



Universidad del Desarrollo
Universidad de Excelencia

MÉTRICAS TOPOLÓGICAS EN REDES DE PROTESTAS EN CHILE

ISRAEL CEBALLOS

**Proyecto de Grado entregado a la Facultad de Ingeniería para optar al grado académico
de Magíster en Data Science**

Profesor guía: CRISTIAN CANDIA

Facultad de Ingeniería
Universidad del Desarrollo
Chile

Julio, 2022

Índice

1. Introducción	4
2. Hipótesis y Objetivos	5
3. Metodología	6
3.1. Fuentes de información	6
3.2. Glosario	6
3.2.1. Variables generales	6
3.2.2. Variables COES	7
3.2.3. Nomenclatura de grafos	7
3.3. Procesamiento de la información	7
4. Definiciones	7
4.0.1. Degree	7
4.0.2. Mean neighborhood degree	8
4.0.3. Clustering coefficient	8
4.0.4. Mean Clustering coefficient	8
4.0.5. Betweenness centrality	8
4.0.6. Eigenvector centrality	8
4.0.7. Bridging centrality	9
4.0.8. Average path length	9
4.0.9. maxelement & removemax	9
4.0.10. Asortatividad & Disortatividad	9
4.0.11. Rich Club	9
5. Estructura topológica de la red	11
5.1. Evolución y centralidad en la red	13
5.2. Rich club phenomenon	16
6. Correlaciones	17
6.1. n° protestas / uso de internet	17
6.2. n° detenidos/ uso de internet	17
6.3. escolaridad / n° detenidos	18
6.4. participación política / n° protestas	18
6.5. escolaridad / subsidios y bonos	19
6.6. ayudas escolares/ bienes propios	19
7. Superficies	20
7.1. Impacto en el entorno	22
7.1.1. Restaurantes	22
7.1.2. Policía	23
7.1.3. Bares	23
7.1.4. Hospitales	24
7.1.5. Tiendas	24
7.1.6. Parques	25

8. Modelos estadísticos	27
8.1. Metodología	27
8.2. Interpretación de resultados	27
8.3. Model 1: $p_{11} \geq 10$	28
8.4. Model 2: $p_{11} \geq 2000$	29
8.5. Model 1b: $p_{23} \geq 20$	30
9. Conclusiones y trabajo futuro	31
10. Apéndice	32
10.1. Fuentes de información	32
10.1.1. Encuesta CASEN	32
10.1.2. REVIEWS	32
11. Bibliografía	33

Resumen

El presente trabajo tiene como objetivo el modelamiento de los actos de protesta en Chile entre los años 2009 y 2019 mediante el uso de la teoría de grafos. Para la construcción de la red, asumiremos que los nodos corresponderán a organizaciones manifestantes y las conexiones se generaran según dos o más organizaciones convoquen al mismo acto de protesta. Aplicaremos una variedad de medidas topológicas al grafo para observar cómo distribuyen estas métricas según el **degree**, dichos valores serán agregados a medidas CASEN y COES para múltiples propositos, ya sea para obtener correlaciones entre variables [6], como también para analizar el impacto que genera una protesta en su entorno a partir de las coordenadas de los focos de protesta y las evaluaciones de **amenities** colindantes a ellas, junto a lo anterior, añadiremos un análisis de superficies y distancias en [7].

En una sección final [8] incluiremos modelos estadísticos de regresión lineal orientados a predecir el número de manifestantes o detenidos de acuerdo a las variables socioeconómicas del entorno y las medidas topológicas aplicables al grafo.

Palabras clave: **grafo, manifestación, protesta.**

1. Introducción

La base del presente estudio nace de la recopilación del Reproducible Research and COES (2020), esta base incluye los actos de protesta registrados por medios de comunicación escritos o digital entre los años 2009 y 2019, dicho registro contiene características propias de una protesta, detallando los nombres de las organizaciones convocantes, la ubicación geográfica, la causa principal, motivación o consigna asociada a la protesta, el número de detenidos, heridos y muertos, entre otros campos.

Diversas líneas de investigación se han planteado a raíz del fenómeno de las revueltas sociales en Chile entre los años 2009 y 2019, parte de estas respuestas se centran en el uso de metodologías propias de las ciencias sociales. El enfoque de este trabajo excluirá conclusiones que incluyan únicamente aspectos ideológicos, políticos o psicológicos del entorno, ya que nos concentraremos en las características granulares de las organizaciones y plantearemos métodos para relacionar a las organizaciones con las características socioeconómicas de su entorno .

La construcción del grafo requiere de una relación entre nodos organizadores y nodos manifestaciones, el vínculo está dado por el hecho de ser convocantes de un acto en común, a partir de ello y mediante una proyección bipartita logramos relacionar a los agentes movilizados entre sí.

Describiremos la topología de la red de protestas presentando múltiples medidas que describan al grafo en su instante final como en su evolución a lo largo del tiempo, incluiremos un análisis de superficies mediante la construcción de polígonos convexos o circunferencias de un radio específico que contengan las coordenadas de los focos de protesta, esto permitirá comprender la relación entre un nodo organizador con su entorno y así evaluaremos los impactos positivos o negativos en los **amenities** circundantes según su evaluación de estrellas en Google Maps. Agregado a lo anterior, desarrollaremos medidas especiales que sean aplicables a las encuestas CASEN de los periodos 2009, 2011, 2013, 2015 y 2017, con el objetivo de describir de mejor forma las condiciones socioeconómicas del entorno.

Finalmente, plantearemos modelos estadísticos que logren estimar el número de manifestantes o detenidos en función de las características propias del convocante y de su entorno, entre las variables a considerar incluiremos métricas topológicas de la red y variables CASEN personalizadas (indicadores de hacinamiento, estado de la vivienda, horas de trabajo semanales, ingreso per cápita, bienes propios, entre otros).

2. Hipótesis y Objetivos

El desarrollo de este trabajo requiere de múltiples hipótesis que pueden llegar a dar explicación al comportamiento de una organización manifestante, todo lo planteado considera que debemos asumir ciertos principios respecto a las organizaciones, entre estos destacamos:

- Las organizaciones manifestantes corresponden a grupos de dos o más personas que no necesariamente deben contar una estructura interna o jerárquica.
- Las organizaciones manifestantes pueden llegar a ser identificadas por medio de un nombre propio o un nombre relacionado a una categoría genérica.
- Las organizaciones manifestantes son evolutivas en el tiempo, su número de adherentes puede aumentar y con ello pueden llegar a definir su estructura, poseer métodos de elección interna y gestionar todo tipo de recursos para el desarrollo de sus actividades.
- Las organizaciones pueden relacionarse entre ellas según compartan ciertos propósitos.
- El número de convocados en una manifestación será asignado para todos los convocantes por igual.

Consideremos que el registro de nombre de una organización es responsabilidad del **COES**, quedando a criterio de ellos el uso de un nombre propio o genérico, a esto agregamos que el registro del número de manifestantes, heridos y muertos dependerá del equipo periodístico encargado de dar cobertura al evento. Si nos enfocamos en las organizaciones como una entidad autónoma, las principales hipótesis que describen las características de sus métricas y el comportamiento son:

- Existe una relación entre el ciclo político de un país y las acciones de protesta.
- Las noticias en tiempo real o las declaraciones emitidas por personas de interés público alteran el desarrollo de una manifestación, este impacto se relacionará con el tono de las declaraciones y la relevancia del sujeto declarante.
- Ciertas organizaciones construyen sus relaciones según la afinidad política de sus pares.
- El **degree** es principal variable de interés y está correlacionada con la mayoría de métricas descritas en 4.
- El grafo presenta el **rich club phenomenon** para ciertos nodos que no corresponden necesariamente a los de mayor **degree**.

Entre los principales objetivos de este trabajo, podemos detallar:

- Describir las métricas topológicas definidas en 4 y comparar la distribución de sus medidas en un instante final (año 2019).
- Diseñar y migrar las bases de datos con información obtenida a partir de los registros del **COES**, estos registros contendrán las métricas topológicas, las coordenadas de cada acto de protesta, información sobre los **amenities** cercanos y los registros de la encuesta **CASEN**.
- Describir las actividades comerciales cercanas a los focos de protestas e investigar el impacto que el desarrollo de una protesta genera en su entorno. Evaluar las variaciones en la dinámica de las protestas según el lugar que pertenece.

3. Metodología

Describiremos las fuentes involucradas y los métodos necesarios para procesar y analizar dicha información. Incluimos un glosario con la terminología usada para este proyecto.

3.1. Fuentes de información

Para las siguientes fuentes se incluirán sus respectivos repositorios en 10.1

- **COES:** Principal fuente de estudio, incluye el registro diario de los actos de protestas registrados por medios de comunicación físicos o digitales entre los años 2009 al 2019. El trabajo de registro se llevo a cabo mediante la revisión manual de un número limitado de medios de comunicación con el proposito de identificar aquellas noticias asociadas a protestas, debemos destacar que el registro de manifestaciones capturadas por RR.SS. es excluido. Consideremos que el trabajo de registro no es encargado a empresas externas, siendo integrantes del COES los responsables de este procedimiento.
- **CASEN:** Encuesta de caracterización socioeconómica del Ministerio de Desarrollo Social, cuyo propósito es modelar las condiciones socioeconómicas de los hogares en Chile, su ejecución es llevada a cabo por el mejor oferente de una licitación pública y su diseño es llevado a cabo por el Min. de Desarrollo Social, dicho diseño involucra la redacción de las preguntas y la elección de las alternativas disponibles como respuesta. En esta investigación se usaron variables presentes en todas las encuestas de los años 2009, 2011, 2013, 2015 y 2017.
- **REVIEWS:** Evaluación de establecimientos o servicios en escala de 0 a 5 estrellas en Google Maps. Estas evaluaciones se obtienen haciendo uso de **Places API Nearby Search** y las coordenadas de los focos de protestas. Este servicio de consulta dispone de los **amenities** cercanos al punto de protesta cuya distancia sea menor a cierto radio de distancia. Cada una de las evaluaciones contiene el nombre del local, las coordenadas de su ubicación geográfica y su evaluación en *estrellas*.
Las categorías evaluadas corresponden a **bares, restaurantes, doctor, parques, tiendas y policía**.

3.2. Glosario

La siguiente terminología tiene su motivación en la simplificación del nombre de las variables, como también en la explicación de aquellas variables CASEN y COES que son de mayor importancia.

3.2.1. Variables generales

- **amenities:** Establecimientos comerciales.
- **neighborhood:** Vecindad de un nodo en el grafo.
- **path:** Camino entre dos nodos de un grafo.
- **transitive:** Transitivo, parcial.
- **endpoints:** Nodos extremos de un camino.
- **Pearson media:** Correlación de Pearson de las medias de cada intervalo (binning).
- **Triángulos:** Unión de tres nodos conectados entre si y pertenecientes a un mismo grafo.

3.2.2. Variables COES

- **p6**: Código INE región.
- **p7**: Código INE provincia.
- **p8**: Código INE comuna.
- **p9**: Lugar específico del acto de protesta, sea este un nombre de calle, villa o sector específico.
- **p11**: Número de manifestantes.
- **p17(a,b,c,d,e,f)** : Nombre de las organizaciones convocantes.
- **p23**: Número de detenidos.
- **p26**: Número de heridos.
- **p28**: Número de muertos.

3.2.3. Nomenclatura de grafos

- **GT**: Grafo Total, incluye todos los actos de protestas registrados por el COES entre los años 2009 y 2019.
- **GCP**: Grafo Componente Principal, corresponde a la componente conexa de mayor extensión contenida en el grafo GT.

3.3. Procesamiento de la información

El proceso de construcción del grafo requiere de los datos del **COES**, cada fila en la base describe un acto de protesta registrado. Las columnas **p17a:p17f** detallan los nombres de las organizaciones convocantes, también puede darse el caso que no existan organizaciones convocantes a una protesta.

ORG1	ORG2	ORG3	ORG4	ORG5	ORG6	IDPROTESTA
------	------	------	------	------	------	------------

Esta estructura corresponde a la forma de las filas del **COES** y permite establecer un vínculo entre las organizaciones convocantes y los actos de protesta, dichas conexiones pueden representarse por medio de un **grafo bipartito** mediante una **proyección bipartita** que permite relacionar organizaciones convocantes entre sí.

En los casos que las variables del numero de detenidos, heridos o muertos de un acto de protesta sean vacíos, dichos valores serán ignorados para el cálculo de promedios, no así su participación en el grafo y las conexiones que establecen, las que serán consideradas de igual forma.

