Beyond Passenger Numbers: Uncovering Hidden Connectivity in the Global Airline Network Using Network Centrality Metrics

A Project Report submitted to Birkbeck, University of London for the degree of M.Sc. in Data Science.

By
Mohd Nafees
Supervised by Dr. Felix Reidl
Department of computing and Mathematical sciences
11th February 2025

Table of Contents

Abstract

This study explores the hidden connectivity of the global airline network using different network analysis methods. Instead of just counting passenger numbers or flight volumes, it looks at centrality measures like degree, betweenness, eigenvector, and harmonic to see what roles airports have in air travel. This analysis uses the OpenFlights dataset. It provides detailed information on airport locations and flight routes. A directed network is constructed to mirror actual flight connections and then Data cleaning and normalization were done carefully so that metric comparisons are fair.

The findings show an unexpected twist: Big international hubs like Frankfurt and London Heathrow lead in connectivity. But some smaller regional airports—like Papa Westray and Westray in the Orkney Islands—emerge with surprisingly high harmonic centrality. This is mostly because the airports are close to each other, which produce large reciprocal values and improve travel efficiency. In other words, these local airports save a significant amount of travel time, so they become important even if they only handle a few flights.

The study shows that airport importance is not determined by traffic volume. It is more about how well it connects with the rest of the network. These findings point out that in planning transportation, better route design can be just as important as adding more flights. Investigating different centrality measures gives a broader view of global air travel. It highlights that it is not just the huge hubs that matter in network connectivity, but also the small, well-placed nodes that play an important role.

Acknowledgements

I would like to express sincere gratitude to my supervisor, Dr. Felix Reidl, for his invaluable support and guidance throughout this research. His clear direction and insightful feedback helped shape the study, offering practical suggestions that made a real difference in how the work evolved. Dr. Reidl’s willingness to discuss ideas and his patient explanations helped clarify complex points and kept the project on track. His mentorship has not only enhanced the quality of this research but also made the journey a rewarding learning experience.

Chapter 1: Introduction

1.1 Background

The global airline network is a huge part of modern connectivity, economies, and cultures, moving billions of passengers every year (ACI, 2021). But figuring out what makes an airport “important” in this massive system is not so simple as it seems. Looking at passenger numbers and flight volumes can give a general idea, but they do not always show the full picture. Some airports might not have huge traffic, but they still play a crucial role—whether it is by linking continents or making regional travel easier. This is where network theory comes in. It offers tool like centrality measures to uncover hidden structure of air travel and how airports impact the global airlines system.

1.2 Research Objectives and Approach

This study examines the global airline network using the OpenFlights dataset (OpenFlights, n.d.) to explore how different measures of network centrality can reveal the unique roles that various airports play. Instead of just focusing on common stats like passenger and flight numbers, airport size, this study takes a deeper look at network-based metrics—including degree, betweenness, eigenvector, and harmonic centralities—to better understand how airports are connected and what functions they serve.

This research intends to uncover the distinct patterns within the airline network. One approach used is degree centrality, which counts an airport’s direct routes and highlights major hubs such as Frankfurt and Charles de Gaulle, known for serving many destinations (Fraport, 2020; Paris Aéroport, 2021). Yet a high number of connections does not necessarily equate to overall network influence. To address this, betweenness centrality is measured to identify airports that act as critical bridges between different parts of the networks. Our analysis identified locations like Keflavik in Iceland and Anchorage in the United States surfaced as key transit points, even though they do not have the huge number of direct routes. These results align with what past studies have found that geographical location plays a big role in making a network stronger and more resilient (Bovet and Makse, 2022).

Eigenvector centrality adds another layer by looking at how important an airport's connections are. This measure reveals how an airport can be important not just because of how many flights it has, but because it is linked to other highly influential hubs. In some cases, an airport's network position matters just as much as the number of direct routes it serves. For example, London Heathrow may not have the highest number of connections, but it connects with other major airports in the world that elevate its score in eigenvector centrality measure (Heathrow Airport, 2022).

At the same time, harmonic centrality looks at how easily an airport can connect to others in the network. What is surprising is that our analysis found small regional airports, like Papa Westray and Westray in the Orkney Islands, ranked quite high when it comes to travel efficiency. Even though they are small and do not have many routes, they still play a key role in reducing travel time. This challenges the usual idea that an airport's importance is based just on the number of flights it has. For some communities, the role of an airport is determined more by how much time it can reduce in travel time than by the number of destinations served.

Using Python and NetworkX to analyse the OpenFlights dataset made it possible to take a closer look at these different centrality metrics. It can be surprising to see the diversity of roles played by different airports. While big international hubs clearly dominate when looking at degree centrality, things start to look a lot more interesting when other measures come into play. The data clearly indicate that connectivity is not a one-dimensional concept; rather, the importance of an airport can vary dramatically depending on whether we consider direct connections, strategic positioning between regions, the influence of its network of neighbours, or travel efficiency.

This research takes a closer look at how different centrality measures compare, both to each other and to real-world airport data. For example, Frankfurt ranks highest in degree centrality because of its huge flight network (Fraport, 2020). But that does not tell the whole story. Keflavik, for instance, ranks much higher in betweenness centrality because its location makes it an important stopover point (Isavia, 2021). Similarly, London Heathrow's strong performance in eigenvector centrality underscores its role as a connector among influential hubs, despite not having the highest raw count of connections. These findings make it clear that no single metric can fully explain how airport connectivity works. It is a lot more complex than just looking at one measure. Previous research has also pointed this out (OAG, 2024; Wasserman and Faust, 1994)

The analysis process not only deepened my understanding of network theory but also gave insight into the practical implications for regional connectivity. It was fascinating to discover how airports that might seem insignificant, such as those in the Orkney Islands, can dramatically enhance local mobility by cutting down travel times—a benefit that goes well beyond mere passenger numbers. This experience really made me realize how network metrics and real-world outcomes are deeply connected, prompting me to advocate for more integrated approaches in transportation planning, rather than just looking at things in isolation.

This research aims to take a detailed look at the global airline network, showing that an airport's true importance lies in a blend of quantitative and qualitative factors. Instead of relying on just one metric, this study uses multiple centralities measures and compares them with real-world data to give a more complete understanding of air transport connectivity.

Chapter 2: Literature Review

2.1 Network Theory and Transportation

Network theory is often used to study complex systems, and we come across a lot of studies that apply it to airline networks. It is fascinating how we can learn so much from connections rather than just looking at route counts. Newman (2010) explains that networks help us understand how things are connected, not just by counting links but by looking at the bigger picture. He also points out that looking at how nodes interact can reveal a lot more about the system than just raw numbers. Barthélemy (2011) adds that spatial factors and economics also shape these networks. This helps us to understand how geography and economic factors can shift the entire shape of an airline network. That is why some airports, even if they do not have tons of flights, can still be important because of their location or regional role.

Researchers like Guimerà et al. (2005) have tried to figure out which airports matter most by looking at centralities such as degree and betweenness. Degree centrality tells us how many direct routes an airport has—useful, but it does not capture everything about how that airport ties various parts of the network together. They also point out that betweenness centrality, which looks at how an airport links distinct parts of the network, is also crucial.

Betweenness centrality is useful as it helps identify airports that play a key role in passenger transfers, even if they do not have a lot of direct routes. Then there is eigenvector centrality. Newman (2010) explains that this method does not just count direct connections but also considers how well-connected an airport's neighbours are. This means an airport can be more important not just because of how many routes it has, but also about who it is connected to—especially if it links to other well-connected airports. Another key useful measure is harmonic centrality, which Boccaletti et al. (2006) describe to calculate how efficiently an airport can connect to all others in the network. Harmonic centrality is quite interesting because it can show how even a small airport can help improve overall travel efficiency in a region.

2.2 Research Gap and Critical Perspective

What can be noticed in reading these different works is that many stick to just one or two centralities, rarely combining them to get a complete picture. That is where we see a gap. Airline networks are so complex—no single metric is enough to grasp why certain airports matter in ways that might not show up in passenger figures. Some researchers, like Guimerà et al. (2005) and Newman (2010), focus heavily on degree or betweenness alone, and it leaves open the question of how airports rank when we layer in eigenvector or harmonic aspects. Very often external influences like local economics or geographic barriers, pointed out by Barthélemy (2011), often get glossed over. If we merged that context with metrics we would see a richer understanding of how airports function day to day.

Overall, these studies have laid a good foundation, but we need more projects that consider multiple centrality measures alongside real-world data on airport performance and local conditions. A single number like degree can make a big hub look dominant, but sometimes a smaller airport is the true lifeline for a region, especially if betweenness or harmonic metrics show it shortens travel immensely.

From existing research, we could gain new insights by weaving quantitative methods together and adding a bit of qualitative detail—interviews with airport managers or airlines—to see what truly drives connectivity and what just looks big on paper. That is the direction of this research aims to take: using several measures at once and putting them side by side with practical evidence to see which airports really hold the network together based on different measures.

2.3 Recent Developments in Centrality Measures.

The works of Newman (2010) and Barthélemy (2011) provide a solid base for understanding network theory, but more recent work shows there is more to the story. For example, Ugurlu 2022b) looks at “Isolating Centrality,” in a way that provide a new perspective on how some nodes contribute uniquely to a network. Whereas Borgatti and Everett (2006) also compare different centrality measures and reveal weaknesses in traditional approaches. All such new studies contribute towards widening the picture, especially when dealing with complex systems such as airline networks.

