



Universidad
de O'Higgins

**DISEÑO DE ALGORITMOS DE DISTRITAJE Y SU APLICACIÓN AL
CASO CHILENO**

LÍA DANAI REYES CARRASCO

TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN CIENCIAS DE LA INGENIERÍA,
MENCIÓN GESTIÓN DE OPERACIONES

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERA CIVIL INDUSTRIAL

PROFESOR GUÍA:
VÍCTOR VERDUGO

PROFESOR CO-GUÍA:
WALDO GÁLVEZ
UNIVERSIDAD DE O'HIGGINS
DIRECCIÓN DE POSTGRADO
ESCUELA DE INGENIERÍA
MAGISTER EN CIENCIAS DE LA INGENIERÍA

RANCAGUA, CHILE
JUNIO, 2024

Resumen

La mayoría de las democracias modernas se basan en órganos representativos, como congresos o parlamentos, para abordar los problemas de la sociedad. Estos órganos suelen organizarse en distritos o regiones, y elegir representantes mediante sistemas electorales específicos. El diseño equitativo de estos distritos es fundamental para garantizar la representatividad y equidad en el proceso democrático.

En Chile, por ejemplo, existen disparidades significativas en la población y la cantidad de representantes en distintos distritos. Esto lleva a una desigualdad en la representatividad política, donde un voto en un distrito menos poblado tiene un peso mayor que un voto en un distrito más poblado. Esta situación plantea cuestionamientos sobre la equidad y legitimidad del sistema democrático.

Para abordar esta problemática, se han propuesto diversos distritajes fundamentados en principios de equidad y contigüidad. Se destacan dos modelos principales: en el primero, el propio plan debe definir los puntos de referencia de los distritos, mientras que en el segundo, se suministran dichos puntos como un componente del modelo. En ambos casos son modelos de factibilidad, ya que no poseen una función objetivo y consisten únicamente en un conjunto de restricciones que deben cumplirse según condiciones específicas para cada ejecución. Para ambos modelos, los puntos de referencia serán llamados “centros” y se refieren a las comunas alrededor de las cuales se generarán los distritos. A partir de estos modelos, se exploran dos escenarios clave: uno que respeta los límites regionales y otro que los elimina. La eliminación de los límites regionales se presenta como una opción con el potencial de mejorar de manera significativa la equidad en la representación. En este modelo, el valor de la métrica que mide la diferencia de población de cada distrito disminuye drásticamente. Este resultado indica que, en términos de distribución de población en los distritos, el nuevo plan distrital es considerablemente más equitativo. Sin embargo, esta mejora plantea interrogantes sobre la relevancia geográfica de los distritos, ya que se están eliminando límites regionales.

Agradecimientos

Agradecer a cada persona que me ha brindado apoyo en estos años resulta una tarea desafiante, pero entre todas ellas, destaco a mi familia como el pilar más significativo. A Angelica, mi madre, ha dado todo por mí, poniéndome por encima de todo y nunca dejarme caer. A Eduardo, mi padre, me ha inculcado la perseverancia, constancia y responsabilidad. Agradecer por el tiempo y esfuerzo que ambos han invertido en mi educación. A Barbara y Jorge, mis hermanos, han sido apoyos fundamentales en mi vida. Su comprensión y paciencia han sido un regalo para mí. En cada desafío, han estado a mi lado, compartiendo risas en los momentos alegres y ofreciendo consuelo en los momentos difíciles. A Tomás y Vicenta, mis sobrinos, gracias por aliviar el estrés de mi vida universitaria con sus locuras y carcajadas, convirtiéndose en una fuente constante de felicidad en mi día a día. A cada uno de ustedes infinitas gracias por siempre estar cuando los necesitaba, escucharme y celebrar mis logros; gran parte de esto es por y para ustedes.

A Sonia, Clementina y Samuel, mis queridas abuelas y abuelo, aunque ya no están para presenciar este momento, no puedo dejar de expresar mi profundo agradecimiento. Desde mi infancia, ustedes me alentaron a estudiar y aprender. Fueron el impulso que necesitaba para ser la mujer que soy hoy. Cada enseñanza y cada gesto de apoyo han dejado una huella imborrable en mi vida, y aunque ya no estén físicamente presentes, su legado vive en mí con amor.

A Emanuel, las palabras se quedan cortas para expresar mi agradecimiento por todo lo que has hecho por mí a lo largo de la carrera. Quiero destacar tu apoyo incondicional, tu fe en mis capacidades, y el constante impulso para que alcanzara metas más altas. Además, agradezco profundamente tu disposición a acompañarme en cada momento que necesité, especialmente en esos momentos nerviosos antes de rendir un examen. Tu presencia y apoyo fueron fundamentales.

A Sayli, Consuelo, Enzo y Tomás, el orden en que los menciono refleja el momento en que formamos nuestra amistad, pero es insignificante al compararlo con el agradecimiento que les tengo. Aprecio profundamente el apoyo y compañía brindada, valorar mi trabajo y alegrarse por mis logros como si fueran propios. La amistad que formé con cada uno de ustedes es una de las cosas más lindas que me ha dejado mi paso por la universidad. No

puedo pasar por alto mencionar a las personas que se cruzaron en mi camino y aportaron alegrías, aprendizajes mutuos, conversaciones y experiencias valiosas durante toda mi etapa universitaria: Víctor, Katherine, Sebastian, Adolfo y Diego. Agradezco a cada uno de ustedes por ser parte importante de mi vida universitaria y personal.

A Víctor Verdugo y Waldo Gálvez, quiero agradecerles por confiar y creer en mis conocimientos. Valoro enormemente las herramientas y oportunidades que me brindaron, las cuales fueron fundamentales para llegar hasta este punto. Ambos han sido guías constantes, y su paciencia junto con el tiempo dedicado a mi desarrollo son invaluable. Estoy agradecida por contar con su apoyo y orientación a lo largo de este camino.

No puedo pasar por alto a los profesores que tuve en estos seis largos años. La profesional en la que me estoy convirtiendo se debe en gran medida a las enseñanzas que me brindaron. Además, no puedo más que admirar la labor que realizan diariamente en el ámbito de la docencia. Agradezco su dedicación y conocimientos que contribuyeron significativamente a mi crecimiento académico y personal.

Quiero agradecer a mis queridas mascotas, Mini, Roma y Mateo; por ser un pilar fundamental en mi vida, colmándola de amor incondicional. Su compañía fue mi refugio en esas largas noches de estudio, y el recibirme con entusiasmo y alegría cada vez que regresaba después de un largo día en la universidad hizo que mi hogar fuera aún más especial.

Mi agradecimiento se extiende a cada una de las personas que han formado parte de mi vida universitaria. Cada gesto de apoyo, palabra amable y acto de generosidad ha dejado una marca en mi corazón, y estoy profundamente agradecida por ello.

Tabla de Contenido

1. Introducción	1
1.1. Trabajo previo	3
1.2. Resultados	6
1.3. Organización	7
2. Tratamiento de datos	8
2.1. Base de datos	8
2.2. Calcular las distancias entre comunas en Chile	10
2.3. Adyacencia entre comunas en Chile	10
2.4. Cálculo de camino más corto entre las comunas de Chile	12
2.5. Calidad de balance poblacional en el distritaje actual para Chile	13
3. Modelos de programación entera	15
3.1. Modelo de factibilidad con centros fijos	17
3.2. Modelo de factibilidad con centros por definir	17
3.3. Restricción de límites regionales	18
3.4. Búsqueda binaria para el balance poblacional óptimo	19
4. Simulaciones y resultados	21
4.1. Respetando límites regionales y 28 distritos	22

4.1.1.	Centros por definir	22
4.1.2.	Centros fijos	22
4.2.	Sin límite regional y 28 distritos	25
4.2.1.	Centros por definir	26
4.2.2.	Centros fijos	30
4.2.3.	Centros por definir integrando función objetivo para imponer compa- cidad	32
4.3.	Centros por definir con límites regionales y variación del número de distritos	33
4.3.1.	Variación entre 16 y 56 distritos	33
4.3.2.	Variación entre 16 y 200 distritos	34
4.4.	Interpretación de los resultados	38
4.4.1.	Efectos de la presencia de límites regionales	38
4.4.2.	Heurística para caso con límites regionales	39
4.4.3.	Número de distritos dónde se alcanza el óptimo para la menor diferen- cia de población	39
4.4.4.	Comparando balance poblacional actual de Chile con los obtenidos en los modelos	40
4.4.5.	Relevancia de restricciones de población en el modelo	41
5.	Conclusión	42
5.1.	Trabajo Futuro	43
	Bibliografía	45
	Apéndice A. Anexo	46
A.1.	Diccionario de funciones utilizadas para el código	46
A.2.	Pseudocódigos	46
A.2.1.	Grafo de adyacencia y matriz de adyacencia	47

A.2.2. Cálculo caminos más cortos	47
---	----

Índice de Figuras

1.1. Gerrymandering, Distrito 12, Carolina del Norte (2014).	2
4.1. Distritaje actual para Chile, respeta límites regionales con $\varepsilon_{BP} = 1,3299$ y $K = 28$	23
4.2. Distritaje actual para la Región Metropolitana, respeta límites regionales con $\varepsilon_{BP} = 1,3299$ y $K = 28$	24
4.3. Distritaje para la Región Metropolitana respetando límites regionales con $\varepsilon_{BP} = 0,8398$, $K = 28$ y centros por definir.	25
4.4. Distritaje para Chile respetando límites regionales con $\varepsilon_{BP} = 0,8398$, $K = 28$ y centros por definir.	26
4.5. Distritaje para la Región Metropolitana respetando límites regionales con $\varepsilon_{BP} = 0,8398$, $K = 28$ y centros fijos.	27
4.6. Distritaje para Chile respetando límites regionales con $\varepsilon_{BP} = 0,8398$, $K = 28$ y centros fijos.	28
4.7. Distritaje para Chile sin límites regionales con $\varepsilon = 0,039$, $K = 28$ y centros por definir.	29
4.8. Distritaje Región Metropolitana sin límites regionales con $\varepsilon = 0,039$, $K = 28$ y centros por definir	30
4.9. Distritaje para Chile sin límites regionales con $\varepsilon = 0,6796$, $K = 28$ y centros fijos.	31
4.10. Variación de ε según el valor de K entre 16 y 56	34
4.11. Variación de ε según el valor de K entre 16 y 200	35

4.12. Comparación del valor de ε según la ecuación despejada y el valor obtenido en el modelo para K entre 16 y 56	36
4.13. Comparación del valor de ε según la ecuación despejada y el valor obtenido en el modelo para K entre 40 y 200	37

Capítulo 1

Introducción

Muchas de las democracias modernas tienen en sus pilares la conformación de órganos representativos, con el fin de poder diseñar soluciones a problemáticas que enfrenta la sociedad¹. Típicamente, estos órganos representativos, expresados en congresos, parlamentos, asambleas u otros, se organizan en base a una representación territorial, a través de regiones o distritos, en los cuales se escoge posteriormente una cierta cantidad de representantes de acuerdo a algún sistema electoral [2]. Usualmente, estos planes distritales se diseñan teniendo en consideración dos dimensiones fundamentales: aspectos geográficos y distribución de la población [11]. Encontrar forma de diseñar los distritos de manera equitativa y justa es esencial para garantizar la representatividad y la equidad en el proceso democrático.

El propósito de un distritaje es particionar el total de comunas en una cantidad determinada de distritos. La determinación de qué comunas conforman cada distrito es un asunto tanto matemático como político. Desde el punto de vista matemático, se puede tratar de agrupar comunas en base a criterios como la contigüidad geográfica, la población o la compacidad [11], con el objetivo de formar distritos de tamaño similar y con características homogéneas. Algoritmos y técnicas de optimización, como “divide y vencerás” o “búsqueda local”, pueden ser empleados para agrupar comunas eficientemente siguiendo ciertos criterios objetivos [11, 13, 12, 9, 5].

Desde el punto de vista político, la delimitación de distritos puede influir en los resultados electorales. La forma en que se agrupan las comunas puede beneficiar a ciertos partidos o candidatos en desmedro de otros, un fenómeno conocido como *gerrymandering* [10], en la Figura 1.1 se observa un ejemplo de este fenómeno. Por ello, es esencial que el proceso de delimitación sea transparente, imparcial y equitativo.

En Chile, para las elecciones parlamentarias, de acuerdo a la nueva ley electoral [2], hay 28 distritos electorales conformados cada uno por un conjunto de comunas. Sin embargo,

¹https://www.bcn.cl/formacioncivica/detalle_guia?h=10221.3/45753.

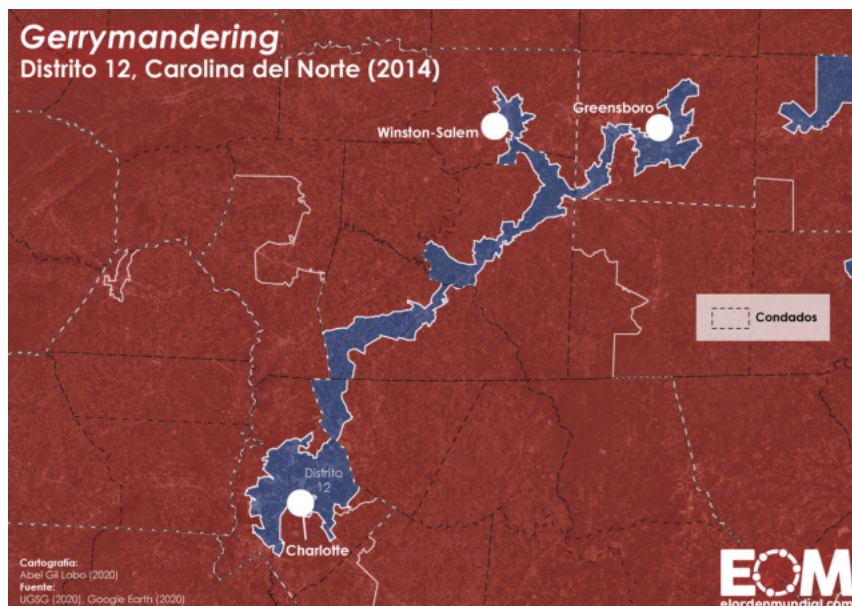


Figura 1.1: Gerrymandering, Distrito 12, Carolina del Norte (2014).

existen grandes diferencias de población entre un distrito y otro. Además, cada distrito elige un número variable de representantes, que va desde tres hasta ocho, según el método proporcional D'Hondt,² mismo método que se utiliza para la repartición de escaños por partido en elecciones de diputados o senadores en Chile.³ Sin embargo, actualmente no existe una regla clara y transparente para decidir qué comunas conforman cada uno de los distritos. Cómo resolver este problema es algo que ha sido discutido y estudiado en la literatura actual.

La mejor forma de dejar en evidencia esta situación es que, según datos del Censo de 2017, el distrito 27, conformado por toda la Región de Aysén, posee una población de 101316 personas y escoge tres representantes, mientras el distrito 8, conformado por ocho comunas de la Región Metropolitana, cuenta con 1438474 personas y elige a ocho representantes [1]. Esto da como resultado una disparidad significativa en la cantidad de votos necesarios para que un representante sea elegido en cada distrito. Para ser electos los representantes del distrito 27 necesitan 33772 votos y los del distrito 8 necesitan 179809 votos.