Para aquellas opiniones o **reviews** donde este especificada la variable **p9**, las coordenadas que usaremos corresponderán a las que obtengamos a partir de **p8**, este resultado dependera exclusivamente de la respuesta del servicio **API**.

4. Definiciones

4.0.1. Degree

En base a la definición de Diestel (2005), dado un grafo $G = (V, U)$ y un vértice $v \in V$, se considera su degree:

$$d_G(v) = |E(v)|$$

donde $|E(v)|$ es el número de edges conectados al vértice v .

Consideremos el average degree de un grafo $G = (V, U)$:

$$d(G) := \frac{1}{|V|} \sum_{v \in V} d_G(v)$$

El número de nodos de un grafo se define por $|V|$ y el número de edges por $|E|$.

4.0.2. Mean neighborhood degree

Dado un grafo $G = (V, U)$ y un nodo $v \in V$, se considera su vecindario : $N(v) = \{u \in V : (u, v) \in E\}$, definimos el grado promedio del vecindario como:

$$d_{ngh}(v) = \frac{1}{|N(v)|} \sum_{u \in N(v)} d_G(u)$$

4.0.3. Clustering coefficient

En base a la definición de Schank and Wagner (2005), dado tres actores $u, v \wedge w$ con una relación entre los nodos $v \wedge u$ como $v \wedge w$, el coeficiente de clustering busca representar la conexión entre los nodos $u \wedge w$.

$$c(v) := \frac{m(v)}{t(v)}$$

Donde $m(v) := |\{\{u, w\} \in E : \{v, u\} \in E \wedge \{v, w\} \in E\}|$ $t(v) := d_G(v)(d_G(v) - 1)/2$

4.0.4. Mean Clustering coefficient

Dado un grafo $G = (V, U)$ y un nodo $v \in V$, se considera su vecindario : $N(v) = \{u \in V : (u, v) \in E\}$, definimos el average clustering coef. como:

$$c_{avg}(v) = \frac{1}{|N(v)|} \sum_{u \in N(v)} c(u)$$

4.0.5. Betweenness centrality

En base a la definición de Gago et al. (2012), dado un grafo $G = (V, U)$ y un vertice $v \in V$ Sea $\sigma_{u,v}$ como el número de *shortest path* entre los vertices u y v , por otro lado, definamos $\sigma_{u,v}(x)$ como el número de *shortest path* entre los vertices u y v , donde $x \neq u, v$, por otro lado, sea $b_{u,v}(x) = \sigma_{u,v}(x)/\sigma_{u,v}$. Definase el betweenness centrality de un nodo x como:

$$B_C(x) = \sum_{(u,v) \in V^2} b_{u,v}(x)$$

4.0.6. Eigenvector centrality

En base a la definición de Gago et al. (2012), dado un grafo $G = (V, U)$, definamos la matriz de adyacencia de un grafo G como:

$$A_{ij} = \begin{cases} 1 & \text{si existe un link entre los nodos } i \wedge j \\ 0 & \text{en caso contrario} \end{cases}$$

4.0.7. Bridging centrality

En base a la definición de Liu et al. (2019), dado un grafo $G = (V, U)$ y un nodo $v \in V$, se considera su vecindario : $N(v) = \{u \in V : (u, v) \in E\}$, definimos el *bridging coefficient* como:

$$BC(v) = \frac{\frac{1}{d_G(v)}}{\sum_{u \in N(v)} \frac{1}{d_G(u)}}$$

A partir de lo anterior, definimos el bridging centrality como:

$$C_R(v) = BC(v)B_C(v)$$

4.0.8. Average path length

Dado un grafo $G = (V, U)$, definimos la distancia promedio de los elementos de un grafo hacia los nodos de un grado k como:

$$\langle L(k) \rangle = \frac{1}{|P(k)|} \sum_{i, j \in P(k) \wedge i \neq j} L(i, j)$$

donde $P(k) = \{\{i, j\} \in V \times V : d(i) = k\}$ y $L(i, j)$ es la distancia entre los nodos i y j .

4.0.9. maxelement & removemax

Definamos las siguientes operaciones sobre conjuntos:

$$maxelement(W) = \{\exists u \in W \forall a \in W : d(u) \geq d(a)\}$$

$$removemax(W) = W \setminus maxelement(W)$$

4.0.10. Asortatividad & Disortatividad

Dado un grafo $G = (V, U)$, $v \in V$, el nodo v se considera asortativo si se conecta en promedio con nodos de mayor grado que él, vale decir:

$$d_G(v) < d_{ngh}(v)$$

Consideramos la disortatividad como la condición contraria a la asortatividad.

4.0.11. Rich Club

En base a la definición de Xu et al. (2007) el de rich-club spectrum $\phi(k)$ definido como el ratio del número de edges entre nodos con un grado $\geq k$ dividido entre el número total de edges posibles entre estos nodos con grado $\geq k$. Dado un grafo $G = (V, U)$ sea:

$$\phi(k) = \frac{2|\{(u, v) \in V^2 : \max(d_G(u), d_G(v)) \geq k\}|}{(|\{u \in U : d_G(u) \geq k\}|)(|\{u \in U : d_G(u) \geq k\}| - 1)}$$

Consideremos que para el siguiente caso es necesario una normalización del coeficiente de rich-club y para ello consideremos el siguiente ratio:

$$\rho(k) = \frac{\phi(k)}{\phi_{unc}(k)}$$

Donde la función analítica $\phi_{unc}(k)$ definida en Colizza et al. (2006) es igual a:

$$\phi(k) = \frac{N \langle k \rangle \int_{x=k}^{x=k_{max}} du \int_{x=k}^{x=k_{max}} P(u, v) dv}{\left[N \int_{x=k}^{x=k_{max}} P(u) du \right] \left[N \int_{x=k}^{x=k_{max}} P(u) du - 1 \right]}$$

5. Estructura topológica de la red

La relación de los agentes convocantes entre los años 2009 y 2019 será descrita mediante el uso de dos grafos: GT y GCP, en números generales, el registro del COES contiene 23398 actos de protesta, de los cuales GT posee 10748 *nodes* y 8449 *edges*, por otra parte, GCP posee 2323 *nodes* y 5845 *edges*, destaquemos que un nodo en GCP posee 3.2 veces más conexiones que un nodo de GT.

Por sus condiciones de compacidad, verificaremos si se cumplen ciertas condiciones en GCP, como que sea **free scale network**, según las condiciones descritas en (Barabasi and Albert (1999)) dicha condición se obtiene si la distribución de grados se ajusta a una *power law* con un valor de $2 < \lambda < 3$. Dada la función $p(k) = k^{-\lambda}$, los valores de ajuste de GCP corresponden a $\lambda_{gcp} = 2,54$, $x_{min} = 8$, a modo de comparación, la red que conecta actores según las películas en que participan posee un $\lambda_{actor} = 2,3$, y la red de conectividad entre paginas web (WWW) posee un $\lambda_{www} = 2,1$.

Según el modo de construcción del grafo, el *degree* de un nodo manifestante dependerá exclusivamente de las conexiones que genere e ignorará todas las circunstancias y consecuencias que tuvo el desarrollo de la protesta. Comparando los grafos GT y GCP según sus distribuciones en [1], su principal diferencia radica en el ruido, lo que se traduce en un número no menor de organizaciones cuyo accionar es limitado, ya sea porque no logran articular una estructura interna organizacional o solo nacen como respuesta a necesidades específicas de una comunidad, a modo de ejemplo una escuela de un sector rural convoca a una protesta por las condiciones de infraestructura del establecimiento, dicha organización responde a la necesidad de una comunidad, pero la permanencia de esta organización dependerá exclusivamente de la respuesta a esta demanda. El ruido contiene todos aquellos nodos que no pertenecen a la componente principal del grafo, que en su mayoría corresponden a nodos de bajo **degree** cuyas ubicaciones geográficas no están relacionadas con las principales ciudades del país.

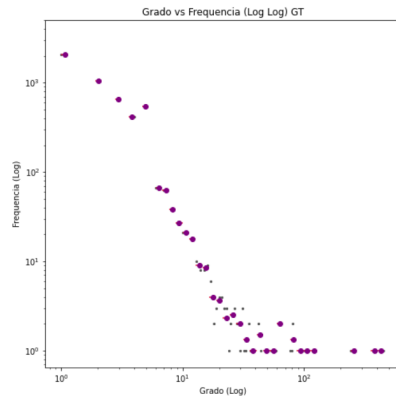


Fig. 1: Distribución degree GT

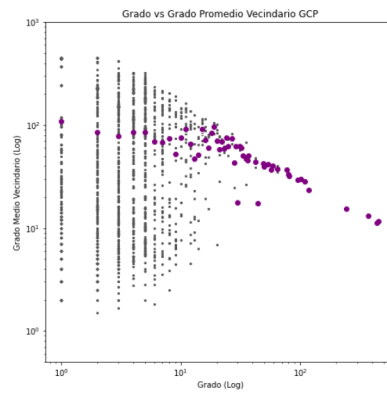


Fig. 2: Distribución degree GCP

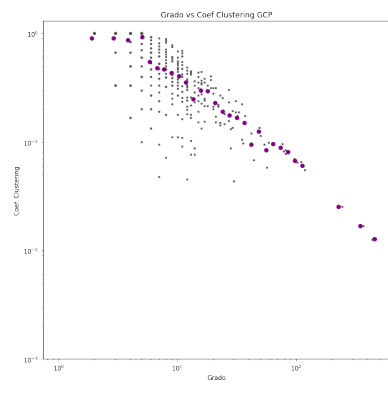


Fig. 3: Degree vs Clustering Coef. GCP

El decaimiento de la curva en [1] se asocia a la vida útil de una organización, aquellos nodos que perduran en el tiempo son capaces de responder a la contingencia nacional mediante la movilización de recursos, esto se hace posible por medio de **bases** que actúan en terreno, de ahí que estos nodos posean mecanismos de democracia y participación interna. Consideremos que el año 2011 participaron 11 veces más organizaciones respecto al año anterior y el año 2019 participaron 1.7 veces más organizaciones que el año 2018. Cabe destacar que las diez organizaciones de mayor *degree* en GT son: **CUT** (450), **ANEF** (440), **COLEGIO DE PROFESORES** (370), **CONFECH** (244), **FENATS** (118), **CONFUSAM** (110), **COORDINADORA NO+AFP** (102), **ACES** (96), **FENPRUSS** (81) y **ESTUDIANTES** (81).

Si consideramos a GT y GCP como una red social, otro fenómeno característico es **small world effect**, dicha condición se cumple si el *average path length* es menor que una red aleatoria y el *clustering coefficient* es considerablemente mayor al de una red aleatoria Watts et al. (2006).

Dada la red GCP, observamos en 3 un decaimiento del *clustering coefficient* a medida que aumenta el grado, lo que es común en redes que con una organización jerárquica, donde un número reducido de nodos toma una posición predominante e influyente en las decisiones del resto de nodos. Si la distribución de valores del *clustering coefficient* se ajusta a $C(v) \sim k^{-1}$ sería un indicador de una topología jerárquica Barabasi et al. (2004). A niveles prácticos, los nodos de bajo *degree* tienden a participar en pocas protestas y como cada registro del COES puede llevar a lo máximo seis participantes, todos ellos estarán relacionados entre sí, lo que implica que el *clustering coefficient* sea igual o cercano a uno. Por otro lado, los nodos de alto *degree* son organizaciones que han participado de múltiples actos de protesta y que sus **vecinos** no necesariamente están relacionados entre ellos.

Consideremos que las propiedades de **scale free network** y **high clustering** pueden coexistir al interior de una red, en particular, el grafo GCP posee un *average path length* $l_{gcp} = 3,724$ y *clustering coefficient* $C_{gcp} = 0,622$ y dada una red aleatoria ($N = 2323$) sus valores de $l_{gcprand} = 4,27$ y $C_{gcprand} = 0,0099$. indican que el grafo cumple la condiciones necesarias de **small world** ($C_{gcp} \gg C_{gcprand}$ y $l_{gcp} < l_{gcprand}$).

En términos relacionales y según lo que observamos en 2, los nodos del grafo **GCP** poseen patrones de asortativos y disortativos [4.0.10], cuyo cohorte se da en *degree* igual a 45. Los nodos v cuyo $d_G(v) \leq 45$ tienden a relacionarse en su entorno con nodos de mayor grado y aquellos nodos v que $45 \leq d_G(v)$ se conectan con nodos de menor grado. Múltiples factores pueden dar explicación a ello, como por ejemplo que ciertas organizaciones de alto **degree** se relacionan con nodos que asumen un rol de *satélite*.