The methodology in this study follows ideas in the comparative analysis at Springer. That study argues good centrality measures should stay the same when the network changes, and they should hold up for networks of all sizes. By linking the use of degree, betweenness, eigenvector, and harmonic centrality to these ideas, a more reliable basis for the method is developed. Using several measures at once is seen to better capture the real function of airports have in a network.

All in all, these newer angles add extra depth to our discussion and support why we used these methods. Classic studies and more recent work both shed light on centrality and mixing them can give us a stronger picture of how airports link up and interact.

Chapter 3: Data and Methodology

3.1 Data Acquisition and Cleaning

This analysis is based entirely on the OpenFlights dataset, which provides detailed and well-documented information on airports and flight routes. In the context of airline networks, where each edge is a real flight connection, centrality measures become very meaningful in understanding the importance of airports. For example, betweenness centrality can highlight an airport's importance as a transfer point, a concept that is considered more tangible in transportation networks than in, say, social networks. The following section outlines the step-by-step methodology implanted in our data analysis– starting from data acquisition and cleaning. It also explains network construction, calculation of centrality measures, normalization, correlation analysis, and visualization –by providing critical insights and relevant mathematical formulations.

There are Two CSV files that form the basis of the analysis sources from Openflights data sets. One file contains node data – that is, information about airports, including attributes such as the airport name, city, country, IATA and ICAO codes, latitude, longitude, altitude, and other variables. The other file contains edge data, which details the flight routes including source and target indices, flight distance, airline, and equipment used. The dataset is chosen for the analysis because it is produced by OpenFlights, a source known for its clarity and high data quality. This is particularly important since many network studies suffer from uncertain data origins, making it difficult to assess reliability (OpenFlights, n.d.).

Pandas is to load and clean the data. It is a python help to work with a large dataset including data cleaning, analysing, exploring and data manipulation. Initially, all column names are standardized by stripping extra spaces, removing any leading characters (such as "#"), and converting them to lowercase. Vague labels (for example, "id" and "source") are renamed to more descriptive labels like "openflights_id" and "src_index" to ensure consistency when matching nodes with edges later, which is also considered practice in data analysis. A snapshot of the nodes data can be seen below:

NodeID	AirportID	AirportName	CityName	CountryName	IATACode	ICAOCode	Latitude	Longitude	AltitudeFT	TimezoneUTC	DSTFlag	PositionVector
1	1	Goroka Airport	Goroka	Papua New Guinea	GKA	AYGA	-6.08169	145.391998	5282	10.0	U	[6.29,1.38]
2	2	Madang Airport	Madang	Papua New Guinea	MAG	AYMD	-5.20708	145.789001	20	10.0	U	[6.31, 1.0]

The edges file has 66,771 rows and 8 columns, is similarly organized, which is critical for correct network analysis.

3.2 Network Construction

A directed graph was built using NetworkX, where each airport is treated as a node that carries details like its name, city, country, latitude, and longitude. Flight routes, on the other hand, are modelled as directed edges carrying a weight corresponding to the flight distance. It is important to use a directed graph here because, in real life, just because a flight going from one airport to another does not mean there is automatically return trip on the same route or with the same frequency. This one-way method really helps capture the actual, sometimes uneven, flow of air traffic. (Freeman, 1978; Da Rocha, 2009).

The reason NetworkX was chosen because it is a well-known and trusted library, that is both reliable and widely used in academic research and industry settings. It is convenient functions such as DiGraph (), add_node (), and add_edge () offer a simple yet robust framework for building and exploring complex networks. The library is reliable and compatible to work with other Python tools like Pandas and Matplotlib for data plotting. These attributes make Network suitable for our analysis. The final graph in our analysis has over 3,200 nodes and about 36,900 edges which accurately show the huge connectivity of the global airline network.

This method makes sense because it matches how air travel works – where the direction of flights matters. It also lines up with previous research done in network analysis. The directed graph gives a realistic interpretation of how airports work as both starting points and destinations which make it easy to identify key hubs and transfer points. In short, the use of a directed graph and NetworkX in combination offers a reliable method for studying the complex structure of airline networks.

3.3 Calculation of Centrality Measures

In airline networks, it is important to figure out which airports have the biggest impact. One way to do this is by using centrality measures, which help us see how connected each airport is.

3.3.1 Degree Centrality

Degree centrality is simply a count of the number of direct flight connections an airport has. In other words, it measures how many other airports a given airport is directly connected to. The formula for degree centrality is written as:

$$C_D(v) = \frac{\text{degree}(v)}{n - 1}$$

Here, degree(v) represents the number of direct links (or flights) from airport v and n is the total number of airports in the network. Degree centrality is calculated using Python’s NetworkX library with the function nx. degree_centrality(G).

In the context of airline networks, an airport that shows a high degree centrality is usually a major hub. This means it has many nonstop routes, which not only makes it a crucial connection but demonstrate its overall importance in keeping the network running smoothly (Freeman, 1978). Understanding these connections can be useful, whether you’re planning new routes or trying to improve the efficiency of an existing network.

3.3.2 Betweenness Centrality

we computed the betweenness centrality of airports to see how much each one serves as a bridge connecting other airports. Basically, betweenness centrality measures how many of the shortest paths between any two airports pass through a given airport. The mathematical expression we used is:

$$C_B(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

Here, $\sigma_{st}$is the total number of shortest paths between airports s and t, while $\sigma_{st}(v)$ is the number of those paths that pass-through airport v.

To compute this in practice, Python’s NetworkX library with the function nx. betweenness_centrality (G, weight='weight') is employed. We used the ‘weight’ parameter to take flight distance into account when calculating paths. This allowed shorter routes to be treated as the “shortest” paths in the network.

This method helped us spot airports that serve as key intermediaries in the in the network. Even if an airport does not have a high number of direct connections, a high betweenness score still indicates that it plays an important role in linking different regions together. These insights are helpful for assessing how efficient the airline network and understanding how disruptions at one airport might affect the whole system (Freeman, 1978).

3.3.3 Harmonic Centrality

In our study, we decided to use harmonic centrality to figure out how quickly and efficiently an airport can reach all the others in the network. Basically, instead of just counting the number of flights, this measure sums up the inverses of the shortest distances between airports. The equation we worked with is:

$$C_H(v) = \sum_{u \neq v} \frac{1}{d(v,u)}$$

where $d(v,u)$ is simply the shortest path distance between airport v and airport u. We calculated this using Python's NetworkX function nx. harmonic_centrality (G, distance='weight'), which means that we factored in the actual flight distances as weights. This approach makes sure that airports with many short, efficient connections end up with a higher score.

What really struck was that harmonic centrality can shed light on the efficiency of smaller, regional airports. Even if these airports do not boast a massive number of direct flights, they can still play a key role by reducing the overall number of stops needed for a journey. This means that, in a network where every minute counts, some airports become surprisingly important for keeping travel routes short and effective (Da Rocha, 2009).

Using this measure allowed us to look beyond the obvious hubs and see the subtle ways in which the network is structured. It gives a fresh angle on how we understand connectivity in airline networks, proving that even less prominent nodes can sometimes wield a surprisingly large influence.

3.3.4 Eigenvector Centrality

Eigenvector centrality is an interesting measure that goes beyond just counting direct flights. Instead, it weighs the importance of an airport by also considering how influential its neighbours are. In simple terms, even if an airport does not have a massive number of direct connections, it might still score high if it's linked to other key hubs. The formula we use is:

$$C_E(v) = \frac{1}{\lambda} \sum_{u \in N(v)} C_E(u)$$

Here, $N(v)$represents the set of airports directly connected to airport v and λ is the largest eigenvalue of the adjacency matrix. Now, one of the tricky parts was that the method we used – the nx. eigenvector_centrality_numpy () function in Python –

requires the network to be fully connected. Since our airline network is not completely connected, we had to convert it into an undirected graph and focus only on the largest weakly connected component. This is not ideal because it might leave out some peripheral airports, but it was necessary to make the calculations work.

This measure is quite interesting because it reflects not just local influence but also the broader integration of an airport into the global system. In practical terms, if a regional hub connects mainly to major airports, it can still be very influential despite having fewer routes overall. However, this approach does have its limits – by only considering the largest connected part, we might overlook how some isolated or less connected airports could play a role in regional connectivity (González, 2009).

3.4 Geographic Distance Calculation

To analyse the efficiency of airport connectivity, it is important to consider not only the number of flight routes but also the physical distances between airports. In our study, we compute the great-circle distances between airports using the haversine formula. This formula calculates the shortest distance over the Earth's surface given the latitude and longitude of two points, and is expressed as:

$$\boldsymbol{d = 2R\arcsin\left(\sqrt{\sin^2\left(\frac{\phi_2 - \phi_1}{2}\right) + \cos(\phi_1)\cos(\phi_2)\sin^2\left(\frac{\lambda_2 - \lambda_1}{2}\right)}\right)}$$

Where $\boldsymbol{\phi}_1$ ***and*** $\boldsymbol{\phi}_2$ the latitudes $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ the longitudes (converted to radians) of the two airports and is the Earth's mean radius (6371 km) (Sinnott, 1984).

