

Popularidad en páginas de Facebook: ¿Qué rol cumplen las variables estructurales de la red?

Popularity in Facebook Pages: What role network structural variables play?

Emiliano M. Gutiérrez¹, Juan M.C. Larrosa²

Cómo citar (APA): Gutiérrez, E.M.; & Larrosa, J.M.C. (2020). Popularidad en páginas de Facebook: ¿Qué rol cumplen las variables estructurales de la red?. AWARI; 1(1). <https://doi.org/10.47909/awari.68>

Recibido: 20-02-2020

Aceptado: 2-06-2020

Editor: Adilson Luiz Pinto

Copyright: © 2020 Gutiérrez & Larrosa. This is an open access article distributed under the terms of the CC BY-NC 4.0 license which permits copying and redistributing the material in any medium or format, adapting, transforming and building upon the material as long as the license terms are followed.

ABSTRACT

Las páginas de Facebook (FP) son puntos de reunión entre eventos y actores en las redes sociales virtuales (RSV) de esta plataforma. Empresas y eventos aglutinan actores que participan a través de los mecanismos de expresión en línea que está habilitada, a través de reacciones, etiquetados, posteo de fotos, comentarios, etc. Estas interacciones forman redes de contacto en línea. Nos proponemos investigar las variables estructurales de las redes capturadas en línea con variables estructurales de estas redes. Encontramos que diversas variables estructurales referidas a concentración o desconexión entre nodos se encuentran asociadas negativamente con la popularidad medida por el número de likes siendo, por otra parte, las FPs más populares asociadas a medios de comunicación (radios y noticias).

Keywords: Facebook; Popularidad de páginas; Red Social

ABSTRACT

Facebook pages (FP) are meeting points between events and actors in virtual social networks (RSV) of this platform. Companies and events bring together actors that participate through the online expression mechanisms that are enabled, through reactions, tagging, photo posting, liking, comments, etc. These interactions form online contact networks. We propose to investigate structural variables of the networks captured online with structural variables of these networks. We found that various structural variables related to concentration or disconnection between nodes are negatively associated with popularity measured by the number of likes, on the other

¹ Departamento de Economía, Universidad Nacional del Sur. Email: emiliano.gutierrez@uns.edu.ar, ORCID: 0000-0002-6424-996X

² Instituto de Investigaciones Económicas y Sociales del Sur. Email: jarrosa@uns.edu.ar, ORCID: 0000-0002-0102-9994

hand, the most popular FPs associated with the media (radio and news).

Keywords: Facebook; Page Popularity; Social Network

1 INTRODUCCIÓN

El proceso de crecimiento de Internet ha sido una invaluable oportunidad para la masificación de comunicaciones. Un interesante fenómeno que puede ser observado desde los albores del siglo 21, es la aparición de las redes sociales virtuales (RSV). En este sentido, las RSV han cambiado nuestras vidas, permitiendo contactar con antiguos compañeros, buscar un trabajo, vender y comprar bienes y servicios, compartir material multimedia o estar en contacto el círculo familiar. En otras palabras, las RSV, han revolucionado los típicos mecanismos de comunicación, y generado un significativo cúmulo de información relevante para las ciencias sociales (Wilson y otros, 2012) y dando lugar a una importante línea de estudio vinculada al análisis de redes sociales (ARS). Antes de este hito, Newman (2003), enfatizaba los recurrentes problemas de poder construir una red social: sesgos en la definición de enlaces, pequeñas muestras y un importante trabajo para la construcción de una red (típicamente mediante encuestas).

Los datos provenientes de las RSV, puede ser considerados como información sumamente valiosa. Resulta de interés recolectar y estudiar relaciones sociales, detectando patrones de interacción provenientes de los entornos digitales. En este sentido, Facebook, resulta ser la RSV más popular del mundo con 1,63 billones de usuarios activos diariamente (Facebook, 2019). Individuos, familias, empresas, políticos, instituciones y sitios de eventos, interactúan a través de esta plataforma. Así, la gran popularidad de Facebook la hace de interés al momento de detectar aspectos sociodemográficos proveniente de los actores online dentro de aquellos campos de interés comunes entre los mismos (Slattery y otros, 2013), tales como deportes, noticias, instituciones educativas, entre otras. Cada uno de estos ejes de interacción puede ser concebido como una red social donde los individuos pueden expresar su apoyo, desinterés o desagrado, frente a una temática específica.

En el caso de Facebook, una forma de analizar este comportamiento es mediante la observación de las Fan Pages (FPs), las cuales pueden ser modelizadas como grafos mediante las herramientas típicas del ARS. Las FP son páginas ofrecidas por la plataforma como base de empresas, causas, blogs, etc que deseen posicionarse dentro de la misma con el acceso a muchos servicios de difusión y publicidad. Esta información podía ser descargada libremente hasta el escándalo de privacidad de datos de Cambridge Analytica (Isaak y Hanna, 2018). Por ello con datos de 2015 y usando herramientas de descarga como NodeXL (Smith y otros, 2010) se pudo observar el fenómeno de las FP para una ciudad entera (Bahía Blanca) durante un año completo de modo de ver qué diversas configuraciones de red las mismas que pueden adoptar. Específicamente, los enlaces de las redes estarán formados por los *likes*, comentarios y compartidos, los cuales también podrían ser utilizados como aproximaciones de la popularidad que recibe una FP respecto de otra.