La centralidad de los nodos en ambos grafos mejora en la medida que aumenta su grado, esto refleja que las organizaciones de mayor grado no se aíslan del entorno y evitan conectarse con nodos *impopulares* cuya actividad no persiste en el tiempo. La actividad de un nodo requiere de una causa demandante, un momento inicial, una respuesta por parte de los aludidos y de un momento concluyente o final. Toda actividad puede extenderse en el tiempo en la medida que la organización demandante sea capaz de presionar durante un periodo hasta lograr una solución a sus demandas.

De acuerdo a 4.0.8, aquella función que mejor describe la distancia hacia un nodo según su **degree** es $L(k) = 4,09 - 0,34 * \log(x)$ ($R^2 = 0,96$) según lo que observamos en 4.

Si agrupamos los valores de **degree** y **p11** según la variable **p8**, podemos observar que aquellas comunas con mayor correlación son: **Ovalle** [0.814], **La Reina** [0.785], **Talcahuano** [0.773], **Estación central** [0.716] y **Osorno** [0.589]. La correlación entre las variables **p11** y **p23** según sus medias es 0.934, lo que lleva a pensar que la presencia de manifestantes trae como consecuencia la presencia de actos vandálicos. Destaquemos que las correlaciones entre todos los valores de **degree** y **p11** [Pearson media: 0.2356] y **degree** y **p23** [Pearson media: 0.357].

De acuerdo a 5, la distribución de distancias entre los vértices de GCP alcanza su mayor valor para $d = 4$ y guarda estrecha relación con 4.

Para profundizar en las características de los nodos presentes en los *neighborhoods* y como distribuyen sus grados, consideremos las siguientes definiciones de acuerdo a 6:

- **h0** (rojo): **nhbd** mean degree.
- **h1** (morado): **nhbd** mean degree excluyendo el nodo de mayor grado.
- **h2** (amarillo): **nhbd** mean degree excluyendo los dos nodos de mayor grado.

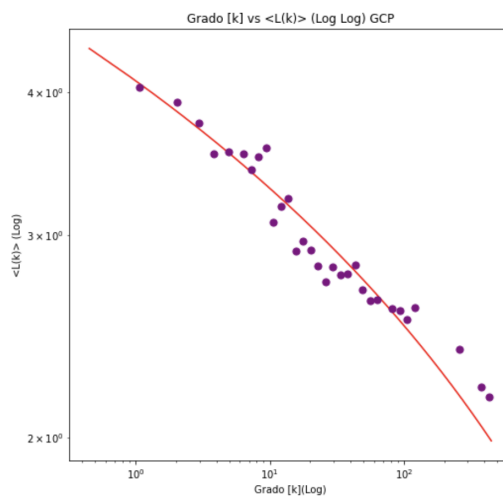


Fig. 4: Shortest path length distribution

Donde la exclusión del nodo de mayor **degree** se ajustará según la definición de *removemax* [4.0.9].

De acuerdo a lo observado en 6, para aquellos nodos v donde $d(v) \leq 10$, los *degree* que integran el *neighborhood* no distribuyen de manera uniforme, sino que su promedio es dependiente del nodo con *degree* mas alto, de ahí el cambio de tendencia en las curvas de color **rojo** y **morado**, por otro lado, para los nodos v cuyo $d(v) \geq 50$ se puede observar una menor amplitud entre las curvas, lo que indica mayor uniformidad del *neighborhood* y que no existe mayor diferencia entre el promedio de los nodos y el mayor **degree** que integra el vecindario, la menor desviación estándar entre estos valores se observa en una mayor cercanía entre las curvas de color **rojo**, **morado** y **amarillo**.

De acuerdo a lo que observamos en 7 a medida que aumenta el $p11$ del nodo, aumenta el $p11$ de su entorno, esto podría deberse a múltiples razones, por ejemplo que los nodos con cierto poder de convocatoria (reflejado en la variable **p11**) tienden a conectarse con nodos de su mismo nivel, por otro lado, puede darse un fenómeno **parasitario** de nodos de bajo **degree** que están conectados a organizaciones con alto poder de convocatoria de **p11**, estos nodos de bajo rango no tienen la capacidad de convocar a un alto número de manifestantes, pero sus conexiones hacen que sean sobre representados.

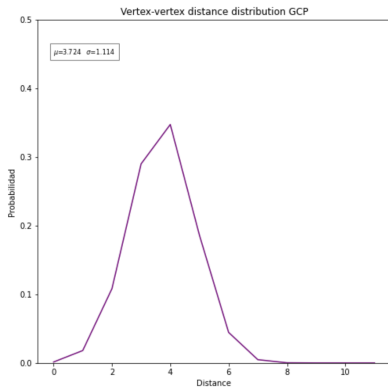


Fig. 5: Vertex-vertex distance GCP

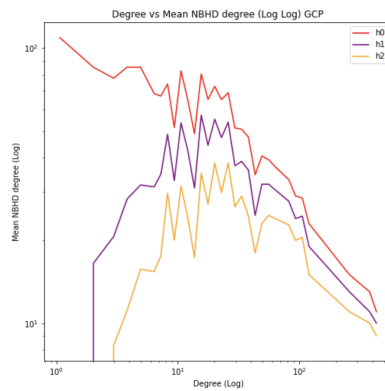


Fig. 6: Degree vs NBHD Degree GCP

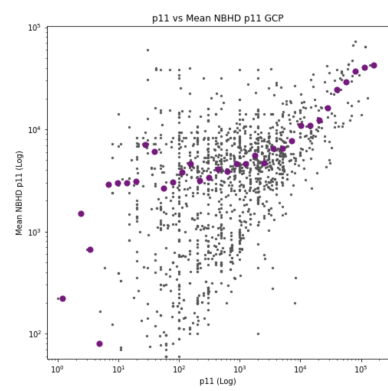


Fig. 7: p11 vs NBHD p11 GCP

En base a la metodología de construcción, a todos los convocantes de una manifestación se les asigna igual número de manifestantes convocados, aun así no debemos obviar el *factor de arrastre* que tienen ciertas organizaciones, ya que algunas son capaces de persuadir y convocar una mayor cantidad de adherentes en comparación a sus pares para un mismo acto de protesta.

5.1. Evolución y centralidad en la red

El principal propósito de esta sección es profundizar en la evolución del grafo a lo largo del tiempo, donde cada acto de protesta implica la adición de nodos y conexiones a GT.

Definamos la sucesión de grafos $G_t = (V_t, E_t)$, donde $G_t \subseteq G_{t+1}$. Cada G_t corresponde a un instante t ($t \leq 23398$) del grafo GT en que se añade un nuevo acto de protesta. Primeramente, evaluaremos la formación de *triángulos* en GT en dos instantes, aquel donde dos nodos ($org1, org2$) están conectados a un nodo $root$ pero no están conectados entre sí, y el instante posterior en que se conectan los nodos ($org1, org2$), para aquello definamos:

$$j = \min\{j : \{org1, org2, root\} \subseteq V_j \wedge \{\{org1, root\}, \{org2, root\}\} \subseteq E_j\}$$

$$k = \min\{k : \{org1, org2, root\} \subseteq V_k \wedge \{\{org1, root\}, \{org2, root\}, \{org2, org1\}\} \subseteq E_k\}$$

Consideremos $timediff = (k - j)$ como el número de actos de protesta necesarios para que se complete la formación de un triángulo. De acuerdo a lo observado en 8, donde comparamos el tiempo de formación de un *triángulo* respecto al *degree* del nodo $root$ en el instante final, observamos una alta densidad de nodos con *degree* menor 100, de lo cual podría concluirse:

- Los nodos de bajo grado tienden a formar más *triángulos* entre sus vecinos, dichas conexiones se crean con mayor velocidad en comparación con otros *root* de mayor grado. Posibles razones pueden ser que las organizaciones pequeñas solo responden a una consigna específica y filtran sus conexiones según sean de su misma clase.
- Los nodos de mayor grado tienen dificultad para formar *triángulos* con los elementos de su vecindario y el tiempo que toman en crear estos *links* tiende a ser mayor que el habitual. Posibles razones para ello puede ser la ubicación geográfica de los nodos, lo que dificulta la coincidencia en los actos.

En la figura 10 comparamos la tasa de crecimiento entre *root* y los nodos (*org1, org2*) en los instantes *j* y *k* (inicial y final respectivamente), para ello definase:

$$RP(m) = \frac{\text{degree}(\text{org1}_m) + \text{degree}(\text{org2}_m)}{2 * \text{degree}(\text{root}_m)}$$

Observamos que la tasa de crecimiento de los nodos (*org1, org2*) es mayor que el crecimiento del nodo *root* entre los instantes *j* y *k*, para analizar esto expondremos dos posibles causas:

- $\text{degree}(\text{root}) \leq 100$: Dado que $\text{timediff} \leq 4000$ y el grado del nodo *root* se estanca según lo que observamos en 9, se podría concluir que los nodos *org1* y *org2* son capaces de crear nuevas conexiones pero no coinciden entre ellos.
- $\text{degree}(\text{root}) \geq 100$: Dado que $\text{timediff} \leq 8000$ y el grado del nodo *root* crece de mucha mayor manera, podríamos asumir que las condiciones geográficas dificultan la generación de nuevas conexiones, dado que los nodos de alto grado tienen presencia en gran parte del territorio nacional, de ahí que el crecimiento de los nodos *org1* y *org2* implique su expansión geográfica.

De modo complementario y según lo que observamos en 9, a los nodos *root* de mayor grado les toma más tiempo completar sus triángulos en comparación a nodos de bajo **degree**, lo que refleja la diversificación de los actos, las organizaciones de alto grado buscan responder a las necesidades de la población.

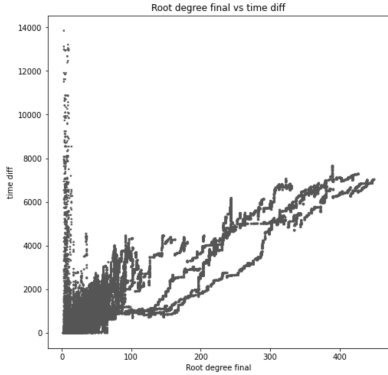


Fig. 8: Root degree final vs timediff

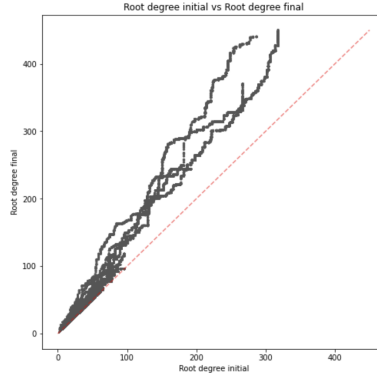


Fig. 9: (Root degree) initial vs final

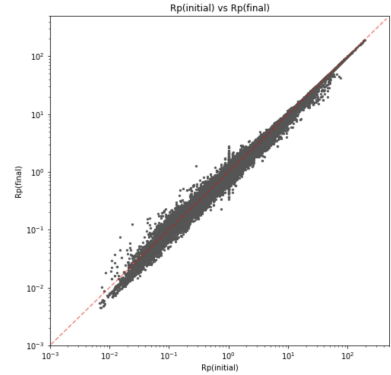


Fig. 10: Rpinitial vs Rpfinal

Dado un grafo $G = (V, U)$, y un *path* p de la sucesión de nodos $v_1, v_2, ..v_n \in G$, donde $(v_i, v_{i+1}) \in V$, definamos el **mean path degree** p_{mean} , el **mean endpoints degree** p_{end} y el **mean transitive degree** $p_{tsv,k}$ según lo siguiente:

$$p_{mean} = \frac{1}{|n|} \sum_{i \leq n} d(v_i)$$

$$p_{end} = \frac{d(v_1) + d(v_n)}{2}$$

$$p_{tsv,k} = \frac{1}{|k|} \sum_{i \leq k} d(v_i)$$

De acuerdo a lo observado en 12, en la medida que p_{mean} aumenta, el valor de p_{end} tiende a bajar, indicando que los caminos entre extremos de bajo grado contienen necesariamente nodos de alto grado, fiel reflejo de la buena centralidad de los nodos de mayor grado. Definamos el promedio transitivo de un *path* como el promedio parcial que se obtiene en la medida que se *camina* a lo largo del *path*. Acá el mayor promedio transitivo se encuentra en el $step = 2$, lo que guarda relación con 4 dado que la menor distancia hacia los nodos de grado más alto corresponde a **2**, vale destacar que a medida que aumenta el largo de un camino disminuye su promedio transitivo, de ahí que los nodos cuyo **degree** es más bajo tienden pertenecer a caminos de mayor extensión.

