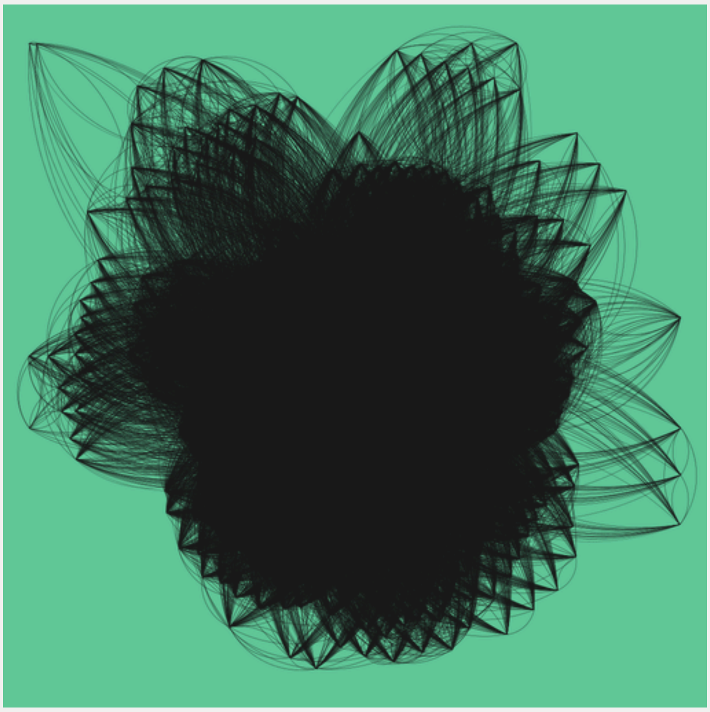
IIIT
DHARWAD

Analysis of AIRLINE TRAVEL REACHABILITY NETWORK

"Mapping Connectivity and Accessibility in Global Airline Networks"



Prepared by

Arpit Nayak 22BDS006

Rajput Ajay 22BDS049

Viraj Surana 22BDS064

Bhabad Ganesh 22BDS067

GitHub Link

Abstract

This project explores the structural properties of a directed network by analyzing its degree distribution, focusing on in-degrees and out-degrees of nodes. The study aims to identify key patterns in connectivity, examining how nodes interact and the prevalence of hubs within the network. By visualizing degree distributions and applying smoothing techniques, we observe a skewed pattern where most nodes have low degrees, while a few act as significant hubs with high out-degrees, suggesting influential or central nodes that serve as connection sources. These insights provide valuable implications for understanding network dynamics, such as identifying influential nodes in social networks or critical points in transportation systems. The project also analyzes the clustering coefficient and its correlation with node degree, illustrating the tendency for higher-degree nodes to have lower clustering. Overall, this report sheds light on network connectivity and informs strategies for enhancing resilience, efficiency, and influence within various networked systems.

INTRODUCTION

1. Background on Network Analysis

Network analysis is a powerful tool for examining relationships and structural properties within a system of interconnected entities. Networks can represent diverse systems, from social networks, where nodes represent individuals and edges represent interactions, to infrastructure networks, such as transportation or telecommunication systems. Directed networks, where edges have a direction (from one node to another), are particularly insightful for understanding asymmetric relationships. For instance, in a railway network, a directed graph can model connections between stations, with each directed edge representing the direction of travel.

In recent years, network science has provided essential methods for analyzing various network properties, including degree distribution, clustering coefficient, and centrality measures. These metrics help in identifying critical nodes, understanding the network's resilience, and characterizing its structure. Degree distribution, for example, reveals how connections are spread among nodes, while clustering coefficients offer insight into the likelihood of nodes to form tightly knit groups, often reflecting redundancy or robustness in the network.
s. Key findings indicate strong engagement rates but opportunities for improvement in content formats and audience targeting. The goal is to assess the effectiveness of our content strategy in reaching and engaging our target audience.

2. Objective of the Project

The primary objective of this project is to analyze the structural properties of a directed network, specifically focusing on its degree distribution and clustering coefficients. Using data from a real-world network (such as the Railway Network dataset), this project examines in-degree and out-degree distributions, clustering tendencies, and connectivity patterns. The goal is to uncover critical insights about the network, such as the presence of hubs—nodes with disproportionately high degrees—and the relationship between a node's degree and its clustering coefficient

By exploring these aspects, the project aims to provide a comprehensive understanding of network behavior, particularly in identifying influential nodes and potential vulnerabilities. This analysis is valuable for applications in optimizing network performance, enhancing resilience, and improving strategic decision-making, such as resource allocation in transportation networks or information flow in communication networks.

- Network Structure Analysis: To examine the underlying structure of the North American Airline Reachability Network, focusing on connectivity, efficiency, and node significance.
- Hub and Route Identification: To identify key hubs and routes that significantly contribute to the network's reachability, efficiency, and resilience.
- Key Network Metrics Calculation: To compute centrality measures, clustering coefficients, and path length distributions that offer insights into network robustness and efficiency.
- Visualization of Network Connectivity: To create a clear visual representation of network connectivity, enabling easy identification of regional clusters, critical hubs, and underserved areas.
- Real-World Inference: To derive practical recommendations from the analysis that can aid in enhancing network resilience, reducing congestion, and supporting improved travel routes.

3. Challenges and Considerations

Analyzing a directed network presents unique challenges. First, directionality adds complexity to the interpretation of degree distribution, as each node has distinct in-degrees and out-degrees. This asymmetry provides more granular insights but also requires careful consideration when analyzing node influence and connectivity. Additionally, directed networks may exhibit different clustering behaviors, depending on the type of connections and their distribution. For example, hubs with high out-degrees may connect to many low-degree nodes, which can affect the clustering coefficient calculations and overall network resilience.

Another consideration is the potential presence of scale-free properties in the network, where a few nodes play a disproportionately large role in maintaining connectivity. Identifying such nodes can be critical for applications where certain nodes serve as pivotal points in communication or transportation networks. In a railway network, for example, a hub with a high out-degree may represent a major station connecting several smaller stations, making it essential for efficient

4. Significance of Findings

The insights gained from this network analysis have practical implications across various domains. In transportation networks, understanding the degree distribution and clustering behavior can inform route optimization, station placement, and resource allocation strategies. In social or communication networks, these metrics can help identify influential users or nodes, guiding information dissemination or containment strategies in case of viral content or misinformation.

Furthermore, the study's findings on degree-clustering relationships contribute to the broader understanding of directed networks and their structural properties. Recognizing patterns in how high-degree nodes behave and cluster can lead to better predictive models for network growth and evolution, benefiting fields such as epidemiology, social network analysis, and infrastructure planning.

METHODOLOGY

The methodology for this project involves analyzing a directed network to gain insights into its structure and connectivity. We use degree distributions, clustering coefficients, and correlation analyses to understand the network's behavior and identify important nodes or communities. Each step below includes relevant concepts, definitions, and formulas.

