Universidad de Valparaíso
CHILE

Facultad de Ciencias
Instituto de Estadística
Ingeniería en Estadística

ANÁLISIS ESPACIAL Y TEMPORAL DEL COVID-19 EN CHILE

Valentina Alejandra Escobar Rivera
7 de enero de 2022

Profesor Guía
Daira Velandia. Dra.
Instituto de Estadística, Universidad de Valparaíso

Profesor Co-Guía
Alejandra Christen. Dra.
Instituto de Estadística, Universidad de Valparaíso

Proyecto de titulación para optar al:

grado académico de: *Licenciado en Estadística*
título profesional de: *Ingeniero en Estadística*
minor en: *Estadística financiera*

Resumen

La pandemia llamada COVID-19 causada por el SARS-CoV-2 sigue amenazando la salud pública en todo el mundo, es por esto que comprender la dinámica espacial y temporal a nivel regional y comunal es muy importante en pro de comprender la propagación de la pandemia y de evaluar las medidas implementadas en el plan Paso a Paso. En el presente proyecto se explora la dinámica espacial y temporal del COVID-19 en Valparaíso, utilizando la variable casos diarios confirmados. Se evalua la asociación espacial a través del índice de Morán (local y global), entre Junio 2020 y Junio 2021, señalando la presencia de un patrón espacial y temporal, dado que el índice de Morán en todas las semanas epidemiológicas consideradas, indicó un p-valor cercano a cero, además fue lo suficientemente alto para indicar dependencia espacial positiva.

Algunas Palabras

Quiero dar agradecimiento primeramente a mis padres José Escobar y Gladys Rivera, quienes fueron mis guías de mi vida. A mi padre gracias por todo el amor que me brindó durante todo el período de enseñanza superior, por los consejos, regaños y apoyo incondicional, por el aliento para seguir siempre adelante y no rendirme en mis peores momentos, y a mi madre gracias por transmitirme valores, por forjarme a la mujer que soy en la actualidad, un beso al cielo mamá. Muchos de mis logros en el que se incluye este se los debo a ustedes. Muchas gracias.

A mis hermanos por estar siempre pendiente de mi ciclo en la universidad. Especialmente a mi hermana Carolina Escobar quien fue el aliento para seguir adelante luego del fallecimiento de mi madre, quien tomó las riendas del puesto de lugar de madre. Gracias por querer siempre lo mejor para mí. Por ser una de las mujeres más esforzadas que conozco te mereces mucho y más. Gracias.

A mi familia Escobar. gracias a mi querida familia, por todo su cariño, empatía, solidaridad y todos esos momentos maravillosos y chistosos que hemos compartido, a todos ellos gracias.

A mi gata Reina, quien fue la compañía en las largas noches de estudio y términos de informes e investigaciones. Quien se mantenía al costado de mi horas y horas frente a la pantalla del computador por tan solo estar conmigo.

A mis profesores de la Universidad de Valparaíso, gracias por transmitirme sus conocimientos y gran sabiduría, gracias a ello he logrado muchos objetivos a lo largo de la carrera.

A mi tutora la doctora Daira Velandia primeramente, por la darme la oportunidad de aceptarme como su estudiante, por enseñarme sus conocimientos y guiarme en mi proyecto. Darle las gracias por su paciencia, tolerancia y gran docente que es. A mi cotutora la doctora Alejandra Christen, por estar presente en mi proyecto, por aconsejarme a lo largo de todo este período, aceptando sus críticas constructivas con el fin de siempre querer un mejor para mi proyecto. Gracias a ustedes profesoras no podría haber finalizado mi investigación.

"*La riqueza de un hombre no se encuentra en la cantidad de dinero que posee, sino en la calidad de su conocimiento y educación*", Javier Herrera.

