```
pip install pyvis
```

```
Collecting pyvis
    Downloading pyvis-0.3.2-py3-none-any.whl.metadata (1.7 kB)
    Requirement already satisfied: ipython>=5.3.0 in /usr/local/lib/python3.10/dist-packages (from pyvis) (7.34.0)
    Requirement already satisfied: jinja2>=2.9.6 in /usr/local/lib/python3.10/dist-packages (from pyvis) (3.1.4)
    Requirement already satisfied: jsonpickle>=1.4.1 in /usr/local/lib/python3.10/dist-packages (from pyvis) (3.3.0)
    Requirement already satisfied: networkx>=1.11 in /usr/local/lib/python3.10/dist-packages (from pyvis) (3.4.2)
    Requirement already satisfied: setuptools>=18.5 in /usr/local/lib/python3.10/dist-packages (from ipython>=5.3.0->pyvis) (75.1.0)
    Collecting jedi>=0.16 (from ipython>=5.3.0->pyvis)
    Downloading jedi-0.19.1-py2.py3-none-any.whl.metadata (22 kB)
    Requirement already satisfied: decorator in /usr/local/lib/python3.10/dist-packages (from ipython>=5.3.0->pyvis) (4.4.2)
    Requirement already satisfied: pickleshare in /usr/local/lib/python3.10/dist-packages (from ipython>=5.3.0->pyvis) (0.7.5)
    Requirement already satisfied: traitlets>=4.2 in /usr/local/lib/python3.10/dist-packages (from ipython>=5.3.0->pyvis) (5.7.1)
    Requirement already satisfied: prompt-toolkit!=3.0.0,!=3.0.1,<3.1.0,>=2.0.0 in /usr/local/lib/python3.10/dist-packages (from ipython>=5.
    Requirement already satisfied: pygments in /usr/local/lib/python3.10/dist-packages (from ipython>=5.3.0->pyvis) (2.18.0)
    Requirement already satisfied: backcall in /usr/local/lib/python3.10/dist-packages (from ipython>=5.3.0->pyvis) (0.2.0)
    Requirement already satisfied: matplotlib-inline in /usr/local/lib/python3.10/dist-packages (from ipython>=5.3.0->pyvis) (0.1.7)
    Requirement already satisfied: pexpect>4.3 in /usr/local/lib/python3.10/dist-packages (from ipython>=5.3.0->pyvis) (4.9.0)
    Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.10/dist-packages (from jinja2>=2.9.6->pyvis) (3.0.2)
    Requirement already satisfied: parso<0.9.0,>=0.8.3 in /usr/local/lib/python3.10/dist-packages (from jedi>=0.16->ipython>=5.3.0->pyvis) (
    Requirement already satisfied: ptyprocess>=0.5 in /usr/local/lib/python3.10/dist-packages (from pexpect>4.3->ipython>=5.3.0->pyvis) (0.7
    Requirement already satisfied: wcwidth in /usr/local/lib/python3.10/dist-packages (from prompt-toolkit!=3.0.0,!=3.0.1,<3.1.0,>=2.0.0->ip
    Downloading pyvis-0.3.2-py3-none-any.whl (756 kB)
    ──────────────────────────────────────── 756.0/756.0 kB 19.3 MB/s eta 0:00:00
    Downloading jedi-0.19.1-py2.py3-none-any.whl (1.6 MB)
    ──────────────────────────────────────── 1.6/1.6 MB 38.0 MB/s eta 0:00:00
    Installing collected packages: jedi, pyvis
    Successfully installed jedi-0.19.1 pyvis-0.3.2
```

```python
import csv
import os
import sys
import json
import pandas as pd
import networkx as nx
from operator import itemgetter
from networkx.algorithms import community
from pyvis.network import Network
from collections import Counter
import numpy as np
import matplotlib.pyplot as plt
```

Start coding or generate with AI.

## Read the Data

```python
from google.colab import drive
drive.mount('/content/drive')
```

```
Mounted at /content/drive
```

```
!ls '/content/drive/My Drive/Abo Akademi/Data science/Data-Miniproject 3'
```

```
artist_institution_network.graphml    interactive_network.html              network_visualization.png
artists.csv                           lib                                   relationships.csv
fig1.jpg                              network.graphml                       schools.csv
fig1.pdf                             network_visualization_high_res.png
institutions.csv                      network_visualization_new.png
```

```python
import os
os.chdir("/content/drive/My Drive/Abo Akademi/Data science/Data-Miniproject 3")
```

```python
artists = pd.read_csv('artists.csv')
artists = artists.dropna(subset=['artistUrl','id','image','nation','title','totalWorksTitle','year'])
artists
```

| | artistUrl | id | image | nation | title | totalWorksTitle | year |
|---|---|---|---|---|---|---|---|
| 0 | /en/ancient-egyptian-painting | 5d230b31edc2c9fb74ae130a | https://uploads2.wikiart.org/00244/images/anci... | Egyptians | Ancient Egypt | 169 artworks | c.3150 BC - c.640 AD |
| 1 | /en/ancient-greek-pottery | 5ced09b1edc2c993247c1ea0 | https://uploads7.wikiart.org/00238/images/anci... | Greeks | Ancient Greek Pottery | 87 artworks | c.900 - 31 BC |
| 2 | /en/ancient-greek-painting | 5ce3c54dedc2c921c4035b7d | https://uploads8.wikiart.org/00237/images/anci... | Greeks | Ancient Greek Painting and Sculpture | 76 artworks | c.700 - c.31 BC |
| 3 | /en/apelles | 5ba232ccedc2c918cc654ea8 | https://uploads2.wikiart.org/00202/images/apel... | Greek | Apelles | 8 artworks | born c.370 BC |
| 4 | /en/fayum-portrait | 5beea1ededc2c915a01acdda | https://uploads7.wikiart.org/00210/images//01.... | Egyptians | Fayum portrait | 76 artworks | c.50 - c.250 |

```
relationships = pd.read_csv('relationships.csv')
relationships = relationships.dropna(subset=['artistUrl', 'friends', 'influenced_by','influenced_on','institution','movements','school','typ
relationships.head()
```

