימד.

Data

הדאטה כולל 2 טבלאות שמכילות את נתוני ההצבעות למפלגות השונות לפי ישובים.

הטבלאות נלקחו מאתר [ועדת הבחירות המרכזית](#).

	סמל ועדה	סמל ישיבה	בוב	נציגים/ערים	פסיכיים	כשרים	אמת	ג	ודעם	ז	דך	זן	טב	י	יו	יג	יק	יר	כ	כן	ל	נחלה	נ	נו	ני	נין	נק	סה	ף	ץ	
שם ישיבה																															
אבו גווייער שט	18	967	2018	595	6	589	13	0	554	0	0	0	0	0	1	0	0	0	0	1	1	4	0	1	0	0	0	8	2	0	
אבו גוש	2	472	4821	2718	36	2682	75	4	2157	0	0	0	5	0	0	5	1	0	1	0	5	280	0	0	0	3	2	123	0	2	
אבו סטאן	6	473	9422	6090	53	6037	92	6	4458	0	1	0	16	1	2	13	2	0	1	1	2	230	190	3	2	1	3	3	752	0	2
אבו שבט	18	958	142	41	1	40	0	0	40	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
אבו קורינטא שט	18	968	3037	1336	8	1328	9	0	1263	0	0	0	1	0	0	3	0	0	1	0	0	8	0	1	2	1	0	33	0	2	
...	
תקומה	18	665	488	416	2	414	5	5	0	0	0	0	163	0	0	0	0	0	0	0	1	161	1	0	0	2	1	63	0	0	
תקוו	2	3563	2158	1619	4	1615	28	33	3	7	0	0	760	0	0	0	0	0	0	5	1	35	583	2	0	0	29	0	85	2	0
תרנגא אצונע שט	18	970	520	48	0	48	0	0	43	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	
תרביא אצאנע שט	18	1346	587	206	0	206	0	0	160	0	0	0	0	0	0	1	0	0	0	0	0	40	0	1	0	0	0	2	1	0	
תרומ	2	778	576	472	1	471	9	7	0	0	0	0	85	0	0	0	0	0	0	0	1	254	1	0	0	2	0	65	0	0	
1214 rows x 37 columns																															

טבלת הצבעות בחירות 2020

Pre-processing

נאחד את טבלאות ההצבעות של שתי מערכות הבחירות ע"י שרשרור של שני הטבלאות לטבלה אחת בצורה אנכית (כלומר לפי המפלגות).

כדי להתמודד עם השינויים בשמות והרכבי המפלגות בין שתי מערכות הבחירות נבצע את הדברים הבאים:
עבור 2020 ניקח רק את 8 המפלגות הגדולות שעברו את אחוז החסימה: הליכוד, כחול לבן, שס, יהדות התורה, ימינה, עבודה-גשר-מרצ וישראל ביתנו. כעת עבור 2021 ניקח את אותן המפלגות, אך נוסיף למפלגות הרלוונטיות את ההצבעות של מפלגות שהתפצלו מהן במערכת הבחירות של 2021; כך למשל למפלגת כחול לבן נוסיף את ההצבעות של מפלגת יש עתיד, ולמפלגת ימינה נוסיף את ההצבעות של מפלגת הציונות הדתית. בטבלה הבאה ניתן לראות את כל השינויים:

2020	2021
הליכוד	הליכוד + תקווה חדשה
כחול לבן	כחול לבן + יש עתיד
הרשימה המשותפת	הרשימה המשותפת + רעם
שס	שס
יהדות התורה	יהדות התורה
העבודה-גשר-מרצ	עבודה + מרצ
ישראל ביתינו	ישראל ביתינו
ימינה	ימינה + הציונות הדתית

הטבלה המאוחדת:

	ב	ל	מחל	שש	ג	כך	אמת	ידעם
שם ישוב								
אבירים	2	2	13	1	0	43	78	3
אדירים	16	0	90	61	4	5	1	0
אושה	7	8	39	0	0	123	92	3
אחוזה	8	3	35	2	0	73	24	0
איבים	14	0	17	1	0	5	0	0
...
תקומה	165	1	161	12	5	63	5	0
תקוע	789	35	583	34	33	85	28	3
תראבין אצאנע שבט	0	0	0	0	0	4	0	43
תראבין אצאנע ישוב	0	0	40	0	0	2	0	160
תרום	87	1	254	47	7	65	9	0

2429 rows × 8 columns

נשים לב ששתי מערכות הבחירות התרחשו בהפרש של שנה בלבד זו מזו, ולכן יש סיכוי טוב שרשימת הישובים נשארה כמו שהיא כך שעבור כל ישוב נקבל 2 שורות בטבלה המאוחדת.

Dimensionality Reduction

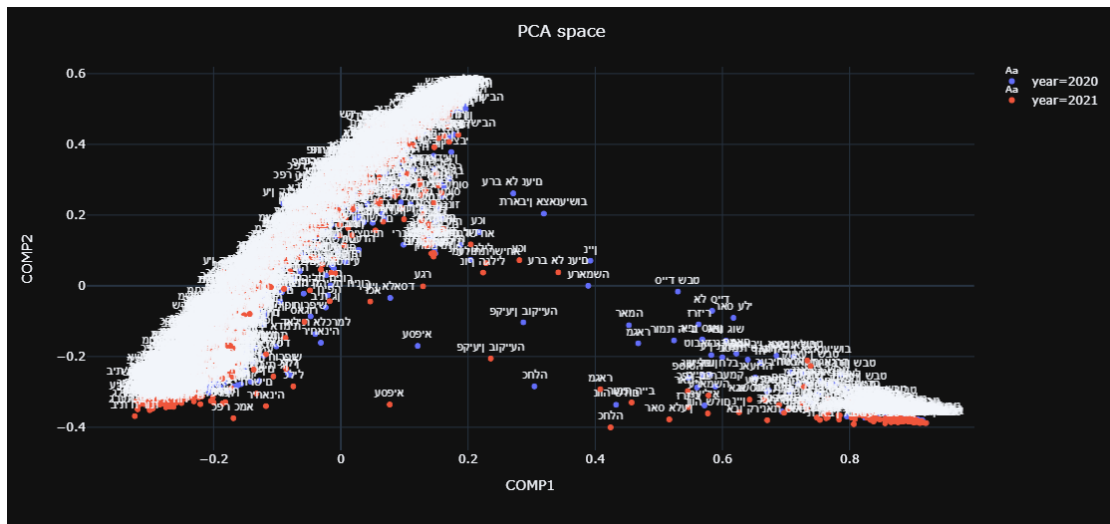
ננרמל את הטבלה המאוחדת ונבצע הורדת מימד ל-2 מימדים באמצעות PCA כאשר מתייחסים לישובים כדגימות ולמפלגות כפיצ'רים. תוצאות:



כל נקודה במרחב מייצגת ישוב מסוים, והישובים נצבעו לפי מערכת הבחירות.

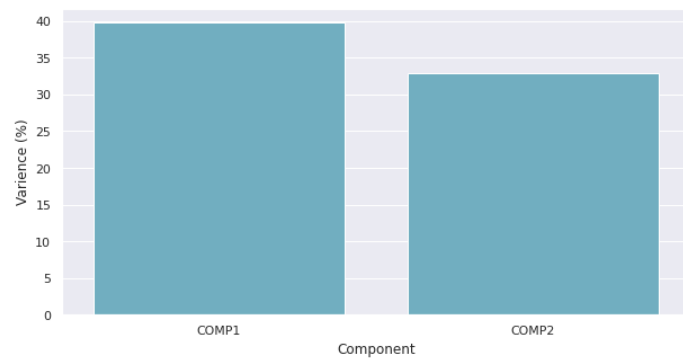
הבנת המרחב

כדי להבין את המרחב שהתקבל, נסתכל ראשית על שמות הישובים:



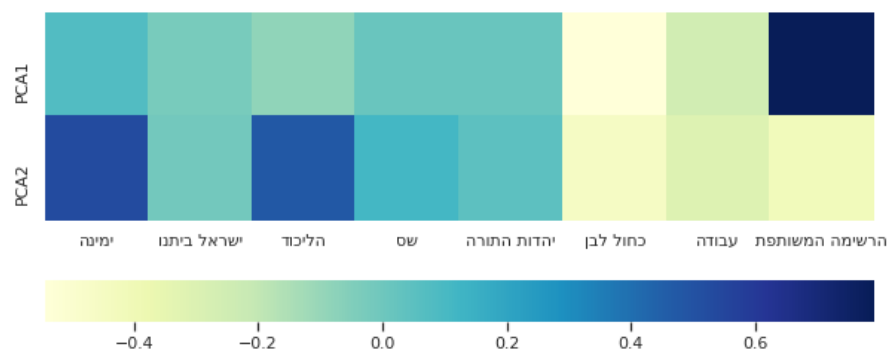
מסקירה שטחית ניתן לראות שגוש הישובים שנמצא למטה מימין מכיל הרבה ישובים שמשויכים למגזר הערבי, והגוש השמאלי הארוך מכיל הרבה ישובים שמשויכים למגזר היהודי.

Scree Plot



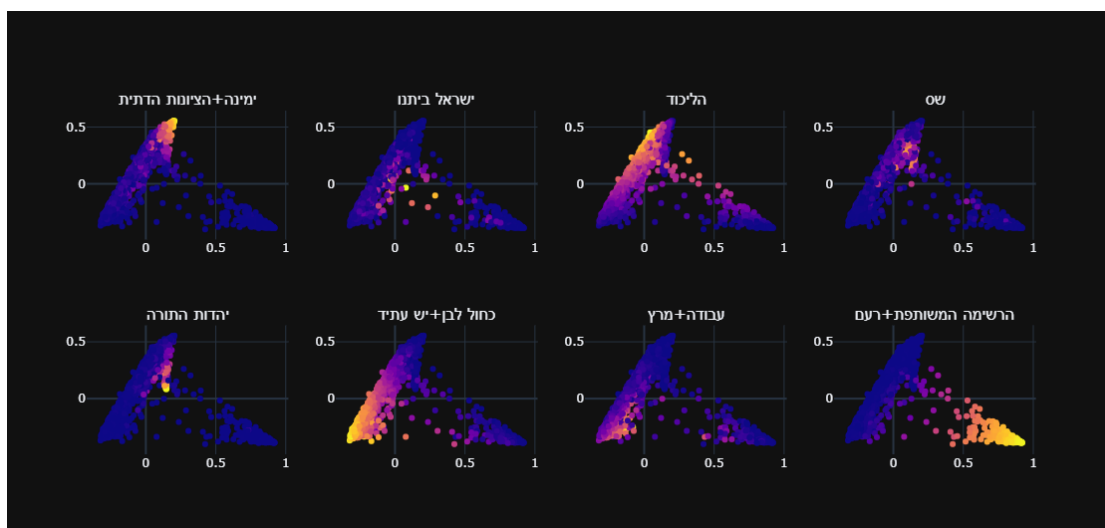
אפשר לראות שב-73% מהשונות בדאטה מוסברת ע"י שני הרכיבים הראשיים המובילים.