Índice general

Resumen **2**

Algunas Palabras **3**

1. Preliminares **5**
1.1. Introducción . . . 5
1.2. Objetivos . . . 7
1.2.1. Objetivo general . . . 7
1.2.2. Objetivos específicos . . . 7
1.3. Pregunta e hipótesis de investigación . . . 7

2. Fundamentación teórica **8**
2.1. Conceptos epidemiológicos . . . 8
2.2. Epidemiología . . . 9
2.2.1. Aplicaciones de la epidemiología . . . 10
2.3. Epidemiología espacial . . . 10
2.3.1. Sistemas de información geográfica . . . 13
2.4. *Cluster* . . . 14
2.4.1. Agrupación jerárquica . . . 15
2.4.2. *Single Linkage* (Vinculación simple) . . . 15
2.4.3. *Average Linkage* (Vinculación promedio) . . . 16
2.4.4. *Complete Linkage* (Vinculación máxima) . . . 16
2.4.5. Medidas de similitud . . . 16
2.5. *Cluster* espacial . . . 18
2.5.1. Índice de Morán . . . 19

3. Resultados **23**
3.1. Base de datos . . . 23
3.2. Análisis descriptivo . . . 25
3.2.1. Análisis descriptivo clásico . . . 25
3.2.2. Análisis espacial . . . 47

4. Conclusión **66**

A. Algoritmos computacionales **67**

Referencias **71**

Capítulo 1

Preliminares

1.1. Introducción

El síndrome agudo respiratorio severo (SARS) es una enfermedad de reciente aparición que se propaga en brotes, se originó en la provincia de Guangdong en China, y los primeros casos se reconocieron a mediados de Noviembre de 2002 (Wong y Hui, 2003). El SARS es causado por el Coronavirus asociado con el SARS-CoV y pertenece a la familia de Coronavirus (es decir la misma familia que puede causar un resfriado común). La familia de los Coronavirus se encuentra muy extendida en el reino animal, y es una de las causas más frecuentes de resfriado en humanos, además de producir infecciones pulmonares en contextos epidemiológicos (Hernández, 2003). Tres de los 7 coronavirus existentes causan infecciones respiratorias en los seres humanos, mucho más graves e incluso a veces mortales, que los demás coronavirus, causando brotes importantes de neumonía mortal en el siglo XXI.

1. El MERS-CoV se identificó en 2012 como la causa del síndrome respiratorio de Oriente Medio (MERS) (Tenisi, 2021).

2. El SARS-CoV fue identificado en 2002 como la causa de un brote de síndrome respiratorio agudo grave (SARS) que comenzó en China hacia finales de 2002 (Tenisi, 2021).

3. SARS-CoV-2 es un nuevo coronavirus identificado como la causa de la enfermedad por coronavirus de 2019 (COVID-19) (Tenisi, 2021).

El COVID-19 es la enfermedad causada por el nuevo coronavirus conocido como SARS-CoV-2. La organizacion mundial de la salud (OMS) reportó por primera vez la existencia de este nuevo virus el 31 de diciembre del 2019, en la localidad de Wuhan-China, expandiéndose ampliamente en todo el mundo (OMS, 2020). La enfermedad por coronavirus (COVID-19) representa una importante amenaza pandémica que se ha extendido a mas de 212 países con mas 7.898.442 casos confirmados en todo el mundo (OMS, 2020), esto ha generando un creciente interés por la comprensión de esta enfermedad desde diversos enfoques. Hasta la fecha, existen seis vacunas efectivas que detienen de cierta forma la propagación de esta pandemia viral emergente. Actualmente, las estrictas medidas de intervención, implementadas en varios países, de detección, control y prevención son las únicas estrategias efectivas para minimizar la propagación del COVID-19. El desafío de controlar la rápida propagación del COVID-19 es su complejo comportamiento espacial y temporal, debido a los rápidos cambios en la dinámica de la población y algunas características demográficas y ambientales, al respecto en algunos países se han realizados estudios de este tipo, que han ayudado un poco a

la comprensión e implementación de nuevas medidas de control.

(Cuartas y cols., 2020) describen la distribución espacio-temporal del COVID-19 en la ciudad de Cali- Colombia, durante el primer mes de epidemia, realizando un análisis exploratorio de datos espaciales y un análisis de densidad de Kernel. Encontraron que la distribución espacial de los casos se concentró inicialmente en en el norte y el sur de la ciudad, con una dinámica cambiante hacia el oriente y occidente. Por otro lado, (Alkhamis y cols., 2020) exploraron la dinámica temporal y espacial del COVID-19 en el estado de Kuwait. Específicamente, cuantificaron los patrones temporales y espacio-temporales de la propagación de la pandemia a nivel nacional y dentro de las comunidades locales. La progresión pandémica del COVID-19 se cuantificó utilizando el Número Reproductivo dependiente del tiempo efectivo (R_t). Se detectaron eventos de agrupación espacio-temporal prospectivos y retrospectivos. Mientras que en Chile, (Vergara-Perucich, 2020) identifican la correlacion espacial entre los determinantes sociales de la salud en la vivienda y la propagación del COVID-19 en Santiago de Chile, realizando un doble exámen/análisis de la correlación espacial utilizando la medida de Pearson para observar cómo el virus se propagó desde zonas con viviendas de alta calidad en la fase inicial del contagio para luego concentrarse en zonas con viviendas de baja calidad. El segundo examen/análisis realizado es una regresión lineal múltiple, identificando los factores de alojamiento que informan sobre la propagación del virus. Las conclusiones sugieren que las políticas públicas deben tratar la calidad de la vivienda como un factor de salud pública y los riesgos para la salud deben abordarse con un enfoque transdisciplinario de la planificación urbana en Chile.