1. Data Collection and Pre-processing

- Objective: To import and pre-process the network dataset, preparing it for analysis.
- Steps:

1. Load Dataset: Use Python libraries such as pandas to load the dataset. Ensure that the data structure represents a directed graph, with nodes (entities) and directed edges (connections between entities).
2. Format Validation: Check if the dataset contains columns specifying node pairs (source and target) to form directed edges. If necessary, rename columns to "source" and "target" for consistency.
3. Data Cleaning: Remove any duplicate rows or self-loops (edges where the source and target are the same) as they might interfere with degree distribution and clustering coefficient calculations.
4. Graph Creation: Use NetworkX's DiGraph function to create a directed graph from the data. Each entry in the dataset is converted to a directed edge between two nodes.

This step ensures that the dataset is clean and properly formatted for further analysis.

2. Degree Distribution Analysis

- Concept: Degree distribution provides insight into the network's connectivity pattern. For directed networks, each node has an in-degree (number of incoming edges) and an out-degree (number of outgoing edges).
- Definitions:
  - In-Degree (k_{in}):The number of edges directed toward a node.
  - Out-Degree(k_{out}): The number of edges directed away from a node.
  - Degree Distribution: The probability distribution of degrees across all nodes.
- Formulas:
  - Let $P(k_{in})$ be the probability that a randomly selected node has in-degree k_in, calculated as:

$$P(k_{in}) = \frac{\text{Number of nodes with } k_{in}}{\text{Total number of nodes}}$$

  - Similarly, $P(k_{out})$ is calculated for out-degrees.

- Steps:
  a. Compute In-Degree and Out-Degree: For each node, calculate k_{in} and k_{out} using NetworkX functions in_degree and out_degree.
  b. Calculate Degree Distribution: Count the frequency of each degree value and divide by the total number of nodes to get probabilities.
  c. Visualize: Use a histogram or log-log plot to visualize the degree distribution for in-degrees and out-degrees. A log-log plot is particularly useful for checking if the network follows a power-law distribution, common in scale-free networks.

This analysis helps in identifying whether the network has a few hubs (nodes with high degrees) and many low-degree nodes.

3. Clustering Coefficient Calculation

- Concept: The clustering coefficient measures the tendency of nodes to form clusters or groups. In a directed network, it reflects the probability that two neighbors of a node are also connected.

- Definitions:
  - Clustering Coefficient: For a node i, the clustering coefficient C_i is calculated based on the number of directed triangles involving i.
- Formulas:
  - For a directed network, the clustering coefficient C_i is:

$$C_i = \frac{T_i}{k_{out}(k_{out} - 1)}$$

where T_i is the number of directed triangles through node iii (connections where two neighbors of i are connected).

- Steps:
  - Calculate Clustering Coefficient for Each Node: Use NetworkX's clustering function to calculate the clustering coefficient for each node. Since it's a directed network, ensure that the function computes clustering specifically for directed graphs.

- Compute Average Clustering Coefficient: Calculate the mean clustering coefficient for all nodes, providing a measure of the overall tendency for clustering in the network.
- Visualize Clustering Coefficient Distribution: Plot the distribution of clustering coefficients across nodes. This can reveal if certain types of nodes (e.g., hubs) have higher or lower clustering tendencies.

This analysis highlights the clustering behavior of the network and whether certain nodes form tightly-knit communities.

4.Correlation Analysis Between Degree and Clustering Coefficient

- Objective: To examine the relationship between a node's degree and its clustering coefficient, which can reveal insights into network structure.
- Hypothesis: In many networks, high-degree nodes have lower clustering coefficients, as they connect to many nodes that are not interconnected.
- Steps:
    a. Prepare Data for Correlation: Create a table or dataset where each row represents a node, with columns for in-degree, out-degree, and clustering coefficient.
    b. Plot Degree vs. Clustering Coefficient: Create scatter plots for in-degree vs. clustering coefficient and out-degree vs. clustering coefficient.
    c. Calculate Correlation: Use statistical functions to calculate correlation coefficients (e.g., Pearson or Spearman) between degree and clustering coefficient. This can confirm or reject the hypothesis of an inverse relationship.
    d. Interpret Results: Analyze the plots and correlation coefficients to determine whether high-degree nodes indeed have lower clustering coefficients.

This step provides insights into the hierarchical structure of the network and the connectivity patterns of high-degree nodes.

5. Identification of Key Nodes (Centrality Measures)

- Concept: Centrality measures help identify the most important or influential nodes in a network.
- Definitions and Formulas:
  - Betweenness Centrality: Measures how often a node appears on the shortest paths between pairs of nodes.

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

where σ_{st} is the total number of shortest paths from node 's' to node 't', and $\sigma_{st}(i)$ is the number of those paths that pass through 'i'.

- In-Degree Centrality (CDin(i)C_{D_{in}}(i)CDin(i)): The proportion of incoming edges directed toward node iii, representing its popularity or importance.

$$C_{Din}(i) = \frac{k_{in}(i)}{N-1}$$

where kin(i)k_{in}(i)kin(i) is the in-degree of node iii, and NNN is the total number of nodes in the network.

- Out-Degree Centrality (CDout(i)C_{D_{out}}(i)CDout(i)): The proportion of outgoing edges from node iii, representing its influence or activity in the network.

$$C_{D_{out}}(i) = \frac{k_{out}(i)}{N-1}$$

  - Closeness Centrality: Measures how close a node is to all other nodes in the network.

$$C_C(i) = \frac{1}{\sum_j d(i,j)}$$

where d(i, j) is the shortest path distance from 'i' to 'j'.

- Eigenvector Centrality
- Concept: Eigenvector centrality is a measure of a node's influence based on the centrality of its neighbors. Nodes connected to other high-centrality nodes receive a higher eigenvector centrality score.
- Formula:
- Eigenvector centrality for a node i is given by:

$$C_E(i) = \frac{1}{\lambda} \sum_j A_{ij} C_E(j)$$

- where AijA_{ij}Aij is the adjacency matrix of the network, CE(j)C_E(j)CE(j) is the eigenvector centrality of node jjj, and λ\lambdaλ is a constant (eigenvalue).
-
- Steps:
  - Calculate Centralities: Use NetworkX functions to compute above mentioned centralities for each node.
  - Identify Key Nodes: Sort nodes by their centrality scores to identify those with the highest influence or connectivity.
  - Visualize: Plot the centrality scores or highlight key nodes within the network visualization.

This step identifies nodes that are critical for maintaining connectivity and may have strategic importance.

6. Visualization of the Network

- Visualizing the network is a crucial step in understanding its structure, relationships, and key nodes. In this project, we employ various tools to create an intuitive and interactive representation of the network.

6.1 NetworkX:

- NetworkX is a powerful Python library for network analysis, providing functionalities to create, manipulate, and study the structure and dynamics of networks. We use NetworkX for:
- Graph Construction: Creating the network based on input data, including nodes and edges, and assigning attributes like weight or type.
- NetworkX's capabilities allow for quick visualization and preliminary analysis, though it is limited in creating interactive or web-based visualizations.

6.2 Dash

- Dash, a framework developed by Plotly, is ideal for building interactive and web-based data applications. It allows for:
- Interactive Visualization: Creating dashboards where users can interact with visualizations, filter data, or explore specific nodes and edges within the network.
- Customizability: Incorporating various UI components like sliders, dropdowns, and buttons to control the view of the network. For example, users can adjust node size or filter nodes based on centrality metrics.
- Integration with Plotly: Leveraging Plotly's graphing library, Dash enables high-quality, interactive visualizations that are accessible via web browsers.
- Dash is particularly useful for sharing network insights with stakeholders, as it allows them to interact with and explore the network without needing technical expertise.

