

Estudio de una red dinámica de adolescentes. Análisis del consumo de sustancias nocivas.

Di Filippo Juan, Catoni Josefina, Yalovetzky Romina

Diciembre 2018

Resumen

En el presente trabajo se analizó la evolución de un grupo de adolescentes a lo largo de tres años. En particular, se hizo foco en el consumo de sustancias nocivas: alcohol, tabaco y cannabis, y cómo influyen estos consumos en la forma de relacionarse entre ellos. Se estudió cuantitativamente la evolución temporal del consumo de las sustancias mencionadas y los mecanismos de selección y asimilación en la red. Finalmente, se analizó la correlación que existe entre los distintos grados de consumos y el sexo de los adolescentes.

1. El dataset utilizado

Se realizó una encuesta entre un grupo de adolescentes de una secundaria del oeste de Escocia que se la considera representativa de la región en términos de la clase social de los estudiantes. La misma encuesta se repitió en 3 años, entre 1995 y 1997, cuando los alumnos tenían entre 13 y 15 años respectivamente. En las encuestas participaron 160 alumnos, de los cuales 134 estuvieron presentes en las tres instancias de medición.[1]

Los estudiantes dieron a conocer su sexo, nombraron hasta 6 amigos y valorizaron su consumo de las siguientes sustancias: Alcohol, tabaco y cannabis otorgando una numeración. Para alcohol: 1 (nunca), 2 (una o dos veces por año), 3 (una vez por mes), 4 (una vez por semana) y 5 (más de una vez por semana). Para tabaco: 1 (nunca), 2 (ocasional), 3 (regular: mas de una vez por semana). Para cannabis: 1 (nunca), 2 (una vez), 3 (ocasional) y 4 (regular).

Construimos una red de amistades donde los nodos representan a los estudiantes y el enlace se da sí hay una relación de amistad. Como cada estudiante podía designar a sus amigos se trata de una red dirigida. Luego, sobre la misma, tendremos para caracterizar los parámetros indegree, outdegree y mutual degree. En Tabla 1 vemos algunos parámetros característicos de la red, mientras que en la figura 1 se visualiza la red el tercer año. El color corresponde al sexo de los nodos, y el grosor del enlace denota el peso en la amistad.

A su vez tomamos como un campo escalar sobre la red la distribución del consumo para cada sustancia de forma que podemos visualizar la evolución del consumo.

| Nodos: 134 | 1er año | 2do año | 3er año |
|--------------|---------|---------|---------|
| Enlaces | 465 | 458 | 490 |
| Max indegree | 11 | 10 | 10 |

Cuadro 1: Caracterización de la red de amistades a lo largo de los 3 años estudiados.

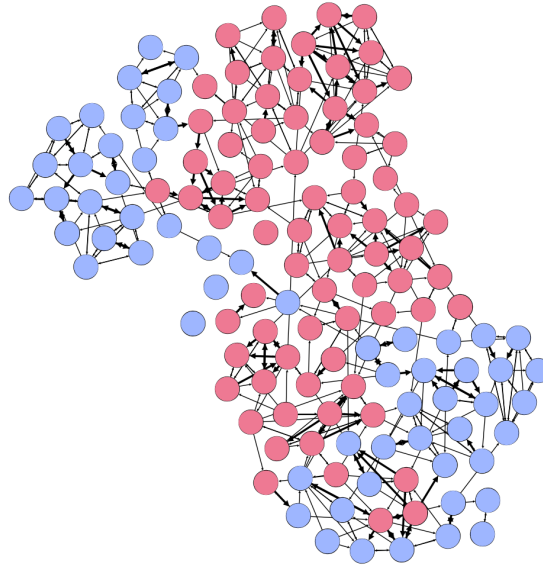


Figura 1: Red de adolescentes en el año 1997. El color denota el sexo de las personas (rojo para mujeres, azul para hombres). El peso de la amistad se representa en el grosor del enlace.

2. Planteo de la problemática e interrogantes

En primer lugar lo que hicimos fue graficar a la red de amistades para cada uno de los consumos tomando a la distribución de grado de consumo como un campo escalar que pesa a los nodos según el nivel de consumo y vimos su evolución en el tiempo. En particular vemos en Fig. 2 para el consumo de tabaco donde los nodos aumentan en tamaño y se oscurecen en color a medida que aumenta la cantidad que consumen. Se observa que por un lado la cantidad de consumo aumenta, mayor cantidad de nodos verde oscuro y de mayor tamaño y que también varios nodos que no consumían en el año previo (color negro) comienzan a consumir. Ésta tendencia se ve para el consumo de las 3 sustancias. Entonces, planteamos la problemática: Se da un incremento de cantidad de consumo e fracción de personas que comienzan a consumir a lo largo de los años en éste grupo de adolescentes.

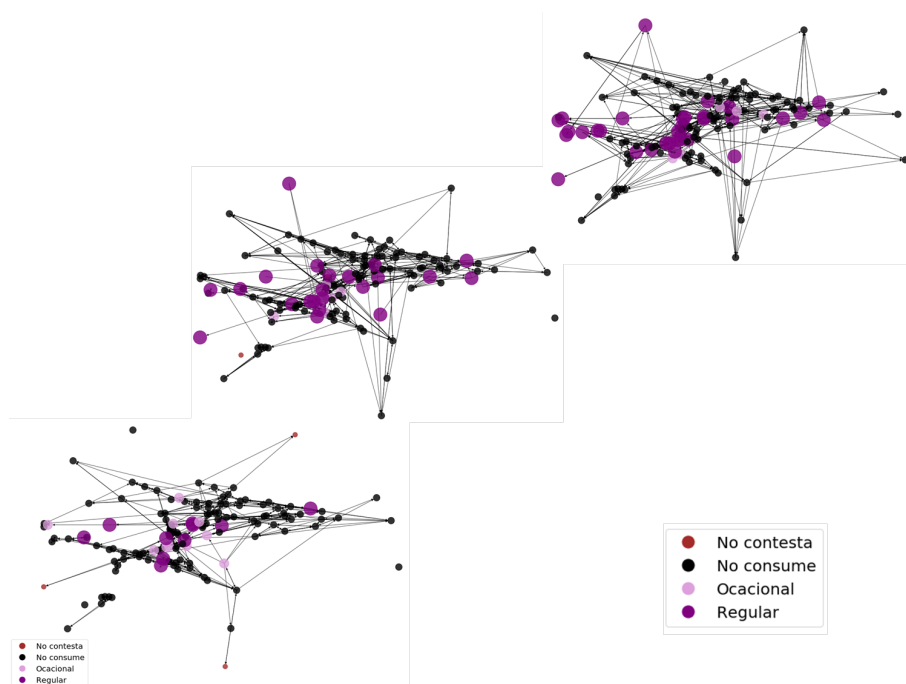


Figura 2: Evolución de la red de amistades para el consumo de tabaco. El tamaño y el color de los nodos está determinado por la cantidad de consumo. Los de color negro no consumen.