La situación mencionada pone de manifiesto una desigualdad evidente en la representatividad política. Esta disparidad en la representación por voto sugiere que un voto en el distrito 27 tiene un peso mucho mayor que un voto en el distrito 8. En términos simples, la voz de un ciudadano de Aysén tiene más impacto en la elección de un representante que la voz de un ciudadano de la Región Metropolitana, al menos en el contexto de estas cifras, pues el voto de una persona en el distrito 27 vale 5.3 veces el voto de una persona en el distrito 8. Esta situación no solo plantea preguntas sobre la equidad en la representación, sino que

²<https://www.bcn.cl/historiadelaley/nc/historia-de-la-ley/3997/>.

³<https://www.servei.cl/2015/08/18/nuevo-sistema-electoral-chileno-metodo-dhont-2/>.

también puede llevar a cuestionamientos sobre la legitimidad del sistema democrático. Si el principio fundamental de la democracia es que cada voto tiene un valor igual, esta diferencia señalada anteriormente en el ‘valor’ de los votos entre distritos, contradice ese principio.

Lo anterior causa una subrepresentación en los distritos más poblados del país. A partir de aquello, nace la idea de desarrollar un distritaje para Chile basada en principios de equidad y representatividad. La finalidad de esta propuesta es que cada voto tenga un peso similar, independientemente del lugar donde se emita, debido a que es fundamental garantizar que todos los ciudadanos tengan una representación equitativa y que el sistema no esté sesgado en favor o en contra de ningún grupo particular. La integridad del proceso democrático depende de que se tomen decisiones justas y transparentes en temas cruciales como este.

1.1. Trabajo previo

Una dirección que ha guiado el trabajo actual es la distribución equitativa de distritos, visto desde distintos ángulos en los estudios [11, 10, 7, 13, 8]. La literatura se centra principalmente en 2 aspectos: En primer lugar, la optimización entre el gerrymandering y la elección social, como también en el poder de manipulación de los partidos sobre los mapas distritales. Y en segundo lugar, en la generación de mapas equitativos y consistentes geoméricamente.

Para comenzar, Garg et al. [10] ponen su centro en la optimización entre el gerrymandering y la elección social. Ellos abordan la necesidad de calcular eficientemente la asignación de escaños por partido en un distrito de tal manera que sean representativos. Los autores concluyen que para evitar manipulaciones extremas se deben asignar múltiples miembros representativos por distrito, es decir, cada distrito debe poseer al menos dos representantes [10].

En cuanto al gerrymandering, la distribución geográfica de cada uno de los votantes puede afectar en la manipulación de un partido al momento de generar distritos electorales. Mientras más concentrados se encuentren los votantes de cada uno de los partidos, se asegura que cada partido obtenga aproximadamente su parte proporcional de los distritos. Sin embargo, este último modelo ignora la compacidad de los distritos, preservación de comunidades y equidad racial. Además, este modelo asume bastantes simplificaciones que lo alejan de la realidad [7].

En la misma línea, Lewenberg et al. [13] presentan un modelo donde los votantes se encuentran dispersos en un mapa geográfico para ser divididos en distritos electorales. A diferencia del artículo anterior [7], los distritos deben ser *compactos*, es decir, con formas uniformes y concentradas, evitando formas dispersas y alargadas. El enfoque principal del estudio es la estrategia de “divide y vencerás” en gerrymandering. Se crean distritos electorales de manera que los votantes de un partido sean divididos y repartidos en varios distritos. Lo anterior disminuye la influencia del partido manipulado y aumenta las posibilidades de

que el partido que controla la manipulación geográfica obtenga más escaños. Se concluye que estas prácticas pueden afectar la representación política justa y la competencia democrática. Además, la manipulación geográfica puede ser utilizada estratégicamente para obtener ventajas electorales, como también el tamaño de cada uno de los distritos puede ser utilizado para manipulaciones [13]. En comparación, Chen et al. [8] estudian dos maneras de gerrymandering. El ‘packing’ (agrupamiento) de votantes de un partido en un número reducido de distritos y también el ‘cracking’ (división) de votantes de un partido en varios distritos. Para poder estudiar el efecto de ambas prácticas se realizan simulaciones de distintos escenarios de distritos electorales compactos y contiguos, comparándolos con la realidad electoral. Finalmente, se concluye que ambas prácticas de gerrymandering pueden beneficiar al partido que controla el proceso de diseño de distritos electorales, pero no cambiaría la composición partidista de un congreso [8].

Gurnee et al. [11], proponen soluciones computacionales que tienen como objetivo la compacidad de los mapas distritales, aunque se menciona que es de mayor importancia la equidad y contigüidad. El modelo se enfoca en generar opciones de redistribución de manera rápida, económica y transparente [11]. En la misma línea de generar mapas distritales, se encuentran Cohen-Addad et al. [9] que proponen un método que en todos sus experimentos, realizados para 50 estados en Estados Unidos, logra respetar los bloques censales, la contigüidad y el equilibrio perfecto de la población. Asimismo, le dan gran importancia a la compacidad de los distritos bajo distintas definiciones que existen por cada uno de los estados. Dicho método consta de dos fases: La primera fase utiliza la búsqueda local para encontrar un plan de distritos idealizado que satisfaga algunas propiedades impuestas por los investigadores en cuanto a población y compacidad. Sin embargo, en esta fase los bloques censales pueden dividirse en más de un distrito y los centros de cada uno de los distritos son escogidos de forma aleatoria. La segunda fase asigna los bloques censales de los límites de cada distrito al centro con menor distancia, preservando la contigüidad y de tal manera que se minimice la diferencia máxima de población entre los distritos resultantes.

En cuanto a generación de mapas distritales como se mencionó anteriormente se tienen dos enfoques de distintos autores [9, 11] para llevar a cabo un plan distrital. En primer lugar, un modelo al cuál se le imponen todas las condiciones para entregar la mejor solución posible sacrificando algunos aspectos. Por ejemplo, el tiempo de ejecución del programa es mayor, pero se gana en simplicidad debido a que solamente es la ejecución de un programa. En segundo lugar, métodos que van adaptando una solución que no cumple todos los objetivos para finalmente llegar a un plan distrital que sea aceptado. Estos métodos tienen un menor tiempo de ejecución, pero se necesitan más fases para poder llegar a una solución que cumpla con los requerimientos.

Por último, también se ha integrado otras técnicas más abstractas al estudio del problema. Abrishami et al. [5] estudian la aplicación de la teoría del transporte óptimo en el campo de las particiones de grafos, la que puede ser utilizada en la generación de planes distritales. La teoría mencionada es la asignación óptima de recursos de un origen a un des-

tino, minimizando el costo total de asignación proporcional a la distancia de transporte sobre las aristas del grafo [5]. Lo llamativo de esta aplicación es que usando un solo problema de programación lineal se pueden calcular estos costos mencionados, para posteriormente realizar las asignaciones minimizando costos. Sin embargo, el costo computacional de calcular dichas distancias varía según qué tipo de distancia se calcule. Por ejemplo, para 11775 pares calcular distancia de Hamming toma alrededor de 7 segundos, y para distancia de transporte alrededor de 15 horas. Lo mencionado anteriormente es otro enfoque para distritaciones que se basa principalmente en las distancias entre las comunas más que en condiciones. Cabe destacar que [5] resuelven todo el problema en una sola instancia a diferencia de [12] que presentan un algoritmo de redistribución de distritos basado en ‘divide y vencerás’. El algoritmo en [12] tiene los siguientes pasos claves:

1. Divide y vencerás:
 - (a) Inicia con el distrito general y el número de particiones restantes, k .
 - (b) Establece la población deseada según si k es par o impar.
 - (c) Redistribuye la población para crear dos distritos en cada iteración.
 - (d) Aplica ‘divide y vencerás’ recursivamente hasta que el número de particiones restantes sea 1.
2. Redistribución de dos distritos:
 - (a) Voronoi: Elige sitios de origen y crea particiones de Voronoi usando un rectángulo delimitador.
 - (b) Intercambio de contigüidad: Intercambia unidades de población para asegurar la contigüidad.
 - (c) Intercambio de población: Intercambia población para minimizar la desviación poblacional.
 - (d) Maximización: Compara opciones de redistribución basadas en diferentes elecciones de sitios de Voronoi y selecciona la mejor.
3. Recursión y optimización: Se selecciona la versión de la partición de dos distritos con la mejor desviación poblacional y se subdivide recursivamente en más distritos, repitiendo los pasos de Voronoi, intercambio de contigüidad, intercambio de población y maximización.

En resumen, el algoritmo en [12] utiliza ‘divide y vencerás’ para subdividir el estado en distritos, y en cada iteración de redistribución de dos distritos, se aplican pasos de Voronoi y varios intercambios para asegurar contigüidad y minimizar la desviación poblacional. Este proceso se repite hasta lograr el número deseado de particiones.

A partir de los artículos estudiados, se concluye que la generación de mapas distritales se enfoca en lograr distritos que sean compactos, contiguos y equilibrados en población,

según se detalla en los artículos [11, 7, 13, 12]. Estos criterios no solo son esenciales para asegurar la equidad y eficacia en la representación política, sino que también desempeñan un papel crucial en el fortalecimiento de la confianza de la población en el proceso de creación de distritos y, por consiguiente, en el sistema democrático. La búsqueda de distritos que cumplan con estas características no solo contribuye a una distribución más justa, sino que también fomenta la transparencia y la legitimidad del sistema electoral.

1.2. Resultados

El principal objetivo de este estudio es la generación de distintos planes distritales para Chile bajo diversas restricciones, como equilibrio de población o contigüidad. Estos resultados consisten en distritajes con asignaciones específicas de comunas a cada uno de los distritos. La información principal que se extrae del modelo incluye:

1. Población por distrito: Se calcula la población total de cada distrito en un plan distrital particular. Esto es esencial para garantizar que cada distrito tenga una representación justa y equitativa.
2. Variabilidad de la población entre cada distrito: Se utiliza una métrica para balance poblacional en cada plan distrital. Este valor refleja la variabilidad en la población de los distritos y destaca las desigualdades en la representación de los ciudadanos en el proceso electoral.
3. Cantidad de comunas por distrito: Se determina cuántas comunas forman cada distrito en un plan distrital dado. Esto es relevante para entender la división geográfica del país en términos electorales.
4. Gráficos: Para visualizar de manera efectiva cómo se distribuyen las comunas en los distintos distritos, se representan los resultados a través de gráficos de Chile y de la Región Metropolitana. Estos gráficos proporcionan una representación visual que permite apreciar la configuración de los distritos en todo el país.

Es importante destacar que, en el Capítulo 4, los resultados optimizan el balance poblacional. A través de esta medición, se realiza una comparación de los diferentes planes distritales, cómo también con el plan distrital actual de Chile, y se analiza cómo varían en términos de representatividad y equidad. Los gráficos de Chile se emplean según sea necesario para proporcionar una representación visual de la distribución de los distritos en el país.

1.3. Organización

En el desarrollo de este trabajo, se abordan diversas etapas fundamentales para el análisis del distritaje en Chile. En el Capítulo 2, nos adentramos en el tratamiento de datos, donde se describen las modificaciones realizadas en la base de datos, el cálculo de las distancias entre las comunas de Chile, la determinación de comunas adyacentes para cada comuna del país y la obtención del camino más corto entre cada par de comunas chilenas. Además, exploramos la calidad del balance de población en el actual plan distrital de Chile.

A continuación, en el Capítulo 3 nos enfocamos en los modelos utilizados para obtener los resultados deseados. Cabe destacar que ambos modelos son de factibilidad, ya que no poseen una función objetivo y consisten únicamente en un conjunto de restricciones que deben cumplirse según condiciones específicas para cada ejecución. En lugar de buscar optimizar una función objetivo, estos modelos se centran en garantizar que la configuración final de los distritos cumpla con ciertos criterios predefinidos, como la equidad poblacional y la contigüidad geográfica. Se presentan dos modelos fundamentales: el primero, conocido como “Modelo de factibilidad con centros fijos”, donde los centros son predefinidos; y el segundo, en el cual el modelo debe determinar los centros de cada distrito, conocido como “Modelo de factibilidad con centros por definir”. En ambos modelos, los “centros” se refieren a las comunas alrededor de las cuales se generarán los distritos. Estos centros desempeñan un papel crucial en el proceso de creación de distritos, ya que actúan como puntos de referencia alrededor de los cuales se llevarán a cabo los distritajes propuestos. A su vez, explicamos la restricción de límites regionales que nos proporciona una variante para ambos modelos mencionados anteriormente. Por último, se explica en detalle la rutina de búsqueda binaria que desempeña un papel esencial en la búsqueda de un aproximado al mínimo error de balance poblacional (ε_{BP}) en cada uno de estos modelos. Aunque es posible crear un distritaje para un valor específico de ε_{BP} en cada modelo, la utilización de esta rutina de búsqueda binaria nos brinda la capacidad de obtener una aproximación al mínimo en el contexto del error de balance poblacional (ε_{BP}).

Posteriormente, el capítulo 4 está dedicado a las simulaciones y resultados, donde se presentan los resultados obtenidos a través de la implementación de los modelos propuestos. Se presentan y analizan los datos junto con conclusiones derivadas de las simulaciones realizadas en el contexto chileno. Esta sección constituye un núcleo esencial del trabajo, donde se desglosan y discuten los hallazgos más significativos.

En la última etapa de este estudio, específicamente en el Capítulo 5, se presenta una conclusión que resume los hallazgos clave y las contribuciones del trabajo. En este capítulo, se destacan las lecciones aprendidas y se ofrecen recomendaciones para investigaciones futuras en el ámbito del distritaje y la representación electoral en Chile.

Capítulo 2

Tratamiento de datos

Antes de profundizar en el análisis y manipulación de datos, es esencial conocer sobre el preprocesamiento de datos. En este capítulo nos centraremos en la preparación y creación de diferentes tipos de datos que serán utilizados como insumos en las etapas posteriores de trabajo.

2.1. Base de datos

La base de datos principal proviene del Instituto Nacional de Estadísticas de Chile (INE) y del Instituto Geográfico Militar de Chile (IGM), y están recopilados en un archivo .csv en el repositorio. [3]. Este archivo contiene datos geográficos de las comunas de Chile, incluyendo su latitud y longitud. El archivo presenta una estructura tabular organizada en columnas, que representan diferentes atributos o características. A continuación, detallamos las columnas que contiene:

1. CUT: El Código Único Territorial es un código numérico que permite identificar de manera única a las diferentes comunas de Chile.
2. Comuna: Contiene el nombre de la comuna. Las comunas son la unidad administrativa básica en Chile. Actualmente, Chile cuenta con 346 comunas distribuidas en todo el territorio.
3. Provincia: Contiene el nombre de la provincia. Las provincias son un conjunto de comunas agrupadas geográficamente y representan una división administrativa intermedia entre las comunas y las regiones.
4. Región: Contiene el nombre de la región. Chile está dividido en regiones, que son las principales divisiones administrativas del país.

5. Población (2017): Se refiere al número de habitantes que reside en cada comuna en el año 2017. Estos datos son obtenidos a través de censos nacionales.
6. Latitud (decimal): Indica la ubicación geográfica de un lugar en relación con el ecuador. En el caso de Chile, al estar ubicado en el hemisferio sur, las coordenadas de latitud son, generalmente, negativas.
7. Longitud (decimal): Representa la ubicación geográfica de un lugar en relación con el meridiano principal, el Meridiano de Greenwich. Al estar Chile en el hemisferio occidental, sus coordenadas de longitud también son, por lo general, negativas.