6.3 Folium

- Folium is a Python library for creating interactive maps, making it an excellent tool for visualizing networks with a geographic component. In this project, Folium is used to:
- Map-Based Visualization: Display nodes and edges overlaid on geographic maps, providing spatial context to the network. For example, nodes can represent physical locations, with edges indicating relationships or interactions between them.
- Marker Customization: Folium allows for customized markers and pop-ups, providing additional information about each node (e.g., centrality metrics, node attributes).
- Integration with Tile Maps: Folium supports different map styles, such as OpenStreetMap and Stamen, enabling visualization at various geographic scales.

By combining Folium with other tools, we create a more comprehensive view of the network's spatial structure, enhancing our ability to analyze geographically distributed networks.

RESULTS AND INFERENCES

1. Generic Results:-
- Nodes (cities): 456
- Type of Graph: Directed
- Edges: 71959

- Average path length: 1.6357
- Maximum path length: 3 round trips
- Diameter: 3 round trips
- Clustering Coefficient : 0.8052

- Average In-Degree: 157.80
- Average Out-Degree: 157.80
- Weighted Average Degree: -54931.40
- Connected components: 1

- Longest airplane route: 2855 min
- Shortest airplane route: 10 min
- Maximum time delay between any two consecutive city: 2835 min
- Minimum time delay between any two consecutive city: 1 min
- Average total airplane route time: 54931.4 min
- Average time delay between consecutive stops: 0.76 min

2. Most "popular" nodes(based on degree centrality value):

City Name	Degree Centrality value
• Los Angeles,CA	1.9472
• San Francisco,CA	1.9142
• Las Vegas,NV	1.8923
• Chicago,IL	1.8879
• Dallas/FortWorth,TX	1.8879

3. Bridge nodes(based on betweenness centrality value):

City Name	Betweenness Centrality value
• Los Angeles,CA	0.017566
• Denver,CO	0.016632
• New York,NY	0.015964
• Toronto,ON	0.015952
• SanFrancisco,CA	0.015807

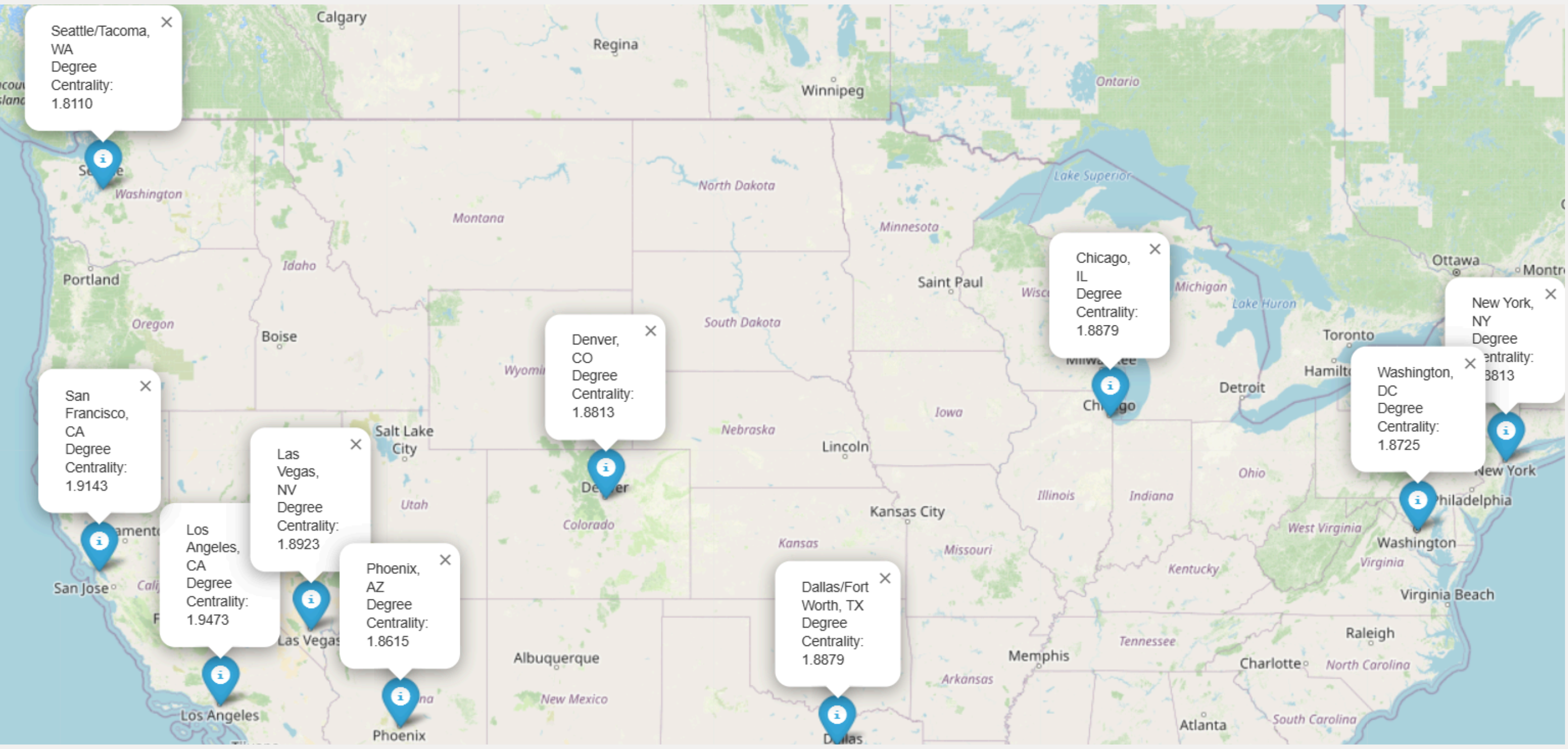
4. Most Influential nodes(based on closeness centrality value):

City Name	Closeness Centrality value
• Los Angeles,CA	0.97430
• San Francisco,CA	0.97014
• Dallas/FortWorth,TX	0.95188
• Chicago,IL	0.94594
• LasVegas,NV	0.94594

5. Degree Centrality Results(Top 10 Cities):

City Name	Degree Centrality
1. Los Angeles, CA	1.9473
2. San Francisco, CA	1.9143
3. Las Vegas, NV	1.8923
4. Chicago, IL	1.8879
5. Dallas/Fort Worth, TX	1.8879
6. Denver, CO	1.8813
7. New York, NY	1.8813
8. Washington, DC	1.8725
9. Phoenix, AZ	1.8615
10. Seattle/Tacoma, WA	1.8110