| | artistUrl | friends | influenced_by | influenced_on | institution | movements |
|---|---|---|---|---|---|---|
| 75 | /en/rogier-van-der-weyden | /en/jacques-daret | /en/jan-van-eyck,/en/stefan-lochner | /en/antonello-da-messina,/en/hans-memling,/en/... | /en/artists-by-art-institution/guild-of-saint-... | Northern Renaissance |
| 87 | /en/pieter-bruegel-the-elder | /en/maarten-de-vos,/en/giulio-clovio | /en/hieronymus-bosch | /en/tobias-verhaecht,/en/peter-paul-rubens,/en... | /en/artists-by-art-institution/guild-of-saint-... | Northern Renaissance |
| 131 | /en/jean-francois-millet | /en/paul-delaroche,/en/constant-troyon,/en/nar... | /en/honore-daumier,/en/michelangelo,/en/nicola... | /en/edward-mitchell-bannister,/en/andres-de-sa... | /en/artists-by-art-institution/cole-des-beaux-... | Realism |
| 169 | /en/ford-madox-brown | /en/william-morris,/en/edward-burne-jones | /en/william-holman-hunt,/en/john-everett-milla... | /en/dante-gabriel-rossetti | /en/artists-by-art-institution/royal-academy-o... | Romanticism |

```
institutions = pd.read_csv('institutions.csv')

institutions
```

| | city | country | title | url |
|---|---|---|---|---|
| 0 | NaN | NaN | Guild of Saint Luke | /en/artists-by-art-institution/guild-of-saint-... |
| 1 | Antwerp | Belgium | Royal Academy of Fine Arts | /en/artists-by-art-institution/royal-academy-o... |
| 2 | Munich | Germany | Akademie der Bildenden Künste München (Munich ... | /en/artists-by-art-institution/akademie-der |
| 3 | Milan | Italy | Accademia di Belle Arti di Brera (Brera Academy) | /en/artists-by-art-institution/accademia-di-be... |
| 4 | Paris | France | Académie des Beaux-Arts | /en/artists-by-art-institution/acad-mie-des-beaux |
| ... | ... | ... | ... | ... |
| 68 | East Sussex | UK | Hastings College of Arts and Technology (HCAT) | /en/artists-by-art-institution/hastings-colleg... |
| 69 | Odessa | Ukraine | Grekov Odessa Art School | /en/artists-by-art-institution/grekov-odessa-a... |
| 70 | Cambridge | MA | American Academy of Arts and Sciences | /en/artists-by-art-institution/american-academ... |
| 71 | London | UK | Central Saint Martins (CSM) | /en/artists-by-art-institution/central-saint-m... |
| 72 | Chicago | IL | School of the Art Institute of Chicago (SAIC) | /en/artists-by-art-institution/school-of-the-art |

73 rows × 4 columns

```
schools = pd.read_csv('schools.csv')
schools
```

| | title | url |
|---|---|---|
| 0 | Flemish School | /en/artists-by-painting-school/flemish-school |
| 1 | Sienese School | /en/artists-by-painting-school/sienese-school |
| 2 | Umbrian school | /en/artists-by-painting-school/umbrian-school |
| 3 | Florentine School | /en/artists-by-painting-school/florentine-school |
| 4 | Parma school | /en/artists-by-painting-school/parma-school |
| ... | ... | ... |
| 215 | Cádiz School | /en/artists-by-painting-school/c-diz-school |
| 216 | Wu School | /en/artists-by-painting-school/wu-school |
| 217 | Bruges School | /en/artists-by-painting-school/bruges-school |
| 218 | Milanese school | /en/artists-by-painting-school/milanese-school |
| 219 | Danube school | /en/artists-by-painting-school/danube-school |

220 rows × 2 columns

## ∨ EDA

```
print("Artists Dataset Overview:")
print(artists.info())
print("Missing Values in Artists Dataset:", artists.isnull().sum())

print("\nRelationships Dataset Overview:")
print(relationships.info())
print("Missing Values in Relationships Dataset:", relationships.isnull().sum())

print("\nInstitutions Dataset Overview:")
print(institutions.info())
print("Missing Values in Institutions Dataset:", institutions.isnull().sum())

print("\nSchools Dataset Overview:")
print(schools.info())
print("Missing Values in Schools Dataset:", schools.isnull().sum())

# Distribution analysis
print("\nNation Distribution among Artists:")
print(artists['nation'].value_counts().head(10))

print("\nActive Year Intervals among Artists:")
print(artists['year'].value_counts().head(10))

print("\nNumber of Artworks Distribution:")
print(artists['totalWorksTitle'].describe())
```

```
Artists Dataset Overview:
<class 'pandas.core.frame.DataFrame'>
Index: 2963 entries, 0 to 2995
Data columns (total 7 columns):
 #   Column           Non-Null Count  Dtype
---  ------           --------------  -----
 0   artistUrl        2963 non-null   object
 1   id               2963 non-null   object
 2   image            2963 non-null   object
 3   nation           2963 non-null   object
 4   title            2963 non-null   object
 5   totalWorksTitle  2963 non-null   object
 6   year             2963 non-null   object
dtypes: object(7)
memory usage: 185.2+ KB
None
Missing Values in Artists Dataset: artistUrl        0
id                 0
image              0
nation             0
title              0
totalWorksTitle    0
year               0
```

```
        dtype: int64

        Relationships Dataset Overview:
        <class 'pandas.core.frame.DataFrame'>
        Index: 58 entries, 75 to 2168
        Data columns (total 8 columns):
         #   Column         Non-Null Count  Dtype
        ---  ------         --------------  -----
         0   artistUrl      58 non-null     object
         1   friends        58 non-null     object
         2   influenced_by  58 non-null     object
         3   influenced_on  58 non-null     object
         4   institution    58 non-null     object
         5   movements      58 non-null     object
         6   school         58 non-null     object
         7   type           58 non-null     object
        dtypes: object(8)
        memory usage: 4.1+ KB
        None
        Missing Values in Relationships Dataset: artistUrl        0
        friends          0
        influenced_by    0
        influenced_on    0
        institution      0
        movements        0
        school           0
        type             0
        dtype: int64

        Institutions Dataset Overview:
        <class 'pandas.core.frame.DataFrame'>
        RangeIndex: 73 entries, 0 to 72
        Data columns (total 4 columns):
         #   Column   Non-Null Count  Dtype
        ---  ------   --------------  -----
```

```
print("Artists Dataset Summary Statistics:")
display(artists.describe())
print('\n\n')
print("Relationships Dataset Summary Statistics:")
display(relationships.describe())
print('\n\n')
print("Institutions Dataset Summary Statistics:")
display(institutions.describe())
print('\n\n')
print("Schools Dataset Summary Statistics:")
display(schools.describe())
```