En Chile, el nuevo coronavirus COVID-19 no solo ha impactado en la mortalidad de la población chilena , sino también en el rubro del trabajo, en la economía y en la salud. Los efectos de la pandemia generaron consecuencias sin precedentes en el rubro laboral, según la organización internacional del trabajo (OIT), además provocó una crisis de gran envergadura en la actividad económica, en el empleo y en los ingresos (OIT, 2020). Mientras que en la salud pública se presentó un colapso de intervenciones con un 97 % de las camas UCI (Unidad de Cuidados Intensivos) ocupadas con fecha 21 de Abril del 2020 (El comercio, 2020). Dado el impacto en los sectores económico, productivo, social y de salud, se considera de gran envergadura establecer cómo ha sido la dinámica espacial y temporal del COVID-19 en Chile para entender su dinámica de transmisión y la implementación de las medidas de intervención.

En el presente trabajo se propone explorar la dinámica espacial y temporal del COVID-19 a nivel de comunal y regional en Chile. En el presente capítulo se plantean los objetivos, pregunta e hipótesis de investigación. Luego se describe la fundamentación teórica para el desarrollo de este trabajo y por último se presenta el capítulo de los resultados obtenidos junto con las respectivas conclusiones.

1.2. Objetivos

1.2.1. Objetivo general

Explorar la dinámica espacial y temporal del COVID-19 en Chile.

1.2.2. Objetivos específicos

- Estudiar la dinámica temporal de la pandemia causada por el COVID-19 en Chile.
- Estudiar la dinámica espacial de la pandemia causada por el COVID-19 en Chile.
- Identificar patrones espacial y temporal de eventos de agrupamiento para evaluar la efectividad de las medidas de intervención (Plan Paso a Paso) (Gobierno de Chile , 2020).

1.3. Pregunta e hipótesis de investigación

Se propone estudiar la dinámica espacial y temporal del COVID-19 en Chile, por lo tanto la pregunta de investigación propuesta es:

¿Es posible análizar la dinámica espacial y temporal del COVID-19 en Chile?.

Y la hipótesis de la investigación:

El SARS-CoV-2 en Chile presenta patrones de comportamiento de acuerdo al tiempo y a la ubicación geográfica de las personas.

Capítulo 2

Fundamentación teórica

En este capítulo se presentan las principales técnicas utilizadas en el análisis. Para empezar se describen algunos conceptos epidemiológicos que son necesarios para entender el desarrollo del trabajo.

2.1. Conceptos epidemiológicos

A continuación algunos conceptos epidemiológicos de interés:

- **SARS-CoV-2**: Coronavirus responsable de la enfermedad COVID-19 (ISC III , 2020).
- **Enfermedad infecciosa**: Enfermedades causadas por microorganismos patógenos como las bacterias, los virus, los parásitos o los hongos (ISC III , 2020).
- **Caso** : Enfermedad confirmada en un individuo. Durante el estudio de un brote epidémico o de una epidemia, las definiciones de caso se van adaptando al conocimiento que se va adquiriendo sobre la epidemiología de la enfermedad. En este caso de la epidemia causada por el COVID-19 (ISC III , 2020).
- **Caso confirmado**: caso que cumple criterio de laboratorio con PCR positivo (ISC III , 2020).
- **Caso probable**: caso cuyos resultados de laboratorio para SARS-CoV-2 no son concluyentes (ISC III , 2020).
- **Caso descartado**: caso cuyos resultados de laboratorio para SARS-CoV-2 son negativos (ISC III , 2020).
- **Caso posible**: caso con infección respiratoria aguda leve sin criterio para realizar test diagnóstico. La detección de infección se realiza básicamente en personas hospitalizadas o que cumplen ingresos de hospitalización, y en personal de servicios esenciales, sobre todo sanitarios. Las personas con síntomas que podrían ser compatibles con COVID-19, pero que, por el motivo que sea, no tienen acceso a las pruebas de PCR, no se incluyen dentro de las cifras que se facilitan diariamente sobre la enfermedad (ISC III , 2020).
- **Incidencia**: Casos recién diagnosticados de una enfermedad que se presentan o se registran en una comunidad determinada a lo largo de un período de tiempo específico (ISC III , 2020).

