

Proyecto final (análisis de redes sociales):
Análisis de red de tags de Steam

Miguel Ferreras Chumillas / Pablo Daurell Marina
(Grupo 3)



STEAM™

1. Introducción:	3
2. Creación de la red:	3
2.1. Obtención de los datos	3
2.2. Construcción de la red	3
3. Análisis de la red	5
3.1 Distribución de grado	7
3.2 Peso de las aristas	10
3.3 Centralidad	13
· Centralidad de intermediación	13
· Centralidad de cercanía	15
4. Comunidades	17
4.1 Comunidades con todos los tags	17
· Comunidad Indie, Acción, Aventura, Casual	18
· Comunidad Simulación, Estrategia, RPG, Multijugador	19
· Comunidades menores	21
4.2 Comunidades sin “Indie, Action, Adventure y Casual”	23
· Comunidad de Software Multimedia	24
· Comunidad Anime	25
· Comunidad Multijugador	27
· Comunidad puzzle y plataformas	29
· Comunidad Simulación, RPG y Estrategia	31
5. Conclusiones finales	33
Anexo: Red de juegos de Steam	34

1. Introducción:

[Steam](#) es una de las mayores plataformas de venta digital de videojuegos para PC, con una gran popularidad y antigüedad frente a otras plataformas similares (como [Origin](#), [GOG](#), o [Epic Games](#) entre otras). Cuenta con un extenso catálogo de juegos de una gran variedad de desarrolladores. Por todo esto, Steam es una muy buena representación de la industria de los videojuegos actualmente.

Steam cuenta con una funcionalidad mediante la cual sus usuarios pueden asignar ciertas etiquetas (o tags) a los juegos para clasificarlos en función del tipo del género al que pertenezcan.

En este proyecto se analizarán las relaciones entre los tags que los usuarios asignan a los juegos de la plataforma, para lo cual construiremos una red, no dirigida, donde los nodos serán los distintos tags y las aristas representarán juegos que poseen ambos tags. Se estudiarán e interpretarán distintas métricas de la red y se analizarán distintas comunidades dentro de esta.

Todo este análisis tendrá como objetivo ver qué tipo de juegos son los que más abundan en Steam, qué tags suelen aparecer en más juegos juntos y en general entender cómo se clasifican los juegos dentro de Steam.

2. Creación de la red:

2.1. Obtención de los datos

Para construir la red utilizamos los datos extraídos de un dataset encontrado en Kaggle (<https://www.kaggle.com/nikdavis/steam-store-games>) que proporciona gran variedad de datos de 27.075 juegos de Steam, extraídos en mayo de 2019, entre los que se encuentran los tags principales de cada juego (en total 339 tags distintos). En concreto se ha utilizado el archivo [Steam.csv](#) que contiene el nombre de cada juego y distintas características de este:

[*appid* | ***name*** | *release_date* | *english* | *developer* | *publisher* | *platforms* | *required_age* | *categories* | *genres* | ***steamspy_tags*** | *achievements* | *positive_ratings* | *negative_ratings* | *average_playtime* | *median_playtime* | *owners* | *price*]

Las columnas que nos interesan son *name* que indica el nombre del juego y *steamspy_tags* que contiene una lista de los principales tags añadidos por los usuarios a ese juego.

2.2. Construcción de la red

La red se ha construido con Python, usando el paquete **Pandas** para tratar los datos del fichero csv y el paquete **Networkx** para construir el grafo.

En primer lugar se extrajeron los datos necesarios del fichero Steam.csv (*name* y *steamspy_tags*) a un dataframe de Pandas para poder tratarlos. A partir del dataframe se procesaron los datos, extrayendo los tags en una lista de tags y creando un diccionario de tags junto a sus respectivos juegos. Finalmente se recorrería el diccionario de tags y juegos para generar un grafo con Networkx de forma que los nodos sean los tags y estén unidos por el número de juegos que comparten mediante una arista con peso

Código para tratar los datos:

```
### Cargar datos sv
df = pd.read_csv("steam.csv")
data = df[['name', 'steamspy_tags']]
data = data.drop(data[data['name'].duplicated() == True].index) # Eliminar nombres duplicados

### Lista de tags
tags = {}
for i in data.tags:
    for t in i.split(';'):
        if(tags.get(t)):
            c = tags.get(t) + 1
            tags.update({t: c})
        else:
            tags.update({t: 1})
tags = sorted(tags, key=tags.get, reverse=True)

### Dict con tags y sus juegos
data_list = data.values.tolist()
tag_dict = {}
for t in tags:
    l = []
    for d in data_list:
        if d[1].split(';').count(t) > 0:
            l.append(d[0])
    tag_dict.update({t: l})

### Función para calcular peso de cada arista
def tags_in_common(a, b):
    n = 0
    for i in a:
        for j in b:
            if i == j:
                n = n+1
    return n
```

Código para generar el grafo:

```
### Generar grafo
G_tags = nx.Graph()
G_tags.add_nodes_from(tags)
tag_list = list(tag_dict.items())
i = 0
while i < len(tag_dict):
    print(i, end='\r')
    t1 = tag_list[i][1] # Lista de tags juego 1
    j = i + 1
```

```

while j < len(tag_dict):
    t2 = tag_list[j][1] # Lista de tags juego 2
    n = tags_in_common(t1, t2)
    if n > 0:
        G_tags.add_edge(tag_list[i][0], tag_list[j][0], weight=n)
    j += 1
i += 1

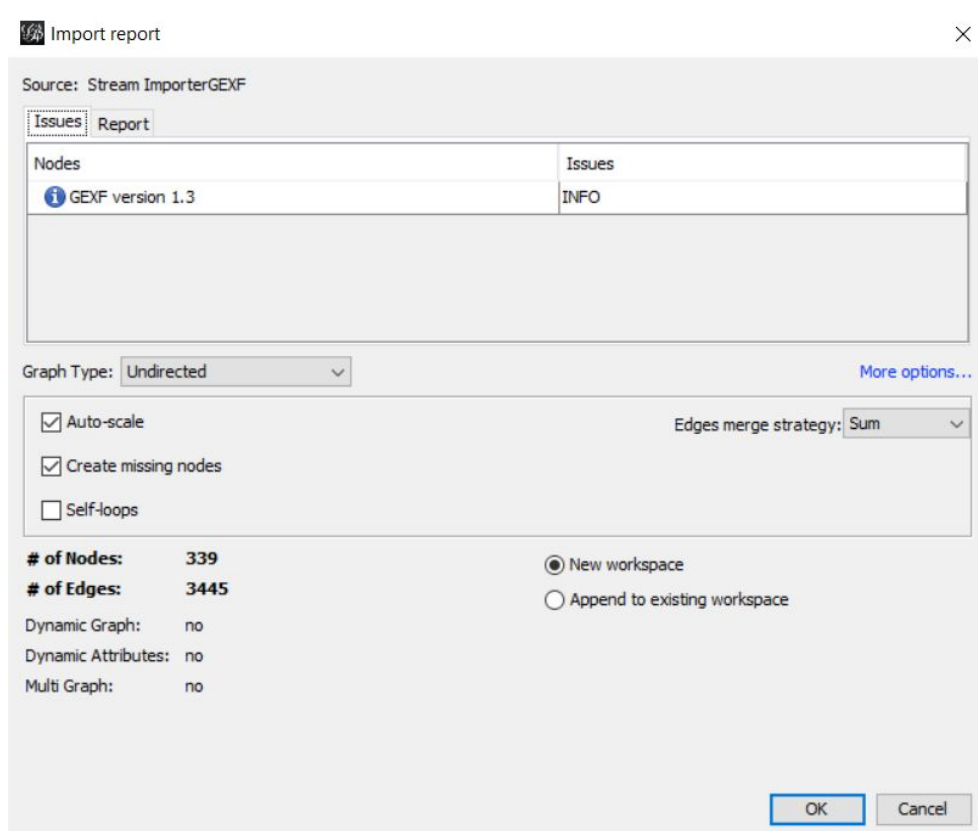
nx.write_gexf(G, 'Tags.gexf')

```

Después de procesar los datos y generar el grafo, exportamos el resultado a un archivo gexf para poder abrir, visualizar y analizar la red con **Gephi**.

3. Análisis de la red

Importando el archivo *Tags.gexf* a Gephi podremos generar un archivo en formato gephi para manejar la red.

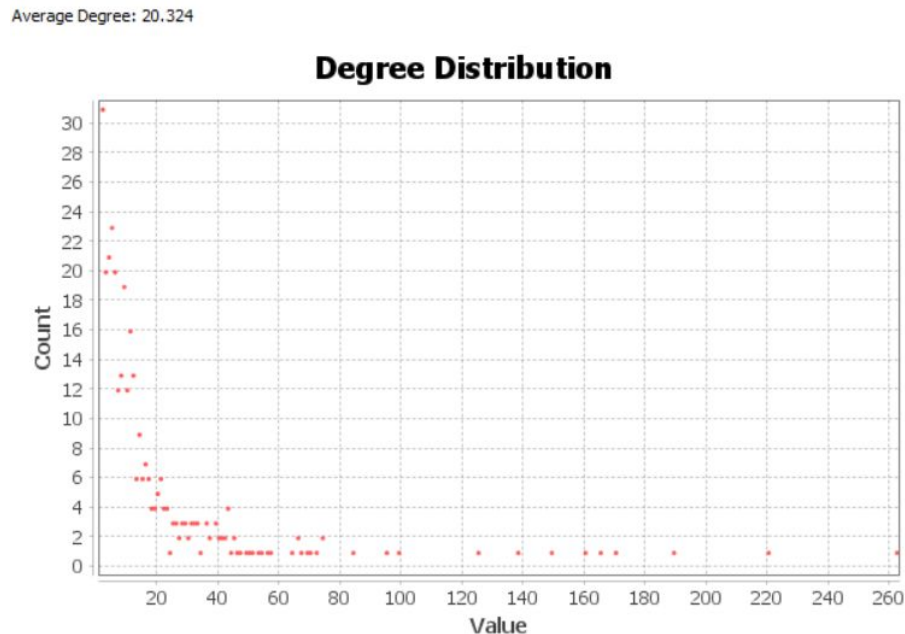


Vemos que la red resultante consta de **339 nodos** (tags) y **3445 aristas** con peso.

Pasamos a visualizarla, aplicando el layout ForceAtlas2 de Gephi y estableciendo que el tamaño y color de los nodos vaya en función de su grado y el grosor de las aristas en función de su peso:

Para entrar más en detalle, analizaremos distintas métricas de la red calculándolas con las herramientas que ofrece Gephi y las exportaremos a un fichero csv para generar ciertas gráficas con Networkx.

