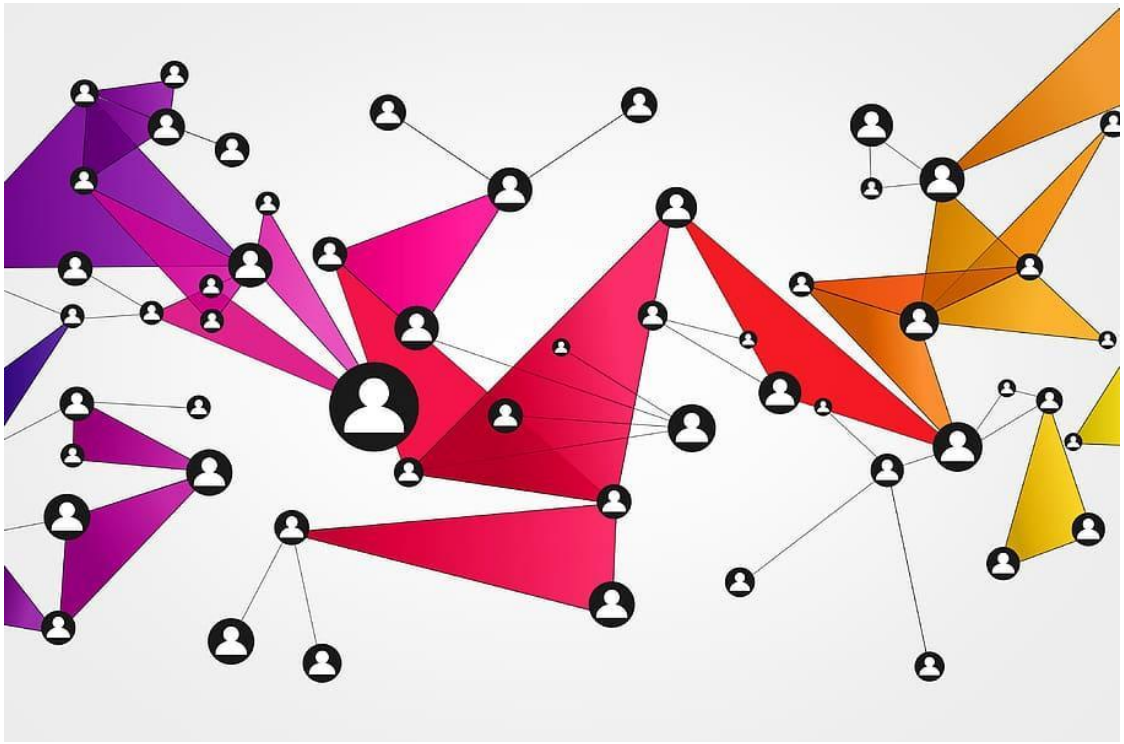
Group - 6

19OH01 - Social and Economic Network Analysis



Team:

19z310 - BHOOVIKA

19z314 - GOKUL

19z334 - NIKHIL

19z337 - PAVITHRA YAZHINI

Social Network Analysis with Genshin Impact

- ## Problem Statement

Examining the Genshin Impact dataset, a well-known game. In this project, we identify the important character and the popular character from a group of characters, the strongest nation from a group of four nations, and make some predictions on rarity.

- ## Dataset description

The Genshin dataset is used in our project and carried out some analysis. Numerous characters from four different countries appear in this dataset. The unique weapon and vision for each character also appeared in the dataset. In this case, we will treat the characters as nodes and the nations around the globe of the game as regions. All the characters in this game are strangers to one another. For instance, there is an edge between two characters if character A knows character B. Rarity is a column in our dataset for all characters which represents the difficulty level of reaching that particular character by other characters in the game.

- ## Tools used

Google collab: It allows anybody to write and execute arbitrary python code through the browser, and is especially well suited to machine learning, data analysis, and education.

Pandas: Library written for the Python programming language for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and time series.

Seaborn: A library that uses Matplotlib underneath to plot graphs. It will be used to visualize random distributions.

Pyplot: An API (Application Programming Interface) for Python's matplotlib that effectively makes matplotlib a viable open-source alternative to MATLAB. Matplotlib is a library for data visualization, typically in the form of plots, graphs, and charts.

Numpy: Library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays.

Standard Scaler: IT removes the mean and scales each feature/variable to unit

variance. This operation is performed feature-wise in an independent way. Standard Scaler can be influenced by outliers (if they exist in the dataset) since it involves the estimation of the empirical mean and standard deviation of each feature.

Polynomial Features: The polynomial features transform is available in the scikit-learn Python machine learning library via the Polynomial Features class.

linear model: Linear model is a class of the sklearn module that contains different functions for performing machine learning with linear models. The term linear model implies that the model is specified as a linear combination of features.

train_test_split: The train-test split is used to estimate the performance of machine learning algorithms that are applicable for prediction-based Algorithms/Applications. This method is a fast and easy procedure to perform such that we can compare our own machine-learning model results to machine results.

networkx: A Python package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks. And NetworkX is a Python library for studying graphs and networks.

json: An open standard file format and data interchange format that uses human-readable text to store and transmit data objects consisting of attribute-value pairs and arrays. It is a common data format with diverse uses in electronic data interchange, including that of web applications with servers.

- ## Challenges Faced

  - ✓ Faced some difficulties in the visualization of the graphs accurately.
  - ✓ Dataset processing the dataset was a bit challenging.
  - ✓ Integrating the code was quite difficult.
  - ✓ Genshin's game was too difficult to understand. Even though it was a popular game, the dataset was quite confusing and had a lot the concept to learn and work on it.