La latitud y longitud de las comunas de Chile se refieren al centro aproximado de cada ciudad, teniendo en cuenta su tamaño y forma. Por otro lado, el archivo mencionado, almacenado en formato *.csv*, requiere de una serie de modificaciones previas para facilitar su procesamiento y análisis. A continuación, se describen los pasos y consideraciones tomadas durante esta etapa de preprocesamiento:

1. Normalización de Caracteres: Se inicia el proceso eliminando las tildes y reemplazando las instancias de la letra ñ por una *n*. Esta acción evita posibles inconvenientes en sistemas o lenguajes de programación que no soporten estos caracteres especiales. Adicionalmente, todos los caracteres se transforman a minúscula para garantizar uniformidad en la base de datos.
2. Manejo de Puntuaciones: Al intentar trabajar con la columna de población, se detecta un inconveniente: los números representativos de la población tienen comas para separar miles, lo que puede causar conflictos, especialmente en lenguajes de programación como Python donde la coma es utilizada para otras funciones. Para solucionar este problema, se eliminan las comas de los valores.
3. Truncado de Números: Al momento de manejar las puntuaciones ocurrió que cualquier número que termine en 0 se trunca, lo que significa que se elimina el cero final. Por ejemplo, un valor de “12,030” no solo vería eliminada su coma, sino que también se truncaría a “1203”. Dada esta situación se realiza una revisión manual.
4. Revisión Manual: Debido a las modificaciones mencionadas, especialmente el truncado, es necesario realizar una revisión manual de la columnas con números para asegurarse de que los datos sigan siendo precisos y no se hayan creado errores.

Los pasos mencionados de preprocesamiento son fundamentales para garantizar que el análisis subsiguiente sea preciso y confiable. Es importante recalcar la necesidad de la revisión manual, ya que cualquier alteración de datos, incluso si se realiza con buenas herramientas, puede introducir errores o inexactitudes que afecten los resultados finales.

2.2. Calcular las distancias entre comunas en Chile

Por otro lado, es necesario calcular las distancias entre todas las comunas, con la finalidad de penalizar la distancia entre comunas asignada en un mismo distrito. Para abordar este requerimiento, se recurre a un código en *Python* diseñado para determinar la distancia entre dos puntos geográficos específicos sobre la Tierra, dados sus respectivos valores de longitud y latitud en formato decimal. Este cálculo se basa en la fórmula conocida como la *Fórmula de Haversine* [14]. Esta ecuación matemática estima la distancia del círculo máximo entre dos puntos ubicados sobre la superficie de una esfera, como nuestro planeta. Utilizando la *Fórmula de Haversine*, es posible calcular con precisión las distancias en kilómetros entre todas las comunas chilenas, proporcionando así un *input* fundamental para los modelos de factibilidad. A continuación, se presenta la *Fórmula de Haversine*:

$$2 \cdot \arcsin \left(\sqrt{\sin \left(\frac{\text{lat2}-\text{lat1}}{2} \right)^2 + \cos(\text{lat1}) \cdot \cos(\text{lat2}) \cdot \sin \left(\frac{\text{lon2}-\text{lon1}}{2} \right)^2} \right) \cdot r_T \quad (2.1)$$

donde:

lat1, lon1 : latitud y longitud del primer punto en radianes,
 lat2, lon2 : latitud y longitud del segundo punto en radianes,
 r_T : radio de la Tierra (aproximadamente 6371 km).

2.3. Adyacencia entre comunas en Chile

Además de la limpieza de datos y las distancias, es fundamental contar con un grafo de adyacencia que represente las conexiones entre comunas del país, donde las comunas serían los nodos y los arcos las conexiones. La presencia de un arco entre dos comunas indicaría que las comunas respectivas son comunas adyacentes. La identificación de comunas adyacentes en Chile requirió una metodología detallada y precisa. Para ello, se recurrió a un método manual mediante la consulta de un mapa detallado del país, donde se pueden apreciar claramente los límites comunales. A través de este mapa, se pudieron apreciar las fronteras que compartían cada una de las comunas con sus vecinas.

A partir de la lista generada, se estructura una matriz de adyacencia, denotada por M . Esta matriz es una representación binaria que describe las relaciones de vecindad entre las comunas:

$$M_{ij} = \begin{cases} 1 & \text{si } i \text{ y } j \text{ son adyacentes,} \\ 0 & \text{si no.} \end{cases} \quad (2.2)$$

Estudiar directamente los límites en un mapa geográfico ofrece una visualización clara y precisa de las adyacencias. Esta metodología manual y meticulosa, aunque puede ser laborio-

2.3. Adyacencia entre comunas en Chile

sa, garantiza un alto grado de precisión, especialmente si el mapa utilizado está actualizado y es de alta resolución. Este trabajo nos proporciona una herramienta valiosa para asegurar la contigüidad en los planes distritales.

En el caso de Chile, se han identificado islas que no tienen comunas adyacentes en términos geográficos, pero en el caso de este trabajo y para fines prácticos, se han vinculado a las comunas más cercanas con mayor población. A continuación, enumeramos los casos en los que se ha aplicado esta técnica por las necesidades de representación de cada una:

1. Antártica con Cabo de Hornos: La Isla de la Antártica, a pesar de su ubicación remota, se considera adyacente a la comuna de Cabo de Hornos debido a la proximidad relativa.
2. Juan Fernández con Valparaíso: El Archipiélago Juan Fernández se asocia a la comuna de Valparaíso, que se encuentra en la costa continental más cercana.
3. Isla de Pascua con Valparaíso: De manera similar la Isla de Pascua se vincula a la comuna de Valparaíso.
4. Guaitecas con Cisnes: El Archipiélago de las Guaitecas se relaciona con la comuna de Cisnes debido a su cercanía.
5. Curaco de Vélez con Castro: Curaco de Vélez se vincula con la comuna de Castro, que es la más grande y poblada en la isla de Chiloé.
6. Puqueldón con Castro: De manera similar, Puqueldón se asocia a la comuna de Castro en la isla de Chiloé.
7. Ancud con Maullín: La comuna de Ancud, ubicada en la Isla de Chiloé, se vincula con la comuna de Maullín, que se encuentra de forma más cercana en el continente.

Esta técnica se utiliza para garantizar que estas islas tengan una representación adecuada en el contexto de la administración y la toma de decisiones locales, a pesar de su distancia de la tierra firme.

Además, es esencial entender que, en el contexto de Chile, se han definido dos listas distintas para identificar comunas adyacentes:

1. Considerando límites regionales: En esta lista, si dos comunas están separadas por un límite regional, no se las considera vecinas, independientemente de su proximidad geográfica.
2. Sin considerar límites regionales: Esta lista omite las fronteras entre regiones, considerando únicamente la contigüidad geográfica. Aquí, todas las comunas que sus fronteras se intersectan directamente son consideradas vecinas, sin importar las divisiones regionales.

En resumen, mientras que la primera lista toma en cuenta las divisiones administrativas entre regiones, la segunda se enfoca únicamente en la geografía para determinar la vecindad entre comunas. Es importante contar con ambas versiones de matrices de adyacencias, porque son variaciones que se consideran al momento de implementar los modelos de factibilidad.

2.4. Cálculo de camino más corto entre las comunas de Chile

Uno de los aspectos cruciales es determinar el camino más corto entre cada par de comunas, porque al momento de garantizar contigüidad entre las comunas de un distrito se utilizan estas rutas. Para llevar a cabo este proceso de manera eficiente, se emprendieron tareas previas de preprocesamiento de datos.

Para calcular estos caminos más cortos, se recurrió a la función `k_shortest_paths` de NetworkX [4], que determina las rutas más cortas entre puntos, y también la función `Snuevo(0)` donde si el camino trazado entre dos comunas consta de más de tres nodos, retornará el predecesor en el camino más corto. Por otro lado, si el camino tiene menos nodos, la función retornará un valor vacío. Lo último indica que las dos comunas en cuestión son vecinas, es decir, no hay comunas intermedias entre ellas. Dada la cantidad de los datos, teniendo en cuenta que la lista de comunas R contiene las 346 comunas de Chile, calcular el camino más corto entre cada par de comunas en cada ejecución consume un tiempo considerable. Se ha optado por elaborar un diccionario para agilizar cálculos recurrentes. En este, cada clave corresponde a una combinación de pares (i, j) presentes en R . Los valores asociados a estas claves son listas que contienen los resultados de la función `Snuevo` para dicha combinación.

Una vez creado, el diccionario se almacena en un bloc de notas para facilitar su acceso y consulta. El diccionario es de tipo $(key, value)$, donde cada llave es un par de comunas. Existen dos versiones de este diccionario: una que respeta los límites regionales al determinar las comunas adyacentes, y otra que ignora dichos límites.

La principal ventaja de este enfoque es la eficiencia. En lugar de reevaluar las distancias en cada instancia, basta con consultar el diccionario para obtener el resultado asociado al par (i, j) . Gracias a esta estrategia, lo que solía requerir horas para calcular, ahora se completa en segundos, lo que representa una mejora sustancial en términos de rendimiento y tiempo de ejecución.

La implementación de las funciones `k_shortest_paths` de Network X y `Snuevo` se pueden consultar en pseudocódigo en el Apéndice A.2.2 .

2.5. Calidad de balance poblacional en el distritaje actual para Chile

Finalmente se requiere el cálculo de una métrica esencial: el balance poblacional, denotado por ε , de un distritaje. Esta métrica nos indica la variabilidad de la población entre los diferentes distritos. La importancia de calcular este valor radica en su utilidad para la toma de decisiones. A partir de aquello, es que se requiere calcular esta métrica para el distritaje actual de Chile que denotaremos por ε_{BPA} . Al comparar el balance poblacional del distritaje actual de Chile (ε_{BPA}) con el obtenido de los modelos de factibilidad, se puede determinar si los modelos propuestos ofrecen una distribución más equitativa de la población o si, por el contrario, presentan desequilibrios mayores.

La determinación del valor del balance poblacional para Chile ε_{BPA} se lleva a cabo tomando como referencia la estructura actual de distritos de Chile [1]. A continuación, se detalla el proceso seguido para su cálculo:

1. Selección del centro de distrito: Inicialmente, para cada distrito se identifica la comuna con la mayor población, designándola como el ‘centro’ del distrito. Consecuentemente, todas las demás comunas de ese distrito se consideran vinculadas o asignadas a ese centro específico.
2. Integración de datos demográficos: Una vez definidos los centros y las asignaciones, se asocia con la información de la base de datos mencionada anteriormente 2.1. Este paso es crucial para determinar la población total de cada distrito. Posteriormente, se calcula la población promedio de los distritos, que servirá como referencia en el cálculo del balance poblacional actual de Chile ε_{BPA} .
3. El valor de ε_{max} se calcula utilizando la siguiente fórmula:

$$\varepsilon_{\text{max}} = \max_{i \in \{1, \dots, k\}} \varepsilon_i \quad (2.3)$$

Con ε_i definido como:

$$\varepsilon_i = \frac{|p_i - \hat{p}|}{\hat{p}} \quad (2.4)$$

- ε_{max} es el valor máximo entre todos los distritos de ε .
- ε_i es el error de balance poblacional específico para un distrito i .
- p_i es la población de un distrito específico.
- $\hat{p}$ es la población promedio de los distritos. Definido como $\frac{p_t}{K}$, donde p_t es la población total del país y K el número de distritos.

2.5. Calidad de balance poblacional en el distritaje actual para Chile

Este cálculo busca identificar el distrito cuya población se desvía más significativamente de la media. En el caso de Chile, el número de distritos como $K = 28$ y la población total como $p_t = 17518991$.

Este indicador es fundamental para asegurar que los derechos y la representación de los ciudadanos sean uniformes a lo largo del país, evitando así zonas sobrerrepresentadas o subrepresentadas. En el caso de Chile se obtiene que $\varepsilon_{\max} \approx 1,3299$, que refleja la variabilidad en la población de cada distrito en el peor caso. Este valor es el que conoceremos como error de balance poblacional actual para Chile.

Capítulo 3

Modelos de programación entera

La programación entera se ocupa para resolver problemas de optimización en los que algunas o todas las variables de decisión deben tomar valores enteros. Los valores enteros de las variables de decisión representan cantidades discretas, como el número de comunas a asignar a un distrito. La restricción de enteridad se puede añadir a un problema de programación lineal para modelar la realidad de que algunas cantidades solo pueden ser enteras. Los problemas de programación entera son generalmente más difíciles de resolver que los problemas de programación lineal. Esto se debe a que el conjunto de soluciones factibles de un problema de programación entera es discreto, mientras que el conjunto de soluciones factibles de un problema de programación lineal es continuo. Esta técnica es especialmente útil cuando se abordan problemas en los que las soluciones fraccionarias no tienen sentido práctico, como en el caso de asignación de bloques a distritos. Por ejemplo, no es posible asignar 1,5 bloques a un distrito, o 0,25 de un bloque a otro. La unidad mínima territorial es un bloque, y esta es indivisible. Por lo tanto, una solución que sugiera fracciones de un bloque no sería aplicable en la realidad.

Cabe señalar que, aunque la programación entera garantiza soluciones aplicables y realistas, también puede ser computacionalmente más desafiante que la programación lineal estándar, debido a que esta última si permite resultados continuos. Los algoritmos para resolver problemas de programación entera suelen ser más complejos y pueden requerir más tiempo, especialmente para grandes conjuntos de datos. Sin embargo, el grafo y la división político administrativa de Chile posee características que pueden favorecer los tiempos de computación. En primer lugar, al momento de realizar una vista desde lejos o “zoom out” del grafo, se observa que el grafo parece una línea o camino. En segundo lugar, Chile, con sus 346 comunas como unidad mínima territorial, presenta un escenario distinto en comparación con Estados Unidos, que cuenta con 3243 condados o equivalentes. Estas diferencias en la estructura territorial tienen importantes implicancias en los modelos estudiados. Modelar problemas con un número menor de nodos, como es el caso de las comunas chilenas, puede ser computacionalmente más sencillo. Un menor número de nodos suele traducirse en menos

restricciones en los modelos, lo que facilita la búsqueda de soluciones óptimas en tiempos más cortos.

Las soluciones serán llamadas *distritajes*. Este término se utiliza para describir el proceso de subdividir un área geográfica o una población con el propósito de organizar y distribuir el ejercicio de los derechos civiles y políticos, así como las funciones públicas y los servicios administrativos.¹ Una solución factible está definido por K distritos y forma una K -partición de las comunas, es decir, donde la unión de todas las subregiones es igual a R . Además, se exige obtener particiones geográficamente contiguas y equilibradas en población. Es decir, el subgrafo definido por un distrito debe ser conexo. Finalmente, se exige también que deben poseer aproximadamente la misma población.