3.1 Distribución de grado

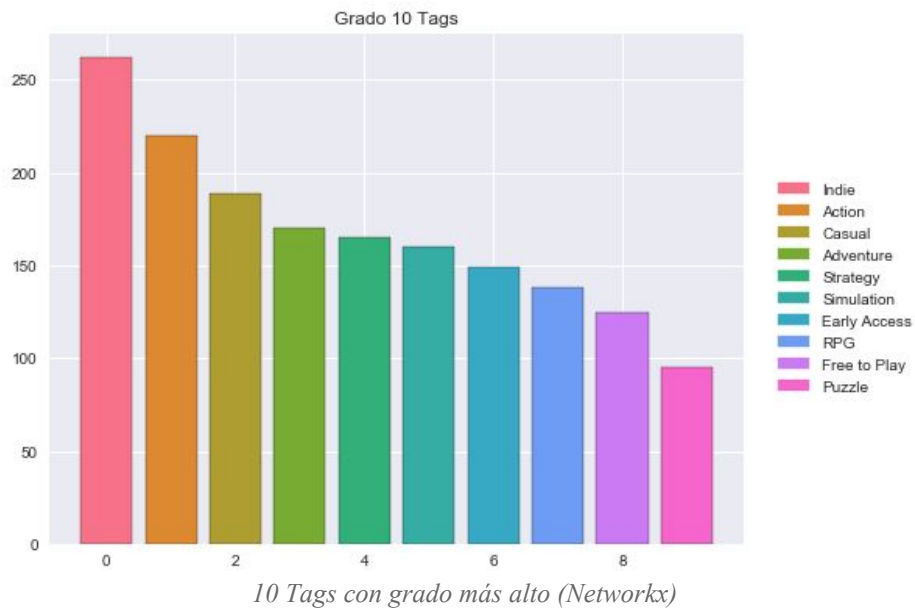
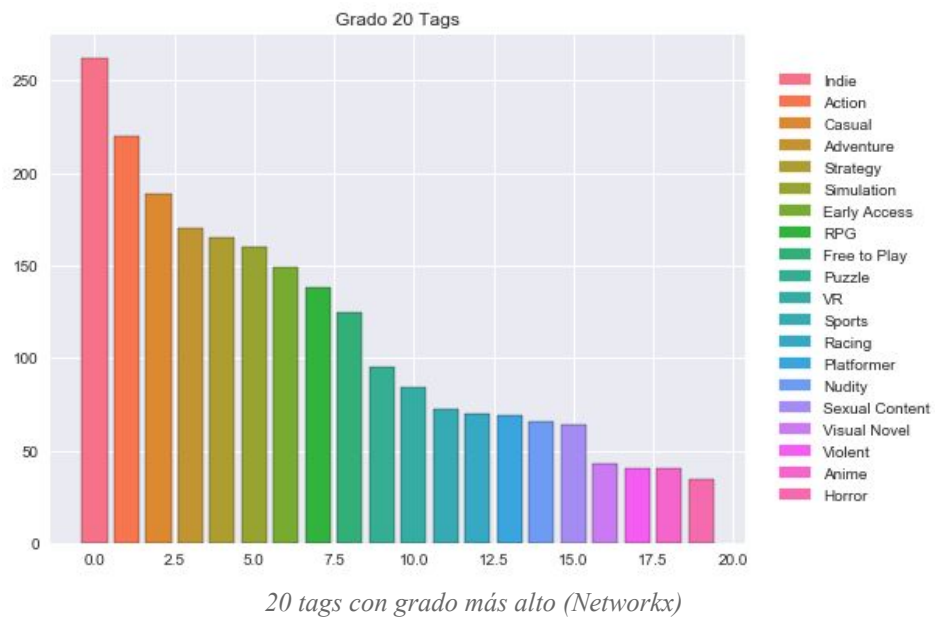


Distribución de grado de la red (Gephi)

Empezamos con la distribución de grado de la red. Como podíamos ver en la visualización del grafo, la mayoría de nodos se concentran en valores bajos o intermedios, mientras que sólo unos pocos nodos poseen valores de grado alto. Además observamos que a pesar de existir nodos con grados entre 100 y 260, el grado medio es de 20,324.

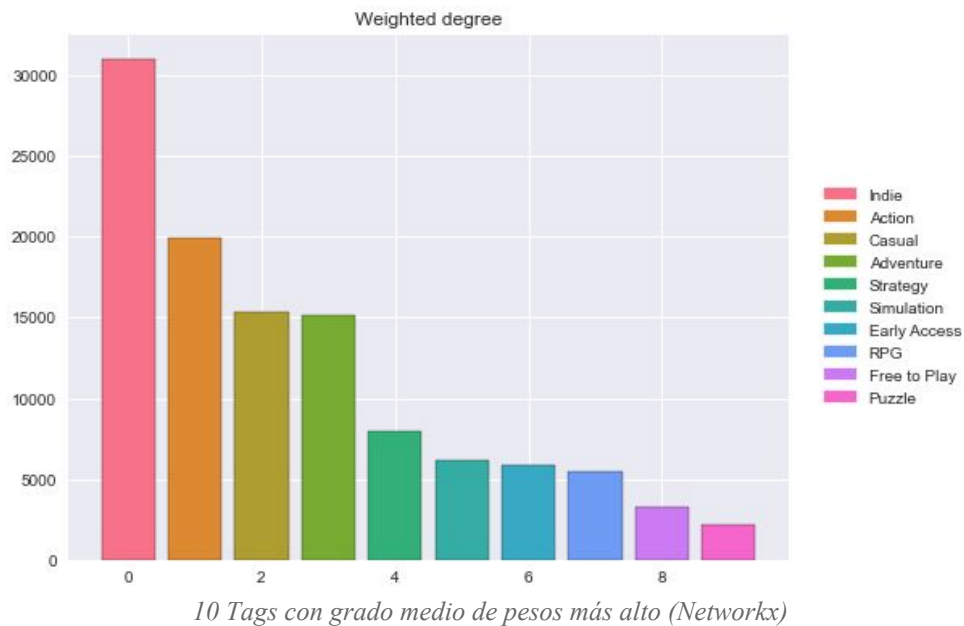
Como establecíamos anteriormente, esto significa que existen unos pocos tags con mucha relevancia dentro de Steam y muchos tags que actuarán de una forma más complementaria.

A continuación, veremos qué tags son los que tienen más número de relaciones y por tanto un mayor grado:



Podemos ver cuáles son los tags con más conexiones de toda la red, destacando sobretudo los géneros Indie y Action.

También podemos ver el grado medio de pesos de cada tag, para ver la media del peso de las aristas de cada nodo lo cual indicará la media de juegos clasificados con ese tag:



Observamos como el orden de relevancia de los tags se mantiene con respecto a la gráfica anterior.

En la gráfica de grados podíamos ver un descenso más progresivo ya que los tags más relevantes se relacionan entre sí mutuamente (por tanto tendrán un grado similar) y, a pesar de estos tags más relevantes, incluso los tags menos importantes también se relacionan bastante con otros tags de la red. Sin embargo en el caso de los grados medios de pesos, el tag Indie tiene una clara superioridad frente al resto, seguido de Acción, Aventura y Estrategia. Esto indica que incluso dentro de los tags más influyentes y con muchas conexiones, Indie es un tag mucho más presente en los juegos de la red.

De estas dos métricas podemos sacar que la importancia de la red se centra en unos pocos tags, siendo estos Indie, Acción, Casual, Aventura, Estrategia, Simulación, Early Access, RPG y Free To Play:

- Indie: Es fácil concluir gracias a la gráfica de grado con pesos que la mayoría de los juegos que alberga la plataforma de Steam no están desarrollados por grandes compañías. Si miramos en Steamworks (herramienta que permite a los desarrolladores publicar sus juegos en steam) podemos ver que cualquiera puede hacerlo. Esto le da sentido a la predominancia del indie en la plataforma.

Por otro lado, en cuanto a la gráfica de grados, nos permite ver que el indie es el género que más se relaciona con el resto de etiquetas. Algo que tiene sentido, al no ser el indie un género como tal, sino más una característica, lo que hace que requiera de otras etiquetas para describir un juego.

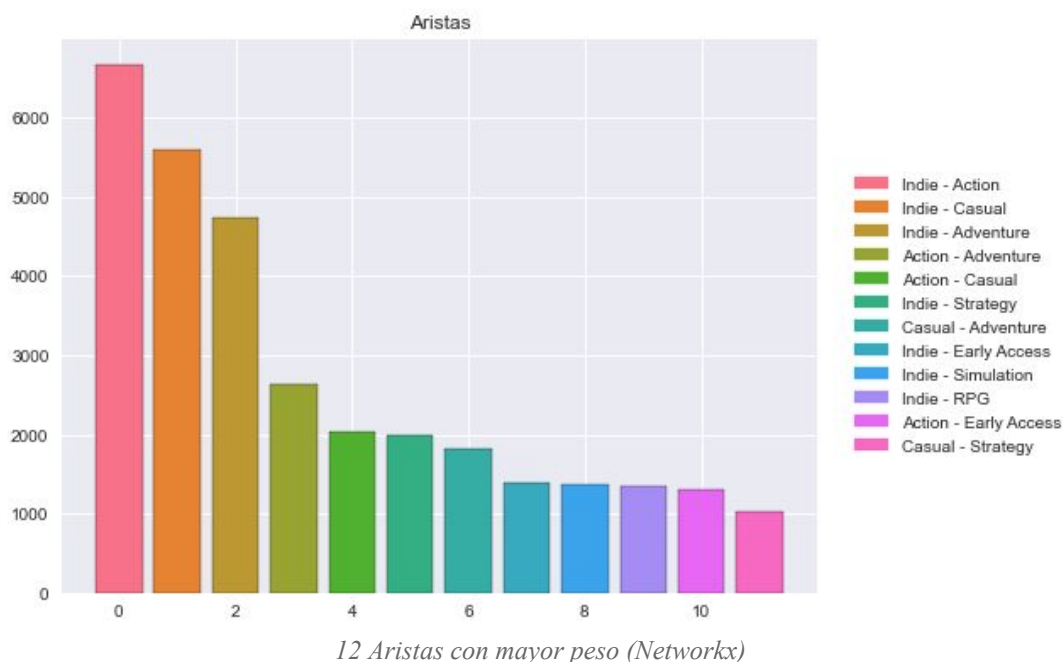
- Casual: Al igual que el indie, no representa una mecánica ni un género en específico. El casual representa un estilo de juego tranquilo, de mecánicas simples y que requiere poca dedicación. Esto explica que tenga un grado tan alto ya que necesita de otros tags para definir correctamente a un juego.

También, al encontrarse como el tercer tag podemos intuir que una gran cantidad de los juegos de steam son relativamente simples. Esto puede deberse a que como hemos mencionado antes, la mayoría de sus juegos están desarrollados por estudios o autores independientes, y presumiblemente sin demasiados medios.

- Acción y aventura: Ambos son géneros muy populares y con un gran mercado, era esperable que estuviesen entre los primeros y las métricas nos lo confirman. Además, aunque estos tags si representan un estilo de juego muy claro, son tan amplios que admiten fácilmente otros tags para definir mejor el estilo del juego.
- Estrategia, simulación y RPG: Aunque a priori parecen géneros tan grandes como los anteriores, en las gráficas vemos que, a pesar de ser tags amplios, necesitan de otros para clasificarse, cuando nos movemos a la de pesos tienen una importancia mucho menor en la red, con muchos menos juegos.
- Early access: Son juegos de acceso anticipado, en versiones alpha, normalmente con la intención de conseguir recaudación para continuar con el desarrollo. La aparición de este tag tan arriba nos vuelve a confirmar la relevancia de los juegos indie detrás de esta plataforma, juegos que no cuentan con una gran empresa o un gran capital que los sustente.

