FLIGHT CONNECTIVITY ACCESS TO THE MOST ANTI-SOCIAL COUNTRY IN SUB-SAHARAN AFRICA

Maths Application in Data Science Final Report

George E. Quaye gequaye@miners.utep.edu

University of Texas, El Paso (UTEP)

Department of Mathematical Sciences

Contents

1 Abstract

Airplanes are essential modes of transportation for both people and goods in today's completely globalized world. Almost everyone benefits from this cutting-edge transportation technology. As a result, having a broad awareness of global airline/airport connectivity is essential, as in knowing how many airlines fly to and from a specific airport in your jurisdiction. As a result, it is necessary to investigate or comprehend whether or not an anti-social country has any kind of flight connectivity globally or within the continent from which they originate. Given such continent, what impact does the country with the most flights connectivity has on this anti-social country and other countries on its continent. In order to address the project's main question, "Are there international flights for anti-social countries?" a graph network analysis was used. If so, where exactly are these connections? For the global network connectivity, a weighted and undirected graph network was obtained. A chosen anti-social country as such Burundi in the Sub-Saharan African geographical area, as well as its connectivity within the global network, were then extracted and observed. In addition, a table of cities, countries, and airports that offered such flight connectivity to the anti-social country was estimated. A graph of network centralities for Sub-Saharan African countries also aided in determining the influence Kenya, which had the largest airports and airlines in the region, has on the total flight connectivity network in the region. Graphical and tabular results suggested that Burundi, an anti-social country, may be reached by plane from four countries and cities: Belgium's Brussels in Europe, Uganda's Entebbe, Rwanda's Kigali, and Kenya's Nairobi in Africa's Sub-Saharan area. However, it was once again highlighted that the failure of Kenya's flight operation jeopardizes the entire flight connectivity in the Sub-Saharan area, rendering nations like Eritrea, Botswana, and Burundi non-functional in terms of flight connectivity, and affecting their economic activity at large.

2 Introduction

This analysis seeks to understand flight connectivity access to the most anti-social country in Sub-Saharan Africa. Also, the effect of the most connected flight country on the Sub-Saharan Africa continent's network.

2.1 Background/rationale

In today's completely globalized world, airplanes are critical modes of transportation for both people and products. This innovative transportation technology benefits almost everyone. As a result, having a broad understanding of worldwide airline/airport connectivity is critical, and knowing which number of airlines are going to and fro a given airport in your territory is also crucial. As a result, it is necessary to investigate or comprehend whether or not an anti-social country has any kind of flight connectivity globally and within the continent from which

they originate, and if so, what impact does a country with the most flights have on this antisocial country and other neighboring countries on this continent given their productivity and economics growth. For this study, a Sub-Saharan African country was studied. Sub-Saharan Africa is the region of Africa south of the Sahara. It includes all African countries and territories that are wholly or partially south of the Sahara, according to the United Nations. Anti-social behaviors are actions that injure others or show a lack of concern for their well-being. It has also been characterized as any sort of activity that violates another person's basic rights, as well as any behavior that is considered disruptive to society. This can be done in a variety of ways, including, but not limited to, purposeful aggressiveness, and covert and overt animosity. Given such definitions, countries such as North Korea, Burundi, and several others were identified as the most anti-social countries globally. However, Burundi in the Sub-Saharan Africa continent was considered for this study.

2.2 Objectives

This report's main objective is to respond to the following research questions.

- Are there international flights for anti-social countries? If so, where exactly?

For a given anti-social country in the Sub-Saharan Africa continent, I investigated its global flight connectivity and identify specifically which airports, cities, and countries provide these connections.

- What effect does a Sub-Saharan African country with the most flight connectivity has on the entire Sub-Saharan Africa flight network?

Using centrality measures, I selected the country with the most flight connectivity from a graph of the Sub-Saharan Africa flight network. Then I disconnected this country from the rest of the network to see how it affects the anti-social country and its neighbors.

- Is there any kind of community structure in the flight network in Sub-Saharan Africa, given the continent's flight connectivity?

In examining various network structures, it is necessary to look for communities inside them. Community detection techniques are useful in this research to locate countries with common flight connections and keep them connected.

Other subsidiary questions that were answered in this report are;

- Globally, which country has the most airports?
- Globally, which country has the most airlines?
- What is the statistical relationship between the number of airports and airlines?

2.3 Data sources

Four different datasets were used for this analysis sourced from the following databases;

- Airport database.
- Flight route database.
- Airline database.
- Country database.

All datasets and their descriptions were all downloaded and for it can assessed from Openflights.org under the Open Database license. OpenFlights is a service that allows users to map their flights across the world, search and filter them in a variety of ways, automatically calculate statistics, and share their flights and vacations with their friends and the rest of the world. The route data had 67663 observations with 9 attritubes. The airline data had 6162 observations with 8 attributes. The airport data had 7750 observations with 14 attributes and the countries data provides information on population, region, area size, infant mortality and more as indicated in the table below. It had 227 observations with 20 attributes.

2.4 Analysis approaches

In order to address the study questions, several descriptive graphs and tables were employed, noticeably graph network modeling (graphs and centralities) and traditional exploratory analysis using bar plots.

2.4.1 Data preparation

A thorough data cleaning and engineering was performed, with the goal of discovering and handling anomalous observations in all four data sets. For further analysis and reporting, the appropriate columns were chosen.

2.4.2 Descriptives

Bar plots were used for descriptive statistics and exploratory analysis of the data. This initial descriptive analysis aided in answering the subsidiary objective questions, such as globally which country has the most airports? which country has the most airlines? and the statistical relationship between airport and airline.

2.4.3 Main Analysis

The following criteria were utilized to investigate the main objective questions of this report:

- Anti-social international flight connectivity

In order to address the project's main question, "Are there international flights for anti-social countries?" a graph network analysis was used. If so, where exactly are these connections? For the global network connectivity, a weighted and undirected graph network was obtained. The chosen anti-social country, as well as its connectivity within the global network, were then extracted and observed. In addition, a table of cities, countries, and airports that offered such flight connectivity to the anti-social was estimated.