The OpenFlights dataset (OpenFlights, n.d.) provides the necessary geographic coordinates for each airport. We cleaned the dataset by removing extra spaces, stripping any leading "#" characters, and converting column names to lowercase. Then, using the "name" column, we filtered for the Orkney Islands airports: Kirkwall, Papa Westray, Westray, Sanday, and Stronsay. The haversine formula was then applied to compute the distances between these airports. Our computed distances are as follows:

• **Kirkwall → Papa Westray:** 43.8 km

• **Kirkwall → Westray:** 43.7 km

• **Kirkwall → Sanday:** 37.5 km

• **Kirkwall → Stronsay:** 26.6 km

• **Papa Westray → Westray:** 2.8 km

• **Papa Westray → Sanday:** 21.6 km

• **Papa Westray → Stronsay:** 26.3 km

These distances illustrate the tight proximity among the Orkney Islands airports. When computing harmonic centrality—which is defined as the sum of the reciprocals of the distances to all other nodes—even short distances (such as 2.8 km between Papa Westray and Westray) yield high reciprocal values, thereby significantly boosting the harmonic centrality of these nodes. In contrast, global hubs such as Frankfurt or Amsterdam cover much larger geographical areas, leading to lower reciprocal values despite their higher connectivity.

3.5 Normalization and Correlation Analysis

Each centrality measure comes out on its own scale – some may range from near zero up to 1, while others have a wider spread. To compare them directly, all measures are normalized to a common 0–100 scale. The formula used is:

$$M_{\text{norm}}(v) = \frac{M(v) - \min(M)}{\max(M) - \min(M)} \times 100$$

In other words, for every airport, the original measure is adjusted by subtracting the minimum value found among all airports and then dividing by the range (which is the maximum value minus the minimum). Multiplying by 100 scales this ratio to a percentage-like scale from 0 to 100. This is particularly useful in our study because it lets us compare, say, degree centrality and betweenness centrality directly, even though they were computed on very different raw scales. As Jolliffe (2002) notes in his study, normalization is a key step when working with diverse datasets to ensure that the comparisons are fair and meaningful.

Once all measures are normalized, the next step is to understand how they relate to each other. To do this, Spearman rank-correlation is computed. Unlike Pearson correlation which compares actual values, Spearman's method compares the rank order of the data. Its formula is:

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

Here, represents the difference in ranks for each airport between two different centrality measures, and is the total number of airports. For example, if an airport is ranked 1st by degree centrality and 2nd by betweenness, the difference in rank for that airport is 1. This method is especially robust when the data do not follow a normal distribution, making it suitable for network measures (Cattell, 1966).

A practical example: if Airport A is ranked 1st for degree and 1st for betweenness, for that airport is zero, contributing nothing to the sum in the numerator. If many airports show similar ranking orders across two measures, the Spearman coefficient will be close to 1, indicating a strong positive correlation. Conversely, if the rankings differ greatly, the coefficient will drop towards -1 or 0, suggesting that the two measures are capturing different network properties.

A Seaborn heatmap is generated based on the computed Spearman correlation matrix for visualisation. The heatmap provides a clear intuition of which centrality measures tend to overlap and which do not. This visualization is important for determining if certain measures might be redundant meaning capturing similar information or if each offers unique insight into the network structure.

This step-by-step normalization and correlation analysis ensures that the all-centrality measures are comparable, and it can give insight into the similarities or differences among them, thereby providing deeper insight into the structure of airline networks.

3.6 Visualization and Verification

To bring the analysis to life, several visualizations were created to help illustrate the network structure and the relationships between centrality measures. One of these is a Seaborn heatmap that shows the Spearman correlation among four centrality measures. In this heatmap (see Figure 1 below in section 4, spearman correlation of centralities), the colour gradients indicate the strength of the correlations: The heatmap uses colour gradients. Darker colours mean stronger correlations, and lighter colours mean weaker correlations.

This visualization makes it easy to see that while degree, betweenness, and eigenvector centralities correlate well, harmonic centrality stands apart—confirming its distinct role in capturing travel efficiency.

Another set of visualizations involves mapping the geographic distribution of the top-ranked airports based on each measure. With Geopandas and Shapely, a map of the entire world was produced with airports displayed as points. In such a map (for example plotted below in Figure 2.2), the most centrally ranked 100 airports (for a chosen centrality metric) have been represented in red, and the most centrally ranked 10 have labels placed over them.

Figure 2.2 : Top 100 by Betweenness Centrality

This map turns abstract centrality scores into real-world locations, mapping network connections gives useful insights into the structure and strength of complex systems. It makes the network easier to understand. These visualization examples not only aid in interpreting the numerical results but also help communicate the broader story of network connectivity and efficiency within the airline system.

So, there is this function created, called verify_top_10_airports_and_suspects, that enable re-check our results. Basically, it prints out the top airports based on each normalized measure and gives extra details—like how many flights come in and go out, plus the overall degree in our NetworkX graph. It also does some random name searches for potential hubs, for example, Heathrow, John F Kennedy, or Dubai, and then shows their centrality scores and rankings. This further verification is really important to make sure our network indicators match up with what we already know about airline operations.

To ensure data integrity by cross-checking the node IDs in both the nodes and edges files, and ensure every flight is properly recorded. Performed few random spot checks, like verifying that the details for Dubai International Airport are the same in both the CSV file and the graph. On top of that, eigenvector centrality was recalculated on the largest connected subgraph to avoid any problem from isolated nodes.

3.7 Justification for Using the OpenFlights Dataset

The choice to use only the OpenFlights dataset rests on two key factors. First off, in transportation networks, every edge really represents an actual flight route. That means things like betweenness centrality are straightforward—a high value basically shows that a particular airport is a key transfer point. Second, the dataset from OpenFlights is notably well-documented and of excellent quality. Unlike many other network datasets that might come with sketchy details or unclear sources, OpenFlights offers clear, reliable information (OpenFlights, n.d.).

Overall, this methodical approach—using tools like Pandas, NetworkX, Matplotlib, Seaborn, Geopandas, and Shapely—helps break down the global airline network into understandable pieces. With all the checks and verifications in place, the analysis feels both trustworthy and insightful.

3.8 Limitations

The analysis depends solely on the OpenFlights dataset, which might not tell the whole story. There are a few gaps—regional routes may be underreported, some airport codes might be outdated, and small airports could be missing entirely. Such gaps can skew centrality results (Kolaczyk, 2009).

In building the network, converting it into an undirected graph and focusing only on the largest connected component was necessary for computing eigenvector centrality. That said, this approach leaves out some isolated nodes that might still matter locally. Also, treating every flight as equal ignores differences in flight frequency or passenger load—a simplification that might not hold up in all cases (Opsahl et al., 2010)

Normalizing the centrality measures onto a 0–100 scale makes comparisons easier, but it also smooths over some details. For instance, harmonic centrality is good for showing travel efficiency, yet it might miss nuances like seasonal route changes or overlapping connections. This is a common challenge with complex networks (Borgatti, 2005).

Finally, the dataset is only a snapshot of a constantly evolving airline network. The metrics we see now could shift with new data. Future work should consider using multiple data sources and possibly weighted models to capture a more complete picture.

Chapter 4: Discussion and Result

4.1 Broad Overview of Measures

When analysing a complex network, such the global airlines network. It is important to have overall understanding of the data. For this reason, we start by examining the distribution and correlation of different centrality measures from a broad perspective before focusing on individual measures. First, we evaluate overall structure to investigate how centrality scores are distributed, and how they relate to each other. We started with the investigation of four centrality measures such Degree, Betweenness, Eigenvector, and Harmonic in the OpenFlights global airline dataset. Our main objective was to see what airports they rank as the 'most important'. Before looking into specific airports list, we want to observe who these measures look across all nodes and if they overlap.