- **Degree Centrality on Map:**

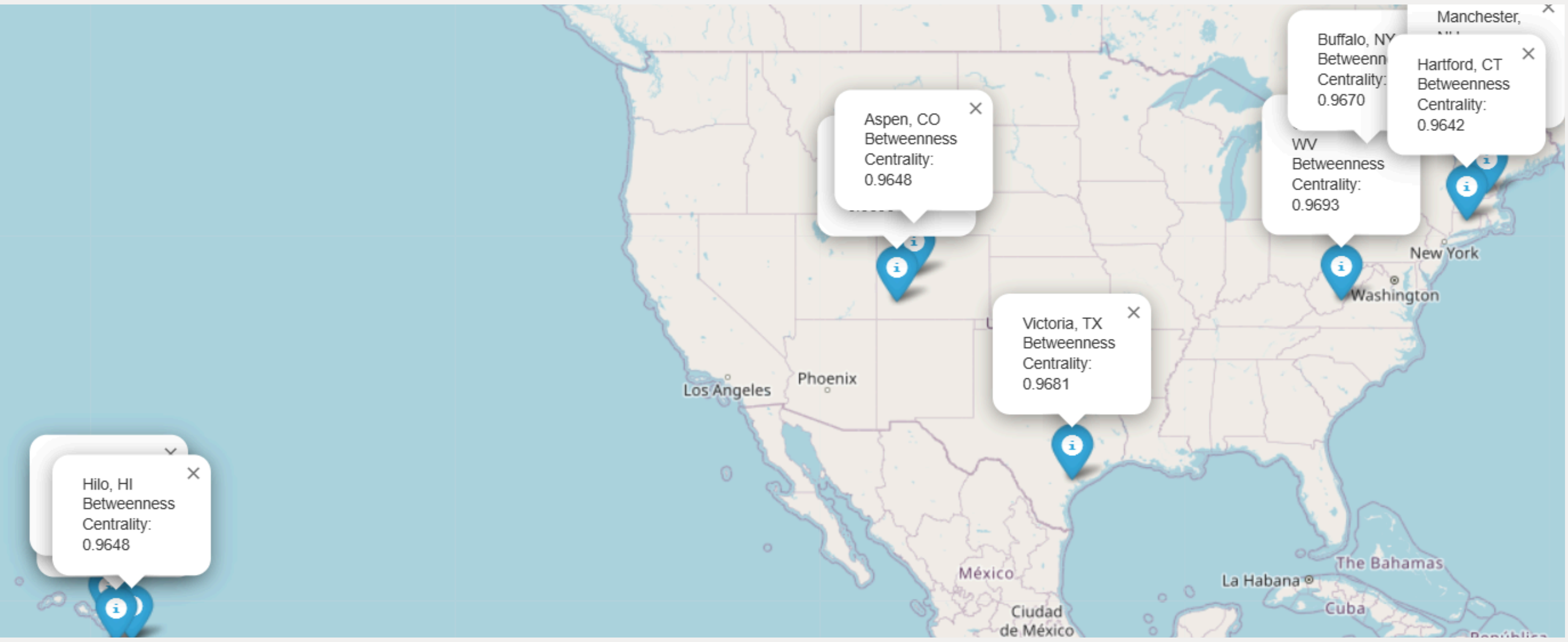


- Inference for Degree Centrality: Cities like Los Angeles, CA, and San Francisco, CA, show high degree centrality, signifying these cities are directly connected to many others. This indicates that these cities serve as major hubs within the network, offering high accessibility and acting as significant connection points in terms of direct links to other nodes.

6. Betweenness Centrality Results(Top 10 Cities):

• City Name	Betweenness Centrality
1. **Kona, HI**	**0.9714**
2. **Kahului, HI**	**0.9699**
3. **Greenbrier, WV**	**0.9693**
4. **Telluride, CO**	**0.9690**
5. **Victoria, TX**	**0.9681**
6. **Buffalo, NY**	**0.9670**
7. **Manchester, NH**	**0.9669**
8. **Aspen, CO**	**0.9648**
9. **Hilo, HI**	**0.9648**
10. **Hartford, CT**	**0.9642**

- **Betweenness Centrality on Map:**

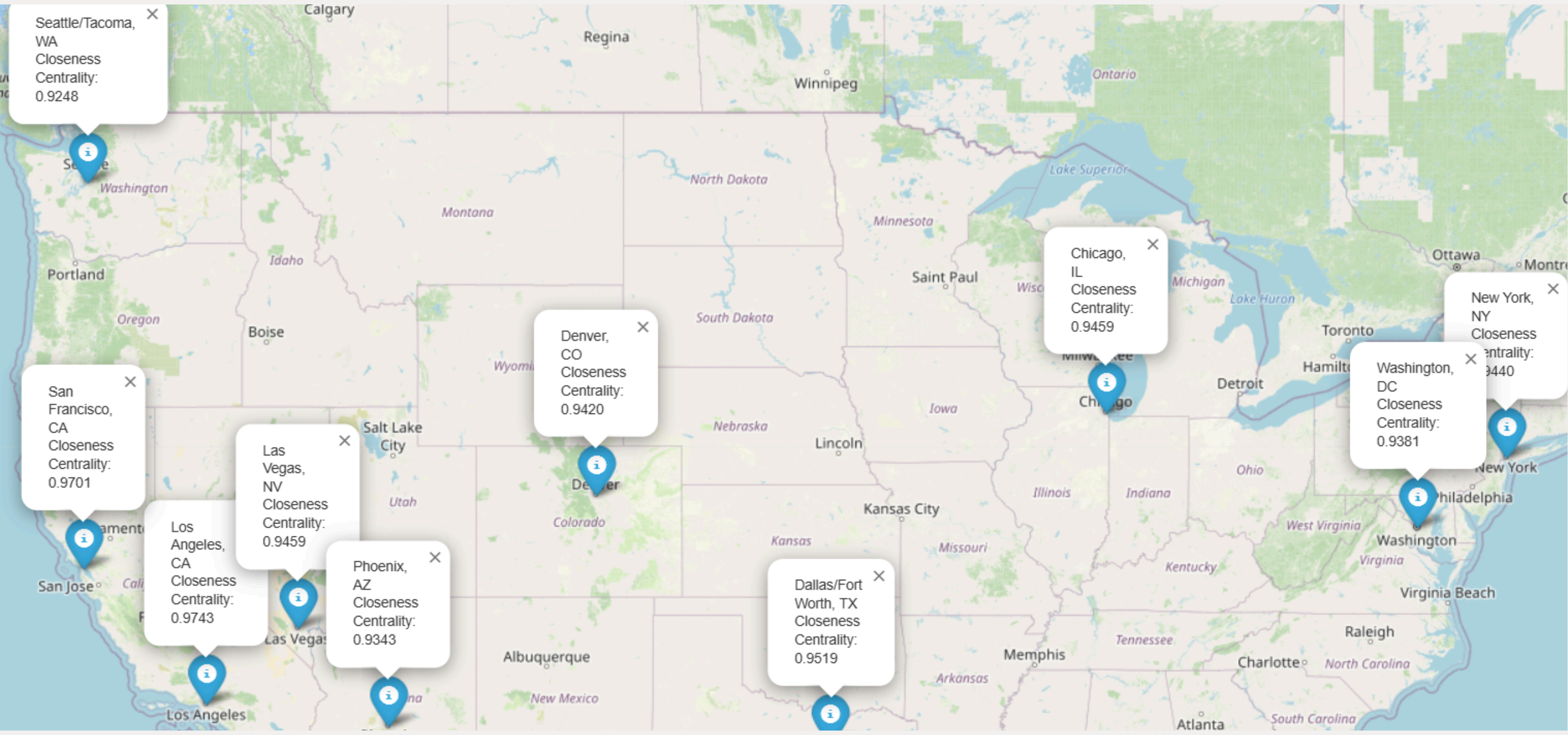


- Inference for Betweenness Centrality: Cities like Kona, HI, and Kahului, HI, have high betweenness centrality, suggesting they act as critical transit points or "gateways" within the network. These cities likely facilitate connections between less central parts of the network, functioning as essential passage points for movement and communication, potentially having a large impact on the overall flow across the network.