La popularidad de las FP ha sido estudiada en diversos trabajos previos puntualizados en grupos focales para análisis temático interpretativo para estudios de alcoholismo en adolescentes como en Goodwin y otros (2016) o Swani y Milde (2017) usando la teoría motivacional psicológica para observar la correlación entre popularidad y ofrecimiento de productos por dar dos ejemplos. Sabate y otros (2014), por su parte, realizar un ejercicio de regresiones múltiples entre diversos posteos en Facebook detectando particularidades respecto a la popularidad de un post específico y su popularidad: posteos con imágenes

en horarios comerciales son más comentados, del mismo modo que imágenes y videos acarrearán más *likes* a una publicación, entre otros. Sin embargo, no es común relacionar la popularidad de un sitio con herramientas de ARS y ese es el propósito de este trabajo. Esta situación manifiesta dos preguntas que en esta investigación pretenderán ser respondidas:

Pregunta 1: ¿Es posible construir una medida que resuma la popularidad que recibe una FP, haciendo uso de indicadores propios de Facebook?

Pregunta 2: ¿Qué relaciones pueden encontrarse entre la popularidad de una FP y las métricas propias del ARS?

A través del enfoque utilizado en este trabajo, el mismo se dividirá en cuatro secciones a saber. En la primera de ellas, se presentan los principales antecedentes de interés vinculados a la temática abordada. En la sección 2, se expondrán los datos a utilizar. En un tercer apartado, se exhiben los resultados provenientes de analizar los vínculos entre la popularidad y las métricas estructurales de red. Finalmente, en la quinta sección se comentarán las principales conclusiones.

2 Antecedentes de relevancia

Típicamente, una red puede ser representada como un grafo compuesto por nodos conectados entre sí mediante enlaces. La disposición en la cual se distribuyen los mismos, da lugar a configuraciones específicas que puede adoptar una red. (Lee y Kalb, 2008; Bounova y De Weck, 2012). Algunas de las topologías más populares son las redes de mundo pequeño (Watts y Strogatz, 1998), libres de escala (Barabási y Albert, 1999), aleatorias (Erdos y Renyi, 1960) y jerárquicas (Ravasz y Barabási, 2003) por mencionar algunas.

Existen diversas maneras de poder modelar una red mediante la información contenida Facebook, siendo este aspecto de profunda relevancia, debido a su influencia al momento de interpretar los resultados obtenidos. Algunas investigaciones recurren al algoritmo de búsqueda BFS (Breadth First Search), para posteriormente construir la red a partir de un nodo inicial (Khadangi y otros 2018). De igual modo, es posible observar que otros trabajos recurren a palabras específicas (Wohlgemuth y Matache, 2014), creación de perfiles sobre una localidad específica (Viswanath y otros, 2009, Hu y Guo, 2012), información obtenida mediante relevamientos anteriores (Brauner y otros, 2009; Traud y otros, 2011) e incluso generación aleatoria de identificadores numéricos que posee esta plataforma (Buccafurri y otros, 2013). Asimismo, merece resaltarse el trabajo de Ugander y otros (2011), quienes, teniendo el acceso a todos los datos de Facebook, construyen una red, mediante los perfiles de aquellos individuos que informan su localización dentro de Estados Unidos.

Por otra parte, a causa de las diferentes modelizaciones sobre las que se plantea el diseño de la red, es de esperarse que no exista un consenso sobre el cual se le atribuya una topología de red específica a Facebook. Sin embargo, un aporte que puede considerarse al menos parcialmente esclarecedor viene dado por Backstrom y otros (2012), quienes al tener acceso a la totalidad de datos de Facebook, realizan un interesante análisis sobre los usuarios de esta plataforma a nivel mundial, siendo hoy día uno de los grafos de ARS de mayor tamaño jamás construido. Concretamente, se analizan 721 millones de usuarios (nodos) y 69 billones de pares de amigos (enlaces). Una de las conclusiones a la que arriban los autores es que Facebook podría ser concebido como un grafo de mundo pequeño con cortas distancias entre muchos pares de nodos.

Asimismo, Slattery y otros (2013) analizan redes de Facebook, mediante la construcción de los grafos provenientes de las interacciones dentro de las FP. Un aspecto interesante que los autores recomiendan que las investigaciones futuras deberían considerar la topología de red al momento de explicar indicadores de audiencia. Precisamente, este tipo de mediciones, remiten a concretamente a la noción de popularidad. En términos de Mohammadi y otros (2017) este concepto refiere a las diferentes métricas tales como visitas, *likes* o votos que una FP recibe en sus contenidos publicados.

Existen una importante cantidad de aportes que pretenden estimar la popularidad que recibe una FP centrado en *likes*, comentarios, compartidos y/o un promedio de los mismos con el fin de analizar diversos aspecto de la vida cotidiana tales como la atención de la ciudadanía en asuntos de gobierno (Díaz y Gutiérrez, 2020; Haro-de-Rosario y otros, 2016), deportes (Achen, 2016; Pronschinske y otros, 2012), estrategias de marketing (De Vries y otros, 2012; Antoniadis y otros, 2018), restaurantes (Kwok y Yu, 2013), agencias de viajes (Sabate y otros, 2014), impacto de publicaciones académicas (Ringelhan y otros, 2015), entre otras cuestiones de relevancia. Debido al enfoque adoptado por este tipo de investigaciones, se analiza la popularidad a partir del tipo de contenido posteado sin hacer uso de herramientas derivadas del ARS.

3 DATOS Y ANÁLISIS

Los datos utilizados en este trabajo provienen de las interacciones de 323 FP mediante el software NodeXL (Smith y otros, 2010). Dicha información pretende capturar la información proveniente de negocios, actividades colectivas, instituciones, entre otras. El denominador común de las FP es que las mismas se hallan localizadas en la ciudad de Bahía Blanca (Argentina). En total, se registró la actividad de más de 100 mil usuarios, desagregados en 29 sectores de influencia. En la figura 1 se expone la composición de FPs por sector.