Con el objetivo de visualizar la problemática vemos en particular el consumo de tabaco. El primer interrogante que nos planteamos fue si existe una relación entre la red de amistades y el hábito de fumar entre los nodos de la red. Como mostramos se trata de una red dirigida entonces tenemos definido in-degree, outdegree y mutua o recíprocal degree. Estudiamos éstas distribuciones de grados considerando como nodos target fumadores, es decir que vimos las distribuciones de grados que presentan sólo los nodos que consumen tabaco. En Fig. 3 comparamos para cada año la distribución del indegree (nodos que eligen a los target fumadores como amigos) y el outdegree que es como los fumadores eligen amigos. Es clara y permanente en los 3 años la tendencia de que los fumadores de tabaco son menos elegidos como amigos en comparación con los amigos que ellos eligen.

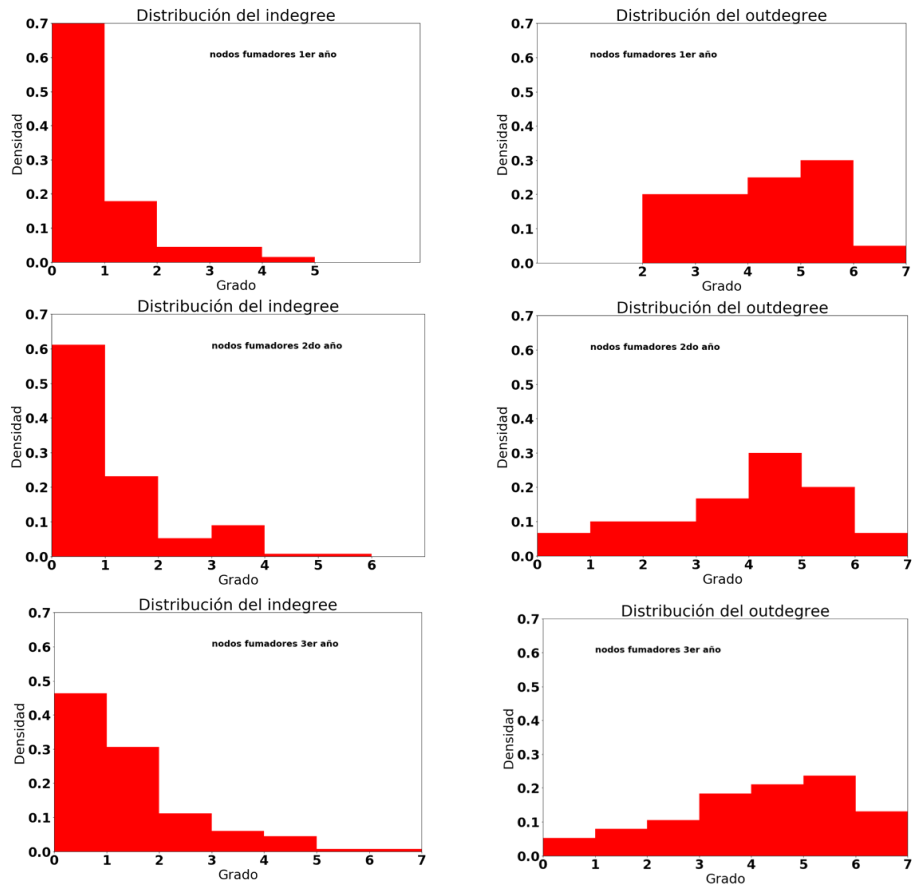


Figura 3: Distribuciones de indegree (izquierda) y outdegree (derecha) para nodos fumadores para los 3 años estudiados.

Por otro lado, estudiamos la distribución del mutua degree para nodos target consumidores de sustancia y su evolución en el tiempo. En Fig. 4 vemos los resultados para el caso particular de nodos fumadores de tabaco para cada uno de los años. En rojo se grafica la densidad para nodos target fumadores y amigos mutuos fumadores y en azul para nodos target fumadores y amigos mutuos no fumadores. Un observable estudiado es el de la reciprocidad que se define como ec. 1 para nodos target con característica p (fumar tabaco) y considerando todos sus vecinos (fumadores como no fumadores). Los resultados se muestran en Tabla. 1. De esa fracción de enlaces se calcula la fracción de enlaces con vecinos no fumadores (2da fila) y fumadores (3ra fila) para los 3 años.

$$\text{reciprocidad}_p = \frac{\sum_i^N \text{enlaces mutuos}_i}{\text{enlaces total de la red}} \quad (1)$$

donde i representa a los nodos target. Se calculan los enlaces mutuos entre el nodo target con característica p y todos sus vecinos y se suma sobre todos los nodos target N .

Aunque no podemos extraer conclusiones de comparación entre el mutual degree fumador-fumador y el mutual fumador-no fumador porque la densidad no está normalizada con respecto a los nodos fumadores sino que con el total, sí podemos afirmar que de un año al otro se ve que aparecen densidades no nulas para valores de grado más alto en el caso del mutual fumador-fumador, mientras que para el mutual degree fumador-no fumador se da una tendencia inversa que es la aparición de densidades no nulas para grados más bajos.