7. Closeness Centrality Results(Top 10 Cities):

City Name	Closeness Centrality
1. **Los Angeles, CA**	**0.9743**
2. **San Francisco, CA**	**0.9701**
3. **Dallas/Fort Worth, TX**	**0.9519**
4. **Chicago, IL**	**0.9459**
5. **Las Vegas, NV**	**0.9459**
6. **New York, NY**	**0.9440**
7. **Denver, CO**	**0.9420**
8. **Washington, DC**	**0.9381**
9. **Phoenix, AZ**	**0.9343**
10. **Seattle/Tacoma, WA**	**0.9248**

- **Closeness Centrality on Map:**

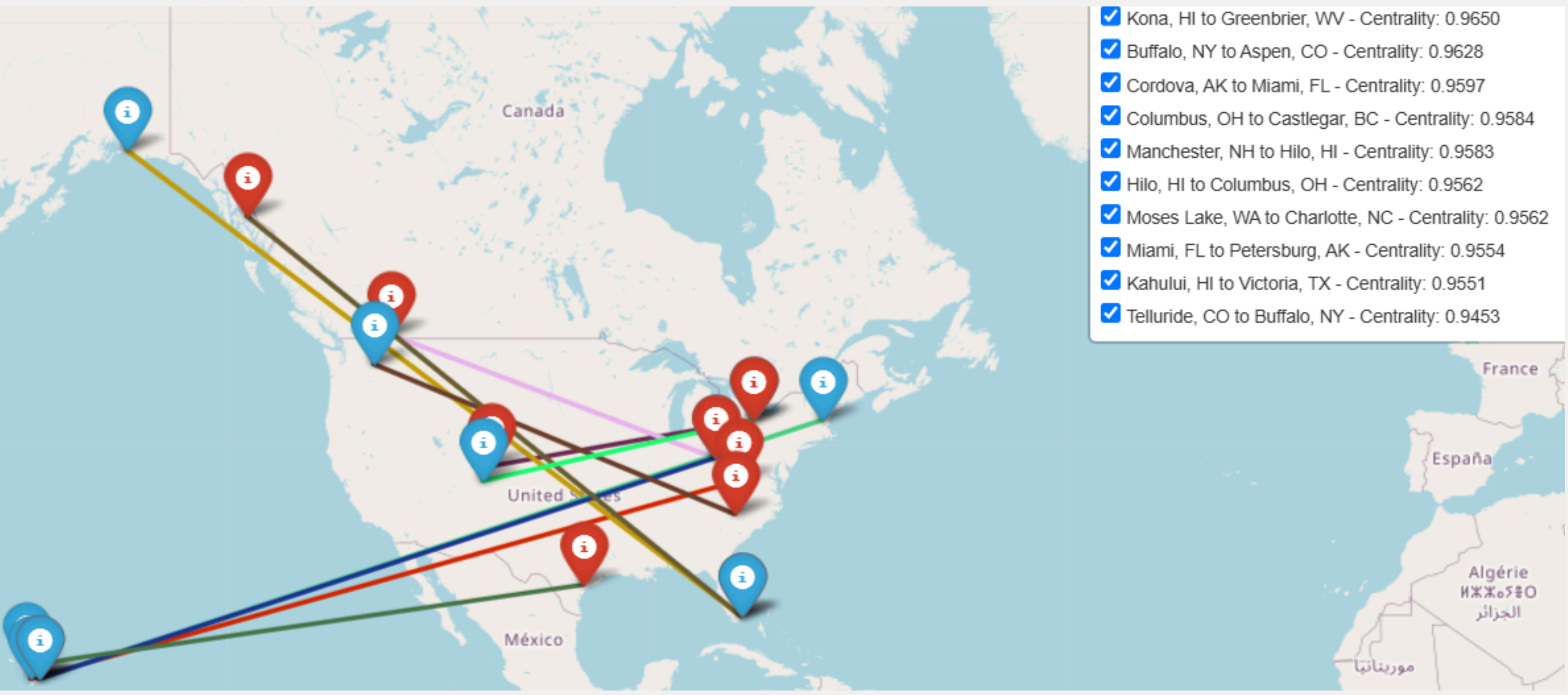


- Inference for Closeness Centrality: Los Angeles, San Francisco, and Dallas/Fort Worth exhibit the highest closeness centrality scores, suggesting they are well-connected hubs. They are likely in positions that allow them to reach other cities quickly, making them central to this network.

8. Edge Betweenness Centrality Results(Top 10 Cities):

City 1 Name	City 2 Name	Edge Betweenness Centrality
1. Kona, HI	Greenbrier, WV	0.9650
2. Buffalo, NY	Aspen, CO	0.9628
3. Cordova, AK	Miami, FL	0.9597
4. Columbus, OH	Castlegar, BC	0.9584
5. Manchester, NH	Hilo, HI	0.9583
6. Hilo, HI	Columbus, OH	0.9562
7. Moses Lake, WA	Charlotte, NC	0.9562
8. Miami, FL	Petersburg, AK	0.9554
9. Kahului, HI	Victoria, TX	0.9551
10. Telluride, CO	Buffalo, NY	0.9453