Artists Dataset Summary Statistics:

| | artistUrl | id | image | nation | title | totalWorksTitle | year |
|---|---|---|---|---|---|---|---|
| **count** | 2963 | 2963 | 2963 | 2963 | 2963 | 2963 | 2963 |
| **unique** | 2963 | 2963 | 2963 | 106 | 2963 | 294 | 2181 |
| **top** | /en/ferdinand-georg-waldm-ller | 57726d96edc2cb3880b4b06a | https://uploads6.wikiart.org/00381/images/ferd... | American | Ferdinand Georg Waldmüller | 15 artworks | born 1937 |
| **freq** | 1 | 1 | 1 | 519 | 1 | 137 | 24 |

Relationships Dataset Summary Statistics:

| | artistUrl | friends | influenced_by | influenced_on | institution | movements | school | type |
|---|---|---|---|---|---|---|---|---|
| **count** | 58 | 58 | 58 | 58 | 58 | 58 | 58 | 58 |
| **unique** | 58 | 53 | 58 | 57 | 45 | 39 | 46 | 1 |
| **top** | /en/rogier-van-der-weyden | /en/jackson-pollock | /en/jan-van-eyck,/en/stefan-lochner | /en/charles-gibbons | /en/artists-by-art-institution/art-students-le... | Expressionism | /en/artists-by-painting-school/pre-raphaelite-... | Artists |
| **freq** | 1 | 6 | 1 | 2 | 4 | 6 | 3 | 58 |

Institutions Dataset Summary Statistics:

| | city | country | title | url |
|---|---|---|---|---|
| **count** | 71 | 71 | 73 | 73 |
| **unique** | 52 | 35 | 73 | 73 |
| **top** | Paris | UK | School of the Art Institute of Chicago (SAIC) | /en/artists-by-art-institution/school-of-the-art |
| **freq** | 7 | 8 | 1 | 1 |

Schools Dataset Summary Statistics:

| | title | url |
|---|---|---|
| **count** | 220 | 220 |
| **unique** | 220 | 220 |

```python
# Making the artist name based on artistUrl in relationships file from artists file
relationships_art = relationships.merge(
    artists[['artistUrl', 'title']],  # Select the 'artistUrl' and 'title' columns from artists
    on='artistUrl',                   # Join based on 'artistUrl'
    how='left'                        # Use 'left' join to keep all relationships data
).rename(columns={'title': 'artist_name'})
```

```python
relationships_art.head()
```

| | artistUrl | friends | influenced_by | influenced_on | institution | movements | |
|---|---|---|---|---|---|---|---|
| 0 | /en/rogier-van-der-weyden | /en/jacques-daret | /en/jan-van-eyck,/en/stefan-lochner | /en/antonello-da-messina,/en/hans-memling,/en/... | /en/artists-by-art-institution/guild-of-saint-... | Northern Renaissance | /sc |
| 1 | /en/pieter-bruegel-the-elder | /en/maarten-de-vos,/en/giulio-clovio | /en/hieronymus-bosch | /en/tobias-verhaecht,/en/peter-paul-rubens,/en... | /en/artists-by-art-institution/guild-of-saint-... | Northern Renaissance | /sc |
| 2 | /en/jean-francois-millet | /en/paul-delaroche,/en/constant-troyon,/en/nar... | /en/honore-daumier,/en/michelangelo,/en/nicola... | /en/edward-mitchell-bannister,/en/andres-de-sa... | /en/artists-by-art-institution/cole-des-beaux-... | Realism | /sch |
| 3 | /en/ford-madox-brown | /en/william-morris,/en/edward-burne-jones | /en/william-holman-hunt,/en/john-everett-milla... | /en/dante-gabriel-rossetti | /en/artists-by-art-institution/royal-academy-o... | Romanticism | / |
| 4 | /en/william-... | /en/dante-gabriel-...ti/...f... | /en/john-...k,/...tt,/...t... | /en/ford-madox-... | /en/artists-by-art-... | Romanticism,Symbolism,Orientalism | / |

```
#split the friends column for seperating the multiple name based on commas
split_fr_df = relationships_art['friends'].str.split(',', expand=True)
split_fr_df.head()
```

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | /en/jacques-daret | None | None | None | None | None | None | None | None | None | None | None | None | None | None | None |
| 1 | /en/maarten-de-vos | /en/giulio-clovio | None | None | None | None | None | None | None | None | None | None | None | None | None | None |
| 2 | /en/paul-delaroche | /en/constant-troyon | /en/narcisse-virgilio-diaz | /en/charles-jacque | /en/theodore-rousseau | /en/camille-corot | None | None | None | None | None | None | None | None | None | None |
| 3 | /en/william-morris | /en/edward-burne-jones | None | None | None | None | None | None | None | None | None | None | None | None | None | None |

```
# Rename columns with a naming pattern or specific names
split_fr_df.columns = [f'Fr_{i+1}' for i in range(split_fr_df.shape[1])]
split_fr_df.head()
```