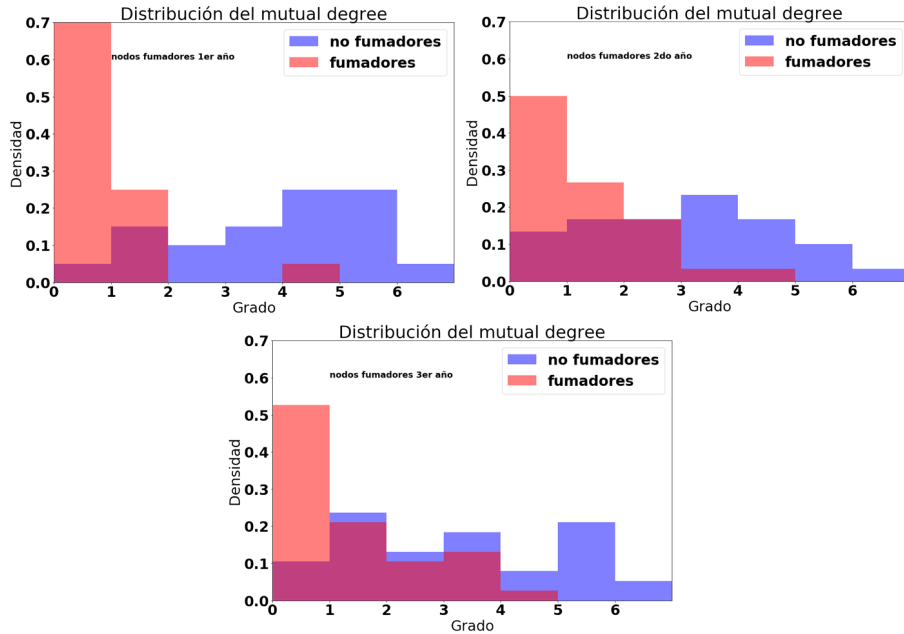


Figura 4: Distribuciones de indegree (izquierda) y outdegree (derecha) para nodos fumadores para los 3 años estudiados.

| Nodos fumadores | 1er | 2do | 3ro |
|------------------------------------|------|------|------|
| Reciprocidad | 0.28 | 0.38 | 0.53 |
| Fracción sobre mutuos que no fuman | 0.89 | 0.75 | 0.75 |
| Fracción sobre mutuos que fuman | 0.11 | 0.25 | 0.25 |

Cuadro 2: Valores de reciprocidad, fracción de enlaces mutuos fumador-no fumador sobre enlaces total y fracción de enlaces mutuos fumador-fumador sobre enlaces total para los tres años.

Esto se ve en Tabla.1 con el aumento de la fracción sobre mutuos que fuman y el descenso de la fracción sobre mutuos que no fuman. Observamos entonces que con el paso de los años los fumadores tienden a generar amistades mutuas con fumadores y disminuir sus amistades con no fumadores.

En las siguientes secciones vamos a cuantificar la problemática presentada y responderemos de forma cuantitativa a los interrogantes planteados definiendo observables:

- 1. ¿El consumo influye en cómo se dan las amistades?: **Observables alter y ego.**
- 2. ¿Amistades mutuas se dan más entre personas con mismo hábito de consumo?. **Homofilia e influencia.**
- 3. ¿Es el consumo de sustancias lo que influye sobre la creación de grupos de amistad o es el grupo de amigos lo que impacta sobre el consumo?. **Homofilia en formación de comunidades.**
- 4. ¿Qué tanto se incrementa el consumo?. **tendencia.**

3. Análisis de la red

Uno de los aspectos fundamentales a cuantificar en la red fue la evolución temporal del consumo de sustancias nocivas. Para ello, definimos un observable, al cual llamamos 'tendencia' (t), que depende del grado de consumo w de cada nodo y se calcula en un periodo comprendido entre dos años, el posterior ($año+$) y el anterior ($año-$):

$$t = \frac{\sum_{i=1} (w_i^{año+} - w_i^{año-})^2 \cdot \text{signo}[w_i^{año+} - w_i^{año-}]}{N}$$

donde N es el total de nodos de la red. Por como esta definida, una tendencia positiva indicaría un aumento (en promedio) en el consumo de una determinada sustancia, mientras que si es de signo negativo, se estaría en presencia de una disminución del mismo.

Se calculo la tendencia para dos periodos: del primer año al segundo y del segundo al tercero. Además, se generó un modelo nulo a partir de dejar fijos los consumos del año anterior y generar una redistribución aleatoria de los consumos del año posterior, manteniendo fija la distribución del grado de consumo de la sustancia a tratar.