- **Tasa de incidencia**: Representa la velocidad de aparición de casos nuevos en la población en riesgo (ISC III , 2020). Para el caso de la tasa de incidencia en Chile, se considera cada 14 días por cada 100.000 habitantes según informe epidemiológico del Ministerio de Salud del gobierno de Chile . Su expresión está dada por:

$$TI = \frac{\alpha}{\beta} * 100000, \tag{2.1}$$

donde α es el número de casos nuevos y β corresponde a la población en riesgo.

- Plan "Paso a Paso nos cuidamos" : Corresponde a las medidas de intervención implementadas por el gobierno de Chile para evitar el alto contagio del COVID-19 entre personas. Se presentó e implementó por primera vez en julio de 2020 como una estrategia gradual para enfrentar la pandemia según la situación sanitaria de cada zona de nuestro país (Gobierno de Chile , 2020). El plan presenta 4 medidas de intervención aplicadas como fases, estas son:

  - Fase 1: Restricción, corresponde a la medida más restrictiva y busca disminuir el desplazamiento dentro y entre las comunas. Dejando la opción de salir solo en caso de necesitar de algún servicio básico con permiso de Carabineros.
  - Fase 2: Transición, se puede salir durante el día sin necesidad de permiso de Carabineros, excepto los fines de semana y festivos.
  - Fase 3: Preparación, término de la cuarentena los fines de semana.
  - Fase 4: Apertura inicial, contempla más libertades, aumentando los aforos en las instancias y establecimientos permitidos.

2.2. Epidemiología

La epidemiología es la disciplina científica que permite describir la frecuencia, distribución y tipos de enfermedades, en distintas epocas y lugares, además de explicar las relaciones existentes entre ellas y las condiciones externas que rodean al individuo y determinan su modo de vida. Su intención final ha sido contribuir a mejorar la capacidad de los servicios de salud para transformar las condiciones de vida y la situación de salud (Ayala, 2014).

La epidemiología se define también como un cuerpo cambiante de conocimientos, una metodología y una forma de pensar. Como disciplina científica dispone de un conjunto de sistemas de hipótesis y conocimientos (el método epidemiológico y el enfoque epidemiológico), y sirve como ciencia básica para la medicina y la salud pública. A pesar de la variedad de definiciones, todas tienen en común el continuo salud-enfermedad, que es su objeto de estudio. La epidemiología se basa en el análisis estadístico, configurando una metodología propia y específica, y emplea e integra algunos procedimientos de las ciencias sociales y de las ciencias básicas. Las amplias posibilidades de aplicación de los procedimientos de análisis epidemiológico puede concretarse en distintos ámbitos: el primero, el estudio de las enfermedades y de las actividades que se desarrollan para su control, de modo que la salud pública y la asistencia sanitaria disponen, a través de la epidemiología, de herramientas para responder adecuadamente a los problemas a los que se enfrentan; el segundo es el de racionalizar las políticas de salud (ámbito de la planificación sanitaria) (Fernández, 2014).

2.2.1. Aplicaciones de la epidemiología

La epidemiología tiene como objetivo proporcionar respuestas válidas y oportunas a cuestiones sobre la comprensión y la mejora de los problemas de salud de la población. Las preguntas que pueden responderse desde, o a través de, la epidemiología se refieren a la importancia de los problemas de salud y a la posibilidad de modificarlos: qué frecuencia y qué tendencia presenta una determinada enfermedad, qué características de utilización y de accesibilidad tienen las prestaciones sanitarias, entre otras. Las principales aportaciones de la epidemiología son su método y su enfoque. El método epidemiológico se basa en el método científico, es decir, se formulan hipótesis y se intenta probar mediante una metodología científica observacional, por otro lado el enfoque epidemiológico se centra en el estudio de grupos humanos (poblaciones, comunidades, etc) definidos por un entorno geográfico o temporal determinado (Fernández, 2014) .

