



**UNIVERSIDAD PERUANA DE CIENCIAS APLICADAS**

**ESCUELA DE POSTGRADO**

**PROGRAMA DE MAESTRÍA EN DATA SCIENCE**

**TRABAJO FINAL**

**CURSO: Redes Complejas**

**AUTORES**

Jose Alexander Lopez Huertos

Juan Jose Agustin Navarro Gonzales

Yoel Yeremias Medina Quispialaya

**Lima, febrero de 2024**

# Índice

Introducción.....	3
Parte 1: Propuesta de red .....	4
Problema o Situación .....	4
Objetivos .....	5
Objetivo General.....	5
Objetivos Específicos .....	5
Descripción de la red.....	5
Nodos .....	5
Aristas.....	6
Importancia de la red .....	6
Parte 2: Descripción de la red .....	7
Características de la red.....	7
Grados de los nodos.....	7
Ponderación del grafo.....	8
Parte 3: Métricas.....	10
Nodos centrales.....	10
Grados de centralidad.....	10
Centralidad de la cercanía .....	11
Centralidad de intermediación .....	11
Centralidad armónica .....	12
Centralidad del vector propio .....	13
Grafos Aleatorios .....	14
Agrupamiento de la red .....	17
Clustering.....	17
Density.....	17
Pagerank.....	17
Comunidades .....	18
Louvain .....	18
Conclusiones y recomendaciones.....	20
Referencias Bibliográficas .....	22

## Introducción

En el ámbito de la logística y el transporte marítimo, la optimización de rutas constituye un pilar fundamental para mejorar la eficiencia operativa y satisfacer las demandas del mercado global. La naviera MSC, reconocida por su amplia presencia internacional, enfrenta el desafío de expandir su red de servicios en Perú, un mercado emergente con un potencial significativo para el comercio marítimo. A pesar de su posición privilegiada, la compañía se ve limitada por un enfoque tradicional que prioriza las rutas establecidas en función de la demanda histórica de sus clientes actuales. Esta aproximación, si bien pragmática, omite oportunidades latentes en rutas no exploradas, demandadas por otros actores del mercado que actualmente son atendidos por competidores.

La propuesta se centra en la implementación de un análisis de redes complejas, utilizando datos proporcionados por la entidad aduanera local, para mapear y evaluar los destinos comerciales más demandados dentro del territorio peruano. Este enfoque metodológico permitirá identificar conexiones entre las naves y los diferentes puertos, revelando así destinos potenciales que MSC podría incorporar a su red de servicios. Al contrastar los destinos existentes con las identificadas a través del análisis, se pretende descubrir brechas en el servicio que, de ser cubiertas, podrían posicionar a MSC como líder indiscutible en el mercado marítimo peruano, aprovechando al máximo las oportunidades de crecimiento en áreas hasta ahora no atendidas.

El resultado esperado de esta investigación es doble. Por un lado, proporcionará a MSC una base de datos analítica robusta sobre la cual tomar decisiones estratégicas para la expansión de sus rutas en Perú. Por otro lado, establecerá un marco de referencia para el uso de análisis de grafos en la industria naviera, demostrando su valor no solo como herramienta para la optimización de rutas existentes, sino como un mecanismo proactivo para la exploración y apertura de nuevos mercados. Esta propuesta no solo busca mejorar la cobertura y eficiencia de MSC, sino también potenciar el comercio marítimo en Perú, generando un impacto positivo en la economía local y en la cadena de suministro global.

## Parte 1: Propuesta de red

### Problema o Situación

La industria naviera juega un rol crucial en la economía global, facilitando el comercio internacional y conectando mercados a través de vastas redes marítimas. En Perú, un país con un litoral estratégicamente ubicado en el Pacífico Sur, el transporte marítimo es esencial para el comercio exterior. Sin embargo, la eficiencia y la cobertura de este servicio enfrentan desafíos significativos, particularmente en la adaptación de las rutas navieras a las cambiantes demandas del mercado. La compañía naviera MSC, a pesar de ser una de las líderes en el sector, ha limitado su expansión en el mercado peruano al no explorar destinos alternativos que podrían satisfacer una demanda creciente y diversificada.

Este escenario revela una desconexión entre las rutas establecidas por MSC y las necesidades reales del mercado peruano. La compañía, enfocándose en las demandas de un grupo selecto de clientes, ha pasado por alto el potencial de integrar nuevos destinos a su red de servicios. Esta situación se agrava por la falta de un análisis competitivo que contemple las operaciones de otras navieras en el país, las cuales han comenzado a captar una porción del mercado al ofrecer servicios a puertos no cubiertos por MSC. Como resultado, existe una oportunidad perdida no solo para la empresa sino también para el comercio peruano, que se ve limitado en sus opciones de transporte marítimo.

La solución a este dilema reside en la implementación de un enfoque más analítico y basado en datos para la planificación de nuevos destinos. Utilizando la información aduanera local, es posible realizar un análisis exhaustivo de los destinos más demandados dentro de Perú, identificando así aquellas áreas que están siendo desatendidas. Este tipo de análisis, centrado en las redes complejas, permite una visión holística del sistema de transporte marítimo, destacando no solo las rutas más eficientes sino también aquellas que podrían generar un mayor retorno comercial al ser incluidas en la red de MSC.

La adopción de esta estrategia analítica no solo beneficiaría a MSC mediante la expansión de su mercado y el aumento de su competitividad, sino que también impulsaría el comercio marítimo en Perú. Al cubrir destinos no explotados y responder de manera más ágil a las demandas del mercado, se facilitaría un mayor flujo comercial, beneficiando a una amplia gama de actores, desde productores locales hasta consumidores finales. Este enfoque, por ende, no solo es una cuestión de mejora operativa para una compañía, sino un paso hacia la optimización de la cadena logística en el país, contribuyendo así al crecimiento económico y a la integración comercial de Perú en el panorama global.

## Objetivos

### Objetivo General

Desarrollar un análisis de redes complejas que permita identificar y proponer nuevos puertos de destino para la naviera MSC en Perú, con el fin de optimizar su red de servicios y responder eficientemente a las demandas no cubiertas del mercado local, mejorando la cobertura y la competitividad de la empresa en el ámbito del transporte marítimo.