Otra medida importante que describe la centralidad e intermediación de los nodos es el **betweenness**, de acuerdo a lo que observamos en [11], en la medida que aumenta el **degree** también lo hace el **betweenness**.

Dado que las conexiones entre organizaciones perduran en el tiempo, podríamos pensar que las organizaciones son capaces de ejercer influencia sobre otras y que esto solo tiene explicación si se asume que la unidad básica de una organización son las personas. Si hipotéticamente removieramos los nodos con alto **betweenness** no solo se verían afectadas las conexiones de la red, sino que incluso se dificultaría el progreso ciertas manifestaciones, ya que se ven involucrados recursos fundamentales (publicidad, transporte, seguridad, exhibición artística).

Acorde a lo descrito anteriormente, la red de protesta es **libre de escala**, esta condición es transitiva para la distribución de **betweenness**, si consideramos la siguiente función:

$$p_b(g) \sim g^{-\eta}$$

El valor de $\eta = 2,15414$ es el de mejor ajuste.

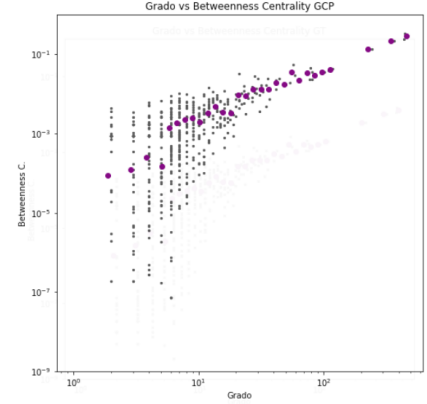


Fig. 11: Degree vs Betweenness C. GCP

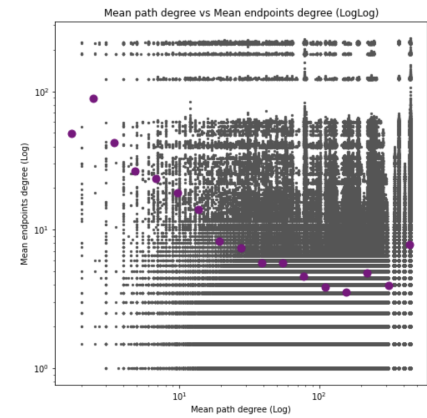


Fig. 12: p_{mean} vs p_{end}

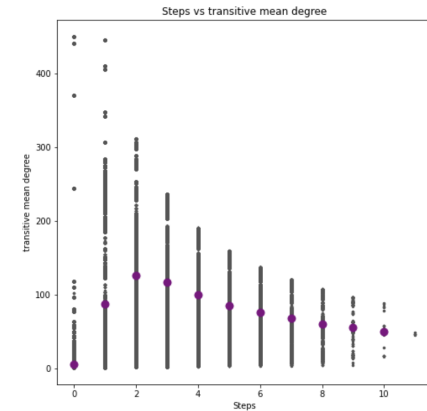


Fig. 13: Steps vs transitive mean GCP

5.2. Rich club phenomenon

El **rich club phenomenon** corresponde a una propiedad topológica presente en ciertos grafos, donde los **rich nodes** están estrechamente conectados con otros **rich nodes**. La condición de **rich** dependerá exclusivamente del contexto de la red, a modo de ejemplo; en una red de conexión de estudiantes universitarios en EE.UU. la condición de **rich** corresponderá al prestigio de la universidad. Para los efectos del grafo de protestas, la condición de **rich** dependerá del **degree** del nodo.

Según la definición de *rich club spectrum* descrita en la sección 4.0.11, observamos un crecimiento monótono de $\phi(k)$ a medida que aumenta el *degree* k , esto podría indicar la presencia del *rich club phenomenon* dentro del grafo GT, dicho fenómeno propone que los nodos de mayor grado tienden a crear relaciones exclusivas con nodos de su misma o mayor condición, dicha condición se cumple para todo nodo cuyo **rich-club** sea mayor que **1**. Detallamos que la condición de *spectrum* mayor a 1 no implica necesariamente que dichos nodos tengan un **rich-club** mayor a 1.

Cada uno de las siguientes distribuciones incluirá a los grafos GT, GCP y GFILT, donde sus variables \mathbf{x} e \mathbf{y} corresponderá a los valores de **degree** y **rich club** respectivamente, y el tamaño del círculo dependerá de las variables **p11**, **p23** y **p26**. Adicionalmente, se analizarán los valores del nodo y el promedio de su entorno, dichos valores serán distinguibles por medio de la circunscripción de dos círculos, donde el valor más alto tomará posición en los extremos.

En general, el valor del *rich-club* disminuye a medida que aumenta el grado, esto podría deberse a la asortatividad y disortatividad de los nodos, ya que los nodos no tienden a conectarse con sus semejantes. Según lo observado en la figura [15] los nodos de menor *degree* tienden a convocar menos que su entorno y en la medida que evaluamos nodos de mayor *degree*, el poder de convocatoria de estas organizaciones aumenta y es mucho mayor en comparación a su entorno. Debemos destacar que el *rich-club* como medida no asegura el éxito en el número de convocados, de ahí que no debemos confundir popularidad con exclusividad.

Situación similar ocurre en [16], no se observan tendencias que hagan suponer un peor comportamiento de los manifestantes según su valor de **rich club**. Los puntos que representan los nodos de mayor **degree** poseen bordes en sus círculos más grandes, lo que sugiere que la proporción entre el número de detenidos según sea el nodo principal y su entorno es mayor a la proporción de estos según su número de convocados. Para cada una de las distribuciones presentadas, observamos que los nodos cuyo *rich club* es mayor que **1** poseen un comportamiento ejemplar en comparación al grafo, ya que su número de manifestantes, detenidos y heridos [17] es mucho menor al resto del grafo.

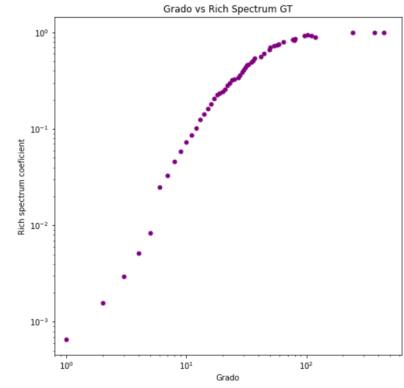


Fig. 14: Degree vs Rich Spectrum GT

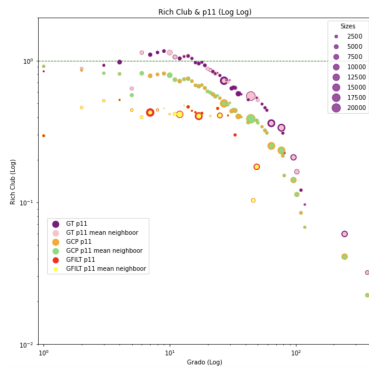


Fig. 15: Degree vs Rich club (p11)

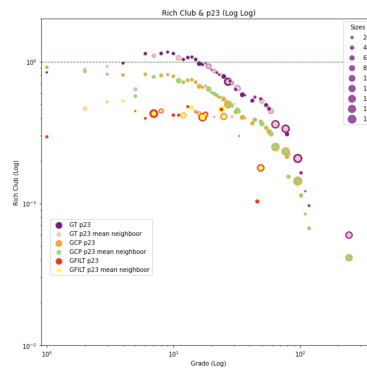


Fig. 16: Degree vs Rich club (p23)

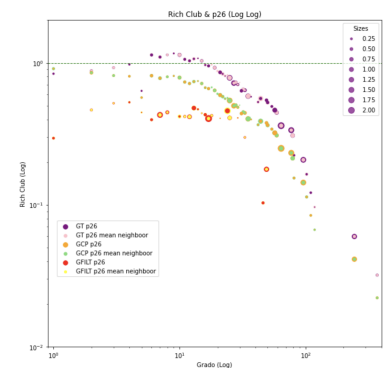


Fig. 17: Degree vs Rich club (p26)

6. Correlaciones

La incorporación de las encuestas CASEN como medida agregada al grafo nos lleva a identificar todas aquellas correlaciones respecto a las medidas del COES. Para el cálculo de las correlaciones se considera como unidad atómica a las organizaciones, quienes ya poseen información propia del COES, por otro lado, dado que las medidas de la encuesta CASEN se agrupan de acuerdo a **p8** (comuna), el valor de la variable corresponderá al promedio ponderado de los valores CASEN según el número de ocurrencias por comuna, para ello proceso asumimos que los participantes de una manifestación son residentes de la comuna donde sucede la manifestación, por ello compartirán las características socioeconómicas registradas en la encuesta. Entre las variables evaluadas, destacamos las siguientes:

6.1. n° protestas / uso de internet

El número de protestas corresponde a la cantidad de manifestaciones en que participa el nodo seleccionado, dicho valor no es lo mismo que el **degree**, dado que mide el número de conexiones que el nodo establece. Para el cálculo de estas correlaciones, consideramos cinco grafos temporales, los cuales contendrán los actos de protestas generados hasta los años 2009, 2011, 2013, 2015 y 2017.

El uso de internet corresponde a una variable numérica, resultante de la ponderación de las respuestas a las siguientes preguntas CASEN:

- ¿Con que frecuencia usa internet?
- ¿Usó internet para buscar información de buscadores como Google u otros?
- ¿Usó internet para informarse por prensa digital o redes sociales?
- ¿Usó internet para entretenimiento o comunicación?
- ¿Tiene conexión a internet móvil, plan de telefonía o banda ancha?

Según avanzan los periodos de evaluación, observamos una mayor correlación entre las variables, inevitablemente la facilidad para contratar planes de internet influye, como también la disminución en los precios de las computadoras y equipos móviles. Para el año 2013 el 86 % de los jóvenes chilenos tenía una cuenta de Facebook y el 52 % se conectaba al menos una vez al día Cabalín-Quijada (2014), por el uso de redes sociales ha sido primordial en la transmisión de contenido por parte de organizaciones.

En virtud a lo expuesto en el estudio **El Chile que viene: Medios y redes sociales** de la encuestadora CADEM (2022) desde Noviembre del 2017 a Enero del 2020 el uso de **WhatsApp** sube de un 58 % a un 72 %, el uso de **Facebook** aumenta de un 48 % a 67 % y el uso de **Twitter** de un 10 % a un 16 %.

2009	2011	2013	2015	2017
0.136175	0.249296	0.2952609	0.348993	0.353473

6.2. n° detenidos/ uso de internet

El número de detenidos corresponde a la cantidad de personas arrestadas promedio para cada organización según cada uno de los cinco periodos donde se evalúa el grafo. El número de detenidos y el uso de internet se relacionan tanto en los días previos a una manifestación como en el final de esta. Podríamos sugerir que el *tono* de la protesta dependerá de los ánimos de los convocantes, quienes son capaces de hacer llamados a la violencia, sean explícitos o no, pueden exacerbar a grupos extremistas o insurgentes. Consideremos que todos los registros audiovisuales vinculados a un acto de protesta que contengan maltratos, detenciones violentas o actos de represión por parte de la policía pueden llevar a una reacción aún más violenta por parte de los manifestantes. Consideremos que en Chile no es común

la aplicación de querrelas por **Incitación al odio, a la subversión**, o la aplicación de la **Ley de Seguridad Interior del Estado**.

La violencia no es tangible sino por medio de sus consecuencias, sean estos daños a la propiedad, dificultad para el libre tránsito o ejercicio de una actividad económica o agresiones físicas hacia civiles o personal policial, de ahí que el único indicador que nos permite estimar dicha violencia es el número de detenidos.

2009	2011	2013	2015	2017
0.1591664	0.268456	0.406227	0.321558	0.441284

6.3. escolaridad / n° detenidos

La escolaridad corresponde al número de años cursados por la persona encuestada desde la enseñanza pre básica hasta su máximo nivel educacional, dicha variable responde a la siguiente pregunta de la encuesta CASEN:

- ¿Cuál fue el nivel educacional más alto alcanzado o el nivel educacional actual?

