Eurovision Song Contest

Introduction

Eurovision is a song contest that was conducted since 1956. There are two semi-finals and one final also the participating countries may differ from year to year. Every country gives points two sets of 1-8, 10 and 12 to participants. The country which get highest score wins.

Our scope is to examine results between 1998 and 2008 for finals. We conducted 6 different approaches in general and in some questions we tried to find solutions by using more than one approaches. Six approaches consists of examining 10 years together, examining 10 years whose weights are more than 7 , examining start year (1998), examining end year (2009), consider links which are greater than 7 for both 1998 and 2008. Our questions were is there any tendency to vote for those who gave vote to them, is gender a deterministic factor for voting, does religion have any influence on scoring, do countries in the same region usually vote for each other, do few countries get higher scores and majority get lower scores. We tried to find solutions by using social network analysis approaches.

Research Methods

The base data was collected from Kaggle. The data includes year, source country, target country, points, and final/semifinal information. [1] The data was enriched by another dataset which includes region, artist, and song name and gender information. [2]The region and gender information is used from second dataset and religion information of countries was also found on the Internet and added to this data set. Before starting our analysis we created multiple data frames and networks for different years and weights which included source, target and weight information.

Results

In first question, is there a tendency to vote for those who gave vote to them? As a result, we evaluated reciprocity for 1998 and 2008 which were beginning and end years of our data. In 1998, the reciprocity was 0.44 but in 2008 that score was 0.28. That means at past people tended to vote for those who gave vote to them. In another perspective, if we narrowed two years data for the weight is over 7 in 1998 the reciprocity score was 0.16 and 0.23 for 2008. That means countries tended to give high points who also gave them high points increased in recent years for high scores.

In second question, is gender deterministic factor for voting process so that we evaluated scores for female, male and groups in four categories which were 1998, 2008, scores are over 7 for 1998 and 2008. In 1998, there were no groups and female singers dominated the scores both for all scores and scores which were over 7. On the other hand, in 2008 there were three types of participants which were males, females and groups. Similarly as in 1998, females dominated the scores. As an interesting finding, both for all scores and scores above 7 the groups were the minority.

In third question, is religion a deterministic factor on voting? However, in 2008 %87 of the participant countries religion was Christianity which was stated in Figure 1. As a result, we could not conclude that question.

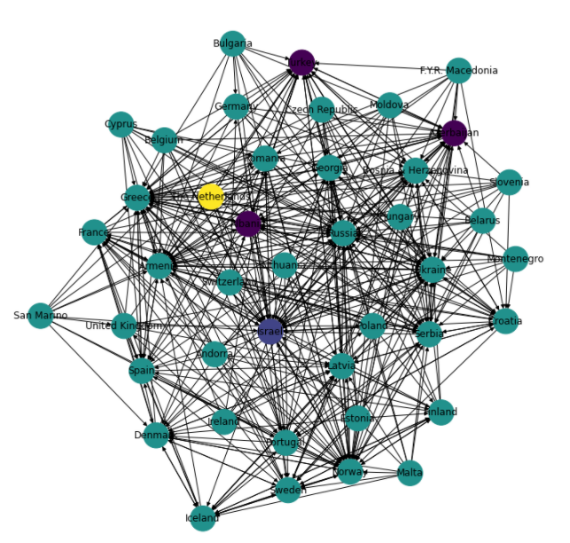
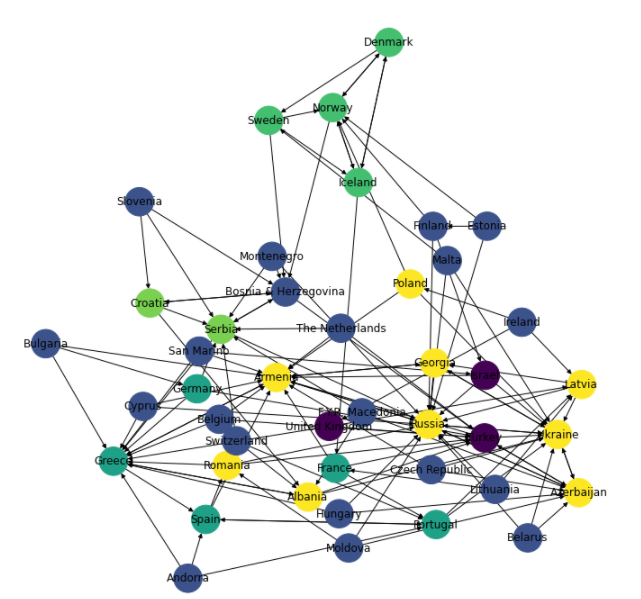