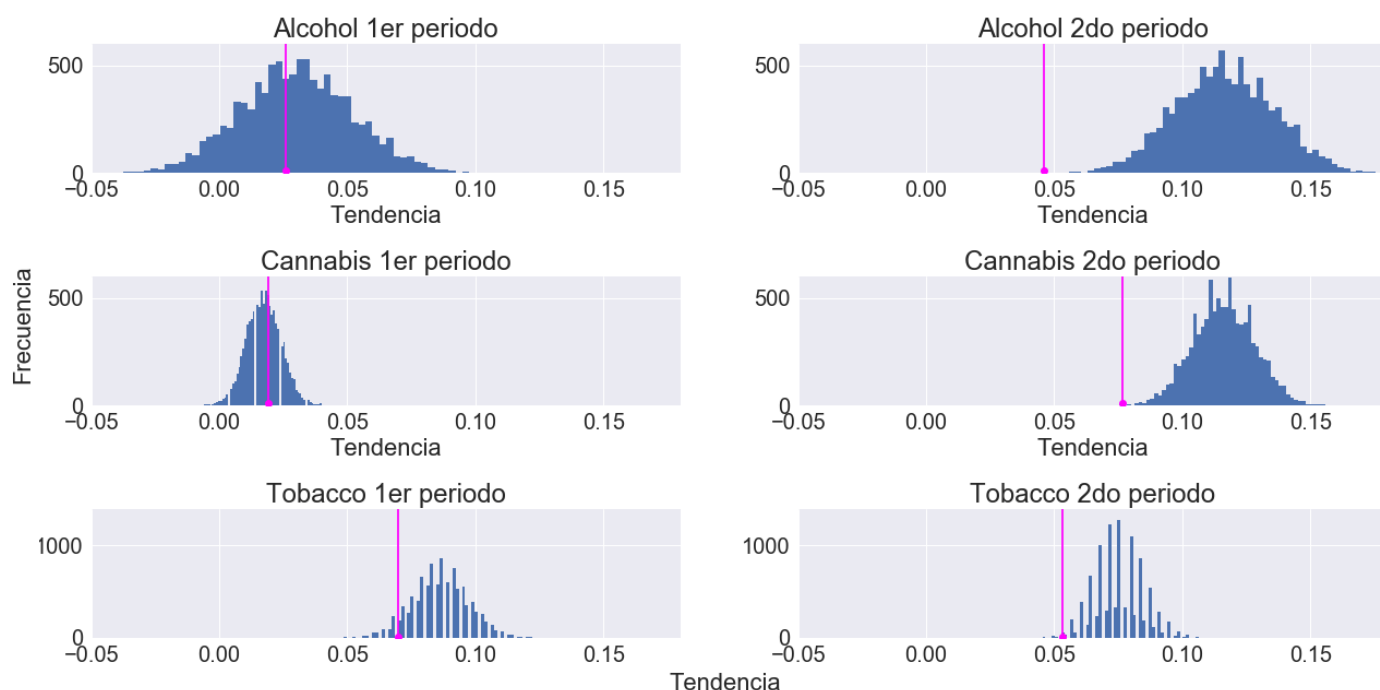


Figura 5: Tendencias calculadas para alcohol,canabis y tabaco, para los dos periodos mencionados. Las líneas continuas indican el valor real de tendencia

En primer lugar, observamos que todas las tendencias fueron positivas, algo esperado a partir de ver la evolución gráfica de la red (Fig. 2), y que para el caso del alcohol y el cannabis, la tendencia del segundo período resulto mayor a la del primero. Sin embargo, es menester denotar que dichas tendencias resultaron menores a las esperadas por azar (dada la distribución fija de consumos). Creemos que esto puede deberse a la presencia del exponente cuadrático, que genera una brecha mayor al calcular la diferencia entre dos grados de consumo.

Para tener una primera noción de si se dan mecanismos de selección (homofilia) o asimilación (influencia) en las amistades, se analizó como se relacionan los nodos según la similitud entre su grado de consumo. Tal como se ilustra en la figura 6, definimos que la red presenta homofilia cuando pares de nodos inicialmente similares pero desconectados tienden a unirse. Definimos que las amistades tiene alta influencia en el consumo cuando los nodos comienzan a parecerse a quienes elijen como amigos.[2]

Es importante destacar que para próximos análisis, no se tuvo en cuenta enlaces que conectaran nodos que no hubiesen respondido las encuestas. En la tabla 3 se tiene cuántos nodos no respondieron la encuesta para cada sustancia en al menos uno de los años comprendidos en cada período, y en consecuencia, qué porcentaje de

los enlaces totales en estos períodos tuvieron que despreciarse ya que alguno de estos nodos estaba involucrado en ellos. Se observa que para el consumo de alcohol se perdió información de gran parte de la red.

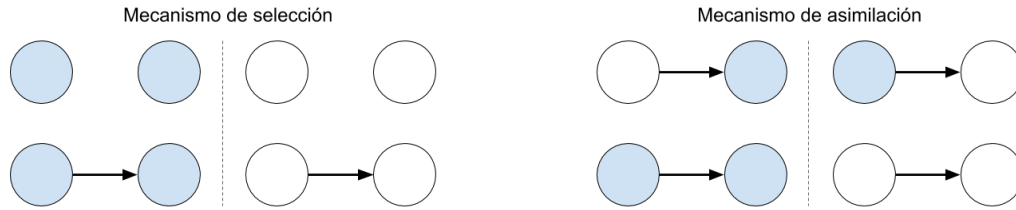


Figura 6: A la izquierda definición de enlace debido a mecanismo de selección u homofilia. A la derecha definición de enlace de influencia en atributo del nodo source que pasa a asemejarse a nodo target.

| | 1er período | | 2do período | |
|----------|-------------|---------|-------------|---------|
| | nodos | enlaces | nodos | enlaces |
| Alcohol | 30 | 37,9 % | 24 | 36,9 % |
| Cannabis | 5 | 3,1 % | 3 | 2,5 % |
| Tobacco | 5 | 3,1 % | 2 | 0,4 % |

Cuadro 3: Cantidad de nodos que no respondieron la encuesta de consumo en cada período, y porcentaje de enlaces que involucraba alguno de estos nodos.

En la figura 7 se tiene para ambos períodos (del primero al segundo año y del segundo al tercero) y para los tres consumos, cuántos enlaces se crearon, perdieron o mantuvieron, en función de cuán similares eran los nodos al principio del período. En todos los casos, se observa que la mayoría de los nuevos enlaces que se crean de un año a otro se da para nodos de alta similitud, lo que podría indicar un carácter de homofilia, es decir que nodos que inicialmente eran parecidos pero no estaban conectados, tienden a conectarse. Sin embargo, se busca una manera de poder cuantificar este comportamiento.