El nivel de escolaridad está altamente conectado al número de arrestos en el sector, podríamos plantear múltiples hipótesis que den explicación a ello, que todo acto de protesta necesariamente conduce a sus participantes a cuestionar el estado de la materia demandante, este espíritu crítico podría ser más frecuente en personas de mejor formación educacional, eso si, no podemos determinar si un mejor nivel educacional implica una mayor probabilidad de realización de actos delictivos, dado que la detención por sí sola puede deberse a estrategias de **copamiento** de un sector o a mejorar en las tácticas policiales, lo que lleva a detenciones masivas. Destaquemos que la implementación de la jornada escolar completa ha ayudado a controlar los delitos asociados a la propiedad, pero han aumentado los delitos con violencia Berthelon and Kruger (2011).

2009	2011	2013	2015	2017
0.253576	0.458321	0.4410784	0.3216197	0.46707

6.4. participación política / n° protestas

La variable de participación política está relacionada a la siguiente pregunta CASEN:

- En los últimos 12 meses, ¿ Ha participado de alguna de las siguientes organizaciones?

Entre las alternativas encontramos una variedad de agrupaciones, ya sean vecinales (junta de vecinos, club del adulto mayor), recreativos (club de deportes), religiosos (congregación, iglesia) o políticos (partido político). Para el cálculo de esta variable consideraremos el número de organizaciones en las que participa el encuestado.

En la medida que las personas se involucran en múltiples organizaciones, tiende a aumentar el número de actos de protesta en los que participan, esto podría relacionarse con que las organizaciones, independiente del propósito por el que fueron constituidas, buscan articular respuestas a sus propias problemáticas.

Dentro de las múltiples organizaciones registradas por la CASEN, cierto número de ellas tendrá un propósito formativo, dado que brindarán contenido valórico, emocional y educativo a sus integrantes, dichos rasgos pueden llegar a influir en la capacidad de discernir del individuo y esto podría llevarlo, en algún momento, a decidir que la manifestación pública es una vía válida de protesta.

2009	2011	2013	2015	2017
0.176606	0.6728182	0.692853	0.40017	0.517786

6.5. escolaridad / subsidios y bonos

La variable CASEN de subsidios y bonos corresponde a los ingresos registrados en los campos **subsidios monetarios**, estos incluyen los aportes subsidiarios provenientes de las municipalidades y gobierno central.

Para los años 2009, 2011 y 2013 las correlaciones toman valores de 0.509, 0.764 y 0.772 respectivamente, lo que indicaría que se entregan más recursos a aquellas familias cuyos nivel educacional es mayor. Este grupo podría ser catalogado como **Clase media**, y ellos tienen la posibilidad de acceder a subsidios para la compra de viviendas (MINVU), subsidios para el creación y desarrollo de una idea de negocio (CORFO), tienen acceso créditos con garantía estatal para cursar estudios superiores.

El hecho que desde el 2015 las correlaciones pasen a ser negativas indica un revés en la focalización de las ayudas sociales, todo esto motivado por los cambios de gobierno. Esta nueva política no excluye de los beneficios a los grupos de menor escolaridad, sino que les brinda mayor asistencia y ayuda económica.

2009	2011	2013	2015	2017
0.50916	0.764986	0.77259	-0.16779	-0.55549

6.6. ayudas escolares/ bienes propios

La variable que representa las ayudas escolares corresponde a un promedio ponderado de las siguientes preguntas en la encuesta **CASEN**:

- En el periodo escolar ¿recibe alimentación gratuita?
- ¿ Ha recibido útiles escolares, texto o material complementario por medio del establecimiento educacional?
- ¿ Ha recibido atención médica o dental en el establecimiento ?

Cada una de estas preguntas fue aplicada a los tres niveles de educación: pre-básica, básica y media. Vale destacar que en los distintos periodos entre 2009 y 2017 la correlación permanece constante, lo que se explicaría por las políticas de focalización en la entrega de ayudas no monetarias y también a la efectividad en su propia implementación, debido a que el filtro que determina quienes son elegibles para estos beneficios se obtiene mediante el registro de la **Ficha de Protección Social** en cada municipio.

2009	2011	2013	2015	2017
-0.76716	-0.771749	-0.748273	-0.734328	-0.749258

7. Superficies

Cada una de las manifestaciones registradas por el COES contiene características geográficas del acto de protesta, además de los códigos INE por comuna, provincia y región, se incluye un campo de **dirección explícita** que corresponde al nombre con que se conoce el lugar, puede ser el nombre de una calle, una intersección, el nombre de un lugar emblemático, villa o población. Con los datos anteriores es posible obtener las coordenadas de la ubicación haciendo uso de **Place search Api**.

Cada una de las coordenadas tendrá la forma $C_n := \{(lat_n, lng_n)\}$ donde n se asocia al acto de protesta, con estos puntos es posible definir el $Pol(org_j)$ como el polígono convexo que contiene todas aquellos focos de protesta de la org_j según las siguientes condiciones:

- **Regionalidad:** El polígono se circunscribe al interior de los límites geográficos de su propia región.
- **Temporalidad:** Los actos de protesta se limitan a espacios limitados de tiempo, que pueden variar entre los años 2009 y 2019.

Para los propósitos de esta sección, consideraremos la subdivisión de manifestaciones por año y serán limitadas a su región, con esto describiremos la influencia que tiene una organización en un territorio según el espacio que cubre.

Segun observamos en [16], el grado y la superficie según sus medias están estrechamente relacionados [0.8 Pearson], lo que implicaría que el aumento de grado en una organización lleva a un crecimiento de sus bases, ya sean estos militantes o simpatizantes, solo esto permite que una organización logre tener presencia regional, por otro lado, una hipótesis del **acarreo** de manifestantes a múltiples puntos dentro de una región no tendría sentido, ya que es necesario conocer el lugar, sus características y a sus habitantes, sin incluir el costo físico y monetario que implicaría el traslado de personas.

La influencia territorial no ha sido desplazada por la influencia de medios, sirviendo como una herramienta para llegar a los rincones más aislados del país.

Para el caso de la Región Metropolitana, la alta densidad poblacional conlleva una inercia de los focos de protesta, [Pearson 0.31], en palabras simples, se repiten los mismos lugares para protestar, caso similar sucede con la Región de Valparaíso [Pearson 0.34], en contraste, aquellas regiones cuya expansión geográfica es mucho mayor y además contienen múltiples ciudades que toman el rol de **gubernaciones provinciales** el nexo entre área y grado es mayor: Antofagasta [0.6443 Pearson], Los Lagos [Pearson 0.7967].

Consideremos la figura [19] que comparará la superficie de influencia de los nodos agrupados de forma regional según sus percentiles 50, 80, 90 y 95. Observamos cierta similitud en los percentiles 80, 90 y 95 de las superficies regionales 5-6-7-8 (Valparaíso, O'Higgins, Maule y Bio Bio), esto nos lleva a plantear que existiría un patrón interregional para la ocupación de espacios, dado que los nodos mejor posicionados de cada región coinciden en sus cifras, esto podría deberse a dos razones:

- **Planificación nacional de organizaciones:** Los nodos de alto grado destinan recursos para lograr un despliegue simultáneo en distintas regiones.

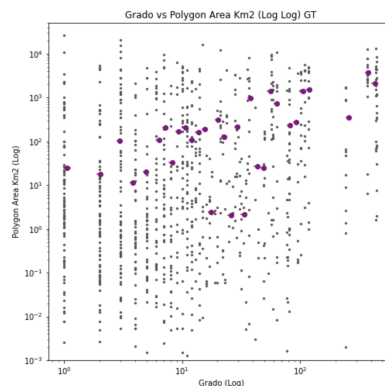


Fig. 18: Degree vs Superficie

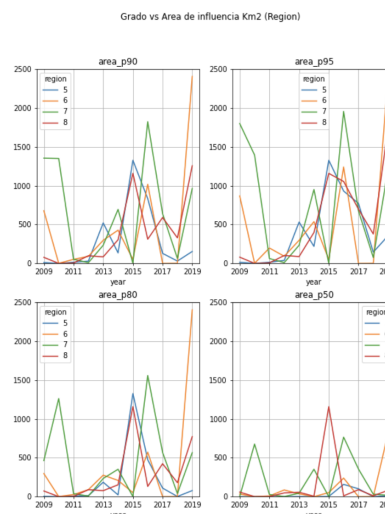


Fig. 19: Degree vs Superficie (Region)

- Coincidencias geográficas del entorno: Las distancias entre las distintas ciudades son semejantes por región, lo que implica una cobertura similar.

La definición del *neighborhood* considera como vecinos de un nodo a todos aquellos con los que posea al menos una conexión. Para replicar la idea de un vecindario según la ubicación geográfica es que consideraremos la siguiente relación:

Dado un acto de protesta t , toda organización que haya convocado en el mismo año del acto t y que haya participado de al menos una manifestación cuya distancia al acto t sea menor o igual al radio r planteado, se considerará como parte del vecindario de los nodos convocantes del acto de protesta t . Los radios usados tendrán los valores de **100mt**, **300mt**, **500mt**, **1km** y **2km**. El valor obtenido de cada una de las circunferencias corresponderá al promedio de **degree**, **p11** y **p23**.

La figura [21] compara el grado de un nodo respecto al promedio de su vecindario geográfico, acá observamos que en los vecindarios de un radio de **100mt** el valor **degree** promedio es mayor y a medida que aumenta el radio tiende a disminuir el **degree** promedio del entorno, esto sugiere que los nodos de mayor **degree** tienden a situarse en espacios de alta densidad de organizaciones, tal como si se tratara de una **disputa por los espacios emblemáticos**, por otro lado, los nodos de menor **degree** se ubican de forma periférica a las grandes concentraciones, estos nodos podrían ser marginados por otras organizaciones o simplemente la construcción de una identidad propia los lleva a diferenciarse del resto.

Similar a lo anterior se observa en la figura [21], dado que para **degree** menores a 100, a medida que el radio disminuye, la variable **p11** aumenta, guardando estricta relación con la idea del uso de **espacios emblemáticos**. La explicación más simple para lo anterior es que a medida que aumenta el **degree** también lo hace **p11**, pero esta última variable podría verse afectada en su recopilación al asumir la **disputa por los espacios emblemáticos**, dado que en estos sectores converge población flotante, comercio formal e informal y servicios de interés público, lo que los hace altamente concurridos.

Por otro lado, observamos que en [22] a medida que aumenta el grado del vecindario, aumenta el poder convocatorio de los nodos contenidos en el radio, pero llega un momento en que el poder convocatorio se estanca y baja (cerca de un valor 100), lo que indica que un entorno a medida que su grado promedio crece, tiene mayor poder convocatorio, pero ese poder se estanca.

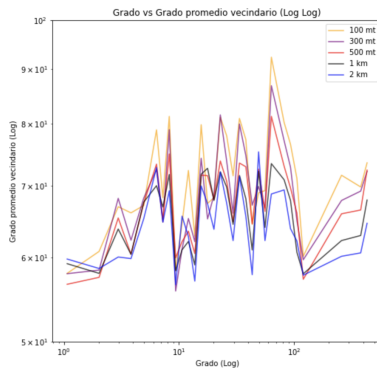


Fig. 20: Degree vs Degree vecindario

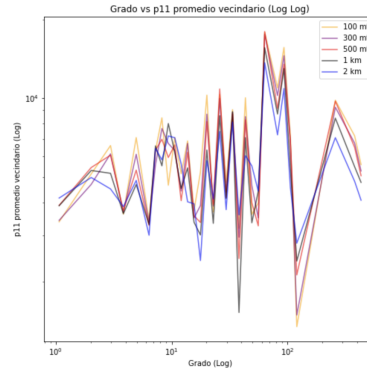


Fig. 21: Degree vs p11 vecindario

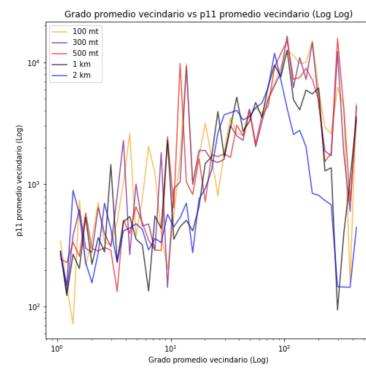


Fig. 22: Degree vs p23 vecindario

7.1. Impacto en el entorno

El propósito de esta sección es describir la dinámica en el desarrollo de una protesta según su entorno geográfico y cómo afecta el normal funcionamiento de las actividades comerciales. Para lograr lo anterior, usaremos las coordenadas asociadas al foco de protesta y mediante el uso de la API **Nearby Search** dispondremos de los nombres, ubicaciones y valoraciones de los siguientes **amenities**: restaurantes, policía, bares, cafes, doctor y tiendas.