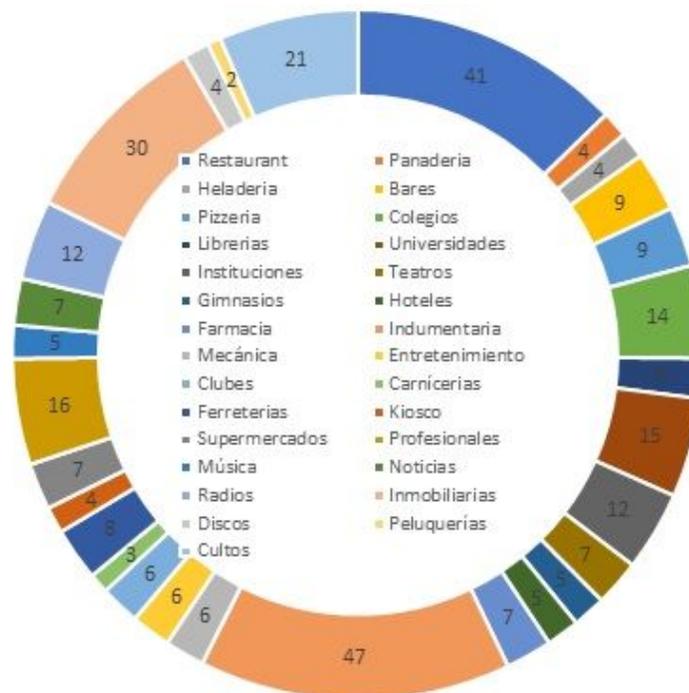


Figura 1. Composición de FPs por sector.

El período temporal bajo análisis fue la totalidad del año 2015, permitiendo relevar gran parte de la actividad digital de la zona geográfica referida anteriormente, cuya población total según el último censo realizado es de alrededor de 300.000 habitantes (INDEC, 2012).

3.1 Popularidad

A fin de construir un indicador que integre la mayor cantidad de información posible, se considera la definición de una métrica alternativa a las típicas mediciones que simplemente consideren *likes*, comentarios o compartidos.

Como respuesta a la pregunta 1 formulada al inicio del trabajo, se pretende buscar una medida alternativa de popularidad y que contenga la mayor cantidad de información posible. Así, se propone que la misma se halle contenida por los comentarios, reacciones, etiquetas y *likes* observados en cada FP, siendo entonces considerados los siguientes atributos:

- Liked Comment: Likes a los comentarios;
- Reacted to post: Reacciones al posteo;
- User tagged in post: Número de usuarios etiquetados en post;
- Commented post: Comentarios realizados por los usuarios dentro de los posts;
- Shared Post: Total de posteos compartidos;
- Commented Comment: Cantidad de comentarios comentados;
- Co-reactor: Reacciones a los comentarios;
- Consecutive commenter: Total de comentarios consecutivos (con respuesta);
- Co-commenter: Usuarios respondiendo comentarios dentro del posteo;
- User Reacted Post: Total de usuarios que reaccionaron al posteo.

De esta forma, mediante un Análisis de Componentes Principales (ACP), se plantea reducir la dimensionalidad de los atributos referidos anteriormente. Dicha técnica resulta de interés para detectar relaciones que a priori no habían sido reconocidas (Zinovyev, 2011; Sabatini, 2009).

Para esto, un primer paso requiere la estandarización de los datos, los cuales pueden ser resumidos en una matriz (X) de n observaciones y p variables. Cada fila representaría a una FP con sus atributos correspondientes, pudiendo esto ser representado tal como:

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \dots & x_{1,p} \\ x_{2,1} & x_{2,2} & \dots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \dots & x_{n,p} \end{bmatrix} \quad (1)$$

Asimismo, se requiere la estimación de la matriz de correlaciones entre las variables, dando lugar a la matriz R de tamaño p x p.

$$R = \begin{bmatrix} 1 & r_{1,2} & \dots & r_{1,p} \\ r_{2,1} & 1 & \dots & r_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p,1} & r_{p,2} & \dots & 1 \end{bmatrix} \quad (2)$$

Sin embargo, R puede ser representada como un producto escalar con una dimensión de p vectores. La aplicación de PCA busca ejes para maximizar la varianza en los datos originales, siendo posible representar cada componente Z como una combinación lineal.

$$Z_i = \sum_{j=1}^p \alpha_{ij} X_j \quad (3)$$

Asimismo, los coeficientes α son denominados como loadings (cargas), y permiten observar el impacto de cada variable dentro de la componente Z. Cada observación tendrá dentro de cada componente específico una valoración, denominada score o puntaje.

En la tabla 1, se representan las loadings asociadas al primer componente. Esta única dimensión, recupera un 59,5% de la varianza total. Asimismo, se detalla el porcentaje de reconstrucción de la variable, resulta ser el cuadrado de la correlación del componente, respecto de la variable original.

El valor de todas las *loadings* relevadas en esta dimensión posee signo positivo, lo cual implicaría que todas las variables utilizadas impactan en la misma dirección sobre el valor final del score de cada FP. Sobresalen los atributos vinculados a los comentarios y sus respuestas como así también los likes a los mismos.

Tabla 1. Resultados de la primera componente.

Variable	Loadings	Porcentaje de reconstrucción
Liked Comment	0,370	81,68%
Reacted to Post	0,371	81,89%
User Tagged in Post	0,138	11,39%
Commented Post	0,392	91,51%
Shared Post	0,062	2,32%
Co-Reactor	0,379	85,42%
Commented Comment	0,252	37,83%
Consecutive Commenters	0,380	85,94%
Co-Commenter	0,320	60,85%
User Reacted Post	0,307	56,25%

Variable	Loadings	Porcentaje de reconstrucción
Varianza explicada		59,51%

Los scores asociados a este componente, pueden ser utilizados como una medida de interés que resuma la popularidad de una FP. Mayores scores, dan lugar a que una FP sea considerada más popular. Sin embargo, al estar utilizando los datos estandarizados, es de esperarse que aquellas FP con un menor nivel de interacción, podrían llegar a obtener scores negativos.