2.4.4 Other Analysis

- What effect does a Sub-Saharan Africa country with the most flight connectivity has on the entire Sub-Saharan Africa flight network?

To investigate this, a directed graph network of flight country connectivity for the Sub-Saharan Africa continent was obtained. With this, I accessed the connectivity of the anti-social country to its neighboring countries. Centrality measures such as degree, betweenness, and Eigenvector were obtained for this network to aid in answering which Sub-Saharan African country has the most airports and most airlines graphically via bar plots. Again, with such an observation made, I examined the effect of this country on the entire flight connectivity network in Sub-Saharan Africa.

- Is there any kind of community structure in the flight network in Sub-Saharan Africa, given the continent's flight connectivity?

In this research, community detection was used to find groups with similar attributes and extract groups for various purposes from the entire graph obtained for the Sub-Saharan Africa continent. The following network community detection algorithms were used;

- Louvain Community Detection:

  This method is based on modularity, which aims to maximize the gap between a community's actual and a predicted number of edges. This algorithm involves two steps processes, and during this process, the Louvain community detection algorithm reveals clusters of communities. It is quite popular due to the algorithm's simplicity of implementation and quickness.

- Walktrap Community Detection:

  Walktrap is also a community discovery method that uses random walks to assess the distance between vertices. According to the algorithm's main premise, random walks on a network tend to get stuck in heavily

connected areas that correlate to communities. Just as in the case of Louvain community detection, the quality of the partitions under Walktrap can be determined by modularity measure.

- Infomap Community Detection:

  A cost function is minimized by the Infomap algorithm. The flow caused by the layout of connections in a network is the basis for partitioning. Also, like the Louvain and the walktrap, this algorithm can also use modularity as a measure.

Modularity

In an identical network with edges set at random, the number of edges falling within groups minus the expected number is modularity. It assumes values in the range of -1 to 1 and is used as an objective function to be maximized by some given community detection algorithms. A positive magnitude indicates the probability of a community structure, zero indicates edges are simply random distribution and a negative magnitude indicates that a community structure must not be present.

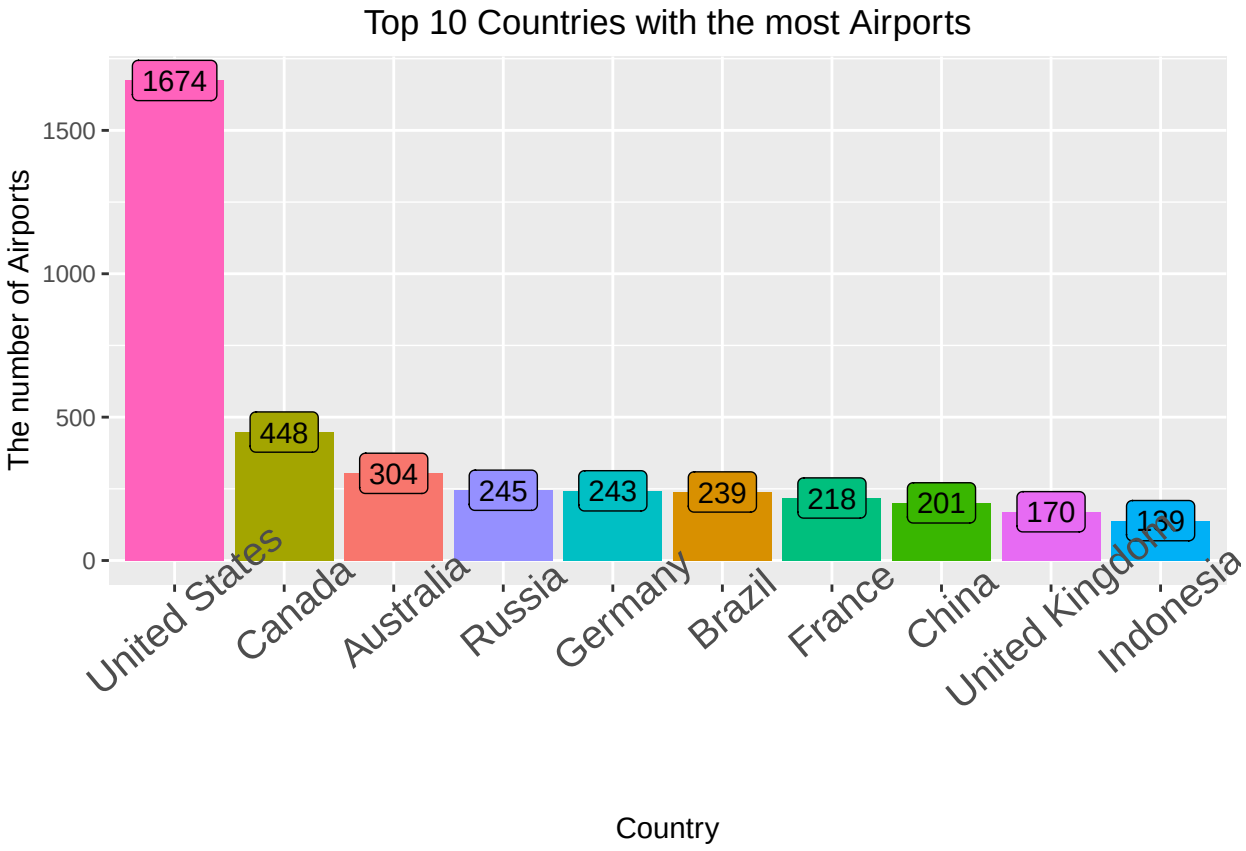
3 Results

The results of exploratory data analysis and graphical network analysis are summarized in this section.