2.3. Epidemiología espacial

El objetivo de la epidemiología espacial consiste en el estudio de los aspectos geográficos o la distribución espacial de los resultados de salud. Por lo general, su foco principal es la incidencia o la prevalencia de la enfermedad. La descripción parsimoniosa de la distribución es a menudo una preocupación importante, está puede implicar la visualización inicial de la distribución y algunas medidas de resumen simples, igualmente puede conducir a un modelado bastante complejo del proceso de la enfermedad. En el caso de una enfermedad infecciosa, la dinámica de transmisión también puede ser de interés (Lawson, Banerjee, Haining, y Ugarte, 2016). El espacio y el tiempo interactuan a menudo en los estudios de salud, por lo general, con una ubicación geográfica, es decir que el estudio de los datos estudiados estén georreferenciados, en estos casos los datos presenta la ubicación espacial/geográfica asociada a ellos (Lawson, 2006). La variación espacio-temporal en la incidencia de enfermedades puede contener importantes pistas etiológicas (estudios sobre las causas) y proporcionan información importante que ayuda a una mejor toma de decisiones en la salud pública (Lawson y cols., 2016).

La naturaleza multifactorial de las enfermedades y la diversidad de las agresiones, que incluyen factores biológicos, físicos, sociales, y factores culturales, combinados con la susceptibilidad genética, sugieren la necesidad de incorporar perspectivas integrales de la investigación epidémiologica multidisciplinaria. Además, existe la necesidad de utilizar herramientas, como los sistemas de información geográfica (GIS) y otros métodos geoespaciales, que pueden integrar factores de multinivel, espaciales y temporales, surgiendo la epidemiología espacial (Lawson y cols., 2016).

(Lawson, 2006) estudia la distribución espacial del cáncer respiratorio, analiza un conjunto de datos con 49 direcciones de certificados de defunción por cáncer respiratorio para el período de 1968-1974 en la pequeña ciudad de Armadale, en el centro de Escocia, situada en una zona industrial cercana a Falkirk (ver figura 2.1). Se registró una tasa de mortalidad estandarizada de 150 para cada uno de los años del período de estudio, este inusual exceso de muertes se denominó "Epidemia de Armadal". Las localizaciones de los casos van acompañadas de la confirmación de la enfermedad coronaria (CHD). Las ubicaciones de los certificados de defunción se utilizaron como control de la aparición de la enfermedad, además una ventana de estudio circular fue implementada para minimizar el sesgo de muestreo direccional. Las realizaciones del caso y del control se muestran en las Figuras 2.2 y 2.3.

Figura 2.1: Localización geográfica de Armadale, Escocia .
Fuente: Statistical Methods in Spatial Epidemiology, (Lawson, 2006) .

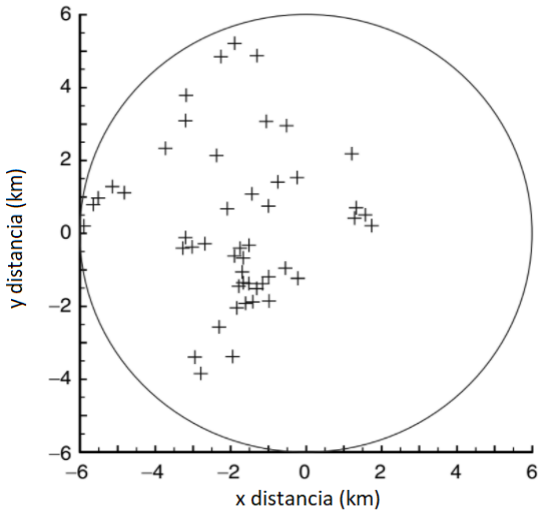

Figura 2.2: Certificados de defunción por cáncer respiratorio dentro de una ventana circular.
Fuente: Statistical Methods in Spatial Epidemiology, (Lawson, 2006).

