

FIFA VISUALIZATION

עילי עוזי 206204588

נועם שמיר 316139070

1. מבוא:

במשחק FIFA לכל שחקן יש דירוג כללי (Overall) ודירוג של כל היכולות האישיות שלו בסקאלה של 1-100.

נושא הפרויקט הוא הקשר בין דירוגי היכולות האישיות של השחקן לבין הדירוג הכללי שלו במשחק לאורך השנים 2017-2022. נרצה לנתח את יכולות אלה על מנת לבדוק כיצד משפיעות יכולות אלה על הדירוג הכולל שלו, בהתאם לעמדה במגרש.

אנו משתמשים במאגר הנתונים "FIFA 2023 OFFICIAL DATASET" מאתר Kaggle המכיל 6 טבלאות נתונים של השנים 2017-2022. כל אחת מהטבלאות מכילה כ-17,000 רשומות, כאשר כל רשומה מייצגת שחקן. בנוסף, כל אחת מהטבלאות מכילה כ-65 עמודות כשחלקן מייצגות נתונים יבשים (למשל, גיל, עמדה, קבוצה, לאום וכו') וחלקן מייצגות דירוגים של יכולותיו האישיות. המפתח הראשי של כל אחת מהטבלאות הוא ID של השחקן.

2. נתונים ומטלות:

2.1 נתונים:

Column name	datatype	Attribute type	Meaning	Cardinality
ID	Item	Categorical	Player ID	More than 17,000 different values
Name	Item	Categorical	Player Name	More than 16,000 different values
Age	Attribute	Quantitative	Age of Player	Values in range 16-54
Nationality	Item	Categorical	Nationality	168 different values
Overall	Attribute	Ordered-Ordinal (Int)	Overall rank	Values in range 0-100

Potential	Attribute	Ordered-Ordinal (Int)	Future Potential Rank	Values in range 0-100
Club	Item	Categorical	Players club	870 different values
Value	Attribute	Categorical – Nominal (String)	Market Value	254 different values
Wage	Attribute	Quantitative	Players' wage	134 different values
Preferred Foot	Attribute	Categorical – Nominal (String)	His strong foot	Right/Left
International reputation	Attribute	Ordered-Ordinal (Float)	International reputation	Values in range 0-5
Weak Foot	Attribute	Ordered-Ordinal (Float)	Rank of his weaker foot	Values in range 0-5
Skill Moves	Attribute	Ordered-Ordinal (Float)	Rank of skill moves	Values in range 0-5
Body Type	Attribute	Categorical – Nominal (String)	Body Type	4 different values
			(Stocky, Lean, Normal, Unique)	
Position	Attribute	Categorical – Nominal (String)	Position (GK, CB, ST, etc.)	16 Different Values
Height	Attribute	Quantitative	Height	Values in range 155-205 cm
Weight	Attribute	Quantitative	Weight	Values in range 60-105 kg
Columns 27-60	Attribute	Ordered-Ordinal (Float)	Contains all Skills attributes.	Values in range 0-100
			each column represents a different skill.	

2.2 מטלות:

שאלה מרכזית :

מה הקשר בין דירוגי היכולות האישיות של השחקן לבין הדירוג הכללי שלו במשחק לאורך השנים 2017-2022 ?

תתי שאלות :

1. לאילו יכולות יש את הקורצליה הגבוהה ביותר עם הדירוג הכללי של השחקן (Overall Rating)?
2. האם פוטנציאל השחקן בשנים מוקדמות מומש בשנים שלאחר מכן?
3. מי יותר טוב בפיפא , מסי או רונאלדו?
4. כיצד מתפלג לאום השחקן בקרב השחקנים הטובים ביותר בכל שנה?
5. מהי קבוצת החלומות?

2.3 עיבוד נתונים מקדים:

בכל שנה, חילקנו את סט הנתונים ל4 תתי סטים ע"פ העמדות השונות על המגרש. למשל עבור שנת 2022:

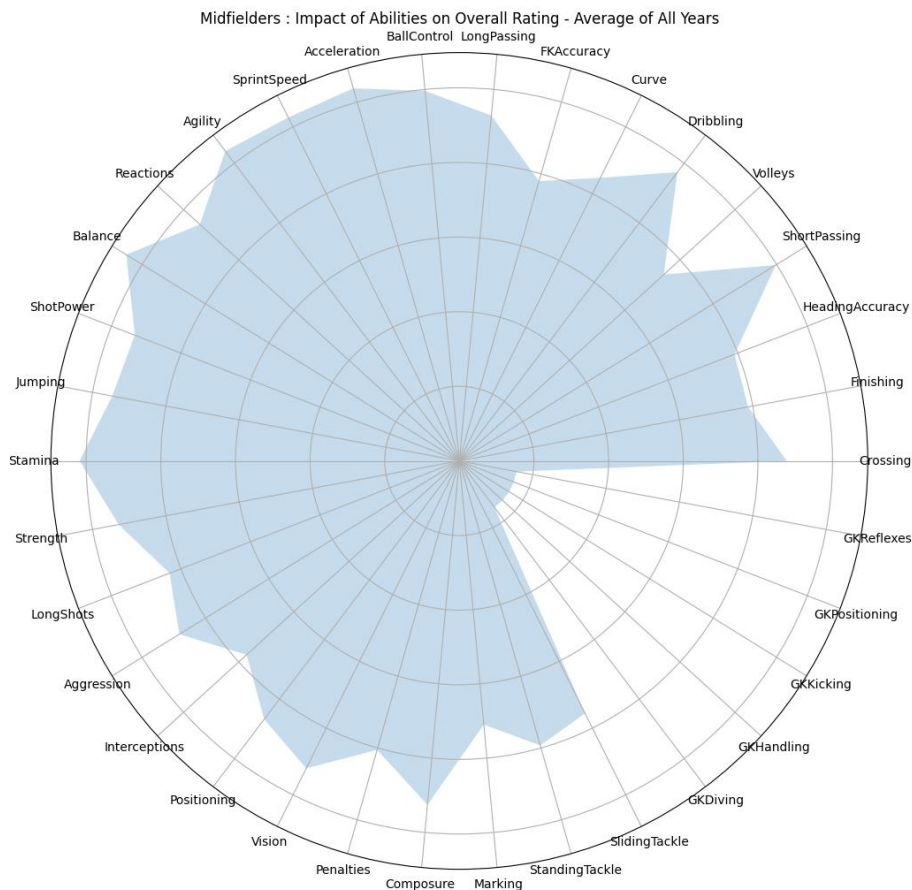
1. Fifa22_defender – סט הכולל רק את שחקני ההגנה.
2. Fifa22_midfielder – סט הכולל רק את שחקני האמצע.
3. Fifa22_forward – סט הכולל רק את שחקני ההתקפה הקדמיים.
4. Fifa22_goalkeeper – סט הכולל רק את השוערים.

עשינו זאת מתוך הנחה שעבור כל עמדה על המגרש, ה overall rating מחושב אחרת מכיוון ששחקן משחק בתפקיד מסוים בהתאם ליכולותיו. למשל, סביר להניח שהפיצ'ר sprintspeed יהיה בעל השפעה רבה יותר בחישוב ה overall rating עבור שחקן התקפה קדמי מאשר עבור שוער. לעומת זאת, יכולות התיקול (tackle) יהיו גבוהות יותר בקרב שחקני ההגנה מאשר שחקני ההתקפה.

3. הסבר על העיצוב שנבחר ועל יישומו

3.1 לאילו יכולות יש את הקורצליה הגבוהה ביותר עם הדירוג הכללי של השחקן (Overall)?

על מנת לענות על שאלה זו, תחילה נציג Radar charts אשר מייצגים את ההשפעה של כל פיצ'ר על ה- Overall. ככל שהפיצ'ר "מושך" אליו את תמונת המכ"מ, כך הוא בעל השפעה גדולה יותר על ה- Overall.



בתרשימים אלו (נספח א') קל לראות בצורה ויזואלית את הפיצ'רים השונים של השחקן והקשר שלהם לoverall rating- (עבור העמדות השונות על המגרש).

קל לראות, בתרשים הראשון למשל, ש- GK Kicking ו- Reactions הם בעלי השפעה גדולה יותר על ה- overall rating של שוער מאשר Dribbling – ו- SprintSpeed.

חסרון מובהק בתרשימים אלו הוא שאין לנו אינדיקציה לגבי גובה הקורלציה. כלומר, אנו רק מקבלים "תחושה" לאילו פיצ'רים יש את הקורלציה הגבוהה ביותר עם ה- Overall, אך אנו לא מקבלים ערך מספרי שמבטא את הקורלציה. (בהמשך הדו"ח נציג ערכים מספריים עבור הקורלציות בתרשים עמודות).

- Marks

ציר זוויתי (Angular) – מיוצג על ידי נקודה על היקף המעגל.

ציר רדיאלי (Radial) – מיוצג על ידי קו ריאלי רציף.

- Channels

ציר זוויתי (Angular) – כל ערך על היקף המעגל (tilt) מבטא יכולת מסוימת של השחקן.

ציר רדיאלי (Radial) – מיוצג על ידי קו רדיאלי רציף, כאשר ככל שמרחק ממרכז המעגל גדול יותר כך הקורלציה גבוהה יותר, אך חשוב לשים לב שמדובר בסולם לא מיושר (unaligned scale) שאינו מעיד על גובה המתאם.

- אקספרסיביות

יכולות השחקן הן ערכים מספרים מסוג quantitative אך בתרשים זה מכיוון שמדובר ברמת המתאם בסולם לא מיושר אל מול הדירוג הכללי ניכר שמדובר במagnitude channels. בנוסף, ככל שהערך רחוק יותר ממרכז המעגל ניתן להסיק על רמת מתאם גבוהה יותר באופן יחסי לתכונות אחרות ולכן מרמז על שוני בין התכונות.