3.1 Descriptives (Exploratory Analysis)

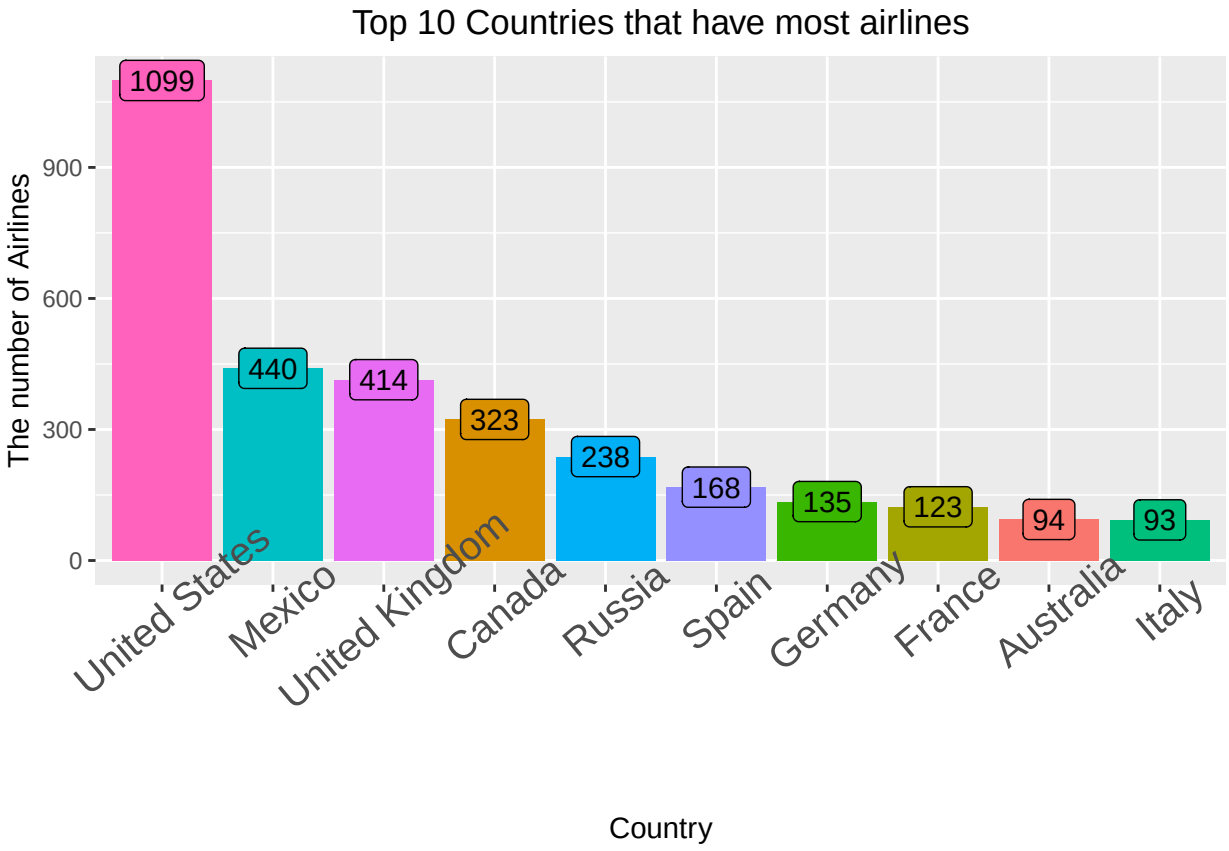
There are a total of 7750 airports in the world, and total of 6162 airlines worldwide.

3.1.1 Which country has the most Airports, globally?



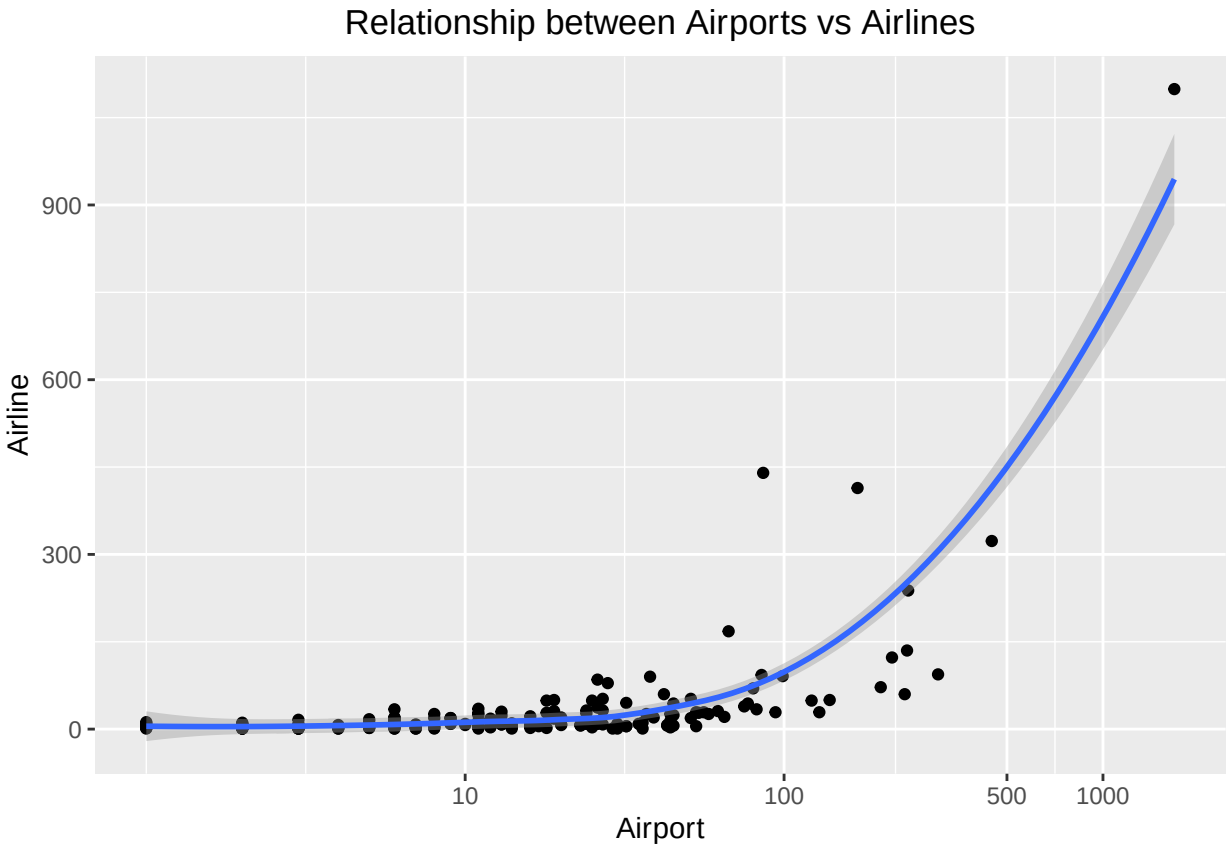
Indicated in the above plot shows that United States has by far the most airports, followed by Canada , Australia and so on.