- ## Contribution of Team Members

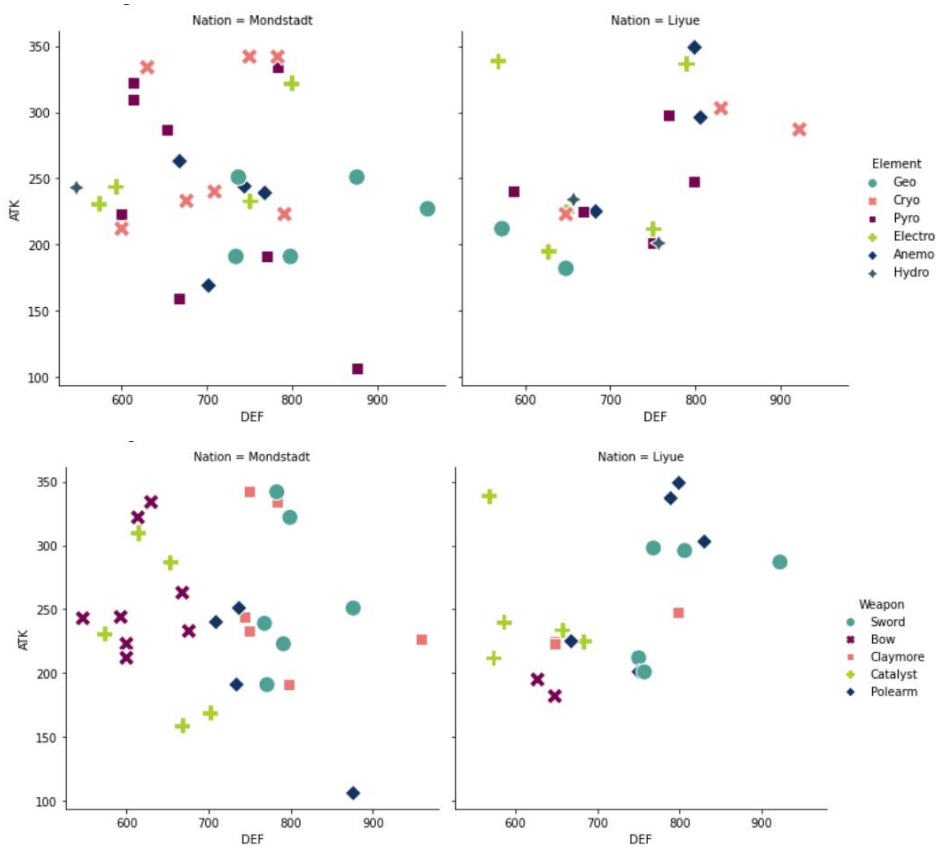
ROLL NO	NAME	CONTRIBUTION
19Z310	Bhoovika	Analysis based on rarity prediction
19Z314	Gokul	Finding the important person + report work
19Z334	Nikhil	Finding the popular person
19Z337	Pavithra Yazhini	Analysis based on stronger nation

- Annexure I: Code

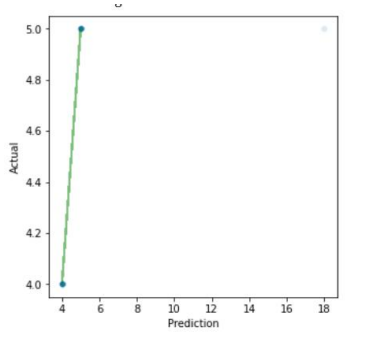
https://github.com/Bhoovika/SENA-project/tree/main/Code

- Annexure II: Output

Output of which nation is strongest:



Output of Prediction:



Output of who is popular and important:

	id	nation	in_degree	out_degree
25	Traveler	NA	0	14
24	Lisa	Mondstadt	15	12
34	Jean	Mondstadt	17	11
5	Kaeya	Mondstadt	13	11
6	Klee	Mondstadt	13	11

	id	nation	in_degree
34	Jean	Mondstadt	17
24	Lisa	Mondstadt	15
5	Kaeya	Mondstadt	13
6	Klee	Mondstadt	13

	id	nation	degree_rank	closeness_rank	betweenness_rank	eigen_rank	average_rank
34	Jean	Mondstadt	1	3	11	1	4.00
5	Kaeya	Mondstadt	3	7	8	3	5.25
6	Klee	Mondstadt	4	6	10	4	6.00
24	Lisa	Mondstadt	2	12	14	2	7.50
3	Xiangling	Liyue	7	2	1	21	7.75
14	Eula	Mondstadt	15	4	3	11	8.25
16	Albedo	Mondstadt	17	5	5	8	8.75
29	Diona	Mondstadt	21	1	2	13	9.25

References

1. https://genshin-impact.fandom.com/wiki/Event
2. https://genshin-impact.fandom.com/wiki/Character#Playable_Characters
3. https://www.sportskeeda.com/esports/genshin-impact-wish-simulator-what-use-it
4. https://videogamesstats.com/genshin-impact/
5. https://drive.google.com/drive/folders/1b0TrYQn52HrzOQD9g_4XixdoqclfVfV?usp=share_link
6. https://gamewith.net/genshin-impact/article/show/23585
7. https://www.kaggle.com/code/faryarmemon/genshin-impact-tevyant-character-analysis/notebook
8. https://cambridge-intelligence.com/eigencentrality-pagerank/
9. https://www.kaggle.com/code/amanb096/genshin-impact
10. https://www.kaggle.com/datasets/genshinplayer/genshin-impact-characters-stats