- אפקטיביות

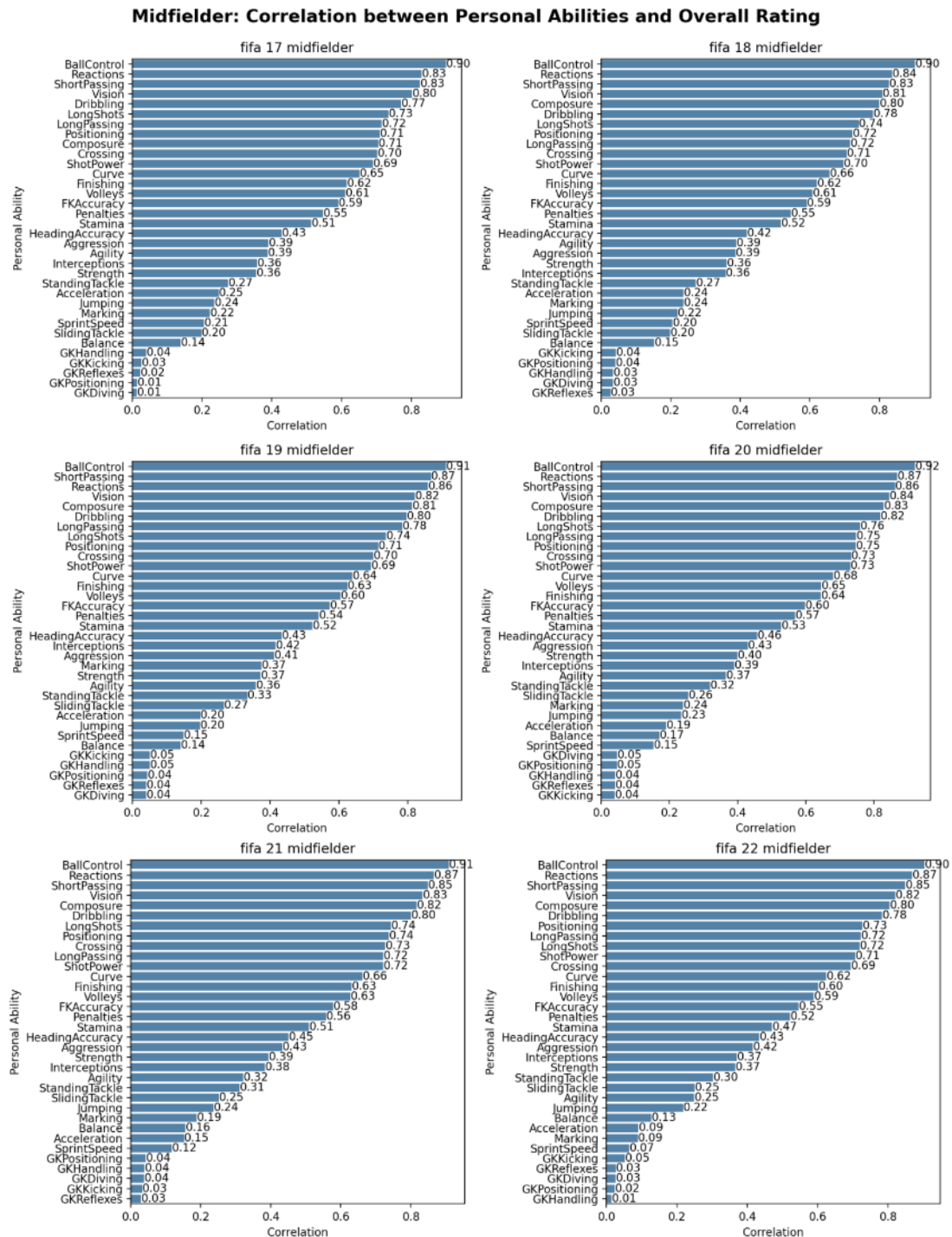
- Accuracy - ניתן להבחין בקלות בשוני בין רמת המתאם של היכולות אך לא נוכל להסיק לגבי ערך המתאם המדויק ולכן זהו חסרון של התרשים.
- Discriminability - כל היכולות ממוקמות על היקף המעגל, בהפרשים סבירים שמאפשרים הבחנה בין היכולות השונות.
- Separability – הציר הזוויתי והציר הרדיאלי אינם קשורים זה בזה ולכן קיימת הפרדה בין channels.
- Popout – באופן מיידי ניתן להבחין בין יכולות בעלות רמת גבוהה ליכולות בעלות רמת מתאם נמוכה, ולזהות את באיזו יכולת מדובר.
- Grouping – כל יכולת בלתי תלויה באחרת ולכן אין שיוך לקבוצות בתרשים זה. עם זאת, קרבה של מספר יכולות סמוכות עם מרחק גבוה/נמוך עלולה לגרום למחשבה שמוטעית שהינן קשורות זה בזה, אך זה המצב.

כעת, על מנת לקבל ערכים מספריים שמייצגים את הקורלציה בין כל פיצ'ר ל-overall וכן, על מנת לראות את השינויים לאורך השנים נציג correlation bar plot. הערכים נעים בין 0-1 כאשר ככל שהערך גדול והגוון שלו אדום יותר, כך הקורלציה בין הפיצ'ר ל-Overall גדולה יותר ולהפך.

3.2 - ניתוח תוצאות עבור השאלה לאילו יכולות יש את הקורצליה הגבוהה ביותר עם הדירוג

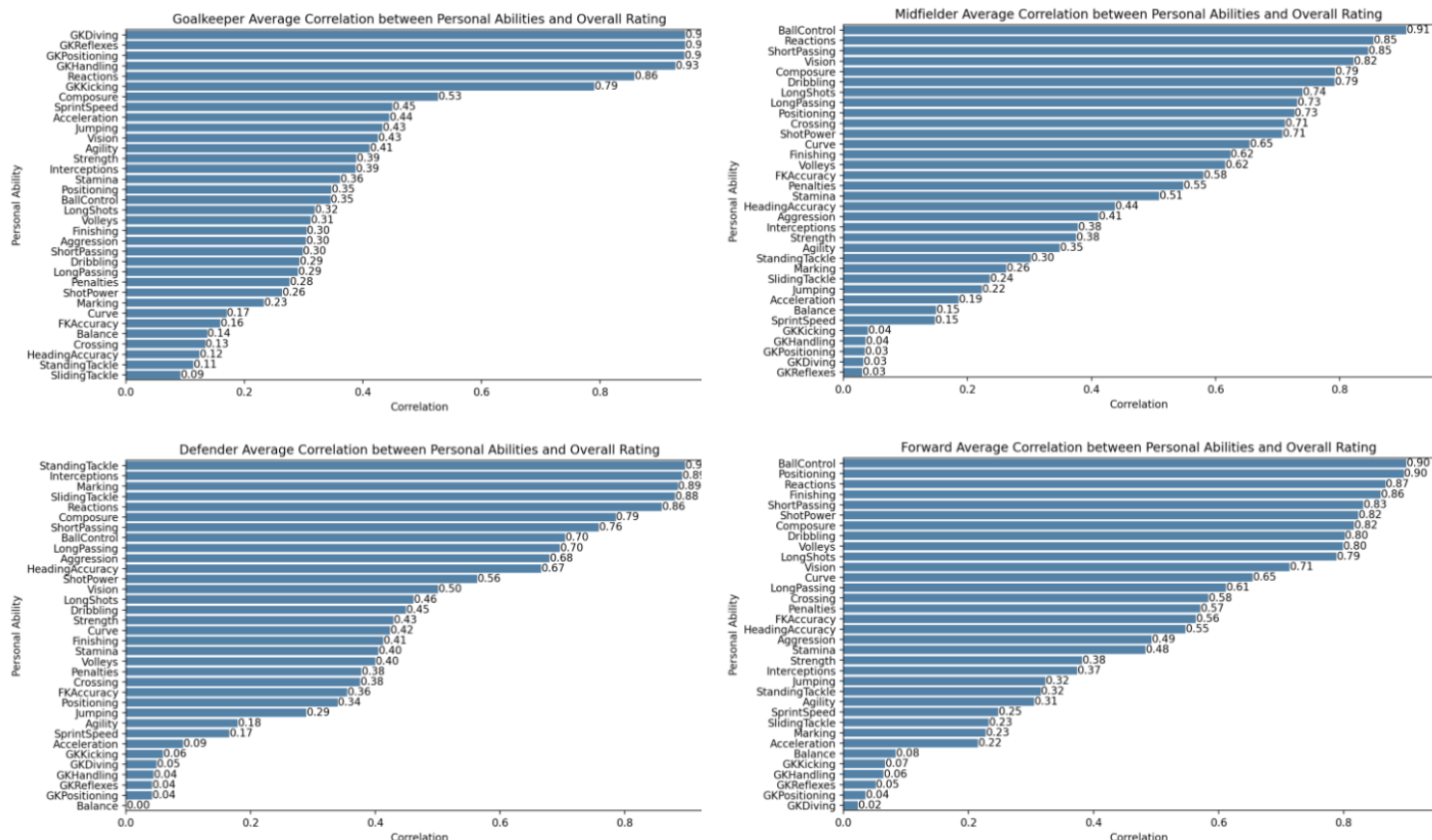
הכללי של השחקן (Overall Rating)?

רצינו לבחון האם חל שינוי לאורך השנים באופן החישוב של הדירוג הכללי של השחקן. כלומר, האם יכולות מסוימת השפיעו יותר בשנים מסוימות לעומת שנים אחרות. מכיוון שחלק גדול מהשחקנים מתאימים לשחק בעמדה מסוימת קיים שוני בין היכולות המזהות עם העמדות השונות, למשל, שחקני קישור יהיו לרוב יותר טכניים מאשר שחקני ההגנה. כדי לקבל תוצאות מדויקות יותר נחלק לפי עמדות: (נספח ב')



כדי לקבל תמונה כללית יותר ללא התמקדות בשנים מסוימות אלא בצורה משוקללת, נחשב את הממוצע של הקורצליות עבור כל עמדה לאורך השנים:

Average Correlation Over The Years Between Personal Abilities And Overall Rating



הצגת הקורלציה באמצעות קווים . כפי שניתן לראות, תצוגה זו מאוד נוחה ויזואלית וניתן לראות עבור כל תפקיד, בכל שנה, את הקורלציה בין הפיצ'רים ל-overall rating.

- Marks

הקווים

- Channels

המיקום האופקי של הקווים (מבטא את ערך המתאם)

אקספרסיביות –

יכולות השחקן הן ערכים מספרים מסוג Ordinal אך בתרשים זה ומכיוון שמדובר ברמת המתאם בסולם לא מיושר אל מול הדירוג הכללי ניכר שמדובר במagnitude channels . בנוסף, ככל שהקו ארוך יותר ניתן להסיק על רמת מתאם גבוהה יותר ובאמצעות הערך המספרי בסופו ניתן לראות מה הוא הערך המדויק.

אפקטיביות-

- Accuracy - ערך המתאם המתקבל מבטא בדיוק את ערך המתאם של היכולת.
- Discriminability - כל היכולות ממוקמות בציר y, בהפרשים סבירים שמאפשרים הבחנה בין היכולות השונות.
- Separability – הציר האנכי בלתי תלוי בציר האופקי ולכן קיימת הפרדה מוחלטת.
- Popout – באופן מיידי ניתן להבחין בין יכולות בעלות רמת מתאם גבוהה ליכולות בעלות רמת מתאם נמוכה יותר, וגם לזהות באיזו יכולת מדובר.
- Grouping – כל אחת מהיכולות בלתי תלויה באחרת ולכן אין שיוך לקבוצות בתרשים זה. מיון היכולות מהגדול לקטן מונע את השיוך לקבוצות.

ניתוח התוצאות:

קישור:

ע"פ תרשים המכ"ם ניתן לראות שהפיצ'רים הדומיננטיים עבור שחקן קישור הם: Acceleration, SprintSpeed, Agility, Reactions, Balance, ShotPower, Stamina, Vision, ShortPassing, LongPassing, Dribbling, BallControl.

לאחר התבוננות בתוצאות של ה correlation bar plot ניתן לראות שהפיצ'רים הדומיננטיים הם: Reactions, BallContrall, ShortPassing, Composure, LongPassing .
מעניין לראות ש- Coposure התגלה כיותר דומיננטי מ Dribbling לאורך השנים.

שוערים : (נספח ב')

ע"פ תרשים המכ"ם ניתן לראות שהפיצ'רים הדומיננטיים עבור שוער הם: GKReflexes, GKPositioning, GK Kicking , GKHandling , GK Diving , Strength , Jumping , Reactions .

לאחר התבוננות בתוצאות של ה correlation bar plot ניתן לראות שהפיצ'רים הדומיננטיים הם: GKReflexes, GkHandling, Reactions, GK Diving, GKPositioning . פיצ'רים אלו מקבלים ערכים גבוהים משמעותית לאורך כל השנים ביחס לפיצ'רים האחרים שהוזכרו לעיל.
בנוסף ניתן לראות שלאורך כל השנים, אלו 5 הפיצ'רים המובילים, עם ערכים גבוהים.

הגנה : (נספח ב')

ע"פ תרשים המכ"ם ניתן לראות שהפיצ'רים הדומיננטיים עבור מגן הם: SlidingTackle , StandingTackle,,Marking Aggression , Interceptions, Strength, Stamina, Jumping, SprintSpeed, Reactions .