3.2 Peso de las aristas

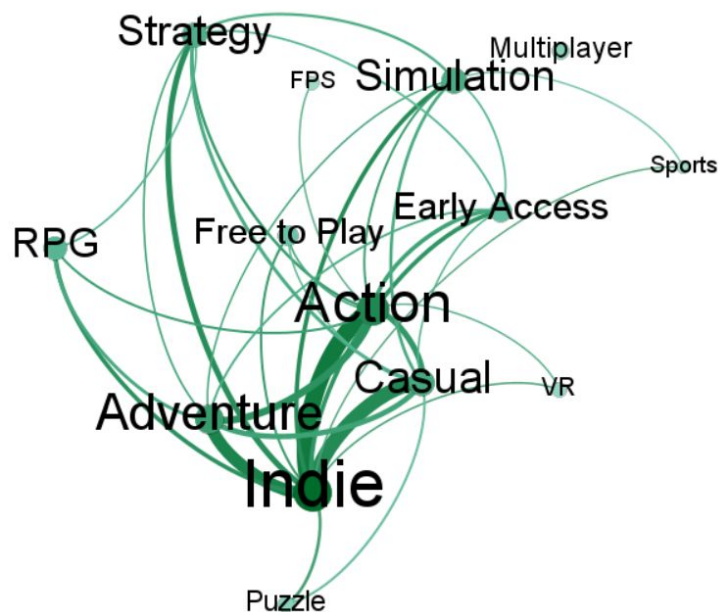
Continuamos el estudio de la red analizando las aristas y sus pesos.



El peso de las aristas de nuestra red representa el número de juegos que contienen los dos tags unidos por esa arista.

Es muy notorio el amplio rango de pesos que hay en las doce aristas más grandes, yendo desde los 1022 hasta los 6660, los tags presentes en estas aristas coinciden con los tags más relevantes, mencionados en el apartado anterior. Esto muestra como estos tags no solo son

importantes para el resto de la red, sino que tienen una gran relación entre ellos, son tan relevantes dentro de la red que tienden a aparecer en muchos juegos juntos.



Aristas más grandes de la red (Gephi)

Source	Target	Weight
Indie	Action	6660.0
Indie	Casual	5590.0
Indie	Adventure	4740.0
Action	Adventure	2632.0
Action	Casual	2030.0
Indie	Strategy	1990.0
Casual	Adventure	1826.0
Indie	Early Access	1379.0
Indie	Simulation	1372.0
Indie	RPG	1341.0
Action	Early Access	1297.0
Casual	Strategy	1022.0
Adventure	RPG	976.0
Casual	Simulation	863.0
Indie	Puzzle	849.0
Action	Strategy	740.0
Strategy	Simulation	706.0
Casual	Puzzle	595.0
Action	RPG	590.0
Indie	Free to Play	551.0
Adventure	Early Access	503.0
Indie	Platformer	496.0
Action	Simulation	479.0
Adventure	Point & Click	455.0
Strategy	Early Access	451.0
Casual	Hidden Object	419.0
Strategy	RPG	396.0
Adventure	Strategy	388.0
Simulation	Early Access	387.0
Adventure	Hidden Object	373.0
Nudity	Sexual Content	370.0
Action	Free to Play	369.0
Indie	VR	355.0
Indie	Racing	339.0

34 aristas ordenadas por peso (Gephi)

A parte de las aristas entre los tags más relevantes, que tienen los valores de pesos más altos de la red, podemos encontrar otras relaciones interesantes y con bastante peso. Veamos algunos ejemplos:

- Indie y Free To Play: Son dos de los géneros más populares, sin embargo es destacable como el tag “Free To Play” no se relaciona tanto con el resto de tags relevantes, solo con “Indie”. Esto demuestra que la mayoría de juegos gratuitos son juegos de pequeños estudios que probablemente opten por otros modelos de financiación (probablemente anuncios o microtransacciones), dejando el acceso a sus juegos gratuitos para tener un mayor alcance.
- Indie y Platformer: Los plataformas son un clásico de los videojuegos y el tag “Platformer” posee un grado bastante alto pero no llega a ser tan relevante como otros tags. Aún así vemos que los juegos Indie apuestan bastante por este género.
- Adventure y Point & Click: El tag “Point & Click” solo tiene un grado de 46 y una media de peso de 948. Podemos ver de que no se trata de un tag demasiado relevante y que gran parte de sus juegos (47,9 %) son también juegos de Aventura. Esto se debe a la relevancia en la industria de los videojuegos de las *Aventuras Gráficas*, juegos que en su mayoría se controlan únicamente con el ratón.
- Nudity y Sexual Content: Dos tags en principio poco relevantes en la red (con 34 y 40 grados respectivamente), sin embargo forman primera arista de peso alto que no incluye ninguno de los tags más influyentes. El motivo es bastante claro ya que los juegos de contenido sexual muy probablemente muestren desnudos.

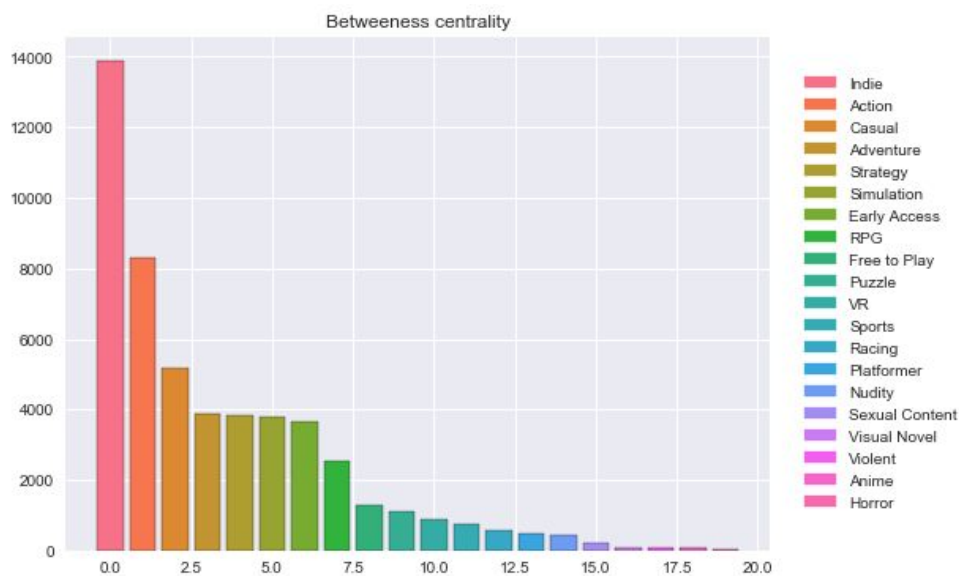
Esta relación entre tags más pequeños nos indica que puede interesante estudiar esta red a nivel de comunidades.

3.3 Centralidad

Otro aspecto a tener en cuenta de la red es su centralidad. Podemos calcular distintos enfoques de esta para poder reafirmar cuáles son los tags más relevantes de la red (y por tanto más centrales) y ver mejor de qué manera se relacionan con el resto de tags.

· Centralidad de intermediación

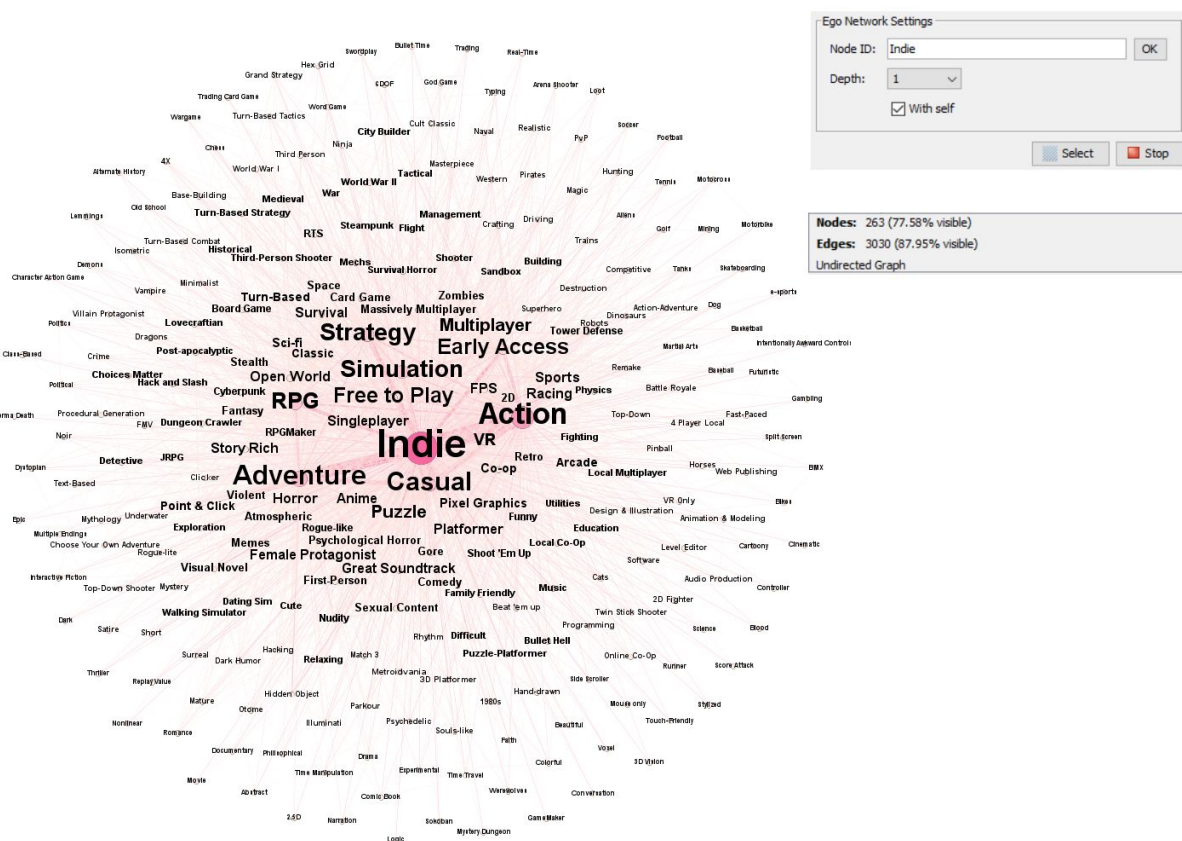
La centralidad de intermediación tiene en cuenta el número de caminos mínimos que atraviesan cada nodo, por tanto nos da información de mediante qué nodos se conectan distintas partes de la red.



20 tags con mayor centralidad de intermediación (Networkx)

Los tags con mayor centralidad de intermediación son aquellos que suponen enlaces importantes entre conjuntos de tags no adyacentes, eliminando estos tags tendríamos una red formada por distintas regiones de tags más separadas unas de otras.