3.1.2 Which Country has the most Airlines, globally?



Once again, it can be seen from above that the United States has a large number of airlines operating. However, certain nations, such as China and Brazil, did not rank among the top ten countries with most airlines, implying that such governments have airline restrictions.

3.1.3 Graphical or statistical relationship between the number of airports and airlines.

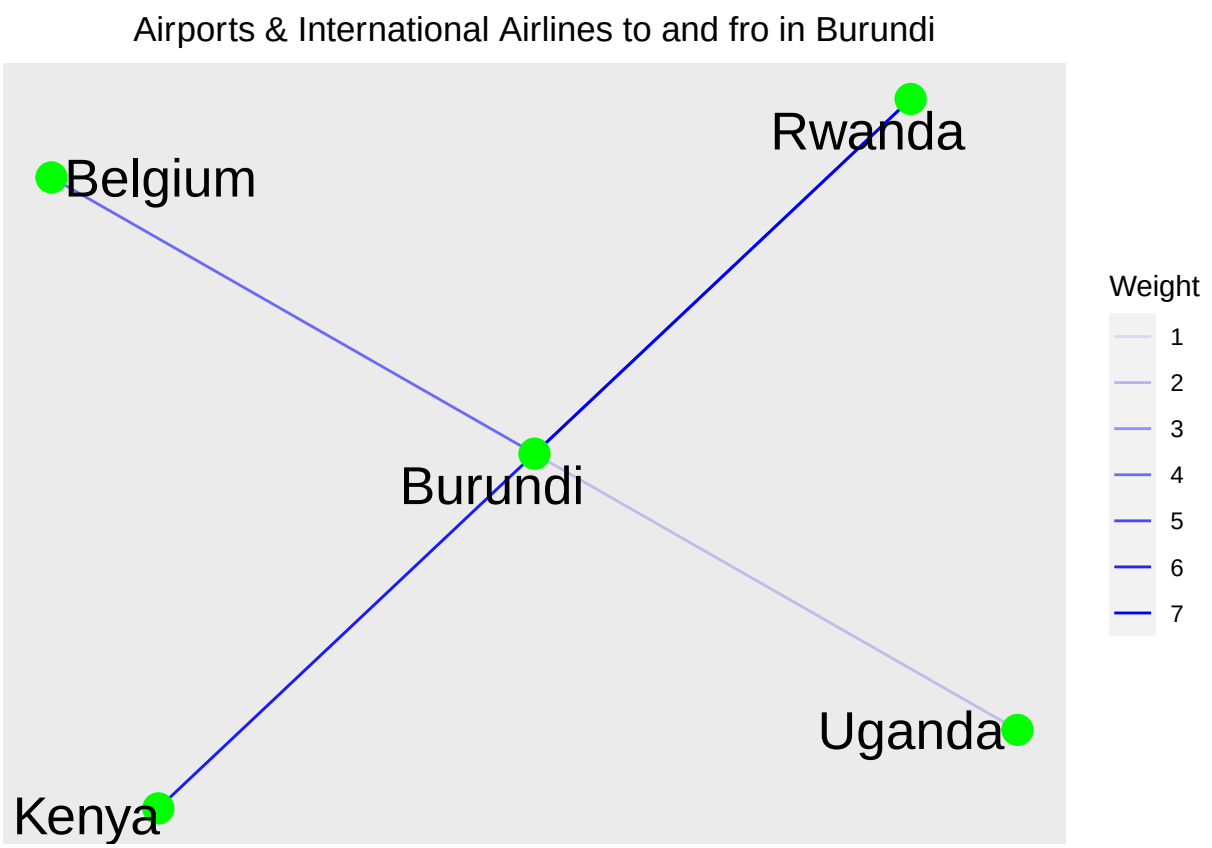


The number of airports and the number of airlines in a given country appear to have a positive correlation (relationship). This means that countries with more airports have a higher likelihood of having a large number of airlines operating in them, which makes intuitive sense.

3.2 Main Objective

Following a preliminary graphical study to identify countries with the most airlines and airports, it was discovered that apart from Africa all continents, had at least a country in the top 10 or 20. This prompted me to look into Africa's flight network as well as flight connections to the continent's most anti-social countries, noticeably Burundi in the Sub-Saharan area.

3.2.1 Are there international flights for anti-social countries?



The output of the graphical network indicates that Burundi has an airport with a few airlines operating there.