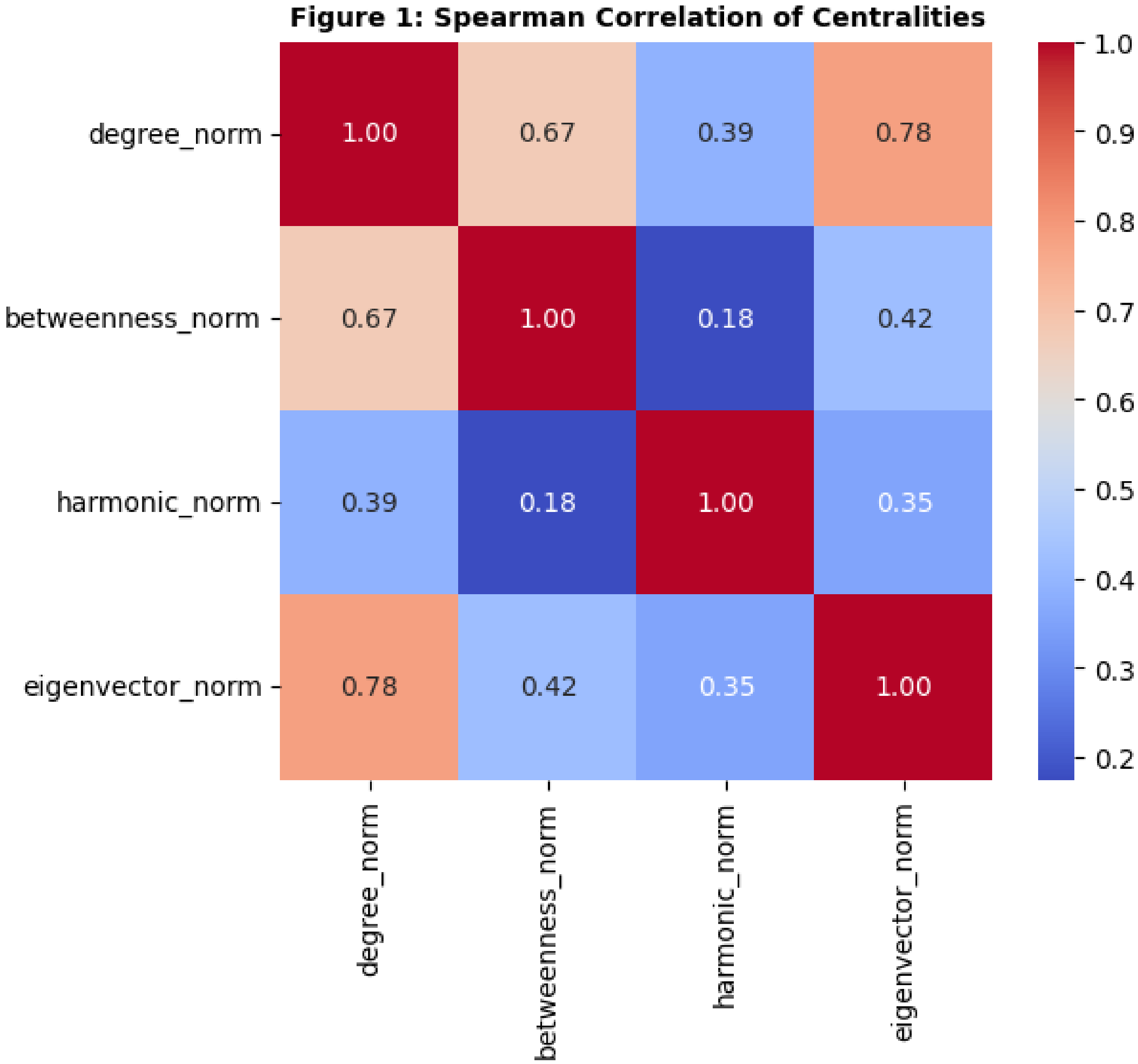
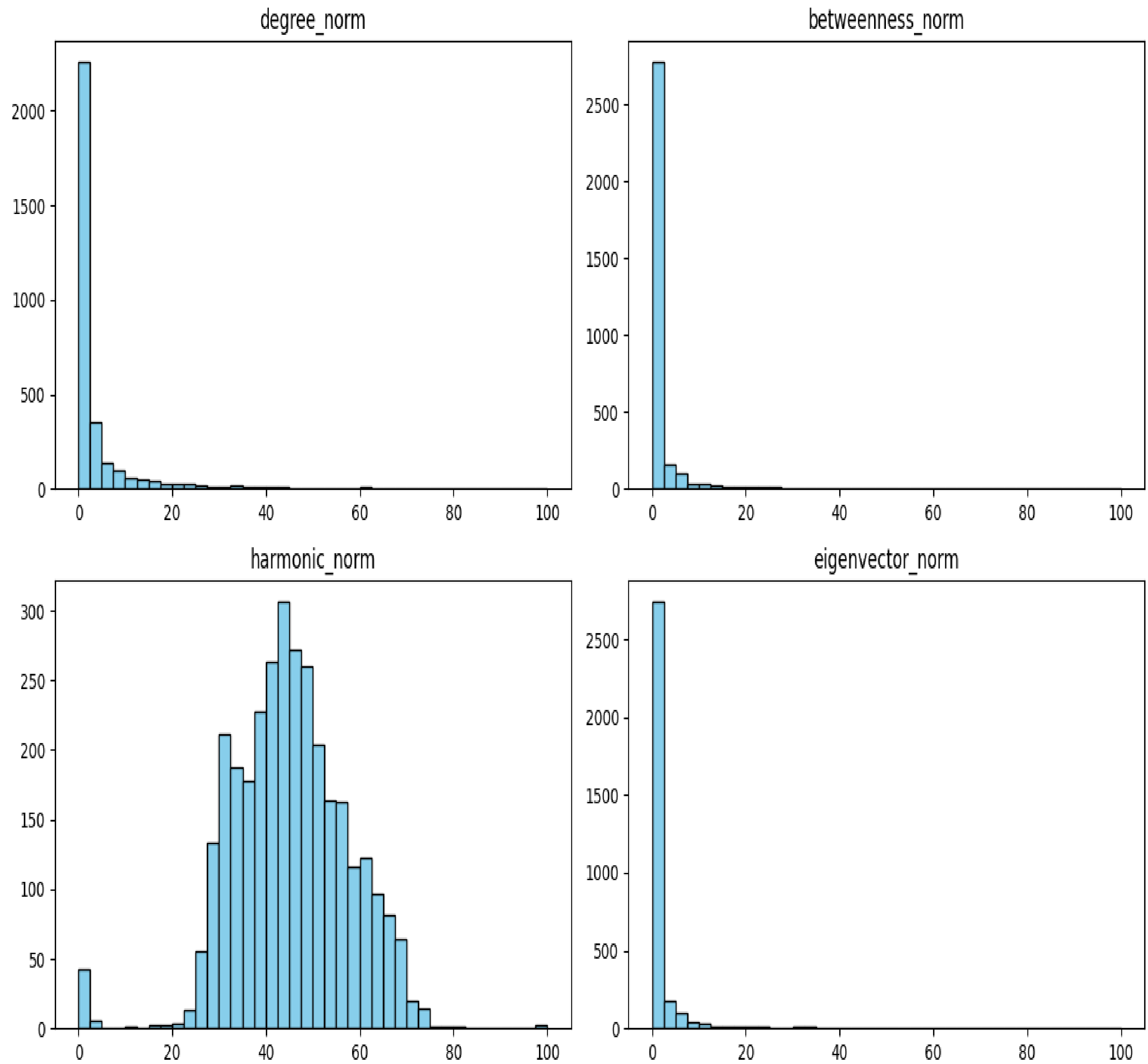


Figure 1: Spearman Correlation of Centralities
degree_norm
betweenness_norm
harmonic_norm
eigenvector_norm
1.00
0.67
0.39
0.78
0.67
1.00
0.18
0.42
0.39
0.18
1.00
0.35
0.78
0.42
0.35
1.00
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2

Figure 1.2: Distribution of Centrality Measures (Normalized)



4.2 Distribution

The histograms in figure in 1.2 distribution of centrality measures reveal degree betweenness and eigen vector are mostly rightly skewed which indicate relatively a small number of airports appear to have scored high and most of them have score low score. However, by contrast harmonic appeared more balanced. It looks quite interesting to see some basic figures for each centrality measure. For degree most of the airport clustered at the lower end of these centrality values and a small number of them leap to the top. If you observe degree around three quarters of all the airports attain score below 25 on the normalized scale, which indicate most airports have relatively few direct routes. However, less than 5% touch beyond 90, these are key mega hubs. Betweenness demonstrate an even sharper till towards the lower end where many airports clustering below 10 and a few critical bridges soars around or above 80. This huge gap suggests how a handful location play an important role in connecting distant region. However, by contrast harmonic shows a broader spread with a mean of roughly 40. This pattern reveals major hubs as well as certain regionally important airports that cut down travel distance for specific routes, despite having minimal long-haul traffic. Meanwhile, Eigenvector mostly stayed in the mid-range for most airports, A smaller subset score above 70, meaning these airports connecting other influential airports. In a nutshell, these patterns demonstrate a network in which a limited number of airports influence connectivity and some regional can command surprisingly high reach if they cut down travel distances for specific routes.

4.3 Rank Correlation

In rank correlation analysis as shown in figure 1, the spearman rank correlation heatmap matrix, we notice there is a strong correlation ($p = 0.78$) between degree and eigenvector centralities. This is expected as airport with many direct connections often connected to other most influential airports, which can push Eigenvector scores. Betweenness, however overlaps partially with Degree ($p = 0.67$) and Eigenvector ($p = 0.42$), which indicates it measures different aspects such as airport that act as bridges across different region. An interesting observation is the weak correlation between harmonic centrality and the other measures. While degree, betweenness, and eigenvector centralities tend to move together—since they focus on direct connections, or the influence of a node's neighbors—harmonic centrality appears to capture something different. As demonstrated by Da Rocha (2009), path-based metrics such as harmonic centrality highlight unique aspects of network structure by focusing on the efficiency with which a node can reach all other nodes. In this case, harmonic centrality is computed as the sum of the inverse shortest path distances from a node to every other node, thereby providing insight into travel efficiency across the network.

