Understanding Pokémon through Networks and Clustering

By: Wong Li Sum

Introduction

This project aims to explore the type dynamics and character profiles within the Pokémon world. Each Pokémon type has specific advantages and disadvantages in the battle, making type interactions a crucial factor in competitive gameplay. Additionally, various attributes such as biometric attributes, battle stats and non-combat attributes of the Pokémon contribute to each Pokémon's unique characteristics, influencing their strategic roles in the game.

This project involves two major analyses using network analysis and clustering:

1. **Network Analysis on Pokémon Type Interactions**:
    - **Objective**: To analyse interactions between different Pokémon types based on their attacking and defending capabilities, and to identify influential types in the attacking and defending scenarios.
    - **Method**:
        1. Construct network graphs for attack and defense interactions among primary Pokémon types.
        2. Calculate PageRank and topic-specific (Personalized) PageRank, to measure the influence of each Pokémon type, highlighting which types are the most influential among the attacking and defending scenarios.

2. **Clustering Pokémon Based on Characteristics**:
    - **Objective**: To discover natural groupings among Pokémon based on their attributes and to interpret these clusters in terms of type, battle strength, catchability and other features.
    - **Method:**
        1. Standardize the features, then use Singular Value Decomposition (SVD) to reduce the dimensionality of the feature space.
        2. Use clustering algorithms such as K-means and agglomerative hierarchical clustering and evaluate the models using the Silhouette score to compare clustering quality.
        3. Interpret the clusters based on the features to understand the natural groupings and their implications.

These analyses provide insights into the effectiveness of Pokémon types and the nature of their characteristics, enhancing our understanding of battle strategy and character roles within the Pokémon world.

Dataset

The complete Pokémon dataset is obtained from Kaggle, which contains comprehensive information on 802 Pokémon, including base stats, performance against other types, biometric data, etc. The information was scraped from Serebii.net and any missing data was filled manually based on information from the same source.

Relevant data fields in the dataset for this project includes:

1. **pokedex_number:** Entry number of the Pokémon in the National Pokédex, indicating the unique identifier for a Pokémon.
2. **name:** The English name of the Pokémon.
3. **type1:** Primary type of the Pokémon.
4. **type2:** Secondary type of the Pokémon, if the Pokémon is dual-type.
5. **against_?:** Eighteen fields denoting the damage multiplier the Pokémon takes when attacked by the particular type. These values indicate type advantages and disadvantages.
6. **height_m:** Height of the Pokémon, in metres.
7. **weight_kg:** Weight of the Pokémon, in kilograms.
8. **capture_rate:** Chance of catching the Pokémon, higher values indicate easier capture.
9. **base_egg_steps:** Number of steps required to hatch the Pokémon's egg, indicating the breeding effort needed.
10. **experience_growth:** Rate at which the Pokémon gains experience, impacting its levelling speed.
11. **base_happiness:** Pokémon's initial happiness, which can influence evolution, as well as the power of moves like Return (high happiness) and Frustration (low happiness).
12. **hp:** Base hit points of the Pokémon, determining how much damage it can withstand before fainting.
13. **attack:** Base physical attack strength of the Pokémon.
14. **defense:** Base physical defense of the Pokémon, determining resistance to physical attacks.
15. **sp_attack:** Base special attack strength of the Pokémon.
16. **sp_defense:** Base special defense of the Pokémon, determining resistance to special attacks.
17. **speed:** Base speed of the Pokémon, influencing move order in battle.
18. **is_legendary:** Denotes if the Pokémon is legendary, where legendary Pokémon are rare and generally more powerful.

Analysis on core features

Core features include biometric attributes such as height and weight, non-combat attributes such as capture rate, base egg steps, experience growth and base happiness and battle stats such as HP, attack, defense, special attack, special defense and speed.

Table 1: Descriptive statistics for biometric ad non-combat attributes.

	height_m	weight_kg	capture_ rate	base_egg_ steps	experience_ growth	base_ happiness
mean	1.16	60.94	98.96	7191.01	1054996.00	65.36
std	1.07	108.51	76.41	6558.22	160255.80	19.60
min	0.1	0.1	3	1280	600000.00	0
25%	0.6	9	45	5120	1000000.00	70
50%	1	27.3	60	5120	1000000.00	70
75%	1.5	63	170	6400	1059860.00	70
max	14.5	999.9	255	30720	1640000.00	140

Table 1 shows significant variability in height and weight, where the average height and weight of a Pokémon are 1.16m and 60.94 kg, respectively. However, the height can range from 0.1m to 14.5m and the weight can range from 0.1kg to 999.9kg, suggesting a broad diversity among Pokémon species and across evolution. The high mean for capture rate indicates that most Pokémon are of relatively easier to catch, as the average is influenced by the common Pokémon that are the majority. Conversely, the high base egg steps and experience growth imply that rarer Pokémon often demand more time and effort to breed and train.

Table 2: Descriptive statistics for battle stats.

	hp	attack	defense	sp_attack	sp_defense	speed
mean	68.96	77.86	73.01	71.31	70.91	66.33
std	26.58	32.16	30.77	32.35	27.94	28.91
min	1	5	5	10	20	5
25%	50	55	50	45	50	45
50%	65	75	70	65	66	65
75%	80	100	90	91	90	85
max	255	185	230	194	230	180

Table 2 shows that most Pokémon generally possess relatively balance battle stats, with the majority of the Pokémon have the battle stats lies within 1 standard deviation from the mean. However, the presence of extreme values suggests that there are Pokémon that are either exceptionally weak or exceptionally strong in specific combat abilities.

The correlation heatmap in Figure 1 shows that height and weight are strongly correlated, aligning with the expectation that larger Pokémon tend to weigh more. Height and weight also show a moderately positive relationship with hit points (HP), indicating larger Pokémon tend to have higher HP, allowing them to withstand more damage during battles. There is a positive correlation between attack and defense stats, suggesting that Pokémon with stronger attack capabilities also tend to have better defense. This indicates the tendency for Pokémon to evolve in a balanced manner, rather than specializing solely as attackers or defenders. Notably, the capture rate has a negative correlation with base stats such as HP, attack, and defense, implying that Pokémon that are harder to catch typically possess stronger combat abilities, highlighting the rarity and power of such Pokémon. However, base egg steps and base happiness shows moderately negative correlation, suggesting that Pokémon which require more breeding effort may have lower initial happiness, which may be the case for rare Pokémon such that more time and effort need to be invested to improve Pokémon's happiness to fully realize their potential.

