

The dataset I chose is 2018_worldcup_v3.csv with link:

<https://www.kaggle.com/datasets/rezaghari/fifa-worldcup-2018>

This dataset includes information about every game in the world cup in 2018. The attributes are year(2018), date, stage, stadium, city, home team name, home team goal, away team goal and away team name and referee. Attributes I chose for the graph are home team name, home team goal, away team goal and away team name. Each team is considered as a node and the games between one another are considered as edges. In order to make a clear and efficient analysis on the graph, I made some changes to the dataset: I only kept the 4 attributes I chose as the rest is not about the graph; I changed the name of teams to double-digit numbers for easy representation. As a result, every node is actually a number representing countries. Since in a game, both sides of the team may have goals so there are two edges with different weights in one game and are both directed. For instance, for a game of Germany vs Japan with score 0:2, it's shown as edge (15,43,0) and (43,15,2) with first two representing teams and last representing goals of first team.

The graph is thus one with many countries linked with their goals to each other. What I am trying to find is the pattern of the team's goal to others using DFS. I first import all the nodes and edges from the cvs form and make it readable and form a graph. And then apply the DFS starting from Russia with number 11, it keeps going to unvisited but related neighborhoods which are teams Russia had a goal with. If there are no more unvisited, this track goes back to the former one and keeps searching other neighbors until there are no more. Even though games between each one can link all countries together somehow, some countries might not have goals for others. 0 goal is actually really common in football games. So some tracks might be stuck in half way but are still generally linked.

It turns out that all teams in the world cup are highly related and we can detect outcomes of football matches of the relationship of every team. The main reason can be the success of Russia advanced from the group making link to other groups.