לאחר התבוננות בתוצאות של ה correlation bar plot ניתן לראות שהפיצ'רים הדומיננטיים הם: StandingTackle, Interceptions, Marking, SlidingTackle, Reactions . פיצ'רים אלו מקבלים ערכים גבוהים משמעותית לאורך כל השנים ביחס לפיצ'רים האחרים שהוזכרו לעיל. בנוסף ניתן לראות שלאורך כל השנים אלו 5 הפיצ'רים המובילים, עם ערכים גבוהים.

התקפה : (נספח ב')

ע"פ תרשים המכ"ם ניתן לראות שהפיצ'רים הדומיננטיים עבור שחקן קישור הם: ShortPassing, Jumping, Acceleration, SprintSpeed, Agility, Balance, ShotPower, Strength, Positioning, Penalties, Coposure, Finishing, Dribbling .

לאחר התבוננות בתוצאות של ה correlation bar plot ניתן לראות שהפיצ'רים הדומיננטיים הם: BallControl, Positioning, Reactions, Finishing, ShortPassing .

מעניין לראות שבשנים מוקדמות יותר ShortPassing היה פחות דומיננטי.

3.3 - האם פוטנציאל השחקן בשנים מוקדמות מומש בשנים שלאחר מכן?

כידוע, לכל השחקן במשחק יש פוטנציאל מסוים שאליו יוכל להגיע במידה ויממש את היכולות שלו. לשחקנים צעירים יותר פוטנציאל גבוה יותר שכן בגילאים מוקדמים חל שיפור דרסטי יותר ביכולות שלהם.

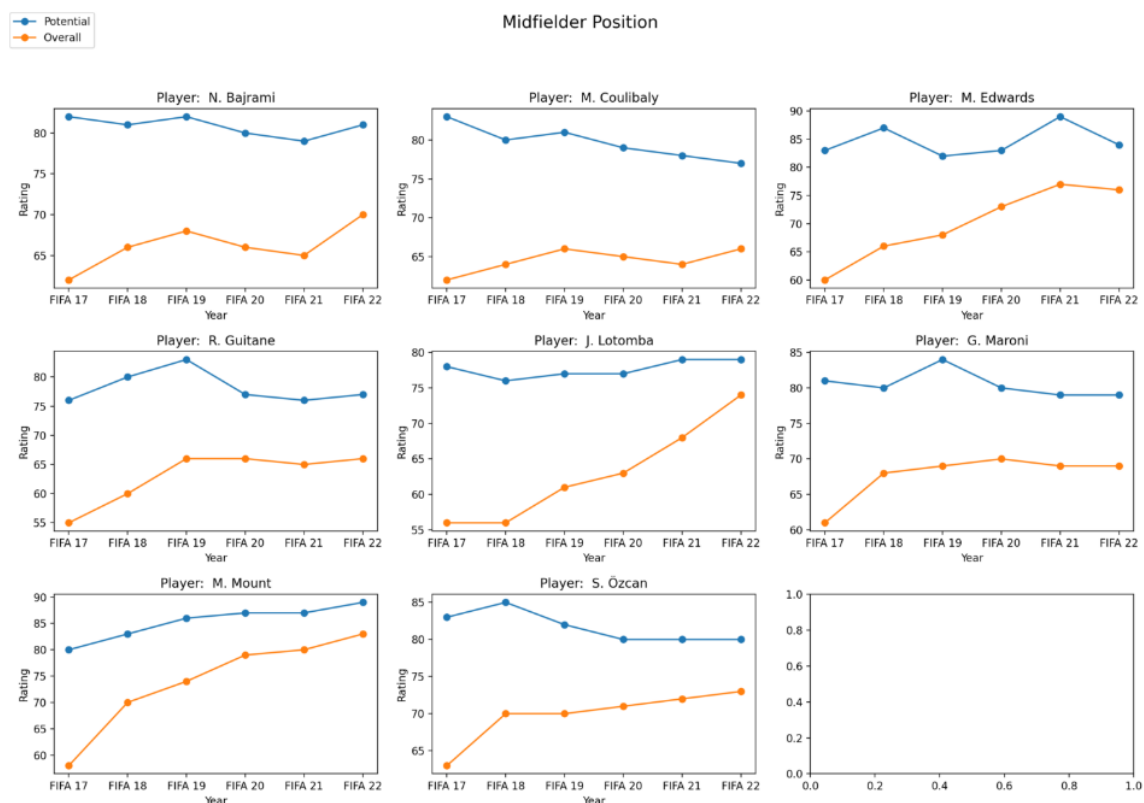
כדי להשיג מטרה זו, בכל אחד מהסטים הוספנו עמודה difference כאשר :

$$difference = Potential - Overall$$

עשינו זאת על מנת לאתר שחקנים שההפרש בין ה-Potential שלהם ל-Overall שלהם הוא הגדול ביותר. ביצענו סינון והשארנו רק שחקנים שה-Potential שלהם גדול מ-75. לאחר מכן השארנו בכל אחד מהסטים את 50 השחקנים אשר ערך ה-difference שלהם הוא הגבוה ביותר.

לבסוף בחנו את ערכי ה-Potential וה-Overall בשנים המאוחרות יותר על מנת לבדוק מי הם השחקנים שאכן מימשו את הפוטנציאל שלהם. נמיינ את השחקנים לפי שם המשפחה.

עשינו זאת עבור כל עמדה בנפרד (נספח ג'). להלן התוצאות עבור הקשרים:



בתרשימים אלו אנו יכולים לראות עבור כל עמדה על המגרש, עבור השחקנים הרלוונטים את מגמת השינוי לאורך השנים בדירוגי ה-Potential ו-Overall.

- Marks

הנקודות והקווים

Channels -

מיקום אופקי ואנכי של הנקודות. צבעים שונים עבור ה- Overall ועבור ה-Potential.

אקספרסיביות –

דירוג Overall וגם דירוג הפוטנציאל (Potential) הם ערכים מספרים מסוג Ordinal. מכיוון שמדובר בתרשים זה בהשוואת ערך ה-overall אל מול ערך ה-Potential שהוא Ordinal ניכר שמדובר במagnitude channels. בנוסף, ככל שהערך גבוה יותר ניתן להסיק על יכולות גבוהות יותר ואף להסיק לגבי הפרש בין ערכים בשנה מסוימת ולבחון מגמת שינוי לאורך השנים.

אפקטיביות-

- Accuracy - הנקודות מבטאות את הערך המספרי של הדירוג, אך לא ניתן להבחין בערך המדויק שלו כאשר הערך הספיציפי אינו מופיע בציר האנכי.
- Discriminability - לכל שחקן קיים plot נפרד ולכן אין חפיפה בגרפים של שחקנים שונים. בנוסף, כל שנה בציר האופקי נבדלת מאחרת והנקודות אינן חורגות מהטווח המוצג בציר האנכי.
- Separability – הציר האופקי (השנים), הציר האנכי (הדירוג) והצבעים בלתי תלויים אחד בשני וקיימת הפרדה מוחלטת ביניהם.
- Popout – באופן מיידי ניתן להבחין בהפרשים בין הדירוגים בשנה מסוימת ולהבחין במגמה מסוימת לאורך השנים (לדוגמה מגמת עלייה של ה-Overall וצמצום ההפרש לינו לבין דירוג הפוטנציאל.
- Grouping – כל שחקן מיוצג בנפרד ולכן בתרשים זה אין נטייה לשייך לקבוצות וגם אין צורך בכך.

ניתוח תוצאות:

עבור כל עמדה, נציג את השחקנים אשר מימשו את הפוטנציאל שלהם לאורך השנים (שחקנים אשר ה- Overall שלהם אכן גדל במשך השנים), ואת השחקנים אשר לא מימשו את הפוטנציאל שלהם לאורך השנים (שחקנים אשר ה- Overall שלהם גדל קצת/לא השתנה/ירד במשך השנים).

שוערים:

שחקנים אשר מימשו את הפוטנציאל שלהם:

- A. Ramsdale – כיום בן 25, משחק בארסנל. בשנת 2017 שיחק ב- Bournemouth
- J. Bijlow – כיום בן 25, משחק ב- Feyenoord בהולנד. בשנת 2017 שיחק ב- Feyenoord
- F. Muller – כיום, בן 25, משחק ב- VfB Stuttgart בגרמניה.

שחקנים אשר לא מימשו את הפוטנציאל שלהם:

- C. Fruchti – כיום בן 23, משחק ב- Austria Wien בגרמניה. בשנת 2017 שיחק ב- Bayern Munich.
- Y. Diouf – כיום בן 23, משחק ב- Reims בצרפת. בשנת 2017 שיחק ב- ESTAC Troyes בצרפת.
- A. Plizzari – כיום בן 23, משחק ב- Pescara באיטליה. בשנת 2017 שיחק ב- Milan.

הגנה:

שחקנים אשר מימשו את הפוטנציאל שלהם:

- C. Cuesta – כיום בן 24, משחק ב- Genk בבליה, שחקן נבחרת קולומביה. בשנת 2017 שיחק ב- Atletico Nacional בקולומביה.
- D. Upamecano – כיום בן 24, משחק ב- Bayern Munich בגרמניה, שחקן נבחרת צרפת. בשנת 2017 שיחק ב- RB Leipzig בגרמניה.
- A. Bastoni – כיום בן 24, משחק ב- Inter באיטליה, שחקן נבחרת איטליה. בשנת 2017 שיחק ב- Atalanta באיטליה.
- E. Konsa – כיום בן 25, משחק ב- Aston villa באנגליה. בשנת 2017 שיחק ב- Charlton Athletic באנגליה.

שחקנים אשר לא מימשו את הפוטנציאל שלהם:

- B. Riveros – כיום בן 25, משחק ב- Brondby בדנמרק, שחקן נבחרת פראגוואי. בשנת 2017 שיחק ב- Basel בשווייץ.
- T. Edwards – כיום בן 24, משחק ב- Barnsley באנגליה. בשנת 2017 שיחק ב- Stoke City באנגליה.
- R. Oxford – כיום בן 24, משחק ב- FC Augsburg בגרמניה. בשנת 2017 שיחק ב- Reading באנגליה.

קישור:

שחקנים אשר מימשו את הפוטנציאל שלהם:

- M. Mount – כיום בן 24, משחק ב- Chelsea, שחקן נבחרת אנגליה. בשנת 2017 שיחק גם כן ב- Chelsea.
 - J. Lotomba – כיום בן 24, משחק ב- Nice בצרפת, שחקן נבחרת שווייץ. בשנת 2017 שיחק ב- Lausanne-Sport בשווייץ.
- שחקנים אשר לא מימשו את הפוטנציאל שלהם:
- M. Coulibaly – כיום בן 24, משחק ב- Ternana באיטליה. בשנת 2017 שיחק ב- Pescara באיטליה.

- M. Bajrami – כיום בן 24, משחק ב- Sassuolo באיטליה. בשנת 2017 שיחק ב- Grasshoppers בשווייץ.

התקפה:

שחקנים אשר מימשו את הפוטנציאל שלהם:

- K. Mbappe – כיום בן 24, משחק ב- Paris Saint-Germain, שחקן נבחרת צרפת. בשנת 2017 שיחק גם כן ב- Paris Saint-Germain.
- E. Haland – כיום בן 22, משחק ב- Manchester City, שחקן נבחרת נורבגיה. בשנת 2017 שיחק ב- Molde בנורבגיה.
- M. Kean – כיום בן 23, משחק ב- Juventus, שחקן נבחרת איטליה. בשנת 2017 שיחק ב- Hellas Verona באיטליה.