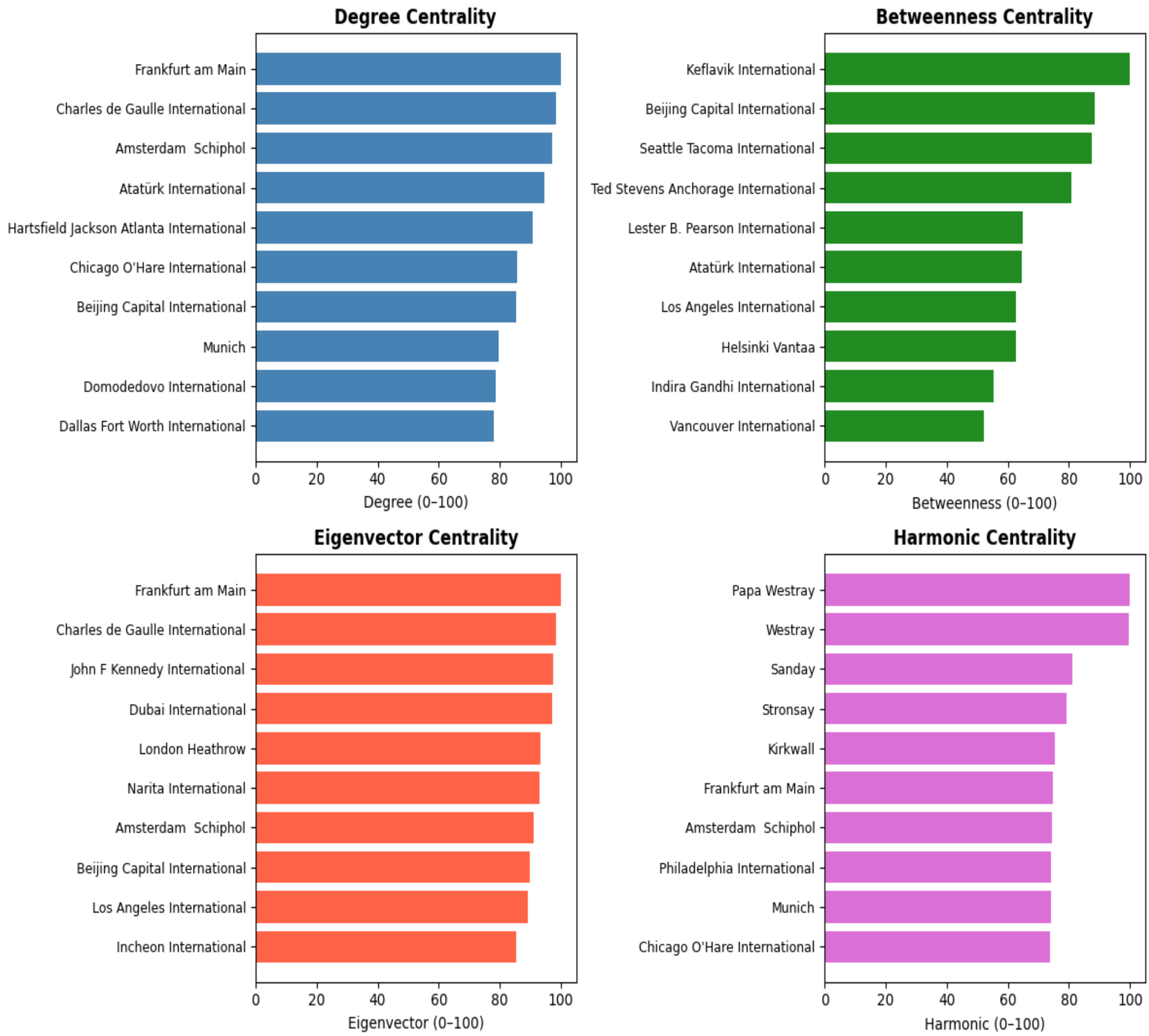
This focus on efficiency means that an airport with a modest number of direct routes might still score high on harmonic centrality if its routes are strategically placed to minimize travel distances. Consequently, the correlation between harmonic centrality and measures that count connections or emphasize bridging roles is relatively weak. In effect, harmonic centrality reveals an efficiency dimension of connectivity that the other metrics tend to overlook.

This observation substantiates that each centrality looks at a different aspect of network structure. Degree and eigenvector highlight the major airports which have many routes, betweenness identify essential airports that connect many different regions, whereas harmonic may point out some smaller airports that cut down the travel time in their region.

4.4 Examining Individual Centrality measures

When we think about airports, we usually imagine busy terminals packed with traveller rushing to catch their flight. But all airports are not similar, every airport has its own characteristics some of them are major international hub while others are small play an essential role of regional connectivity. To understand how airports into the global airlines network and their significance, we can investigate relevant centrality measures: degree, betweenness, eigenvector, and harmonic centrality. Each airport has its own importance and play different role, whether it is the number of flights it handles, its significance in bringing between regions, or how well they are connected to other most influential airports.

Figure 1.3: Top 10 Airports by Centrality Measures



4.4.1 Degree Centrality

Looking at Degree simply counts how many direct routes each airport has. In Figure 1.3 (top-left bar chart), Frankfurt am Main (FRA) leads the list, serving more than 300 destinations and around 69 million passengers in 2019 (Fraport, 2020). It is the Germany's largest airport witnessed a high demand for a medium haul flight to destination having warm weather including Spain, Greece, Italy and Portugal. In

2024 around 61.6 million passengers travelled through Frankfurt am Main highlighting growth of nearly 3.7 percent compared to 2023. It surpasses world's leading airports in terms of hub connectivity since it provides numerous destinations to transfer passengers on direct connecting flight (Fraport, 2020). Charles de Gaulle (CDG) comes in right behind it, topping 76 million in 2019, it serves 328 destinations in 119 countries, which make it a leading airport in France and is Europe's second largest airport. With nine passenger terminals spreading over three large terminal and two independent runways can fit in huge aircraft. The layout of airport Favor scheduling 120 movements per hour (Paris Aéroport, 2021). Amsterdam Schiphol (AMS) is next, well known for its wide network across Europe, North America, and beyond. It serves around 70 million passenger and constantly rank third for being busiest airport in Europe and dominate the list of top European airport in terms of traffic and number of flights. It has 6 runways and 100 of flight operate from this airport and is ranked 6th largest for handling international traffic (Amsterdam Airport, n.d.)

Atatürk International (IST) also scores highly, although Istanbul's main hub has moved to the new Istanbul Airport since these data were collected. Hartsfield-Jackson Atlanta (ATL), despite carrying over 110 million passengers (ACI, 2021), lands fifth in Degree, likely because it's more focused on domestic U.S. routes than an array of global links. Rounding out the top 10 are Chicago O'Hare (ORD), Beijing Capital (PEK), Munich (MUC), Domodedovo (DME), and Dallas Fort Worth (DFW).

Figure 2.1 : Top 100 by Degree Centrality

Turning to Figure 2.1, which maps the top 100 airports by Degree, you immediately see these hubs clustering in North America, Europe, and East Asia. That pattern makes sense, because places like the United States, Germany, France, and China run busy flight corridors, producing a ton of direct connections. Airports like Atlanta, Chicago, and Dallas anchor big domestic networks in the U.S., while Frankfurt, Paris, and Beijing each link to hundreds of destinations worldwide. This concentration shows how Degree centrality really does highlight “mega-hubs” with the sheer number of routes—an outcome that also reflects the fact that large economies tend to support more flights and bigger route networks.

4.4.2 Betweenness Centrality

Now, if you switch to Betweenness, which captures how often an airport lies on shortest paths between other airports, you see different names at the top. In Figure 1.3 (top-right), Keflavik (KEF) in Iceland surprisingly takes first place, thanks to its spot between North America and Europe. It only serves around 10 million passengers (Isavia, 2021), but many transatlantic itineraries stop there. Beijing Capital (PEK) shows up again, bridging Asia, Europe, and North America.

Anchorage (ANC) is next, a huge cargo refueling point for transpacific flights (ACI, 2021). Other airports like Seattle-Tacoma (SEA), Toronto Pearson (YYZ), Los Angeles (LAX), Helsinki Vantaa (HEL), Delhi (DEL), and Vancouver (YVR) each gain a Betweenness boost from how they shorten routes across big regions.

Figure 2.2 : Top 100 by Betweenness Centrality

Looking at Figure 2.2, you can spot those bridging airports scattered around the edges of major continents or across key crossroads. Keflavik is way out in the North Atlantic, Anchorage in far-north Alaska, Helsinki perched in northern Europe, yet all of them show up in bright red as top Betweenness nodes. Their geography makes them crucial connectors, even if airports do not match the raw traffic of Frankfurt or Chicago. This highlights how Betweenness is less about how many routes airports have and more about how vital their position is for linking distant places.

4.4.3 Eigenvector Centrality

Eigenvector focuses on how well-connected an airport's neighbors are, emphasizing ties to other influential airports. In Figure 1.3 (bottom-left), Frankfurt (FRA) and

Charles de Gaulle (CDG) show up again, but we also see John F. Kennedy (JFK) in New York, Dubai International (DXB), and London Heathrow (LHR). JFK handles over 60 million passengers (Port Authority of NY & NJ, 2023), and Dubai hits about 86 million (Dubai Airports, 2019). While neither topped the Degree centrality list, they link up with a bunch of other high-profile hubs, pushing them up in Eigenvector. Heathrow is a similar story: it’s historically one of the world’s busiest, around 80+ million passengers (Heathrow Airport, 2022), but it doesn’t necessarily break records for total destinations. Eigenvector rewards that “club” effect, where influential airports connect to others that are also influential.