Ambos modelos a presentar se sustentan en lo realizado por Gurnee et al. [11], y se necesita como *input* un conjunto R de bloques o unidades básicas censales, un conjunto P con la población de cada uno de dichos bloques, D con la distancia entre cada uno de los bloques censales y lista de adyacencia de cada uno de los bloques. Se deben entender los siguientes parámetros:

- $d_{i,j}$: Distancia entre comunas i y j .
- p_j : Población de comuna j .
- $\hat{p}$: Población promedio según la cantidad de distritos que se está estudiando.
- ε_{BP} : Es el error de balance poblacional. Este valor es el que señala la desigualdad en la representación. Se entrega como insumo al modelo para que genere un distritaje factible si es posible, y sino dirá que es infactible para dicho valor.
- $pred_j(i)$: Comuna antecesora, es decir, ubicada justo antes de la comuna i en el camino más corto entre i y j . Se calcula con las funciones `Snuevo` y `k_shortest_path` [ver Sección 2.4 y Anexo A.1].

Los modelos de factibilidad tienen el mismo propósito de crear distritajes que cumplan con una serie de condiciones, como la contigüidad geográfica y equilibrio poblacional. Sin embargo, se diferencian en algunos aspectos importantes:

- Modelo de factibilidad con centros por definir: En este modelo, los centros son definidos por el modelo en su ejecución. Esto significa que el modelo tiene que considerar dos variables de decisión:
 - $x_{i,j}$: Variable binaria que indica si la comuna i está asignada a la comuna j .
 - y_j : Variable binaria que indica si la comuna j es un centro.

¹<https://dle.rae.es/distrito?m=form>

3.1. Modelo de factibilidad con centros fijos

- Modelo de factibilidad con centros fijos: En este modelo, los centros son definidos previamente fuera del modelo. Esto significa que el modelo solo tiene que considerar una variable de decisión:

– $x_{i,j}$: Variable binaria que indica si la comuna i está asignada al centro j .

3.1. Modelo de factibilidad con centros fijos

El modelo debe dividir una población total en un número específico de distritos. La cantidad de distritos se determina en función de una lista predefinida de centros geográficos proporcionados como entrada al modelo. Cada centro geográfico actúa como un punto de referencia desde el cual se extiende un distrito para incluir una porción de la población total. A continuación, se presenta el modelo:

$$\sum_{j \in C} x_{i,j} = 1 \quad \text{para toda comuna } i \in R, \quad (3.1)$$

$$\sum_{i \in R} p_j x_{i,j} \leq \hat{p}(1 + \varepsilon_{BP}) \quad \text{para todo centro } j \in C, \quad (3.2)$$

$$\sum_{i \in R} p_j x_{i,j} \geq \hat{p}(1 - \varepsilon_{BP}) \quad \text{para todo centro } j \in C, \quad (3.3)$$

$$x_{i,j} \leq x_{\text{pred}_j(i),j} \quad \text{para toda comuna } i, \text{ para todo centro } j \text{ cuando } \text{pred}_j \neq \emptyset. \quad (3.4)$$

$$x_{i,j} \in \{0, 1\} \quad \text{para todo centro } j \in C, \quad \text{para toda comuna } i \in R, \quad (3.5)$$

donde (3.1) impone que cada bloque este asignado a exactamente un centro. Las restricciones (3.2) y (3.3) aseguran el equilibrio de población. Se permite una variación entre $(1 + \varepsilon_{BP})$ y $(1 - \varepsilon_{BP})$ en la población de cada distrito. La restricción (3.4) hace cumplir la contigüidad. La explicación es, para que una comuna i esté conectada a un centro j , todas las comunas en el camino más corto entre i y j también deben estar conectadas. No se permite tener una conexión directa entre una comuna y un centro sin asegurar que todas las comunas en el camino más corto también estén conectadas. Por último, (3.5) es la naturaleza de la variable, donde se indica que la variable de decisión puede tomar el valor de 1 si es que la comuna i está asignada al centro j , y 0 sino.

3.2. Modelo de factibilidad con centros por definir

El propósito de este modelo es llevar a cabo la división de una población total en un número determinado de distritos. La cantidad de distritos, representada por la constante K ,

3.3. Restricción de límites regionales

es parte esencial de la entrada del modelo. En este contexto, el objetivo principal del modelo es identificar y establecer los K centros geográficos óptimos para generar los distritos que cumplan con las restricciones especificadas.

$$\sum_{j \in R} x_{i,j} = 1 \text{ para toda comuna } i \in R, \quad (3.6)$$

$$\sum_{i \in R} p_i x_{i,j} \leq \hat{p}(1 + \varepsilon_{BP}) y_j \quad \text{para todo centro } j \in R, \quad (3.7)$$

$$\sum_{i \in R} p_i x_{i,j} \geq \hat{p}(1 - \varepsilon_{BP}) y_j \quad \text{para todo centro } j \in R, \quad (3.8)$$

$$x_{j,j} = y_j \text{ para todo centro } j \in R, \quad (3.9)$$

$$\sum_{j \in R} y_j = K, \quad (3.10)$$

$$x_{i,j} \leq y_j \text{ para todo centro } j \in R, \text{ para toda comuna } i \in R, \quad (3.11)$$

$$x_{i,j} \leq x_{\text{pred}_j(i),j} \text{ para toda comuna } i, \text{ para todo centro } j \text{ cuando } \text{pred}_j \neq \emptyset. \quad (3.12)$$

$$x_{i,j} \in \{0, 1\} \quad \text{para todo centro } j \in R, \quad \text{para toda comuna } i \in R, \quad (3.13)$$

$$y_j \in \{0, 1\} \quad \text{para todo centro } j \in R \quad (3.14)$$

La restricción (3.6) impone que cada bloque este asignado a una subregión. Las restricciones (3.7) y (3.8) aseguran el equilibrio de población. Se permite una variación entre $(1 + \varepsilon_{BP})$ y $(1 - \varepsilon_{BP})$ en la población de cada subregión. La restricción (3.9) impone que si un bloque es centro debe estar asignado al distrito del cuál es centro. La restricción (3.10) impone que los centros deben ser igual a K (un parametro fijo y dado para el modelo). La restricción (3.11) impone que en caso de asignar bloques a un centro, ese bloque debe estar abierto como centro. La restricción (3.12) hace cumplir la contigüidad de manera análoga al modelo anterior 3.1. Por último, se tiene la naturaleza de las variables en (3.13) y (3.14). En el caso de (3.13) indica que la variable de decisión puede tomar el valor de 1 si es que la comuna i está asignada al centro j , y 0 sino. Mientras que, en el caso de (3.14), su valor es 1 cuando indica que la comuna j es un centro, y 0 sino.

3.3. Restricción de límites regionales

Para los modelos previamente expuestos, que permiten la formación de distritos con comunas de diferentes regiones, es necesario incorporar una restricción adicional cuando se busca limitar los distritos a conformaciones que incluyan únicamente comunas de la misma

3.4. Búsqueda binaria para el balance poblacional óptimo

región. La restricción se plantea de la siguiente manera:

$$x_{i,j} = 0 \text{ para todo } j \in C, \text{ para toda } i \in R, \text{ que pertenezcan a distintas regiones.} \quad (3.15)$$

Se debe verificar, para cada comuna i y cada centro potencial j , si la región de la comuna es diferente a la región del centro. En caso de que pertenezcan a regiones distintas, se debe establecer que la variable binaria que indica la asignación de la comuna al centro debe ser igual a cero. Esto asegura que las comunas solo pueden ser asignadas a centros dentro de su misma región.

Es importante destacar que la restricción 3.15, tal como está formulada, es válida para la sección 3.1 del modelo, ya que en ese contexto se itera sobre el conjunto de centros predefinidos C . Sin embargo, para que esta restricción sea aplicable a la sección 3.2, es necesario realizar una modificación que podemos observar a continuación:

$$x_{i,j} = 0 \text{ para todo } j \in R, \text{ para toda } i \in R, \text{ que pertenezcan a distintas regiones.} \quad (3.16)$$

En este caso, en lugar de utilizar el conjunto de centros C , se debe utilizar el conjunto de todas las comunas, que es representando como R (recordemos que en este caso el modelo debe definir los centros dentro del conjunto de todas las comunas al momento de la ejecución). Esto garantizará que la restricción sea coherente con el contexto en el que se aplica.

3.4. Búsqueda binaria para el balance poblacional óptimo

A continuación, se detalla un pseudocódigo que describe una estrategia de búsqueda binaria para encontrar un aproximado al valor mínimo de ε_{BP} que hace factible los problemas de factibilidad abordados anteriormente en las secciones 3.1 y 3.2. Esta estrategia se basa en calcular promedios iterativamente dentro de un rango de valores establecidos inicialmente, hasta que se encuentre el mínimo bajo cierta convergencia. La convergencia de ε_{BP} se determina por la cercanía entre el valor actual de ε_{BP} y el valor anterior, dentro de una tolerancia específica para el problema, o por otro lado cuando se alcanza un cierto número predeterminado de iteraciones.

Algorithm 1 Factibilidad mínima de ε_{BP}

Input: $\varepsilon_{BP}, \varepsilon_0 \triangleright$ Errores de balance poblacional entre los cuales se comprobará factibilidad

Output: $lista_fact$ $\triangleright$ Lista con valores factibles e infactibles entre ε_{BP} y ε_0

```
1:  $\varepsilon_{BP} \leftarrow 2$ 
2:  $\varepsilon_0 \leftarrow 0$ 
3: Crear lista vacía llamada  $lista\_fact$ 
4: Resolver el problema de factibilidad para  $\varepsilon_{BP}$ 
5: if el problema de factibilidad para  $\varepsilon_{BP}$  es factible then
6:   Agregar  $\varepsilon_{BP}$  a  $lista\_fact$  como factible
7:   while  $\varepsilon_{BP}$  no converja do
8:      $promedio \leftarrow (\varepsilon_{BP} + \varepsilon_0)/2$ 
9:     Resolver el problema de factibilidad para  $promedio$ 
10:    if el problema para  $promedio$  es factible then
11:      Agregar  $promedio$  a  $lista\_fact$  como factible
12:       $\varepsilon_{BP} \leftarrow promedio$ 
13:    else
14:      Agregar  $promedio$  a  $lista\_fact$  como infactible
15:       $\varepsilon_0 \leftarrow promedio$ 
16:    end if
17:  end while
18: else
19:   print 'El modelo es infactible'
20: end if
```

Capítulo 4

Simulaciones y resultados

Actualmente, Chile ha implementado un distritaje que respeta los límites regionales, lo que significa que un distrito no puede abarcar comunas pertenecientes a regiones diferentes. Además, posee un error de balance poblacional $\varepsilon_{BP} \approx 1,3299$, que refleja la variabilidad en la población de los distritos en el peor caso y señala la desigualdad en la representación. Para entenderlo mejor, se busca que cada distrito tenga, idealmente, la misma población para que cada persona tenga igual representación. Sin embargo, puede que algunos distritos tengan más personas y otros menos. Entonces, ε nos ayuda a medir esta diferencia. Si el valor de ε es 0, significa que todos los distritos tienen exactamente la misma cantidad de gente (representación perfectamente igualitaria). Si ε es un número mayor que 0, nos indica que hay una diferencia entre el distrito con más personas o el distrito con menos personas, respecto a la población promedio según la cantidad de distritos.

Por otro lado, Chile ha estructurado su sistema electoral en 28 distritos [1]. Estos distritos son unidades territoriales que sirven como base para las elecciones parlamentarias, permitiendo la elección de representantes para el Congreso Nacional. Cada distrito elige un determinado número de diputados dependiendo de su población. Los 28 distritos de Chile son una pieza fundamental del sistema democrático del país, jugando un rol crucial en la forma en que se toman decisiones a nivel nacional.

Las simulaciones realizadas fueron ejecutadas en un computador equipado con un procesador Intel(R) Core(TM) i5-7200 CPU @ 2.50 GHz, 8 GB de memoria RAM. Se utilizó Python 3.11 y Gurobi 10.0.1 para llevar a cabo los análisis y cálculos pertinentes.

Es importante destacar que al representar gráficamente los distritajes propuestos, se opta por excluir Isla de Pascua y Juan Fernández, esta decisión se basa en la considerable distancia geográfica que estas localidades mantienen con respecto al resto del país, lo que resulta en la generación de grandes espacios vacíos en el mapa. La exclusión de estas regiones contribuye a una representación más significativa de los distritos, mejorando la claridad de la visualización

de los distritajes propuestos.

Además, es importante notar que, en algunas de las figuras presentadas puede haber similitud entre los colores utilizados. No obstante, esta similitud no implica la existencia de distritos desconexos, más bien es resultado de colores cercanos en la paleta utilizada.

4.1. Respetando límites regionales y 28 distritos

En esta oportunidad se utilizan los límites regionales, es decir, se respeta la demarcación actual de las regiones de Chile al conformar los distritos. Para lograr respetar los límites regionales se debe utilizar la matriz de adyacencia que respeta dichos límites y agregar la restricción 3.15, que fuerza sólo asignaciones de la misma región. Esta consideración se hace con el objetivo de que la comparativa sea más ajustada a la estructura actual del país. En la misma línea de ajustar la comparativa es que se establece la cantidad de distritos en 28. El distritaje actual de Chile se puede observar en la Figura 4.1 y en la Figura 4.2 para la Región Metropolitana.

4.1.1. Centros por definir

Cuando se establecen 28 distritos donde el modelo es el encargado de escoger los centros y respetando límites regionales, el modelo arroja un distritaje factible con un error de balance poblacional de $\varepsilon_{BP} \approx 0,8398$ en aproximadamente 234 segundos. Este nuevo valor refleja una notable reducción en la variabilidad de la población entre distritos en comparación con el sistema actual de Chile. En la Figura 4.3 se observa el distritaje propuesto bajo estas condiciones para la Región Metropolitana y en la Figura 4.4 para todo Chile.

La magnitud de esta mejora en equidad de representación, que es aproximadamente $2/3$ del valor actual para Chile, abre un debate respecto a, ¿Qué consideraciones son las que se han tenido al momento del diseño de distritos? Es conocido que se hace en términos de igualdad y proporcionalidad, pero esto evidencia que no se aproxima a la solución óptima y que existen herramientas para que las decisiones sean tomadas con respaldo.