La metodología de cálculo es similar a la sección anterior, donde a partir de la coordenada de un foco de protesta es posible establecer radios de búsqueda de **100mt**, **300mt**, **500mt**, **1km** y **2km**. Con cada circunscripción se calcula el promedio de las valoraciones de rating del *establecimiento* según el número de opiniones que haya recibido el establecimiento.

7.1.1. Restaurantes

De acuerdo a lo observado en [23], comparamos el **degree** de un vecindario y la valoración de los **restaurantes** según cada uno de los radios evaluados. A medida que estos locales son más cercanos a los focos de protesta tienden a tener una peor evaluación por parte de los *reviews*, esto podría deberse a que la presencia de organizaciones afecta la experiencia de los clientes. Convengamos que el indicador principal de una protesta es el número de manifestantes que logra convocar, según 24 observamos que independiente del numero de convocados la valoración de los *restaurantes* a 100 mt tiende a ser más desfavorable y a medida que aumenta el radio, las evaluaciones logran estabilizarse. Lo anterior puede deberse a las consecuencias que nacen de una protesta; como rayados, destrucción de mobiliario público, ruidos molestos o los métodos de respuesta empleados por la policía, tales como el uso de gas lacrimógeno o el carro lanza-agua.

Si examinamos 25, la valoración de los *restaurantes* disminuye a medida que aumenta el número de detenidos. donde el umbral de variación llega hasta un 7%. Aquellos establecimientos dentro del radio de 100 mt son los que se ven más afectados en su valoración. Consideremos que el número de detenidos se relaciona con actos violentos y aumenta la sensación de inseguridad del sector, esto dificulta la experiencia del cliente al consumir alimentos, restringe el horario del funcionamiento del lugar y produce una sensación de estrés en el personal del *restaurante*.

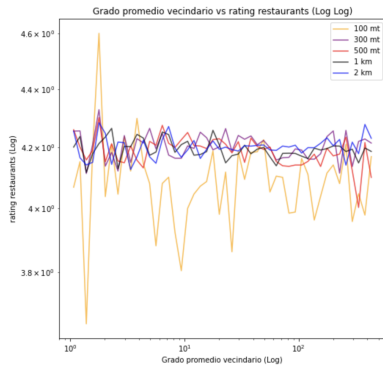


Fig. 23: Degree nhbd vs rating

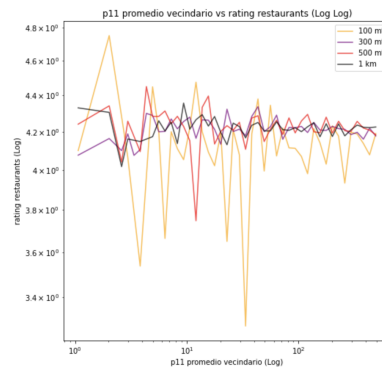


Fig. 24: Degree vs p11

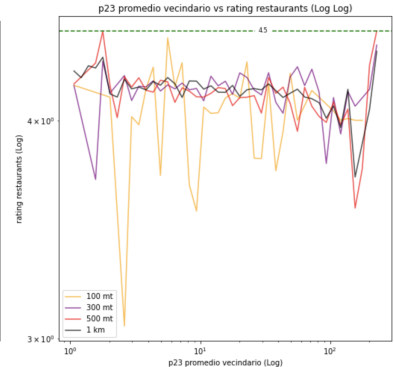


Fig. 25: Degree vs p23

7.1.2. Policía

Para los propósitos del análisis, la definición de policía incluirá a los siguientes puntos de interés: **comisarías, subcomisarías, retenes, brigadas, tenencias y jefaturas.**

Según lo descrito en [26], la evaluación de la policía considera solo aquellos centros cuyo número de **reviews** sea mayor que **30**, todo esto para agregar mayor confianza al análisis. A medida que aumenta el rating de la policía en radios de **2 km**, **1 km** y **0.5 km** disminuye el número de manifestantes, entre sus posibles causas se encuentran:

- **Métodos disuasivos:** La policía posee los recursos necesarios para mitigar la aparición de protestas, dichos recursos dependen directamente de la municipalidad, de ahí que existan sectores con mejor seguridad, ya que los municipios con mejor presupuesto disponen de seguridad ciudadana, vehículos, cámaras, atención telefónica y digital, lo que mejora los tiempos y efectividad en la respuesta.
- **Condiciones del entorno:** Aquellos sectores que evalúan de mejor manera a la policía no acostumbran a manifestarse, ya sea porque son grupos de mejor condición socioeconómica o porque prefieren recurrir a conductos regulares que permiten la solución de conflictos (audiencias con la autoridad, envío de cartas).

Otro ámbito de interés es conocer la distancia que existe desde un foco de protesta hasta el recinto policial más cercano, según lo que observamos en 27 los nodos que más convocan no tienden a alejarse de la policía, sino que se mantienen a una distancia cercana a **1 km**, por otro lado, los nodos más alejados de la policía convocan menos personas, esto se podría deber a las dificultades de acceso a zonas periféricas de la ciudad, también podría deberse a que son protestas radicalizadas, lo que desmotiva el interés del público general.

7.1.3. Bares

La categoría de bares incluye a todos aquellos locales que realicen venta de alcoholes, entre los cuales se encuentran **terrazas, bares, discotecas, licorerías, galerías de arte o clubes.**

Según lo que observamos en 28, a medida que aumenta el número de bares en un radio de **2 km** también lo hace el número de manifestantes, caso contrario para radios de **100mt**, **300mt** y **500mt**, podríamos plantear que la alta densidad de locales en ciertos polos de entretenimiento dificulta el surgimiento de protestas, dado que estos locales cuentan con servicios de seguridad privada, los cuales velan por la integridad y el normal funcionamiento del negocio, vale decir, evitar a toda costa una mala reputación del local y una fuga de clientes. Añadido a lo anterior, la mayoría de este tipo de locales se ubican cerca de avenidas, estaciones de metro o centros comerciales, lo que dificulta aún más el desarrollo de una protesta, dada la alta afluencia de público. Algunas razones del aumento de protestantes en radios de **2 km**. puede deberse a que los manifestantes pertenecen a un segmento cuya preferencia de consumo (estilo de vida) son compatibles con las propuestas de valor asociados de los bares y la vida nocturna.

Según lo que observamos en 29 el rating y el número de manifestantes aumentan de forma proporcional, siendo el máximo valor de **p11** obtenido para un rating de **4.1**, para ratings mayores el número de convocados baja, lo que se relaciona con la exclusividad en ciertos sectores cuyos ingresos promedio son mayores. Vale destacar la amplitud entre las curvas para ratings mayores a **4.1**, esta alta desviación podría indicar una menor densidad de bares en los sectores con mayor rating, se podría suponer que

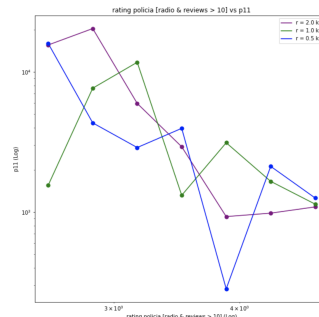


Fig. 26: Rating policia vs p11

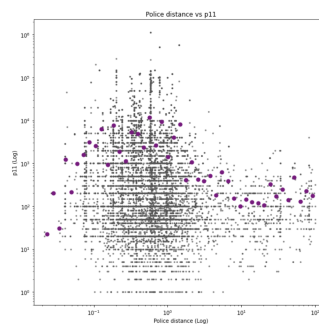


Fig. 27: Distancia policia vs p11

están rodeados de sectores residenciales o parques, consideremos que la valoración de los bares no solo se relaciona con los consumidores, sino con su entorno y sí incomoda o no a los vecinos del sector.

Por otro lado, observamos en [30] que el número de detenidos alcanza su tope en rating cercano a 4.4, cabe resaltar la diferencia de amplitud en el número de detenidos en ambos extremos del punto de inflexión, lo que indica una diferencia de trato en los procedimientos policiales.

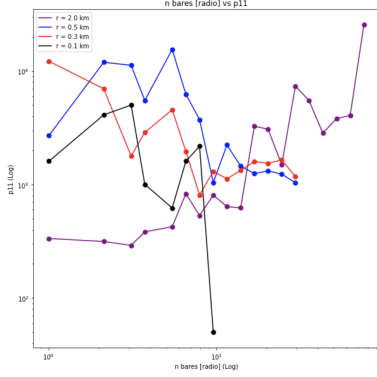


Fig. 28: n bares vs p11

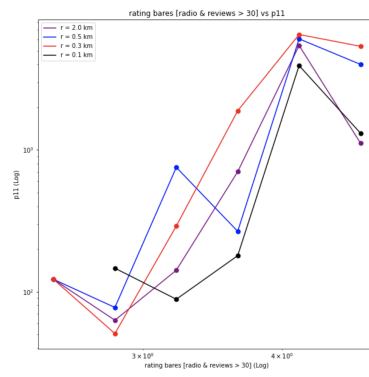


Fig. 29: rating bares ≥ 30 vs p11

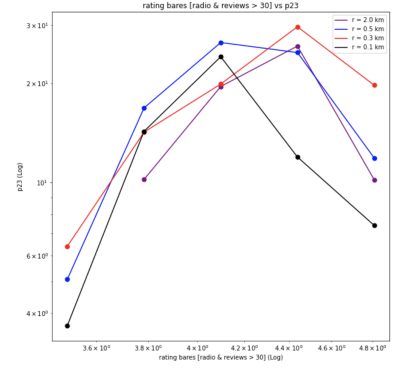


Fig. 30: rating bares ≥ 30 vs p23

7.1.4. Hospitales

Dentro de esta sección se incluirán los siguientes establecimientos: **hospitales, clínicas, centros médicos, consultas particulares** entre otros.

De acuerdo a lo observado en la figura [31] el número de manifestantes disminuye a medida que el radio baja, esto sugiere que los polos de protestas prefieren tomar distancia de *hospitales*, dicha decisión puede deberse a la contribución positiva que realizan estos centros a la comunidad y a la imposibilidad de acceso a atención médica privada por parte de ciertos grupos vulnerables de la población.

Si comparamos el rating con el número de detenidos en [32] distinguimos un punto de inflexión para el valor de rating **2.7**, dicha condición sería un umbral que separa a aquellos usuarios que se sienten conformes o disconformes con el servicio médico brindado. Este cambio puede deberse a la distribución geográfica de los centros médicos, dado que es más común encontrar **CESFAM, SAPU** y **SAR** en sectores de menores ingresos, consideremos que la atención brindada por estos establecimientos no es la mejor, lo que bajaría su rating y por otro, dichas construcciones tienden a estar emplazadas en sectores residenciales cuya población tiende a ser de mayor edad, lo que explicaría un menor número de detenidos. Por otro lado, para rating mayores a **2.7** el número de detenidos aumenta, esto se debe exclusivamente a razones externas al servicio prestado, ya sea por la ubicación geográfica de los **Hospitales** y **Clínicas**, como por la cercanía que poseen a grandes avenidas o centros comerciales.

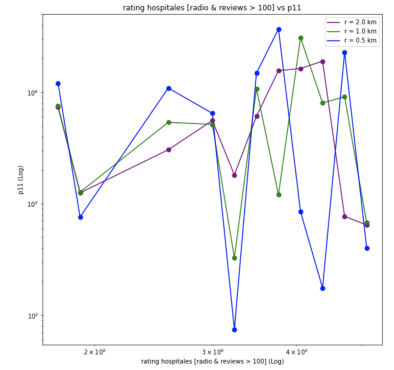


Fig. 31: Rating hospital vs p11

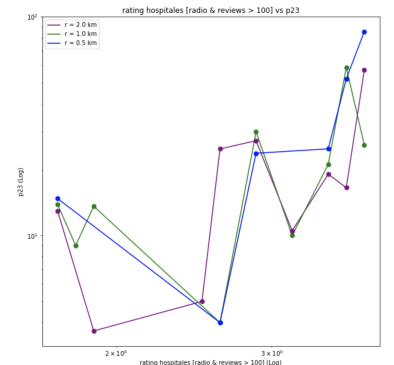


Fig. 32: Rating hospital vs p23

7.1.5. Tiendas

Dentro de esta categoría se incluyen todo tipo de establecimientos comerciales, ya sean **supermercados, farmacias**, Según lo que observamos en [33], a medida que aumenta el número de tiendas,

disminuye el número de manifestantes, esto se cumple para los radios de **1 km**, **0.5 km**, **0.3 km** y **0.1 km**. Las posibles razones que explican esta tendencia son la alta afluencia de público, la presencia de seguridad privada y el alto número de vehículos que recorren las calles aledañas. Para el caso de evaluación de locales, filtraremos las **tiendas** cuyo número de **reviews** sea mayor a 30, así consideraremos sólo establecimientos consolidados y las opiniones vertidas pueden llegar a ser reflejo de la calidad del lugar.