שחקנים אשר לא מימשו את הפוטנציאל שלהם:

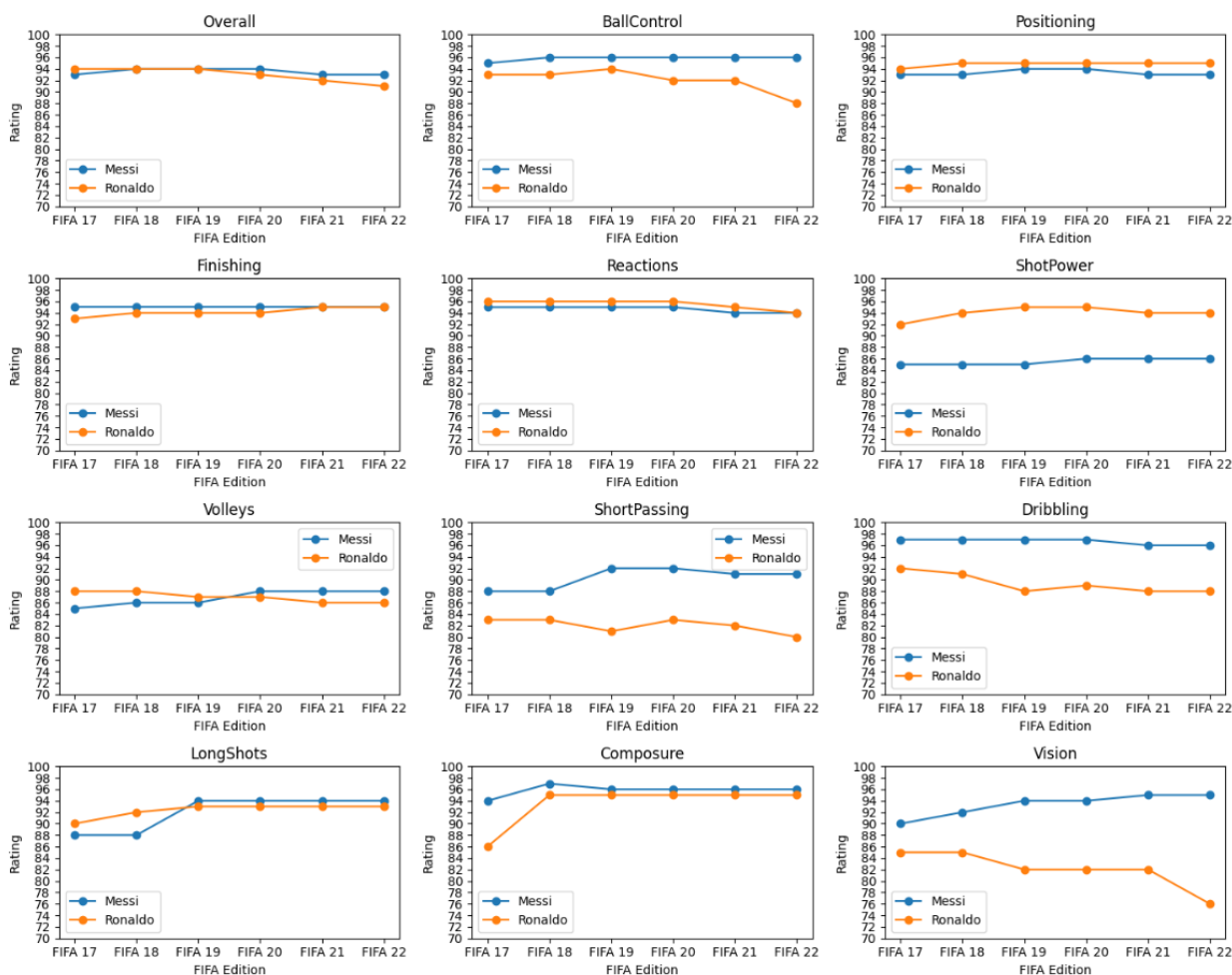
- Hang Kwang Song – כיום בן 24, לא שחקן פעיל. בעונת 2020-2021 שיחק ב- Al-Duhail בקטאר. בשנת 2017 שיחק ב- Cagliari באיטליה.
- R. Brewster – כיום בן 23, משחק ב- Sheffield United באנגליה. בשנת 2017 שיחק ב- Liverpool.

3.4 - מי יותר טוב בפיפא, מסי או רונאלדו?

בשנים האחרונות שאלה זו מעסיקה רבים בתחום הכדורגל ורצינו להשוות ביניהם במשחקים לאורך השנים, בתקווה למצוא התשובה לשאלה.

לצורך כך בחרנו את 10 הפיצ'רים עם הקורלציה הגבוהה ביותר עבור שחקני ההתקפה (אליה השניים משתייכים) והשוונו את הדירוגים של מסי ורונאלדו בפיצ'רים אלו לאורך השנים. בנוסף השוונו בין דירוג ה-Overall של שניהם לאורך השנים.

Lionel Messi vs Cristiano Ronaldo - Ratings Comparison



בתרשים זה ניתן לראות עבור כל פיצ'ר את הדירוג של מסי ורונאלדו, את מגמת ההתקדמות במהלך השנים של כל אחד מהשחקנים וכן ניתן להשוות בין שני השחקנים.

- Marks

הנקודות והקווים

- Channels

מיקום אופקי ואנכי של הנקודות. צבעים שונים עבור כל שחקן.

אקספרסיביות –

דירוג היכולות של השחקנים הם ערכים מספרים Ordinal. מכיוון שמדובר בתרשים זה בהשוואת יכולת סוימת בין שני השחקנים ניכר שמדובר ב magnitude channels . בנוסף, ככל שהערך גבוה יותר ניתן להסיק על יכולות גבוהות יותר ואף להסיק לגבי הפרש בין ערכים בשנה מסוימת ולבחון מגמת שינוי לאורך השנים.

אפקטיביות-

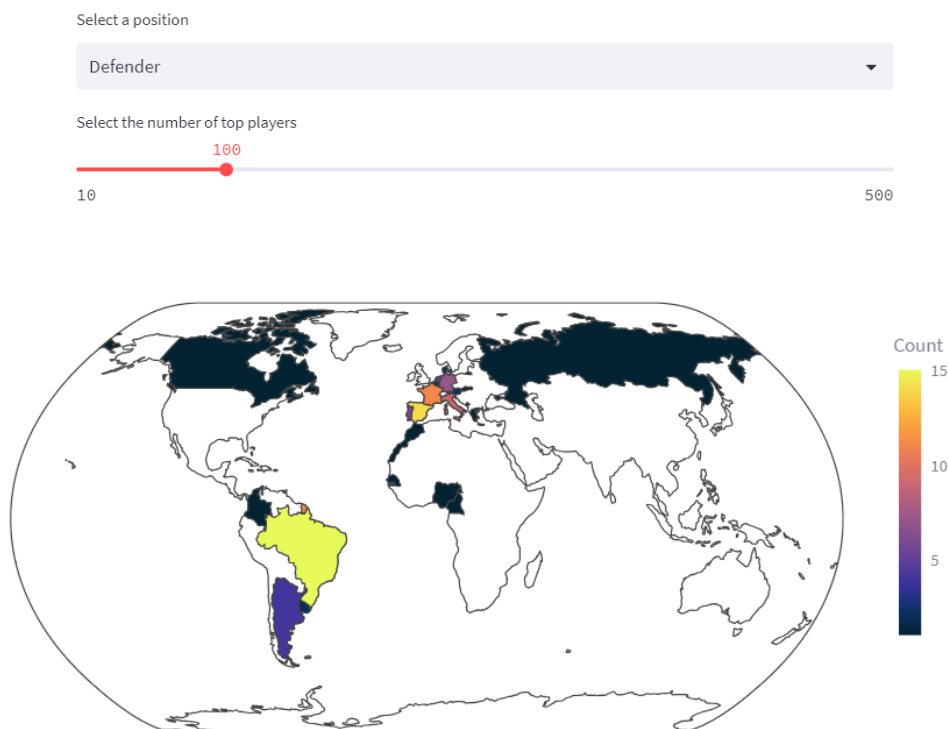
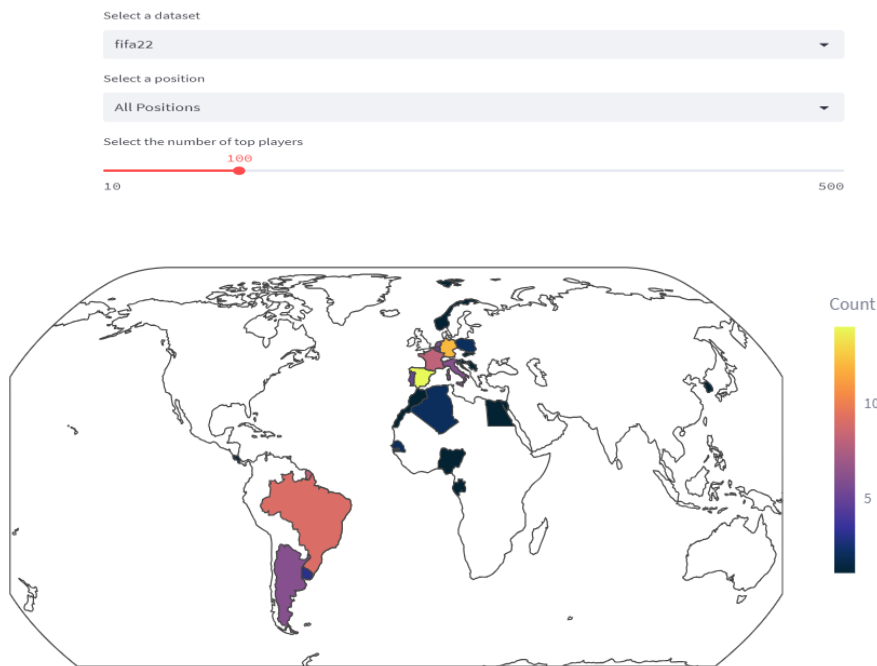
- Accuracy - הנקודות מבטאות את הערך המספרי של היכולת, אך קשה יותר להבחין בערך המדויק שלו כיוון שלא מופיע גל גבי הנקודה עצמה.
- Discriminability - לכל יכולת נבחנת קיים plot נפרד ולכן אין חפיפה בגרפים של יכולות שונות. בנוסף, כל שנה בציר האופקי נבדלת מאחרת והנקודות אינן חורגות מהטווח המוצג בציר האנכי.
- Separability – הציר האופקי (השנים), הציר האנכי (ערך מספרי של היכולת) והצבעים בלתי תלויים אחד בשני וקיימת הפרדה מוחלטת ביניהם.
- Popout – באופן מיידי ניתן להבחין בהפרשים בין היכולות בשנה מסוימת ולהבחין במגמה מסוימת לאורך השנים.
- Grouping – כל אחת מהיכולות היא בגרף נפרד ולכן אין נטייה לשייך לקבוצות את התכונות. הצבעים מבטאים את אותם שחקנים ולכן ניתן לייחס את אותו הצבע בין הגרפים השונים.

ניתוח תוצאות

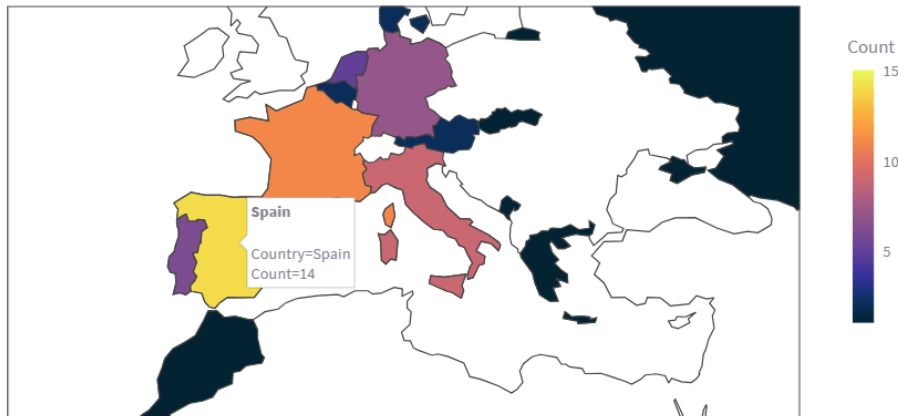
ניתן לראות כי ה- Overall של מסי לאורך השנים היה יותר גבוה. בנוסף ברוב הפיצ'רים מסי מדורג מעל רונאלדו לאורך השנים אך קיימים יתרונות מסוימים לרונאלדו על פני מסי בעיקר בהיבטים הפיזיים והעוצמות. באופן כללי, מסי בעל יכולות טכניות ושליטה בכדור גבוהות ונסיק מכך שמסי יותר טוב מרונאלדו ב FIFA . האם זה ככה גם במציאות? תלוי את מי שואלים...

3.5 – כיצד מתפלג לאום השחקן בקרב השחקנים הטובים ביותר בכל שנה?

נבחן עבור כל אחת מהשנים, מה מספר השחקנים מכל מדינת לאום בקרב השחקנים הטובים ביותר על פי דירוג ה-Overall של השחקן. מדובר בתרשים אינטרקטיבי שבו ניתן לבחור את השנה הרצויה ואת מספר השחקנים אותם נרצה לבחון. נציג מספר תוצאות אפשריות להמחשה בשנת 2022:



III. Zoom in והצגת המידע עבור המדינה הנבחרת:



– Marks

מפת העולם, מדינות והשחקנים (נתונים גאומטרים)

– Channels

מיקום המדינות על גבי המפה וצביעת המדינות בהתאם למספר השחקנים שלה.

– אקספרסיביות

מספר השחקנים שמייצגים כל מדינה הוא ערך Ordinal.

ומכיוון שמדובר בתרשים זה בהשוואת מספר השחקנים בין מדינות שונות ניכר שמדובר

in magnitude channels.

בנוסף, ככל שהערך גדול יותר ניתן להסיק שהמדינה דומיננטית יותר בעמדה הנבחרת ואף ולבחון

מגמת שינוי לאורך השנים, תוך שליטה עם מספר השחקנים שנבחרו להציג.

– אפקטיביות

- Accuracy - ערך הcount של כל מדינה מייצג באופן מדויק את מספר השחקנים המייצגים כל מדינה, אך ללא סימון המדינה קשה להבחין בערך המדויק. המדינות ניתנות לאבחנה ביתר קלות על ידי ציור של מפת העולם והגבולות הגאוגרפים שלה.
- Discriminability - אין מדינות שקיימות בנתונים ולא יופיעו במפת העולם, ולא קיימים שחקנים שאינם מייצגים מדינה מסוימת ולכן לא תתפספס אף פיסת מידע.
- Separability - המיקום הגאוגרפי של המדינה אינו תלוי במספר השחקנים המייצגים אותה במשחק ולכן קיימת הפרדה מוחלטת ביניהם.
- Popout - באמצעות הצבעים ניתן להבחין באופן מיידי במדינות בעלות מספר שחקנים רב יותר ואף ניתן להבחין בשינויים המתרחשים לאורך השנים ולא שינוי במספר השחקנים הנבדקים.

- Grouping – אמנם מספר השחקנים בכל מדינה בלתי תלוי במספר במדינה אחרת, אך צביעת מדינות מסוימות באותו צבע (אותו מספר שחקנים) עלול לגרום לחשוב שמדובר בקבוצה מסוימת אך מלבד היותם בעלי אותו מספר שחקנים אין משמעות לקבוצה.

ניתוח תוצאות:

בתרשים זה קיימות אינספור מגמות שנוכל לבחון מכיוון שמדובר ב-3 פרמטרים אותם ניתן לשנות. לצורך ההמחשה בחרנו להתמקד בבחינת השחקנים המדורגים ב-top 100 במדינות אירופה בעמדות השונות בשנת 2017 מול 2022 (נספח ד').

שוערים:

ניתן לראות שהשוערים הספרדים שולטים באירופה, כאשר אחריה עם מספר שוערים שסובב ה-10 נמצאות איטליה וצרפת. לעומת זאת גרמניה, אוקראינה, רוסיה וטורקיה במגמת ירידה במספר השוערים בטופ 100 בשנת 2022. המדינות הצבעים הכהים משתנות וכנראה פחות יציבות עקב מספר נמוך של שוערים בטופ 100.

הגנה:

גם בקרב שחקני ההגנה ניתן לראות שספרד שולטת באירופה (כ-15 שחקנים), אחריה נמצאות צרפת וגרמניה, אך זו ירדה בשנת 2022 כאשר איטליה תפסה את מקומה. גם כאן, שאר המדינות בעלות מספר שחקנים נמוך מ-5 פחות יציבות כשחלקן מופיעות וחלקן אינן מצופיעות בשנת 2022 לעומת 2017.

קישור:

מעניין לראות שכאן הפערים הולכים וגדלים כאשר ספרד שולטת ביד רמה עם כ-15 שחקנים ולעומת זאת שאר המדינות באירופה בעלות ייצוג נמוך מ-10 שחקנים ברשימת הטופ 100. בשנת 2017, פורטוגל במקום השני אך בשנת 2022 גרמניה עוקפת אותה וצרפת מיד לאחר מכן.

התקפה:

בשנת 2017 ספרד רק במקום השני כשמקום הראשון נמצא צרפת עם כ-15 שחקנים. ניתן להבחין במדינות חדשות שהתווספו בקטגוריה זו כמו למשל המדינות הסקנדינביות שהצליחו שחלקן הצליחו לשמור על ייצוג בשנת 2022 גם כן.

סיכום:

נראה שמדינות השולטות באירופה הן מדינות מרכז אירופה כאשר הייצוג המשמעותי הוא של נבחרות כמו ספרד, איטליה, צרפת וגרמניה שנחשבות בעיני רבים לליגות הטובות והאיכותיות בעולם, כאשר פורטוגל דולקת אחריהן. שחקני הקישור בעיקר מייצגים את מדינות המרכז כאשר בקטגוריות האחרות ובמספרים נמוכים יותר ניתן לראות מדינות מזרח אירופה ואזור הבלקן. תופעות מעניינות נוספות הן חוסר הייצוג של האי הבריטי שנחשבת בעיני מומחי כדורגל לליגה העשירה והטובה בעולם. בנוסף, הייצוג של מדינות צפון אפריקה שמעלות שאלות של תופעות הגירה

ופוליטיקה כמו התיישבות של צפון אפריקאים במדינות מרכז אירופה ומתן אזרחויות ליוצאי המדינות האלה או שמא מדובר בפיתוח והתפתחות של השחקנים על ידי המדינות הללו.

3.6 מהי קבוצת החלומות ?

רצינו לבדוק מהו הרכב החלומות (ההרכב הטוב ביותר) ע"פ מספר מערכי משחק שונים. להלן התוצאות עבור FIFA22 (שאר התוצאות בנספח ד')

FIFA22 Dream Team with Formation: 433



FIFA22 Dream Team with Formation: 4231



FIFA22 Dream Team with Formation: 442



– Marks

מלבן המייצג את מגרש הכדורגל והעמדות בהתאם למערך הנבחר.

– Channels

מיקום השחקן על המגרש וצבעי המגרש.

אקספרסיביות –

יכולת ה-Overall של השחקנים היא תכונה Quantitative אך בשיבוץ במערך המשחק מדובר בערך Ordinal מכיוון שאין משמעות לערך המספרי של היכולת אלא רק למי נמצא במקום הראשון והשני (תלוי במערך שנבחר). מכיוון שמדובר בתרשים זה בבחירת השחקנים הטובים לפי תכונת ordinal magnitude channels . ניכר שמדובר ב-

אפקטיביות-

- Accuracy - ה Overall של כל שחקן מייצג באופן מדויק יכולתו במשחק ולכן השחקן הטוב ביותר בכל עמדה שנבחרה יהיה הנבחרת.
- Discriminability - אין חפיפה בין העמדות וכל העמדות על המגרש מוצגות באופן ברור וגלוי. צבעי המגרש בולטים לעין ולא יוצרים בלבול.
- Separability – במערך שנבחר העמדות אינן תלויות אחת בשניה מכיון שכל עמדה בפני עצמה ונבחרה במיוחד כדי שלא ליצור חפיפה בין שני תפקידים (למשל RM ו-RW)
- Popout – ניתן להבחין באופן מיידי בשחקנים שנבחרו למערך הנבחר.
- Grouping – צביעת כלל ההרכב באותו צבע מסייעת להבנה שמדובר בקבוצה אחת ואינה מופרדת לתפקידים במבחינת הצביעה אלא רק מבחינת העמדות.

ניתוח תוצאות:

בעמדת השוער באופן קבוע נבחר Oblak שמשחק באתלטיקו מדריד ונעל פי המשחק מדובר בשוער הטוב בעולם.

באדר לחוליית ההגנה , צמד הבלמים Ramos ו Van Dijk נשאר ללא שינוי בשנים אלה כאשר בשנת 2022 המגנים של ליברפול Robertson ו Alexander Arnoldi תפסו את מקומם של Alba ו Carvajal מריאל מדריד וברצלונה בהתאמה .

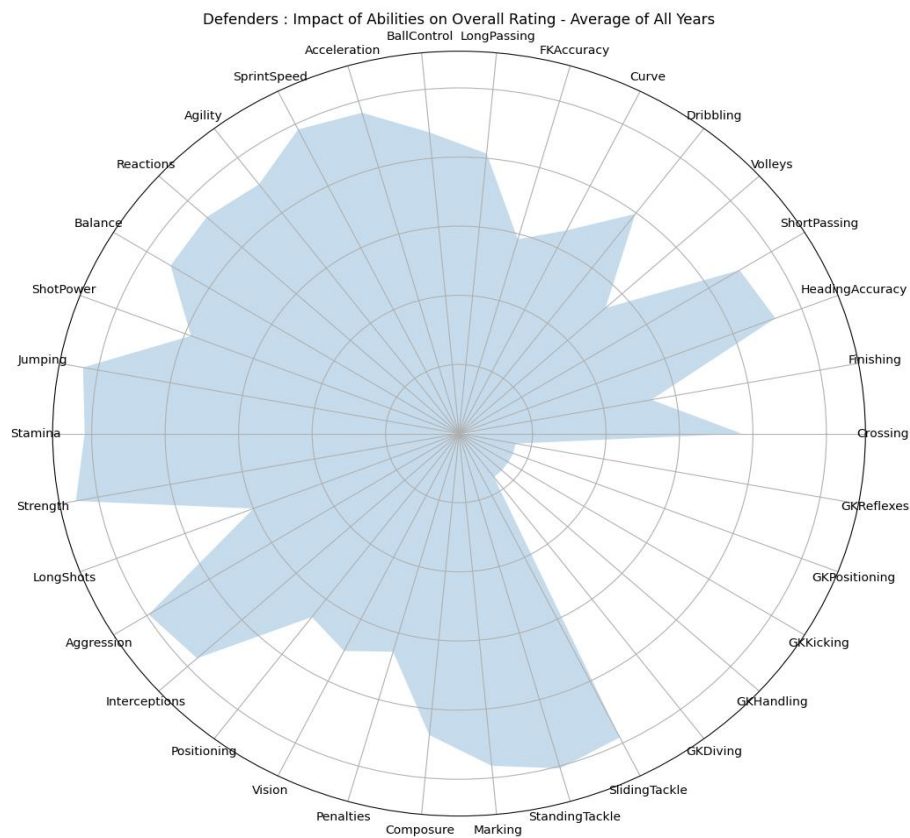
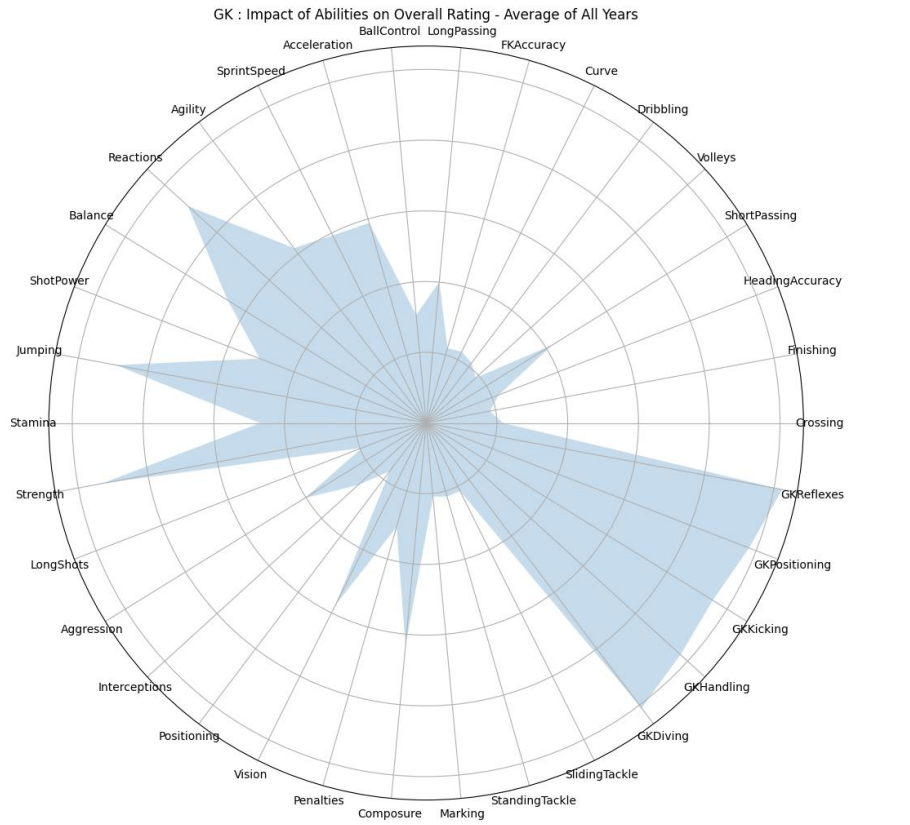
חוליית הקישור משתנה בין המערכים כאשר Kross ו Modric מריאל מדריד במרכז הקישור ובשנת 2022 De Bruyne מוסט לעמדה אחורית יותר משנים קודמות . באגפים Neymar ו Messi מוצבים ובחוד 3 החלוצים הטובים ביותר הם Benzema , Lewandowski ו C. Ronaldo .

תיאור הקוד בקצרה

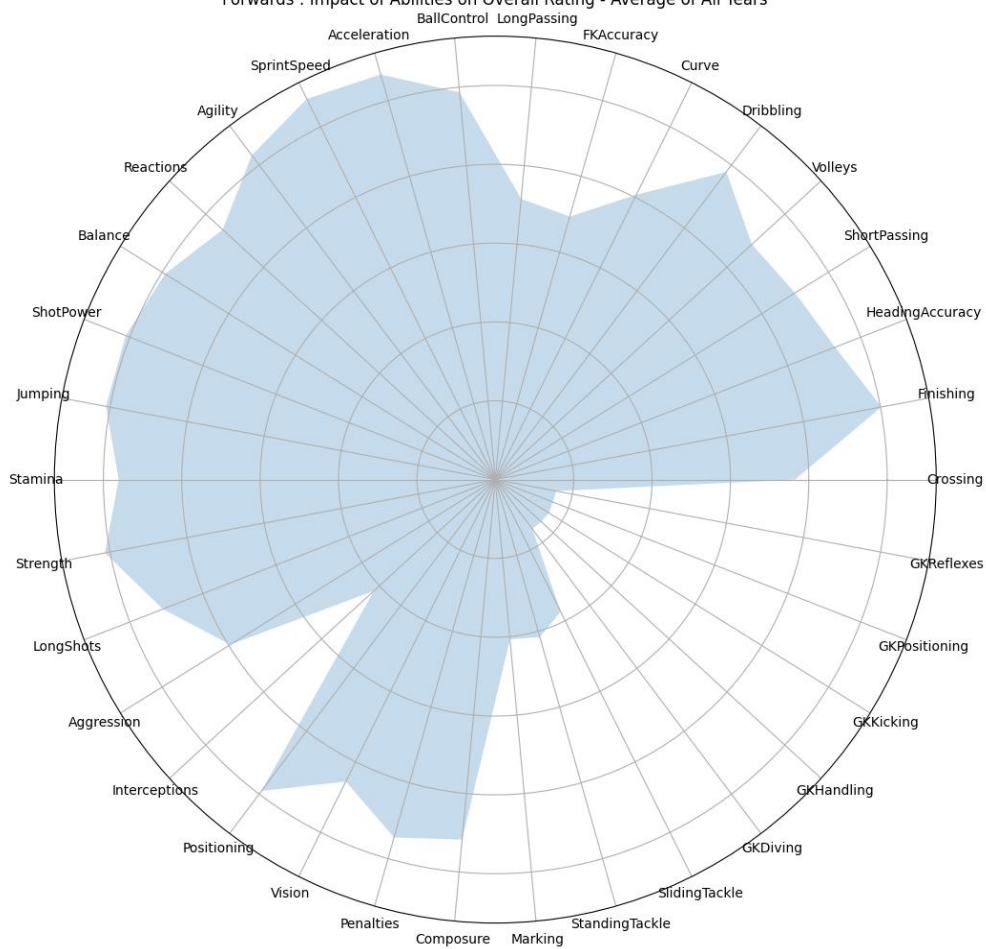
תחילה עבדנו במחברת colab. הורדנו את הטבלאות מ-kaggle וביצענו preprocessing על הנתונים. המטרה העיקרית בשלב זה הייתה ליצור טבלאות נתונים עבור כל עמדה וכל שנה (לדוגמא fifa17_defender, fifa17_midfielder, fifa17_forward, fifa17_goalkeeper). לאחר מכן הורדנו את הטבלאות האלה מהקולאב כקבצי csv וטענו אותן לפרוייקט ב-pycharm . ב-pycharm יצרנו את הויזואליזציה הסופית ע"י שימוש בספריית stremlit . עבור כל שאלה שמופיעה בתחילת הדו"ח יצרנו פונקצייה שמוציאה פלט כתשובה לשאלה. השתמשנו בספריות numpy, seaborn, streamlit, pandas, mplsoccer, matplotlib, plotly . השתמשנו בדוגמא קיימת שמצאנו באינטרנט עבור הויזואליזציה של "קבוצת החלומות" (שימוש בספריה mplsoccer על מנת להציג מגרש כדורגל), אך ביצענו התאמות כמו מציאת הקוארדינטות המתאימות לשיבוץ השחקנים בעמדות על המגרש בהתאם למערך שנבחר.

נספחים

נספח א'

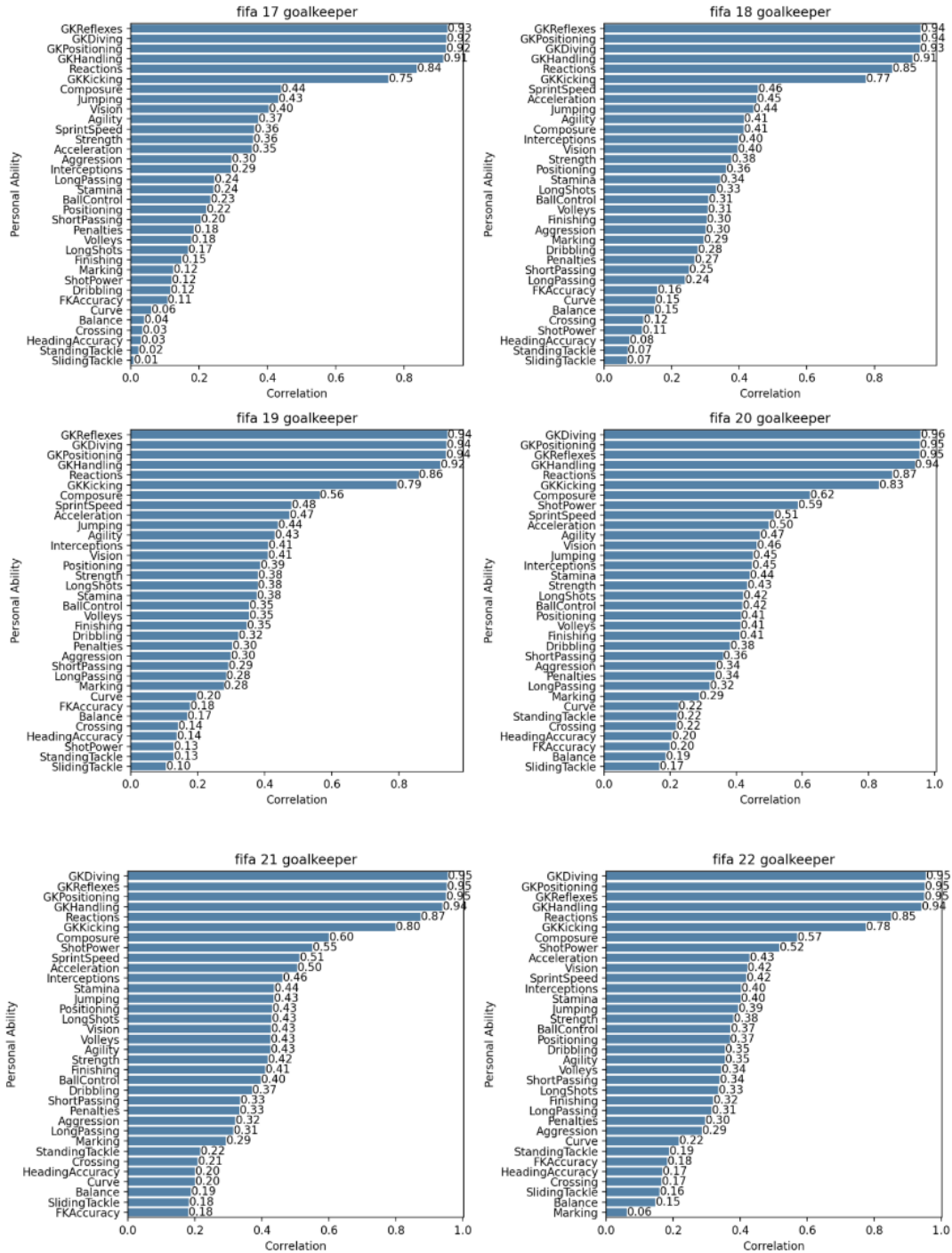


Forwards : Impact of Abilities on Overall Rating - Average of All Years

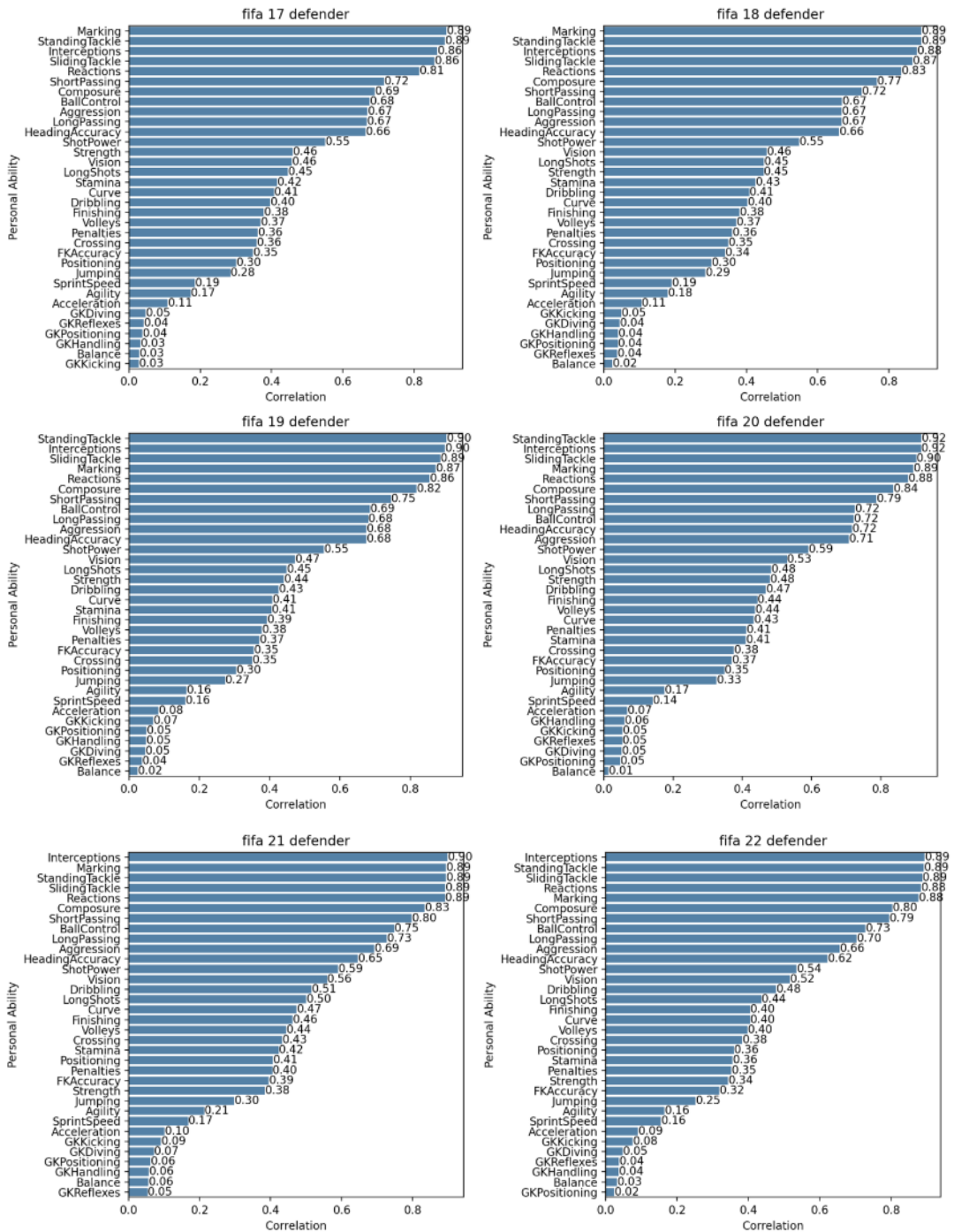


נפח 1

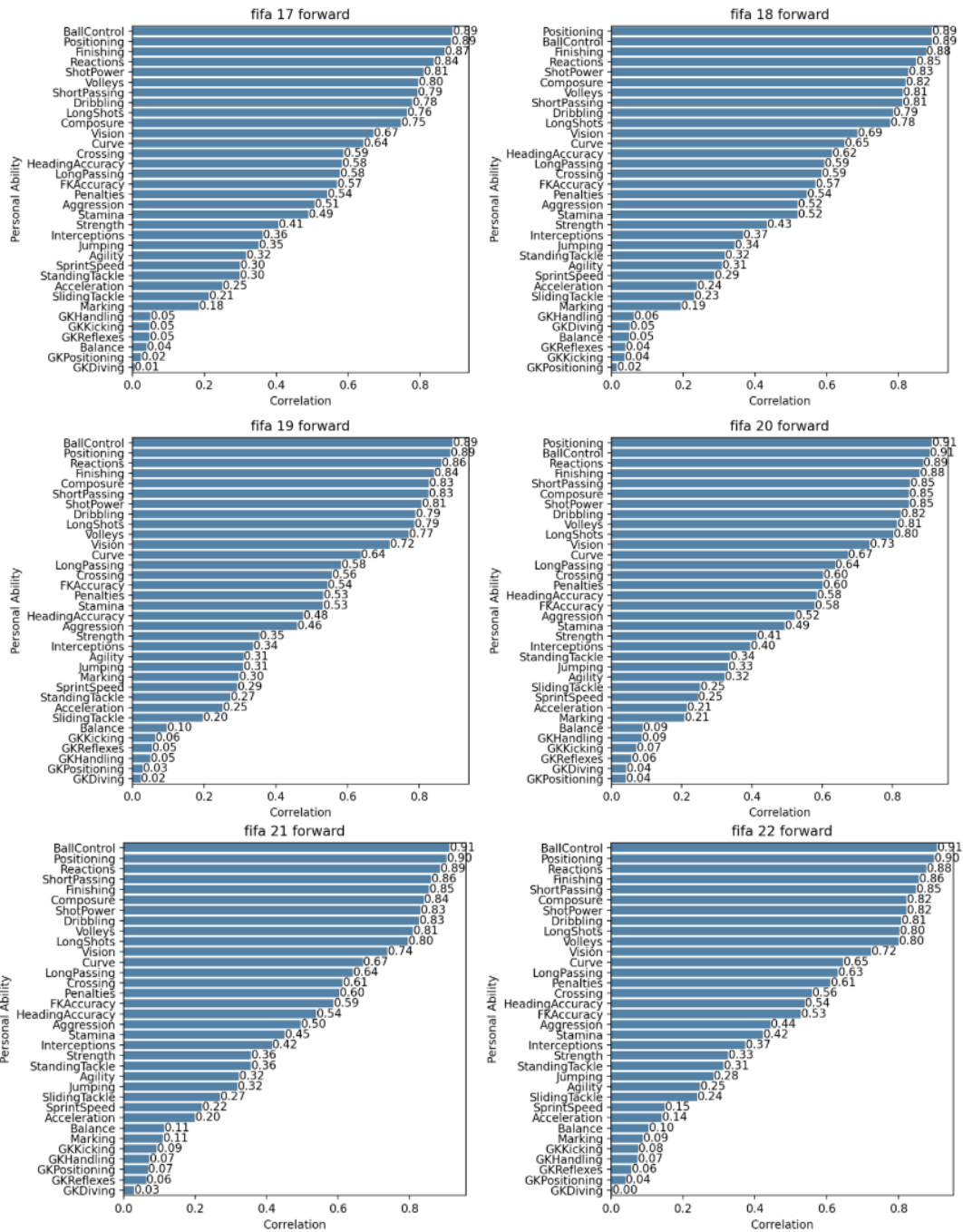
Goalkeeper: Correlation between Personal Abilities and Overall Rating



Defender: Correlation between Personal Abilities and Overall Rating



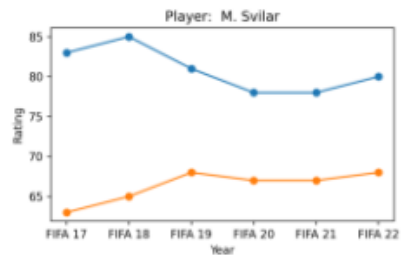
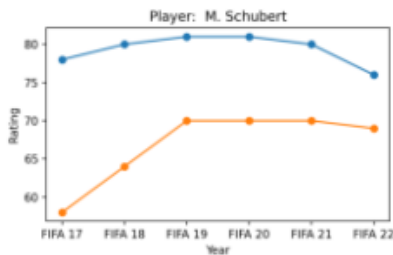
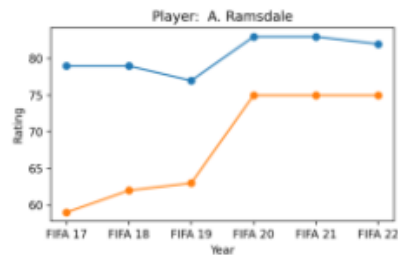
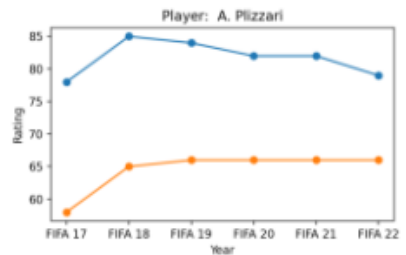
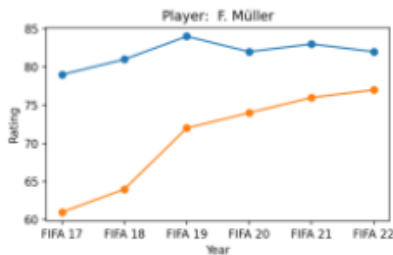
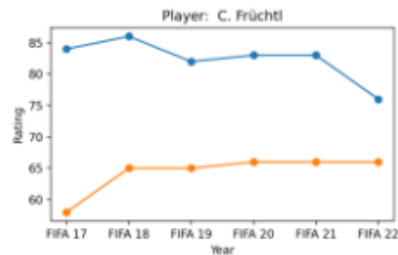
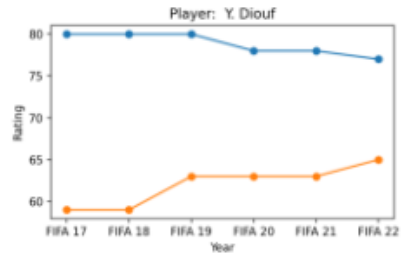
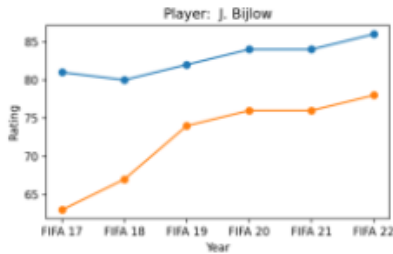
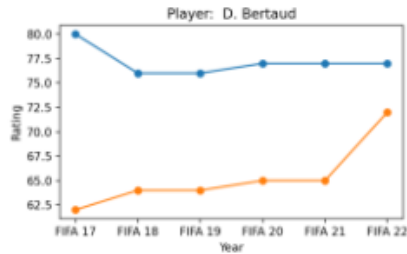
Forward: Correlation between Personal Abilities and Overall Rating