4.1.2. Centros fijos

Para este caso se conserva la misma cantidad de distritos(es decir, 28) y se establecen centros fijos previamente para cada uno de los distritos. Estos centros están determinados por las comunas con la mayor población dentro de los actuales distritos en el país, la lista de centros es la siguiente:

4.1. Respetando límites regionales y 28 distritos

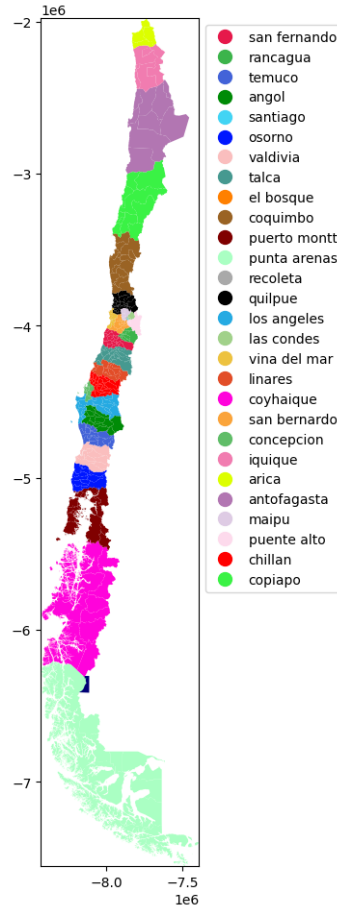


Figura 4.1: Distritaje actual para Chile, respeta límites regionales con $\varepsilon_{BP} = 1,3299$ y $K = 28$

- | | | | |
|----------------|----------------|----------------|----------------|
| • Arica | • Maipú | • Rancagua | • Angol |
| • Iquique | • Recoleta | • San Fernando | • Temuco |
| • Antofagasta | • Santiago | • Talca | • Valdivia |
| • Copiapó | • Las Condes | • Linares | • Osorno |
| • Coquimbo | • Puente Alto | • Chillan | • Puerto Montt |
| • Quilpué | • El Bosque | • Concepción | • Coyhaique |
| • Viña del mar | • San Bernardo | • Los Ángeles | • Punta Arenas |

El resultado en esta ocasión es $\varepsilon_{BP} \approx 0,8398$ en 24 segundos. Esta cifra resulta inespablemente igual al resultado obtenido en el caso de la Sección 4.1.1. Aunque los modelos se basan en diferentes criterios para seleccionar centros, porque en un caso el modelo debe definir los centros y en el caso actual los centros son un insumo del modelo, el impacto global

4.1. Respetando límites regionales y 28 distritos

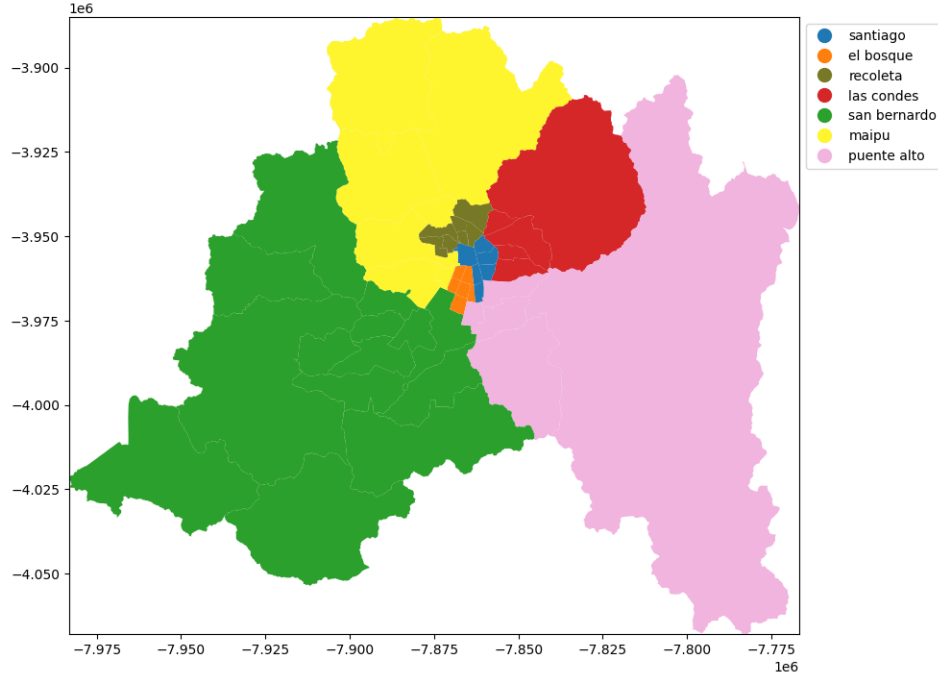


Figura 4.2: Distritaje actual para la Región Metropolitana, respeta límites regionales con $\varepsilon_{BP} = 1,3299$ y $K = 28$

sobre el equilibrio poblacional es idéntico. Sin embargo, la composición de cada uno de los distritos es diferente. Dada la igualdad alcanzada en el error de balance poblacional ε_{BP} , se visualizará el distritaje propuesto, comenzando con la Región Metropolitana en la Figura 4.5 y luego extendiéndose a todo el país en la Figura 4.6.

En el caso de la Región Metropolitana se puede notar de manera más evidente la diferencia de composición en el distritaje pese a que el error de balance poblacional ε_{BP} es el mismo. En el caso donde el modelo debe definir los centros 4.1.1 se establecen 9 distritos (ver Figura 4.3) y con los centros fijos se establecen 7 distritos (ver Figura 4.5). Existen varias hipótesis que podrían explicar esta igualdad en el error de balance poblacional ε_{BP} :

- Proximidad geográfica de los centros: Es posible que para ambos modelos se hayan seleccionado centros que, aunque diferentes, estén geográficamente cerca unos de otros. Si estos centros están en ubicaciones con densidades poblacionales similares, podrían resultar en valores de ε_{BP} cercanos.
- Distribución poblacional: Si las áreas con alta población están distribuidas de manera que hay varios centros potenciales cercanos que pueden servir a esas poblaciones, entonces diferentes modelos podrían seleccionar centros ligeramente diferentes pero todavía alcanzar resultados similares en términos de equilibrio poblacional.
- Capacidad de los centros: A causa de que para ambos casos el valor de la población

4.2. Sin límite regional y 28 distritos

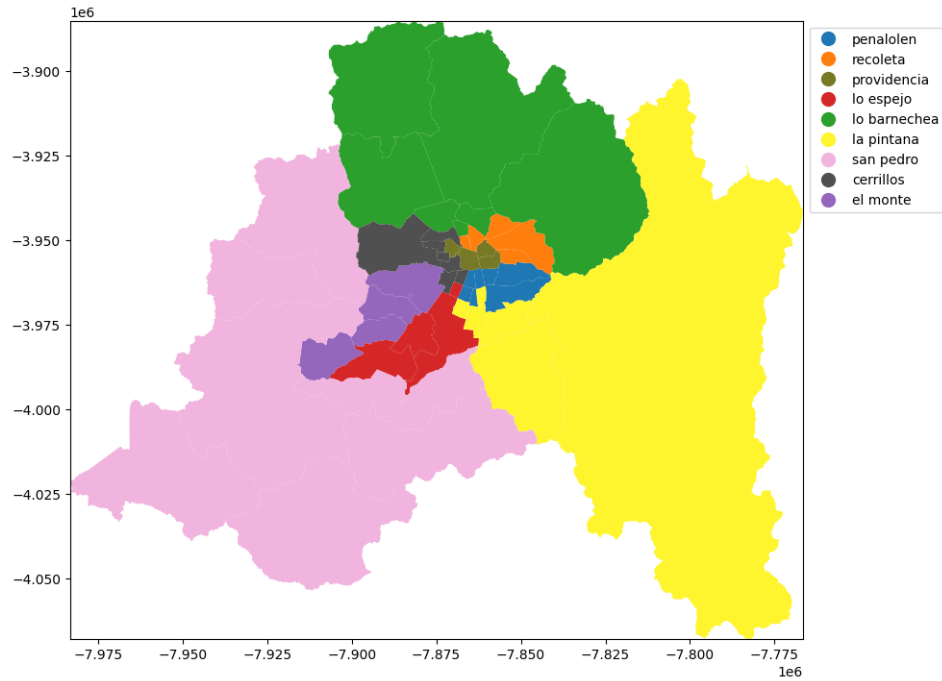


Figura 4.3: Distritaje para la Región Metropolitana respetando límites regionales con $\varepsilon_{BP} = 0,8398$, $K = 28$ y centros por definir.

promedio es igual, la capacidad que cada centro posee también debería ser similar por lo que pueden servir a poblaciones comparables, llevando a ε_{BP} similares.

- Restricción del peor caso: La Región de Aysén es la región del país que tiene la menor población [1] e impone el límite inferior al momento de equilibrar la población. Puede ser que en ambos casos sea esta restricción la que define el error de balance poblacional ε_{BP} .

La situación anterior sugiere que, al menos en este escenario, hay cierta flexibilidad en la selección de centros sin comprometer el equilibrio poblacional o mejor dicho el error de balance poblacional ε_{BP} , ya que de igual manera se obtienen resultados de manera óptima. La diferencia recae sobre el tiempo debido a que cuando el modelo debe asignar los centros 4.1.1 le toma diez veces más tiempo que para el caso actual. De cualquier manera esto sugiere que, hay margen para mejorar la equidad en la representación actual del país.

4.2. Sin límite regional y 28 distritos

Con modelos de factibilidad sin considerar el límite regional, se abre la posibilidad de formar distritos que combinen comunas de diferentes regiones, esto permite más opciones de

4.2. Sin límite regional y 28 distritos

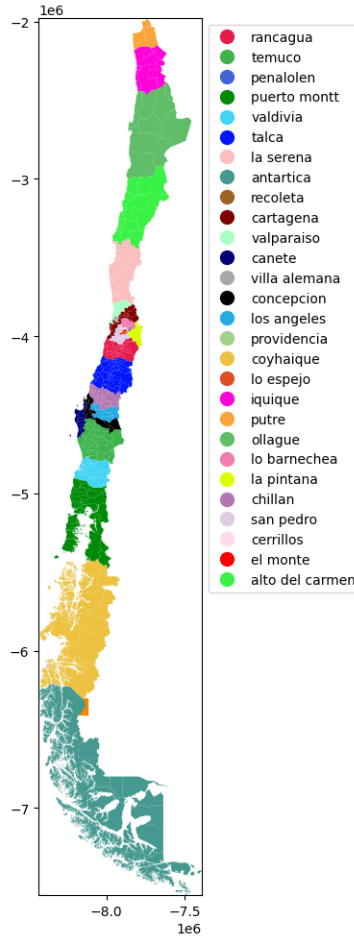


Figura 4.4: Distritaje para Chile respetando límites regionales con $\varepsilon_{BP} = 0,8398$, $K = 28$ y centros por definir.

como conformar cada uno de los distritos. Sin embargo, se mantienen el número actual de distritos para poder tener punto de comparación con el sistema actual.

4.2.1. Centros por definir

La implementación de este caso arroja un distritaje factible con un error de balance poblacional $\varepsilon_{BP} \approx 0,039$. Este valor es sorprendentemente pequeño. Sin embargo, en un modelo de factibilidad donde no existen límites regionales y el propio modelo es responsable de seleccionar los mejores centros, se espera una mejora significativa en ε_{BP} , esto se debe a que se abren nuevas posibilidades:

- Libertad de Selección de Centros: Al eliminar las restricciones de límites regionales, el

4.2. Sin límite regional y 28 distritos

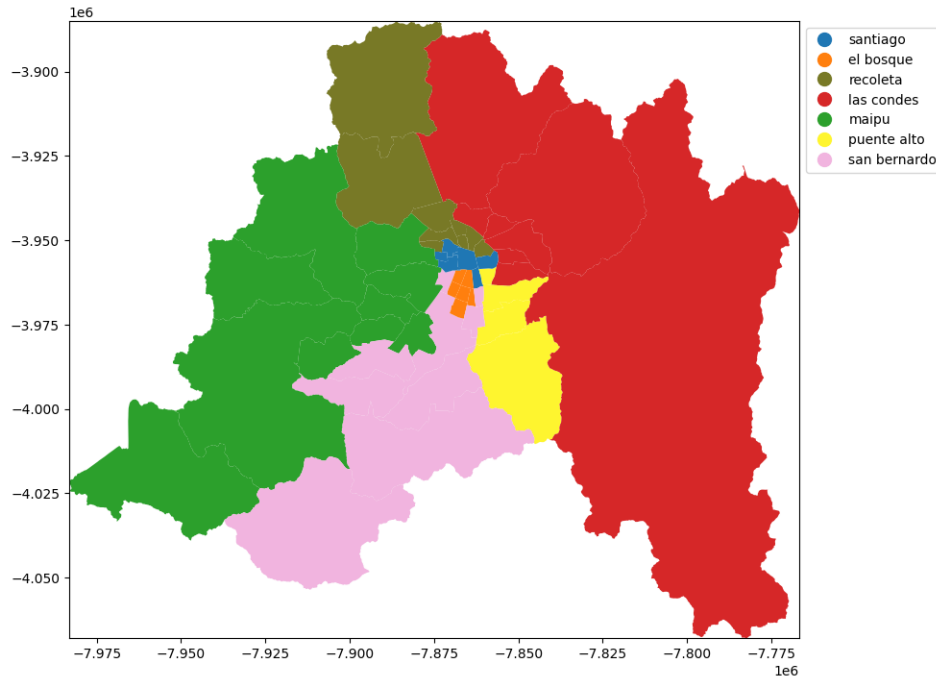


Figura 4.5: Distritaje para la Región Metropolitana respetando límites regionales con $\varepsilon_{BP} = 0,8398$, $K = 28$ y centros fijos.

modelo tiene una mayor flexibilidad para escoger centros en ubicaciones estratégicas, sin estar acotado por límites geográficos. Esto significa que puede optar por centros que estén exactamente en lugares óptimos para equilibrar la distribución poblacional.

- Optimización Global en Lugar de Local: En muchos casos, las restricciones regionales pueden conducir a una optimización “local”, donde se buscan soluciones óptimas dentro de una región particular, pero no necesariamente a nivel global. Al eliminar estos límites, el modelo puede considerar el panorama global, resultando en una mejor solución general.
- Eliminación de Múltiples Centros Cercanos: En un escenario con restricciones regionales, es posible que se requieran múltiples centros en áreas cercanas simplemente porque están en diferentes regiones. Al eliminar estas restricciones, el modelo puede identificar y evitar redundancias, optando por ubicaciones que sirvan a áreas más amplias y evitando la cercanía innecesaria de centros.

Un $\varepsilon_{BP} \approx 0,039$ indica una diferencia pequeña en la distribución de la población entre distritos. Además, el valor obtenido indica que la variación de la población entre distritos es cercana al mínimo, por lo que la representación está cerca de ser perfectamente igualitaria. En la Figura 4.7 se observa como se ve el distritaje para todo Chile.

En esta etapa de creación de distritos, es evidente que se han generado distritos con

4.2. Sin límite regional y 28 distritos

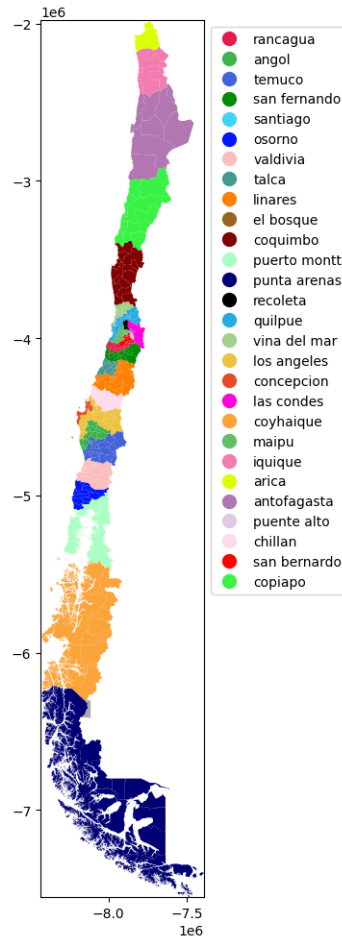


Figura 4.6: Distritaje para Chile respetando límites regionales con $\varepsilon_{BP} = 0,8398$, $K = 28$ y centros fijos.

formas irregulares y notables diferencias entre ellos. Es especialmente notable que en el sur del país existe un distrito que agrupa alrededor de tres regiones. Esta agrupación se debe principalmente a la baja población en dichas regiones, siendo una de ellas la que cuenta con la menor población entre las 16 regiones del país. De manera similar, en el norte del país, también encontramos regiones agrupadas en un solo distrito debido a las pequeñas poblaciones que albergan.

Es difícil observar la zona central en la Figura 4.7 que representa todo el país. Sin embargo, es de gran interés observar lo que ocurre en la Región Metropolitana, considerando las condiciones y libertades mencionadas. Por lo tanto, en la Figura 4.8, observaremos el distritaje para la Región Metropolitana.