| | Fr_1 | Fr_2 | Fr_3 | Fr_4 | Fr_5 | Fr_6 | Fr_7 | Fr_8 | Fr_9 | Fr_10 | Fr_11 | Fr_12 | Fr_13 | Fr_14 | Fr_15 | F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | /en/jacques-daret | None | None | None | None | None | None | None | None | None | None | None | None | None | None | |
| 1 | /en/maarten-de-vos | /en/giulio-clovio | None | None | None | None | None | None | None | None | None | None | None | None | None | |
| 2 | /en/paul-delaroche | /en/constant-troyon | /en/narcisse-virgilio-diaz | /en/charles-jacque | /en/theodore-rousseau | /en/camille-corot | None | None | None | None | None | None | None | None | None | |
| 3 | /en/william-morris | /en/edward-burne-jones | None | None | None | None | None | None | None | None | None | None | None | None | None | |

```
art_fr = artists.rename(columns={'artistUrl':'Fr_1'})
art_fr.head()
```

| | Fr_1 | id | image | nation | title | totalWorksTitle | year |
|---|---|---|---|---|---|---|---|
| 0 | /en/ancient-egyptian-painting | 5d230b31edc2c9fb74ae130a | https://uploads2.wikiart.org/00244/images/anci... | Egyptians | Ancient Egypt | 169 artworks | c.3150 BC - c.640 AD |
| 1 | /en/ancient-greek-pottery | 5ced09b1edc2c993247c1ea0 | https://uploads7.wikiart.org/00238/images/anci... | Greeks | Ancient Greek Pottery | 87 artworks | c.900 - 31 BC |
| 2 | /en/ancient-... | 5cc2e54dedc2c921e4035b7d | https://uploads8.wikiart.org/00237/images/anci... | Greeks | Ancient Greek Painting and... | 76 artworks | c.700 - |

```
# Making the friend name based on Url
relationships_fr = split_fr_df.merge(
```

```python
    art_fr[['Fr_1', 'title']],  # Select the 'artistUrl' and 'title' columns from artists
    on='Fr_1',                  # Join based on 'artistUrl'
    how='left'                          # Use 'left' join to keep all relationships data
).rename(columns={'title': 'Fr_1_name'})
#for Fr2 column
art_fr = artists.rename(columns={'artistUrl':'Fr_2'})
relationships_fr = relationships_fr.merge(
    art_fr[['Fr_2', 'title']],  # Select the 'artistUrl' and 'title' columns from artists
    on='Fr_2',                  # Join based on 'artistUrl'
    how='left'                          # Use 'left' join to keep all relationships data
).rename(columns={'title': 'Fr_2_name'})
# for fr3 column
art_fr = artists.rename(columns={'artistUrl':'Fr_3'})
relationships_fr = relationships_fr.merge(
    art_fr[['Fr_3', 'title']],  # Select the 'artistUrl' and 'title' columns from artists
    on='Fr_3',                  # Join based on 'artistUrl'
    how='left'                          # Use 'left' join to keep all relationships data
).rename(columns={'title': 'Fr_3_name'})
# for fr4 column
art_fr = artists.rename(columns={'artistUrl':'Fr_4'})
relationships_fr = relationships_fr.merge(
    art_fr[['Fr_4', 'title']],  # Select the 'artistUrl' and 'title' columns from artists
    on='Fr_4',                  # Join based on 'artistUrl'
    how='left'                          # Use 'left' join to keep all relationships data
).rename(columns={'title': 'Fr_4_name'})
relationships_fr
```