### Objetivos Específicos

1. Evaluar las capacidades y limitaciones de la infraestructura portuaria peruana, incluyendo la capacidad de manejo de carga de los puertos, la conectividad terrestre y las facilidades logísticas disponibles, para determinar su adecuación a las nuevas rutas identificadas.
2. Realizar un análisis competitivo que permita comprender las operaciones de otras navieras en el mercado peruano, identificando los destinos a los que van los competidores y evaluando las oportunidades de mercado para MSC a través de la diferenciación de servicios o la cobertura de destinos no explorados.
3. Desarrollar un modelo de red compleja que integre los datos de comercio, infraestructura portuaria y análisis competitivo, para simular diferentes escenarios de expansión de rutas y determinar las opciones más viables y rentables para MSC. Este modelo ayudará a visualizar las conexiones entre diferentes puertos y a identificar patrones de comercio que pueden ser aprovechados por la compañía.
4. Formular recomendaciones estratégicas para la expansión de la red de servicios de MSC en Perú, basadas en el análisis realizado. Esto incluirá planes de acción específicos para la implementación de nuevas rutas y destinos, consideraciones sobre alianzas estratégicas o acuerdos con operadores portuarios locales, y estrategias para la gestión de la capacidad y la optimización de las operaciones logísticas.

## Descripción de la red

### Nodos

- Naves: Este campo puede incluir el nombre de la nave o un código único que la identifica. Es esencial para distinguir entre las diferentes naves que forman parte de la flota de una compañía naviera y participan en el transporte marítimo.
- Puertos: Identificación del Puerto: Código o nombre que identifica al puerto. Esto es crucial para mapear las rutas de transporte y los puntos de carga y descarga de mercancías.

## Aristas

- Viajes de Naves: Son las interacciones realizadas de la nave a cada uno de los destinos registrados.

## Importancia de la red

La construcción de una red detallada que los destinos y los viajes de las naves, con una especial atención en la cantidad de contenedores transportados entre los puertos y las naves en el contexto peruano, tiene una importancia crítica en varios aspectos. Perú, con su estratégica ubicación en la costa del Pacífico Sur, es un país cuyo comercio exterior y crecimiento económico dependen significativamente del eficiente funcionamiento de sus puertos y de la optimización de sus rutas marítimas. La implementación de una red de este tipo permite un entendimiento profundo de las dinámicas comerciales y logísticas, esencial para fortalecer la posición de Perú en el comercio internacional.

Desde una perspectiva económica, esta red facilita la identificación de los destinos con mayor demanda, permitiendo a los operadores y a la industria naviera ajustar sus operaciones para maximizar la eficiencia y reducir los costos. Para una economía emergente como la peruana, donde sectores como la agricultura, minería y manufactura son pilares de exportación, optimizar el transporte marítimo puede traducirse en una ventaja competitiva sustancial en los mercados globales. Esto no solo mejora la balanza comercial del país, sino que también estimula la inversión en infraestructura portuaria y logística.

En el ámbito social y de desarrollo, una red optimizada contribuye significativamente a la creación de empleo y al desarrollo regional. Al identificar y explotar eficientemente los destinos, se fomenta el desarrollo de los puertos y las ciudades circundantes, generando oportunidades económicas y mejorando la calidad de vida de las comunidades locales. Además, al facilitar un comercio más ágil y económico, se promueve la inclusión de pequeños y medianos productores en la cadena de exportación, integrándolos de manera más efectiva en la economía global.

Desde la perspectiva de la sostenibilidad y el medio ambiente, la optimización de las rutas marítimas a través de una red detallada puede contribuir a la reducción de la huella de carbono del sector transporte, al minimizar las distancias navegadas y mejorar la eficiencia del combustible.

En conclusión, la importancia de desarrollar y utilizar una red de transporte marítimo en Perú radica en su capacidad para impulsar el comercio exterior, promover el desarrollo económico y regional, y contribuir a la sostenibilidad ambiental. Este enfoque no solo beneficia a la industria naviera y al sector de comercio exterior, sino que también tiene un impacto positivo en la economía nacional y el bienestar social, posicionando a Perú como un actor competitivo y responsable en el escenario internacional del comercio marítimo.

## Parte 2: Descripción de la red

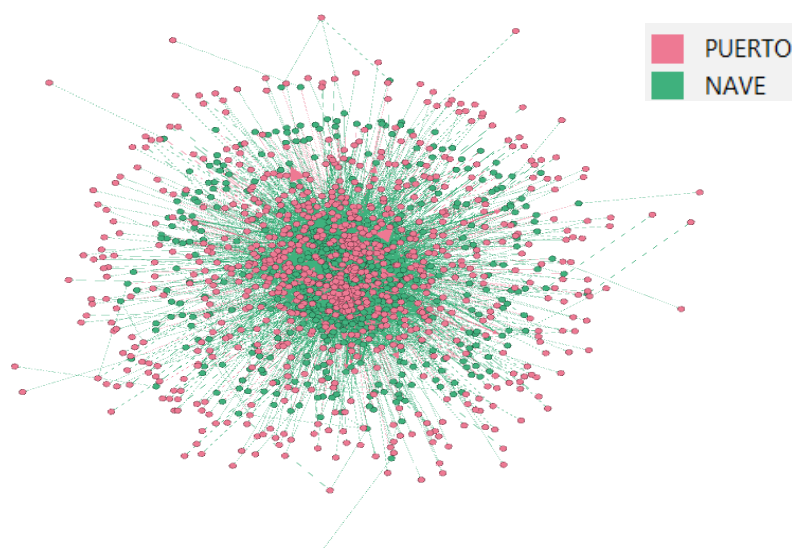
La red nos permite analizar los destinos de cada una de las naves que partieron del puerto del Callao el año 2023, en ese sentido, cada nave está vinculada a una empresa, pudiendo ver a qué destinos llegan las empresas y cuántos contenedores descargan en cada uno de los puertos a los que llegan.

### Características de la red

La red consta de 2 tipos de nodos: puertos (838) y naves (517), los cuales tienen 14015 aristas.