Airports that have flight access to Burundi

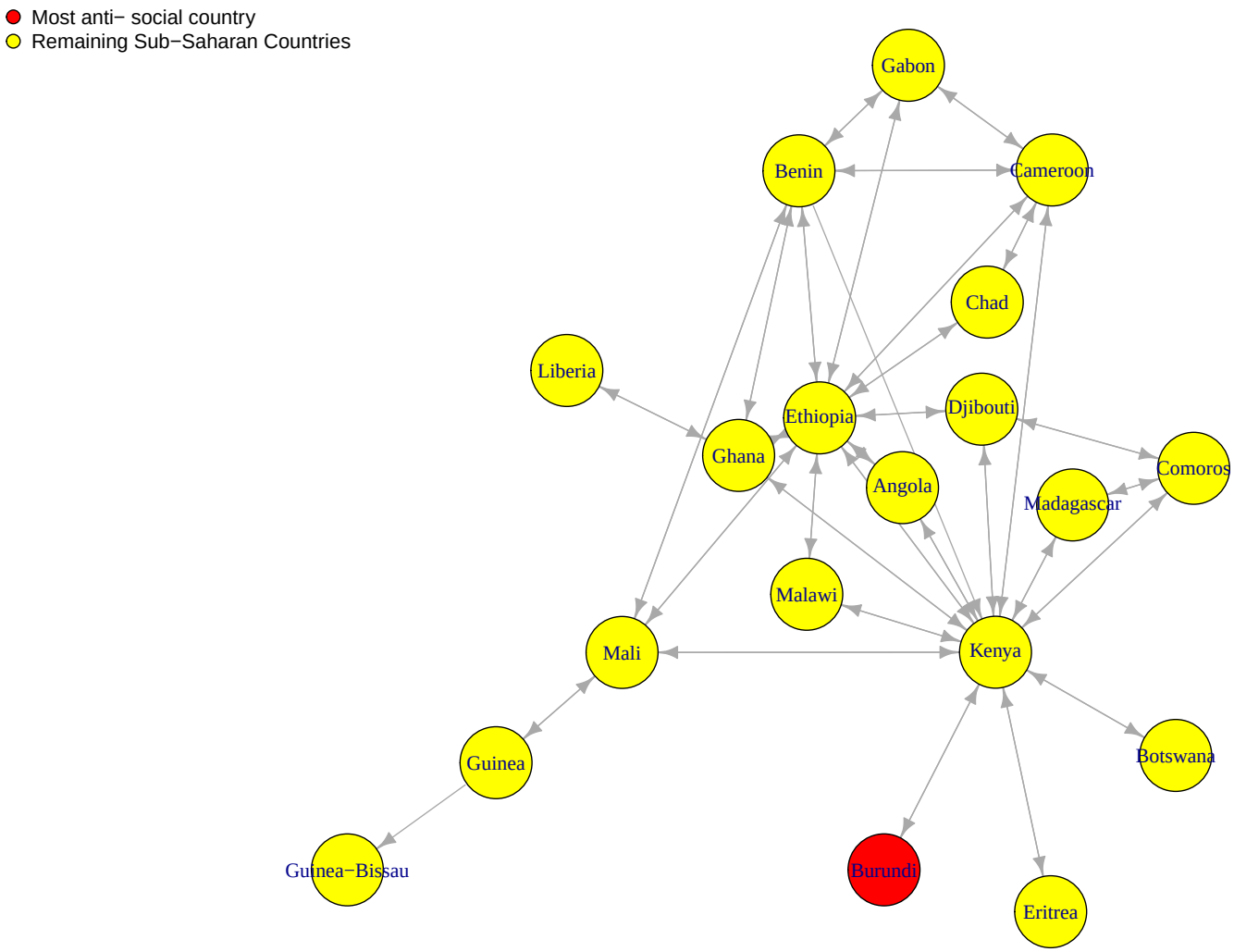
Table 1: Airports and cities with flight connection to Burundi

Airport_Name	Country	City	Latitude	Longitude
Brussels Airport	Belgium	Brussels	50.901401	4.48444
Entebbe International Airport	Uganda	Entebbe	0.042386	32.44350
Kigali International Airport	Rwanda	Kigali	-1.968630	30.13950
Jomo Kenyatta International Airport	Kenya	Nairobi	-1.319240	36.92780

It turns out that Burundi is accessible by plane from four nations and cities. Belgium's Brussels, Uganda's Entebbe, Rwanda's Kigali, and Kenya's Nairobi are the four cities in which flights to Burundi are available. The number of airlines flying to Burundi is indicated by the edge weight. As can be seen, Kenya and Rwanda have more flights to Burundi than Uganda, which only has one flight.

3.3 Other Analysis

3.3.1 Sub-saharan Africa Countries' Flights Network



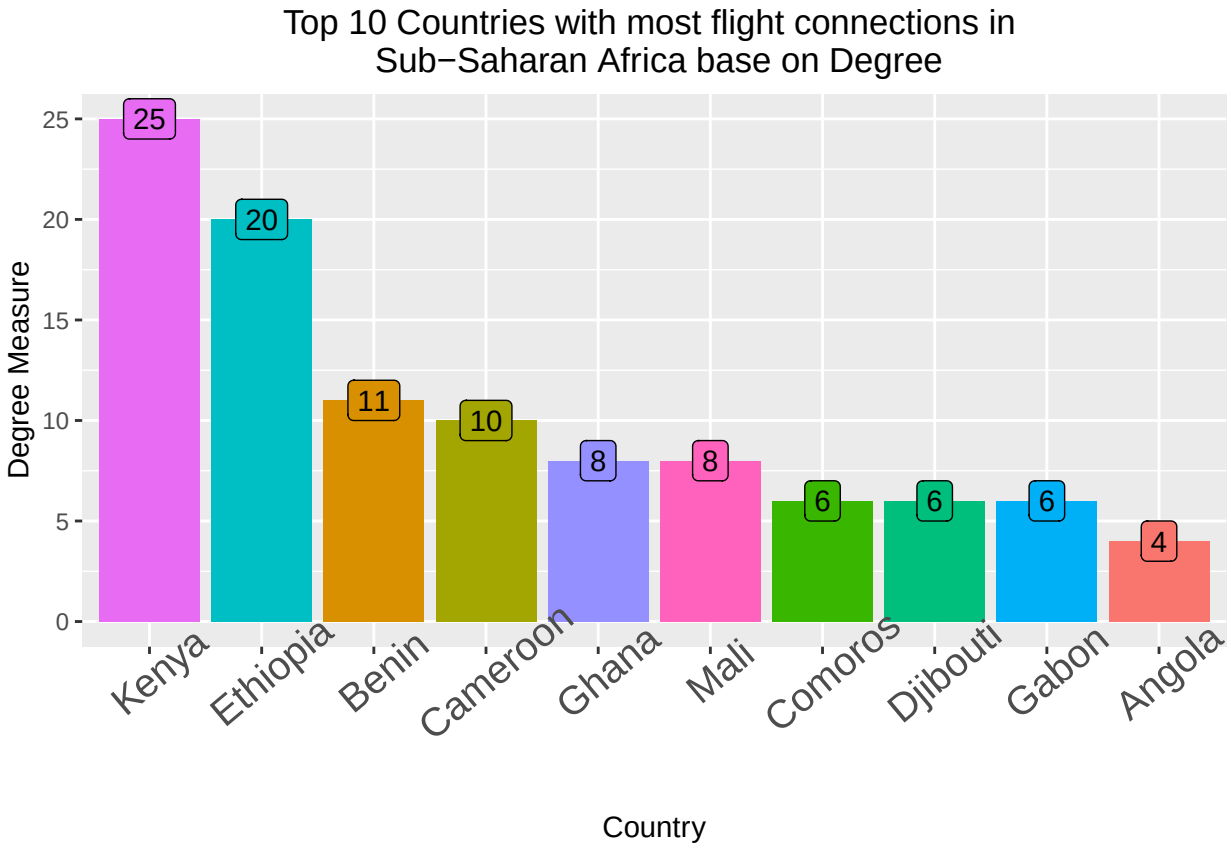
Given the graph of Sub-Saharan Africa flight connection, it is clear that Burundi only has one direct linking country, Kenya.