Figure 1: Directed network colored by religion

In fourth question, do countries in same region usually give points to each other so that we created graphs and colored by their regions for 1998, 2008 and weights above 7 for 1998 and 2008. Comparing 1998 and 2008 revealed that most countries tend to give points in same regions. However, in order to prevent confusion, for 1998 and 2008 the weights which are greater than 7 was added to the paper and shown in Figure 2. As a result of this part, mainly there was a direct arrow between same colors.

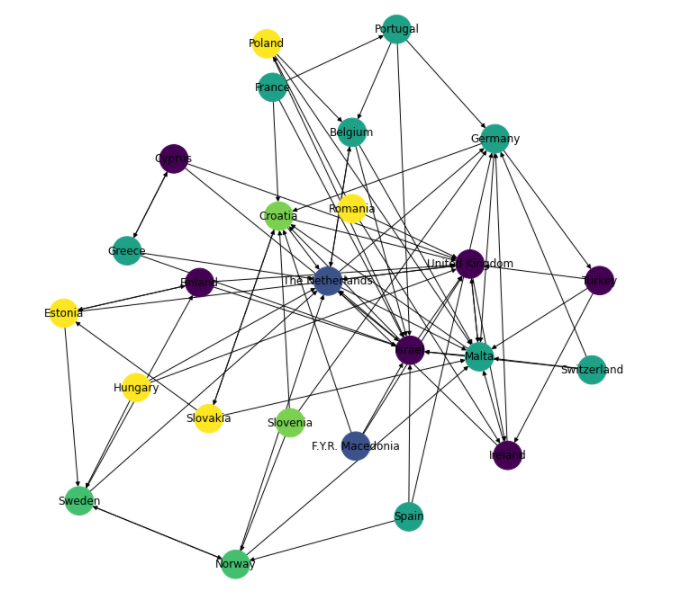
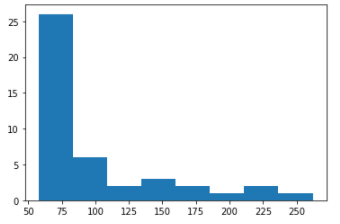


Figure 2: Left graphic for 1998 right graphic for 2008 which the directed graphs colored by region

In fifth question, do few countries get higher scores and majority get lower scores? The directed network converted to undirected network to calculate degrees. Degrees of all countries are plotted as histogram in order to understand if there is a power law or the network formed as random. In 2008, we can observe that less countries got high scores and majority got lower scores so there is a power law. On the other hand, in 1998 the numbers were distributed randomly so there is no power law. The histograms for 1998 and 2008 are visualized in Figure 3. Also high weight version is similar so that they were not added to the paper.



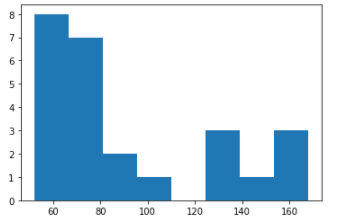


Figure 3: Left histogram for 1998 right histogram for 2008

Conclusion

As a result, even if reciprocity score is higher in 1998, in weight eliminated version 2008’s result is higher so that for high scores countries tend to give high scores where they received. Gender is a deterministic factor on scoring thus female participants mostly got higher scores. Religion was not an important attribute. On the contrary, region is important in voting decisions. Countries tend to give scores in their same region. It is understood from the findings that minority gets the higher scores which were 8, 10, 12 and majority gets lower scores.

Further Research

In further research, for every year the political relationship between countries can be added in dataset. The reason is, in recent years it is debated that instead of evaluating performance of the singer, countries tend to vote for countries they are politically close. As a result the politic relationship can be examined using community detection or visualizing using colored graphs.

References

1. <https://www.kaggle.com/datagraver/eurovision-song-contest-scores-19752019?select=eurovision_song_contest_1975_2019.xlsx>
2. <https://www.kaggle.com/c/Eurovision2010/data>