| | Fr_1 | Fr_2 | Fr_3 | Fr_4 | Fr_5 | Fr_6 | Fr_7 | Fr_8 | Fr_9 | Fr_10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | /en/jacques-daret | None | None | None | None | None | None | None | None | None |
| 1 | /en/maarten-de-vos | /en/giulio-clovio | None | None | None | None | None | None | None | None |
| 2 | /en/paul-delaroche | /en/constant-troyon | /en/narcisse-virgilio-diaz | /en/charles-jacque | /en/theodore-rousseau | /en/camille-corot | None | None | None | None |
| 3 | /en/william-morris | /en/edward-burne-jones | None | None | None | None | None | None | None | None |
| 4 | /en/dante-gabriel-rossetti | /en/john-everett-millais | /en/edward-robert-hughes | /en/herbert-gustave-schmalz | None | None | None | None | None | None |
| 5 | /en/william-holman-hunt | None | None | None | None | None | None | None | None | None |
| 6 | /en/henri-fantin-latour | /en/edouard-manet | None | None | None | None | None | None | None | None |
| 7 | /en/ilya-repin | /en/ivan-kramskoy | /en/grigoriy-myasoyedov | /en/vladimir-makovsky | /en/mikhail-vrubel | /en/taras-shevchenko | /en/nikolai-ge | /en/petro-kholodny-elder | None | None |
| 8 | /en/vasily-polenov | /en/alexandre-benois | /en/mykola-murashko | None | None | None | None | None | None | None |
| 9 | /en/paul-cezanne | /en/vincent-van-gogh | /en/stanislaw-wyspianski | /en/emile-bernard | /en/meijer-de-haan | /en/charles-laval | /en/roderic-oconor | /en/ferdinand-du-puigaudeau | None | None |
| 10 | /en/julian-alden-weir | /en/henry-tonks | /en/martin-rico | /en/frank-omeara | /en/william-logsdail | None | None | None | None | None |
| 11 | /en/mykola-murashko | /en/valentin-serov | /en/konstantin-korovin | None | None | None | None | None | None | None |
| 12 | /en/edward-e-simmons | None | None | None | None | None | None | None | None | None |
| 13 | /en/paul-klee | /en/franz-marc | /en/alexej-von-jawlensky | /en/vadym-meller | None | None | None | None | None | None |
| 14 | /en/xavier-martinez | /en/emilie-charmy | /en/henri-manguin | None | None | None | None | None | None | None |
| 15 | /en/albert-julius-olsson | None | None | None | None | None | None | None | None | None |
| 16 | /en/marc-chagall | /en/fernand-leger | /en/jan-toorop | /en/theo-van-doesburg | /en/roman-selsky | /en/margit-selska | None | None | None | None |
| 17 | /en/olga-rozanova | /en/lyubov-popova | /en/aleksandra-ekster | /en/marc-chagall | /en/laszlo-moholy-nagy | /en/vadym-meller | /en/pavel-filonov | None | None | None |
| 18 | /en/josef-albers | /en/wassily-kandinsky | /en/august-macke | /en/vadym-meller | None | None | None | None | None | None |
| 19 | /en/georges-braque | /en/constantin-brancusi | /en/man-ray | /en/max-jacob | /en/amedeo-modigliani | /en/christopher-wood | /en/mykhailo-boychuk | /en/c-r-w-nevinson | /en/valentine-hugo | /en/lee-miller |
| 20 | /en/vsevolod-maksymovych | None | None | None | None | None | None | None | None | None |
| 21 | /en/george-bellows | /en/rockwell-kent | /en/charles-e-burchfield | None | None | None | None | None | None | None |
| 22 | /en/marc-chagall | /en/piet-mondrian | /en/le-corbusier | /en/robert-delaunay | /en/constantin-brancusi | /en/roman-selsky | /en/alice-bailly | None | None | None |
| 23 | /en/edvard-munch | /en/august-macke | /en/wassily-kandinsky | /en/erich-heckel | None | None | None | None | None | None |
| 24 | /en/oleksandr-murashko | /en/pablo-picasso | /en/diego-rivera | /en/victor-palmov | /en/vasyl-yermylov | /en/oleksandr-bogomazov | /en/alexander-khvostenko-khvostov | None | None | None |
| 25 | /en/robert-delaunay | /en/alice-bailly | None | None | None | None | None | None | None | None |
| 26 | /en/diego-rivera | /en/david-alfaro- | None | None | None | None | None | None | None | None |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | rivera | siqueiros | | | | | | | | |
| 27 | /en/constantin-brancusi | /en/diego-rivera | /en/chaim-soutine | /en/moise-kisling | /en/jacques-lipchitz | /en/max-jacob | /en/pablo-picasso | /en/jacob-epstein | /en/juan-gris | /en/gino-severini |
| 28 | /en/egon-schiele | None | None | None | None | None | None | None | None | None |
| 29 | /en/franz-marc | /en/paul-klee | None | None | None | None | None | None | None | None |
| 30 | /en/amedeo-modigliani | /en/saturnino-herran | /en/robert-montenegro | /en/jose-clemente-orozco | /en/david-alfaro-siqueiros | /en/mykhailo-boychuk | /en/alice-rahon | /en/jacqueline-lamba | None | None |
| 31 | /en/robert-delaunay | /en/fernand-leger | /en/piet-mondrian | /en/kazimir-malevich | None | None | None | None | None | None |
| 32 | /en/alfred-stieglitz | /en/arthur-dove | /en/charles-demuth | /en/beauford-delaney | None | None | None | None | None | None |
| 33 | /en/amedeo-modigliani | /en/maruja-mallo | None | None | None | None | None | None | None | None |
| 34 | /en/paul-klee | None | None | None | None | None | None | None | None | None |
| 35 | /en/gustav-klimt | /en/oskar-kokoschka | None | None | None | None | None | None | None | None |
| 36 | /en/constantin-brancusi | /en/salvador-dali | /en/marcel-duchamp | /en/francis-picabia | /en/pablo-picasso | /en/alfred-stieglitz | /en/maurice-tabard | /en/lee-miller | None | None |
| 37 | /en/amedeo-modigliani | /en/juan-gris | None | None | None | None | None | None | None | None |
| 38 | /en/amedeo-modigliani | /en/michel-kikoine | /en/pinchus-kremegne | None | None | None | None | None | None | None |
| 39 | /en/frederic-edwin-church | /en/john-trumbull | /en/asher-brown-durand | None | None | None | None | None | None | None |
| 40 | /en/jackson-pollock | None | None | None | None | None | None | None | None | None |
| 41 | /en/jackson-pollock | None | None | None | None | None | None | None | None | None |
| 42 | /en/jackson-pollock | None | None | None | None | None | None | None | None | None |
| 43 | /en/jackson-pollock | None | None | None | None | None | None | None | None | None |
| 44 | /en/lee-krasner | /en/robert-motherwell | /en/mark-rothko | /en/willem-de-kooning | /en/barnett-newman | /en/philip-guston | /en/ad-reinhardt | None | None | None |
| 45 | /en/kenneth-noland | /en/jack-tworkov | /en/arshile-gorky | /en/david-alfaro-siqueiros | None | None | None | None | None | None |
| 46 | /en/jackson-pollock | /en/rosemarie-beck | None | None | None | None | None | None | None | None |
| 47 | /en/jackson-pollock | None | None | None | None | None | None | None | None | None |
| 48 | /en/jackson-pollock | None | None | None | None | None | None | None | None | None |
| 49 | /en/boris-kustodiev | /en/filipp-malyavin | /en/oleksandr-bogomazov | /en/heorhiy-narbut | /en/mykhailo-boychuk | None | None | None | None | None |
| 50 | /en/francis-bacon | None | None | None | None | None | None | None | None | None |
| 51 | /en/morris-louis | None | None | None | None | None | None | None | None | None |
| 52 | /en/morris-louis | /en/jose-clemente-orozco | /en/diego-rivera | /en/antonio-berni | None | None | None | None | None | None |
| 53 | /en/jasper-johns | /en/robert-rauschenberg | None | None | None | None | None | None | None | None |
| 54 | /en/larry-rivers | /en/philip-pearlstein | /en/jean-michel-basquiat | /en/francesco-clemente | /en/salvador-dali | /en/steve-kaufman | /en/jack-armstrong | None | None | None |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 55 | /en/cy-twombly | None | None | None | None | None | None | None | None | None |
| 56 | /en/blinky-palermo | /en/sigmar-polke | /en/joseph-beuys | /en/georg-baselitz | /en/manfred-kuttner | /en/konrad-lueg | None | None | None | None |
| 57 | /en/gerhard-richter | /en/manfred-kuttner | /en/konrad-lueg | None | None | None | None | None | None | None |