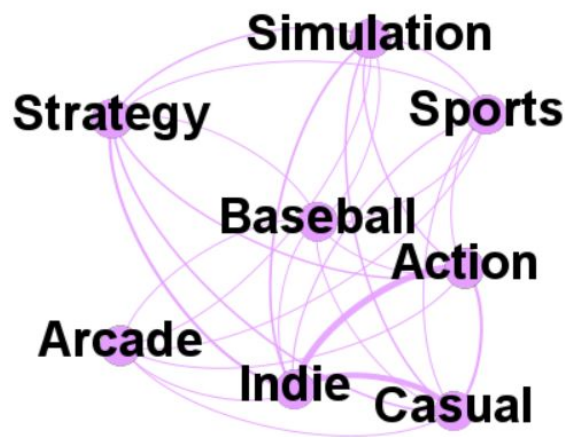
Los tags con mayor centralidad de intermediación coinciden con los tags con mayor grado. Indie destaca claramente por encima del resto debido a que es una característica tan amplia que puede relacionarse con gran variedad de tags, podemos ver como aún filtrando los tags no directamente conectados a Indie seguimos teniendo un 77.5% de la red:



Tags conectados directamente con "Indie" (Gephi)

De esto podemos concluir que los tags más relevantes de toda la red, sobretodo Indie y Action al definir características tan poco restrictivas se aparecen mucho junto a cualquier tipo de tag, en contraposición a tags mucho más específicos que por tanto sólo tiene sentido que aparezcan con cierto tipo de tags, uno de los muchos ejemplos de esto puede ser el tag Baseball:

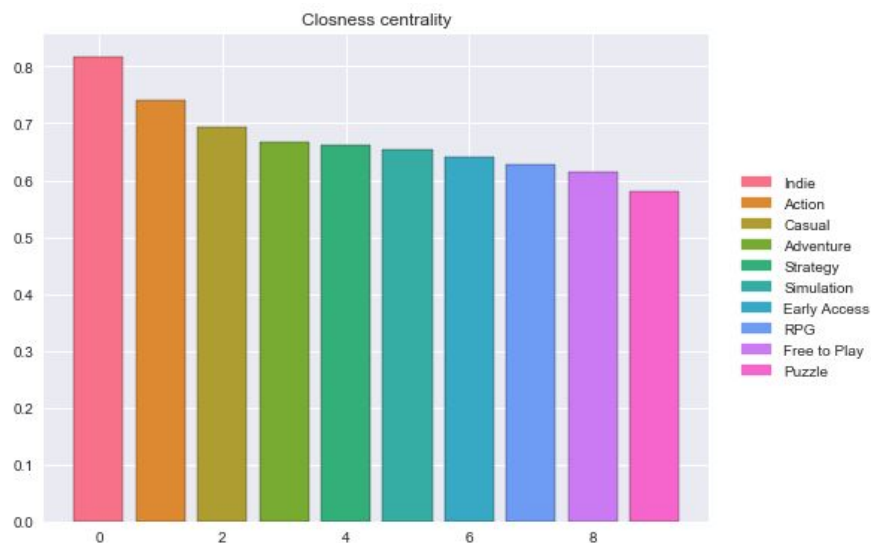
Id	Label	Betweenness Centrality
Baseball	Baseball	0.157118



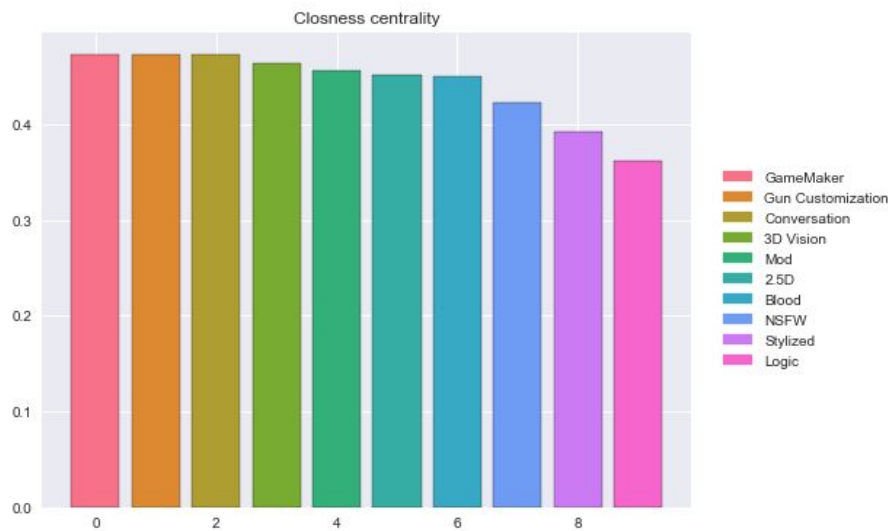
Tags directamente conectados a "Baseball" (Gephi)

· Centralidad de cercanía

Por otro lado, la centralidad de cercanía mide la facilidad con la que se puede llegar de un nodo al resto de nodos de la red. Con esto podemos saber qué tags son los que más se relacionan con la red en general.



Centralidad de cercanía de los tags más relevantes (Networkx)



Centralidad de cercanía de los tags menos relevantes (Networkx)

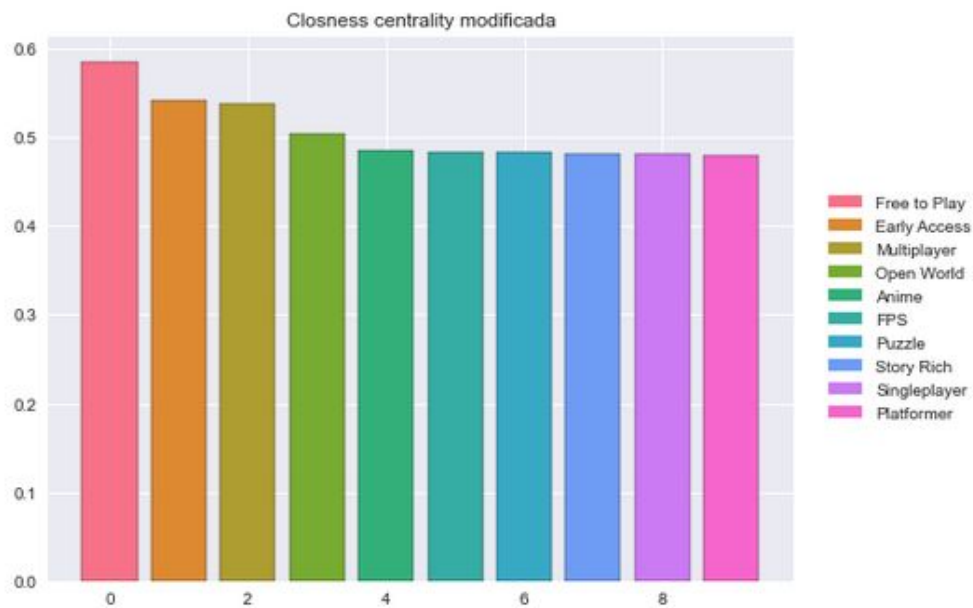
Como vemos, entre los tags más relevantes y los menos relevantes no hay una diferencia abismal como ocurría con otras medidas, oscilando entre 0,8 y 0,4 aproximadamente. Esto se debe a que no es una red excesivamente grande y está muy conectada (el diámetro de la red es de 4 solamente lo cual nos dice que no es una red para nada dispersa).

El hecho de que un tag tenga una centralidad de cercanía muy alta puede significar que describe un género muy amplio y por tanto precisa de otros tags para definir mejor a un juego.

Por otro lado, que los tags pequeños no tengan valores muy bajos puede deberse a que:

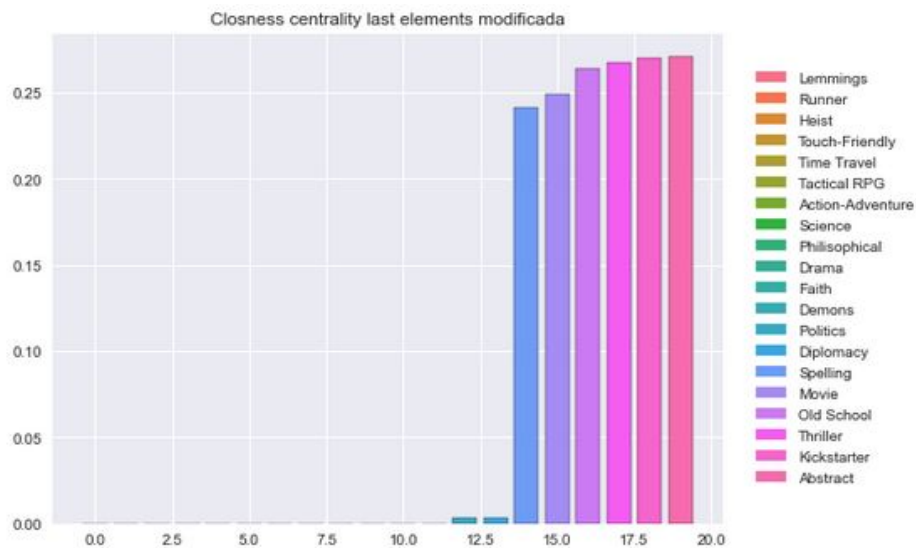
- No se relacionen tanto con otros tags pequeños sin embargo sí se relacionen con alguno de los tags grandes que a su vez sí se relacionan con el resto de tags pequeños.
- O que sí que se relacionen mucho entre sí.

Para aclarar esto volvimos a tomar las medidas quitando los nodos con valores más más grandes.



Centralidad de cercanía sin los tags con mayor valor (Networkx)

Quitando los siete primeros nodos podemos ver que en la parte superior no hay un cambio muy apreciable.

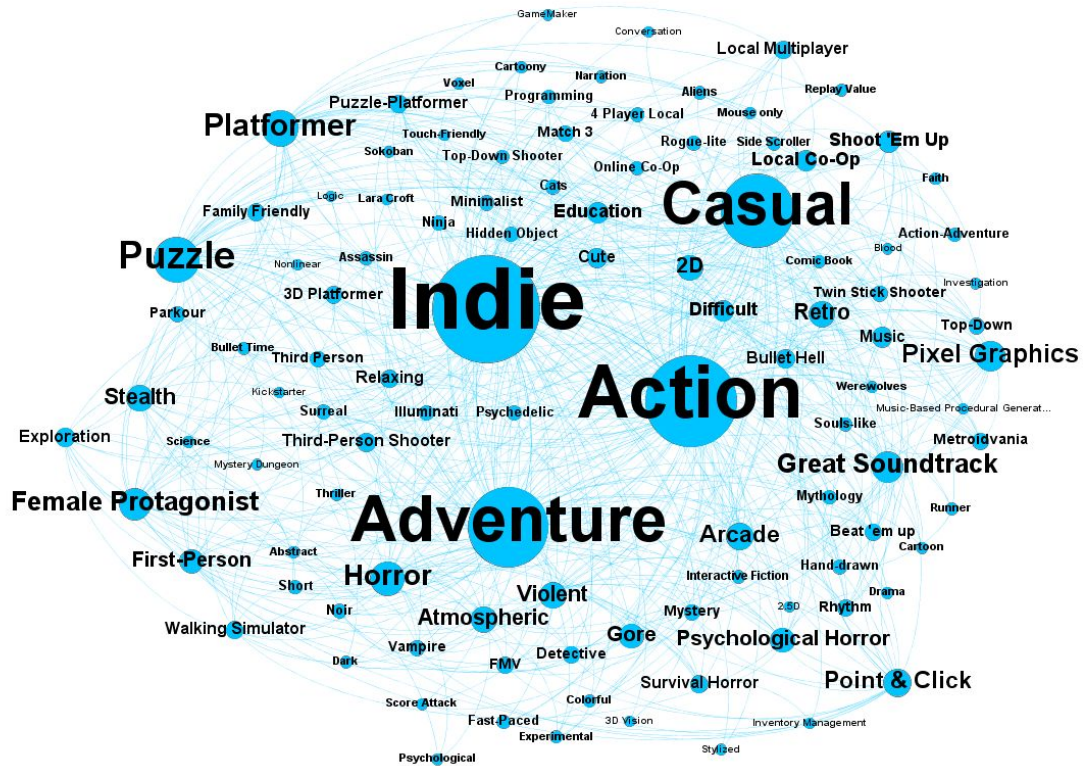


Centralidad de cercanía de los tags menos relevantes, sin los tags con mayor valor (Networkx)

Sin embargo al mirar en la parte de abajo podemos ver como muchos nodos se desconectan del resto de la red o bajan en mucho su centralidad, lo que indica que es una red muy dividida en comunidades de tags que están muy relacionados entre ellos, pero que sin embargo algunos de sus nodos solo se conectan con el resto de la red mediante los géneros más grandes. Para ver esto estudiamos la estructura de comunidades de la red.