Effect of the features on each component



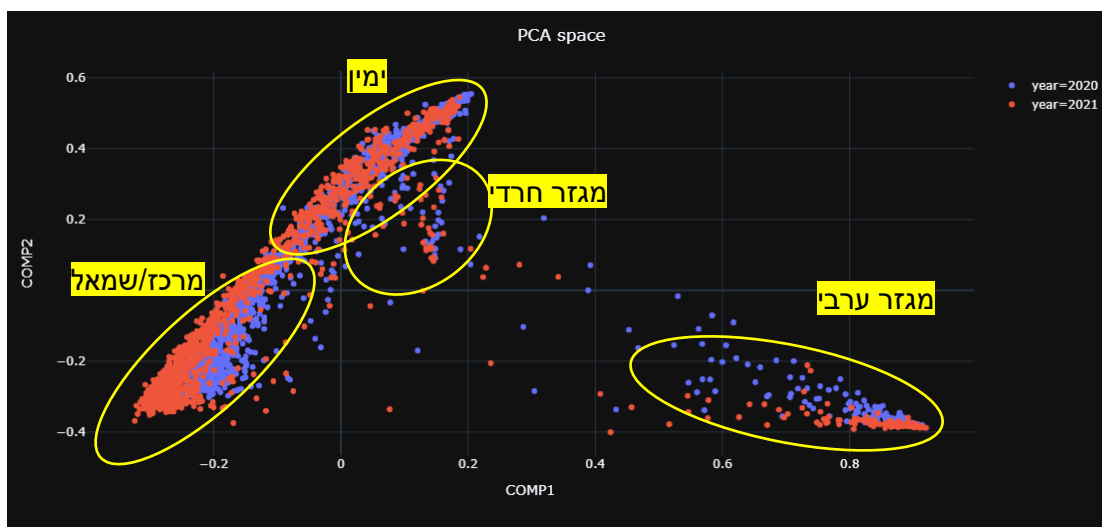
ניתן לראות שבמרכיב הראשי הראשון הרשימה המשותפת משפיעה באופן חזק.

צביעת המרחב לפי אחוזי ההצבעה לכל מפלגה:



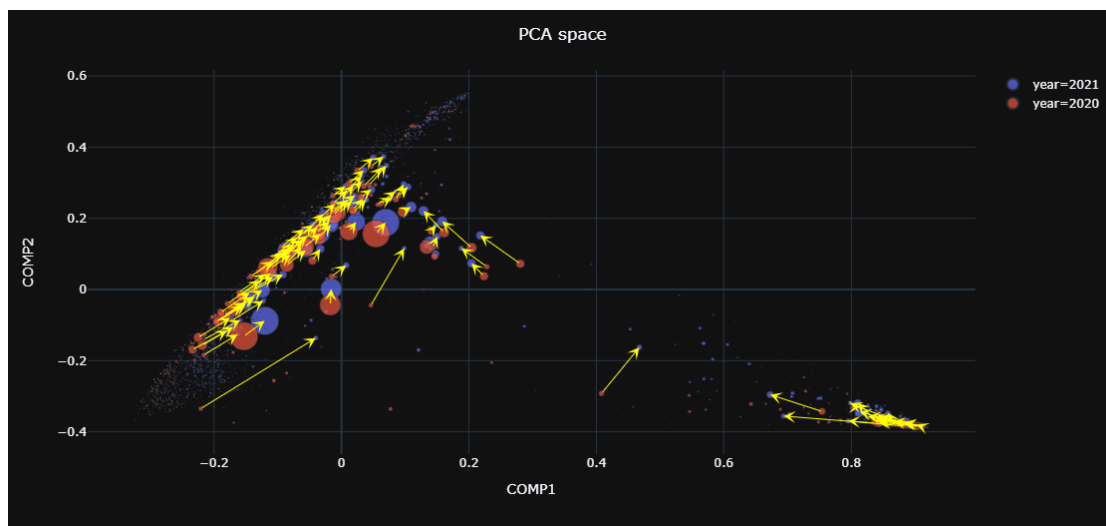
ככל שהצבע של נקודה יותר "חם" כך התמיכה של הישוב במפלגה עולה. ניתן לראות למשל שבגוש הישובים שזיהינו ששייכים למגזר הערבי התמיכה במפלגות הרשימה המשותפת ורע"ם אכן גבוהה. בנוסף אפשר לראות שבגוש השמאלי הארוך של הישובים יש ביטוי לנטייה פוליטית – בקצה העליון יש תמיכה במפלגות ימניות וככל שיוודים יש יותר תמיכה במפלגות מרכז/שמאל.