En el gráfico 1, se exhiben los diagramas de caja asociados al score de cada una de las FP, desagregadas por sector. Radios y noticias son los sectores cuyas FP poseen mayores niveles de participación, lo cual podría explicarse por la interacción que genera la difusión de contenidos de las mismas. Asimismo, otros sectores como farmacia y panadería presentan importantes dispersiones, debido a que en ambos casos existe una FP marcadamente más popular respecto de las otras.

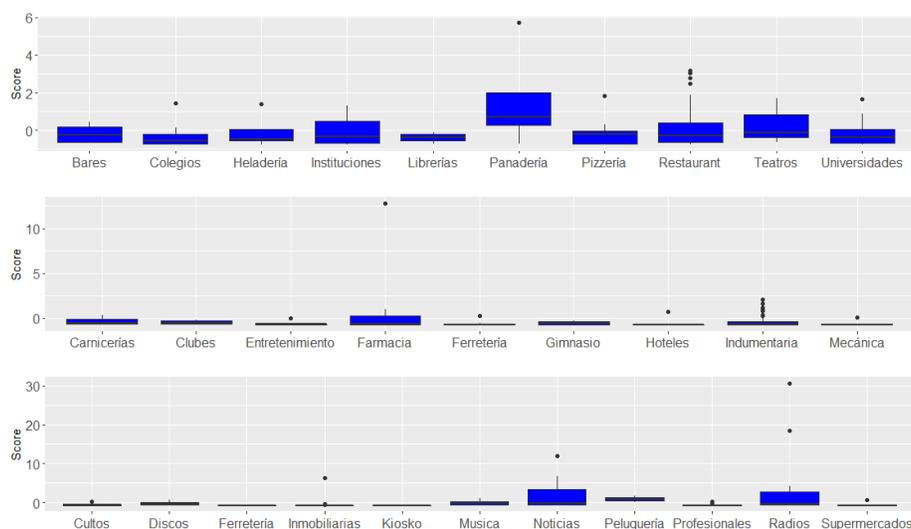


Gráfico 1. Diagrama de caja de scores obtenidos por sector.

Respondiendo a la primera pregunta de investigación formulada al inicio de este trabajo, se considera pertinente caracterizar a esta primera dimensión relevada como un indicador de popularidad, reflejando entonces el nivel de atención que recibe una FP, mediante los scores asociados a la misma.

Métricas estructurales de red

Considerando a cada FP, como un grafo es posible observar las métricas propias del ARS para cada una de ellas. En la tabla 2, se exhiben las 23 métricas consideradas, las cuales resultan ser relevadas para cada una de las FP analizadas.

Métrica	Descripción
Nodos	Cantidad total de nodos del grafo

Métrica	Descripción
Enlaces únicos	<i>Enlaces que no se repiten en destino y origen</i>
Enlaces duplicados	<i>Enlaces cuyo origen y destino se repiten</i>
Enlaces totales	<i>Cantidad total de enlaces del grafo</i>
Auto-enlaces	<i>Enlace cuyo origen y destino es el mismo nodo</i>
Coefficiente de pares de nodos recíprocos	<i>Proporción de nodos adyacentes</i>
Coefficiente de pares de enlaces recíprocos	<i>Proporción de enlaces recíprocos</i>
Máximo de nodos en un componente conectado	<i>Cantidad de nodos en el componente gigante</i>
Máximos de enlaces en un componente conectado	<i>Cantidad de enlaces en el componente gigante</i>
Diámetro	<i>Máxima distancia de conexión entre dos nodos</i>
Distancia geodésica media	<i>Promedio de distancia de la distancia más corta entre pares de nodos</i>
Densidad	<i>Proporción de enlaces respecto de una red completa</i>
Asortatividad de grado entrada-entrada	<i>Correlación de los enlaces de los nodos de entrada respecto de los enlaces de los nodos de entrada</i>
Asortatividad de grado entrada-salida	<i>Correlación de los enlaces de los nodos de entrada respecto de los enlaces de los nodos de salida</i>
Asortatividad de grado salida-entrada	<i>Correlación de los enlaces de los nodos de salida respecto de los enlaces de los nodos de entrada</i>
Asortatividad de grado salida-salida	<i>Correlación de los enlaces de los nodos de salida respecto de los enlaces de los nodos de salida</i>
Índice de Zagreb	<i>Cuadrado de la suma del grado de enlaces de los nodos</i>

Métrica	Descripción
M1	
Grado medio	<i>Promedio de enlaces de cada nodo</i>
Número de pares inalcanzables	<i>Cantidad de pares de nodos que no poseen un sendero que los una</i>
Centralización de intermediación de red	<i>Variación en centralidad de intermediación de los nodos respecto de la máxima variación de centralidad de intermediación posible en una red del mismo tamaño.</i>
Modularidad	<i>Proporción los enlaces que caen dentro de grupos dados menos el valor esperado que esta proporción hubiese recibido si los enlaces se hubiesen distribuido aleatoriamente.</i>
Coefficiente de agrupamiento Watts-Strogatz	<i>Promedio simple de los coeficientes de agrupamiento local de todos los nodos con grado al menos 2.</i>
Coefficiente de agrupamiento de red (Transitividad)	<i>Promedio ponderado de los coeficientes de agrupamiento local de todos los nodos con grado al menos 2.</i>

Tabla 2. Métricas estructurales de red relevadas.