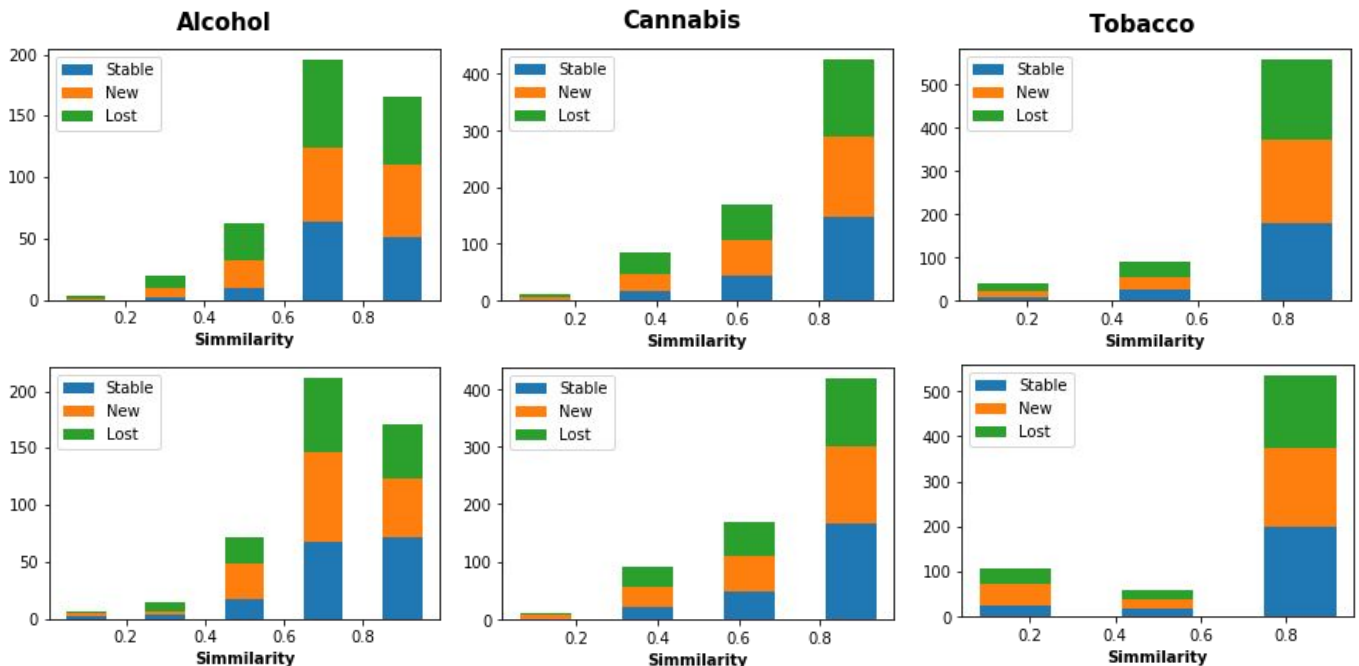


Figura 7: Enlaces nuevos, perdidos o constantes de un año a otro para cada período, según la similitud de los nodos en cuando al grado de consumo de cada sustancia al inicio de cada período.

A continuación, con el objetivo de tener una manera más concreta de cuantificar este comportamiento, se analizaron por separado los casos de enlaces nuevos y estables. Se definió el parámetro de homofilia (ver

figura 6) como la fracción de los enlaces nuevos que se dieran para nodos que mantuvieran la misma alta similaridad de un año al otro. De manera análoga, se definió el parámetro de influencia, como la fracción de enlaces estables que se diera entre nodos cuya similaridad aumenta de un año al otro, en particular el caso en el que el nodo 'target' mantiene constante su consumo mientras que el nodo 'source' se asemeja más a éste, tal como se muestra en la figura 6.

En la figura 8 se tiene la relación de similaridad entre nodos para enlaces que se mantienen constantes o que aparecen de un año a otro. Si la similaridad aumenta, decimos que los nodos se acercan. Si la similaridad disminuye, decimos que se alejan. En el caso de que se mantenga constante, se separan en dos casos: Se mantienen alejados (poco similares) o se mantienen cerca (altamente similares). El mismo mecanismo se repitió para ambos períodos para las tres sustancias.

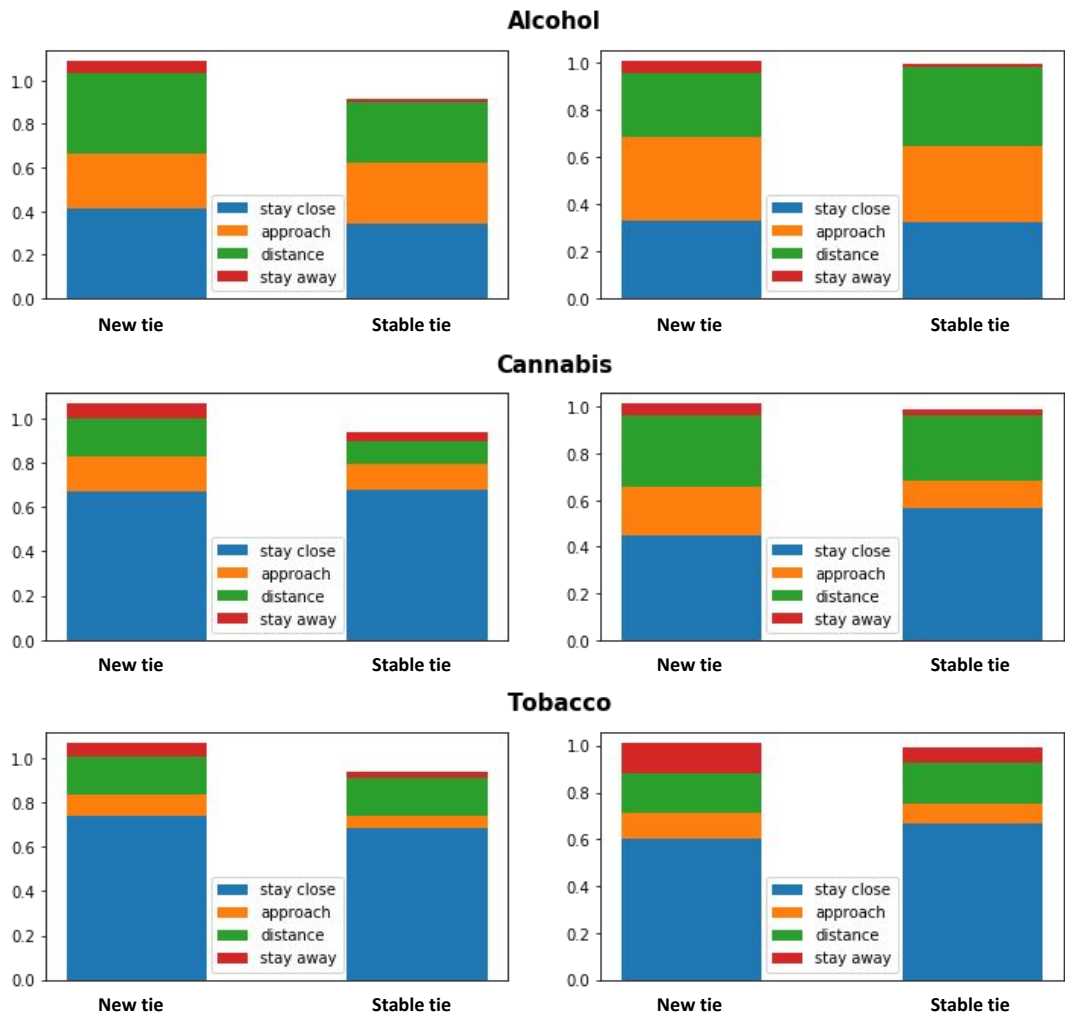


Figura 8: Para enlaces nuevos y constantes de un año a otro para cada período, se grafica el porcentaje de pares de nodos que se acercan en semejanza, que se alejan, que se mantienen altamente similares, o alejados para cada sustancia.