La visualización es la siguiente (generada con Gephi):

**Figura 1: Visualización de la red de naves y puertos de destino**



### Grados de los nodos

Los grados de los nodos son el número de conexiones que tiene un nodo específico en un grafo. Se define como el número de aristas que están conectadas a ese nodo. De acuerdo con la red propuesta en el presente trabajo se calculó el top 10 por tipo de nodo (diferenciando a naves y a puertos), al respecto se tiene lo siguiente:

Top 10 de Naves:

1. Nodo: 145, Nombre: COSCO SHIPPING VOLGA, Grado: 260
2. Nodo: 491, Nombre: VALPARAISO EXPRESS, Grado: 246
3. Nodo: 109, Nombre: CMA CGM CARL ANTOINE, Grado: 233
4. Nodo: 223, Nombre: GUAYAQUIL EXPRESS, Grado: 229
5. Nodo: 328, Nombre: MONTEVIDEO EXPRESS, Grado: 226
6. Nodo: 113, Nombre: CMA CGM ESTELLE, Grado: 215
7. Nodo: 72, Nombre: CALLAO EXPRESS, Grado: 205
8. Nodo: 437, Nombre: SANTOS EXPRESS, Grado: 200
9. Nodo: 106, Nombre: CMA CGM ARKANSAS, Grado: 185
10. Nodo: 124, Nombre: CMA CGM OHIO, Grado: 184

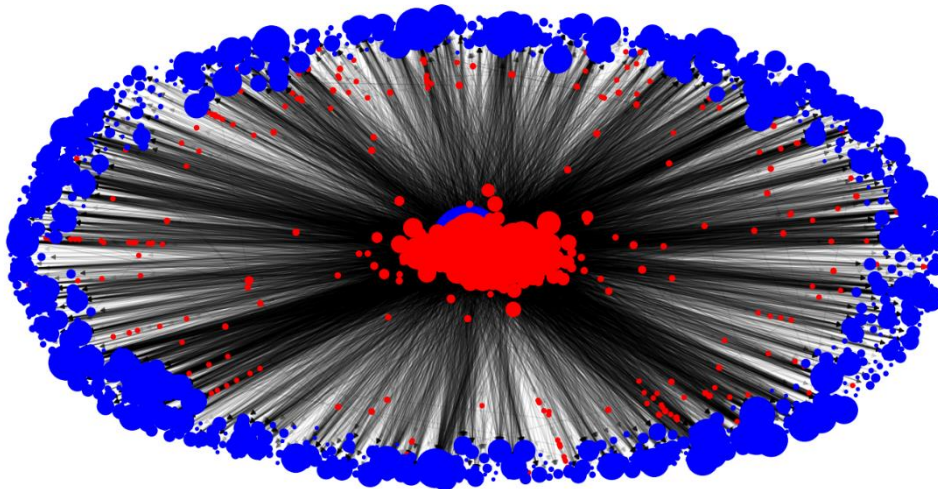
Top 10 de Puertos:

1. Nodo: 1, Nombre: PECLL, Grado: 517
2. Nodo: 620, Nombre: CLSAI, Grado: 170
3. Nodo: 654, Nombre: CNSHA, Grado: 161
4. Nodo: 946, Nombre: KRPUS, Grado: 149
5. Nodo: 661, Nombre: CNTAO, Grado: 132
6. Nodo: 649, Nombre: CNNGB, Grado: 129
7. Nodo: 834, Nombre: GTPRQ, Grado: 120
8. Nodo: 839, Nombre: HKHKG, Grado: 119
9. Nodo: 1216, Nombre: USLAX, Grado: 107
10. Nodo: 646, Nombre: CNLYG, Grado: 106

Los resultados muestran que, por el lado de las naves con mayor número de conexiones, se observa un número mayor grado de en comparación con los puertos, es decir las naves muestran más interacciones que los puertos. Sin embargo, hay que exceptuar al puerto del Callao (PECLL), ya que el mismo es el puerto de partida de todas las naves a cada uno de los destinos en la red.

Asimismo, se hizo una visualización de los nodos, donde el tamaño de cada nodo era proporcional al grado de este, teniendo lo siguiente:

**Figura 2: Visualización de la red de naves y puertos de destino, de acuerdo con el grado**



### Ponderación del grafo

En primer lugar, la variable elegida para realizar la ponderación del grafo es el número de containers descargados, es decir, la cantidad de contenedores que cada una de las naves dejaron en los puertos de destino. Sin embargo, la información correspondiente a los contenedores podría no ser la necesaria para el análisis, ya que, al ser información mostrada en términos absolutos, no permite comparar o tener en cuenta el factor del número de veces en las que se realizó el viaje al puerto de destino, por lo que una nave podría haber descargado la misma cantidad de contenedores que otra nave, pero en menos viajes, lo que la hace el puerto más importante.



En ese sentido, y con la finalidad de realizar una ponderación que sirva como herramienta para la toma de decisiones, se creó una razón, definida por lo siguiente:

$$C_{Viaje} = \frac{N^{\circ} \text{ de Contenedores descargados}}{N^{\circ} \text{ de viajes}}$$

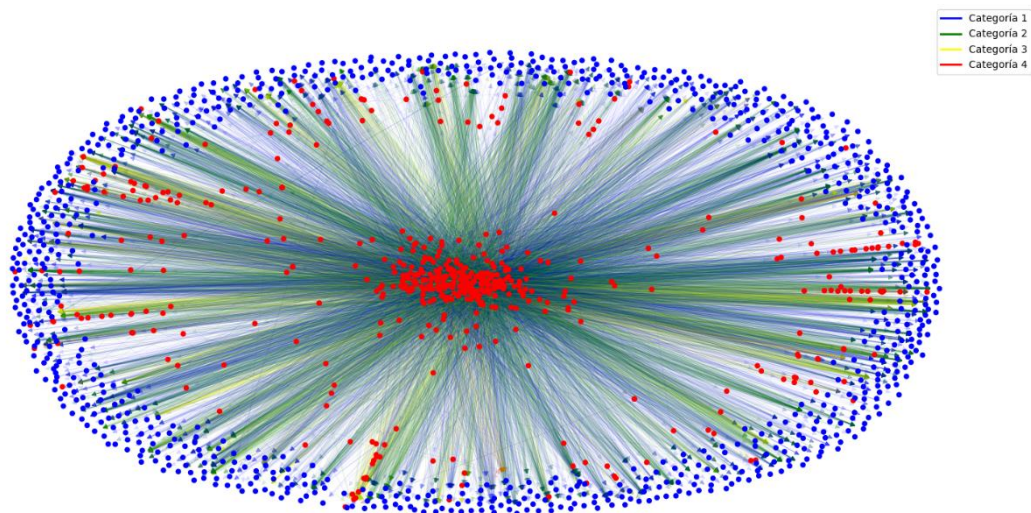
Dicha razón permite visualizar el número de contenedores que recibió el puerto X por parte de la nave Y en promedio por cada viaje realizado, identificado en el manifiesto.