Así las cosas, resulta de interés evaluar la existencia de alguna relación existente entre las métricas estructurales de la red respecto de la popularidad de que presentan las FP, aspecto que será abordado en la próxima sección.

4 RESULTADOS

Frente a la segunda pregunta planteada en la introducción de este artículo, puede señalarse que una evaluación de la relación existente entre popularidad y métricas de red podría ser estudiada mediante el uso de Mínimos Cuadrados Ordinarios (MCO), a través la típica regresión multivariada, la cual puede ser planteada como:

$$Popularidad = b_0 + \sum_{i=1}^{23} b_i \gamma_i + \varepsilon \quad (4)$$

Cada γ_i representa una métrica en cuestión. En la tabla 3, se exponen los resultados de esta estimación. Sin embargo, la misma presenta síntomas que en simultáneo revelarían la existencia de multicolinealidad, tales como una elevada bondad de ajuste ($R^2=0,9015$), con pocas variables significativas, (8 de las 23 consideradas) y altos Factores Inflacionarios de Varianza (FIV), donde en el caso de estos últimos, 5 regresores resultan ser mayores a 10.

Tabla 3. Estimación mediante MCO (Nota: $R^2 =0,9015$; R^2 ajustada= $0,894$; *: Denota significatividad al 95%).

	Coefficiente	Error estándar	P-valor	FIV
(Intercepto)	-1,84300	0,58	0,00*	
Nodos	0,00123	0,00	0,00*	13,11
Enlaces únicos	0,00001	0,00	0,01*	39,62
Enlaces duplicados	0,00006	0,00	0,00*	6,87
Enlaces totales	-0,00004	0,00	0,00*	70,61
Auto-enlaces	0,00002	0,00	0,78	1,14
Coefficiente de pares de nodos recíprocos	-3,16500	4,04	0,43	82,69
Coefficiente de pares de enlaces recíprocos	1,15800	2,81	0,68	88,24
Máximo de nodos en un componente conectado	0,00001	0,00	0,41	1,17
Máximos de enlaces en un componente conectado	-0,00001	0,00	0,00*	4,80
Diámetro	0,02113	0,07	0,75	4,35
Distancia geodésica media	-0,05406	0,30	0,86	6,50
Densidad	0,81630	0,71	0,25	3,86
Asortatividad de grado entrada-entrada	-0,44380	0,49	0,37	4,57
Asortatividad de grado entrada-salida	0,06040	0,42	0,89	1,78
Asortatividad de grado salida-entrada	-0,21840	0,51	0,67	3,60
Asortatividad de grado salida-salida	-0,11790	0,39	0,76	4,95
Índice de Zagreb M1	0,00001	0,00	0,00*	8,02
Grado medio	0,00158	0,00	0,33	8,32
Número de pares inalcanzables	0,0000002	0,00	0,01*	3,07

	Coeficiente	Error estándar	P-valor	FIV
Centralización de intermediación de red	-0,33940	0,41	0,41	1,89
Modularidad	0,32170	0,49	0,51	2,68
Coeficiente de agrupamiento Watts-Strogatz	1,88900	0,52	0,00*	2,01
Coeficiente de agrupamiento de red (Transitividad)	0,17090	0,56	0,76	3,15

La utilización entonces de una regresión lineal multivariada, queda en este caso, parcialmente invalidada debido a la presencia de multicolinealidad. Dicho aspecto si bien no altera las propiedades de los estimadores, reduce el número de variables significativas. Sobresalen algunos signos esperados, como la presencia de una mayor cantidad de nodos, que explicaría la atención que recibe la red. Misma situación se replica con el caso del coeficiente de Watts-Strogatz, el cual denota la existencia de una relación positiva entre agrupamiento y popularidad, pudiendo ser este explicado por la presencia de comunidades específicas dentro de las diferentes FP.

En lo que respecta a los enlaces e impacto en la popularidad, esta estimación refleja una relación positiva para los enlaces únicos y duplicados respecto de la variable dependiente. En el primer caso, podría ser justificado por la existencia de enlaces asociados a eventos con una menor probabilidad de ocurrencia, tales como compartir o etiquetar usuarios, mientras que el segundo, aludiría al acuerdo o reconocimientos constantes hacia ciertos nodos concretos. De manera contraria sorprende el signo del atributo que releva los enlaces totales, pudiendo ser explicado porque dicha variable no distinguiría entre diversas acciones que impactan de manera desigual en la popularidad, tales como *likes*, comentarios y compartidos.

Asimismo, en el caso del índice de Zagreb M1, esta métrica refleja directamente una complejidad dentro que puede poseer una red, lo cual es de esperarse que ocurra en aquellas FP con mayor popularidad, donde el grado de los nodos integrantes de estas redes podría ser superior.

Sin embargo, a fin de capturar toda la información proveniente de las métricas estructurales de red y observar su relación con la popularidad que puede recibir una FP y, simultáneamente subsanando la presencia de multicolinealidad es factible recurrir a la Regresión de Componentes Principales (RCP). En el caso de RCP es posible la aplicación de MCO, pero donde las variables explicativas resultan ser los componentes principales de los atributos utilizados anteriormente (Hastie y otros, 2009). De esta manera, la estimación se realiza en dos pasos. En primer lugar, se aplica ACP sobre las métricas estructurales ya definidas. Posteriormente se seleccionan los componentes de relevancia, calculándose los scores correspondientes a cada dimensión para finalmente estimarse una regresión lineal cuya variable dependiente es la popularidad de la FP, lo cual puede ser expresado como:

$$Popularidad = \alpha_0 + \sum_{k=1}^{23} \alpha_k CP_k \quad (5)$$

Donde CP equivale a cada componente principal que contiene a las métricas estructurales. El máximo de CP a utilizar es de 23, lo cual correspondería a hacer uso de toda la información

contenida en las variables explicativas. Sin embargo, si bien se soluciona la multicolinealidad con la utilización de RCP, también es cierto que se adiciona una dificultad vinculada a la interpretación de cada componente.