- **Edge Betweenness Centrality on Map:**

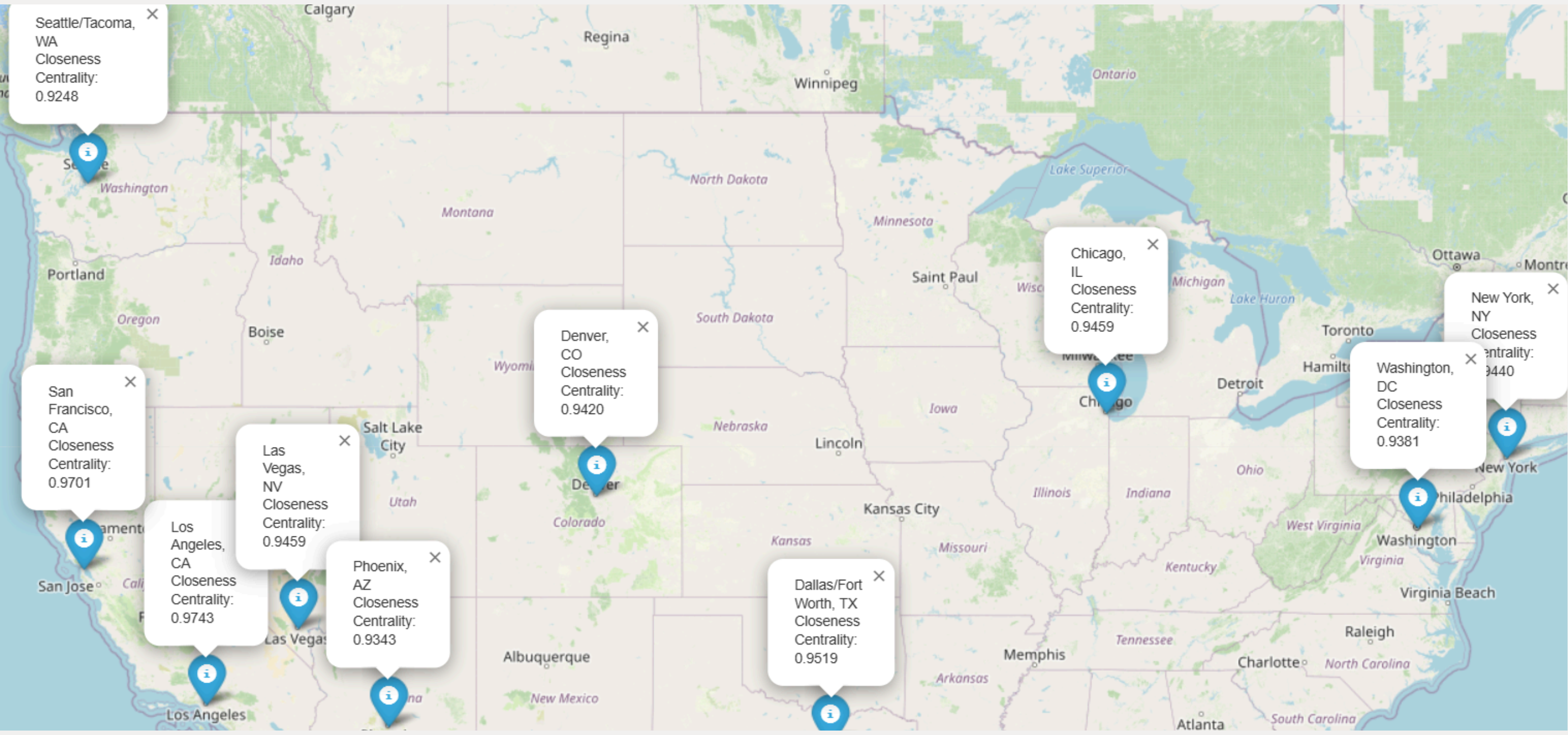


- Inference for Edge Betweenness Centrality: The edges between cities such as Kona and Greenbrier, and Buffalo and Aspen have the highest edge betweenness centrality, indicating that they play key roles in facilitating communication or travel between different parts of the network. These connections are pivotal "bridges" that link otherwise distant or less connected regions and are likely critical for ensuring network connectivity.

8. Eigenvector Centrality Results(Top 10 Cities):

City Name	Eigenvector Centrality
1. Los Angeles, CA	0.9743
2. San Francisco, CA	0.9701
3. Dallas/Fort Worth, TX	0.9519
4. Chicago, IL	0.9459
5. Las Vegas, NV	0.9459
6. New York, NY	0.9440
7. Denver, CO	0.9420
8. Washington, DC	0.9381
9. Phoenix, AZ	0.9343
10. Seattle/Tacoma, WA	0.9248

- **Eigenvector Centrality on Map:**

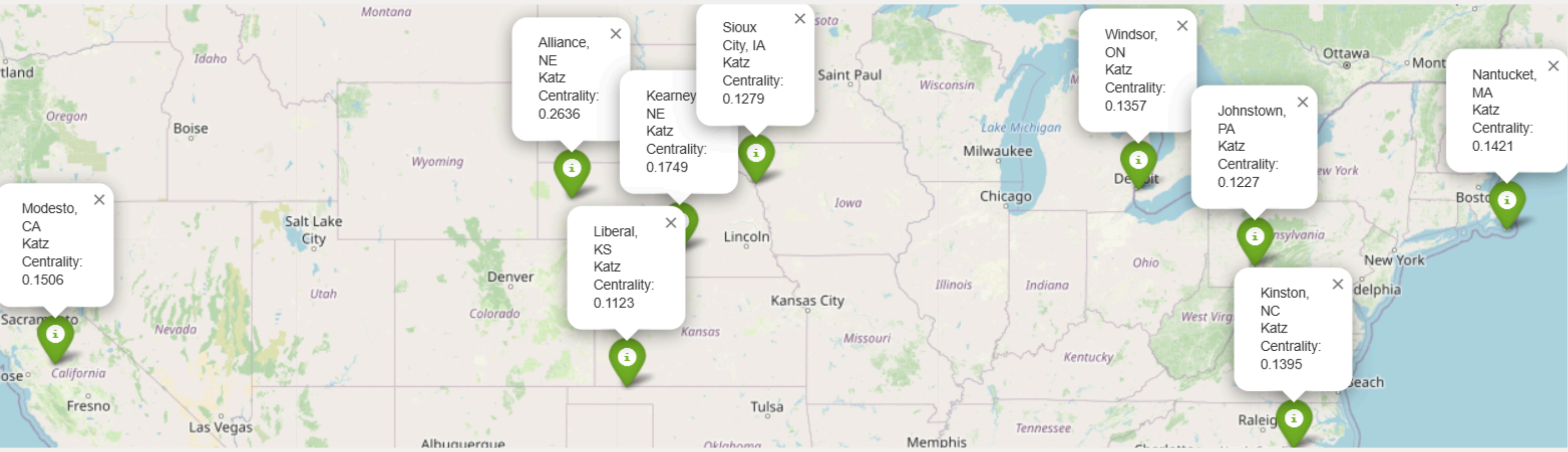


- Inference for Eigenvector Centrality: Los Angeles, San Francisco, and Dallas/Fort Worth exhibit the highest closeness centrality scores, suggesting they are well-connected hubs. They are likely in positions that allow them to reach other cities quickly, making them central to this network.

9. Katz Centrality Results(Top 10 Cities):

City Name	Katz Centrality
1. **Alliance, NE**	**0.2636083315033205**
2. **Kenai, AK**	**0.1840938071323797**
3. **Kearney, NE**	**0.1748748180545355**
4. **Modesto, CA**	**0.1506024645953630**
5. **Nantucket, MA**	**0.1421094658098384**
6. **Kinston, NC**	**0.1395459750630394**
7. **Windsor, ON**	**0.1356999356576483**
8. **Sioux City, IA**	**0.1278511318426642**
9. **Johnstown, PA**	**0.1226887372282293**
10. **Liberal, KS**	**0.1123072108683634**