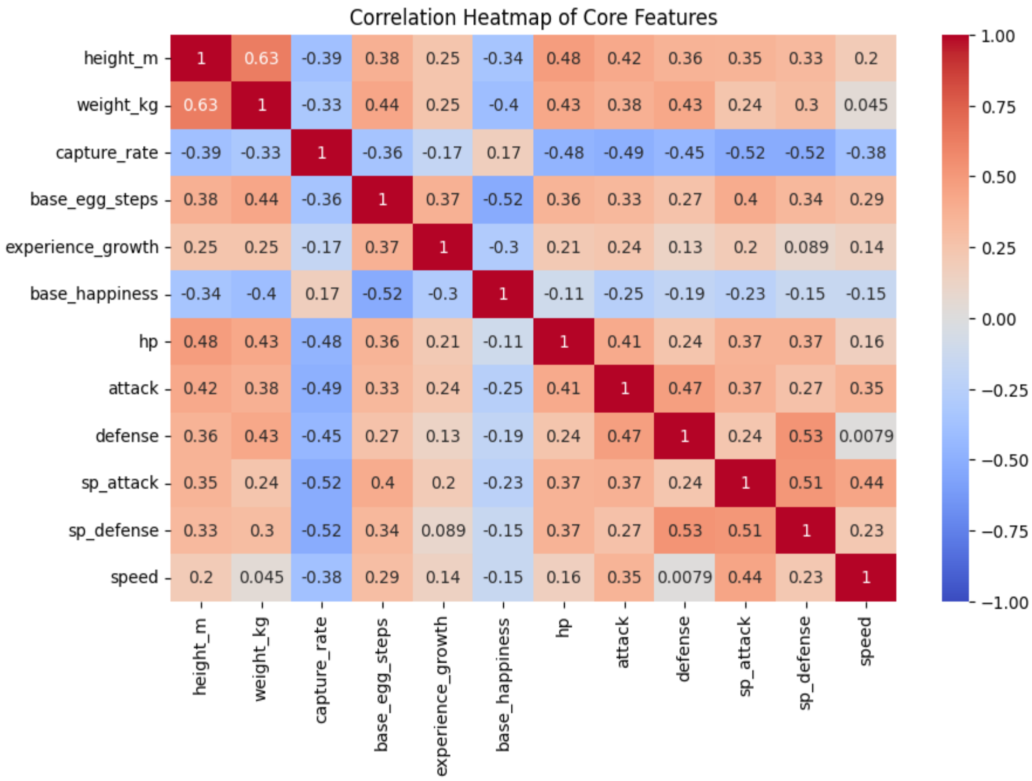

Figure 1: Correlation heatmap of core features.

Analysis on Pokémon type

Pokémon can have either a single type or two distinct types, with 18 possible types in total. In dual-type Pokémon, the order of types does not matter, meaning a Grass/Poison type Pokémon is considered the same as a Poison/Grass type Pokémon. Single-type and dual-type Pokémon are nearly equally represented in the Pokémon world, with a slight majority of dual-type Pokémon at 51.2%, as shown in Figure 2.

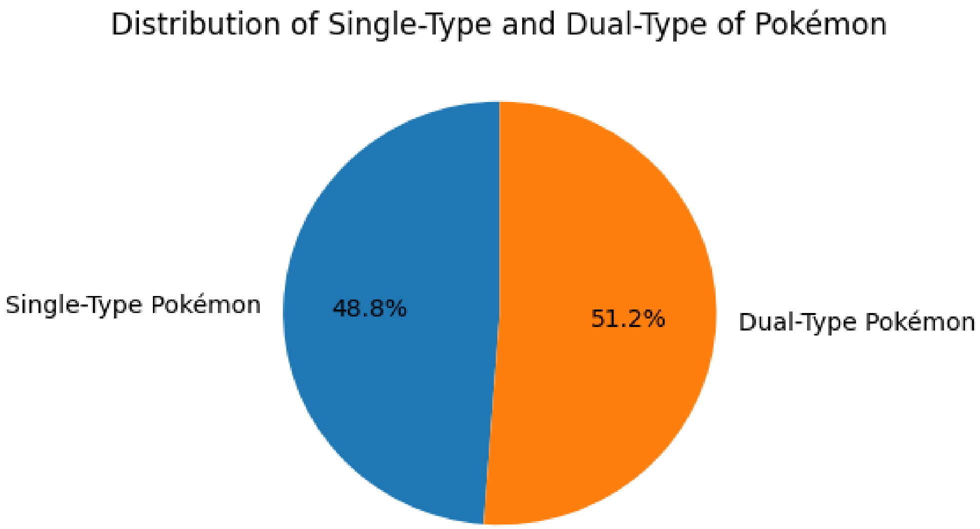

Figure 2: Distribution of single-type and dual-type of Pokémon.

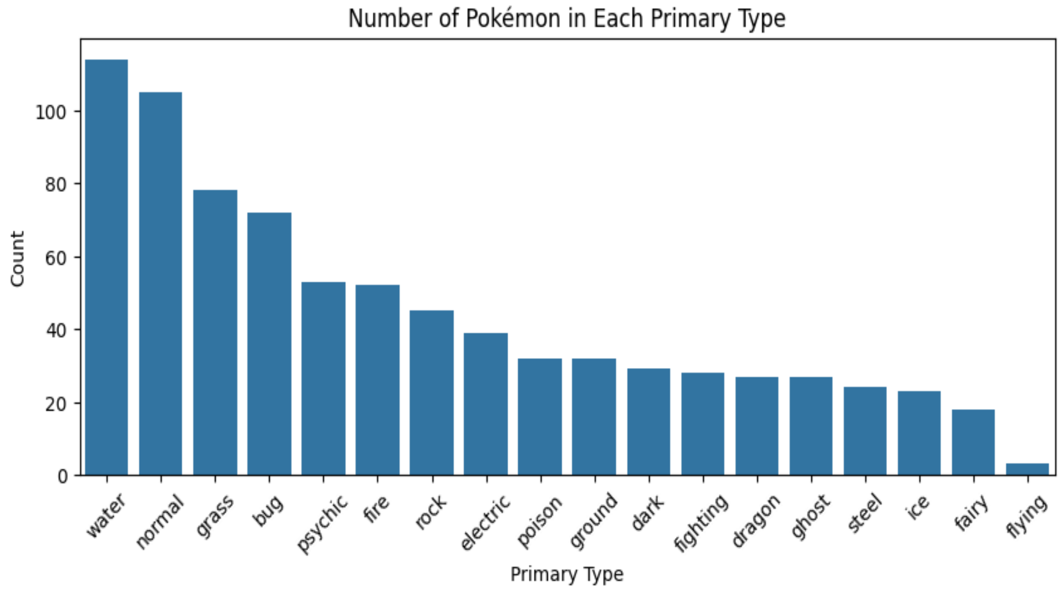

Figure 3: Number of Pokémon in each primary type.

While the order of types has no impact on a Pokémon's type effectiveness, we will assess the dual-type Pokémon according to the order documented in the National Pokédex, where National Pokédex records information on all known Pokémon. As shown in Figure 3, the most common primary types are Water and Normal, while the least common primary type is Flying.

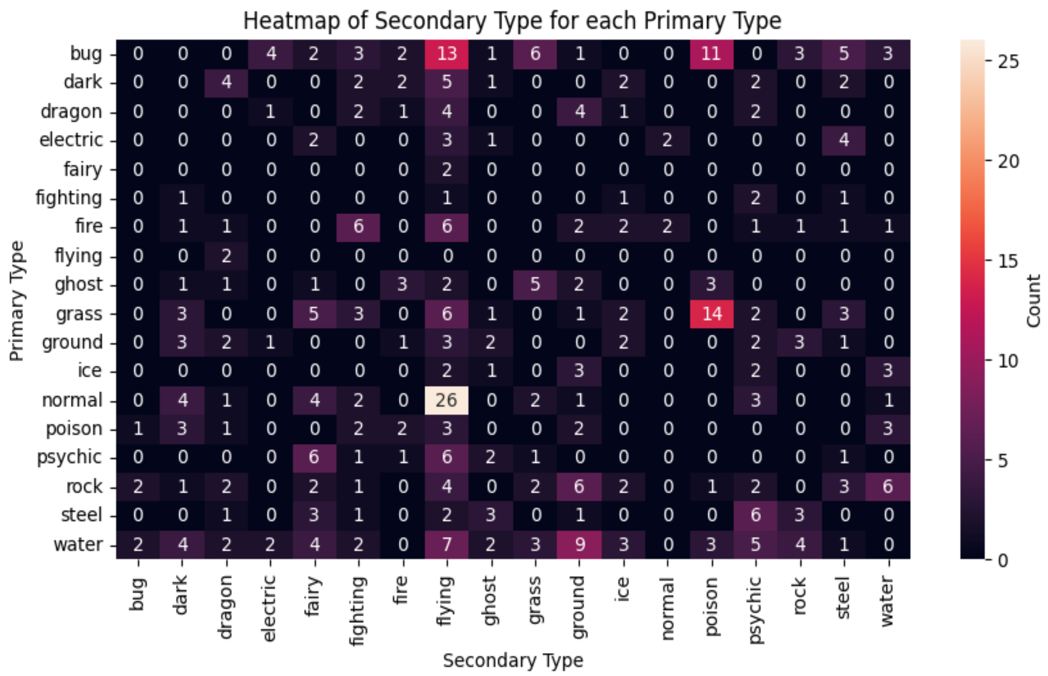

Figure 4: Heatmap of secondary type for each primary type.