La elección de cuantos componentes utilizar es también un aspecto de discusión debido a que la reducción en dimensión implica una pérdida en la bondad de ajuste del modelo. En el gráfico 2, se observa la relación existente entre el R^2 y el número de componentes utilizados. Evidentemente, la adición de componentes, otorga mayor ajuste al modelo. Sin embargo, debido al trade-off existente entre ajuste e interpretación, se hará uso de las dos primeras dimensiones. Dicho umbral se encuentra representado en la gráfica por la línea roja puntada. Esta cantidad de componentes, da lugar a un R^2 de 0.54.

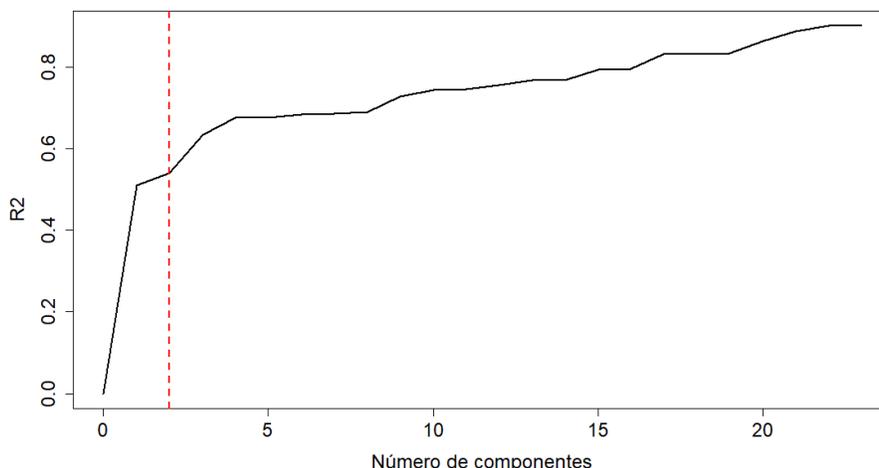


Figura 3. Relación entre bondad de ajuste y número de componentes.

En la tabla 4, se muestran los loadings asociados las dos primeras componentes, optándose por suprimir aquellos cuyo valor absoluto resulta ser inferior a 0.1. En el caso del primer componente (CP1), observa que tanto las variables asociadas a relevar la cantidad de nodos y enlaces, poseen el mismo signo. Estas variables, resultan de importancia para poder observar el tamaño de la red. Sin embargo, en dirección opuesta, aquellas métricas asociadas a la reciprocidad de enlaces y adyacencia de los nodos resultan tener un coeficiente positivo. De igual modo, las conexiones al componente gigante, junto con ciertas tipologías de asortatividades, la modularidad y el índice de Zagreb también poseen un signo negativo. Estas últimas métricas, se encuentran íntimamente ligadas con el agrupamiento que pueda existir en una red y la complejidad de la misma, debido al número de interacciones.

Esta componente parecería reflejar la característica de desconexión que puede tener una red, donde la influencia del tamaño de la misma (Nodos y enlaces), la existencia de enlaces recíprocos, con escasas conexiones del resto de la red como así también las métricas destinadas a caracterizar el agrupamiento y complejidad de la red, resultarían encontrarse en la misma dirección que la cantidad de nodos y enlaces. Así las cosas, se considera como la *desconexión*, como el principal denominador en torno a este componente.

	CP 1	CP 2
Nodos	-0,328	
Enlaces únicos	-0,342	-0,106
Enlaces duplicados	-0,25	
Enlaces totales	-0,359	
Auto-enlaces		
Coefficiente de pares de nodos recíprocos	0,105	0,13
Coefficiente de pares de enlaces recíprocos	0,107	0,144
Máximo de nodos en un componente conectado		
Máximos de enlaces en un componente conectado	-0,302	
Diámetro	-0,251	0,193
Distancia geodésica media	-0,215	0,339
Densidad	0,16	-0,365
Asortatividad de grado entrada-entrada		-0,415
Asortatividad de grado entrada-salida	-0,152	
Asortatividad de grado salida-entrada	-0,253	
Asortatividad de grado salida-salida		-0,401
Índice de Zagreb M1	-0,288	-0,159
Grado medio	-0,293	-0,194
Número de pares inalcanzables	-0,146	
Centralización de intermediación de red		0,315
Modularidad	-0,178	0,167
Coefficiente de agrupamiento Watts-Strogatz		
Coefficiente de agrupamiento de red (Transitividad)		-0,321

	CP 1	CP 2
<i>Varianza explicada (componente)</i>	27,14%	15,60%
<i>Varianza explicada (acumulada)</i>	27,14%	42,74%

Tabla 4. Loadings de las dos primeras componentes.

La segunda componente (CP2), destaca por la dirección de las asortatividades, como así también la densidad y el índice de Zagreb. De manera opuesta, se destaca la centralización de intermediación de red y la distancia geodésica media, esta situación sería característica de redes donde existen importantes nodos que concentran mucha información, debido a la escasez de enlaces. Asimismo, la negatividad del coeficiente de agrupamiento de red, también podría resultar explicativa en lo que respecta a que esta dimensión estaría indicando la existencia de pocos agentes con un gran poder de intermediación dentro de la red. Se propone entonces como denominar este componente como *concentración*.

Una vez estimados los componentes, resulta de interés analizar el coeficiente que adoptan los mismos dentro de la regresión, tal como se observa en la Tabla 5.