Cabe destacar que, pese a las similitudes en algunos colores utilizados, representan distintos distritos. Además, se puede observar que en la simbología se encuentran nombres de

4.2. Sin límite regional y 28 distritos

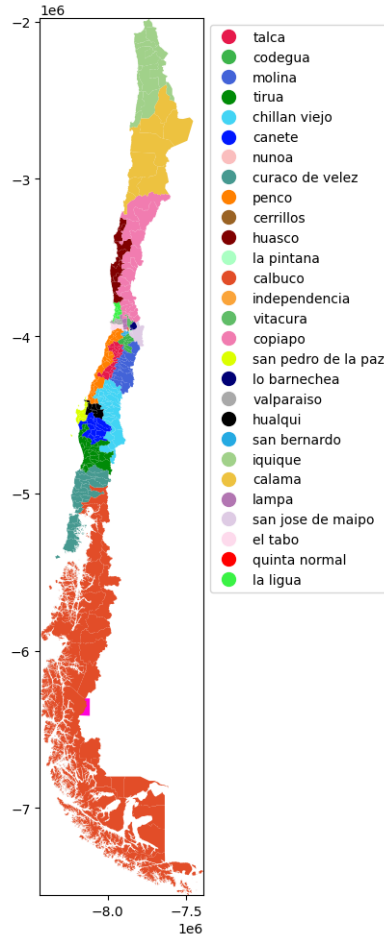


Figura 4.7: Distritaje para Chile sin límites regionales con $\varepsilon = 0,039$, $K = 28$ y centros por definir.

comunas que no pertenecen a la Región Metropolitana. Esto se debe a que la flexibilidad de permitir la creación de distritos sin necesariamente respetar los límites regionales conlleva a que, en algunos casos, el modelo pueda considerar que la posición óptima para establecer un centro se encuentre en una región diferente, pero se permite que incluya comunas de la Región Metropolitana. Además, se puede mencionar que la zona central concentra la mayoría de los distritos, llegando a observarse 13 de ellos en la Figura 4.8, lo que equivale a casi la mitad de los distritos disponibles para distribuir en todo el país. Sin embargo, dichas asignaciones tienen bastante sentido porque esta área cuenta con la mayor densidad de población en el país, con numerosas comunas densamente pobladas.

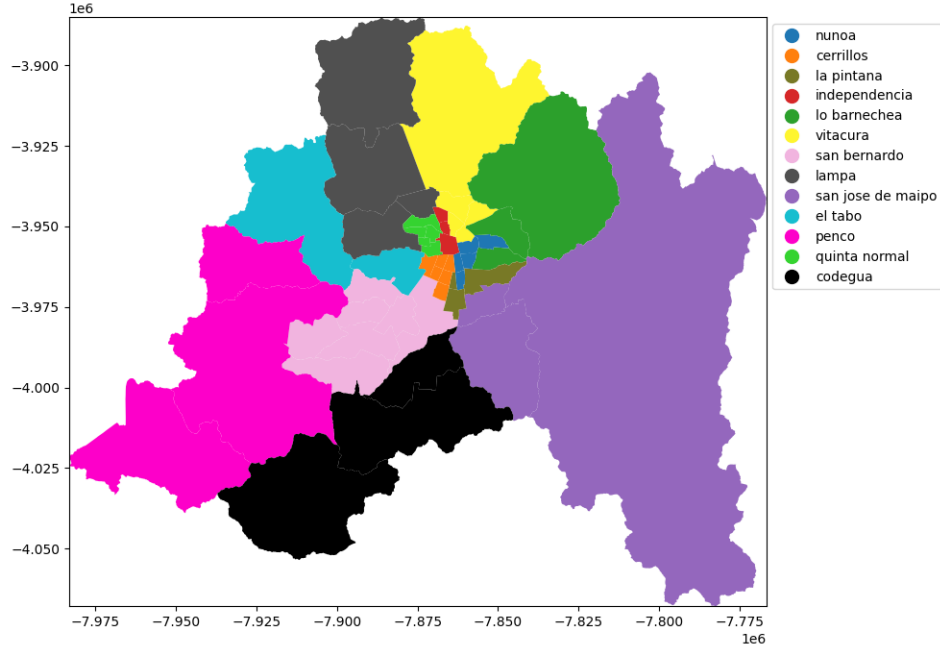


Figura 4.8: Distritaje Región Metropolitana sin límites regionales con $\varepsilon = 0,039$, $K = 28$ y centros por definir

4.2.2. Centros fijos

En esta sección, se establecen centros fijos para cada uno de los distritos. Estos centros están determinados nuevamente por las comunas con la mayor población dentro de los actuales distritos en el país, del mismo modo que en la Subsección 4.1.2.

La implementación de este nuevo modelo se completó en un tiempo breve de aproximadamente 13.03 segundos, y el resultado es un distritaje con $\varepsilon_{BP} \approx 0,6796$. El error de balance poblacional ε_{BP} para el distritaje indica una diferencia más significativa en la distribución poblacional entre los distritos en comparación al escenario previamente mencionado en la Subsección 4.2.1. Esta mayor disparidad, aunque no es ideal, es comprensible dadas las restricciones impuestas al modelo al definir los centros previamente. La decisión de definir centros y que sea uno de cada distrito actual del país, impone que haya al menos uno por región lo que introduce un nivel de rigidez en la solución, pese a que no esté agregado como restricción que debe existir un centro o distrito por región. Esto puede llevar a resultados subóptimos por varias razones:

- **Distribución Desigual de la Población:** Algunas regiones son considerablemente más pequeñas en términos de población que una sola comuna de la Región Metropolitana. Por ejemplo, toda la región de Aysén posee una población de 103158 y solamente la comuna de Puente Alto en la Región Metropolitana posee 568106 habitantes. Al

4.2. Sin límite regional y 28 distritos

imponer un centro en cada región, sin tener en cuenta estas grandes diferencias de población, se está creando un desequilibrio inherente.

- Limitaciones en la Optimización y Rigidez: Al dar centros fijos con las características mencionadas se introducen restricciones adicionales en el modelo, por lo que se reduce el espacio de soluciones posibles. En consecuencia, el modelo podría no ser capaz de encontrar la solución verdaderamente óptima debido a estas restricciones que obligan a la presencia de un centro por región.

Al observar la Figura 4.9 de Chile, no existen agrupaciones notorias de más de una región en un solo distrito. Esto se debe a que, como se mencionó en la selección de centros, se ha asignado una comuna de cada distrito actual del país, por lo que existe al menos un centro por región. A continuación, se presenta el gráfico mencionado.

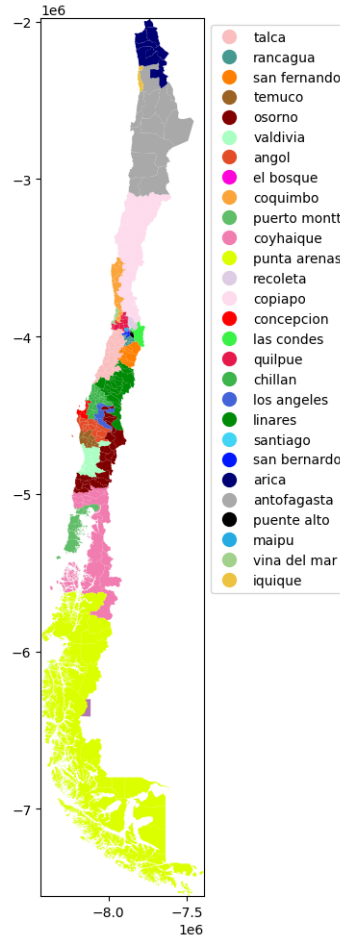


Figura 4.9: Distritaje para Chile sin límites regionales con $\varepsilon = 0,6796$, $K = 28$ y centros fijos.

4.2.3. Centros por definir integrando función objetivo para imponer compacidad

La optimización de los distritajes con el objetivo de lograr una forma visualmente compacta para cada uno de los distritos presenta desafíos computacionales significativos. Se ha propuesto la inclusión de una función objetivo, al igual que en Gurnee et al. [11], de la forma.

$$\min \sum_{i \in R} \sum_{j \in C} (d_{ij})^\alpha p_i x_{ij} \quad (4.1)$$

Dicha función objetivo busca minimizar la suma de las distancias elevadas a la potencia α , ponderadas por la población de cada comuna y la variable de decisión asociada. El parámetro α controla la influencia de las distancias en la función objetivo. La inclusión de este término elevado a la potencia α tiene como objetivo penalizar de manera más significativa las asignaciones que resulten en distritos con formas menos compactas con la intención de favorecer la formación de distritos más circulares y evitar formas irregulares, como puntas o comunas atrapadas.

No obstante, la aplicación de esta función objetivo ha resultado en una ineficiencia computacional. Los tiempos de ejecución por cada error de balance poblacional ε_{BP} , es decir, por cada iteración, han alcanzado hasta 456280 segundos, con un promedio de 59199 segundos, lo que equivale a aproximadamente 16 horas. Este tiempo prolongado de ejecución representa un desafío importante, especialmente cuando se busca explorar diversas configuraciones para mejorar la compacidad de los distritos.

En el caso específico de Chile, los resultados obtenidos al incorporar la función objetivo no han demostrado cambios significativos que justifiquen el esfuerzo computacional adicional. A pesar de la inclusión de la función objetivo, aún se observan distritos alargados, similar a los resultados obtenidos sin su aplicación. Es importante destacar que estas pruebas se llevaron a cabo en un escenario en el cual el modelo debe definir los centros, sin límites regionales, esta decisión es debido a que se ha observado que la inclusión de estos límites proporciona cierta compacidad inherente a los resultados.

En resumen, aunque la idea de la función objetivo para fomentar la compacidad de los distritos es válida, los desafíos computacionales sugieren la necesidad de explorar enfoques alternativos o mejoras en la eficiencia del algoritmo para garantizar una aplicación práctica en situaciones de planificación distrital.

4.3. Centros por definir con límites regionales y variación del número de distritos

El propósito de este análisis es entender cómo la elección del número de centros, representado por K , influye en el balance poblacional, representada por ε_{BP} . Específicamente, estamos interesados en observar cómo cambia ε_{BP} a medida que ajustamos el valor de K . Para cada valor de K , se llevó a cabo el problema de factibilidad para determinar los centros óptimos y el valor de ε_{BP} más pequeño. Para encontrar este último valor se realiza la rutina de búsqueda binaria señalada en la Sección 3.4. Esta rutina se realiza fijando una tolerancia de $T = 0,005$.

4.3.1. Variación entre 16 y 56 distritos

En este caso, se desea llevar a cabo un análisis que abarque desde 16 distritos hasta un máximo de 56 distritos. La elección de estos valores se fundamenta en dos consideraciones. En primer lugar, se toma en cuenta el mínimo de 16 distritos, ya que esto coincide con el número de regiones en Chile y es el mínimo de distritos factible bajo la condición de respetar límites regionales. El establecer 16 distritos asegura que cada región se convierta en su propio distrito, respetando los límites regionales. Por otro lado, se establece un límite superior de 56 distritos, lo que representa el doble del número de distritos actualmente implementados. Esta amplia variación de distritos proporciona un rango adecuado para el análisis en el contexto chileno.

Para cada uno de los casos, se han obtenido resultados interesantes. Con el propósito de ofrecer una visualización más clara de estos resultados, se han representado en el gráfico de la Figura 4.10:

Es importante señalar que, aunque hay una tendencia decreciente general, no todos los distritos consecutivos tienen una reducción en el valor de ε_{BP} . Además, inicialmente el valor de ε_{BP} es elevado debido a la limitación que impone el escenario con 16 distritos, donde no se permite dividir ninguna región, por lo que necesariamente cada región debe constituir un distrito independiente. Esto conlleva a marcadas disparidades de población entre los distritos resultantes.

A medida que el valor de K se incrementa, se ajusta progresivamente la subdivisión de las poblaciones en los K distritos necesarios para minimizar las desviaciones respecto a la población promedio, lo que a su vez reduce el valor de ε . Sin embargo, en el caso particular de $K = 20$, se evidencia un aumento en el valor de ε en comparación con su antecesor. Este incremento podría estar relacionado con la población promedio para el caso y la limitación de no dividir comunas y respetar los límites regionales.

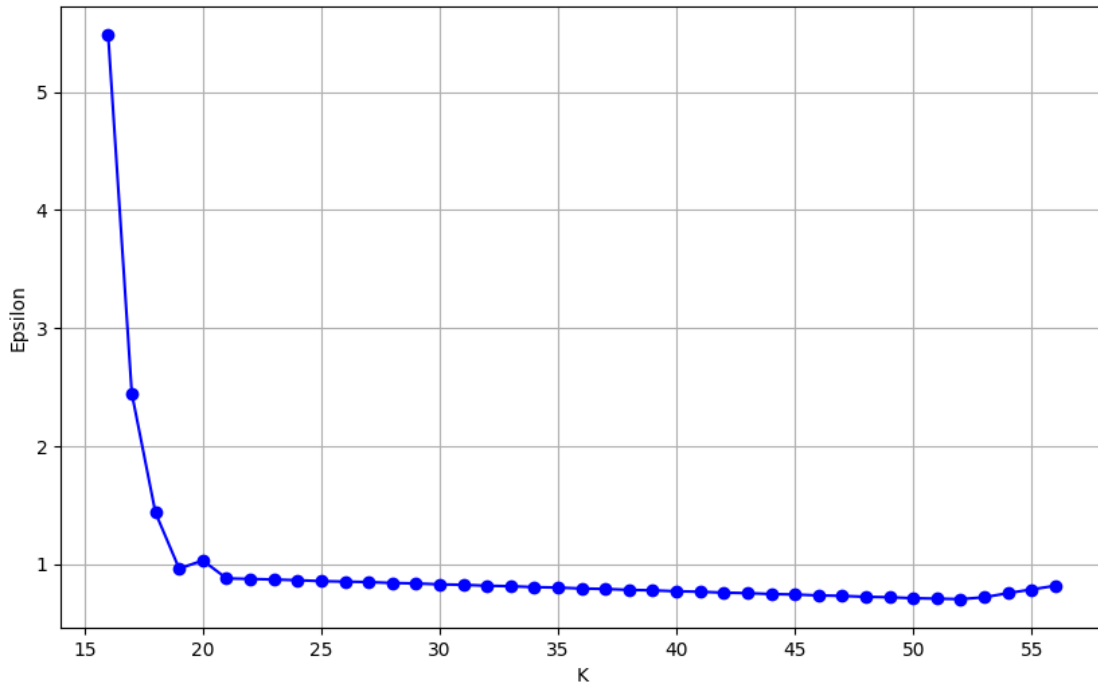


Figura 4.10: Variación de ϵ según el valor de K entre 16 y 56

4.3.2. Variación entre 16 y 200 distritos

A raíz de la situación anterior, se ha tomado la decisión de evaluar a un mayor número de distritos. En el caso anterior, se exploraron de manera secuencial las opciones que abarcaban desde 16 hasta 56 distritos. A partir de este punto, se procedió a realizar evaluaciones adicionales en incrementos de 10 distritos, abarcando casos desde 60, 70... 190, y finalmente, 200 distritos. Este enfoque escalonado es para observar cómo se comportan las soluciones en función del número de distritos. A continuación, se observa el valor de ϵ entre 16 y 200 distritos:

Como se mencionó antes, se observa una clara tendencia decreciente en los valores de ϵ_{BP} a medida que aumenta el número de distritos, hasta llegar a $K = 52$. Es particularmente significativo destacar que el punto en el que se alcanza el menor valor de ϵ_{BP} corresponde a $K = 52$ en este contexto específico. Lo anterior sugiere que, este resultado destaca la importancia de elegir la cantidad de distritos adecuada, mostrando que, en este caso, $K = 52$ representa la configuración óptima para minimizar el error de balance poblacional asociado a ϵ_{BP} .