The heatmap in Figure 4 shows that Flying is the most common secondary type, especially when paired with Normal as the primary type. This suggests that Flying types more frequently appear in dual-type Pokémon, explaining why they are less common as a primary type. Conversely, it is relatively rare for Pokémon to have Normal or Bug as a secondary type. Additionally, Pokémon with Water and Bug as their primary types are more likely to have a secondary type, likely due to the overall higher frequency of Pokémon in these categories.

Network Analysis on Pokémon Type Interactions

A Pokémon's type has a direct impact on its battle performance, influencing both the damage it causes and the damage it receives when interacting with other types. Dual-type Pokémon have combined advantages and disadvantages which bring complexity to this analysis of type interactions. Therefore, for simplicity, we will focus on the type interactions among single-type Pokémon, where the strengths and weaknesses of each type are straightforward.

We will analyse the interactions among the different Pokémon types in both attack and defense scenarios using the against_? variable, which represents the damage multipliers taken by a Pokémon when attacked by a specific type. For example, against_dragon indicates the damage a Pokémon sustains from Dragon-type attacks. The variables are interpreted as follows:

- Value < 1: Less damage than usual, suggesting the Pokémon's type defends well against that particular type.
- Value = 1: Normal damage, suggesting neutral interaction where the type neither effectively defends nor is particularly vulnerable to that particular type.
- Value > 1: More damage than usual, suggesting the type is more vulnerable to attack from that particular type. Conversely, this suggests an attacking ability for the attacking type.

This variable is used to identify which types excel in attack or defense scenarios, uncovering the natural strengths and weaknesses of each type. We consider types that cause more damage than usual as the effective attacking types, while types that cause less damage than usual as the effective defending types.

Attack scenario

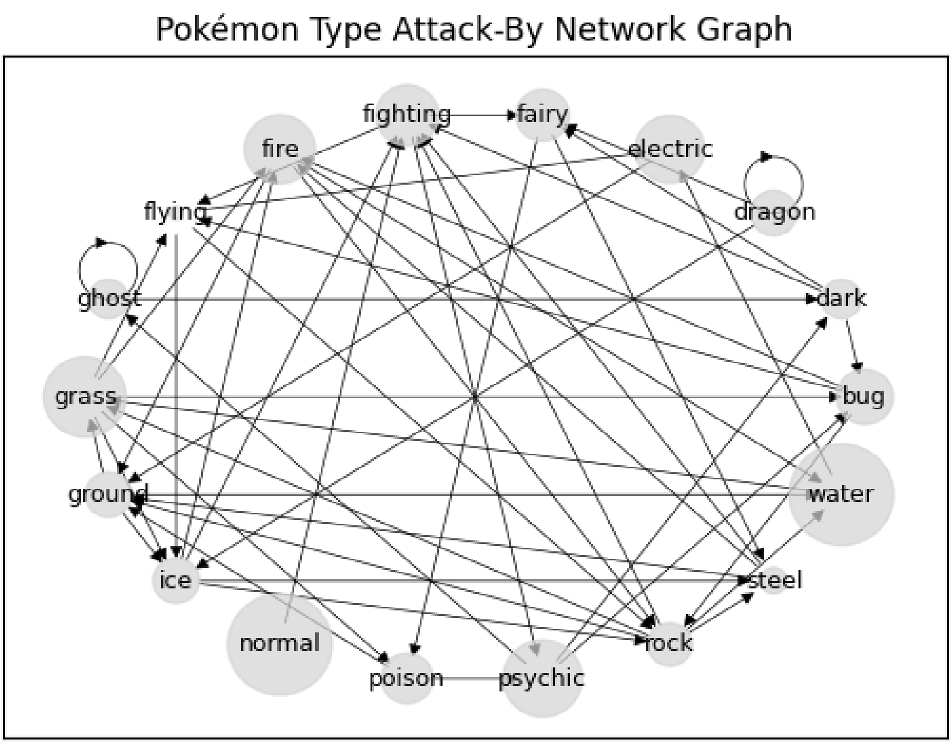

Figure 5: Pokémon type attack-by network graph.

We analyse type interactions where the against_? value is greater than 1 in the attacking scenario, indicating that the attacking type causes more damage to the receiving type. In the network graph shown in Figure 5, each node represents a Pokémon type, and the directed edges represent the effective attack relationships. For example, an edge from Fire to Water indicates that Fire type is effectively attacked by Water type, which mean it causes more damage than usual to Fire type when battle against Water type. The node size is proportional to the weight, which represents the distribution of Pokémon in each type. Larger nodes correspond to more common types, while smaller nodes correspond to rarer types.

Table 3: Pokémon Type Interaction Metrics in attack network.

Type	Weight	In-Links	Out-Links	PageRank	Personalized _PageRank	PR_Rank	PPR_Rank
ground	0.0307	5	3	0.1199	0.1160	1	1
grass	0.0972	3	5	0.0860	0.0954	2	2
water	0.1560	3	2	0.0738	0.0866	4	3
fighting	0.0563	5	3	0.0698	0.0757	6	4
ice	0.0307	4	4	0.0760	0.0700	3	5
fire	0.0716	4	3	0.0677	0.0674	7	6
rock	0.0281	4	5	0.0726	0.0659	5	7
electric	0.0691	2	1	0.0555	0.0617	9	8
flying	0.0026	3	3	0.0559	0.0512	8	9
psychic	0.0895	2	3	0.0454	0.0511	12	10
bug	0.0460	3	3	0.0463	0.0465	11	11
steel	0.0102	3	3	0.0546	0.0439	10	12
fairy	0.0409	3	2	0.0419	0.0382	13	13
poison	0.0384	2	2	0.0408	0.0382	14	14
dark	0.0230	2	3	0.0369	0.0312	15.5	15.5
ghost	0.0230	2	2	0.0369	0.0312	15.5	15.5
normal	0.1560	0	1	0.0083	0.0234	18	17
dragon	0.0307	1	3	0.0116	0.0064	17	18

Table 3 shows the metrics related to the interaction and importance of different Pokémon types, based on the behaviour in the attack-by network graph. In-links indicate the number of types a Pokémon type can effectively attack while out-links indicate the number of types a Pokémon type can effectively attacked by. For example, Ground type has 5 in-links and 3 out-links indicating that it can effectively attack 5 distinct types of Pokémon, while it can be effectively attacked by 3 distinct types of Pokémon. Ground and Fighting types have the most in-links, suggesting that they are the most effective attackers. In contrast, Grass and Rock types have the most out-links, indicating they are more vulnerable and susceptible to attacks from the greatest number of types. Although Normal type cannot effectively attack any Pokémon type, it can be effectively attacked by the fewest types, making it less vulnerable overall.