Como mencionamos anteriormente, los tags más importantes de la red (Indie, Action, Adventure, Casual, Strategy, etc) han influido bastante en la creación de de comunidades, dando lugar a dos grandes grupos, uno marcado por los tags Indie, Action, Adventure y Casual (comunidad azul) y otro por Strategy, Simulation, Early Access y Free To Play. Estas dos comunidades son quizás demasiado generales y tienen poco margen de análisis.

· Comunidad Indie, Acción, Aventura, Casual



Comunidad de la red con todos los tags (Gephi)

Nodos: 107

Aristas: 660

Grado medio

12,336

Grado medio con pesos

661,065

A primera vista podemos ver que es una comunidad muy grande, que corresponde a casi un tercio de la red completa y que junta géneros totalmente distintos, como pueden ser “pixel graphics” con “walking simulator” y “conversation”. Aunque estos géneros puedan compartir algún juego, consideramos que el juntarlos a todos no se corresponde con una distribución adecuada de las comunidades, ya que se podrían separar mucho mejor.



· Comunidad Simulación, Estrategia, RPG, Multijugador

[illegible]

Comunidad de la red con todos los tags (Gephi)

Nodos: 210
Aristas: 1359

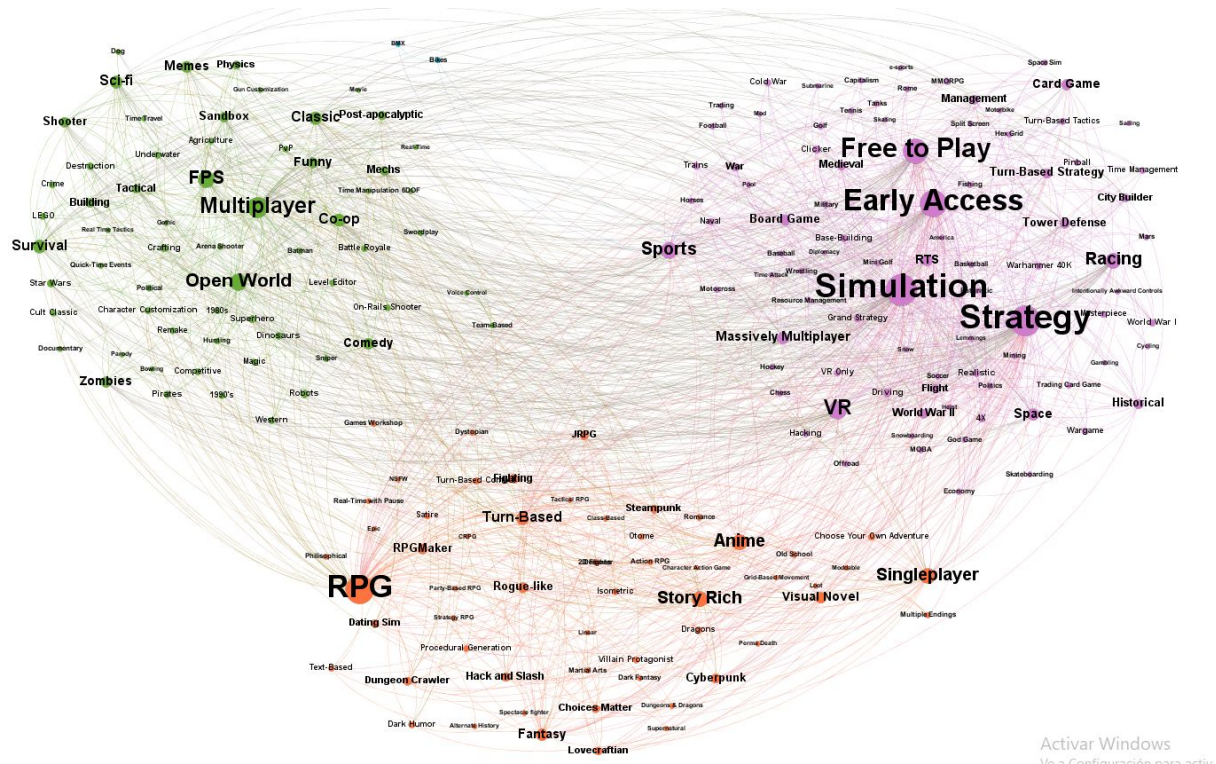
Grado medio

12,943

Grado medio con pesos

233,305

Una comunidad que ocupa dos tercios de la red es difícil de analizar, por eso la dividimos en subcomunidades



Subcomunidades de la comunidad más grande red completa (Gephi)

Primero vemos es una subcomunidad de solo dos nodos (arriba a la izquierda) formada por “Bikes” y “BMX” que no es demasiado relevante.

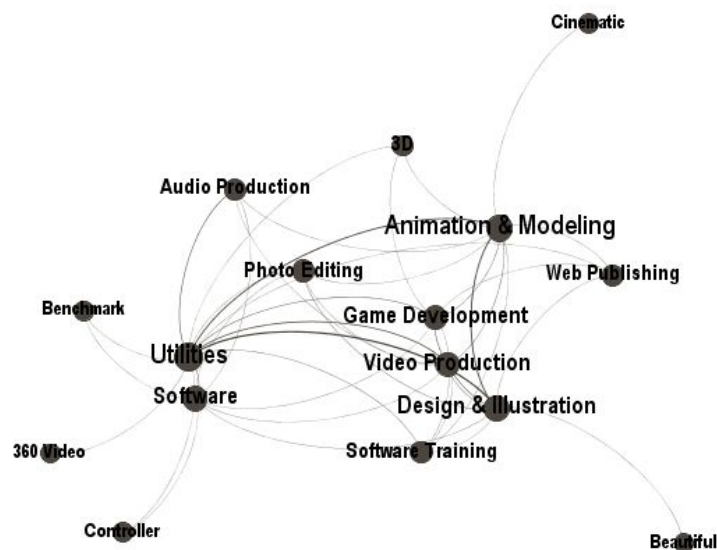
El resto de subcomunidades, a diferencia de en la anterior comunidad, tienen algo de sentido, aunque haya una serie de nodos que nos descuadran un poco como juntar “RPG” con “Dating Simulation”.

En general los nodos morados representan tanto simulación deportiva como estrategia más enfocada a la economía en los juegos o a la creación de ciudades, algo que podríamos ver en un juego como *Age of Empires* por ejemplo. Por otra parte, los verdes representan juegos de disparos multijugador y los naranjas los RPG con hack and slash o turn based.

En definitiva aunque esta comunidad no esté tan influida por los nodos más grandes como sí hacía la anterior, corresponde a una porción demasiado grande de la red y solo se vuelve interesante su análisis cuando se divide.

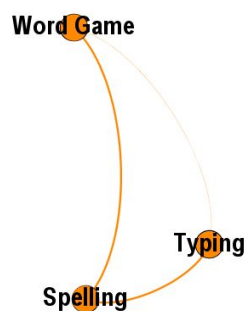
· Comunidades menores

A parte tenemos otras tres comunidades notablemente más pequeñas:



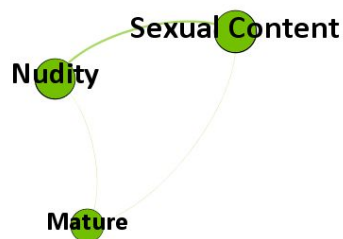
Comunidad de la red con todos los tags (Gephi)

- La primera y más grande está relacionada con software multimedia (Steam aparte de vender videojuegos tiene un pequeño catálogo de este tipo de programas). Destaca el nodo de “Utilities” sobretodo en el peso de sus aristas, debido probablemente a que no es un tag muy cerrado, menos dentro de esta comunidad, y quizás es destacable el hecho de que el nodo “Education” no esté aquí dentro así que es posible que la influencia de los grandes tags lo haya excluido de esta comunidad.



Comunidad de la red con todos los tags (Gephi)

- La segunda comunidad es muy clara con tres nodos relacionados en el ámbito de las palabras y la escritura. Los tres nodos tienen pocas aristas que los conecten fuera de la comunidad.

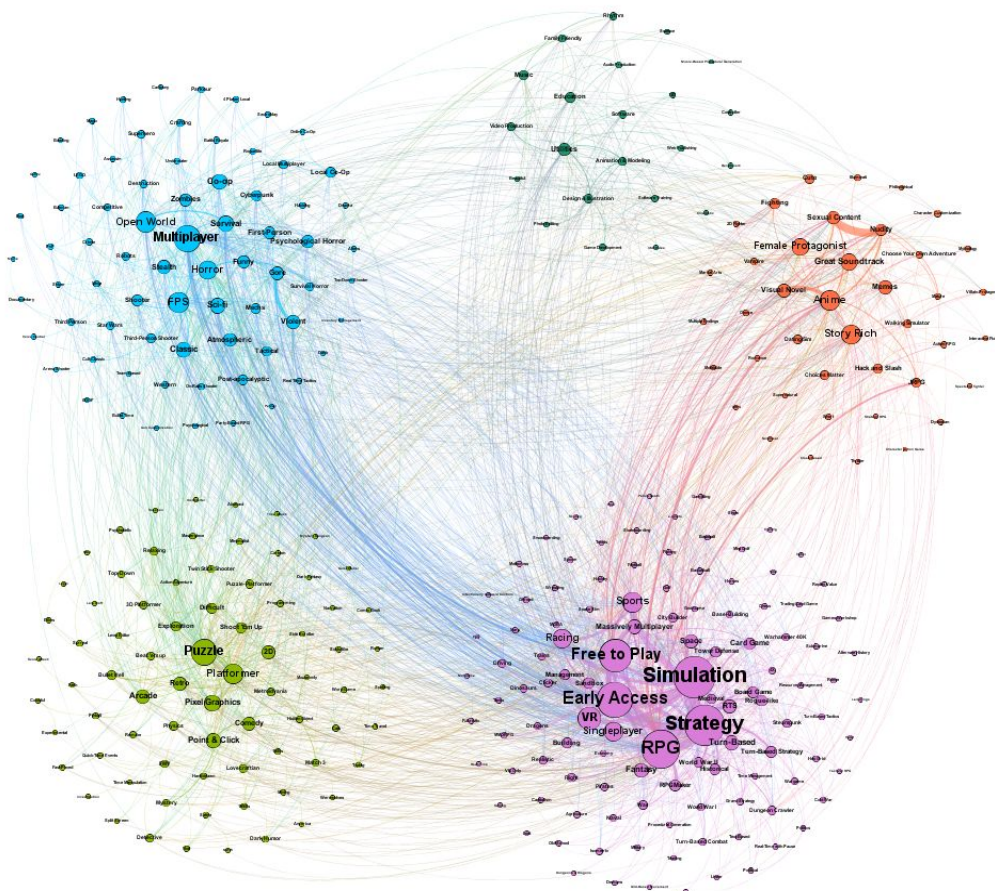


Comunidad de la red con todos los tags (Gephi)

- La última comunidad, aunque también tiene un concepto muy claro no nos acaba de convencer, creemos que estos tags que representan un contenido adulto están más asociados también a otros géneros en concreto, y una comunidad tan pequeña no nos permite ver esto.