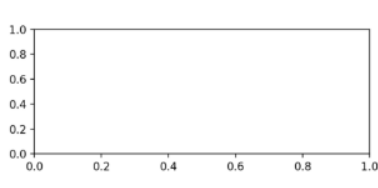
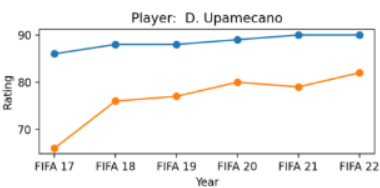
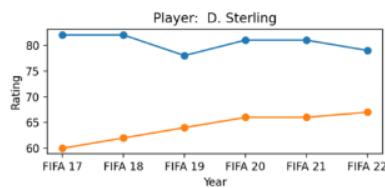
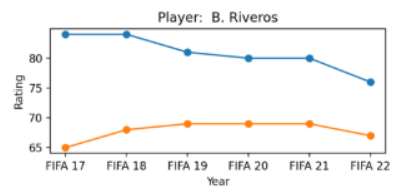
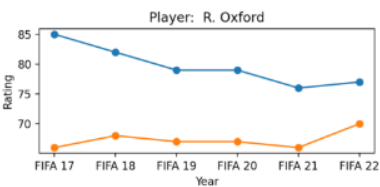
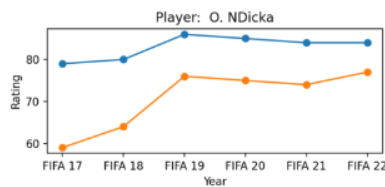
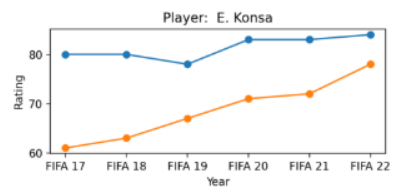
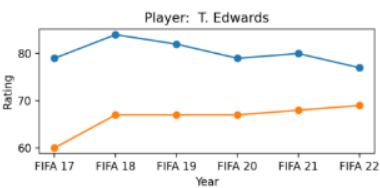
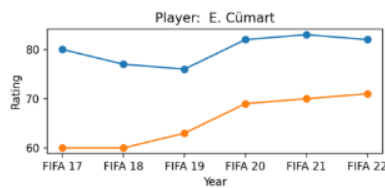
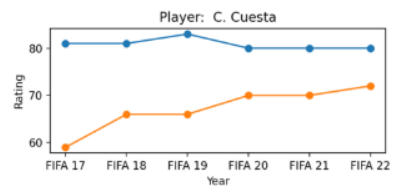
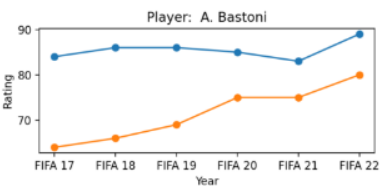
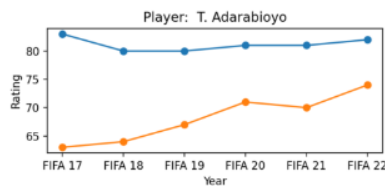
λ ηφ01



Goalkeeper Position

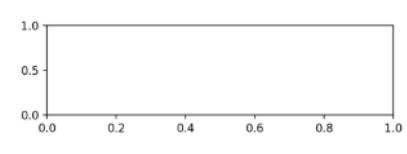
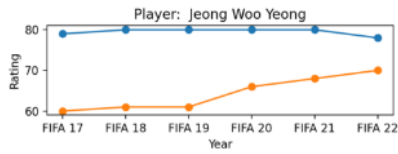
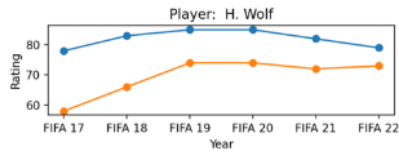
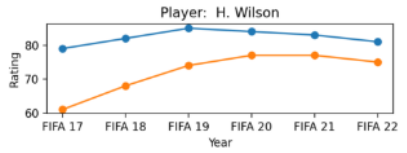
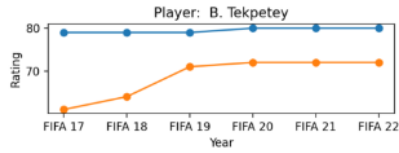
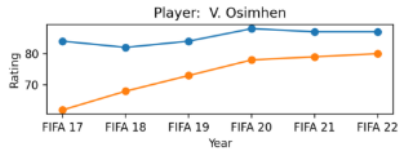
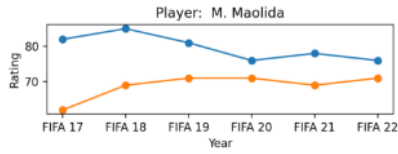
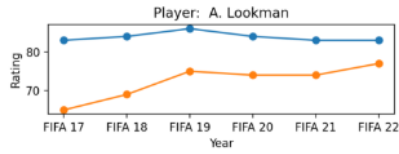
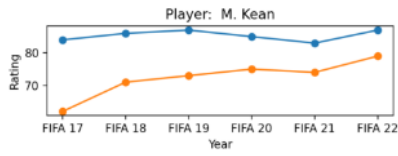
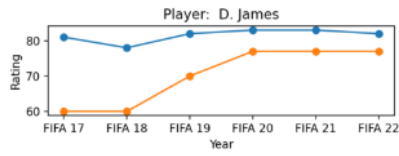
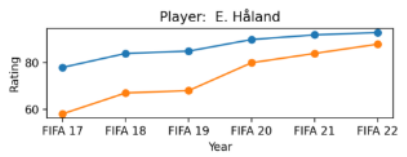
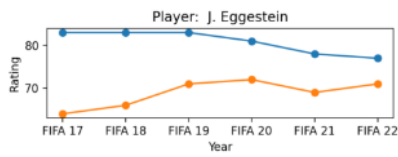
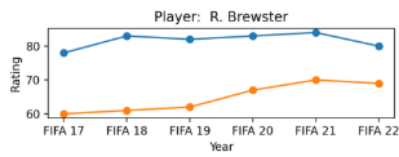


Defender Position





Forward Position



T nƧ01

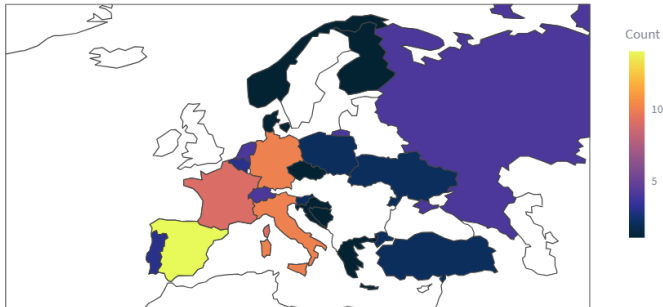
Select a dataset

fifa17

Select a position

Goalkeeper

Select the number of top players



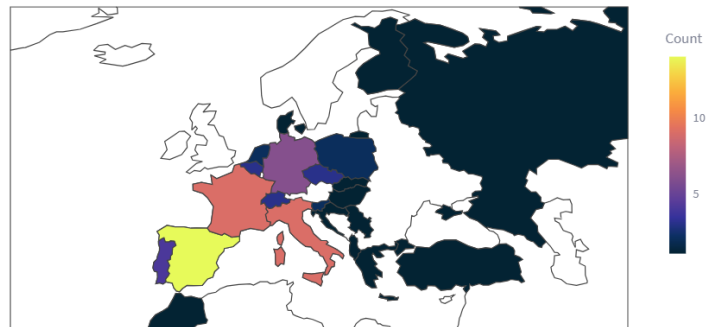
Select a dataset

fifa22

Select a position

Goalkeeper

Select the number of top players



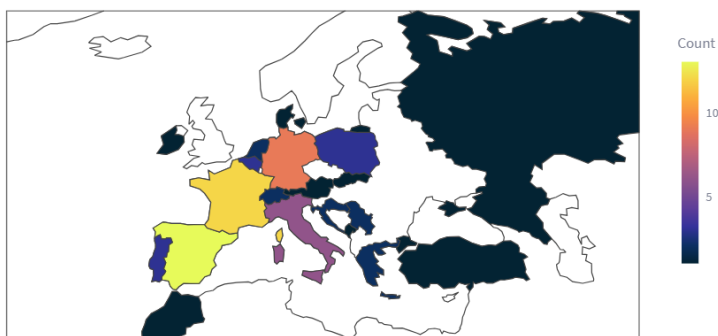
Select a dataset

fifa17

Select a position

Defender

Select the number of top players



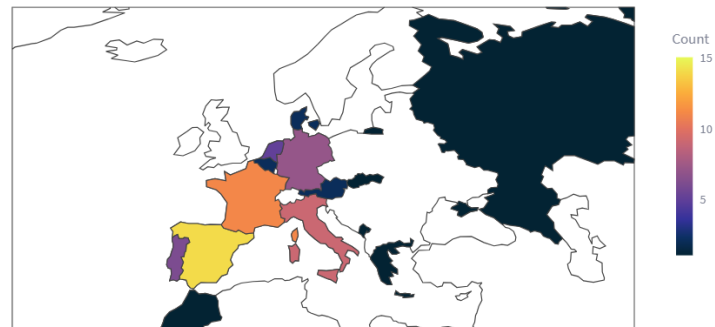
Select a dataset

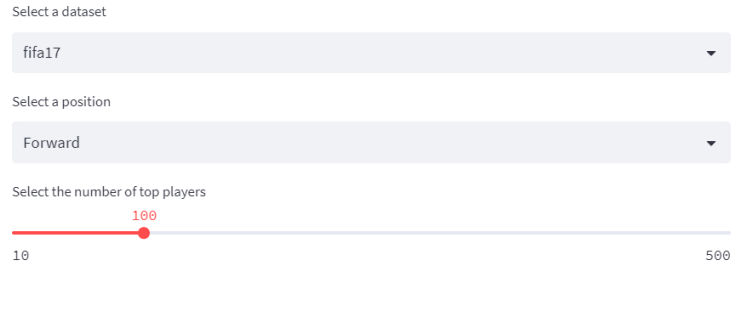
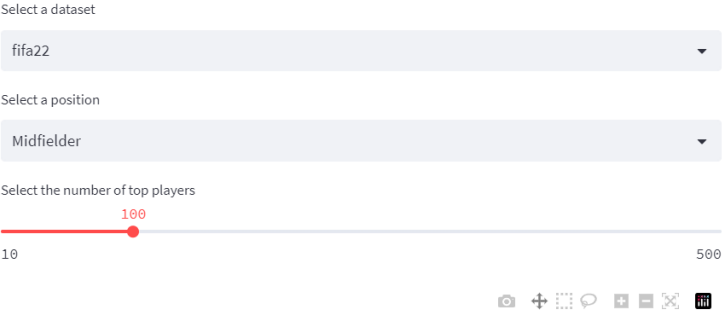
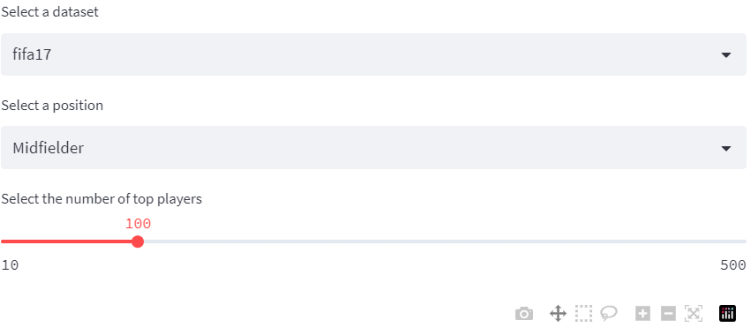
fifa22

Select a position

Defender

Select the number of top players





נספח ה'

FIFA21 Dream Team with Formation: 433



FIFA21 Dream Team with Formation: 4231



FIFA21 Dream Team with Formation: 442



FIFA20 Dream Team with Formation: 433



FIFA20 Dream Team with Formation: 4231



FIFA20 Dream Team with Formation: 442