- **Katz Centrality on Map:**

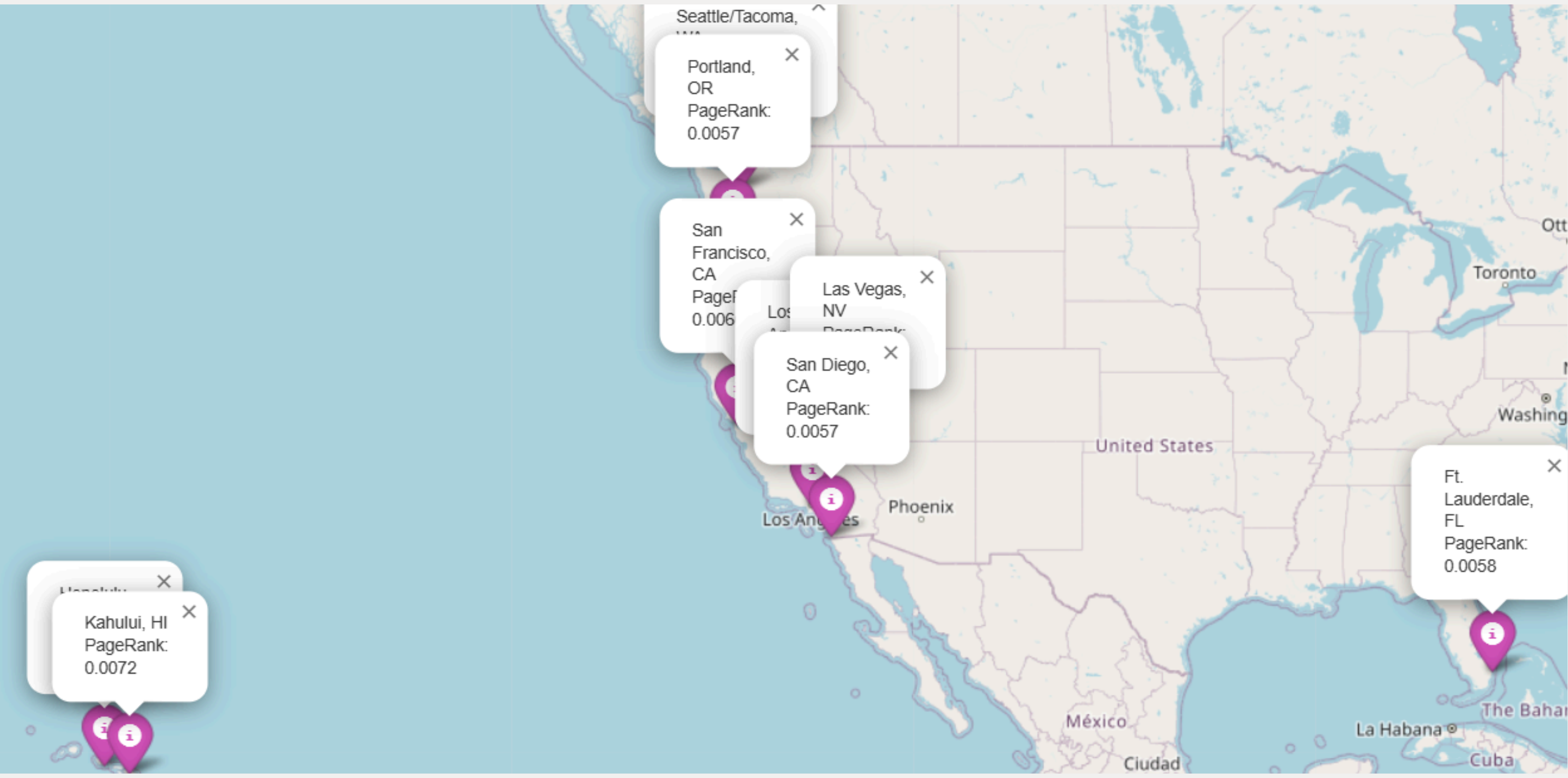


- Inference for Katz Centrality: Nodes such as Alliance and Kenai have relatively high centrality values, signifying that they are influential not necessarily because of direct connections but due to their proximity to other well-connected nodes. These nodes may not connect to many destinations directly but are effectively placed near key hubs, allowing influence and communication to propagate effectively through them. Cities with high Katz Centrality likely play a role in amplifying or facilitating indirect connections, often bridging outlying areas with more central nodes. As a result, they contribute significantly to the cohesion of the network, helping to integrate more isolated areas by connecting through central nodes.

10. Page Rank Results(Top 10 Cities):

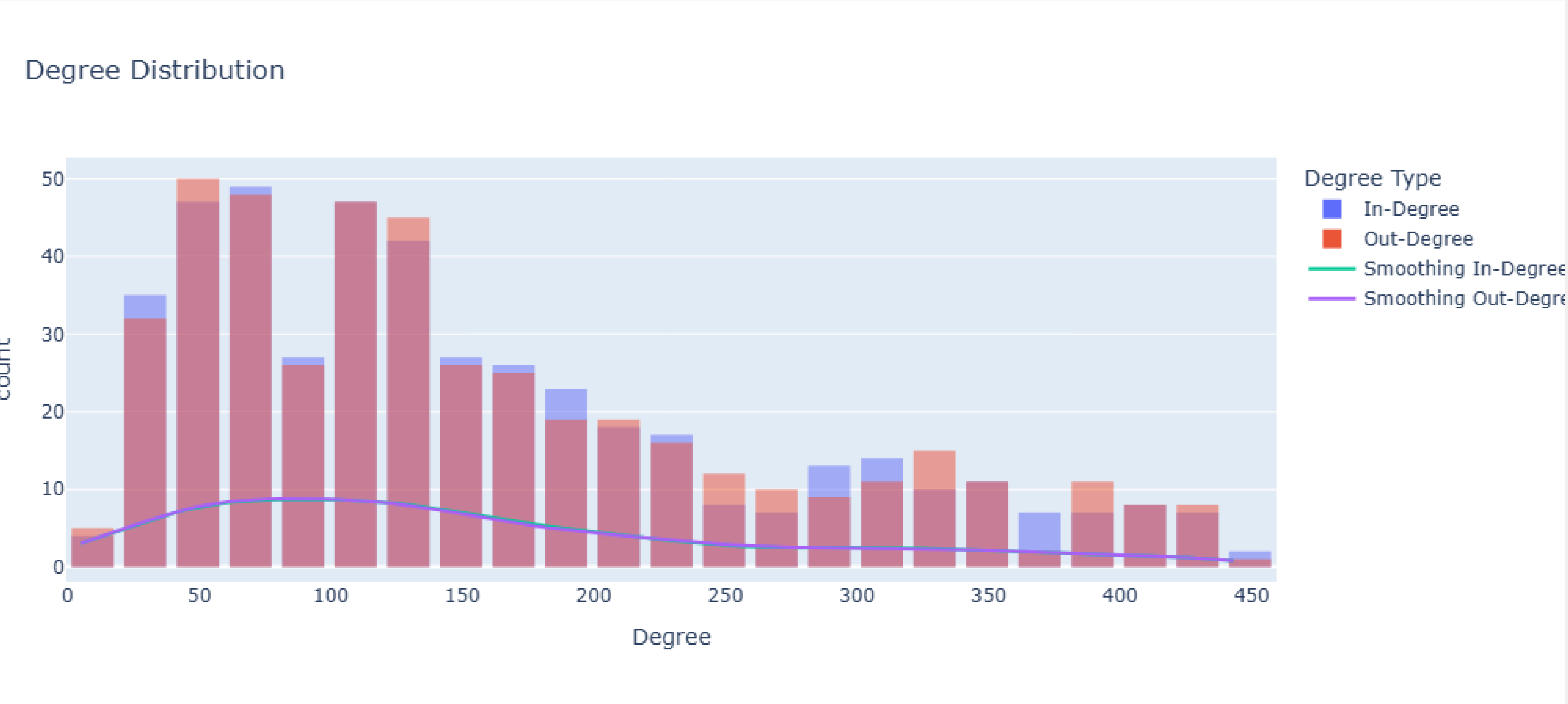
City Name	Page Rank
1. Honolulu, HI	0.0082
2. Kahului, HI	0.0072
3. San Francisco, CA	0.0066
4. Anchorage, AK	0.0065
5. Los Angeles, CA	0.0062
6. Ft. Lauderdale, FL	0.0058
7. Las Vegas, NV	0.0058
8. Seattle/Tacoma, WA	0.0058
9. Portland, OR	0.0057
10. San Diego, CA	0.0057