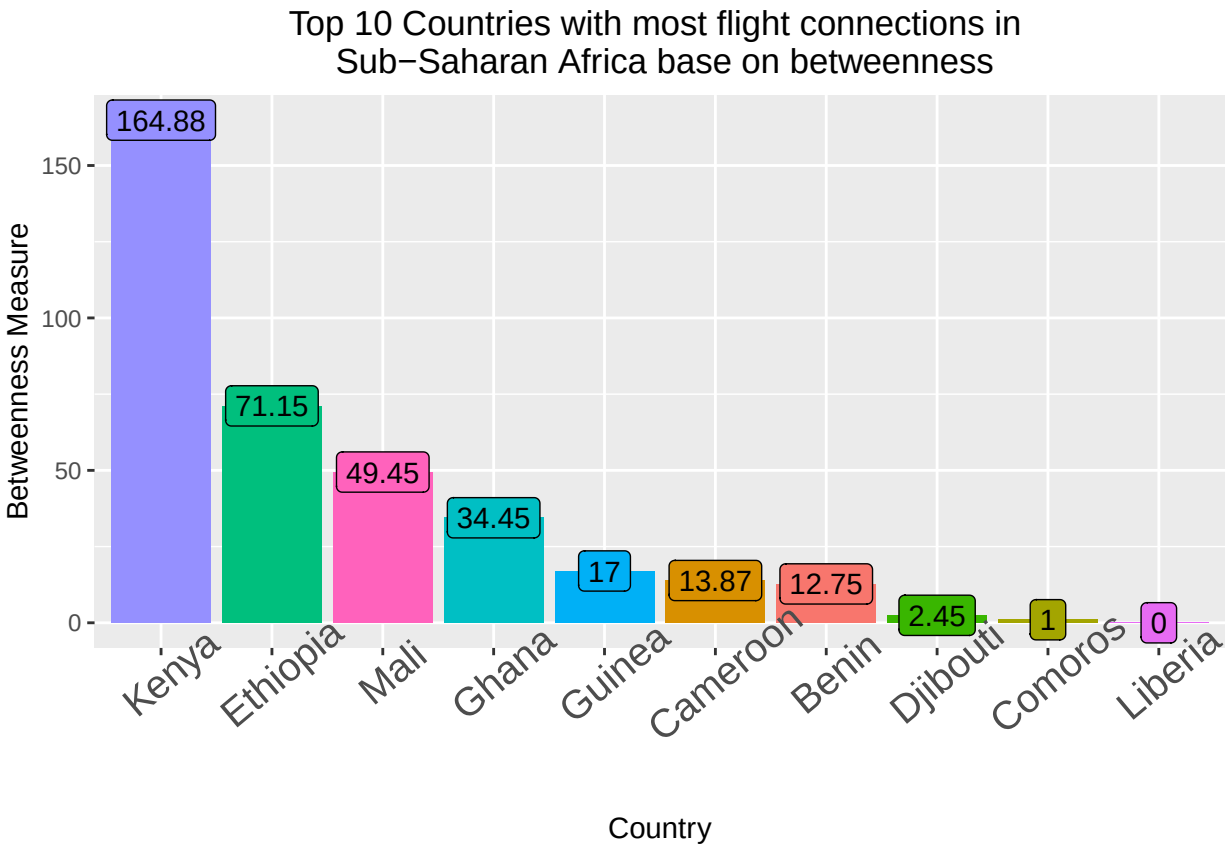
- Centrality measures and analyses

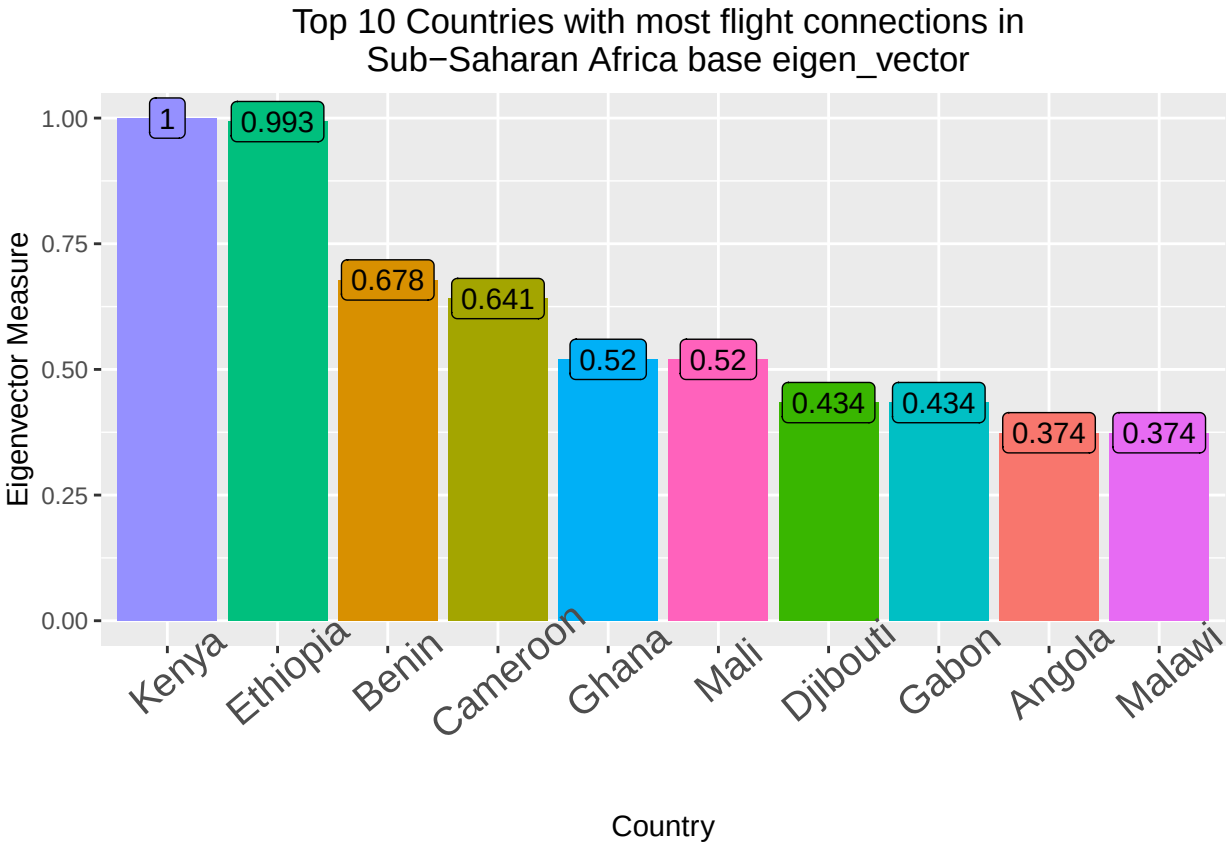
In this part, I looked at centrality measures including degree, betweenness, and eigenvector to see which Sub-Saharan African country has the most flight connections.