Tabla 5³: Estimación RCP sobre la popularidad (Nota: $R^2=0,54$; R^2 ajustada=0,538; *: Denota significatividad al 95%).

	Coefficiente	Error estándar	P-valor
<i>Desconexión (CP1)</i>	-0,697	0,037	0,00*
<i>Concentración (CP2)</i>	-0,227	0,049	0,00*

Ambos coeficientes resultar ser negativos. En el caso de la componente señalada como desconexión, es de esperarse que aquellas redes con este tipo de características, vean reducida su popularidad, debido a que los nodos no posean enlaces constantes ni presenten algún interés relevante que los induzca a interactuar. La concentración (CP2), también tiene una relación con negativa con la popularidad debido a que al existir pocos nodos con gran poder de información, resulta plausible considerar la ausencia vínculos fluidos sobre la red.

5 CONCLUSIONES

El ARS, para comprender fenómenos sociales vinculados a las plataformas digitales, tiene un vasto camino por desarrollar. En el caso de este trabajo, se intentó analizar la vinculación existente entre la configuración de una red y su popularidad. Previo a esto se construyó una medida de popularidad que resume la cantidad de información vinculada a las interacciones que pueda recibir una FP.

3 Dado la ortogonalidad de las matrices proyectadas al momento de la proyección realizada mediante RCP, el FIV es similar a 1 mientras que la estandarización de las variables, da lugar a un intercepto igual a 0 por lo que se omiten dichos resultados.

Dos tipos de estimaciones fueron realizadas a fin de establecer los vínculos entre popularidad y las métricas de red. La primera de ellas consistió en una regresión lineal multivariada donde se detectó que ciertas métricas específicas repercuten positivamente mientras que otras lo hacen de forma negativa. Sin embargo, de manera complementaria, se recurrió a la RCP, donde se observó que tanto aquellas redes cuya información se encuentra concentrada o existe desconexión entre los nodos, se encuentran asociadas negativamente con la popularidad que puede tener una FP. Precisamente, las FPs más populares se encuentran asociadas a medios de comunicación (radios y noticias), lo cual explicaría que estos sectores, exista un gran tamaño de red, con marcados agrupamientos de usuarios y donde la concentración debería ser menor.

Finalmente, se consideran como futuras líneas de investigación, la indagación de los comportamientos de usuarios dentro otras RSVs (Twitter, Instagram, etc.) y con los indicadores de popularidad típicos de las mismas a fin de observar la existencia de similitudes (o no) de los resultados obtenidos en esta investigación.

BIBLIOGRAFÍA

Achen, R. (2016). "Examining the Influence of Facebook Fans, Content, and Engagement on Business Outcomes in the National Basketball Association". *Journal of Social Media for Organizations*, 3 (1), 1-15. Disponible en: <http://www2.mitre.org/public/jsmo/pdfs/03-01-nba-business-outcomes.pdf>

Antoniadis, I., Simos, P. y Assimakopoulos, C. (2018). "Factors affecting engagement and reactions of posts in Greek Facebook brand pages". *4th International Conference on Contemporary Marketing Issues*.

Backstrom, L., Boldi, P., Rosa, M., Ugander, J., y Vigna, S. (2012). "Four degrees of separation". *Proceedings of the 3rd Annual ACM Web Science Conference on - WebSci '12*.

Barabási, A.-L. y Albert R. (1999). "Emergence of Scaling in Random Networks". *Science*, 286 (5439): 509-512. (DOI: 10.1126/science.286.5439.509)

Brauner, K., Kocheturov, A., Pardalos, P. M. (2018). "Network Analysis of Page Likes from Facebook User Profiles". *Journal of Undergraduate Research* 18(2). Disponible en: <https://ufdc.ufl.edu/UF00091523/00851>

Bounova, G., y De Weck, O. (2012). "Overview of metrics and their correlation patterns for multiple-metric topology analysis on heterogeneous graph ensembles". *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 85(1). Disponible en: <https://doi.org/10.1103/PhysRevE.85.016117>

Buccafurri, F., Lax, G., Nocera, A., y Ursino, D. (2013). "Internetworking Assortativity in Facebook". *International Conference on Cloud and Green Computing 2013*.

De Vries, L., Gensler, S., y Leeflang, P. S. H. (2012). "Popularity of Brand Posts on Brand Fan Pages: An Investigation of the Effects of Social Media Marketing." *Journal of Interactive Marketing*, 26(2), 83-91. Disponible en: <https://doi.org/10.1016/j.intmar.2012.01.003>

Díaz, L., y Gutiérrez, E. (2020). "Utilización de la red social Facebook por parte del gobierno de Bahía Blanca en tiempos de coronavirus. Análisis de factores que influyen las

interacciones en las publicaciones”. En *La Investigación en Ciencias Sociales en tiempos de la Pandemia por Covid-19* (p. 264–278). Bahía Blanca: IIESS UNS-CONICET.

Erdős, P. y Rényi, A. (1960), “On the Evolution of Random Graphs”. *Institute of Mathematics, Hungary Academic of Sciences* 5, 17-61. Disponible en: <http://snap.stanford.edu/class/cs224w-readings/erdos60random.pdf>

Facebook (2019), Key facts. Disponible en: <https://newsroom.fb.com/company-info>

Goodwin, I., Griffin, C., Lyons, A., McCreanor, T. y Moewaka Barnes, H. (2016) “Precarious Popularity: Facebook Drinking Photos, the Attention Economy, and the Regime of the Branded Self”, *Social Media + Society* 2(1). Disponible en: <https://doi.org/10.1177/2056305116628889>

Haro-de-Rosario, A., Sáez-Martín, A., y del Carmen Caba-Pérez, M. (2016). “Using social media to enhance citizen engagement with local government: Twitter or Facebook?” *New Media & Society*, 20(1), 29-49. Disponible en: <https://doi.org/10.1177/1461444816645652>

Hastie, T., Tibshirani, R., y Friedman, J. (2009). *The Elements of Statistical Learning*. En Springer Series in Statistics. Springer New York.