4.2 Comunidades sin “Indie, Action, Adventure y Casual”

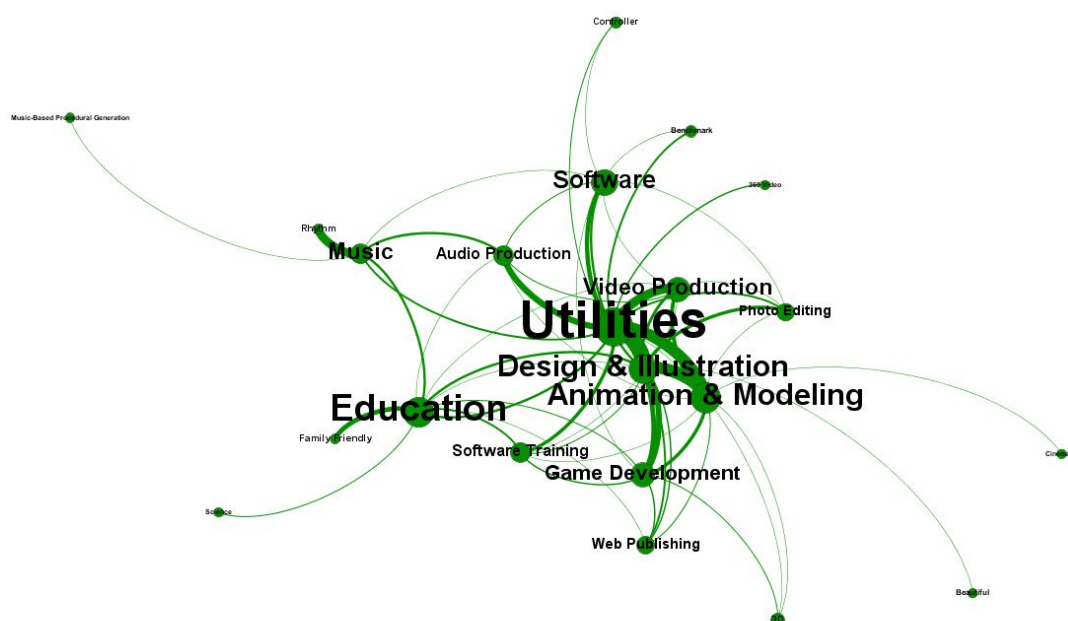
Como hemos visto en el apartado anterior, los tags principales de la red influyen demasiado en la generación de comunidades. Procederemos a eliminar los tags Indie, Action, Adventure y Casual:



Comunidades de la red sin los tags principales (Gephi)

Al quitar estos nodos se puede apreciar rápidamente el cambio en las comunidades. Aunque siguen siendo 5 esta vez tienen tamaños mucho más parejos y los tags están mejor distribuidos.

· Comunidad de Software Multimedia



Comunidad de programas multimedia (Gephi)

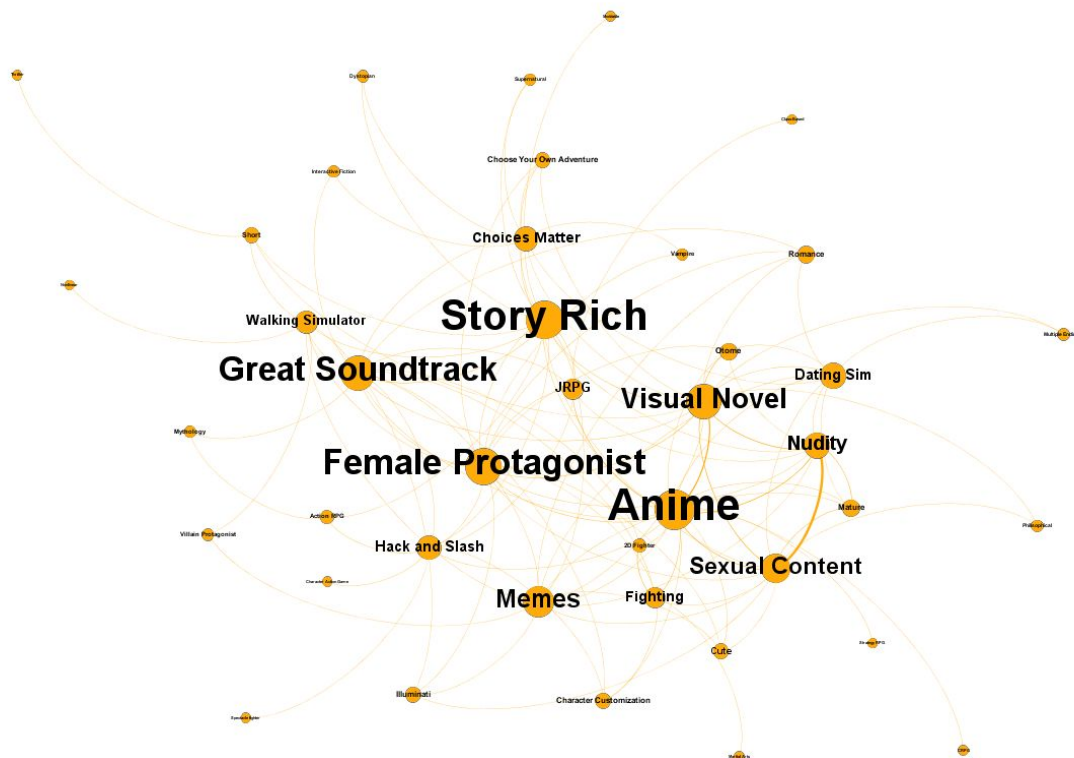
Nodos:	22
Aristas:	57
Grado medio	5,182
Grado medio con pesos	27,455

Es una comunidad relativamente pequeña y muy similar a la mostrada en la división con todos los nodos. Como hemos aclarado antes, surge de que en Steam no solo hay videojuegos sino también una pequeña cantidad de programas multimedia.

Destacan los tags de “Utilities”, “Design & Illustration”, “Animation & Modeling” y “Education”, que esta vez sí entra en la comunidad. También hay tags menores como “Science”, que solo se conecta con educación. O el tag “Beautiful”, que solo se conecta con diseño e ilustración, algo que llama la atención, ya que nos da la impresión de que podría ser utilizado fácilmente para describir juegos con una estética llamativa.

En definitiva es una comunidad dentro de Steam que no se centra en los juegos, si no en el resto de software, aplicaciones que aporten utilidad al usuario o que le sirvan para aprender.

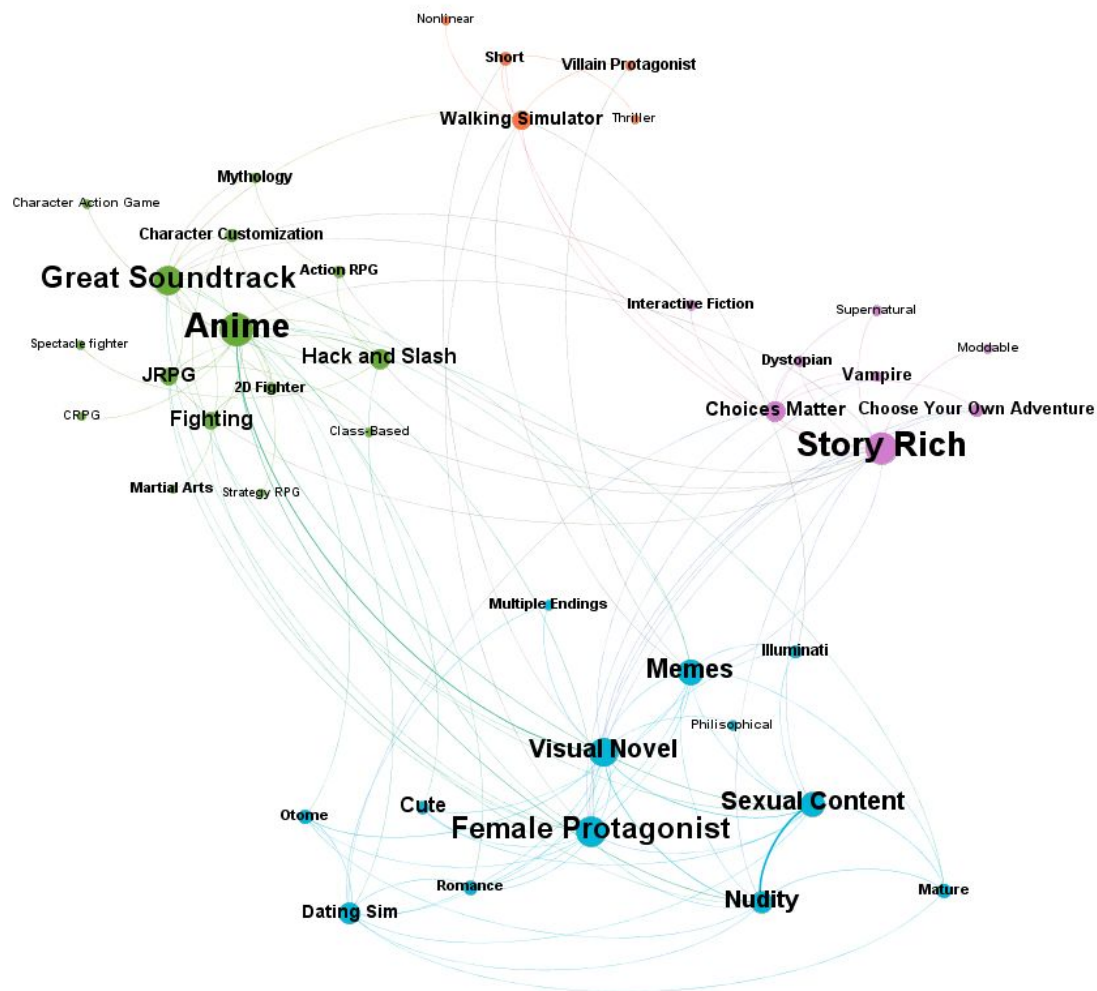
· Comunidad Anime



Comunidad Anime (Gephi)

Nodos:	41
Aristas:	116
Grado medio	5,659
Grado medio con pesos	64,878

Esta comunidad es una mezcla de varios conceptos. Está la idea general de una temática anime (dibujos de animación procedentes de Japón), historias interactivas y contenido adulto. Estos géneros suelen ser comunes en juegos visuales, orientados a público adulto / adolescente, centrados sobre todo en narrar una historia. Sin embargo sigue sin quedar claro cómo se conectan bien estos géneros. Por eso a pesar de ser una comunidad pequeña de 41 nodos, decidimos ver cómo eran sus subcomunidades:



Subcomunidades de la comunidad Anime (Gephi)

Separándolo de esta forma es más fácil ver porqué la comunidad está establecida así:

Por un lado tenemos los nodos verdes, donde destacan dos tipos, los juegos de lucha (Fighting), con artes marciales como podrían ser el *Tekken* o el *Street Fighter*, donde es habitual tener una cantidad considerable de personajes femeninos y normalmente sexualizados. Y por otro lado los JRPG (japanese RPG), con tags como “Hack & Slash” o “Class Based” representados por juegos como *Final Fantasy* centrado sobretodo en narrar una historia y luchar contra enemigos usando distintos personajes, como la mayoría de estos juegos. Además todos los juegos representados en los nodos verdes están caracterizados por una temática Anime, y una gran banda sonora.

En los nodos morados tenemos a su vez que los juegos con una gran historia también se corresponden con aquellos de escoger tu propio camino donde las decisiones importan (“Choice Matter”), géneros que normalmente están representados en novelas visuales y algunos de ellos con protagonistas femeninos.