Table 2: Centrality measures

Country	degree	betweenness	eigen
Benin	11	12.75	0.678
Liberia	2	0.00	0.098
Ethiopia	20	71.15	0.993
Kenya	25	164.88	1.000
Mali	8	49.45	0.520
Djibouti	6	2.45	0.434
Angola	4	0.00	0.374
Ghana	8	34.45	0.520
Gabon	6	0.00	0.434
Malawi	4	0.00	0.374
Cameroon	10	13.87	0.641
Chad	4	0.00	0.307
Guinea	3	17.00	0.099
Madagascar	4	0.00	0.247
Comoros	6	1.00	0.316
Burundi	2	0.00	0.188
Botswana	2	0.00	0.188
Eritrea	2	0.00	0.188
Guinea-Bissau	1	0.00	0.009





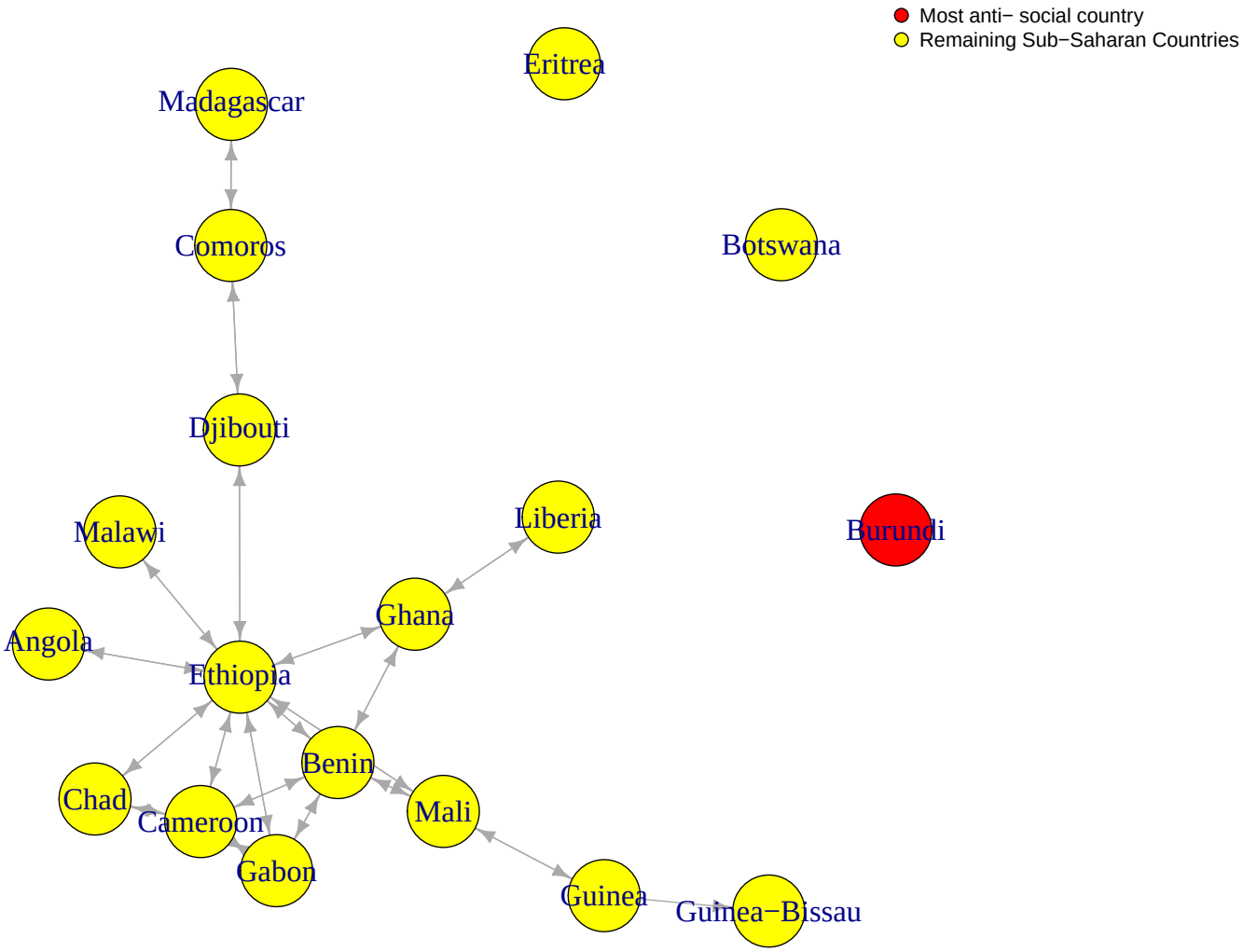


According to the three centrality metrics, Kenya has the most flight connections in Sub-Saharan Africa, followed

by Ethiopia.