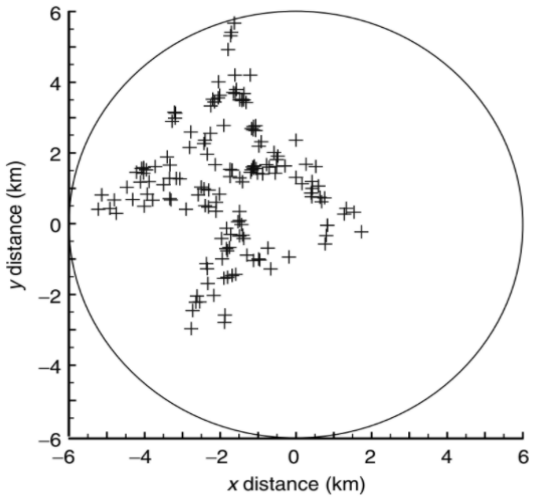

Figura 2.3: Realización de direcciones de certificados de defunción de CHD, dentro de la ventana circular.

Fuente: Statistical Methods in Spatial Epidemiology, (Lawson, 2006) .

Por otra parte, en (Lawson, 2006), analizan un conjunto de datos de leucemia y linfoma infantil, denominados"Datos sobre leucemia y linfoma en Humberside", en la región del norte de Humberside en Inglaterra para el período 1974-1986. Como "control" de la población "de riesgo" de la zona, se obtuvo una amplia muestra del registro de nacimientos para la región y el período. Las Figuras 2.4 y 2.5 muestran los mapas de eventos y de control. El propósito original de la investigación fue examinar la tendencia de la agrupación de los eventos del caso.

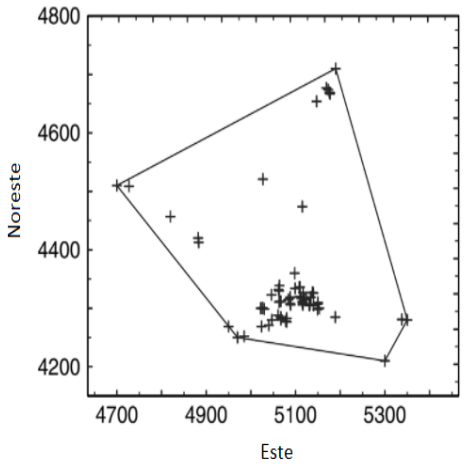

Figura 2.4: Humberside: mapa de casos de leucemia y linfoma (1974-1986).

Fuente: Statistical Methods in Spatial Epidemiology, (Lawson, 2006) .

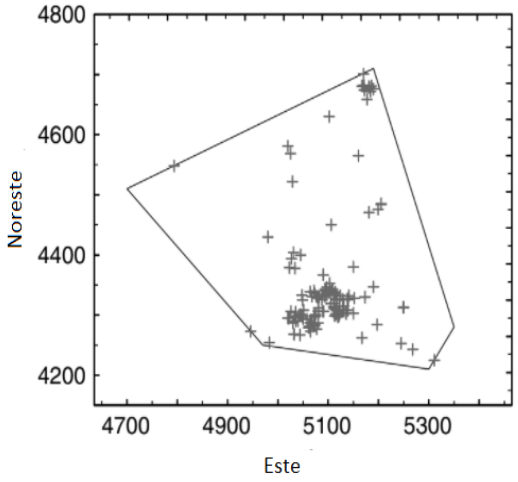

Figura 2.5: Humberside: mapa de eventos de control de leucemia y linfoma (1974-1986).
Fuente: Statistical Methods in Spatial Epidemiology, (Lawson, 2006) .

Entre las técnicas estadísticas ampliamente utilizadas para la comprensión de enfermedades de transmisión desde el enfoque espacial y espacio-temporal se encuentran los estudios de conglomerados *cluster*, el escaneo espacio-temporal, entre otros.

2.3.1. Sistemas de información geográfica

Un sistema de información geográfica (GIS) es un marco de trabajo para reunir, gestionar y análizar datos. Arraigado en la ciencia geográfica, GIS integra diversos tipos de datos. Análiza la ubicación espacial y organiza capas de información para su visualización. Prácticamente en todos los campos se utiliza un GIS para crear mapas que comunican, análizan, comparten información y resuelven problemas complejos alrededor del mundo (Aeroterra, 2021).

Un GIS puede contribuir a:

- Identificar problemas.
- Gestionar y dar respuesta a emergencias.
- Establecer prioridades.
- Monitorear cambios.
- Realizar predicciones.
- Comprender tendencias.