סיכום הבנת המרחב



מגמות בדפוסי הצבעה

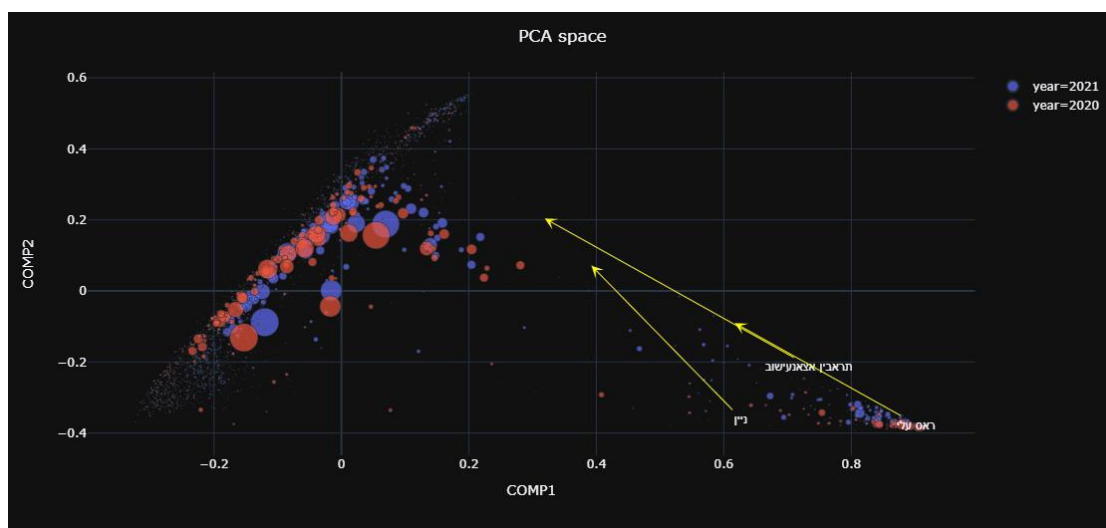
כעת נצייר חץ מכל ישוב בבחירות 2020 לאותו ישוב בבחירות 2021. נבצע זאת רק לישובים "מספיק גדולים" שמספר הבוחרים בהם גדול מ-10,000. בנוסף נציג את המרחב כך שגודל כל נקודה פרופורציונלי לגודל הישוב שהיא מייצגת.



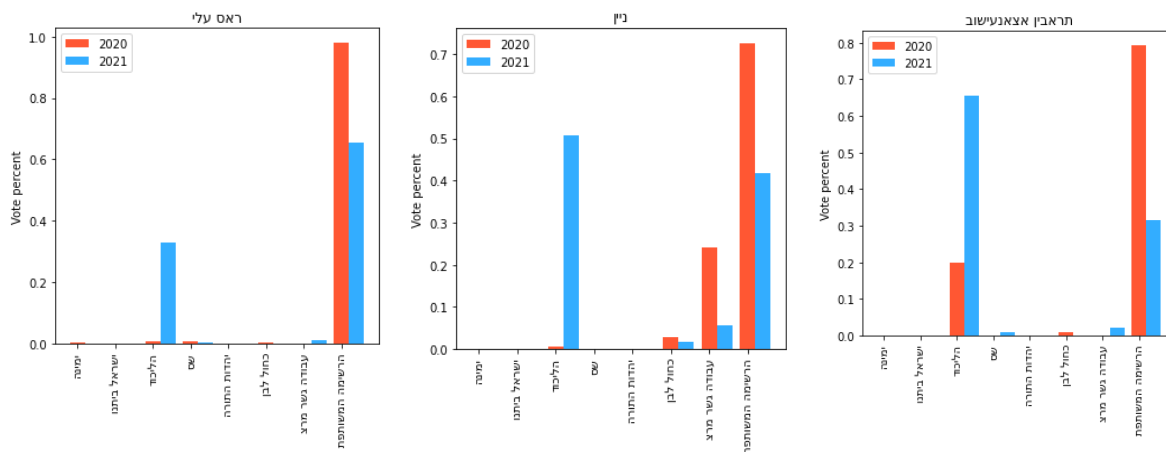
ראשית ניתן לראות שבאופן כללי ישובים גדולים חזו מעט לעומת ישובים קטנים, כלומר דפוסי ההצבעה בהם לא השתנו באופן דרמטי.