The PageRank score measures the importance of a type within the attack, considering both the number of links and the ranks of the nodes that link to it. A high PageRank indicates that the type is influential in the attacking network, where it is capable of attacking more types or attacking other influential types. Based on the PageRank scores, Ground and Grass types are the most influential followed by Ice type, indicating that they can attack a wide range of Pokémon types and also types that are important in the network.

However, PageRank only measures general attacking popularity. To account for the distribution of Pokémon type, we use topic-specific (personalized) PageRank, which incorporates the weight of each Pokémon type. For example, even if a type can attack many other types, it may be less important if the number of Pokémon in those types are little, as the chance encountering them in the battle are small. Ground and Grass types remain the most influential, but Water type now ranks higher than Ice type. This is because Water type can attack high-ranking and commonly occurring types like Ground, Fire, and Rock, whereas Ice type primarily attacks rarer types like Flying and Dragon. Therefore, when accounting for Pokémon type distribution, Water type proves to be more influential in the attack network.

Defense scenario

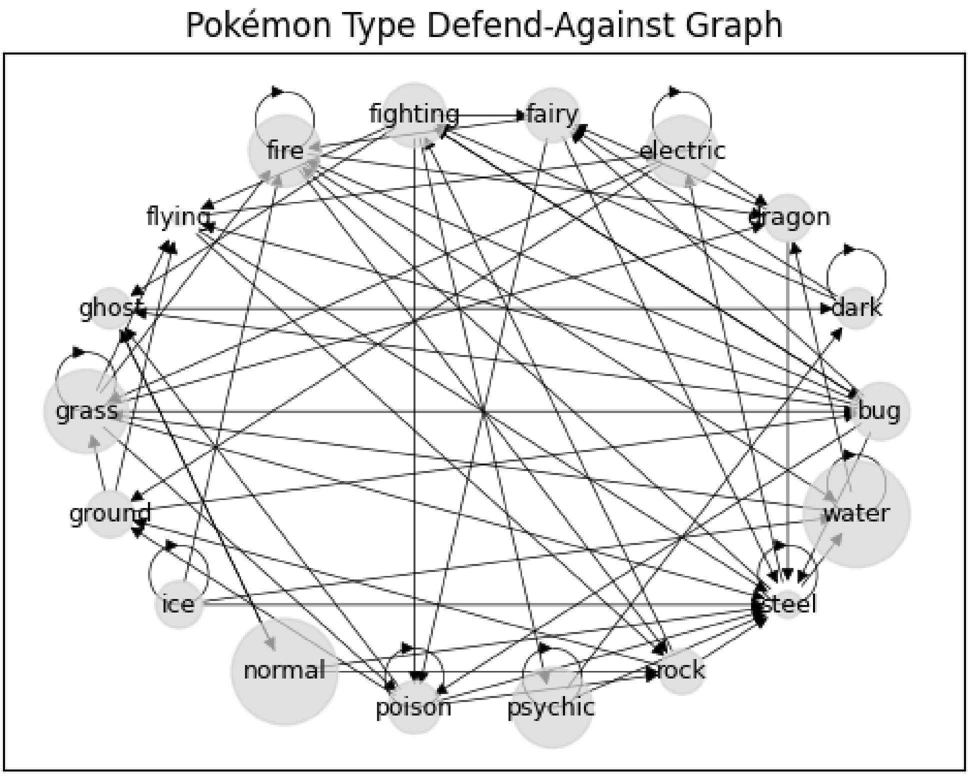

Figure 6: Pokémon type defend-against graph.

For the defense scenario, we analyse the type interactions where the against_? value is less than 1, indicating that the receiving type takes less damage from the attacking type. In the network graph shown in Figure 6, each node represents a Pokémon type, and the directed edges represent the effective defense relationships. For example, an edge from Fairy to Poison indicates that Poison type can effectively defend against Fairy type, which mean Poison type receives less damage than usual when battle against Fairy type. Similar to the attack network, the node size is proportional to the weight, with larger nodes representing more common types.

Table 4: Pokémon Type Interaction Metrics in defense network.

Type	Weight	In-Links	Out-Links	PageRank	Personalized_ PageRank	PR_Rank	PPR_Rank
steel	0.0102	11	4	0.1529	0.1464	1	1
water	0.1560	4	3	0.0876	0.1053	3	2
fire	0.0716	6	4	0.0929	0.0929	2	3
grass	0.0972	4	7	0.0689	0.0793	5	4
dragon	0.0307	4	2	0.0753	0.0774	4	5
electric	0.0691	3	4	0.0662	0.0638	6	6
fairy	0.0409	4	3	0.0613	0.0578	7	7
rock	0.0281	4	3	0.0548	0.0516	8	8
poison	0.0384	5	5	0.0530	0.0494	9	9
ground	0.0307	3	3	0.0469	0.0412	10	10
normal	0.1560	1	3	0.0226	0.0368	16	11
fighting	0.0563	3	6	0.0393	0.0367	12	12
dark	0.0230	3	3	0.0392	0.0338	13	13
bug	0.0460	3	7	0.0356	0.0334	14	14
ghost	0.0230	4	2	0.0336	0.0315	15	15
flying	0.0026	4	3	0.0399	0.0309	11	16
psychic	0.0895	2	3	0.0194	0.0260	17	17
ice	0.0307	1	4	0.0106	0.0058	18	18

Table 4 shows the metrics related to the interaction and importance of Pokémon types in the defend-against network graph. In-links indicate the number of types a Pokémon type can effectively defense while out-links indicate the number of types that can defend against it. For example, Steel type has 11 in-links and 4 out-links indicating that it can effectively defends against 11 distinct types, while only 4 distinct types can defend against attack from steel type, suggesting that it is the strongest defenders. In contrast, Grass and Bug types have the most out-links, indicating greater vulnerability as more types are able to defend against their attacks.

The PageRank score measures the type's importance in the defense network. A high PageRank indicates that the type is influential in the defending network, where it is can defend against more types or defend against influential types. Based on the PageRank scores, Steel and Fire types are the most influential followed by Water type, indicating that they can defense a wide range of Pokémon types and also types that are important in the network.

Similarly to the attack network, the distribution of Pokémon type is considered using the personalized PageRank. While Steel remains the top defender, Water type surpasses Fire in influence when weighted by the distribution across the Pokémon type, suggesting a stronger defense for Water type against common types in battle.

Summary

PageRank scores in both networks highlight types that not only interact widely but also with other influential types, indicating their central roles. Ground and Grass types are highly influential in the attack network, while Steel type dominates in the defense network. This suggests that they serve as pivotal types within their respective attacking and defending strategies. Personalized PageRank refines this perspective by incorporating type distribution. Water type gains higher importance in both networks when accounting for its common presence, making it a well-rounded choice that can perform well in both attacking and defending scenarios against common matchups.

When crafting a Pokémon team, incorporating high-ranking types from each network ensures broad coverage and resilience. For example, Ground, Steel, and Water types provide a versatile team that can maximizing flexibility in different battle situations. Meanwhile, although types like Ghost and Dark have lower PageRank scores, they can add strategic depth with their unique interactions, covering niche matchups that are not easily addressed by more common types.

However, the analysis did not consider the dual-type Pokémon which may missed out the specific interactions that is unique to the dual-type. The analysis also did not consider individual battle stats such as the attack strength, defense strength and hit points, which could further impact the type's effectiveness. These can reveal types that are powerful in battle strength though less central in the network. Therefore, using network influence alongside individual stats, with consideration of dual-type Pokémon would yield a comprehensive strategy for building competitive Pokémon teams.

Clustering Pokémon Based on Characteristics

In the Pokémon world, there are many features in addition to the type of the Pokémon that determines the unique characteristics of a Pokémon. To uncover natural groupings of the Pokémon, we will consider the core features to cluster the Pokémon. The core features considered include biometric attributes such as height and weight, non-combat attributes such as capture rate, base egg steps, experience growth and base happiness and battle stats such as HP, attack, defense, special attack, special defense and speed.