La tecnología de los GIS aplica la ciencia geográfica a través de herramientas para la comprensión y la colaboración, algunas de estas son:

- **Análisis.**
  El análisis espacial permite realizar evaluaciones en términos de adecuación y capacidad,

estimar y predecir, interpretar y comprender, y mucho más, ofreciendo nuevas perspectivas al conocimiento y toma de decisiones (Aeroterra, 2021).

- **Mapas.**
  Los mapas son el contenedor geográfico para las capas de datos y análisis con los que se quiera trabajar (Aeroterra, 2021).

- **Datos.**
  Los GIS integran diversos tipos de capas de datos que utilizan la ubicación espacial. La mayoría de los datos tienen un componente geográfico. Los datos GIS incluyen imágenes, atributos y mapas base (Aeroterra, 2021). La definición correspondiente de base de datos de los GIS es:
  "Cualquier conjunto de procedimientos informáticos utilizados para almacenar y manipular datos referenciados geográficamente".

El análisis espacial incluye el análisis cartográfico y topológico (por ejemplo, la medición de distancias y áreas, y las relaciones espaciales entre las observaciones, incluidas las operaciones de superposición y amortiguación), el modelado matemático (por ejemplo, el análisis de redes y superficies) y la estadística espacial.

En las siguientes secciones se introducen los métodos de asociación y agrupamiento utilizados en la estadística clásica y espacial.

2.4. *Cluster*

Existe gran variedad de enfoques para definir un *cluster*. Knox en 1964 expresa que "es un grupo geográficamente delimitado de ocurrencias de tamaño y concentración suficientes como para que sea improbable que se produzcan por casualidad" (Kanox, 1964). Una definición mas básica es: "cualquier área dentro de la región de estudio de riesgo elevado significativo", esta definición suele denominarse agrupación de puntos calientes. Esencialmente, cualquier zona de riesgo elevado, independientemente de su forma o extension, podría considerarse una agrupación, siempre que la zona cumpla algunos criterios estadísticos (Lawson, 2006).

Los estudios de conglomerados o cluster se utilizan habitualmente en tres ámbitos:

1. Como parte de una respuesta de salud pública a una alegación presentada por ciudadanos preocupados (Lawson y cols., 2016).

2. En estudios confirmatorios que buscan una asociación entre una supuesta exposición ambiental y un resultado sanitario (Lawson y cols., 2016).

3. En un marco de comprobación de hipótesis para avanzar sistemáticamente en el conocimiento científico (Lawson y cols., 2016).

El objetivo del análisis *cluster* es describír la estructura latente de los datos y determinar patrones no observables directamente, este agrupa los objetos en conglomerados, de tal forma que los objetos de un mismo conglomerado sean mas parecidos entre si, que los de otros conglomerados. Para lograr esta clasificación se maximiza la homogeneidad de los objetos dentro de los conglomerados mientras que a la vez se maximiza la heterogeneidad entre los grupos o *cluster*. Los grupos

son desconocidos a priori, y por ello su detección es el objetivo primario del análisis de *cluster*. Por otro lado esta es una técnica descriptiva, ateórica y no inferencial, no tiene bases estadísticas sobre las cuales realizar inferencia estadística para una población a partir de una muestra, es un método basado en criterios geométricos y se utiliza fundamentalmente como técnica exploratoria, descriptiva pero no explicativa. Para definir la estructura de los datos agrupados de acuerdo a su semejanza es necesario definir una medida que evalúe las diferencias o similitudes entre los individuos (Fuente, 2016) . En otras palabras el análisis de *cluster* busca patrones en un conjunto de datos agrupando las observaciones (multivariantes) en conglomerados. En el análisis de clasificación, se asignan las observaciones a un número conocido de grupos o poblaciones predefinidas mientras que en el análisis de conglomerados, ni el número de grupos ni los propios grupos se conocen con anterioridad. Para agrupar las observaciones en *clusters*, muchas técnicas se basan en las similitudes entre todos los pares de observaciones, en muchos casos las similitudes se fundamentan en alguna medida de distancia. Otros métodos de agrupación utilizan una elección preliminar de los parámetros de los *clusters* o una comparación de la variabilidad dentro y entre los *clusters*, además es posible usar a correlación como una medida de similitud (Rencher, 1934).

Entre los métodos de análisis de conglomerados se describirá el método de agrupamiento jerárquico.