A partir del análisis de la Figura 4.11, se destacan dos observaciones significativas. En primer lugar, al observar la variación del error de balance poblacional ϵ_{BP} , se identifica un intervalo inicial en el cual parece decrecer de manera constante. Este descenso puede estar

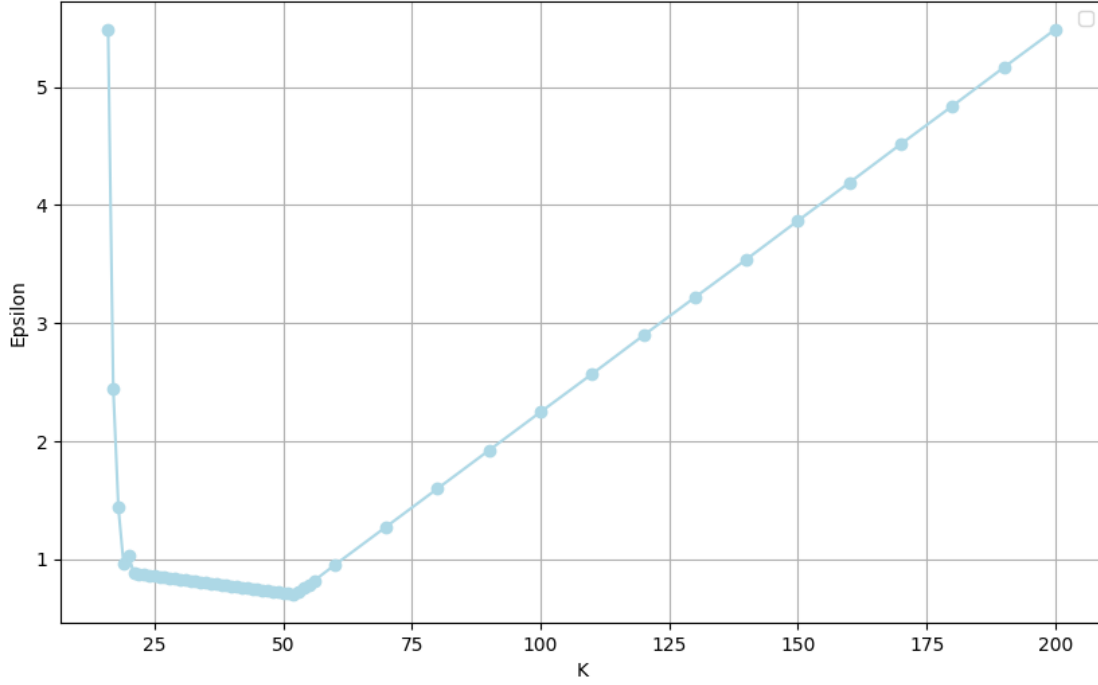


Figura 4.11: Variación de ε según el valor de K entre 16 y 200

delimitado por la región de menor población, sugiriendo una influencia directa de este factor en la variabilidad de ε_{BP} . En segundo lugar, a partir de $K = 60$, se observa una marcada tendencia creciente en los errores de balance poblacional ε_{BP} , aparentemente constante.

El comportamiento mencionado de decrecer y crecer de forma constante sugiere la presencia de restricciones fuertes que determinan en ε_{BP} . En particular, las restricciones (3.8) y (3.9) emergen como candidatas clave en la explicación de esta tendencia.

Ambas situaciones resaltan la importancia de estas restricciones en la determinación de ε_{BP} , y su comprensión es esencial para interpretar los patrones observados en la variación de este parámetro en el contexto del problema en cuestión. A continuación, se procederá al despeje de ε_{BP} a partir de las restricciones mencionadas.

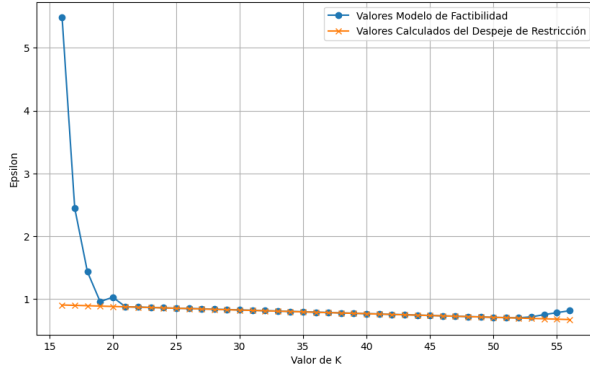
A partir de la siguiente desigualdad comienza el despeje para (3.8):

$$p_i \geq \frac{p_t}{K} \cdot (1 - \varepsilon_{BP}) \quad (4.2)$$

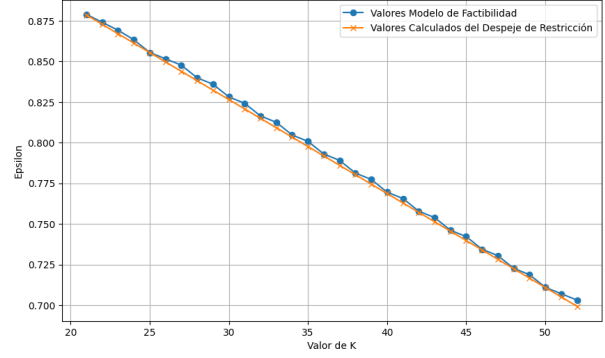
Obteniendo como resultado la siguiente desigualdad que puede estar definiendo el valor de ε_{BP} en el intervalo decreciente:

$$\varepsilon_{BP} \leq 1 - \frac{p_i \cdot K}{p_t} \quad (4.3)$$

4.3. Centros por definir con límites regionales y variación del número de distritos



(a) Comparación para K entre 16 y 56



(b) Comparación para K entre 21 y 52

Figura 4.12: Comparación del valor de ε según la ecuación despejada y el valor obtenido en el modelo para K entre 16 y 56

Y para (3.9) es a partir de la siguiente desigualdad:

$$p_i \leq \frac{p_t}{K} \cdot (1 + \varepsilon_{BP}) \quad (4.4)$$

Obteniendo como resultado la siguiente desigualdad que puede estar definiendo el valor de ε_{BP} en el intervalo creciente:

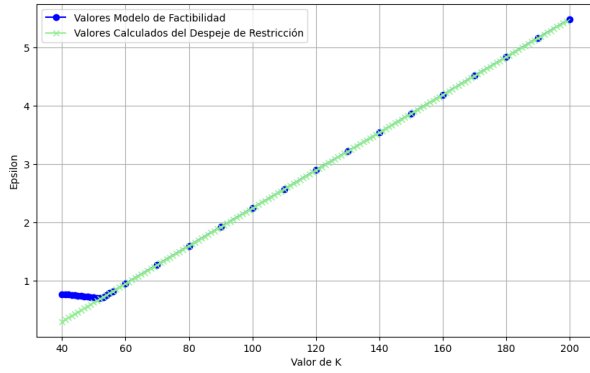
$$\varepsilon_{BP} \geq \frac{p_i \cdot K}{p_t} - 1 \quad (4.5)$$

En el caso específico de Chile, se destaca la notable disparidad poblacional entre las regiones, evidenciada por la Región de Aysén, que cuenta con la menor población, registrando un total de 101316 habitantes. Por otro lado, la comuna de Puente Alto se posiciona como la de mayor población con un total de 568106 habitantes. Estos extremos ilustran la diversidad demográfica presente en el país.

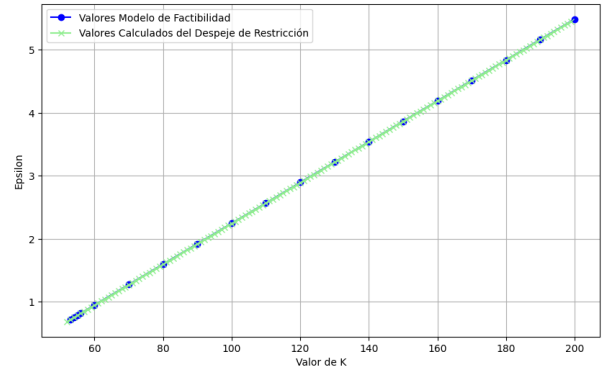
Para explorar la relación entre la población de la Región de Aysén y la población total de Chile en el contexto de la ecuación (4.3), se han utilizado los valores antes mencionados. La ecuación proporciona datos que se reflejan en el gráfico adjunto, el cual ofrece una representación visual de la influencia de la población de la Región de Aysén en el comportamiento de la variable de interés (ε_{BP}).

En este contexto, la simbología utilizada en la figura 4.12 se divide en dos conjuntos: los “Valores Modelo de Factibilidad” y los “Valores Calculados del Despeje de Restricción”. Los primeros surgen directamente del modelo de factibilidad, el cual incluye límites regionales y centros por definir. Por otro lado, los segundos resultan de sustituir las poblaciones específicas de la Región de Aysén y de Chile en la ecuación (4.3). En la Figura 4.12a se observa de forma más completa la situación mientras que en la Figura 4.12b se observa en mayor detalle el intervalo comprendido entre $K = 21$ y $K = 52$.

4.3. Centros por definir con límites regionales y variación del número de distritos



(a) Comparación para K entre 40 y 200



(b) Comparación para K entre 53 y 200

Figura 4.13: Comparación del valor de ε según la ecuación despejada y el valor obtenido en el modelo para K entre 40 y 200

A partir de lo anterior, se evidencia que los valores generados por la ecuación (4.3) se ajustan bastante a los obtenidos a través del modelo de factibilidad que considera límites regionales y centros por definir. Sin embargo, este ajuste es significativo exclusivamente dentro del intervalo específico de $K = 21$ a $K = 52$, como se ilustra claramente en la Figura 4.12. La razón fundamental detrás de este fenómeno es la activación de restricciones específicas en función del valor de K . Entre $K = 21$ y $K = 52$, la restricción (3.8), que impone un límite inferior a la población permitida, entra en acción debido a la Región de Aysén. Este hecho conduce a un ajuste notorio en los errores de balance poblacional ε_{BP} de acuerdo con la ecuación (4.3).

Es esencial destacar que tanto antes como después de este rango específico, la ecuación (4.3) deja de ajustarse de manera adecuada. Este fenómeno sugiere la existencia de otra restricción que entra en juego y ejerce influencia sobre el error de balance poblacional ε_{BP} , fuera del rango mencionado. La identificación y comprensión de esta restricción adicional son cruciales para una interpretación completa de los resultados obtenidos. Para continuar con el análisis, se explorará la relación entre la población de la comuna más poblada, Puente Alto, y la población total de Chile, en el contexto de la ecuación (4.5). Utilizando los valores poblacionales previamente mencionados, se generan datos que se presentan en el gráfico adjunto para comparar nuevamente los “Valores Modelo de Factibilidad” y los “Valores Calculados del Despeje de Restricción”.

A partir de la observación de la Figura 4.13, se evidencia un ajuste consistente entre los valores generados por la ecuación (4.5) y aquellos obtenidos a través del modelo de factibilidad, que considera límites regionales y centros por definir. Sin embargo, este ajuste es notablemente significativo únicamente en el intervalo creciente, abarcando desde $K = 52$ hasta $K = 200$, según se muestra de manera evidente en la Figura 4.13.

Esta conclusión permite inferir que, en el caso de valores crecientes de K , existe una

relación directa con la restricción (3.9), utilizando a Puente Alto como la comuna dominante en la situación. En situaciones hipotéticas donde la población de Puente Alto disminuiría, otra comuna con mayor población podría asumir el rol dominante, señalando esta transición como el cuello de botella del problema.

Es crucial destacar que muchos de los valores de K estudiados están estrechamente vinculados a las restricciones de población (3.8) y (3.9), las cuales definen en la mayoría de los casos el error de balance poblacional ε_{BP} . Sin embargo, se observa que existen valores que no pueden explicarse de manera precisa mediante ninguna ecuación. Este fenómeno resalta la complejidad y la interacción dinámica de las restricciones en el modelo, las cuales influyen en la variación de ε_{BP} en diferentes segmentos del rango de K .

4.4. Interpretación de los resultados

En esta sección se presenta la interpretación general de los resultados obtenidos previamente en el Capítulo 4. Dichos resultados se agrupan en diferentes subsecciones para facilitar su comprensión.

4.4.1. Efectos de la presencia de límites regionales

Un dato relevante a destacar es la diferencia entre los dos grandes escenarios propuestos. Cuando se eliminan los límites regionales, ya sea cuando los centros son fijos o cuando los centros se deben definir, se logra un error de balance poblacional ε_{BP} más bajo, evidenciando que los límites regionales conllevan ciertas restricciones que dificultan alcanzar la equidad en la representación.

La magnitud de estas mejoras en equidad de representación resalta la importancia de la restricción de límites regionales en la configuración de distritos. La eliminación de esta restricción podría ofrecer una representación más equitativa en términos de población, pero también plantea preguntas sobre si los representantes deben ser electos respetando esta restricción, ya que los límites regionales no son arbitrarios, sino que reflejan historias y problemáticas compartidas por las comunas de una misma región. Esto podría llevar a situaciones donde los representantes electos no estén completamente alineados con las particularidades y necesidades de todas las comunas que representan. Además, podrían surgir conflictos de intereses si diferentes comunas dentro de un mismo distrito tienen objetivos o necesidades opuestas debido a sus diferentes contextos regionales. La decisión de si los representantes deben ser electos respetando la restricción de límites regionales debería considerar tanto la equidad en representación como la pertinencia geográfica de los distritos.

Este análisis abre un debate importante sobre qué priorizar en la definición de distri-

tos: ¿Se debe buscar una representación lo más equitativa posible desde el punto de vista poblacional, aun si eso significa cruzar límites regionales? ¿O es más adecuado mantener la integridad de las regiones, reconociendo que esto puede resultar en ciertas desigualdades en la representación?

La reflexión final sería que, aunque la equidad en términos de representación poblacional es un objetivo esencial, también es importante considerar las historias y particularidades de las regiones. Encontrar un equilibrio entre estos dos aspectos es un desafío central en el diseño de cualquier sistema electoral.

4.4.2. Heurística para caso con límites regionales

Uno de los hallazgos más destacados del trabajo realizado es la sorprendente similitud en el error de balance poblacional ε_{BP} cuando hay límites regionales ya sea para el caso de centros fijos como en el caso de centros por definir. Esto sugiere que el problema de factibilidad con centros fijos y con límites regionales podría aproximarse de manera eficiente a través de una heurística que conduce a una solución factible y razonable. Lo que es aún más notable es la drástica diferencia en los tiempos de cómputo entre ambos enfoques. En particular, el tiempo requerido para el caso con centros fijos es aproximadamente diez veces menor que el tiempo necesario para el caso en el que los centros son definidos por el modelo.