Es así como, se realizó la visualización del grafo de acuerdo con la ponderación, donde también se categorizó en función del indicador que sirve como variable de ponderación:

- Categoría 1: 1 contenedor por viaje.
- Categoría 2: Entre 2 y 20 contenedores por viaje.
- Categoría 3: Entre 21 y 500 contenedores por viaje.
- Categoría 4: 500 contenedores por viaje a más.

La visualización será la siguiente:

**Figura 3: Visualización de la red de naves y puertos de destino, ponderado por el número de contenedores por viaje en el puerto de destino**



## Parte 3: Métricas

### Nodos centrales

El cálculo de medidas de centralidad en el grafo analizado servirá para observar la importancia de los nodos en la red que ha sido representada. En ese sentido, se calcularon las siguientes medidas de centralidad:

#### Grados de centralidad

Los grados de centralidad mostrarán la importancia de los nodos en función de la conectividad, es decir mostrar el número de interacciones que tiene un nodo, es decir, cuantas aristas tiene dicho nodo.

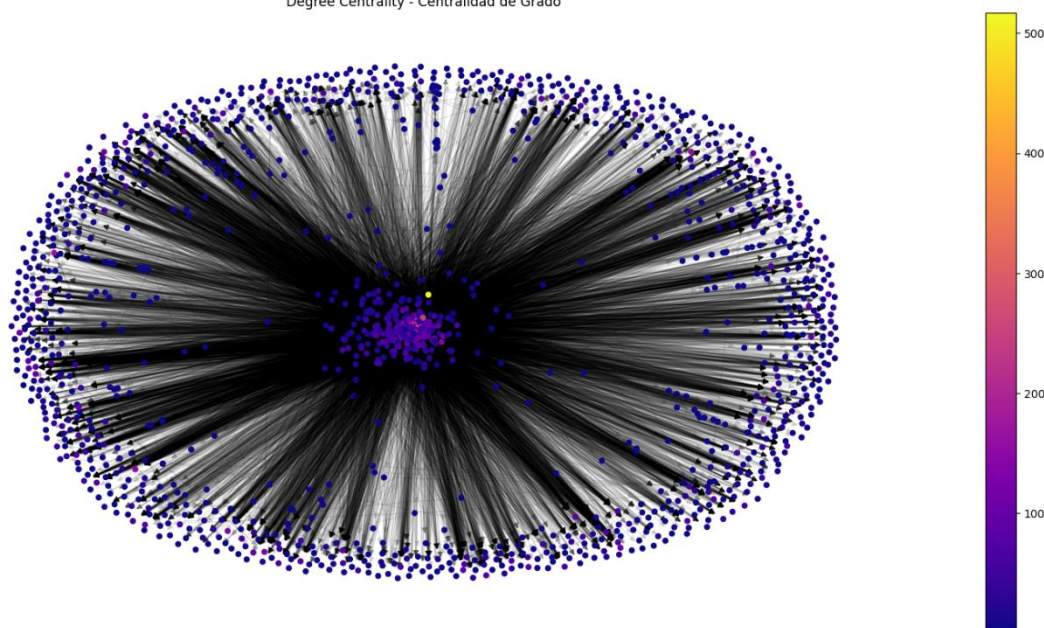
De acuerdo con lo calculado, los nodos con mayor grado son los siguientes:

Nodo	Grado	Nombre
1	164814	Puerto del Callao (Puerto)
670	21744	CNYTN (Yantian, China)
654	17400	CNSHA (Shanghai, China)
614	16049	CLLQN (Lirquen, Chile)
657	13958	CNSHK (Shekou, China)

La información se visualizará de la siguiente forma:

**Figura 4: Visualización de la centralidad del grado**

Degree Centrality - Centralidad de Grado



## Centralidad de la cercanía

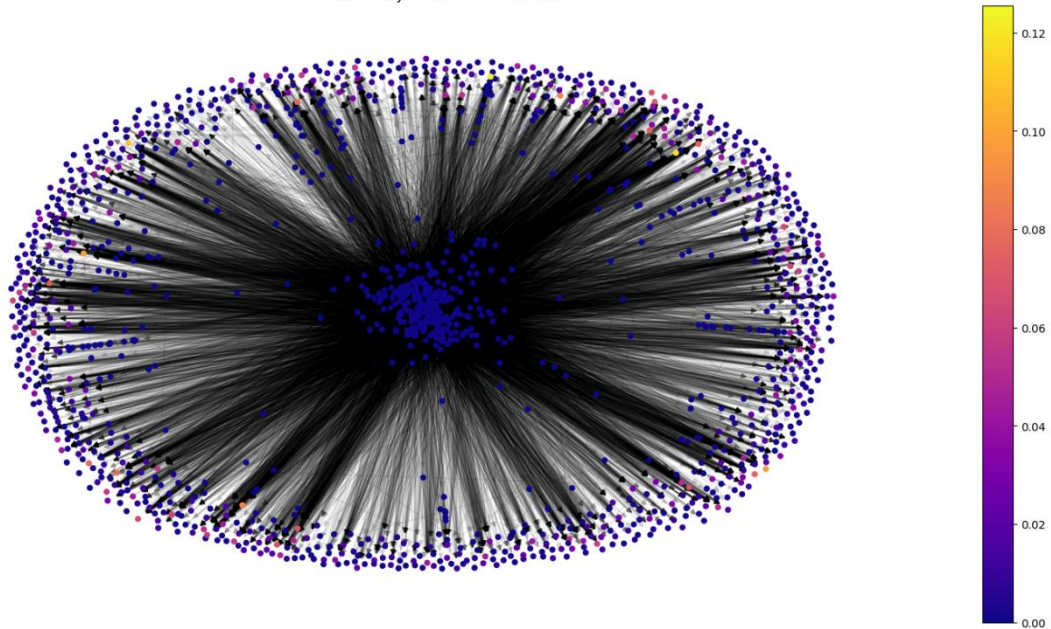
La centralidad de la cercanía es una medida que evalúa la proximidad de un nodo a los demás. En el caso del presente análisis el top 5 de nodos ordenados de acuerdo con su centralidad a la cercanía es el siguiente:

Nodo	Grado	Nombre
1351	0.030	ZADUR (Durban, RSA)
683	0.029	CRCAL (Caldera, Costa Rica)
546	0.028	AUMEL (Melbourne, Australia)
673	0.027	CNZJG (Zhangjiagang, China)
1103	0.022	SVAQJ (Acajutla, El Salvador)

La visualización es la siguiente:

**Figura 5: Visualización de la centralidad a la cercanía**

Closeness Centrality - Centralidad de Cercanía



## Centralidad de intermediación

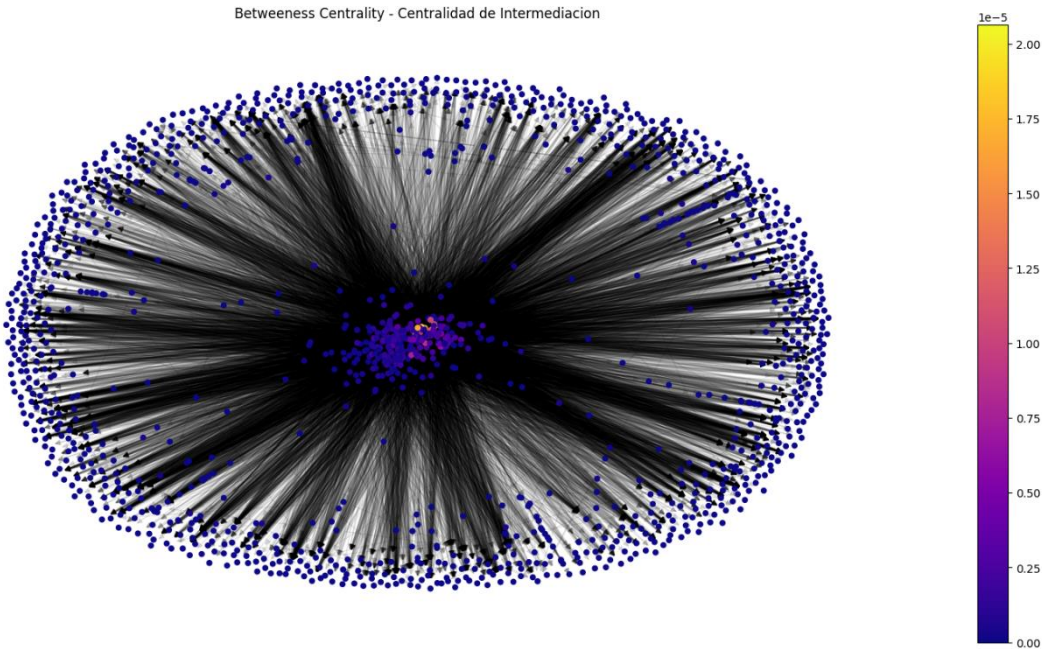
La centralidad de la cercanía es una medida que evalúa la importancia de un nodo como puente o intermediario del enlace entre otros nodos de la red.

El top 5 de nodos ordenados de acuerdo con su centralidad a la cercanía es el siguiente:

Nodo	Grado	Nombre
12	1.91e-05	ALGOL (Terminales Marítimos SAC)
145	1.36e-05	COSCO SHIPPING VOLGA
252	1.31e-05	JENS MAERSK
253	1.14e-05	JEPPESEN MAERSK
437	1.14e-05	SANTOS EXPRESS (Terminales Marítimos SAC)

La visualización es la siguiente:

**Figura 6: Visualización de la centralidad de la intermediación**



**Centralidad armónica**

La centralidad armónica es una medida que evalúa la proximidad de un nodo a los demás evaluando la capacidad de llegar a otros nodos de la red a través de caminos cortos en lugar de usar la distancia promedio. El top 5 de nodos ordenados de acuerdo con su centralidad a la cercanía es el siguiente:

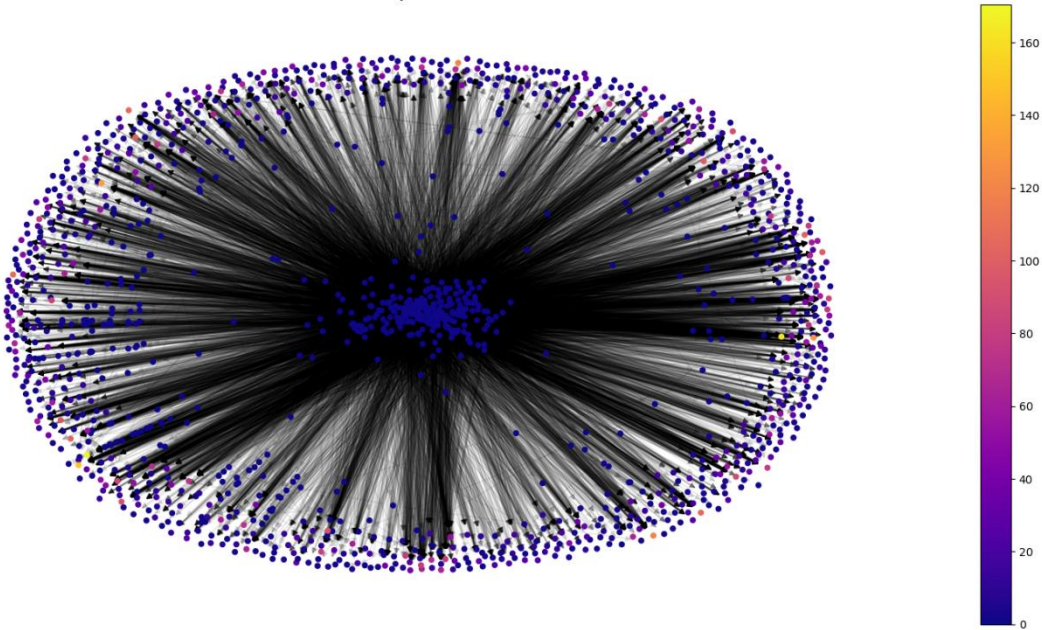
Nodo	Grado	Nombre
620	170.5	CLSAI - San Antonio (Chile)
654	161.5	CNSHA – Shangai (China)
946	149.5	KRPUS – Busan (Corea del Sur)
661	132.5	CNTAO – Qingdao (China)
649	129.5	CNNGB – Ningbo (China)



La visualización es la siguiente:

**Figura 7: Visualización de la centralidad armónica**

Harmonic Centrality - Centralidad Armónica



**Centralidad del vector propio**

La centralidad de la cercanía es una medida que evalúa la importancia a los nodos en función de la importancia de sus conexiones con otros nodos importantes en la red. En otras palabras, un nodo tiene una alta centralidad del vector propio si está conectado a otros nodos que a su vez también son importantes.