De acuerdo a lo observado en [34], el principal punto de inflexión del número de protestantes se da para un rating de **3.7** y para valores mayores a **4.3** el número de protestantes disminuye considerablemente. Las tiendas con mejor rating son aquellas que ofrecen mejores productos, sus vendedores tratan con amabilidad al público y cuentan con instalaciones acordes al rubro desempeñado. Aquel estatus que alcanzan las tiendas debe ser resguardado por ellas mismas para que perdure en el tiempo, de ahí que la disminución del número de protestas nasca como respuesta de los administradores de locales para proteger los intereses de la **tienda**. Fenómeno similar se cumple para la variable **p23** en 35, observemos que comparte el mismo punto de inflexión en **3.7** como consecuencia de la correlación entre **p11** y **p23**, cabe resaltar la menor desviación entre las curvas para rating mayores a **4.3**, volviendo a retratar un mejor actuar policial para el control de manifestaciones.

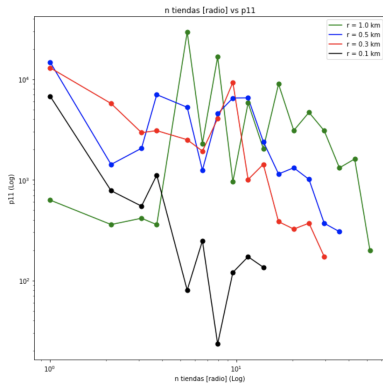


Fig. 33: n tiendas vs p11

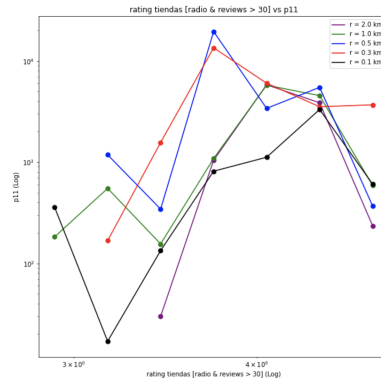


Fig. 34: rating tiendas ≥ 30 vs p11

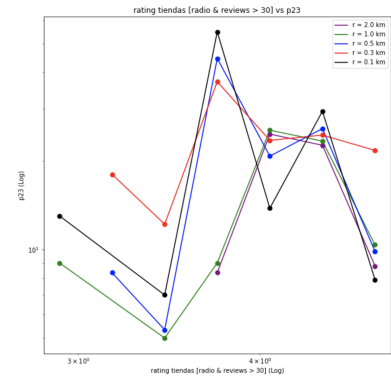


Fig. 35: rating tiendas ≥ 30 vs p23

7.1.6. Parques

A diferencia de los casos anteriores, según lo observado en 36 a medida que aumenta el número de parques en radios de **1 km** y **2 km** también aumenta el número de manifestantes, esto puede relacionarse a la superficie disponible del entorno, lo que permite un mayor número de asistentes, agreguemos a lo anterior que la comodidad, las áreas verdes y la recreación son atributos valorados por la sociedad en general. Caso contrario, dichas cifras podrían deberse a problemas de estimación por parte de la autoridad, dada la dificultad para distinguir entre manifestantes y visitantes del parque. Misma tendencia se observa para la variable **p23**, lo que guarda se explicaría con que la efectividad de las policías aumenta en espacios libres por efectos de una mejor visibilidad. Consideremos la figura [37] que evalúa el rating de los parques según el número de manifestantes, donde observamos tres puntos de inflexión en los valores **2.9**, **3.2** y **4.3**. Aquellos parques cuyo rating pertenece a (**2.9**, **3.2**) tienden a convocar menos personas, estos corresponden a un grupo cuyas características demográficas influyen en la convocatoria, las cuales podrían relacionarse con **factores etarios**, lo que se conoce como barrios para jóvenes y barrios para adultos, también puede influir la dificultad de acceso a estos parques. Para todos los radios evaluados se observa una disminución en el número de convocados desde el rating **4.3** en adelante, aquí influye el factor de exclusividad del entorno y las condiciones socioeconómicas de los vecinos colindantes.

Fenómeno distinto se observa para **p23** en 38 ya que a medida que aumenta el radio, disminuye el número de detenidos, indicando que las posibilidades de fuga aumentan en parques, dado que son espacios de mayor extensión. Para ratings superiores a **4.1** la diferencia de detenidos entre los distintos

radios disminuye, lo que sugiere un trato preferencial para aquellos **parques** de mayor rating, en los cuales el número de detenidos tiende a ser uniforme.

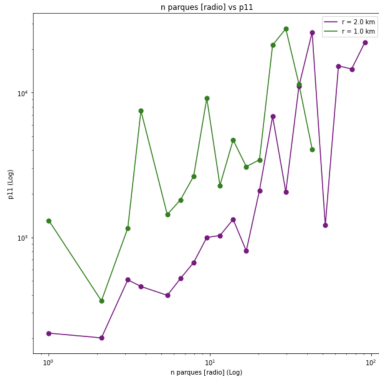


Fig. 36: n parques vs p11

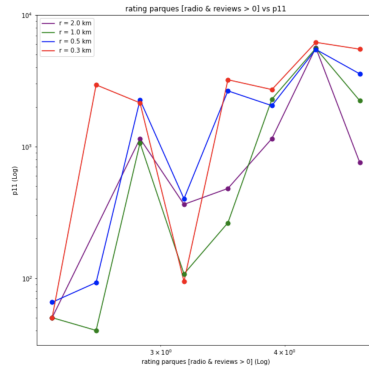


Fig. 37: rating parques ≥ 0 vs p11

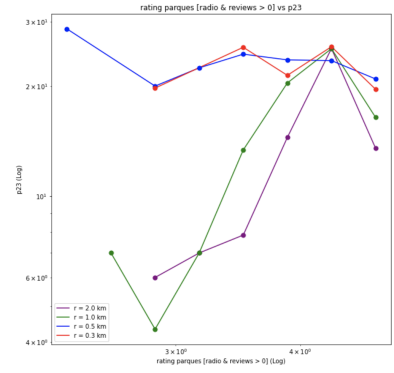


Fig. 38: rating parques ≥ 0 vs p23

8. Modelos estadísticos

A raíz de las múltiples métricas disponibles, ya sean generadas a partir del grafo o de los datos proporcionados por la encuesta **CASEN**, es posible plantear modelos predictivos con el propósito de estimar el número de manifestantes (**p11**) o el número de detenidos (**p23**). Presentamos tres modelos de regresión lineal multivariado, donde la variable principal del **COES** corresponderá al **degree**, dada la alta correlación que posee con el resto de variables, por otro lado, las variables de control corresponderán a métricas **CASEN** que logren representar las condiciones del entorno en que se desenvuelve la protesta. Para el desarrollo de estos modelos, filtraremos las organizaciones según su número de convocados o detenidos sea mayor o igual a un umbral **u**. Lo anterior parte de la base que las organizaciones con mayores números están más arraigadas con su entorno y comparten de mejor forma las características socioeconómicas, educacionales de la comuna en que se desarrolla el acto de protesta.

8.1. Metodología

Para el desarrollo de cada modelo, consideraremos cinco instantes de evaluación para la encuesta **CASEN** (2009, 2011, 2013, 2015 y 2017), en cada instante construiremos el grafo a partir de la proyección bipartita del periodo acumulado, vale decir, el **grafo 2013** incluirá todos los actos de protesta entre los años 2009 y 2013, para luego aplicar las métricas descritas en la sección 4 a cada uno de estos objetos. Sumado a lo anterior, consideramos la encuesta **CASEN** que incluirá información socioeconómica y de entorno a nivel comunal. Dado que la información disponible está a nivel comunal y el propósito del desarrollo es usar datos a nivel de organizaciones, es que nos vemos en la obligación de realizar los siguientes pasos:

- Las métricas topológicas serán aplicadas a cada grafo temporal y los valores de **p11** y **p23** corresponderán al promedio total de manifestantes y detenidos comprendidos en el periodo evaluado. En aquellos casos donde los valores de **p11** o **p23** sean nulos, estos serán ignorados y no influyen en el cálculo del promedio.
- Las métricas resultantes de la encuesta **CASEN** corresponderán al promedio ponderado de los valores según la comuna donde se produce la manifestación, esto se basa en la hipótesis que los manifestantes comparten ciertas características socioeconómicas, educacionales y de entorno que coinciden con la comuna en la que se desarrolla el acto.

La construcción de los modelos considera como variable única **COES** el **degree**, dada su alta correlación con el resto de variables, esto evita que el **variance invariance factor** sea mayor a **5** como restricción de multicolinealidad (Akinwande et al. (2015)). De las variables **CASEN**, consideramos un número limitado de variables menor o igual a **cinco**, cuya elección será en función de maximizar el valor de **r2-ajustado** y reducir el tiempo de cómputo.

8.2. Interpretación de resultados

A partir de los resultados podremos determinar cada uno de los coeficientes que acompañan los coeficientes lineales del polinomio. Cada coeficiente expresará el peso o influencia, sea positiva o negativa, de cada variables, para cada tipo de coeficiente, consideremos:

- Positiva: La variable se considera un *acelerante*, donde estas variables de entorno condicionan el aumento de **p11** o **p23**. Dichas variables pueden estar relacionadas con el poder convocante de una organización, como también con condiciones socioeconómicas que se vinculan directamente con el malestar social, siendo estas consignas propias en el desarrollo de ciertas protestas (salud, educación, vivienda, trabajo).

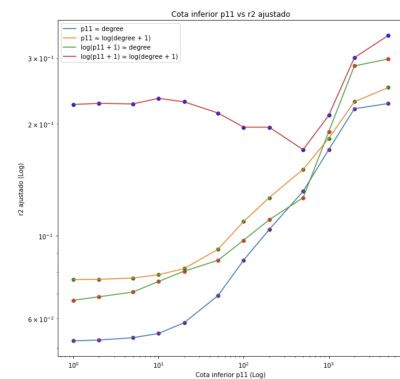


Fig. 39: Cota p11 vs r2 adj

- Negativa: Se consideran variables *mitigantes*, estas variables se pueden relacionar a características culturales (políticas) o monetarias que influyen en el número de convocantes, sea por acción directa o por omisión.

Definamos la **cota inferior p11** como el menor valor para el cual se considerarán organizaciones que al menos hayan convocado un número mayor o igual de manifestantes.

Evaluaremos los distintos escenarios para la variable p11 con cotas inferiores desde 1 hasta 5000. Según las condiciones anteriormente descritas de elección de variables, observamos en [39] el máximo valor de r2 ajustado según la evaluación de todos los posibles modelos, en términos generales, los modelos predictivos que incluyen **log** tiene mejor performance, esto puede deberse a la expansión que genera la aplicación del logaritmo en una variable, además la mejor performance del modelo se obtiene en valores altos de **p11**, lo que indica múltiples cosas, entre ellas, que los modelos tienen la capacidad de predecir en la medida que se cumplan patrones a lo largo del tiempo, aquellas organizaciones que convocan mucho más son aquellas que permanecen en el tiempo, logran construir su identidad e identificar a sus manifestantes con una causa.

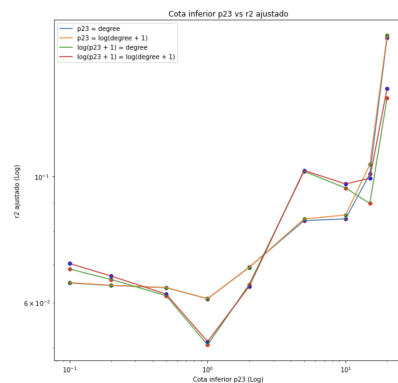


Fig. 40: Cota p23 vs r2 adj

8.3. Model 1: $p11 \geq 10$

En el siguiente modelo, la variable de *ayudas_escolares* corresponde a todo tipo de asistencia médica, dental, psicológica o a la entrega de útiles escolares, tiene un factor de **-0.37**, indicando que en la medida que los gobiernos locales tienen mayor presencia en escuelas y liceos prestando atención a sectores vulnerables, el número de manifestantes tiende a disminuir. Dada la municipalización de los establecimientos educacionales, podría asumirse que la atención brindada por el municipio es óptima en múltiples aspectos, a modo de ejemplo, furgones de traslado de estudiantes, becas de alimentación, transporte, manutención, talleres extraprogramáticos o la entrega de colaciones a los estudiantes.