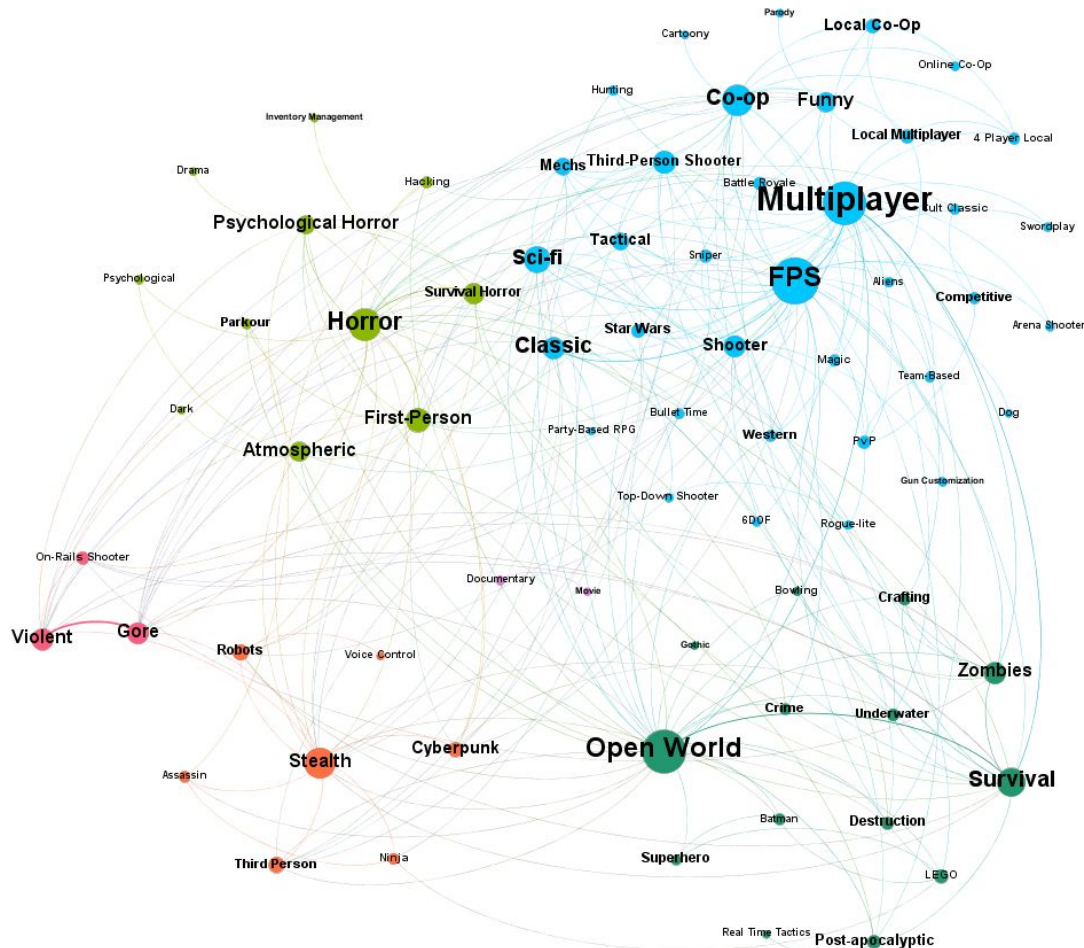
En los nodos azules podemos ver a las novelas visuales, que se separan en dos tipos. Aquellas con una historia rica y que están muy influidas por los tags morados, y aquellas mucho más

7,26

30,932

Esta es una comunidad más grande, con 73 nodos y un grado medio de 7, pero mucho más clara en concepto, con juegos multijugador en los cuáles la tendencia suelen ser disparos en primera persona y mundo abierto.

Al ser una comunidad más grande podemos también desglosarla en subcomunidades:



Subcomunidades de la comunidad Multijugador (Gephi)

Si nos fijamos en la comunidad hay juegos claros que juntan todos los géneros, como pueden ser *Call of duty*, juego multijugador de disparos en primera persona (“FPS” y “Shooter”) con modos competitivos (“Competitive”) o más cooperativos (“Co-Op”); o el *Battlefront 2*, juego de disparos ambientado en el universo de Star Wars y con disparos en tercera persona (“Third Person Shooter”). Muchos de estos juegos cuentan con mundos amplios (“Open World”).

También es habitual entre estos juegos o que aparezcan zombies, con un cierto tono de terror. Un ejemplo puede ser el juego *Left 4 Dead 2*, un multijugador, cooperativo de disparos y zombies.

En definitiva todos los juegos aquí recogidos son el claro exponente de lo que denominaríamos un juego de acción, un género que se usa de una forma mucho más amplia en la red completa.

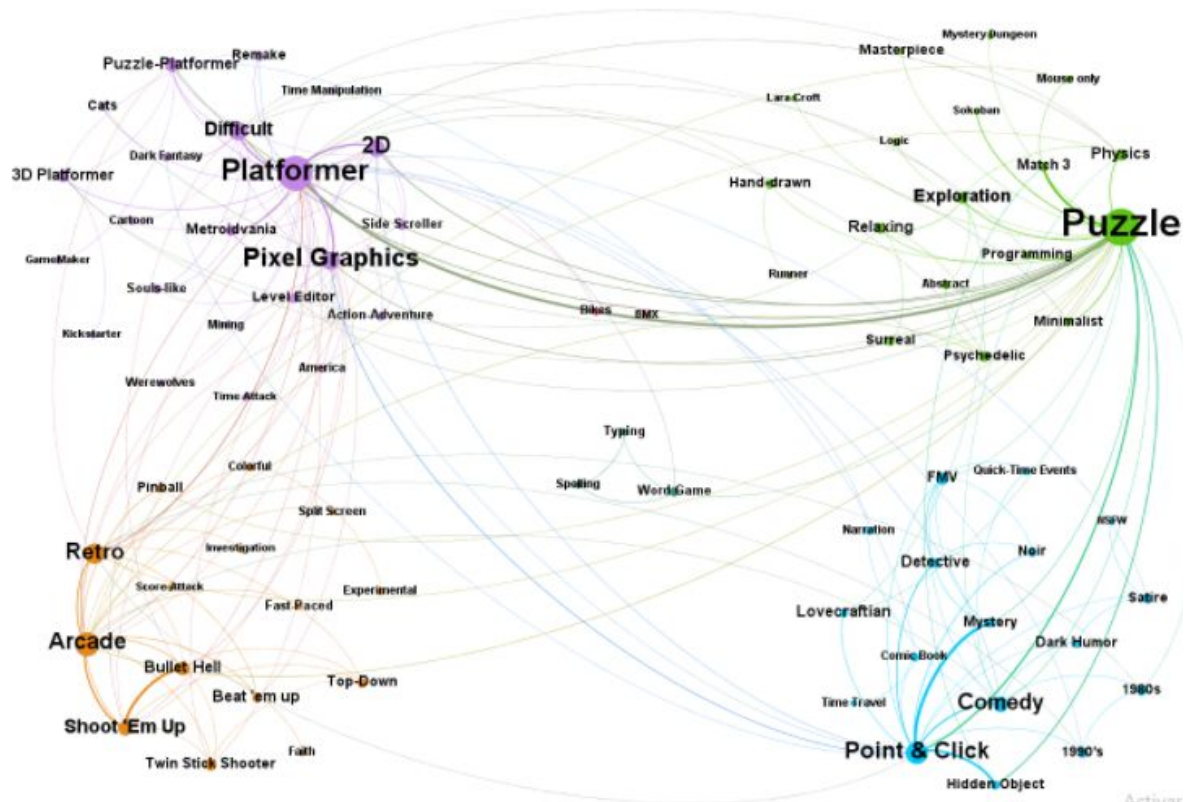
- Comunidad puzzle y plataformas



Esta comunidad tiene dos géneros principales, que son el “Puzzle” (juegos enfocados en resolver problemas de una forma lógica) y las “Plataformas” (Juegos basados en moverse por el escenario a través de obstáculos y plataformas), géneros muy comunes en muchos juegos como *Portal* o *Braid*, de hecho podemos ver que existe “Puzzle-Platformer” como un solo tag. A su vez hay una gran cantidad de géneros más pequeños relacionados con ellos de forma independiente.

Esto se ve reflejado en el grado medio, que a pesar de tener 77 nodos la comunidad, apenas llega al 4,7. Como hemos dicho hay unos géneros principales que unen toda la red y que a su vez arrastran una serie de subgéneros muy conectados al tag principal pero muy poco al resto.

Esto permite separar la comunidad de una forma muy clara.



Subcomunidades de la comunidad Puzzle y Plataformas (Gephi)

Por una parte está el tag “Puzzle”, con una variante muy clara como es el tag “Point and Click”, donde la mecánica básica del juego consiste en como dice el nombre apuntar a algún lugar de la escena y clicar, una mecánica muy simple y muy relacionada con la resolución de puzzles o las aventuras gráficas. Estos juegos tienen todo tipo de estéticas, y así lo demuestran otros tags como “Comedy”, “Lovecraftian”, “Comic book” o “Hidden object” entre otros.

La otra parte de los puzzles es mucho menos clara y depende más de la mecánica, aunque no hay ninguna que de tanto juego en la comunidad como point and click. Podemos encontrar tags como “Programming”, “Exploration” o “Sokoban”, ejemplos claros de tags a los que sería muy difíciles de ver separados de la etiqueta puzzle.

El otro gran género de la comunidad son las plataformas, en su mayoría son juegos en 2D que suelen tener un arte más “pixelado” (“Pixel Graphics”).

A parte también tenemos ciertos juegos de este género en 3D.

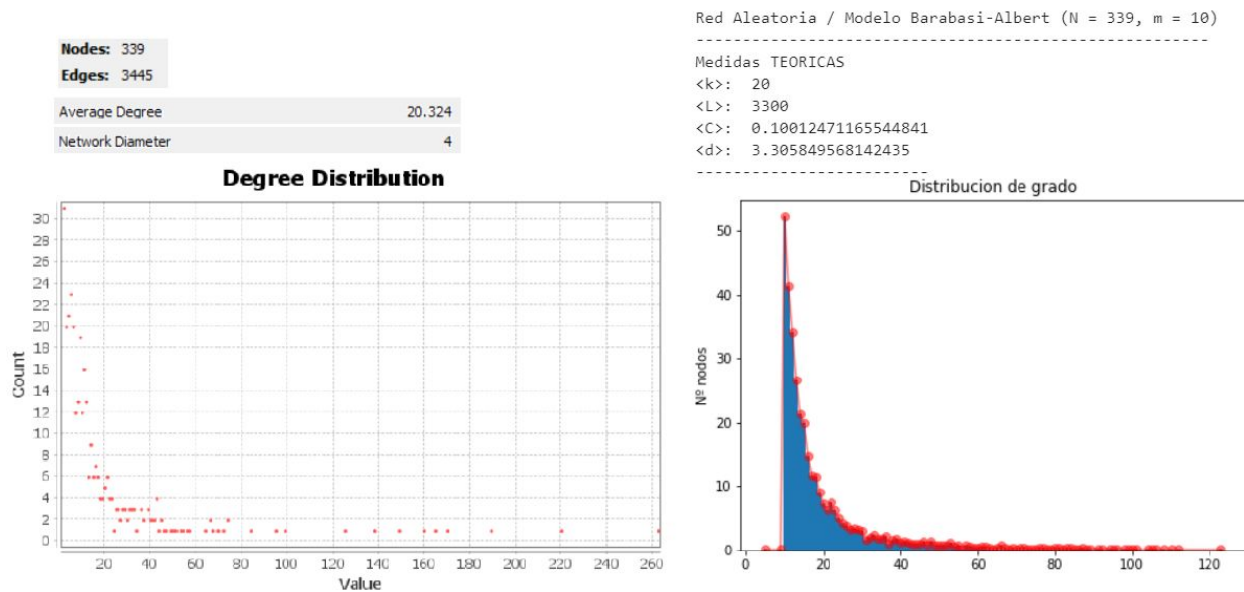
También es interesante destacar el tag “GameMaker”, ya que una gran cantidad de estos juegos de plataformas dejan al usuario desarrollar sus propios niveles.