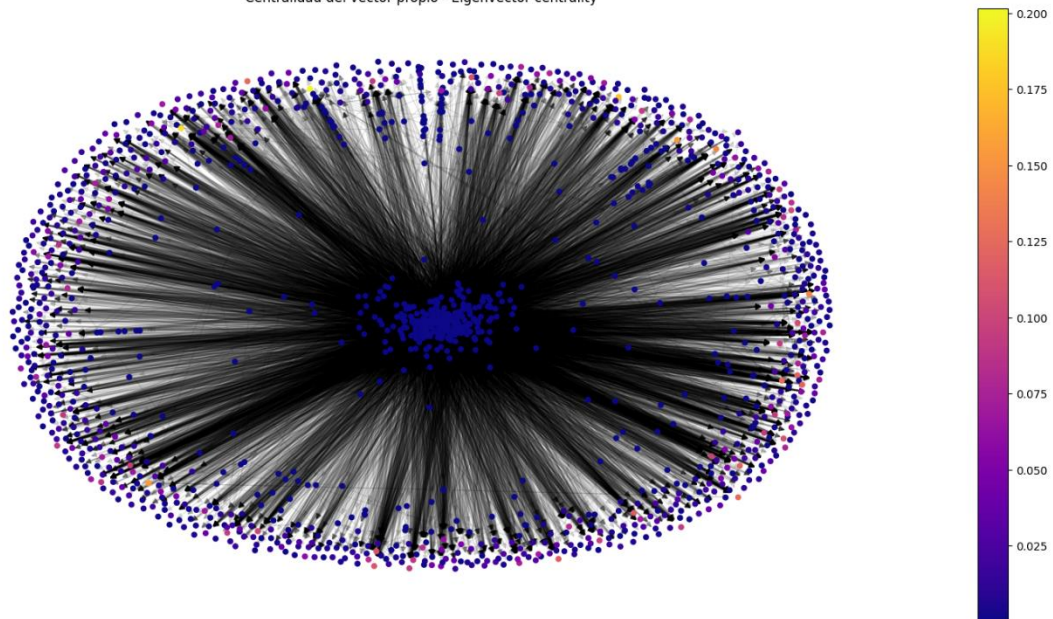
El top 5 de nodos ordenados de acuerdo con su centralidad a la cercanía es el siguiente:

Nodo	Grado	Nombre
670	0.529	CNYTN (Yantian, China)
654	0.453	CNSHA (Shanghai, China)
657	0.370	CNSHK (Shekou, China)
614	0.349	CLLQN (Lirquen, Chile)
615	0.286	CLMJS (Mejillones, Chile)

La visualización es la siguiente:

**Figura 8: Visualización de la centralidad del vector propio**

Centralidad del vector propio - Eigenvector centrality



## Grafos Aleatorios

Con la información de la red de las naves y los puertos destino de cada uno de los manifiestos representados en un grafo, se procedió a crear un grafo aleatorio. En ese sentido, al hacer uso del modo de Erdos-Renyi se construirá un grafo con un número fijo de nodos  $n$ , y cada par de nodos tiene una probabilidad  $p$  de estar conectados por una arista. Esto significa que cada arista se añade independientemente con una probabilidad  $p$ .

La aproximación que se utilizará para el valor crítico será la siguiente:

$$\frac{\log(n)}{n}$$

Donde  $n$  representa al número total de nodos en el grafo, los cálculos en Python fueron los siguientes:

```
n = len(list(rutas_nd.nodes()))
pstar = np.log(n)/n

pstar
```

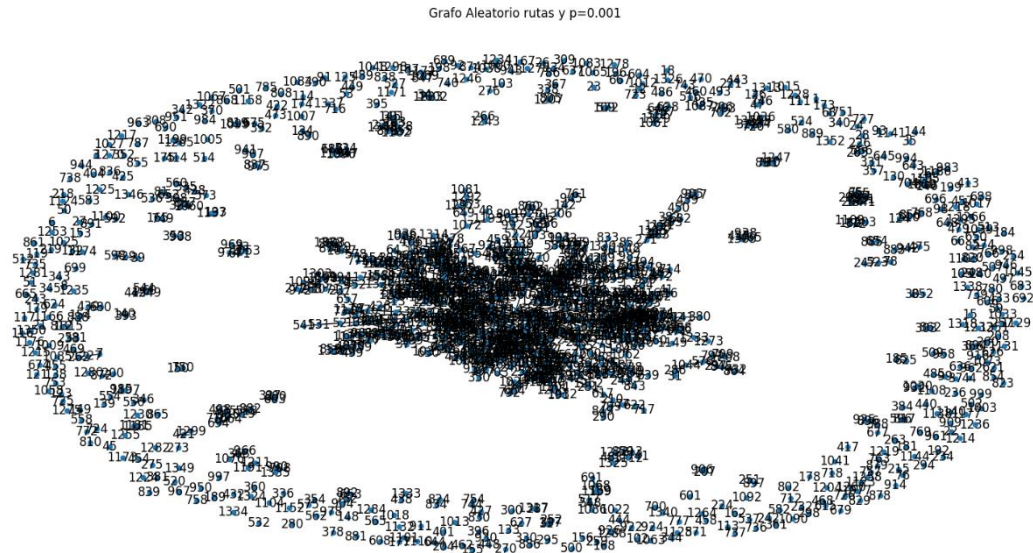
```
0.005322182091006496
```

El valor de 0.0053 muestra que la probabilidad de que una arista conecte a cualquier par de nodos es de 0.53%, este valor muestra entonces que hay una baja densidad de conexión ya que la mayoría de los nodos no están interconectados entre sí.

Asimismo, este valor crítico también muestra que la red es dispersa, por lo que de forma relativa los nodos tienen pocas conexiones.