Observando los resultados de la figura 8 y a partir de la definición del parámetro de homofilia e influencia, se aprecia que el consumo de cannabis y tabaco presentan la mayor tasa de homofilia a la hora de creación de nuevas amistades, mientras que la influencia resulta ser mayor para el consumo de alcohol. Para verificar esta hipótesis, se compararon los resultados obtenidos para estos parámetros con una distribución al azar. La hipótesis nula constó de distribuir de manera aleatoria los atributos sobre los nodos para el segundo año de cada período, conservando la cantidad total de personas que consumiesen en cada grado y recalculando cuánto darían estos parámetros. Es necesario resaltar que la redistribución de consumos se hizo sobre la red original, es decir antes de quitar los nodos que no hubiesen respondido la encuesta, por lo que en cada iteración fueron distintos enlaces los que se despreciaron por falta de información.

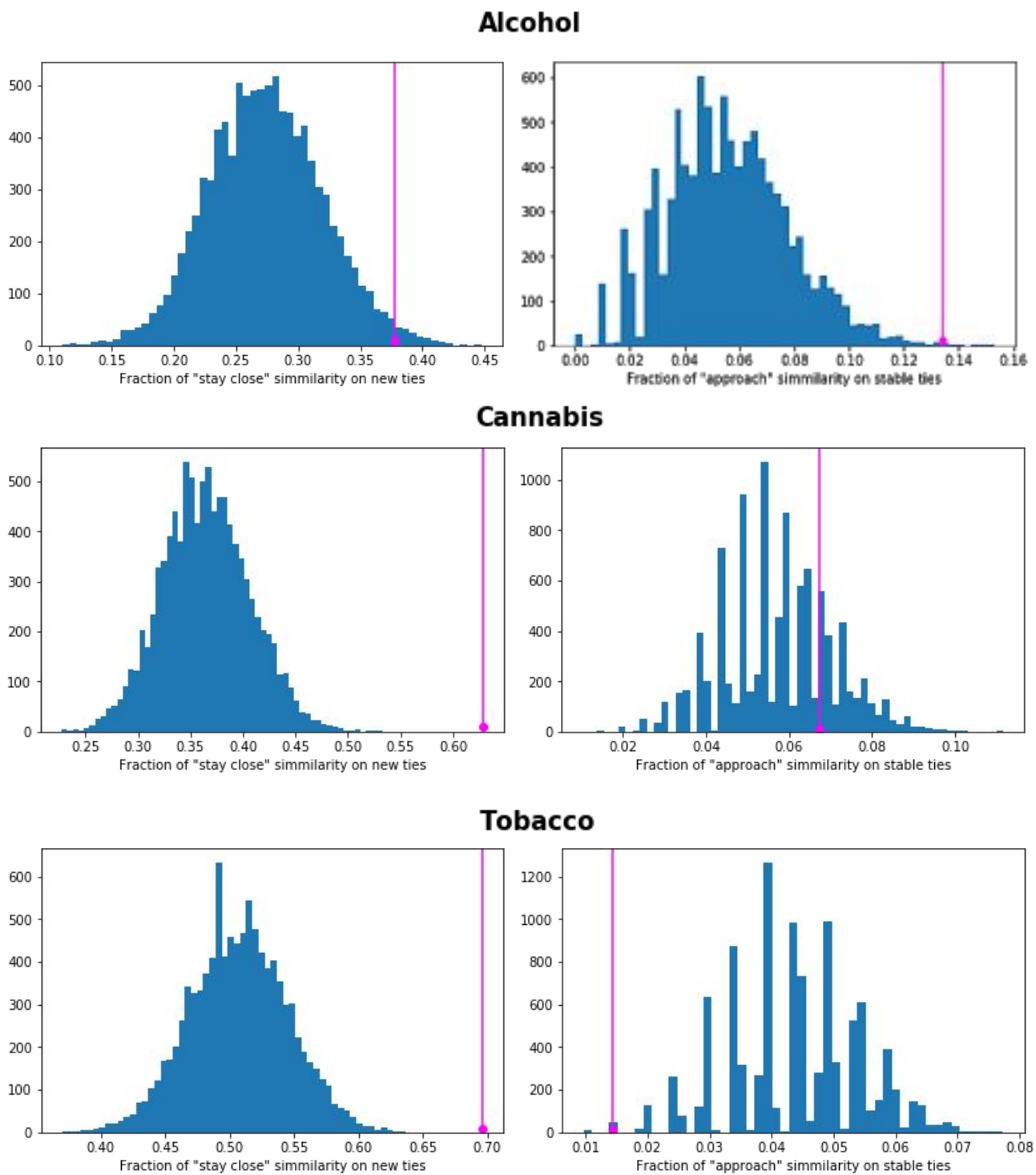


Figura 9: Para el consumo de cada sustancia en el primer período se tiene a la izquierda el parámetro de homofilia calculado para la red real en rosa y la distribución nula en azul. A la izquierda, el parámetro de influencia.

En la figura 9 se ilustran los resultados para el primer período. Para cada consumo, se tiene a la izquierda el parámetro de homofilia de la red y la distribución nula, y a la derecha el parámetro de influencia. Para el segundo período se obtuvieron resultados similares. En la tabla 4 se registraron los valores de dichos parámetros junto con el p-valor calculado a partir de la comparación con la hipótesis nula.

Tal como se había intuído, las relaciones de amistad tienen el mayor grado de homofilia cuando se trata de consumo de cannabis y tabaco, mientras que la mayor influencia entre pares se da en el consumo de alcohol en el primer período y cannabis en el segundo. Este resultado es coherente con lo que uno visualiza en la red, tal como se puede ver en la figura 10. Ésta, muestra como cambia el consumo de cannabis en el segundo período, es decir del segundo al tercer año. Se puede notar como nodos que inicialmente consumían, se rodean de más

consumidores al año siguiente.

| | | 1er período | | 2do período | |
|----------|------------|-------------|---------|-------------|---------|
| | | valor | p-valor | valor | p-valor |
| Alcohol | homofilia | 0,377 | 0,016 | 0,329 | 0,21 |
| | influencia | 0,133 | 0,0005 | 0,110 | 0,19 |
| Cannabis | homofilia | 0,629 | <0,0001 | 0,444 | <0,0001 |
| | influencia | 0,067 | 0,24 | 0,058 | 0,019 |
| Tobacco | homofilia | 0,696 | <0,0001 | 0,597 | <0,0001 |
| | influencia | 0,014 | 0,0016 | 0,043 | 0,0001 |

Cuadro 4: Valores de los parámetros de homofilia e influencia calculados y p-valores con respecto a la distribución nula para las tres sustancias en ambos períodos.