58 rows × 21 columns

```python
relationships_art[['Fr1','Fr2','Fr3','Fr4']]=relationships_fr[['Fr_1_name','Fr_2_name','Fr_3_name','Fr_4_name']]

relationships_art.head()
```

| | artistUrl | friends | influenced_by | influenced_on | institution | movements | |
|---|---|---|---|---|---|---|---|
| 0 | /en/rogier-van-der-weyden | /en/jacques-daret | /en/jan-van-eyck,/en/stefan-lochner | /en/antonello-da-messina,/en/hans-memling,/en/... | /en/artists-by-art-institution/guild-of-saint-... | Northern Renaissance | / sc |
| 1 | /en/pieter-bruegel-the-elder | /en/maarten-de-vos,/en/giulio-clovio | /en/hieronymus-bosch | /en/tobias-verhaecht,/en/peter-paul-rubens,/en... | /en/artists-by-art-institution/guild-of-saint-... | Northern Renaissance | / sc |
| 2 | /en/jean-francois-millet | /en/paul-delaroche,/en/constant-troyon,/en/nar... | /en/honore-daumier,/en/michelangelo,/en/nicola... | /en/edward-mitchell-bannister,/en/andres-de-sa... | /en/artists-by-art-institution/cole-des-beaux-... | Realism | / sch |
| 3 | /en/ford-madox-brown | /en/william-morris,/en/edward-burne-jones | /en/william-holman-hunt,/en/john-everett-milla... | /en/dante-gabriel-rossetti | /en/artists-by-art-institution/royal-academy-o... | Romanticism | / |
| 4 | /en/william-holman-hunt | /en/dante-gabriel-rossetti,/en/john-everett-mi... | /en/john-ruskin,/en/giotto,/en/ambrogio-lorenz... | /en/ford-madox-brown | /en/artists-by-art-institution/royal-academy-of | Romanticism,Symbolism,Orientalism | / |

Start coding or generate with AI.

```python
# art = artists_red.rename(columns={'artistUrl':'friends'})
# art
```

```python
# relationships_fr = relationships_art.merge(
#     art[['friends', 'title']],  # Select the 'artistUrl' and 'title' columns from artists
#     on='friends',                # Join based on 'artistUrl'
#     how='left'                   # Use 'left' join to keep all relationships data
# ).rename(columns={'title': 'fr_name'})
```
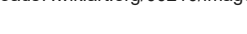
```python
# relationships_fr
```

```python
relationships_art.head()
```

| | artistUrl | friends | influenced_by | influenced_on | institution | movements | |
|---|---|---|---|---|---|---|---|
| 0 | /en/rogier-van-der-weyden | /en/jacques-daret | /en/jan-van-eyck,/en/stefan-lochner | /en/antonello-da-messina,/en/hans-memling,/en/... | /en/artists-by-art-institution/guild-of-saint-... | Northern Renaissance | / sc |
| 1 | /en/pieter-bruegel-the-elder | /en/maarten-de-vos,/en/giulio-clovio | /en/hieronymus-bosch | /en/tobias-verhaecht,/en/peter-paul-rubens,/en... | /en/artists-by-art-institution/guild-of-saint-... | Northern Renaissance | / sc |
| 2 | /en/jean-francois-millet | /en/paul-delaroche,/en/constant-troyon,/en/nar... | /en/honore-daumier,/en/michelangelo,/en/nicola... | /en/edward-mitchell-bannister,/en/andres-de-sa... | /en/artists-by-art-institution/cole-des-beaux-... | Realism | / sch |
| 3 | /en/ford-madox-brown | /en/william-morris,/en/edward-burne-jones | /en/william-holman-hunt,/en/john-everett-milla... | /en/dante-gabriel-rossetti | /en/artists-by-art-institution/royal-academy-o... | Romanticism | / |
| 4 | /en/william-holman-hunt | /en/dante-gabriel-rossetti,/en/john-everett-mi... | /en/john-ruskin,/en/giotto,/en/ambrogio-lorenz... | /en/ford-madox-brown | /en/artists-by-art-institution/royal-academy-of | Romanticism,Symbolism,Orientalism | / |

```
relationships_art.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 58 entries, 0 to 57
Data columns (total 13 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   artistUrl     58 non-null     object
 1   friends       58 non-null     object
 2   influenced_by 58 non-null     object
 3   influenced_on 58 non-null     object
 4   institution   58 non-null     object
 5   movements     58 non-null     object
 6   school        58 non-null     object
 7   type          58 non-null     object
 8   artist_name   58 non-null     object
 9   Fr1           55 non-null     object
 10  Fr2           38 non-null     object
 11  Fr3           31 non-null     object
 12  Fr4           24 non-null     object
dtypes: object(13)
memory usage: 6.0+ KB
```

```
# reduce the artists's node for better showing the graph
artists_red = artists.drop(artists.index[80:])
artists_red
```

| | artistUrl | id | image | nation | title | totalWorksTitle | year |
|---|---|---|---|---|---|---|---|
| 0 | /en/ancient-egyptian-painting | 5d230b31edc2c9fb74ae130a | https://uploads2.wikiart.org/00244/images/anci... | Egyptians | Ancient Egypt | 169 artworks | c.3150 BC - c.640 AD |
| 1 | /en/ancient-greek-pottery | 5ced09b1edc2c993247c1ea0 | https://uploads7.wikiart.org/00238/images/anci... | Greeks | Ancient Greek Pottery | 87 artworks | c.900 - 31 BC |
| 2 | /en/ancient-greek-painting | 5ce3c54dedc2c921c4035b7d | https://uploads8.wikiart.org/00237/images/anci... | Greeks | Ancient Greek Painting and Sculpture | 76 artworks | c.700 - c.31 BC |
| 3 | /en/apelles | 5ba232ccedc2c918cc654ea8 | https://uploads2.wikiart.org/00202/images/apel... | Greek | Apelles | 8 artworks | born c.370 BC |
| 4 | /en/fayum-portrait | 5beea1ededc2c915a01acdda | https://uploads7.wikiart.org/00210/images//01.... | Egyptians | Fayum portrait | 76 artworks | c.50 - c.250 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| | /en/hans- | | | | | | |

```
# pip install nltk
```

```
# from nltk.corpus import stopwords
# import nltk

# # Download stopwords if you haven't yet
# nltk.download('stopwords')

# # Load the English stopwords list
# stop_words = set(stopwords.words('english'))

# # Remove stopwords from the specified column
# artists_red['title'] = artists_red['title'].apply(lambda x: ' '.join([word for word in x.split() if word.lower() not in stop_words]))

# # Strip whitespace and remove any potential unwanted characters
# artists_red['title'] = artists_red['title'].str.strip()

# artists_red
```

```
# # Check for duplicate titles
# print(artists_red['title'].duplicated().sum())
# # Add only unique titles as nodes
# unique_titles = artists_red['title'].unique()
```

```
# artists_red['title'] = artists_red['title'].astype(str)
```