El Arcade, quizás sería el tercer gran género de la comunidad, muy relacionado con Retro debido a los juegos procedentes de las antiguas máquinas recreativas. También hay

5. Conclusiones finales

Con este estudio hemos podido ver como la gran variedad de tags que hay en Steam no se relacionan de una manera arbitraria, claramente hay ciertos tipos de juegos más abundantes que otros. Por esto, la red formada tiene una estructura bastante centralizada, principalmente influenciada por los tags Indie, Acción, Aventura, Casual, Estrategia y Simulación.

Este comportamiento se puede relacionar con el modelo de redes de Barabási-Albert, que establece que los nodos añadidos a una red tenderán a conectarse a los nodos más relevantes de esta. Podemos ver una estructura similar en una red Barabási-Albert con $m = 10$:



Distrib. grado red tags de Steam (izquierda) / Distrib. grado red BA con $m = 10$ (derecha)

Que esta red cumpla el principio del modelo de Barabási-Albert es debido a que los tags más minoritarios tienen una mayor probabilidad de conectarse con los más relevantes ya que estos últimos definen conceptos más generales en los que se pueden encajar muchos más juegos. Además, a la hora de crear nuevos juegos para Steam, es probable que los desarrolladores busquen crear juegos cercanos a los tags más populares para tener un alcance mayor (p.e. Si los juegos de Acción no fuesen populares, los estudios evitarían este género y este tag no sería tan común).

Por otro lado, hemos podido profundizar en las relaciones entre todos los tags viendo que comunidades se forman.

Un conjunto de tags puede definir mucho mejor a un juego que el uso de un solo tag, por eso cada comunidad representa con mucha más precisión los distintos tipos de videojuegos. Ignorando los tags más grandes de la red hemos visto cómo las comunidades más grandes son, por un lado los juegos de puzzles y los juegos de plataformas, por otro lado los juegos de estrategia y simulación; y finalmente los juegos multijugador, generalmente de disparos. Dentro de estas comunidades hemos podido ver otros subgéneros más específicos.

En resumen. Steam es una plataforma con juegos de una gran variedad de géneros. Hay tags que suelen aparecer juntos mucho más a menudo, formando comunidades claramente definidas. Además todas estas comunidades están muy relacionadas entre sí debido a la existencia de unos pocos tags muy amplios y por tanto con mucha relación con toda la red.

Anexo: Red de juegos de Steam

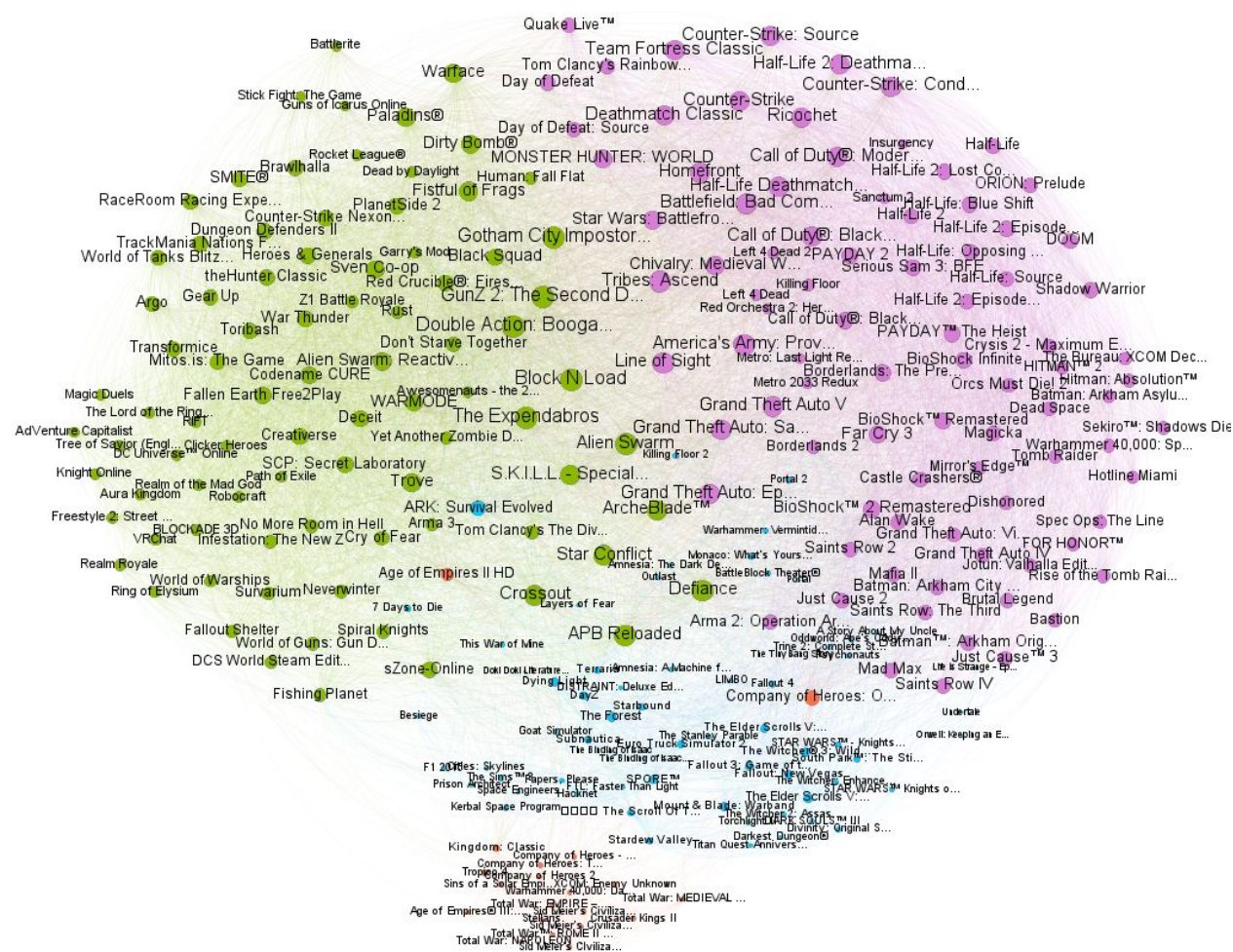
Los datos utilizados para el estudio anterior nos permiten generar otra proyección distinta de la red, en la que los nodos serían los juegos y las aristas serían los tags que tienen en común dos juegos distintos.

El análisis de esta red podría servir para ver que juegos son más influyentes en Steam, a que tags corresponden o estudiar comunidades de juegos. Sin embargo debido a problemas relacionados con el gran tamaño de la red no se ha podido llevar a cabo el análisis ya que las herramientas de visualización de grafos no eran capaces de tratar con archivos tan pesados.

Primero se intentó generar una red con todos los juegos unidos por sus tags en común, el resultado fue una red de más de 20 millones de aristas la cuál Gephi era incapaz de abrir.

Tras esto se filtró el número de juegos para tener en cuenta sólo aquellos con un número elevado de ventas (obteniendo 8475 juegos de los 27075 iniciales), ya que son los que más podían representar la realidad de Steam, sin embargo el resultado también fue una red intratable de 14 millones de aristas.

Filtrando aún más por número de ventas y quedándonos solo con los juegos que más ventas tienen obtenemos una red bastante más reducida y manejable (206 juegos y 10799 aristas)

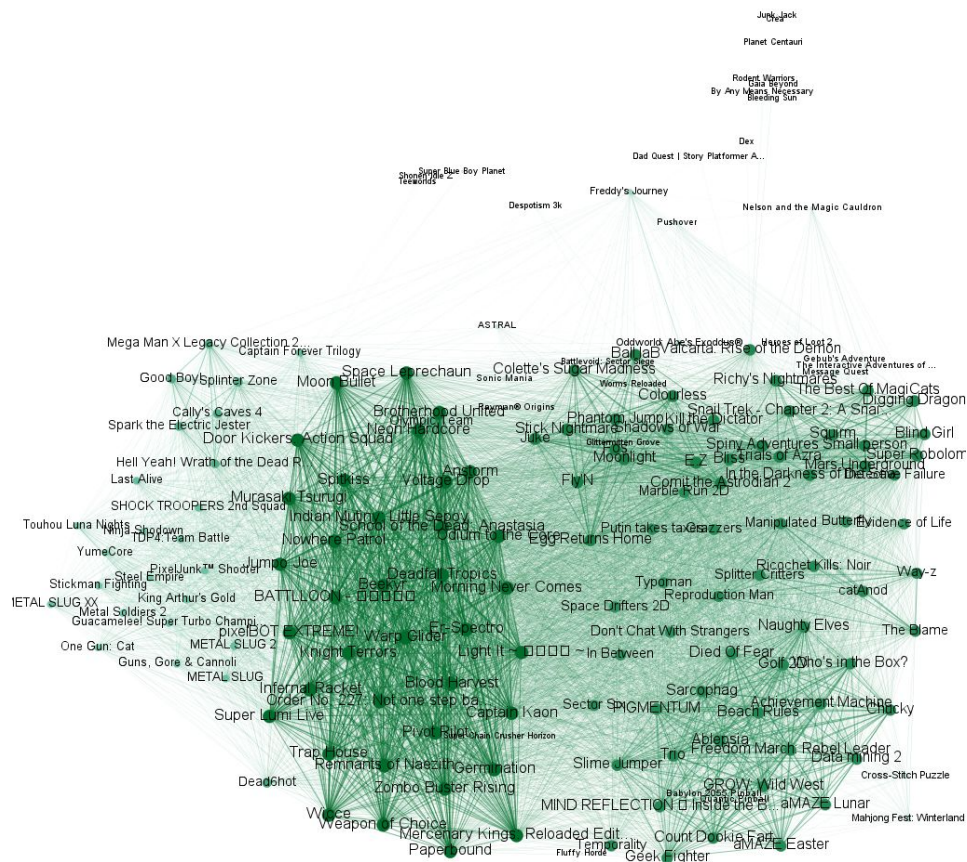


Red de juegos (juegos con más de 2.000.000 de ventas)

Viendo la red resultante podemos ver como destacan algunos juegos más que otros e incluso podemos sacar varias comunidades, la verde englobaría juegos indie, la morada juegos de acción, la azul engloba juegos RPG y la naranja juegos de estrategia. Todo esto encaja con los géneros mayoritarios vistos en el estudio de la red de tags. Sin embargo tampoco podemos sacar muchas más conclusiones ya que estando tan filtrados los datos se aleja bastante de las estadísticas reales de la red completa. Podemos ver este sesgo en el tamaño de los nodos como el grado de los juegos es bastante similar dentro de sus comunidades cuando en la red real debería haber juegos más relevantes dentro de cada comunidad, esto se debe a la falta de enlaces que estos juegos tendrían con otros juegos con menos ventas.

Otro enfoque fue intentar generar redes únicamente de los juegos pertenecientes a cierto tag unidos entre sí por el resto de tags que poseen en común. Aun así, las redes de los tags más relevantes (Indie, Acción, Aventura, etc..) seguían teniendo millones de aristas y un tamaño intratable. También se generaron redes de tags más pequeños (como *2D*, *Sports* o *VR*), pero el análisis de estos grafos tan reducidos resulta bastante irrelevante para el estudio de la red general.

Podemos ver un ejemplo de la red generada por los juegos con tag *2D*:



Ejemplo: Red de juegos con tag '2D'

En conclusión, debido a las limitaciones de las herramientas utilizadas para tratar con grafos demasiado grandes y la información relevante que se pierde al filtrar tanto los datos, no hemos podido realizar un análisis con detalle de la red de juegos.