שנית, אפשר לראות שבגוש השמאלי של הישובים רוב החיצים פונים כלפי מעלה ושבגוש הימני רוב החצים פונים שמאלה, וזה עשוי להעיד על כך שהתמיכה במפלגות הימניות בבחירות 2021 גדלה מבחירות 2020.

שלושת הישובים שחזו הכי הרבה (ללא תלות בגודל הישוב):



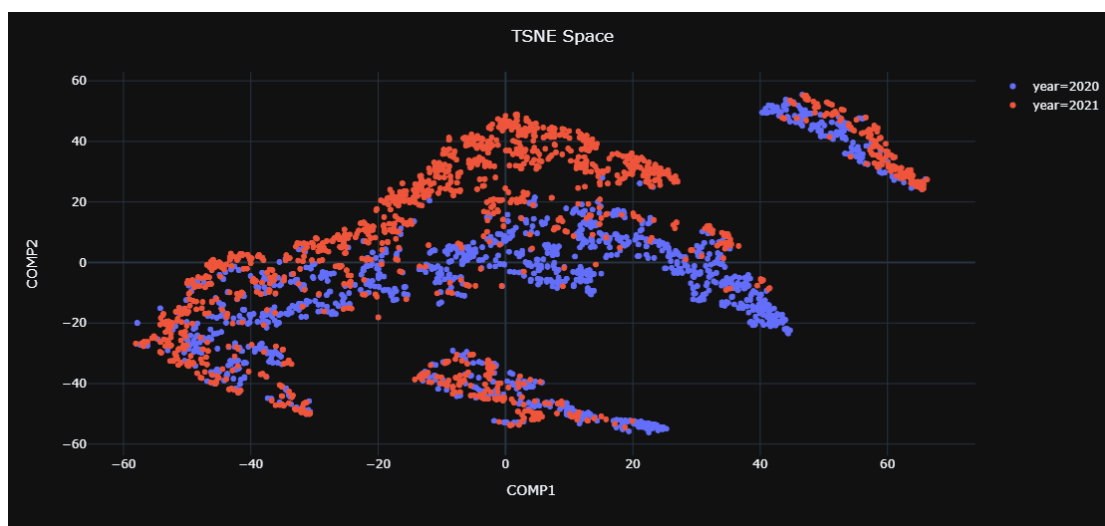
הישובים ראש עין, וניין שייכים למגזר הערבי והישוב תראבי אכאענישוב למגזר הבדואי. על פי התרשים נראה כי בישובים אלו גברה התמיכה במפלגות ימניות על חשבון המפלגות הרשימה המשותפת ורע"ם. ואכן אפשר לראות זאת בבירור בתרשימים הבאים שמציגים את אחוזי ההצבעה בישובים אלה:

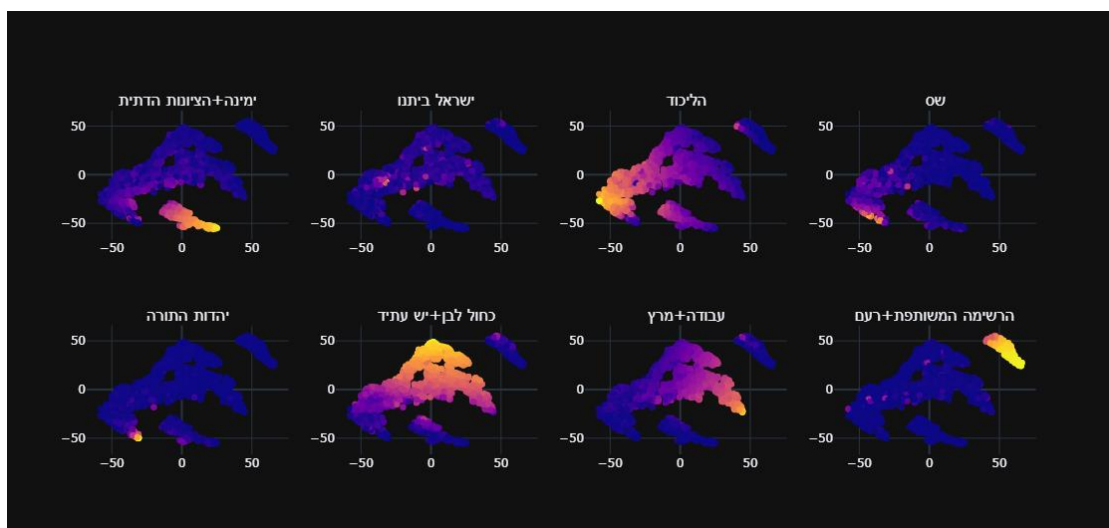


מעניין להבין מדוע חל שינוי דרמטי בדפוסי ההצבעה בישובים אלה. סיבה אפשרית היא שקמפיין "החיבוק" של הליכוד הצליח להשפיע על בוחרים (ראו כאן למשל).

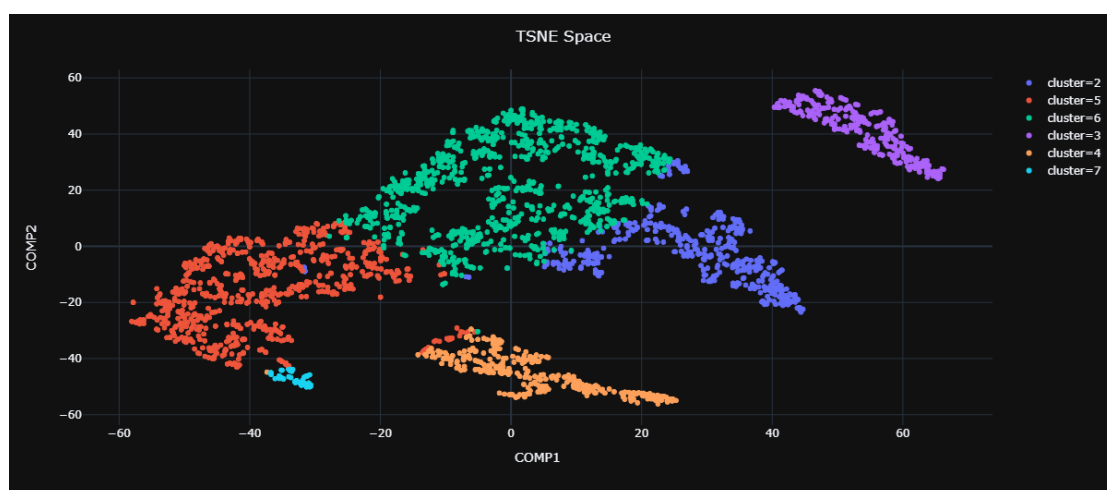
Further Analysis

אפשר לנסות שיטות להורדת מימד אחרות ובפרט שיטות לא לינאריות (manifold learning) כגון T-SNE, LLE וכו'. כל שיטה עשויה לחשוף צדדים אחרים של הדאטה ועשויה לספק תובנות חדשות. שימוש ב-TSNE למשל מוביל לאבחנה ברורה יותר בין המגזרים וכן חשיפה של מגזרים אחרים כגון מפד"ל:





ביצוע k-means clustering עם $k=6$:



אפשר לזהות:

