**FASE1 DEL PROYECTO DEL PROYECTO**

1. **Carga de datos**

Para la fase inicial del análisis en Gephi, donde el objetivo es identificar actores relevantes de la red usando medidas de centralidad y transitividad se toma la decisión de no incluir ningún atributo, en los siguientes pasos del proyecto se hará la selección de atributos relevantes para la red de inversionistas.

Utilizando Python se crea un notebook para transformar y limpiar tanto el archivo de nodos como el de arcos, a continuación, se detalla los pasos aplicados:

1. **Importación de Datos:** El primer paso consistió en cargar los archivos de datos desde Excel

* Nodos.xlsx : Contiene información sobre las empresas e inversionistas.
* Arcos.xlsx: Representa las conexiones entre los actores de la red.

1. **Limpieza de Texto en Identificadores:** Para estandarizar los valores en los identificadores, se aplicó una función de limpieza de texto al campo Id en el archivo de nodos y a los campos Source y Target en el archivo de arcos.

Las Transformaciones aplicadas fueron: conversión texto a mayúsculas, eliminación de caracteres especiales (excepto números, comas y espacios) y eliminación de espacios extra.

1. **Se renombra el identificador de nodos:** se renombra "Id" y se renombró a "Label" para que Gephi la reconozca.
2. **Adicción de “Type” en los arcos:** Se añade en el archivo de arcos la columna “Type” indicando que la relación es "directed" (dirigida).

Una vez que los datos fueron transformados y exportados en formato CSV, se cargaron en Gephi para su análisis visual. Desde la herramienta se importaron los archivos nodos.csv y arcos.csv, se verificó que los nodos y arcos estuvieran correctamente identificados y se definió la red como dirigida, ya que representa relaciones de inversión.

Seguidamente se aplicó el algoritmo Force Atlas como primer paso en la distribución de los nodos, ya que, según la literatura, es adecuado para redes pequeñas o medianas al generar una distribución clara y destacar comunidades de inversión y relaciones clave. Sin embargo, al notar que la red seguía densa, se utilizó Yifan Hu, un algoritmo que optimiza la disposición de los nodos, facilitando la identificación de patrones y mejorando la visualización cuando Force Atlas no logra una separación óptima. Finalmente, se aplicó la opción Expand para ajustar la escala del gráfico sin alterar la disposición de los nodos, mejorando la legibilidad y evitando la superposición de elementos.

Para definir el tamaño de los nodos, se utilizó la métrica de grado de salida, permitiendo identificar a los inversionistas más activos dentro de la red. En cuanto a las etiquetas, se decidió mostrar el nombre de las empresas e inversionistas para facilitar la interpretación de las conexiones y destacar a los actores clave en la red.

1. **Análisis Visual de la red**

Se realiza el análisis visual de la red y se identifican sus actores relevantes. Para ello, se presentan imágenes de la red debidamente descritas y se identifican visual y analíticamente los actores relevantes de la red. En el análisis se argumenta la elección de diferentes medidas de centralidad y transitividad y cómo estas permiten identificar a los actores relevantes en la red.

**INCLUIR EN MEDIDAS CENTRALIDAD:**

yose-GRADO DE ENTRADA

yose -GRADO DE SALIDA ( TAMAÑO DEL NODO PORQUE NOS PERMITIRA VISUALIZAR UN INVERSIONISTA LIDER.

yose -BETWENESS: INVERSIONISTA O EMPRESA TIENEN INTERMEDIACION CLAVES EN LA CONECTIVIDAD DE LA RED DE INVERSIONES.

pipe-CLOSENESS CENTRALITY : INVERSIONISTA O EMPRESA ES INFLUYENTE EN TODA LA RED.

**INCLUIR EN MEDIDAS TRANSITIVIDAD:**

pipe -MODULARIDAD: CLUSTER POR COLORES

pipe -DENSIDAD DE LA RED:CUAN CONECTADA ESTA EL SISTEMA.

nico-DIAMETRO DE LA RED: UTIL PARA ENTENDER LA DISPERSION DE LAS INVERSIONES

nico -CLUSTERING COEFFICIENTE: MIDEN SI LAS EMPRESAS FORMAN GRUPOS CONECTADOS ENTRE SI.

1. **Conclusiones**

nico

Se realiza el análisis visual de la red y se identifican sus actores relevantes. Para ello, se presentan imágenes de la red debidamente descritas y se identifican visual y analíticamente los actores relevantes de la red. En el análisis se argumenta la elección de diferentes medidas de centralidad y transitividad y cómo estas permiten identificar a los actores relevantes en la red.