Por otro lado, el coeficiente de *bienes_propios* es **-0.4**, dicho valor corresponde a los bienes muebles que posea el núcleo familiar, podríamos considerar que su uso provee cierta condición de status, confort y seguridad para el núcleo familiar. El coeficiente de *status_hogar* es de **-0.44**, este indicador representa el estado de los materiales, conexiones, cimientos y estructuras de la vivienda, a medida que mejor estado tiene la vivienda, el número de manifestantes tiende a disiparse, conengamos que el estado del hogar influye directamente en las condiciones descanso (calefacción, iluminación, sistemas de desagüe).

La *escolaridad* tiene un factor positivo de **1.76**, indicando que la toma de conciencia y la visión de mundo influye, dentro de esta variable se incluyen todos los tipos de educación superior, sean de técnico

<i>Dependent variable: log(p11 + 1)</i>	
(1)	
<i>ayudas_escolares</i>	-0.37370*** (0.02830)
<i>bienes_propios</i>	-0.40804*** (0.03184)
<i>log(status_hogar + 1)</i>	-0.44038*** (0.06991)
<i>log(escolaridad + 1)</i>	1.76023*** (0.12868)
<i>const</i>	4.17320*** (0.04486)
<i>log(degree + 1)</i>	0.93172*** (0.01711)
Observations	11,064
R^2	0.23425
Adjusted R^2	0.23390
Residual Std. Error	1.43033(df = 11058)
F Statistic	676.54413*** (df = 5.0; 11058.0)

Note: *p<0.1; **p<0.05; ***p<0.01

superior, universitario o diferenciales, de ahí que posibles razones como el adoctrinamiento por grupos subversivos se descarta.

Finalmente, el **degree** influye positivamente en el polinomio, dado su coeficiente de **0.93**, esto viene a reafirmar la idea que una organización crece en la medida que crea nuevas conexiones y es capaz de sostener una base de simpatizantes que adhieren a sus causas, lograr construir su orgánica y puede elegir a sus dirigentes por medio de procesos democráticos.

8.4. Model 2: $p11 \geq 2000$

Consideremos en el siguiente modelo que el coeficiente de *horas_trabajo* es de **-0.04**, lo que indica que a medida que es mayor la jornada laboral, menor es el número de manifestantes, destaquemos la neutralidad de esta variable, dado que las horas de trabajo no implican necesariamente un mejor salario o una mejora en las condiciones económicas del núcleo familiar.

El coeficiente que acompaña a la variable de *ingreso_percapita* es bajo, **0.0000019**, debemos considerar que su orden de magnitud es mucho mayor comparado con el resto de variables **CLP**, dicha variable representa los ingresos mensuales de cada integrante del hogar, resaltamos que el sueldo mínimo en Chile por ley es de 380000 pesos chilenos. De ahí que en la medida que los ingresos aumentan, aumenta el número de protestas, destaquemos que las protestas no son un fenómeno exclusivo de clases populares y de los sectores de menores ingresos, sino que históricamente la presencia de sectores de clase media y de sectores más ilustrados han propiciado el surgimiento de manifestaciones.

El coeficiente de $\log(part_politica + 1)$ corresponde a **-0.189**, lo que indica algo contrario a lo esperado, su explicación podría deberse a dos razones:

- Despolitización: La desconfianza en la clase política ha llevado a la creación de movimientos autonomistas.
- Militancia: La convocatoria a ciertas manifestaciones puede ser frenada por intereses políticos, dado que el hecho de participar en ciertas coaliciones de gobierno hace cambiar la postura respecto al uso de la **protesta** como un medio de presión.

El coeficiente del $\log(degree + 1)$ es igual a **0.141**, lo que reafirma el poder de convocatoria de una organización para masas críticas sobre **2000** manifestantes.

<i>Dependent variable: log(p11 + 1)</i>	
(1)	
<i>horas_trabajo</i>	-0.0407684*** (0.0093247)
<i>ingreso_percapita</i>	0.0000019*** (0.0000005)
<i>escolaridad</i>	0.2430704*** (0.0447703)
<i>const</i>	7.9565012*** (0.1155203)
$\log(part_politica + 1)$	-0.1895220*** (0.0416449)
$\log(degree + 1)$	0.1415564*** (0.0300014)
Observations	935
R^2	0.3044738
Adjusted R^2	0.3007304
Residual Std. Error	0.7568293(df = 929)
F Statistic	81.3358838*** (df = 5.0; 929.0)

Note: *p<0.1; **p<0.05; ***p<0.01

8.5. Model 1b: $p23 \geq 20$

La variable de *ingreso-percapita* es positiva, pero a diferencia del caso anterior, aquellas comunas que tienen mejores ingresos económicos, tienden a tener mejor presencia policial y por ende, controlan de mejor manera los disturbios. Junto a lo anterior, la variable *bienes-propios* es positiva **30.76**, lo que indica que en aquellos lugares donde hay mayor cantidad de bienes propios, aumenta el número de detenciones, la presencia de bienes propios indica vínculo mucho más potente con la comuna y el entorno, y probablemente esté asociado a población con vivienda definitiva y no con población flotante. Recordemos que ciertas comunas disponen de un mejor presupuesto para contratar personal de seguridad ciudadana (policía privada) lo que facilita los procesos de investigación, ya que cuentan con mejor material audiovisual y colaboran en las detenciones.

La variable $\log(\text{num_piezas} + 1)$ corresponde al número piezas dividido por el número de personas en una vivienda, dicha variable está estrechamente relacionada con el concepto de **hacinamiento**, a medida que aumenta el hacinamiento en un hogar, mayor es la violencia en el núcleo familiar, sin lugar a dudas, esto impacta en el desarrollo personal e influye en el surgimiento de actos vandálicos que derivan en detenciones.

<i>Dependent variable: p23</i>	
(1)	
$\log(\text{ayudas_escolares} + 1)$	-32.34227*** (11.84703)
$\log(\text{bienes_proprios} + 1)$	30.76288*** (10.46528)
<i>ingreso-percapita</i>	0.00007*** (0.00001)
$\log(\text{num_piezas} + 1)$	-12.47710*** (3.37521)
<i>const</i>	20.06772 (14.46206)
$\log(\text{degree} + 1)$	1.87547 (1.99634)
Observations	266
R^2	0.19254
Adjusted R^2	0.17701
Residual Std. Error	30.02422(df = 260)
F Statistic	12.39916*** (df = 5.0; 260.0)

Note: *p<0.1; **p<0.05; ***p<0.01

9. Conclusiones y trabajo futuro

Las principales ideas que podemos resaltar al final de este trabajo guardan relación con el vínculo existente entre el desarrollo de una protesta y su entorno, las características socioeconómicas del lugar influyen en el impacto y desarrollo de una protesta. Cabe destacar que fue posible plantear modelos estadísticos que lograron predecir el número de manifestantes o detenidos en una manifestación, esto abre la puerta para que en un futuro puedan plantearse nuevos modelos a partir de los datos generados.

Proyectando este trabajo al futuro, faltan múltiples bases de datos por agregar al análisis, ya sean el registro de **GDELT** que contiene las totalidad de noticias publicadas en medios digitales, lo que permitiría crear nuevas métricas que detallan el impacto y las consecuencias de una protesta, también dispondremos de un registro sobre la opinión ciudadana o editorial transmitida. Se vuelve necesario agregar información sobre RR.SS. de aquellas organizaciones relevantes, con ello conoceremos la opiniones vertidas por sus líderes y la forma en que llaman a protestar. Sumado a lo anterior, toda la información de presupuestos municipales, registros policiales y pruebas de educación aplicadas a la población escolar (SIMCE, PSU) podría ser utilizada para detallar aún más las características del entorno. Debemos incluir visualizaciones que faciliten el acceso a la información, siguiendo los lineamientos de este proyecto, usaremos **Looker** de **GCP**.

10. Apéndice

10.1. Fuentes de información

10.1.1. Encuesta CASEN

- CASEN 2009: <http://observatorio.ministeriodesarrollosocial.gob.cl/encuesta-casen-2009basedatos>
- CASEN 2011: <http://observatorio.ministeriodesarrollosocial.gob.cl/encuesta-casen-2011>
- CASEN 2013: <http://observatorio.ministeriodesarrollosocial.gob.cl/encuesta-casen-2013>
- CASEN 2015: <http://observatorio.ministeriodesarrollosocial.gob.cl/encuesta-casen-2015basedatos>
- CASEN 2017: <http://observatorio.ministeriodesarrollosocial.gob.cl/encuesta-casen-2017>

10.1.2. REVIEWS

Places search API: <https://developers.google.com/maps/documentation/places/web-service/search>

Nearby search API: <https://developers.google.com/maps/documentation/places/web-service/search-nearby>

Bases de datos: Contiene informacion sobre los **amenities**: bares, restaurantes, doctor, parque, policia, cafes. <https://github.com/imceballos/nearysearchdata>

11. Bibliografía

Referencias

- M. O. Akinwande, H. G. Dikko, and A. Samson. Variance inflation factor: As a condition for the inclusion of suppressor variable(s) in regression analysis. *Open Journal of Statistics*, 05(07):754–767, 2015. doi: 10.4236/ojs.2015.57075. URL <https://doi.org/10.4236/ojs.2015.57075>.
- A.-L. Barabasi and R. Albert. Albert, r.: Emergence of scaling in random networks. *science* 286, 509-512. *Science (New York, N.Y.)*, 286:509–12, 11 1999. doi: 10.1126/science.286.5439.509.
- A.-L. Barabasi, E. Regan, and Z. Oltvai. *Hierarchical Organization of Modularity in Complex Networks*, pages 46–65. 03 2004. ISBN 978-3-540-40372-2. doi: 10.1007/978-3-540-44943-0_4.
- M. E. Berthelton and D. I. Kruger. Risky behavior among youth: Incapacitation effects of school on adolescent motherhood and crime in chile. *Journal of Public Economics*, 95(1-2):41–53, Feb. 2011. doi: 10.1016/j.jpubeco.2010.09.004. URL <https://doi.org/10.1016/j.jpubeco.2010.09.004>.
- C. Cabalín-Quijada. Online and mobilized students: The use of facebook in the chilean student protests. *Comunicar*, 22(43):25–33, July 2014. doi: 10.3916/c43-2014-02. URL <https://doi.org/10.3916/c43-2014-02>.
- V. Colizza, A. Flammini, M. Serrano, and A. Vespignani. Detecting rich-club ordering in complex networks. *Nat Phys*, 2, 03 2006. doi: 10.1038/nphys209.
- R. Diestel. *Graph Theory (Graduate Texts in Mathematics)*. Springer, August 2005. ISBN 3540261826. URL <http://www.amazon.ca/exec/obidos/redirect?tag=citeulike04-20{&}path=ASIN/3540261826>.
- S. Gago, J. Hurajová, and T. Madaras. Notes on betweenness centrality of a graph. *TEC2005-03575*, 62, 02 2012. doi: 10.2478/s12175-011-0065-7.
- W. Liu, M. Pellegrini, and A. Wu. Identification of bridging centrality in complex networks. *IEEE Access*, PP:1–1, 07 2019. doi: 10.1109/ACCESS.2019.2928058.
- C. f. S. C. Reproducible Research and C. S. COES. Observatory of Conflicts - Cumulative Dataset, 2020. URL <https://doi.org/10.7910/DVN/GKQXBR>.
- T. Schank and D. Wagner. Approximating clustering coefficient and transitivity. *Journal of Graph Algorithms and Applications*, page 2, 01 2005. doi: 10.7155/jgaa.00108. URL https://i11www.iti.kit.edu/_media/projects/spp1126/files/sw-acct-05.pdf.
- D. Watts, D. J, Strogatz, and S. H. Collective dynamics of 'small world' networks. pages 301–303, 01 2006.
- X. Xu, J. Hu, and F. Liu. Complex network study of asian go players. *Chaos (Woodbury, N.Y.)*, 17, 07 2007. doi: 10.1063/1.2722040.