Figure 2.3 : Top 100 by Eigenvector Centrality

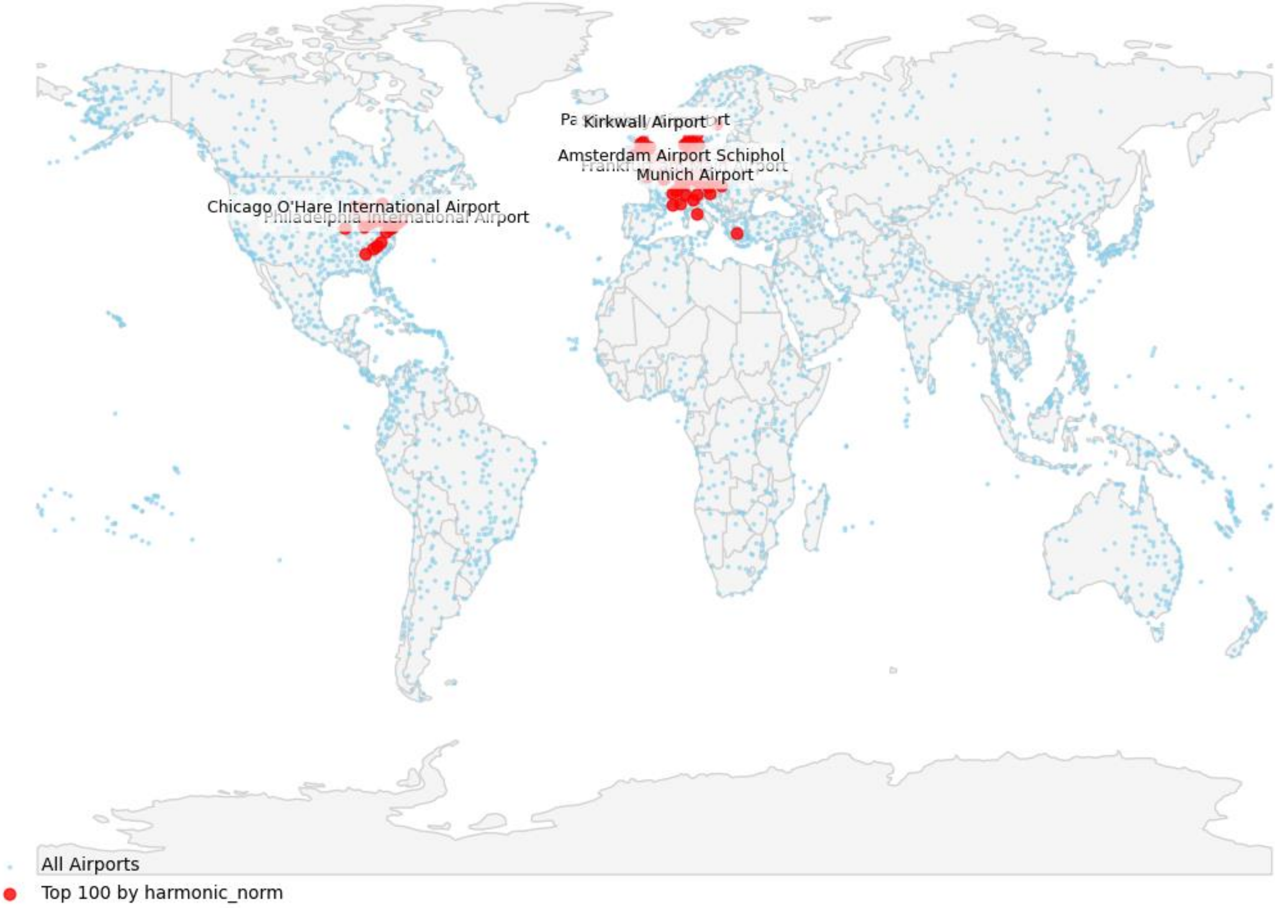
In Figure 2.3, these top Eigenvector airports are spread across multiple continents but form a sort of global backbone. You can see how JFK, LAX, Heathrow, Dubai, and Beijing Capital all appear as red dots in major international corridors. It is less about having the largest number of routes and more about who you’re connected to—if you are tied into an exclusive network of big airports, you score well in Eigenvector.

4.4.4 Harmonic Centrality

Finally, Harmonic centrality measures how quickly an airport can reach every other airport in the system, basically how many flights it saves on average. That is why, in Figure 1.3 (bottom-right), you see a bunch of tiny Scottish airports Papa Westray, Westray, Sanday, Stronsay, and Kirkwall—sitting near the top. Although they each only serve a few thousand passengers (Loganair, 2019), they slash travel times for people in the Orkney Islands. In Orkney, ferries can be delayed by weather, and traveling between islands can take hours if you rely on boats and connecting flights elsewhere. Papa Westray and Westray, each operating a runway of just a few hundred meters (Civil Aviation Authority, 2021). They collectively host what is widely called the world's shortest scheduled commercial flight, often clocking in at under two minutes in the air (Loganair, 2022). For residents, that hop cuts down an otherwise tedious journey by a tremendous margin, making life drastically easier.

Right after that local group, you get the expected major hubs: Frankfurt, Amsterdam, Philadelphia, Munich, and Chicago O'Hare. Philadelphia (PHL) doesn't appear in the top lists for other measures, but it sits in a heavily travelled U.S. East Coast region, giving it short average flight paths to lots of destinations.

Figure 2.4: Top 100 by Harmonic Centrality

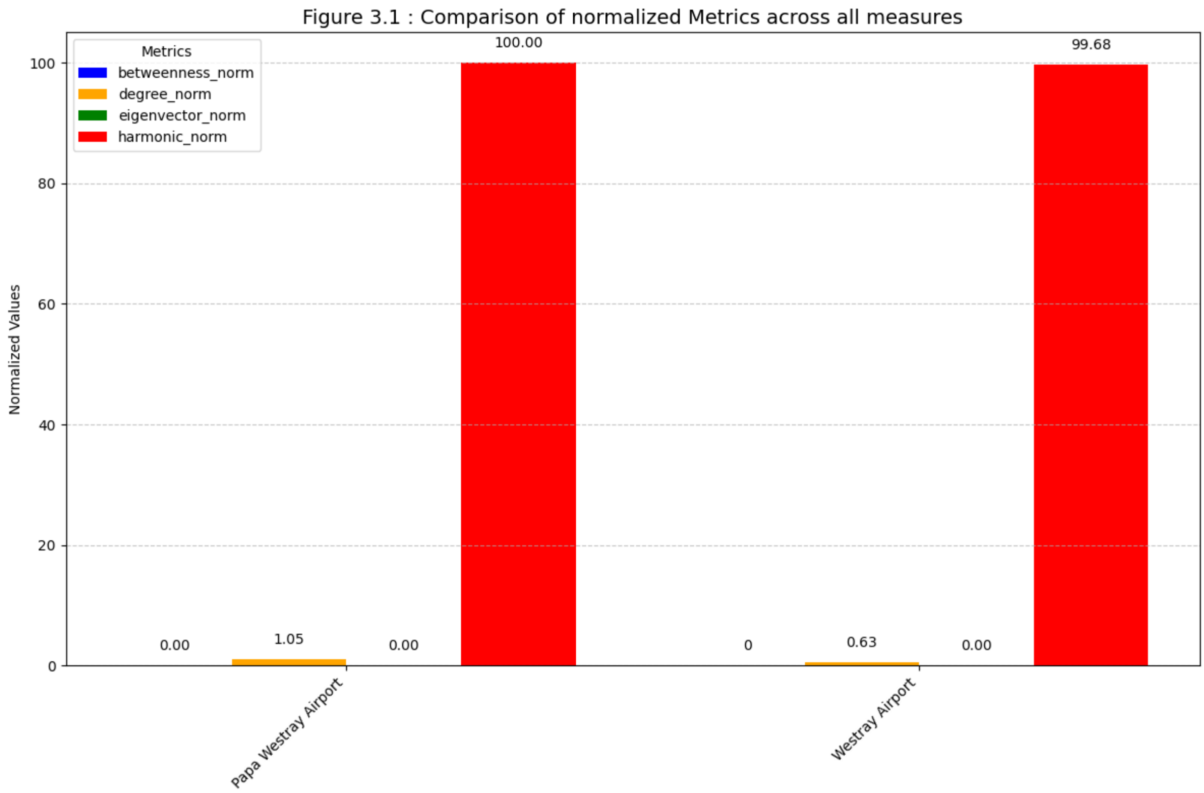


If you check out Figure 2.4, it’s immediately clear that the highest-ranked Harmonic airports are either gigantic hubs in North America and Europe, or those little island airports up by Scotland. This might seem strange seeing Papa Westray next to Chicago O’Hare, but it shows how an airport can be “crucial” if it spares travellers lengthy connections, even if the total passenger count is small.

Overall, each of these measures Degree, Betweenness, Eigenvector, and Harmonic tells a different perspective about how airports fit into the global network. Mega-hubs like Frankfurt and Beijing Capital excel across multiple categories, while specific outliers (Keflavik for bridging, Papa Westray for local connections) illustrate that “importance” depends on what exactly we measure. After we verify these results against actual passenger stats and cargo data (ACI, 2021; Fraport, 2020; Paris Aéroport, 2021; Port Authority of NY & NJ, 2023; etc.), it’s clear that traditional giants really do handle huge volumes, smaller connectors are positioned in just the right spots, and local airstrips can be lifesavers for their own region.

4.5 Spotlight and Further Justification on Outliers: Why Papa Westray and Westray Top the Harmonic Centrality Ranking?

Before this analysis, it was assumed that giants' hubs London Heathrow, Frankfurt, JKF, Dubai would dominate across all the metrics. After looking at the top 10 airports in each measure, a couple of airports popped up unexpectedly. Further investigation by comparing all the measures and identifying airports rank highly in one measure and scoring very low in other measures Papa Westray and Westray dominate Harmonic measure (Figure 3.1). Their prominence in Harmonic centrality demonstrates the critical significance of regional connectivity and this contrast reveal their unique importance in the global airlines network.



Papa Westray and Westray look like nothing more than tiny airstrips in the Orkney Islands as can be seen in Figure 3.2, yet on Harmonic centrality, they surpass even the largest global hubs. At first, it seems puzzling: how could two short runways in remote islands surpass some of the busiest airports on Earth that handle tens of millions of passengers? —but it makes perfect sense once you realize how Harmonic works. This measure cares about how many “hops” an airport can save across the entire network, not just how many destinations it serves or how many travellers it moves. This means that even if an airport has only a few routes, it can still rank highly if those routes are optimally placed to reduce travel time across the network.