- **Page Rank on Map:**



- Inference for Page Rank: The PageRank scores indicate the relative importance of each node within the network, with higher scores suggesting that certain locations are more "influential" in terms of connectivity and accessibility. The leading PageRank scores in this network belong to cities like Honolulu and Kahului implying that these cities are key hubs that likely receive significant traffic and serve as vital transit points. Their influence extends to facilitating connections not only within their immediate surroundings but also across broader parts of the network. Although the differences in PageRank are relatively small, this distribution implies a slightly hierarchical network structure. Cities with higher PageRank scores, such as Honolulu and San Francisco, act as primary connection points, enhancing network cohesion. Meanwhile, cities with lower, yet still similar, PageRank scores (like Las Vegas and Seattle) serve as secondary hubs. This structure allows for both strong connectivity and resilience within the network, as multiple influential nodes can absorb traffic or connections if others become less accessible.

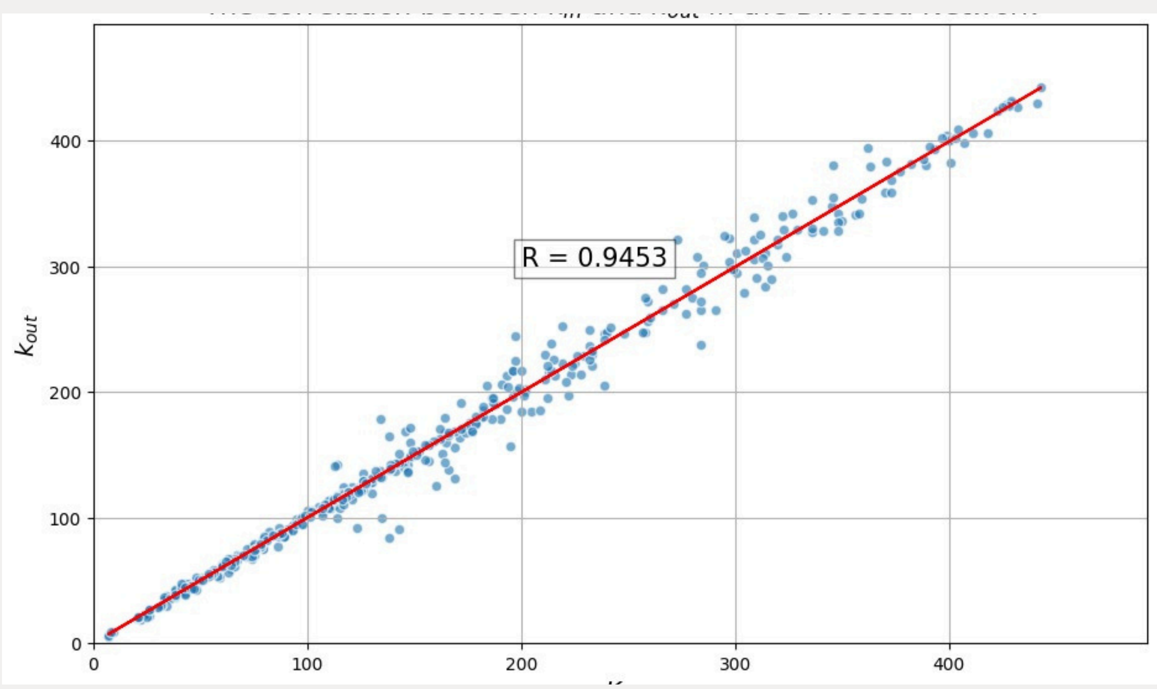
10. In-Degree V/S Out-Degree:



- **Key Insights:**

1. Out-Degree (in red) appears to dominate across most bins, indicating that nodes generally have more outgoing connections than incoming ones.
2. In-Degree (in blue) has relatively lower counts but remains consistent with Out-Degree trends.
3. Smoothing Lines: The smoothed lines provide insights into the distribution patterns for each degree type.
4. Both In-Degree and Out-Degree distributions taper off as the degree increases, showing that nodes with higher degrees are less common in the network.
5. The Out-Degree smoothing line (purple) is generally higher than the In-Degree (green), reaffirming the dominance of outgoing connections.
6. Distribution Shape: The distribution is right-skewed, which is common in real-world networks, where most nodes have few connections, and only a few nodes are highly connected.

11. Correlation Between k_{in} V/S k_{out}:-



Key Insights:

- Hub-and-Spoke Structure: The weak correlation between in-degrees and out-degrees reveals a hub-and-spoke pattern in the network. Cities often specialize as either departure hubs or arrival destinations, indicating an unbalanced role distribution across the network.
- High Out-Degree Departure Hubs: Cities with high out-degrees and low in-degrees, such as Atlanta (ATL) and Dallas (DFW), serve as major departure hubs. They are critical for routing passengers but are not primary destinations themselves.
- High In-Degree Destination Cities: Cities like New York City (JFK) and Chicago (ORD), with high in-degrees and low out-degrees, act as major arrival points, emphasizing their role as key destinations for flights originating from multiple locations.
- Low Reciprocity: The low reciprocity (R = 0.9453) indicates a predominantly unidirectional network structure, where routes are often driven by demand in one direction (e.g., business or tourism). This characteristic aligns with typical airline networks that cater to asymmetric travel flows.
- Network Vulnerability: The network's reliance on a few high-degree hubs makes it vulnerable to disruptions. If a central hub faces operational issues, many lower-degree cities dependent on these hubs could experience reduced connectivity, highlighting the need for robust network planning.

CONCLUSION

In this project, we effectively analyzed and visualized a network to identify key nodes and understand the structural characteristics of the network. Using centrality measures such as degree centrality and eigenvector centrality, we identified influential nodes that play significant roles within the network. The visualization tools, including NetworkX, Dash, and Folium, enabled us to gain insights into the network's structure and geographic context, while interactive features allowed for deeper exploration of relationships and clusters. This approach demonstrates the importance of network analysis in uncovering hidden patterns, optimizing connections, and making informed decisions in real-world applications. Future work could involve applying advanced algorithms or machine learning techniques to further refine our insights and predictions based on network dynamics.

REFERENCES

- **Hagberg, A., Swart, P., & S Chult, D. (2008). Exploring Network Structure, Dynamics, and Function using NetworkX. Los Alamos National Laboratory (LANL), LA-UR-08-05495.**
- **Natural Earth. (2023). ne_110m_admin_0_countries: Countries dataset. Retrieved from https://www.naturalearthdata.com/downloads/110m-cultural-vectors/110m-admin-0-countries/**
- **Folium. (2023). Folium Documentation: Interactive Maps. Retrieved from https://python-visualization.github.io/folium/**
- **Python Software Foundation. (2023). Dash: Analytical Web Applications. Retrieved from https://dash.plotly.com/**
- **Hunter, J.D. (2007). Matplotlib: A 2D Graphics Environment. Computing in Science & Engineering, 9(3), 90-95.**