Hu, H. B. y Guo, J.L., (2012). “A comparative research on Facebook networks in different institutions”. *Advances in Complex System*, 15. Disponible en: <https://doi.org/10.1142/S0219525912500300>

Instituto Nacional de Estadísticas y Censos (2012). Censo Nacional de Población, Hogares y Viviendas 2010.

Isaak, J. y Hanna, M. J. (2018), "User Data Privacy: Facebook, Cambridge Analytica, and Privacy Protection," *Computer* 51(8): 56-59. Disponible en: <https://doi:10.1109/MC.2018.3191268>

Khadangi, E., Bagheri, A. y Zarean, A. (2018). “Empirical analysis of structural properties, macroscopic and microscopic evolution of various Facebook activity networks”. *Quality and Quantity*, 52, 249–275. Disponible en: <https://doi.org/10.1007/s11135-016-0465-4>

Kwok, L., y Yu, B. (2012). “Spreading Social Media Messages on Facebook”. *Cornell Hospitality Quarterly*, 54(1), 84-94. Disponible en: <https://doi.org/10.1177/1938965512458360>

Lee, D. y Kalb, J. (2008). *Network Topology Analysis* (Reporte técnico). Nuevo México, Estados Unidos. Sandia National Laboratories

Mohammadi, S., Farahbakhsh, R. y Crespi, N. (2017). “Popularity evolution of professional users on Facebook”. *2017 IEEE International Conference on Communications (ICC)*.

Newman, M.E. (2003), “The Structure and Function of Complex Networks”, *SIAM Review* 45, 167-256. Disponible en: <https://epubs.siam.org/doi/pdf/10.1137/S003614450342480>

Pronschinske, M., Groza, M. y Walker, M. (2012). “Attracting Facebook 'fans': The importance of authenticity and engagement as a social networking strategy for professional sport teams”. *Sport Marketing Quarterly*, 21: 221-231.

- Ravasz, E., y Barabási, A.-L. (2003). "Hierarchical organization in complex networks." *Physical Review E*, 67(2). Disponible en: https://regan.med.harvard.edu/pdf/publications/Hierarchical_org_in_Complex_Netw_PRE.pdf
- Ringelhan, S., Wollersheim, J., y Welppe, I. M. (2015). "I Like, I Cite? Do Facebook Likes Predict the Impact of Scientific Work?" *PLOS ONE*, 10(8). Disponible en: <https://doi.org/10.1371/journal.pone.0134389>
- Sabate. F., Berbegal-Mirabent, J., Cañabate, A. y Lebherz, P.R. (2014) "Factors influencing popularity of branded content in Facebook fan pages", *European Management Journal* 32 (6): 1001-1011. Disponible en: <https://doi.org/10.1016/j.emj.2014.05.001>
- Sabatini, F. (2009). "Social capital as social networks: A new framework for measurement and an empirical analysis of its determinants and consequences". *Journal of Socio-Economics*, 38(3), 429–442. Disponible en: <https://doi.org/10.1016/j.socec.2008.06.001>
- Slattery, R.E., McHardy R.R., y Bairathi R. (2013), "On the Topology of the Facebook Page Network". *Proceedings from Engagement Science: Theory and Methods: Insights from the Emergence*.
- Smith, M., Ceni A., Milic-Frayling, N., Shneiderman, B., Mendes Rodrigues, E., Leskovec, J., Dunne, C. (2010). NodeXL: a free and open network overview, discovery and exploration add-in for Excel 2007/2010/2013/2016, <http://nodexl.codeplex.com/> from the Social Media Research Foundation. Disponible en: <http://www.smrfoundation.org>
- Swani. K. y Milne, G.R. (2017), "Evaluating Facebook brand content popularity for service versus goods offerings", *Journal of Business Research* 79: 123-133. Disponible en: <https://doi.org/10.1016/j.jbusres.2017.06.003>
- Traud, A. L., Mucha, P. J. y Porter, M. A. (2012). "Social structure of Facebook networks". *Physica A: Statistical Mechanics and Its Applications*, 391(16), 4165-4180. Disponible en: <https://doi.org/10.1016/j.physa.2011.12.021>
- Ugander, J., Karrer B., Backstrom L. y Marlow C. (2011). "The Anatomy of the Facebook Social Graph". arXiv:1111.4503 [cs.SI]
- Viswanath, B., Mislove, A., Cha, M. y Gummadi, K.P (2009). "The evolution of user interaction in Facebook". *Proceedings of the 2nd ACM Workshop on Online Social Networks*, 37–42. Disponible en: <https://conferences.sigcomm.org/sigcomm/2009/workshops/wosn/papers/p37.pdf>
- Watts, D.J. y Strogatz, S.H. (1998). "Collective dynamics of 'small-world' networks". *Nature*, 393 (6684). Disponible en: <https://doi.org/10.1038/30918>
- Wilson, R. E., Gosling, S. D. y Graham, L.T. (2012). "A Review of Facebook Research in the Social Sciences". *Perspectives on Psychological Science*, 7(3), 203-220. Disponible en: <https://doi.org/10.1177/1745691612442904>
- Wohlgemuth, J. y Matache, M.T. (2014), "Small-world properties of Facebook group networks". *Complex Systems*. 23(3): 197-225. Disponible en: https://www.complex-systems.com/abstracts/v23_i03_a01/

Zinovyev, A. (2011). Data visualization in political and social sciences. International Encyclopedia of Political Science. Disponible en: <https://arxiv.org/abs/1008.1188>