

תיאור המשימה:

המעבדה עוסקת בהצגה גלובלית של מפלגות וקלפיות באמצעות הורדת מימד בשיטת Principal Component Analysis (PCA).

1. א. הציגו תמונה של 30 **המפלגות** במישור הנפרש ע"י שני המרכיבים הראשיים המובילים (PCA) עבור **קובץ הקלפיות** בבחירות 2020 ע"י scatter plot. כלומר, כל מפלגה j ניתן לייצג ע"י וקטור x_j באורך 11179 המתאר את מספר הקולות שקיבלה המפלגה בכל קלפי, ולאחר טרנספורמצית PCA המפלגה תיוצג ע"י וקטור באורך 2, כנקודה במישור, בעזרת scatter-plot. הוסיפו טקסט עם שם המפלגה ליד כל נקודה. האם הקרבה בין מפלגות מצופה/מפתיעה? האם תוכלו לתת משמעות לשני הצירים?
ב. חזרו על התהליך אבל הפעם נרמלו קודם כל וקטור x_j כך שיהיה באורך 1 ועשו את טרנספורמצית PCA לוקטורים המנורמלים. האם הקרבה בין המפלגות והמשמעות של הצירים השתנו?
2. א. הציגו תמונה של **הקלפיות** במישור הנפרש ע"י שני המרכיבים הראשיים המובילים (PCA) עבור **קובץ הקלפיות** בבחירות 2020, לאחר נרמול. כלומר, כל קלפי ניתנת לייצג ע"י וקטור באורך 30 המכיל את שכיחויות ההצבעה לכל המפלגות הגדולות בקלפי. יש לנרמל כל וקטור כזה כך שיהיה באורך 1 ואז לבצע טרנספורמצית PCA שאחריה הקלפי תיוצג ע"י וקטור באורך 2, כנקודה במישור. הציגו כל קלפי בצבע/סמל המראה את האשכול החברתי-כלכלי אליו שייכת הקלפי (לשם כך יש לאחד את קובץ הקלפיות עם הנתונים החברתיים הכלכליים שהם על פי ישובים).
ב. חזרו על התהליך אבל הפעם צרו 8 גרפים שונים (4×2) כאשר בכל גרף צבעו את הקלפיות על פי מידת התמיכה במפלגה אחרת מבין 8 הגדולות. (כל גרף יראה כמו הגרף בסעיף א' פרט לצביעת הנקודות שתהיה שונה). תארו עבור 2 מפלגות את דפוסי התמיכה בהן על פי הגרף.
ג. חזרו על התהליך אבל הפעם בחרו 3 ערים עם לפחות 10 קלפיות והציגו רק את הקלפיות שלהן בנקודות, בצבע אחר עבור כל עיר. האם ישנן קלפיות הנראות חריגות ביחס לעיר?
3. א. נרמלו את 2 ה-PCs הראשונים שקיבלתם בשאלה 2 לקבלת וקטורים באורך 1. הציגו את 2 הוקטורים המנורמלים ב-bar-plot (עמודות ב-2 צבעים שונים) בו רואים את הגודל של הרכיב עבור כל מפלגה. אילו מפלגות הן דומיננטיות עבור 2 ה-PCs הראשונים?
ב. הציגו גרף של השונות המוסברת היחסית של הנתונים כפונקציה של מספר הרכיבים הראשיים, החל מ-1 ועד למספר המפלגות. לאיזה מימד הייתם ממליצים להוריד את הנתונים על פי גרף זה?

הערות:

- חשבו על עיצוב הגרפים. תנו כותרת לצירים, שימו לב לאורך הצירים.
- השתמשו בצבעים, עובי נקודה, וכו' כדי להדגיש נקודות חשובות.
- מותר לכם להיות יצירתיים; נסו לחשוב על שיטות אחרות לצביעת הקלפיות - למשל על פי ההצבעה לאחד הגושים, על פי מיקום גאוגרפי, ..