Standardize and Reduce Dimensionality of the Features

The features are standardized by removing the mean and scaling to unit variance. This ensures that each feature contributes equally to the cluster formation, regardless of its original scale. Singular Value Decomposition (SVD) is applied to reduce the dimensionality of the feature space. This compressed representation captures the most of the original data's variance while discarding noise or less informative components, making it easier to visualize data and perform clustering, with reduced computational cost and minimal information loss.

To determine the number of components to retain from the SVD, we compute the explained variance ratio, which indicates the percentage of variance explained by each component, as illustrated in Figure 7. Based on the plot, it is observed that the individual explained variance decreases significantly beyond 2 components. However, the cumulative explained variance reaches 80% with at least 6 components. Therefore, we will fit the clustering with the dimensionality-reduced data, reducing from the original 12 features to 2 and 6 features respectively, to compare which reduction provide a better clustering.

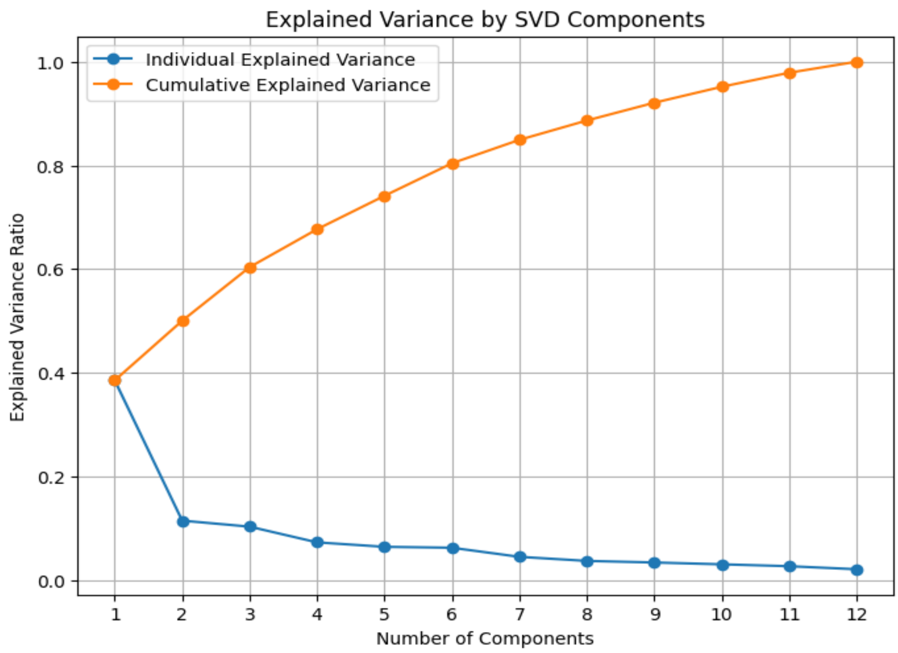

Figure 7: Explained variance by SVD Components

K-Means Clustering

The elbow method is used to determine the optimal number of clusters (k) by identifying the "elbow" point at which increasing the number of clusters no longer significantly improves the model's performance. The model's performance is measured by the within-cluster sum of squares (WCSS), also known as inertia, which measures how closely the data points within a cluster are grouped around the centroid.

The elbow method was applied to three different sets of features: the non-reduced standardized features, the reduced standardized features with 6 dimensions, and the reduced standardized features with 2 dimensions. All three sets of features resulted in the same optimal number of clusters, i.e., k = 3, based on the elbow method. This consistency suggests that the optimal clustering structure in the data remains the same, regardless of the dimensionality of the feature set. The elbow method which illustrates this result is shown in Figure 8.

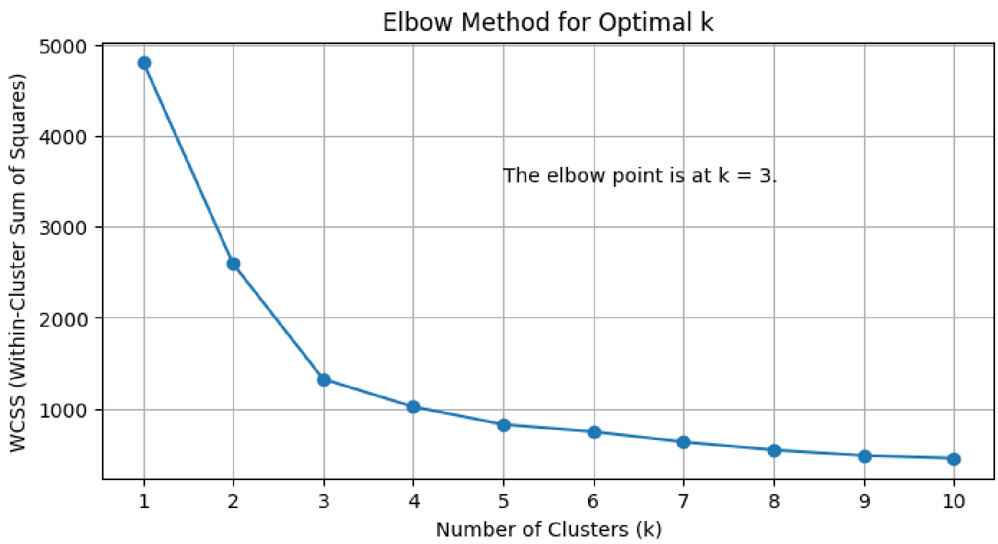

Figure 8: Elbow method for Optimal k.

K-Means algorithm partition data into k-clusters by iteratively assigning points to the nearest centroid and updating the centroids to minimize the within-cluster variance. The algorithm stops when the centroids stabilize. The K-means clustering is performed on the three different sets of features mentioned. Silhouette score is used to compare the clustering quality across the different sets where the clusters are assumed to be convex-shaped. Silhouette score measures how similar each data point is to its own cluster compared to other clusters and it is ranged from -1 to 1, where

- score near to +1 indicates that the data points are well-clustered, with each point being closer to its own cluster's centroid than to other clusters.
- score near to 0 indicates that the data points are very near to the decision boundary between 2 clusters, and
- a score near to -1 indicates that the data points are likely to been assigned to the wrong clusters, with the data points being closer to the centroids of other clusters than to their own.

The Silhouette score for the non-reduced standardized features, the reduced standardized features with 6 dimensions, and the reduced standardized features with 2 dimensions are 0.233, 0.297 and 0.513 respectively. Therefore, the K-Means clustering model fitted with the reduced standardized features with 2 dimensions achieved the highest Silhouette score, indicating that it has the most effective and accurate clustering result.