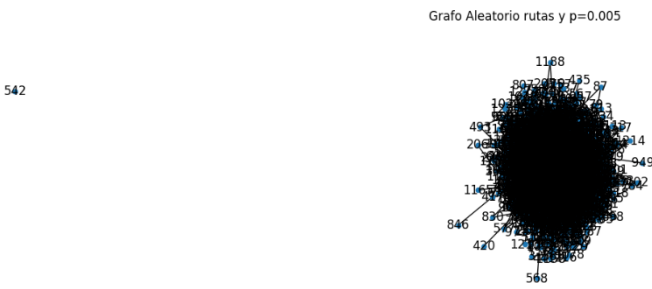
Se creó el primero grafo, con un valor menor al valor crítico ( $p=0.001$ ). Los resultados muestran que al ser el  $p$  menor al  $p^*$ , la probabilidad de que las aristas estén conectadas es menor, por lo que no será un grafo conectado, la visualización será la siguiente:

**Figura 9: Grafo aleatorio con  $p=0.001$**



El segundo grafo aleatorio tiene un  $p=0.005$ , valor similar al  $p^*$  del grafo analizado. La visualización será la siguiente:

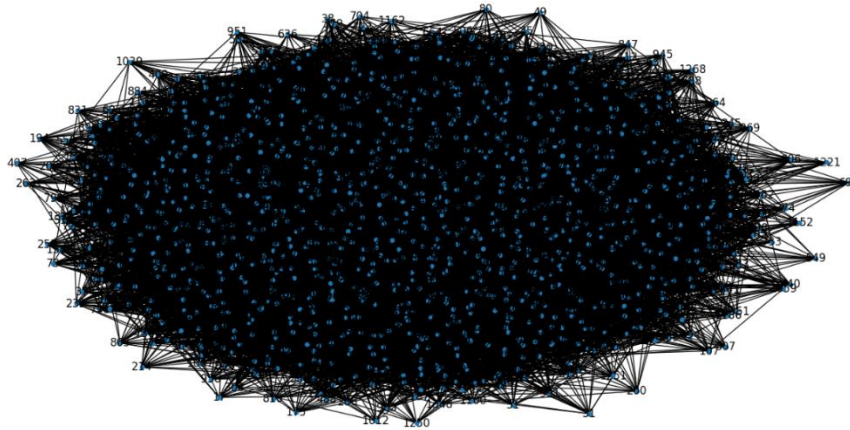
**Figura 10: Grafo aleatorio con  $p=0.005$**



El tercer grafo aleatorio tiene un  $p=0.02$ , valor mayor al de  $p^*$ . La visualización será la siguiente:

**Figura 11: Grafo aleatorio con  $p=0.02$**

Grafo Aleatorio rutas y  $p=0.02$



En el grafo, cuando la probabilidad de conexión entre dos nodos es menor que 0.005, esto quiere decir que es poco probable que exista un componente gigante en el grafo. Por otro lado, cuando la probabilidad de conexión es mayor que 0.005, es probable que exista un componente gigante en el grafo.

En el caso de  $p=0.001$ :

- Ya que el valor  $p$  es menor al crítico que ha sido calculado, se estima que es probable que existan nodos no conectados y/o componentes pequeños.
- Debido a que la probabilidad de conexión es baja y menor al valor crítico, la densidad de conexiones en el grafo será baja en comparación con otros grafos aleatorios con valores de conexión superiores al valor crítico.

En el caso de  $p=0.02$ :

- Ya que el valor es alto y muy superior al valor crítico, es muy probable que el grafo tenga un componente gigante. Esto significa que la mayoría de los nodos estarán conectados de alguna manera directa o indirectamente.
- En ese sentido, habrá una alta probabilidad de conexión entre nodos implica que la mayoría de los pares de nodos estarán conectados entre sí. En consecuencia, la densidad de conexiones en el grafo será alta en comparación con un grafo con un valor de conexión más bajo.
- Finalmente, ya debido al alto nivel de conexión entre nodos del grafo, es probable que existan caminos cortos entre la mayoría de los nodos. Esto podría hacer que el grafo tenga un alto nivel de conectividad y no existan distancias tan largas entre nodos.



Por otro lado, al visualizar el grafo con el valor  $p$  similar al  $p^*$  (0.005), se puede observar que hay una diferencia significativa en el número de aristas, por lo que se puede concluir que el grafo aleatorio no tiene el mismo comportamiento que la red inicial.

## Agrupamiento de la red

Las medidas de agrupamiento de la red indican qué tan conectados están sus vecinos entre sí.

### Clustering

Se calculó el coeficiente de clustering para cada nodo y luego del promedio para toda la red para poder tener información importante sobre la estructura de la red.

Los resultados mostraron que el coeficiente de clustering para cada nodo de la red fue igual a 0, lo que quiere decir que no existen conexiones entre los vecinos de cada uno de los nodos. Por consiguiente, el valor de del promedio del clustering de la red será de 0.

### Density

La densidad de un grafo es una medida de qué tan densamente conectado está el grafo. Se definirá entonces como el número de conexiones presentes en el grafo dividido por el número total de conexiones posibles entre los nodos del grafo.

El valor calculado fue de 0.007, lo que representa un valor relativamente bajo y quiere decir que este valor representa la proporción de aristas en el grafo en comparación el número de aristas posibles.

### Pagerank

El PageRank es un algoritmo utilizado para medir la importancia relativa de los nodos en una red, como un grafo dirigido o una red de enlaces web. Asigna a cada nodo de la red un valor numérico que representa su importancia relativa. Este valor se calcula en función de la estructura del grafo y del número y calidad de los enlaces que apuntan hacia un nodo en particular.

El algoritmo considera que un nodo es más importante si es enlazado por otros nodos importantes, lo que refleja la idea de que la relevancia de una página web aumenta si es enlazada por otras páginas web populares o relevantes.

El PageRank se puede interpretar como una medida de la "autoridad" o "importancia" de un nodo en una red. Se utiliza ampliamente en la clasificación de páginas web en motores de búsqueda, pero también tiene aplicaciones en otros campos, como la detección de

comunidades en redes sociales, la clasificación de documentos y la recomendación de productos.

En el contexto del presente análisis, podemos calcular el Pagerank para ver los puertos y naves que más importancia tienen, en ese sentido se calculó el Pagerank y el top 10 de nodos con mayor Pagerank:

Nodo	Grado	Nombre
620	0.0116	CLSAI - San Antonio, Chile
613	0.0087	CLIQQ - Iquique, Chile
1048	0.0081	PEZZD
807	0.0065	GBMLF - Puerto de Milford Haven, UK
634	0.0064	CNFAN - Fangcheng, China
661	0.0060	CNTAO - Qingdao, China
615	0.0048	CLMJS - Mejillones, Chile
654	0.0046	CNSHA - Shanghai, China
1042	0.0045	PABLB - Balboa, Panamá
834	0.0043	GTPRQ - Puerto Quetzal, Guatemala

## Comunidades

A partir del grafo analizado, se hará un análisis de las comunidades usando dos métodos: el de Louvain y el de Girvan-Newman:

### Louvain

Este método tiene como objetivo encontrar particiones en el grafo donde los nodos dentro de una partición estén densamente conectados entre sí, pero escasamente conectados con nodos fuera de esa partición.

```
louvain_com = algorithms.louvain(rutas_nd, weight='weight')

print('La cantidad de comunidades es {}'.format(len(louvain_com.communities)))
for c in louvain_com.communities:
    print(c)

La cantidad de comunidades es 5

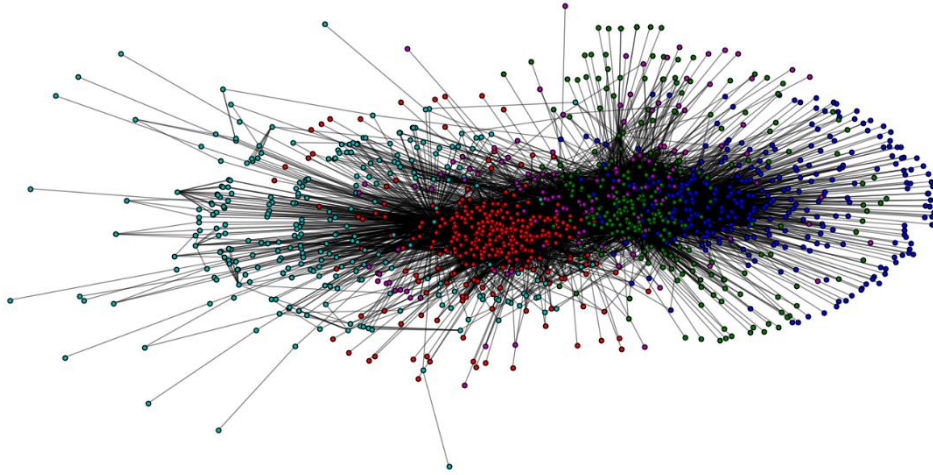
#mostrar la modularidad general
print(louvain_com.modularity_density())
#mostrar la modularidad de cada comunidad
print(louvain_com.internal_edge_density(summary=False))

FitnessResult(min=None, max=None, score=31.879713641769413, std=None)
[0.08063255152807391, 0.030730215559036326, 0.060780336475360935, 0.01047319296722379,
0.06346899224806202]
```

De acuerdo al algoritmo de Louvain, se encontraron cuatro comunidades, las mismas que tienen la modularidad mostrada en punto anterior. Los valores encontrados muestran que cada una de las comunidades no es significativa ya que es un valor cercano a 0.

La visualización de las comunidades será la siguiente:

**Figura 12: Grafo de rutas por comunidades**



## Conclusiones y recomendaciones

Debido al foco de la compañía en las demandas de un grupo selecto de clientes, se ha pasado por alto el potencial de integrar nuevos destinos a su red de servicios. Esta situación se agrava por la falta de un análisis competitivo que contemple las operaciones de otras navieras en el país, las cuales han comenzado a captar una porción del mercado al ofrecer servicios a puertos no cubiertos por MSC.

En ese sentido, el objetivo del trabajo es proponer un análisis de redes complejas que permita identificar y proponer nuevos puertos de destino para la naviera MSC en Perú, con el fin de optimizar su red de servicios y responder eficientemente a las demandas no cubiertas del mercado local.

La ponderación de la red se propuso inicialmente con el número de contenedores descargados, sin embargo, se identificó que dicha variable, mostrada en términos absolutos, no da un índice de la real capacidad de descarga de las naves en los respectivos puertos, por lo que se creó un índice que calcula el número de contenedores descargados por nave a cierto puerto específico y con dicho índice se hizo la ponderación respectiva.

En el análisis de centralidad se detectaron los puertos más importantes en relación con su grado, a la cercanía, a la intermediación, entre otros. Cabe resaltar que en el caso de los nodos identificados como más importantes con relación a la centralidad de la intermediación, los resultados mostraron a las diversas naves que sirven como enlace entre el puerto del Callao y los diferentes destinos registrados en el año 2023 y no se identificó en el top 5 a ninguna nave de la compañía en la que se centra el análisis.

Se analizó por medio de grafos aleatorios el comportamiento de la red, encontrando que el grafo aleatorio con el valor crítico similar a la red propuesta tiene un número mucho mayor de aristas, mostrando un comportamiento diferente en comparación a la red inicial.

Se calculó el coeficiente de clustering para cada nodo y luego del promedio para toda la red para poder tener información importante sobre la estructura de la red. Los resultados mostraron que el coeficiente de clustering para cada nodo de la red fue igual a 0, lo que quiere decir que no existen conexiones entre los vecinos de cada uno de los nodos. Por consiguiente, el valor de del promedio del clustering de la red será de 0.

En relación con el análisis de densidad del grafo, los resultados mostraron un valor de 0.007, lo que representa un valor relativamente bajo y quiere decir que este valor representa la proporción de aristas en el grafo en comparación el número de aristas posibles.

Finalmente, se proponen dos estrategias para poder mejorar la competitividad de la empresa dentro del mercado peruano. Ambas estrategias estarán basadas en la optimización de destinos para una mayor rentabilidad de la empresa.

La primera propuesta de solución al problema de acuerdo con lo analizado en la red sería generar un instrumento de detección de posibles destinos que priorice aquellos que tienen un alto grado de entrada, este se haría de la siguiente forma:

- Identificar los nodos con mayor grado de entrada, estos nodos serían los puertos que reciben la mayor cantidad de naves, ponderado con el número de contenedores descargados por viaje.
- Identificar a cuáles de estos destinos las naves de la empresa MSC no están llegando.
- Proponer y optimizar las rutas que ya tiene la empresa basada en aquellos destinos identificados.

La segunda propuesta de solución al problema de acuerdo con lo analizado en la red es también generar una herramienta para identificar los posibles destinos:

- Identificar los nodos con mayor centralidad de intermediación, que serán los nodos más importantes que sirven como enlace en la red.
- Identificar los destinos a los cuáles están llegando las naves que tienen mayor centralidad de intermediación.
- Proponer y optimizar las rutas que ya tiene la empresa basada en aquellos destinos identificados.

## Referencias Bibliográficas

Barabási, A.-L. (2016). Network science. Cambridge University Press. doi:10.1017/CBO9781316567637.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, 40(1), 35-41. doi:10.2307/3033543.

Girvan, M., & Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12), 7821-7826. doi:10.1073/pnas.122653799.

Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32(3), 245-251. doi:10.1016/j.socnet.2010.03.006.

Pizzuti, C. (2017). Python for network analysis. Apress. doi:10.1007/978-1-4842-3235-4.