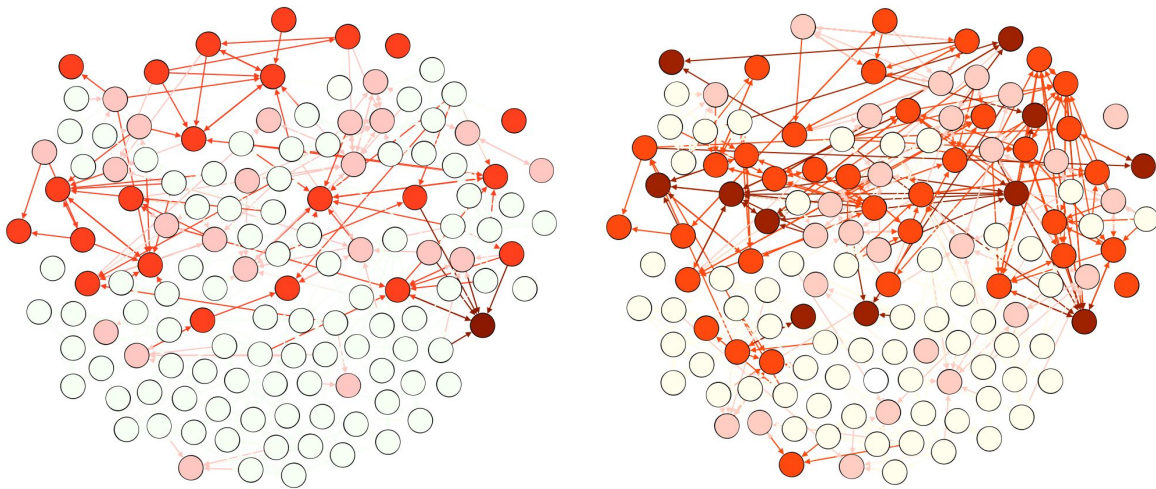


Figura 10: A la izquierda y derecha se tiene la red de adolescentes en el segundo y tercer año respectivamente. El color del nodo corresponde al grado de consumo de cannabis. En color blanco se tienen los nodos que no consumen o no contestaron la encuesta. Los nodos de color presentan mayor consumo cuanto más oscuro su color.

Se puede visualizar cómo este mecanismo de selección (homofilia) se da en la formación de comunidades. En la figura 11 se tiene la red en el último año. El color del nodo corresponde a la comunidad a la que pertenece (éstas fueron calculadas mediante el método de Louvain) mientras que el tamaño denota el grado de consumo de los mismos. Se observa que tanto para cannabis como para tabaco, el tamaño de los nodos dentro de cada comunidad es similar.

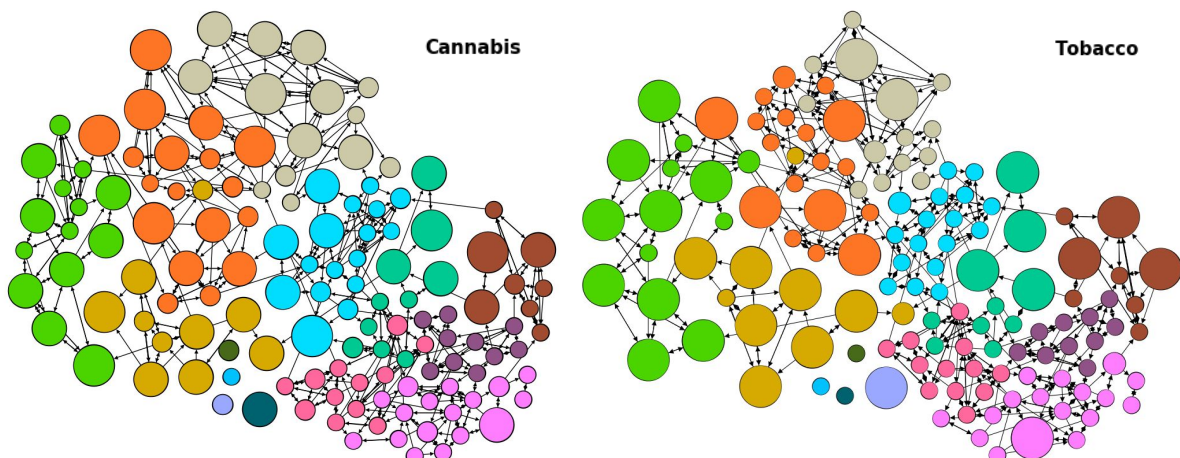


Figura 11: Red de adolescentes en el tercer año. El color de los nodos según comunidad de pertenencia. El tamaño del nodo según grado de consumo de cannabis y tabaco respectivamente.

Otro de los aspectos analizados previamente fue la relación entre como los estudiantes elijen a sus amigos y que tan atractivos son ellos mismos para ser elejidos como amigos por el resto (Fig. 3). Para cuantificar este comportamiento nos limitamos a calcular la diferencia entre el in-degree y el out-degree medios para consumidores y no consumidores:

$$alter/ego = \frac{\sum_i InDegree_i}{M} - \frac{\sum_i OutDegree_i}{M}$$

donde M representa el total de nodos consumidores/no consumidores. En la tabla 5 se muestran los resultados obtenidos para cada año, para las tres sustancias nocivas.

| | 1er año | | 2do año | | 3er año | |
|----------|---------|-------|---------|-------|---------|-------|
| | X | ✓ | X | ✓ | X | ✓ |
| Alcohol | 1.13 | -0.73 | 0.01 | -0.65 | -4.0 | -0.38 |
| Cannabis | -0.30 | 0.59 | -0.16 | 0.13 | -0.45 | 0.05 |
| Tobacco | -0.04 | -0.3 | -0.01 | -0.33 | -0.01 | -0.26 |

Cuadro 5: Valores de alter/ego calculados para las distintas sustancias nocivas y los tres años de consumo. Las X representan los no consumidores y los ✓ los consumidores.

Para el caso de los consumidores de tabaco en los tres años se ve que, en comparación con los no fumadores, tienden a elegir mas amigos que los que los elijen a ellos, mientras que para los consumidores de cannabis, sucede todo lo contrario, resultando estos últimos máss atractivos como amigos. Para el caso del alcohol no hay una norma que se de para todos los años.