## ˅ Network

```
# Initialize the graph
G = nx.Graph()
```

## ˅ Artists and their freinds network

In the relationships dataframe, friends column has urls. That's why I firstly trying to find the exact url's name from the artist column. In the freinds column, there are multiple values seperated by commas and I have to seperate the column and then doing naming. Also, there are more rows approximately 3000. I reduce this node number for better showing the network in colab small interface.

```
# Add each title as a single node, ensuring no accidental splits occur
for title in artists_red['title']:
    # print("Adding node:", title)  # Check output to confirm full title
    G.add_node(title, type='artist')

print("Total nodes after addition:", len(G.nodes))
```

⤓ Total nodes after addition: 80

```
print(G.nodes)
```

⤓ ['Ancient Egypt', 'Ancient Greek Pottery', 'Ancient Greek Painting and Sculpture', 'Apelles', 'Fayum portrait', 'Orthodox Icons', 'Wu Da

◄ ▬▬▬▬▬▬▬                                                                                                                              ▶

```
for _, row in relationships_art.iterrows():
    artist = row['artist_name']

    # Check if 'Fr1' column has valid (non-NaN) data
    if pd.notna(row['Fr1']):
        # Split the string into individual friends if comma-separated
        friends = row['Fr1'].split(',')
        for friend in friends:
            G.add_edge(artist, friend.strip(), relationship='friend')
    if pd.notna(row['Fr2']):
        # Split the string into individual friends if comma-separated
        friends = row['Fr2'].split(',')
        for friend in friends:
            G.add_edge(artist, friend.strip(), relationship='friend')
    if pd.notna(row['Fr3']):
        # Split the string into individual friends if comma-separated
        friends = row['Fr3'].split(',')
        for friend in friends:
            G.add_edge(artist, friend.strip(), relationship='friend')
    if pd.notna(row['Fr4']):
        # Split the string into individual friends if comma-separated
        friends = row['Fr4'].split(',')
        for friend in friends:
            G.add_edge(artist, friend.strip(), relationship='friend')
```

```
print(G.edges)
```

⤓ [('Rogier van der Weyden', 'Jacques Daret'), ('Giulio Clovio', 'Pieter Bruegel the Elder'), ('Pieter Bruegel the Elder', 'Maarten de Vos

◄ ▬▬                                                                                                                                   ▶

```
print("Total edges after addition:", len(G.edges))
```

⤓ Total edges after addition: 124

```
degree_seq = [d for n,d in G.degree()]
aver_deg = sum(degree_seq)/len(degree_seq)
```

```
print("Average_degree: ",aver_deg)
```

```
Average_degree:  1.1221719457013575
```

```
# art_dic = {}
# inst_dic = {}
# school_dic = {}
```

```
# for node in G.nodes:
#     art_dic[node[0]]=node[1]
#     inst_dic[node[0]]=node[2]
#     school_dic[node[0]]=node[3]
```

```
density = nx.density(G)
density
```

```
0.005100781571369807
```

```
# Visualization using networkx
plt.figure(figsize=(70, 70))
pos = nx.spring_layout(G, k=.2)
nx.draw(G, pos, node_size=180, with_labels=True, font_size=30)
plt.title("Basic Network Visualization")
# plt.savefig("network_visualization_high_res.png", format="PNG", dpi=300)
plt.show()
```
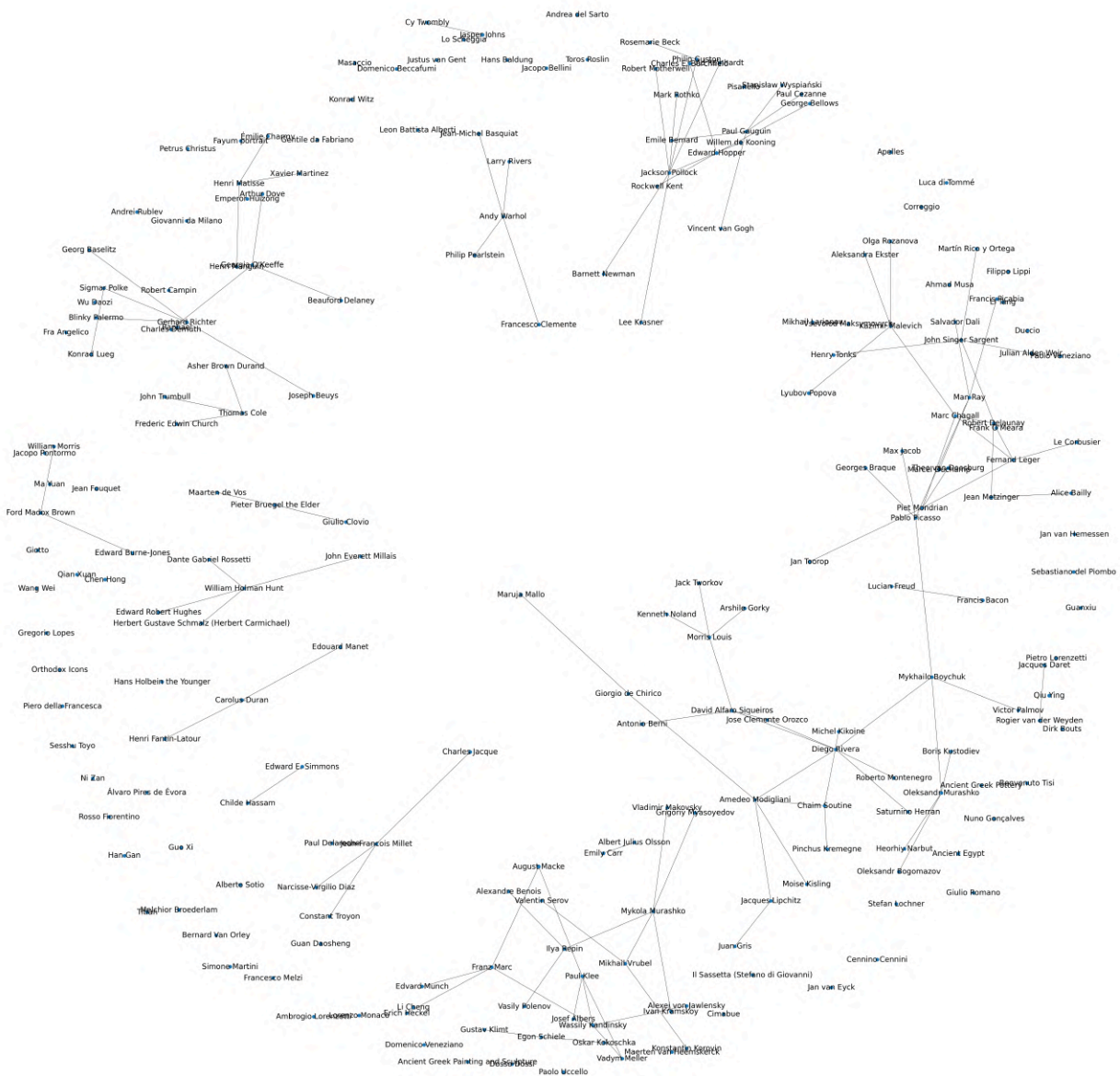
```
plt.savefig("network_visualization_high_res.png", format="PNG", dpi=300)
```