Figure 3.2: Airports with High Harmonic but Low in Other Measures

Figure 3.2

Kirkwall Airport as the central hub of the Orkney Islands—a busy, well-connected node in a tight-knit network. In our analysis, we were interested in not just how many routes each airport has, but also in how the actual physical distances between them impact overall connectivity. To capture this, we use a measure called harmonic centrality. (We explained the formula earlier in our methodology.)

The core idea behind harmonic centrality is that it rewards not only the number of connections an airport has but also the efficiency of those connections—in other words, how quickly and efficiently an airport can reach all other nodes in the network. This is done by taking the reciprocal of the distances. Shorter distances yield much higher reciprocal values than longer distances.

For our analysis, the distances were computed using the geographic coordinates from the OpenFlights dataset. Our calculations produced the following distances for the Orkney Islands:

- **Kirkwall → Papa Westray:** 43.8 km
- **Kirkwall → Westray:** 43.7 km
- **Kirkwall → Sanday:** 37.5 km
- **Kirkwall → Stronsay:** 26.6 km
- **Papa Westray → Westray:** 2.8 km
- **Papa Westray → Sanday:** 21.6 km
- **Papa Westray → Stronsay:** 26.3 km

Let us break down what these numbers mean using Harmonic formula

For Kirkwall:

- The connection to Papa Westray is 43.8 km long, so its reciprocal is about $\frac{1}{43.8} \approx 0.0228$.
- The link to Westray is 43.7 km and its reciprocal is $\frac{1}{43.7} \approx 0.0229$

• The route to Sanday is 37.5 km its reciprocal $\frac{1}{37.5} \approx 0.0267$

• And the link to Stronsay is 26.6 km its reciprocal $\frac{1}{26.6} \approx 0.0376$

Adding these together:

0.0228 + 0.0229 + 0.0267 + 0.0376 ≈ 0.1099

Thus, Kirkwall's total reciprocal contribution to harmonic centrality is about 0.11

For Papa Westray:

• Its connection to Kirkwall (43.8 km) again contributes about $\frac{1}{43.8} \approx 0.0228$

• However, it is very short link to Westray is only 2.8 km, giving a reciprocal of roughly $\frac{1}{2.8} \approx 0.3571$

• The connection to Sanday (21.6 km) contributes about $\frac{1}{21.6} \approx 0.0463$

• And its link to Stronsay (26.3 km) adds approximately $\frac{1}{26.3} \approx 0.0380$

Summing these:

0.0228 + 0.3571 + 0.0463 + 0.0380 ≈ 0.4642

So, Papa Westray's overall harmonic centrality is around 0.46—a value that is much higher than Kirkwall's. This is largely due to that very short 2.8 km link to Westray, which alone gives a reciprocal of about 0.3571.

4.5.1 Compounding the Effect of Local Proximity

The key takeaway is that harmonic centrality is highly sensitive to the length of connections. Shorter distances yield much larger reciprocal values. In the Orkney cluster, the very short 2.8 km link between Papa Westray and Westray (yielding a reciprocal of approximately 0.3571) dominates the overall efficiency measure. This is not just a matter of one connection being short; it compounds the overall harmonic centrality because an airport benefits from all its direct links. If an airport is connected to another node that itself has short, efficient connections to the rest of the network, then that efficiency is compounded in the harmonic measure.

In contrast, consider global hubs like Frankfurt or Amsterdam. Although they have many connections, these connections cover much larger geographic areas. The longer distances result in much smaller reciprocals, which means that despite high connectivity, their overall harmonic centrality is lower. Thus, the compact geography of the Orkney Islands drives a significant “compounding effect” that boosts the harmonic centrality of small, closely located airports like Papa Westray and Westray.

The detailed calculations show that while Kirkwall’s cumulative reciprocal contribution is only about 0.11, Papa Westray’s is around 0.46—primarily due to its extremely short 2.8 km connection to Westray. This “compounding effect” of local efficiency demonstrates that in tightly clustered networks, even small regional airports can exert a disproportionate influence on overall connectivity. Thus, while traditional measures might focus on the number of routes or passenger volume, harmonic centrality reveals the critical role of proximity and efficiency in determining an airport’s true importance.

Degree centrality, on the other hand, emphasizes raw route counts—an area where these islands cannot hope to compete with big players like London Heathrow or Frankfurt. Betweenness Favors places that bridge large regions or handle cargo on long transits, but Papa Westray and Westray mainly exist to serve a small set of communities, so they hardly register there either. Eigenvector captures how connected your neighbours are, which again Favors major hubs linking to other influential nodes.

So, even though Papa Westray and Westray might not impress you with high passenger volumes or huge route maps, they’re classic examples of how “importance” can mean different things depending on the measure. Our analysis reveals that their true importance lies in the efficiency of their local connections. When measured by harmonic centrality—which rewards the ability to reach other nodes quickly by using the inverse of the distance resulting a small local airport can outperform international gateways. From a global perspective, it is almost surprising to see them overshadow mega-hubs in any category, but from a local standpoint, it makes total sense.

4.6 Heathrow, Dubai, and JFK Below the Top 10 in Degree Centrality: Why It's Surprising and Unexpected?

It can feel surprising to see global giants like Heathrow (LHR), Dubai (DXB), and John F. Kennedy (JFK) airports land outside the top 10 in Degree centrality- all of which handle tens of millions of passengers every year—show up **outside** the top ten in Degree centrality (Heathrow Airport, 2022; Dubai Airports, 2019; Port Authority of NY & NJ, 2023). Heathrow, for example, lands at **12th** place with a normalized Degree of **71.22**, while Dubai takes **11th** (76.26), and JFK is **17th** (67.44). Meanwhile, Frankfurt (FRA) sits comfortably at **#1** with a perfect **100.00**.

Airport	Degree (Norm)	Rank	Outgoing	Incoming	NetX Degree
Frankfurt	100.00	1	497	493	477
Heathrow	71.22	12	525	522	340

Table 1: Degree centrality rank for Frankfurt and Heathrow

By quickly looking at the raw data in our Openflights analysis, you might believe Heathrow is more connected because it has 525 outgoing routes and 522 incoming. However, when we convert these routes into a simple NetworkX DiGraph, duplicates and reverse listings (e.g., LHR → FRA vs. FRA → LHR) merge into fewer unique edges. Frankfurt list 497 outgoing and 493 incoming collapses into remarkable 477 distinct nodes and Heathrow's actual count of unique endpoints boil down to 340. It turns out Heathrow serves many of the same major city pairs repeatedly (sometimes for multiple airlines or flight numbers), whereas Frankfurt spreads out its routes across a broader spectrum of endpoints.

Centrality Measure	Normalized Score	Rank
Degree Centrality	71.22	12
Betweenness Centrality	39.88	21
Harmonic Centrality	69.27	58
Eigenvector Centrality	93.40	5

Table 1.1: London Heathrow Airport Centrality Measures

Heathrow's business model (Heathrow, 2022) mainly prioritizes maximising long-haul passenger loads, also limit the total number of smaller or regional routes due to runway constraints. Schiphol, for example, focuses on attracting budget and regional airlines. This helps it add more direct connections. Similarly, Frankfurt offers a broader array of mid-range flight (Fraport, 2020), or secondary cities connecting to multiple European destination that offer more unique nodes in our DiGraph analysis. In comparison, an airport that mainly focuses on long-haul flights wouldn't have as many short-haul routes. Therefore, in Degree centrality Heathrow fall behind some of its European and American counterparts, despite its massive passenger's count. Recent data from OAG's Megahubs Index (OAG, 2024) shows that Heathrow ranks high for international connectivity. But at the same time, it doesn't always surpass hubs like Frankfurt. Frankfurt covers a wider range of European destinations, which helps it build more unique direct connections. This supports our finding that having a lot of passengers doesn't necessarily mean an airport will rank high in Degree centrality. It's not just about how busy the airport is. What really matters is how many different places it connects to. The more unique routes an airport has, the more diverse its network. So, in the end, route diversity can be just as important as passenger traffic.

Centrality Measure	Normalized Score	Rank
Degree Centrality	76.26	11
Betweenness Centrality	26.60	52
Harmonic Centrality	50.75	977
Eigenvector Centrality	97.30	4

Table 1.2: Dubai International Airport Centrality Measures

You find the same thing happening with Dubai. It ranks 11th in Degree centrality scoring (76.26) and handles around 86 million passengers (Dubai Airports, 2019). But instead of reaching a huge number of destinations, it mainly focuses to connect Europe, Africa, and Asia. The focus is on essential travel corridors rather than spreading out to multiple smaller airports.

John F Kennedy International Airport Centrality Measures

Centrality Measure	Normalized Score	Rank
Degree Centrality	67.44	17
Betweenness Centrality	32.70	33
Harmonic Centrality	67.96	88
Eigenvector Centrality	97.38	3

Table 1.3: John F Kennedy International Airport Centrality Measures

JFK is another example of this. It's one of the busiest airports for transatlantic and global flights (Port Authority of NY & NJ, 2023). But compared to airports like Atlanta (ATL) or Chicago O'Hare (ORD), it does not have the same massive domestic network. That is a big reason why its degree centrality is not as high as those major U.S. hubs.

Even though these airports do not rank at the very top for Degree centrality, but if you notice they move up when we look at Eigenvector centrality. Heathrow ranks 5th (93.40), Dubai takes 4th (97.30), and JFK comes in at 3rd (97.38). This is because Eigenvector centrality does not measure just about how many routes an airport has. It looks at who they are connected to instead. So, it less about having distinct multiple destinations and more about being connected to other important hubs.

Studies on air transportation networks suggest that Eigenvector Centrality is more about the importance of connections rather than just the total number of routes (Centralities based on centralities, 2024; A change of perspective in network centrality, 2018). In other words, it rewards airports that are linked to other major hubs instead of just counting how many destinations they serve. That is why Heathrow scores high in Eigenvector centrality. It has strong links with big airports like New York JFK and Hong Kong, which helps boosting it eigenvector score, even if it trails in the number of total connections (OAG, 2024; Bovet and Makse, 2022). There is also a 0.78 correlation between Degree and Eigenvector centrality between degree and Eigenvector observed in our analysis, which is common. Airports with large routes of network usually connect to each other. But sometimes, an airport puts more focus on a few important destinations, which can change how it ranks.

Furthermore, OpenFlights data mostly treats all connections as unweighted, which can make it seem like all routes are equal, even though some flights runway more often or carry way more passengers (Centrality anomalies for the domestic air transportation networks in the USA, 2023). As mentioned earlier Heathrow mainly focuses on premium long-haul flights, so it does not have as many total destinations as some other airports. But in real world context, the connections it does have are extremely important. This paradox aligns with previous studies that unweighted models can sometimes produce “anomalies” where an airport’s Eigenvector and Degree rankings don’t match up as expected (A change of perspective in network centrality, 2018). In the end, Heathrow’s case proves that one single measure is not enough to fully capture an airport’s importance. Its degree rank might be lower than Frankfurt or Amsterdam, but its high Eigenvector score shows that it still plays a major role by connecting to some of the most important hubs in the world.

Chapter 5: Conclusion

This study uncovered some unexpected insights into complex links between airports. Instead of simply counting flights or passengers, the research investigates various network measures—degree, betweenness, eigenvector, and harmonic centrality—to explore more what really makes an airport significant.

The analysis results show that no one metric can explains the importance an airline networks. Rather, a blend of these factors gives a more complete picture of airport importance. Big airports like Frankfurt and Beijing Capital may have numerious direct connections, yet smaller hubs like Keflavik and Anchorage play a critical role in linking distant regions. Their importance goes beyond sheer volume; it’s all about where they sit in the network.

One of the most surprising parts of this research was discovering how important some of the smaller regional airports are—take Papa Westray and Westray, for example. Even though they handle only a small number of the traffic seen at major international airports in the world, these tiny airstrips can considerably cut down travel times for the people living in the region, these are more than just transit points and are a lifeline, showing that importance is not just about size or volume rather where strategically they are located.

5.1 Implications for Air Network Design

The study shows that when an airport has a lot of connections, those connections tend to be influential as well. In other words, a high number of direct links usually comes with a strong impact on the network. Harmonic centrality gives a different perspective by focusing on travel efficiency. The takeaway is clear that enhancing the air network is not merely increasing the number of routes, but also carefully designing how those routes connect.

Although the study relied on a single data source and required certain simplifications, the analysis provide a more detailed picture of global air travel. The findings make it clear that understanding an airport’s importance is a multi-layered challenge. The intricate nature of air travel networks cannot be restricted to just one number; instead, the value of an airport depends on a blend of factors that determine how well it fits into the overall global system.

In short, airport significance turns out to be a multi-faceted issue. By examining several centrality measures, it is clear that the story is not just about the massive international hubs. This broader perspective not only our enriches academic knowledge but also provides practical insights for enhancing the efficiency and resilience of the global airline network.

References

1. **Airports Council International (ACI). (2021).**

World Airport Traffic Report.

Available at: https://aci.aero/data-centre/ (Accessed: 21 January 2025).

2. **Amsterdam Airport (n.d.).**

Frankfurt, Istanbul, and Paris Charles de Gaulle are the world's best-connected airports.

Available at: https://www.amsterdamairport.info/ (Accessed: 20 January 2025).

3. **Barthélemy, M. (2011).**

'Spatial networks', *Physics Reports*, 499(1–3), pp. 1–101.

https://doi.org/10.1016/j.physrep.2010.11.002

4. **Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. and Hwang, D.U. (2006).**

'Complex networks: Structure and dynamics', *Physics Reports*, 424(4–5), pp. 175–308.

https://doi.org/10.1016/j.physrep.2005.10.009

5. **Borgatti, S.P. (2005).**

'Centrality and network flow', *Social Networks*, 27(1), pp. 55–71.

https://doi.org/10.1016/j.socnet.2004.11.008

6. **Borgatti, S.P. and Everett, M.G. (2006).**

'A graph-theoretic perspective on centrality', *Social Networks*, 28(4), pp. 466–484.

https://doi.org/10.1016/j.socnet.2005.11.005

7. **Bovet, A. and Makse, H.A. (2022).**

"Centralities in Complex Networks," in *Statistical and Nonlinear Physics*. Springer.

(If you have a specific DOI, include it.)

8. **Civil Aviation Authority. (2021).**

UK Regional Airport Runway Data.

Available at: https://www.caa.co.uk/ (Accessed: 19 January 2025).

9. **Da Rocha, L.E.C. (2009).**

'Harmonic Centrality in Complex Networks', *Journal of Statistical Mechanics: Theory and Experiment*, 2009(04), P04020.

https://doi.org/10.1088/1742-5468/2009/04/P04020

10. **Dubai Airports. (2019).**

Dubai International Traffic Statistics.

Available at: https://www.dubaiairports.ae/ (Accessed: 24January 2025).

11. **Fraport AG. (2020).**

Frankfurt Airport Annual Report 2019.

Available at: https://www.fraport.com/ (Accessed: 27 January 2025).

12. **Freeman, L.C. (1978).**

'Centrality in Social Networks: Conceptual Clarification', *Social Networks*, 1(3), pp. 215–239.

https://doi.org/10.1016/0378-8733(78)90021-7

13. **Guimerà, R., Mossa, S., Turtschi, A. and Amaral, L.A.N. (2005).**

'The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles', *Proceedings of the National Academy of Sciences*, 102(22), pp. 7794–7799.

https://doi.org/10.1073/pnas.0407994102

14. **Heathrow Airport. (2022).**

Annual Passenger Figures.

Available at: https://www.heathrow.com/ (Accessed: 25 January 2025).

15. **Isavia. (2021).**

Keflavik Airport Traffic Data.

Available at: https://www.isavia.is/en/corporate (Accessed: 25 January 2025).

16. **Jolliffe, I.T. (2002).**

Principal Component Analysis. 2nd edn. New York: Springer-Verlag.

https://doi.org/10.1007/b98835

17. **Kolaczyk, E.D. (2009).**

Statistical Analysis of Network Data: Methods and Models. New York: Springer.

https://doi.org/10.1007/978-0-387-88146-1

18. **Loganair. (2019).**

Shortest Flight Statistics.

Available at: https://www.loganair.co.uk/ (Accessed: 21 January 2025).

19. **Loganair. (2022).**

Passenger Statistics and Orkney Flight Schedules.

Available at: https://www.loganair.co.uk/ (Accessed: 23 January 2025).

20. **Newman, M.E.J. (2010).**

Networks: An Introduction. Oxford: Oxford University Press.

ISBN: 9780199206650

21. **OAG. (2024).**

Megahubs Index.

Available at: https://www.oag.com/megahubs-airports-2024 (Accessed: 14 January 2025).

22. **OpenFlights. (n.d.).**

OpenFlights Dataset.

Available at: https://openflights.org/data.html (Accessed: 1 February 2025).

23. **Paris Aéroport. (2021).**

Charles de Gaulle Passenger Statistics. Groupe ADP.

Available at: https://www.parisaeroport.fr/ (Accessed: 27 January 2025).

24. **Port Authority of New York & New Jersey. (2023).**

JFK Annual Traffic Data.

Available at: https://www.panynj.gov/ (Accessed: 22 January 2025).

25. **'A change of perspective in network centrality'. (2018).**

Scientific Reports, 8, 11559.

https://doi.org/10.1038/s41598-018-33336-8

26. **Centralities based on centralities. (2024).**

Advanced Topics in Network Science.

Available at: https://skojaku.github.io/adv-net-sci/m06-centrality/eigencentrality.html (Accessed: 14 January 2025).

27. **Centrality anomalies for the domestic air transportation networks in the USA. (2023).**

EPJ Special Topics, 232(1), pp. 1–14.

https://doi.org/10.1140/epjp/s13360-023-04003-3

28. **Ugurlu, O. (2022b).**

'Comparative analysis of centrality measures for identifying critical nodes in complex networks', *Journal of Computational Science*, 55, 101738.

https://doi.org/10.1016/j.jocs.2022.101738

29. **Wasserman, S. and Faust, K. (1994).**

Social Network Analysis: Methods and Applications. Cambridge: Cambridge University Press.

ISBN: 9780521387071

30. **Opsahl, T., Agneessens, F. and Skvoretz, J. (2010).**

'Node centrality in weighted networks: Generalizing degree and shortest paths', *Social Networks*, 32(3), pp. 245–251.

https://doi.org/10.1016/j.socnet.2010.03.006

31. Sinnott, R.W. (1984) 'Virtues of the Haversine', *Sky & Telescope*, 68, pp. 159–163.