Finalmente, para completar el análisis realizado, se estudió la correlación entre los distintos consumos de sustancias. Para ello, se procedió de la siguiente forma: Considerando dos sustancias A y B, calculamos la cantidad de adolescentes que consumen ambas sustancias simultaneamente y calculamos la fracción sobre la cantidad de adolescentes que consumen la sustancia A, por un lado, y sobre los que consumen B, por otro. Luego, pesamos esa fracción con la esperada por azar:

$$CORR(A, B) = (A \cap B)/A - \frac{n_A}{N} \frac{n_B}{N}$$

$$CORR(B, A) = (A \cap B)/B - \frac{n_A}{N} \frac{n_B}{N}$$

donde N es la cantidad total de nodos y n_A y n_B son la cantidad de nodos que consumen la sustancia A y B, respectivamente.

En las tablas 6 a 8 se observan los resultados obtenidos para los distintos años, las sustancias nocivas tratadas, y el sexo de los alumnos (F/M):

| | Alcohol | Cannabis | Tabaco | Sexo(F) |
|----------|---------|----------|--------|---------|
| Alcohol | 0.0 | -0.52 | -0.53 | -0.6 |
| Cannabis | 0.34 | 0.0 | 2.04 | 0.37 |
| Tabaco | 1.54 | 4.81 | 0.0 | 2.00 |
| Sexo(F) | -0.16 | 0.02 | 0.17 | 0.0 |

Cuadro 6: Correlaciones calculadas para el primer año

| | Correlación 2do año | | | |
|----------|---------------------|----------|--------|---------|
| | Alcohol | Cannabis | Tabaco | Sexo(F) |
| Alcohol | 0.0 | -0.58 | -0.54 | -0.6 |
| Cannabis | -0.30 | 0.0 | 0.11 | -0.24 |
| Tabaco | 0.50 | 1.16 | 0.0 | 0.28 |
| Sexo(F) | -0.14 | -0.01 | -0.16 | 0.0 |

Cuadro 7: Correlaciones calculadas para el segundo año

| | Correlación 3er año | | | |
|----------|---------------------|----------|--------|---------|
| | Alcohol | Cannabis | Tabaco | Sexo(F) |
| Alcohol | 0.0 | -0.55 | -0.46 | -0.56 |
| Cannabis | 0.04 | 0.0 | 0.88 | 0.08 |
| Tabaco | 0.89 | 1.82 | 0.0 | 0.68 |
| Sexo(F) | -0.12 | -0.06 | -0.03 | 0.0 |

Cuadro 8: Correlaciones calculadas para el tercer año

Para una mejor visualización de la correlación, se puede observar un mapa de calor, realizado para el promedio de las correlaciones en los tres años, puesto que se observan comportamientos similares:

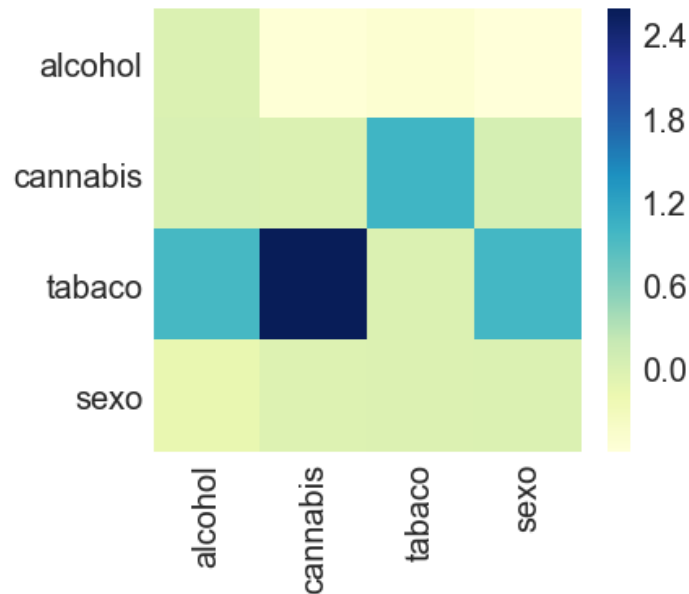


Figura 12: Mapa de calor para la correlación promedio de los tres años

Se observa, por un lado, que la intensidad de las correlaciones positivas son mucho mayores a las negativas, y en particular, que la gente que consume tabaco tiene una correlación positiva con respecto a las demás sustancias, particularmente con el cannabis, siendo esta la mayor de todo el mapa. Además, se observa que la mayor parte de los consumidores de tabaco son mujeres. Por otro lado, las correlaciones negativas se dan para el caso del alcohol, con lo que se ve que en ese caso el porcentaje de jóvenes que consumen alcohol y también otra sustancia es menor al esperado por azar.

4. Conclusiones

A partir del análisis de la red de adolescentes, se halló que en los lazos de amistad que se forman a lo largo del tiempo se dan mecanismos de homofilia e influencia con respecto al consumo de sustancias mayores a los que se esperarían por azar. Este comportamiento se pudo visualizar en la red tanto para influencia como para homofilia en formación de comunidades.

Además, se observó que para el caso de los consumidores de tabaco y cannabis, la forma en la cual los consumidores se relacionan con su ambiente es distinta para los no consumidores. En particular, los consumidores de cannabis resultan ser mas atractivos como amigos, mientras que los fumadores de tabaco suelen elegir mas amigos de los que los elijen a ellos.

Finalmente, se estudió la correlación que existe entre el consumo de alcohol, cannabis y tabaco y el sexo de los alumnos, donde la correlación mas notoria se da entre los consumidores de tabaco y cannabis.

A partir de los presentes resultados, se podrían fomentar campañas de intervención basados tanto en los mecanismos de propagación de consumo como en las correlaciones hayadas.

Referencias

- [1] http://www.stats.ox.ac.uk/~snijders/siena/Glasgow_data.htm
- [2] Pearson, Michael, Christian Steglich, and Tom A. B. Snijders. 2006. "Homophily and Assimilation Among Sport-Active Adolescent Substance Users."
- [3] Steglich, C, Snijders, T and Pearson, M. 2004. Dynamic Networks and Behaviour: Separating Selection from Influence.