Este hallazgo tiene implicaciones importantes, ya que demuestra la eficacia de la heurística aplicada al problema. Además, resalta la importancia de considerar el tiempo de cómputo al tomar decisiones en situaciones de distritaje, donde la eficiencia en la generación de distritos es un factor crítico. En resumen, hemos observado que, independientemente del método utilizado para definir los centros, el error de balance poblacional ε_{BP} se mantiene constante, lo que sugiere que bajo las condiciones señaladas y para el caso de Chile, la heurística conlleva a una solución óptima y factible, además de una diferencia significativa en la eficiencia computacional.

Además, cabe destacar que la heurística encontrada es un ajuste del distritaje actual del país, ya que en base a los distritos que ya están conformados se selecciona la comuna con más población y entorno a dicha comuna se generan los nuevos distritos con los cuales se obtiene un error de balance poblacional ε_{BP} óptimo.

4.4.3. Número de distritos dónde se alcanza el óptimo para la menor diferencia de población

Es relevante destacar el número de distritos para el cual el error de balance poblacional ε_{BP} alcanza su punto óptimo, lo que significa el valor más bajo de ε_{BP} que puede lograrse

bajo las condiciones especificadas, como respetar los límites regionales y permitir que el modelo defina los centros de los distritos, además de las restricciones que posee el modelo. En este caso, ese número es 52, lo que representa aproximadamente el doble de la composición actual de 28 distritos. El error de balance poblacional ε_{BP} óptimo en esta configuración es de 0.7031.

Comparando esto con el caso de 28 distritos, donde el mejor error de balance poblacional ε_{BP} que se puede alcanzar es 0.8398, notamos una diferencia de 0.1367. Aunque esta diferencia no parece ser significativa a primera vista, es importante considerar que en términos de representatividad o valor de voto, podría tener un impacto significativo. Por lo tanto, la elección del número de distritos y el ε_{BP} son decisiones cruciales en el proceso de creación de distritos, ya que afectan directamente la representación y el equilibrio de poder en el sistema político.

4.4.4. Comparando balance poblacional actual de Chile con los obtenidos en los modelos

Al inicio del Capítulo 4, se menciona que actualmente Chile tiene un error de balance poblacional ε_{BPA} cercano a 1,3299 en su actual sistema de distritaje. Bajo las mismas condiciones que rigen el distritaje chileno actual, es decir, respetando los límites regionales y con un total de 28 distritos, el óptimo del error de balance poblacional ε_{BP} se reduce significativamente a aproximadamente 0.8398. Esto indica que existe margen para mejorar la representatividad en el sistema actual.

Sin embargo, es importante destacar que se pueden lograr aún mayores niveles de optimización al cambiar las condiciones del distritaje. Por ejemplo, al eliminar por completo los límites regionales, el error de balance poblacional ε_{BP} disminuye drásticamente a aproximadamente 0.039. Esto sugiere que la eliminación de los límites regionales permitiría una representación mucho más equitativa y justa.

Por otro lado, al crear 52 distritos y aún manteniendo los límites regionales, se obtiene un ε_{BP} de aproximadamente 0.7031. Esto muestra que es posible lograr una mejora sustancial en la representatividad incluso sin eliminar los límites regionales, al simplemente aumentar el número de distritos.

Estos resultados demuestran que existen múltiples enfoques y oportunidades para mejorar la representatividad en el sistema de distritaje chileno, y que la elección de las condiciones y el número de distritos juega un papel crucial en este proceso.

4.4.5. Relevancia de restricciones de población en el modelo

La relevancia de las restricciones de población en el modelo se destaca a través de las observaciones realizadas en la Subsección 4.3.2. Los errores de balance poblacional ε_{BP} presentan una asociación notable con las restricciones (3.8) y (3.9). Estas restricciones imponen límites significativos, siendo la Región de Aysén el límite inferior y la comuna de Puente Alto el límite superior que dominan en la mayoría de los errores de balance poblacional ε_{BP} para los distintos valores de K .

El comportamiento de ε_{BP} dentro del rango estudiado refleja la influencia directa de estas restricciones de población, evidenciando un patrón claro de variación asociado a la demografía de la Región de Aysén y la comuna de Puente Alto. La Figura 4.13 proporciona una representación visual la relación con Puente Alto, donde se puede apreciar cómo la ecuación (4.5) se ajusta de manera coherente en el intervalo creciente de $K = 52$ a $K = 200$. Mientras que, la Figura 4.12 representa el caso para la Región de Aysén y cómo la ecuación (4.3) se ajusta en el intervalo decreciente de $K = 21$ a $K = 52$.

Es crucial señalar que, aunque la mayoría de los valores de K pueden explicarse directamente mediante estas restricciones, existen instancias, como se observa entre $K = 16$ y $K = 20$, donde la asociación no es clara. Específicamente, el valor obtenido para $K = 20$ plantea desafíos en la interpretación, sugiriendo la posible influencia de factores adicionales no capturados por las restricciones poblacionales predominantes. Esta complejidad resalta la necesidad de una comprensión más profunda de las dinámicas demográficas y de las restricciones en juego.

Capítulo 5

Conclusión

En el Capítulo 4 de simulaciones y resultados, se llevaron a cabo extensos análisis y se obtuvieron conclusiones importantes en relación con el diseño de distritos en Chile. Aquí resumimos las conclusiones más destacadas:

1. Impacto de los Límites Regionales: Los límites regionales tiene un impacto significativo en la equidad de representación. Cuando se eliminan los límites regionales, se logra una representación más equitativa en términos de población, pero esto también plantea preguntas sobre la pertinencia geográfica de los distritos.
2. Heurística Eficaz: Se descubrió que, para el caso de Chile y bajo las condiciones mencionadas, cuando se aplican límites regionales, la heurística utilizada en el caso de centros fijos, definidos por la comuna con más población de cada uno de los distritos actuales del país, conduce a una solución óptima, factible y con eficiencia computacional.
3. Óptimo en el Número de Distritos: Se determinó que el número óptimo de distritos para mantener los límites regionales y minimizar la diferencia de población es 52. Esto sugiere que hay un equilibrio crítico entre la equidad en la representación y la cantidad de distritos.
4. Asociación de los errores de balance poblacional ε_{BP} : Se ha observado que las restricciones de población están dominadas o influenciadas por la comuna de Puente Alto y la Región de Aysén, pues es en base a ellas que se determinan y explican la mayoría de los errores ε_{BP} .

5.1. Trabajo Futuro

Este estudio proporciona una base sólida para futuras investigaciones y abre una serie de interrogantes para investigarse. Algunos posibles temas de trabajo futuro incluyen:

1. Efectos Políticos: Sería interesante investigar cómo los cambios en el diseño de distritos podrían afectar la representación política y la toma de decisiones a nivel nacional. A partir de los distritos propuestos, investigar como quedaría la asignación de escaños o cuántos escaños son necesarios para una representación justa.
2. Caso Fácil para Chile: A lo largo de la investigación surge una interrogante respecto a qué particularidad del grafo de Chile es la que permite que resolver los modelos sea fácil, es decir, si existe alguna característica que permite la eficiencia computacional. Pese a que el tratamiento y el preprocesamiento de los datos es parte fundamental, se ha demostrado que el problema es NP-hard en grafos generales [6], incluso en el caso de dos centros. Entonces, entender qué tipo de grafo es Chile podría permitir reducir dicha complejidad en otros contextos relacionados.

Finalmente, el diseño de distritos en Chile es un tema complejo que involucra una serie de consideraciones importantes. Este estudio proporciona una visión general de las implicaciones de diferentes enfoques y abre la puerta a investigaciones adicionales para mejorar la representación política en el país.

Bibliografía

- [1] Conformación de nuevos distritos y circunscripciones, reforma al sistema binominal. <https://www.bcn.cl/siit/actualidad-territorial/conformacion-de-nuevos-distritos-y-circunscripciones/>, 2015.
- [2] Ley 20840. <https://www.bcn.cl/leychile/navegar?idNorma=1077039>, 2015.
- [3] Comunas de chile. <https://github.com/altazor-1967/Comunas-de-Chile/blob/master/Latitud-Longitud%20Chile.csv>, 2020.
- [4] Networkx: Network analysis in python. <https://networkx.org/documentation/>, 2023.
- [5] Tara Abrishami, Nestor Guillen, Parker Rule, Zachary Schutzman, Justin Solomon, Thomas Weighill, and Si Wu. Geometry of graph partitions via optimal transport. *SIAM Journal on Scientific Computing*, 42(5):A3340–A3366, 2020.
- [6] Nicola Apollonio, Isabella Lari, Federica Ricca, Bruno Simeone, and Justo Puerto. Polynomial algorithms for partitioning a tree into single-center subtrees to minimize flat service costs. *Networks*, 51:78–89, 01 2008.
- [7] Allan Borodin, Omer Lev, Nisarg Shah, and Tyrone Strangway. Big city vs. the great outdoors: Voter distribution and how it affects gerrymandering. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI-18*, pages 98–104. International Joint Conferences on Artificial Intelligence Organization, 7 2018.
- [8] Jowei Chen and David Cottrell. Evaluating partisan gains from congressional gerrymandering: Using computer simulations to estimate the effect of gerrymandering in the u.s. house. *Electoral Studies*, 44:329–340, 2016.
- [9] Vincent Cohen-Addad, Philip N. Klein, and Neal E. Young. Balanced power diagrams for redistricting. *CoRR*, abs/1710.03358, 2017.

- [10] Nikhil Garg, Wes Gurnee, David Rothschild, and David Shmoys. Combatting gerrymandering with social choice: The design of multi-member districts. pages 560–561, 07 2022.
- [11] Wes Gurnee and David B. Shmoys. Fairmandering: A column generation heuristic for fairness-optimized political districting. *CoRR*, abs/2103.11469, 2021.
- [12] Harry A. Levin and Sorelle A. Friedler. Automated congressional redistricting. *ACM J. Exp. Algorithmics*, 24, apr 2019.
- [13] Yoad Lewenberg, Omer Lev, and Jeffrey S. Rosenschein. Divide and conquer: Using geographic manipulation to win district-based elections. In *Adaptive Agents and Multi-Agent Systems*, 2017.
- [14] G. Van Brummelen. *Heavenly Mathematics: The Forgotten Art of Spherical Trigonometry*. Princeton University Press, 2012.

Apéndice A

Anexo

A.1. Diccionario de funciones utilizadas para el código

- `dist`: Recibe dos comunas y entrega la distancia en kilometros entre ellas.
- `pob`: Recibe una comuna y entrega la población de dicha comuna.
- `calcular_poblacion_total`: Recibe una lista de comunas y entrega la población total de todas ellas.
- `obtener_region`: Recibe una comuna y entrega la región a la cuál pertenece.
- `obtener_comunas`: Recibe una región y entrega las comunas que pertenecen a dicha región.
- `codigo_com_cut`: Recibe una comuna y entrega el CUT de ella.
- `k_shortest_path`: Recibe un grafo G , un nodo fuente, un nodo destino, y un número k , y devuelve los k caminos más cortos desde el nodo fuente al nodo destino en el grafo G .
- `Snuevo`: Recibe V , u y v . V es un conjunto de nodos para crear un subgrafo inducido; se busca el camino más corto entre u y v en ese subgrafo, y se retorna el penúltimo nodo de ese camino si su longitud es al menos 3; de lo contrario, retorna una lista vacía.

A.2. Pseudocódigos

En la presente sección se presentan algunos pseudocódigos con el fin de mejorar el entendimiento de la implementación de algunas de las funciones mencionadas en el código.

Estos pseudocódigos proporcionan una descripción de cómo se crea el grafo de adyacencia y cómo se encuentran los caminos más cortos entre comunas en el contexto del problema. Estos pseudocódigos son útiles para comprender mejor la implementación de estas funciones.

A.2.1. Grafo de adyacencia y matriz de adyacencia

A continuación, se observa de qué manera se implementa la creación del grafo de adyacencia utilizando la lista de adyacencias como insumo. Finalmente, una vez que el grafo de adyacencia está construido, se procede al cálculo de la matriz de adyacencia.

Algorithm 2 Creación del Grafo de Adyacencia

Input: *diccionario_de_adyacencia_por_comunas* ▷ Diccionario con clave el nombre de cada comuna y como llave sus adyacencias.

Output: *grafo_adyacencia*

Output: *matriz_adyacencia*

- 1: Crear una lista para almacenar los identificadores de las comunas
 - 2: Inicializar *id_local* como una lista vacía
 - 3: Inicializar un grafo de adyacencia como un grafo vacío
 - 4: **for** cada comuna, adyacencia en *diccionario_de_adyacencias_por_comuna* **do**
 - 5: Agregar el identificador de la comuna a *id_local*
 - 6: **for** cada comuna_ady en adyacencia **do**
 - 7: Agregar una arista entre comuna y comuna_ady en el grafo de adyacencia
 - 8: **end for**
 - 9: **end for**
 - 10: Calcular la matriz de adyacencia a partir del grafo de adyacencia
 - 11: Convertir la matriz de adyacencia a una matriz NumPy y redondear los valores
-

A.2.2. Cálculo caminos más cortos

Al buscar los caminos más cortos entre dos comunas, se emplea la matriz de adyacencia y se utilizan dos funciones clave: *k_shortest_paths* y *Snuevo*. La función *k_shortest_paths* se encarga de encontrar los *k* caminos más cortos entre dos comunas, considerando un peso asociado a las aristas del grafo. Por otro lado, la función *Snuevo* se utiliza para obtener caminos específicos entre dos comunas de interés. Estas funciones desempeñan un papel fundamental en el contexto del problema tratado.

A partir de lo mencionado en Anexos [A.2.1](#) y en la sección actual, se lleva a cabo el cálculo de los caminos más cortos entre cada par de comunas. Este proceso es fundamental para garantizar la contigüidad en la creación de distritos.

Algorithm 3 Cálculo de caminos más cortos(*k_shortest_paths*)

Input: G $\triangleright$ grafo inducido sobre el cual se calcularán los caminos más cortos

Input: *comuna_origen* y *comuna_destino*

Input: k $\triangleright$ cantidad de caminos más cortos que quiero que la función me devuelva

Input: *peso* $\triangleright$ es el peso que se dará a cada arista

Output: *resultado* $\triangleright$ lista con los nodos en el camino más corto

- 1: Crear una lista resultados
 - 2: **for** cada camino más corto en el grafo G desde *comuna_origen* hasta *comuna_destino* con peso **do**
 - 3: Agregar el camino a resultados
 - 4: **end for**
 - 5: **return** Devolver resultados
-

Algorithm 4 Cálculo de caminos entre dos comunas de interés (Snuevo)

Input: V $\triangleright$ Lista de nodos para crear un grafo

Input: u y v $\triangleright$ nodo origen y nodo destino, respectivamente

Output: *nodo_intermedio* $\triangleright$ Devuelve el nodo antecesor (respecto al nodo destino) entre el camino del nodo origen y el nodo destino

Output: *lista_vacia*

- 1: Crear un subgrafo H como subgrafo inducido de grafo_adyacencia utilizando los nodos en V
 - 2: Encontrar el camino más corto entre u y v en el subgrafo H .
 - 3: **if** el camino tiene al menos 3 nodos **then**
 - 4: **return** Devolver el segundo nodo del camino (nodo intermedio o antecesor)
 - 5: **else**
 - 6: **return** Devolver una lista vacía
 - 7: **end if**
-