3.3.2 Effect of Kenya on the entire Sub-Saharan flight network

Again, because Kenya is the most connected flight country in Sub-Saharan Africa, I considered how it would affect the entire network in Sub-Saharan Africa if it becomes non-operational.

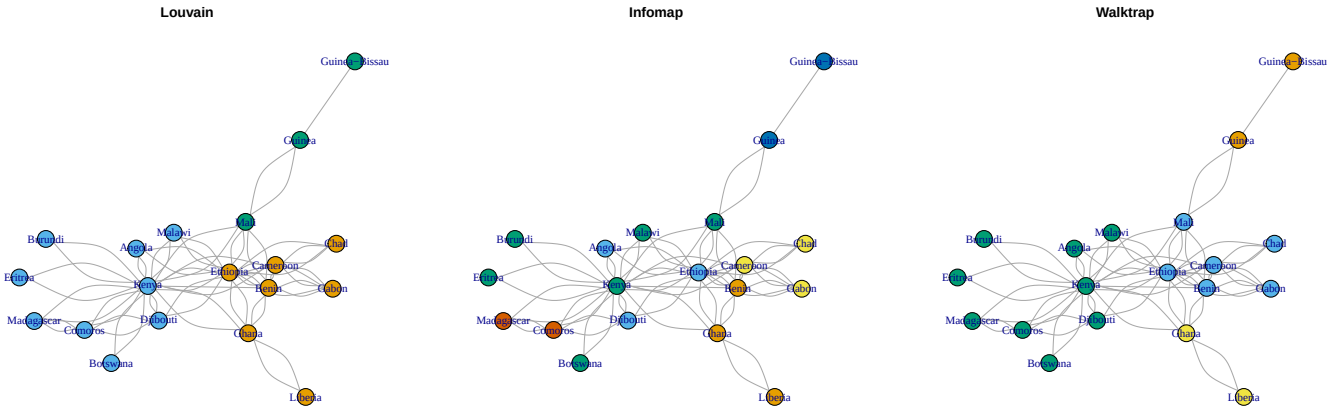


The plot above shows that if Kenya is disconnected or non-operational, countries like Eritrea, Botswana, and Burundi will be severely impacted. This implies that such countries will not have flight connectivity inside Sub-Saharan Africa; nevertheless, other countries such as Belgium and Uganda may still have connectivity to Burundi.

3.3.3 Community detection within the Sub-Saharan Africa's flight network

Communities are referred to as groups, clusters, coherent subgroups, or modules; in a social network, community identification is defined as detecting sets of nodes where the connections between nodes within a set are greater than their connections to other network nodes. In this section, I investigate by three algorithms, if such groups are

possible given the country flight connectivity of the Sub-Saharan Africa region.



Each algorithm generates a community; nevertheless, it is necessary to ensure that these clusters are not simply edge randomness.

Table 3: Algorithms and Modularity measure

Algorithm	Communities	Modularity
Louvain	3	0.283
Infomap	6	0.164
Walktrap	4	0.268

Positive numbers for each algorithm in the table above suggest the likelihood of a community structure; however, their lesser values away from 1 and closer to 0 imply that these structures may be attributable to random edge distributions in the networks.

4 Discussion

4.1 Key results

The United States was classified as the leading country in terms of airports and airlines by the descriptive, which is likely due to the United States' extensive military outposts around the world. Countries with larger regions, such as Russia and Canada, have numerous airports because they require connectivity to remote cities. The number of airports a country has is significantly influenced by its size and economy as well as the number of airports determines the number of available airlines in most cases. Again, In order to assess whether or not an anti-social country like Burundi in the Sub-Saharan Africa continents has any worldwide flight connectivity. Graphical and tabular results suggested that Burundi, an anti-social country, may be reached by plane from four countries and cities: Belgium's Brussels in Europe, Uganda's Entebbe, Rwanda's Kigali, and Kenya's Nairobi in Africa's Sub-Saharan area. Kenya was identified as the country with the most flight connectivity, or the country with the most airports and airline operations, in Sub-Saharan Africa. However, it was once again highlighted that the failure

of Kenya's flight operation economy jeopardizes the entire flight connectivity in the Sub-Saharan area, rendering nations like Eritrea, Botswana, and Burundi non-functional in terms of flight connectivity. Countries with similar flight connectivity characteristics are detected as a group. Specifically, numerical findings showed that there is a probability of group existence; however, this presence could be attributable to the randomness of the airlines(edges) inside such structured communities.

4.2 Interpretation

Burundi was ranked among the top ten least developed countries in the world by the United Nations (UN)(Eijdenberg and Montfort 2017). As previously discussed in this study, the extent of the connectivity, or the number of airports and airlines in airlines, is mostly determined by the country's economic development and geographical size. This provides more proof of Burundi's limited connectivity in the Sub-Saharan region and worldwide. It can also be argued that Burundi is not developing at a higher rate than other Sub-Saharan African nations such as Kenya and Ethiopia because of insufficient flight connectivity for the transportation of products and services to maximize productivity and growth of the economy.

4.3 Generalisability

Even though Botswana and Eritrea are not among the top 20 most anti-social countries according to this study, however in terms of flight connectivity, these countries stand a chance of having the same faith as Burundi in terms of productivity and economic development should Kenya within the Sub-Saharan Africa continent go out of flight connectivity operation. The most recent pandemic epidemic, in which countries closed their borders to avoid flight operations, had a significant impact on Burundi, Eritrea, and Bostwana, as no airlines were operating in such underdeveloped countries.

5 Reference

Eijdenberg, Emiel L, and Kees van Montfort. 2017. "Explaining Firm Performance in African Least Developed Countries: Evidence from Burundi and Tanzania." *International Review of Entrepreneurship* 15 (3).

6 Appendix

6.1 Codes

Codes are available upon request!