## ⌄ The goal is to answer the following questions:

1. Which were the most influential artists?
2. Which were the most influential movements?
3. Which were the most influential institutions?
4. Which nationalities concentrate the majority of artists?
5. Which are the biggest communities in the network?

```python
# Initialize the graph again
G1 = nx.Graph()
```

```python
# Add artist nodes
for _, row in artists.iterrows():
    G1.add_node(row['title'], type='artist', nation=row['nation'], active_years=row['year'], artworks=row['totalWorksTitle'])

# Add institution nodes
for _, row in institutions.iterrows():
    G1.add_node(row['url'], type='institution', country=row['country'], city=row['city'])

# Add school nodes
for _, row in schools.iterrows():
    G1.add_node(row['url'], type='school')

# Add nation nodes
for _, row in artists.iterrows():
    G1.add_node(row['nation'], type='nation')
#add movements nodes
for _, row in relationships.iterrows():
    G1.add_node(row['movements'], type='movement')
```

```python
print("Total nodes after addition:", len(G1.nodes))
```

⏏ Total nodes after addition: 3401

```python
print(G1.nodes)
```

⏏ ['Ancient Egypt', 'Ancient Greek Pottery', 'Ancient Greek Painting and Sculpture', 'Apelles', 'Fayum portrait', 'Orthodox Icons', 'Wu Da

```python
# Add edges based on relationships
for _, row in relationships_art.iterrows():
    artist = row['artist_name']

    if pd.notna(row['friends']):
        for friend in row['friends'].split(','):
            G1.add_edge(artist, friend.strip(), relationship='friend')

    if pd.notna(row['influenced_by']):
        for influencer in row['influenced_by'].split(','):
            G1.add_edge(artist, influencer.strip(), relationship='influenced_by'

    if pd.notna(row['influenced_on']):
        for influence in row['influenced_on'].split(','):
            G1.add_edge(artist, influence.strip(), relationship='influences_on')

    if pd.notna(row['institution']):
        for institution in row['institution'].split(','):
            G1.add_edge(artist, institution.strip(), relationship='studied_at')
    if pd.notna(row['school']):
        for institution in row['school'].split(','):
            G1.add_edge(artist, institution.strip(), relationship='studied_at')
    if pd.notna(row['movements']):
        for mov in row['movements'].split(','):
            G1.add_edge(artist, mov.strip(), relationship='movement')

#add nation edges
for _, row in artists.iterrows():
    artist = row['title']
```

```
        if pd.notna(row['nation']):
            for nat in row['nation'].split(','):
                G1.add edge(artist, nat.strip(), relationship='nation')
```

```
print("Total edges after addition:", len(G1.edges))
```

Total edges after addition: 3973

```
print(G1.edges)
```

[('Ancient Egypt', 'Egyptians'), ('Ancient Greek Pottery', 'Greeks'), ('Ancient Greek Painting and Sculpture', 'Greeks'), ('Apelles', 'G

◀ ⬤                                                                                                                                      ▶

```
degree_seq = [d for n,d in G1.degree()]
aver_deg = sum(degree_seq)/len(degree_seq)
print("Average_degree: ",aver_deg)
```

Average_degree:  2.08392341987936

```
density = nx.density(G1)
density
```

0.0005466745592548164

```
# Remove nodes of a certain type, e.g., nodes without edges (isolated)
isolated_nodes = [node for node, degree in G1.degree() if degree == 0]
G1.remove_nodes_from(isolated_nodes)
```

```
# Visualization using networkx
plt.figure(figsize=(80, 80))
pos = nx.spring_layout(G1, k=.1)
nx.draw(G1, pos, node_size=180, with_labels=True, font_size=30)
# plt.savefig("network_visualization_new.png", format="PNG", dpi=300)
plt.title("Network Visualization")
plt.show()
```