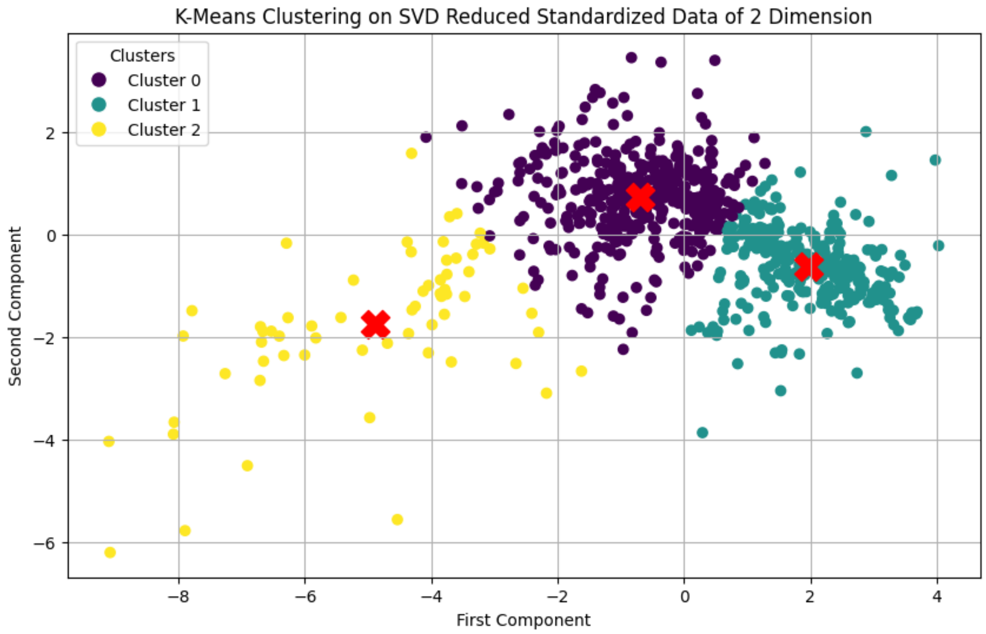

Figure 9: K-Means clustering on SVD reduced standardized data of 2 dimension.

Figure 9 shows a scatterplot with 801 Pokémon plotted on the first and second components of the SVD reduced standardized data in 2 dimensions. The K-Means clustering separated the Pokémon into 3 clusters, as shown by the different colors. Cluster 0 and 1 appear to be more compact, which suggests that the Pokémon within these clusters share more similar attributes. In contrast, Cluster 2 is more spread out suggesting greater variability among the Pokémon in this group. The scatterplot shows some overlap at the boundaries between the clusters, which reveals areas where the clustering algorithm has difficulty distinguishing between Pokémon based on the chosen features.

Agglomerative Hierarchical Clustering

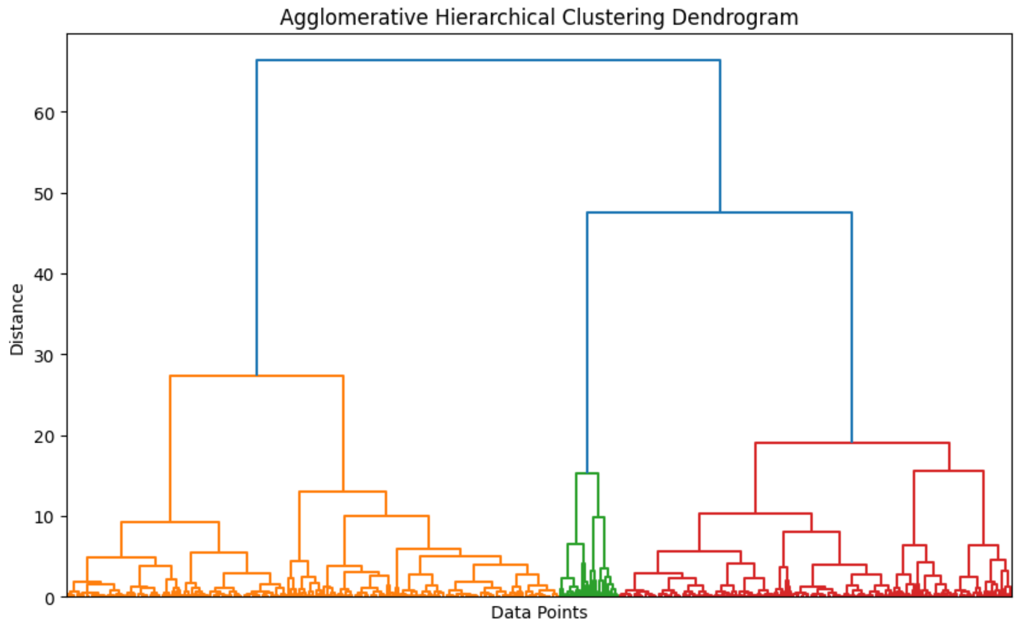

Figure 10: Hierarchical clustering dendogram

Agglomerative hierarchical clustering algorithm begins by treating each data point as a cluster and iteratively merges the closest clusters using the Ward's linkage method, which merge clusters by minimizing the increase in total variance. The algorithm continues until all data points are in a single cluster. Figure 10 shows the hierarchical structure in a dendogram, revealing 3 distinct clusters at the distance of around 30. This aligns with the optimal cluster count determined through the K-Means clustering.

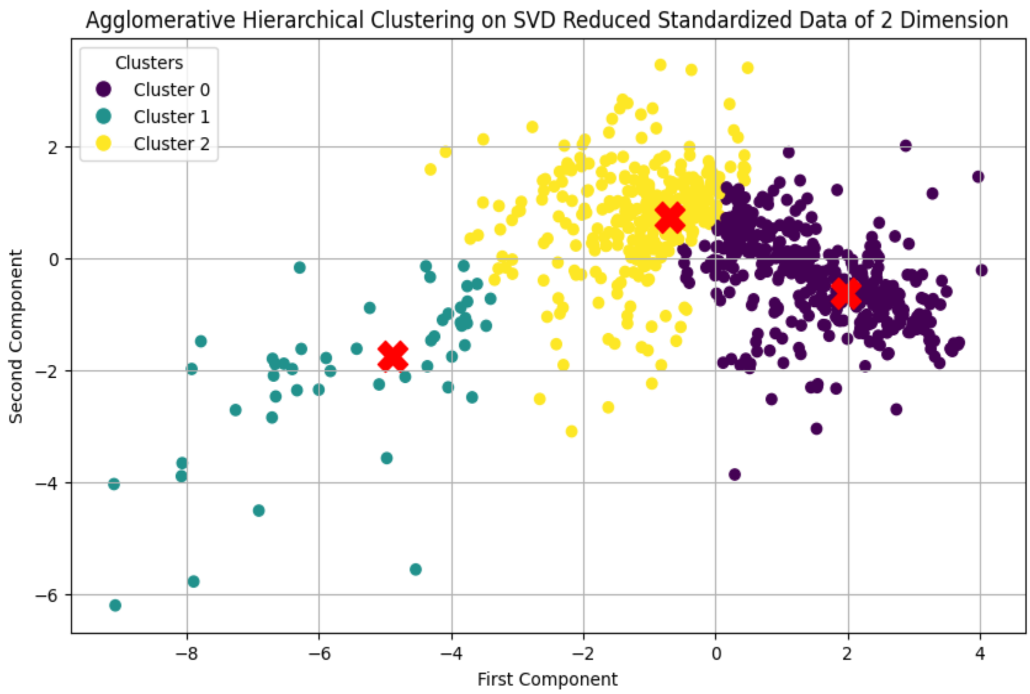

Figure 11: Hierarchical clustering on SVD reduced standardized data of 2 dimension

Similar to Figure 9, Figure 10 shows a scatterplot with 801 Pokémon plotted on the first and second components of the SVD reduced standardized data in 2 dimensions. The agglomerative hierarchical clustering separated the Pokémon into 3 clusters, as shown by the different colors.

The clusters determined by hierarchical clustering are similar to the one by K-Means clustering, however there are some differences at the cluster boundaries. The Silhouette score for the agglomerative hierarchical clustering is 0.454, which is lower than the Silhouette score of 0.513 obtained from K-Means clustering with the reduced standardized features in 2 dimensions, indicating that the K-Means model provides a more accurate clustering result.

Analysis of clustering result

We will analyse the clusters based on the K-Means clustering with the reduced standardized features in 2 dimensions, in an attempt to understand the natural grouping of Pokémon.

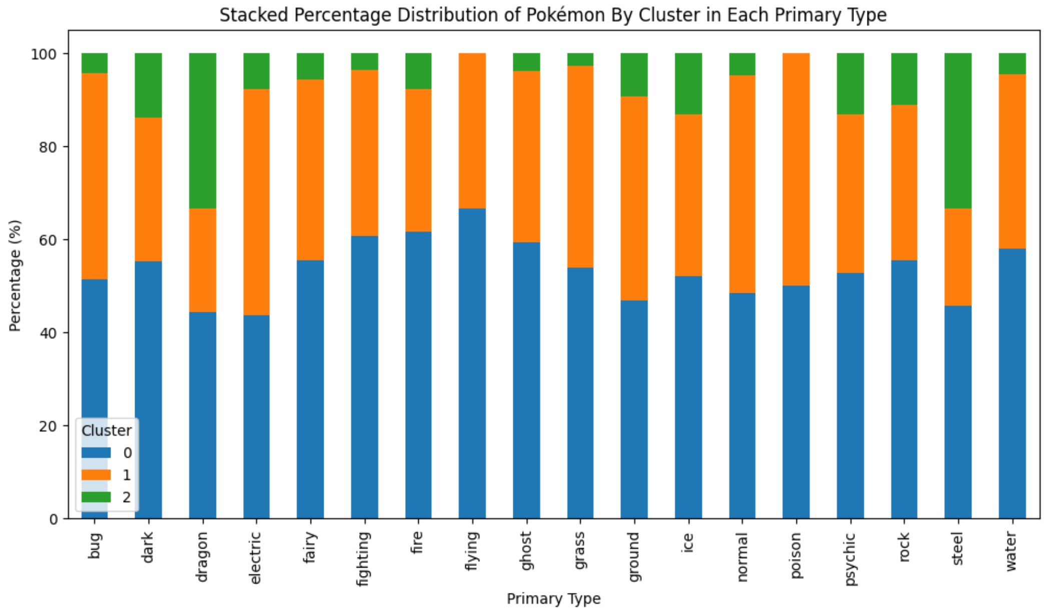

Figure 12: Stacked percentage distribution of Pokémon by cluster in each primary type.

As observed in Figure 9 and Figure 12, Cluster 2 have the least number of Pokémon, while Cluster 0 has a greater number of Pokémon than Cluster 1. Figure 12 further reveals that Dragon and Steel type are more commonly classified to Cluster 2 compared to other Pokémon types, while there are little to no Flying and Poison types in Cluster 2. A Chi-squared test is conducted to test if there is an association between the clusters and the Pokémon types. The test yielded a Chi-squared statistics of 72.01 and a p-value of 0.00015, which is significantly less than 0.5. Hence, we can conclude that the cluster and primary type are significantly associated.

Density plots in Figure 13 provide insights into the distribution of Pokémon attributes within each cluster. For both height and weight, the Pokémon in Cluster 0 and 1 are more concentrated, with Pokémon in Cluster 0 having slightly higher mean compared to Cluster 1 in general. However, the height and weight distribution for Cluster 2 are more spread out, with some Pokémon exhibiting higher values. In terms of HP and speed, Custer 3 have the highest mean values for both attributes, followed by Cluster 0, then Cluster 1.

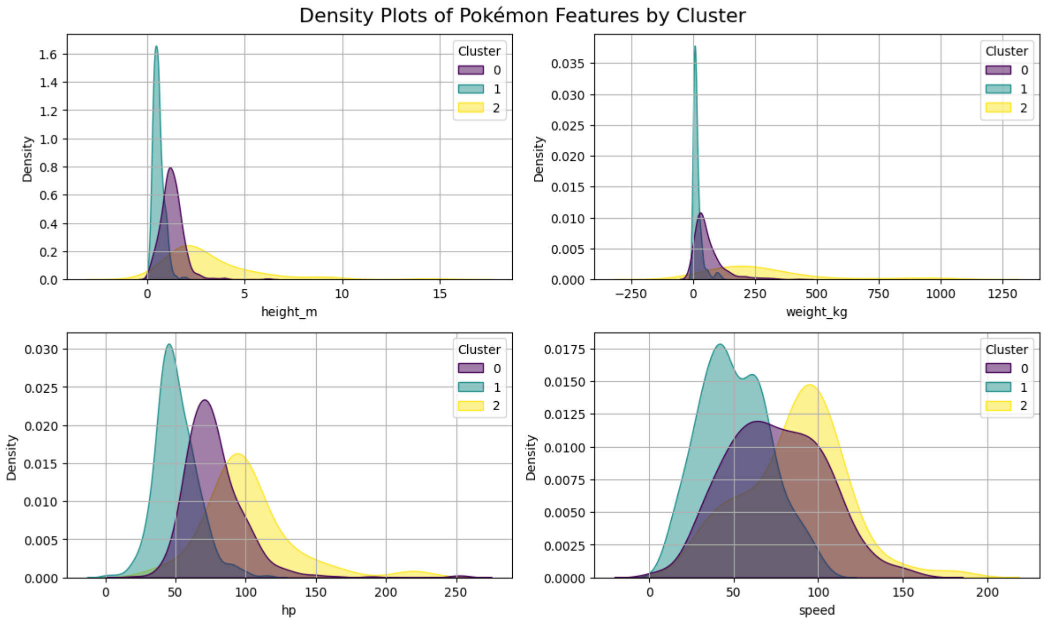

Figure 13: Density plots of Pokémon features by cluster

The scatterplot in Figure 14 shows the distribution of Pokémon based on their attack and defense strength across the different clusters. Although there is no clear boundary between the clusters in term of attack and defense, there are some patterns discovered. Cluster 2 are generally more spread out along the higher end of the attack and defense scales, suggesting that Pokémon in Cluster 2 exhibit high variability in these attributes, and have a greater potential for high attack and defense abilities. Cluster 1 tends to be positioned at the lower end of both the attack and defense scales, indicating that Pokémon in this cluster is generally weaker in terms of their attack and defensive capabilities compared to Pokémon in Cluster 0 and Cluster 2.

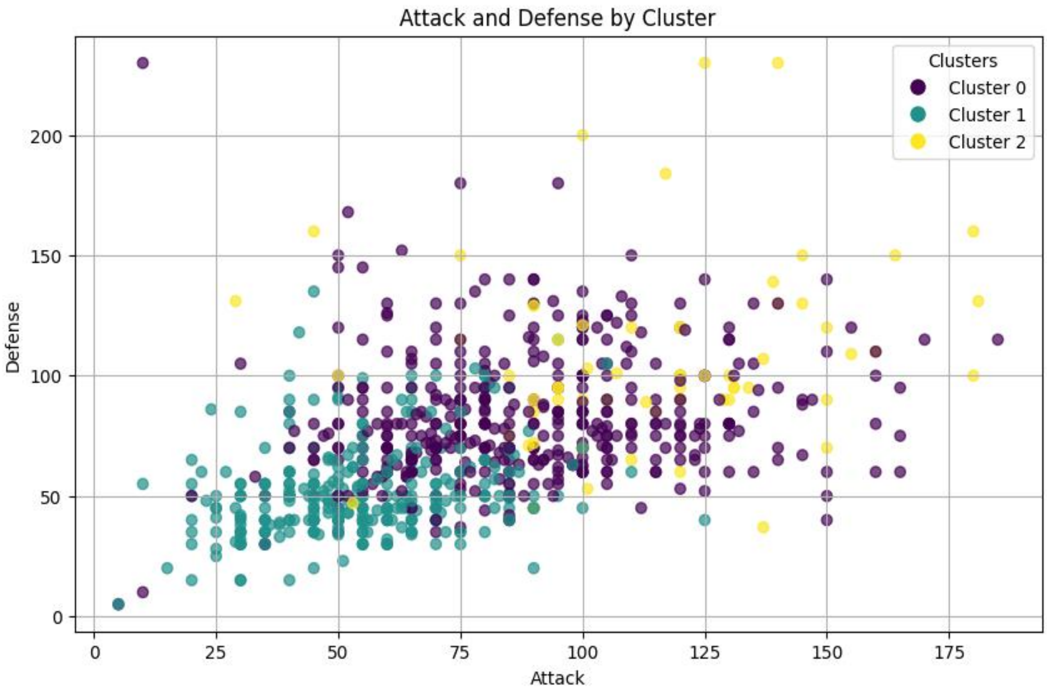

Figure 14: Attack and defense by cluster

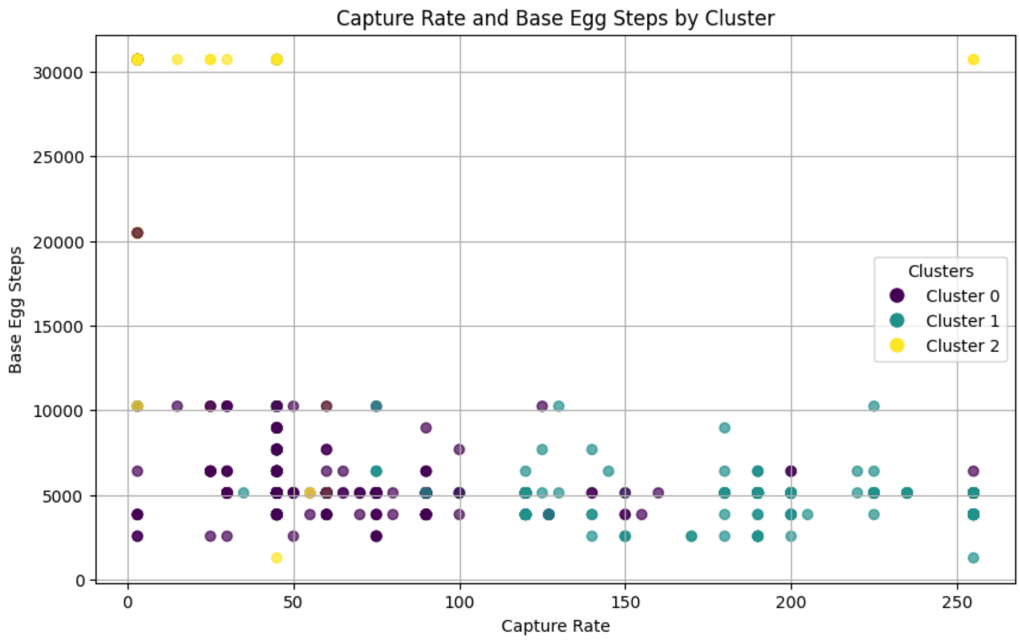

Figure 15: Capture rate and base egg steps by cluster

The scatterplot in Figure 15 shows the distribution of Pokémon based on their capture rate and base egg step across the different clusters. It is observed that Pokémon in Cluster 2 tend to have low capture rates and is the only cluster that has extremely high base egg steps, suggesting that these Pokémon may be the rarer and require more effort to capture and breed. Pokémon in Cluster 0 generally have lower capture rate, while Pokémon in Cluster 1 have higher capture rate, with both clusters require similar levels of breeding effort, as reflected by the base egg steps.

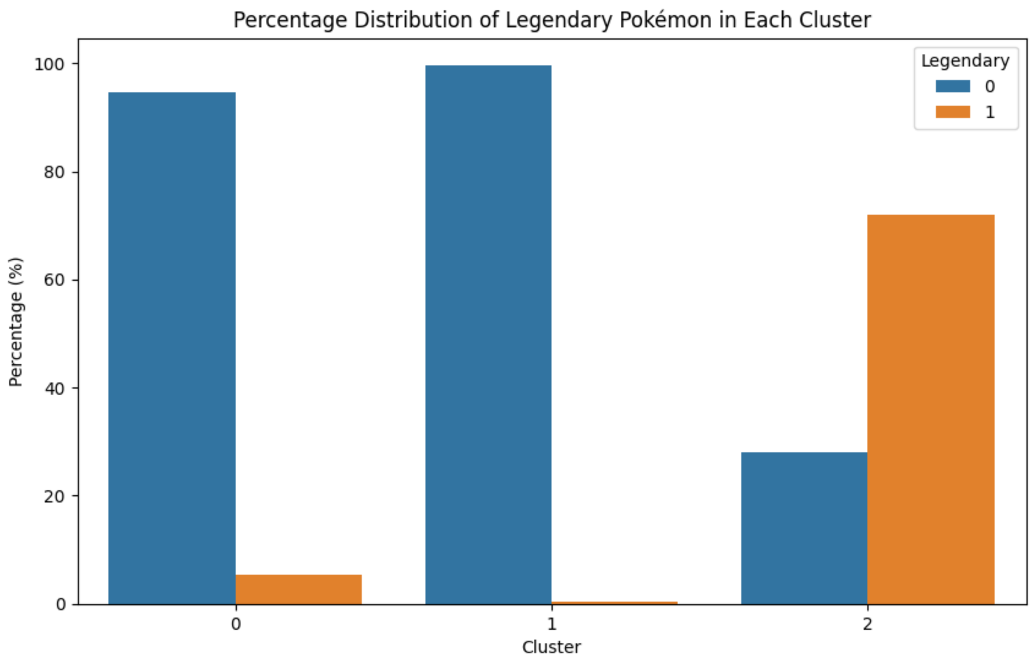

Figure 16: Percentage distribution of legendary Pokémon in each cluster

From Figure 16, it is evident that majority of Cluster 2 consists of legendary Pokémon, which are typically rare and powerful. This observation aligns with previous analyses where Cluster 2 have a wide distribution of height and weight, as well as higher average values for HP and speed.

Additionally, Cluster 2 exhibits greater variability in speed and defense, with potential for high attack and defense abilities, along with low capture rate and high breeding effort, which are the characteristics associated with rare and powerful Pokémon

Based on the above analyses, it suggests that Pokémon are naturally grouped into 3 clusters. Cluster 2 is generally composed with rare and powerful Pokémon such as the legendary Pokémon. In contrast, Cluster 1 contains the more common and weaker Pokémon, while Cluster 0 includes common Pokémon that have a middle to strong ability. This clustering provides insight into how Pokémon can be grouped by their characteristics and battle potential. However, there are also some boundary confusions, where certain Pokémon near the cluster boundary are not clearly distinguishable by the clustering model, indicating possible limitations in separating groups with subtle differences.

Conclusion

In conclusion, the two analyses conducted have provided valuable insights into the Pokémon world in terms of their type dynamics and character profiles. The first analysis has successfully evaluated the dynamics of primary types in terms of their attacking and defending relationships within a network, using PageRank and topic-specific (personalized) PageRank. This revealed the most influential types in the attacking and defending network, providing information on the strategic interactions between Pokémon types.

The second analysis performed K-means clustering and agglomerative hierarchical clustering on the dimensionally reduce features set to group Pokémon into three clusters. These clusters represented natural groupings of Pokémon, such as the common and weak Pokémon, the common and middle to strong Pokémon and the rare and strong Pokémon. The clustering results align with the idea that Pokémon possess varied battle strengths and characteristics which influence their roles in battle.

The results from the two analyses provide insights that can influence strategic decisions in forming a Pokémon team and also planning in the battle.