2.4.1. Agrupación jerárquica

Los métodos jerárquicos y otros algoritmos de clustering representan un intento de encontrar "buenos" *clusters* en los datos, utilizando una técnica computacionalmente eficiente. Por lo general, no es factible examinar todas las posibilidades de agrupación de un conjunto de datos, especialmente si es grande (Rencher, 1934). El número de formas de dividir un conjunto de n elementos en g *clusters* está dado por:

$$N(n,g) = \frac{1}{g!}\sum_{k=1}^{g}\binom{g}{k}(-1)^{g-k}k^n \tag{2.2}$$

Los algoritmos de clustering jerárquico implican un proceso secuencial, en cada paso del enfoque jerárquico aglomerativo, una observación o un grupo de observaciones se fusiona con otro grupo. En este proceso, el número de *clusters* se reduce y los *clusters* propios aumentan. Comenzando con n conglomerados (elementos individuales) y terminando con un único conglomerado que contiene todo el conjunto de datos. En cada paso de un enfoque jerárquico aglomerativo, los dos conglomerados más cercanos se fusionan en un solo conglomerado nuevo. El proceso es, por tanto, irreversible en el sentido de que los dos elementos que se agrupan en un *cluster* no se pueden separar más adelante, cualquier error inicial no puede ser corregido. Dado que un procedimiento jerárquico aglomerativo combina dos *clusters* más cercanos en cada paso, se debe considerar la cuestión de medir la similitud o disimilitud de los dos *clusters*. Los diferentes enfoques para medir la distancia entre *clusters* dan lugar a diferentes métodos jerárquicos (Rencher, 1934).

2.4.2. *Single Linkage* (Vinculación simple)

En el método de enlace simple, la distancia entre dos clusters A y B se define como la distancia mínima entre un punto de A y un punto de B:

$$D(A,B) = min[d(y_i, y_j),\ \ para\ \ y_i\ \ en\ \ A\ \ y\ \ y_j\ \ en\ \ B], \tag{2.3}$$

donde $d(y_i, y_j)$ es la distancia euclidiana o alguna otra distancia entre los vectores y_i e y_j. Este enfoque también se denomina método del vecino más cercano. En cada paso del método de enlace único se encuentra la distancia para cada par de *clusters*, y los dos *clusters* con menor distancia se fusionan. El número de conglomerados se reduce a 1. Después de fusionar dos *clusters*, el procedimiento se repite, para el siguiente paso se calculan de nuevo las distancias entre todos los pares de *clusters* y el par con la distancia mínima se fusiona en un único *cluster*. Los resultados de un procedimiento de agrupación jerárquica pueden representarse gráficamente mediante un diagrama de árbol, también conocido como dendrograma, que muestra todos los pasos del procedimiento jerárquico, incluidas las distancias a las que se fusionan los *clusters* (Rencher, 1934).

2.4.3. *Average Linkage* (Vinculación promedio)

En el enfoque de *Average Linkage*, la distancia entre dos *clusters* A y B se define como la media de las n distancias $n_A n_B$ entre los n puntos n_A de A y los n_B de B:

$$D(A, B) = \frac{1}{n_A n_B} \sum_{i=1}^{n_A} \sum_{j=1}^{n_B} d(y_i, y_j), \tag{2.4}$$

donde se suma sobre todos los y_i en A y todos los y_j en B. En cada paso, se unen los dos *clusters* con la menor distancia (Rencher, 1934).

2.4.4. *Complete Linkage* (Vinculación máxima)

En el enfoque de vinculación completa, también llamado método del vecino más lejano, la distancia entre dos conglomerados A y B se define como la distancia máxima entre un punto de A y un punto de B:

$$D(A, B) = max[d(y_i, y_j), \ para \ \ y_i \ \ en \ \ A \ \ y \ \ y_j \ \ en \ \ B], \tag{2.5}$$

En cada paso, se encuentra la distancia para cada par de *clusters*, y los dos *clusters* con la menor distancia se fusionan (Rencher, 1934).

2.4.5. Medidas de similitud

Para unir variables o individuos es necesario tener algunas medidas numéricas que caractericen las relaciones entre las variables o los individuos. Cada medida refleja asociación en un sentido particular y es necesario elegir una apropiada para el problema en estudio. La medida de asociación puede ser una distancia o una similaridad. Cuando se elige una distancia (por ejemplo, la distancia euclídea) los grupos formados contendrán individuos parecidos, de forma que la distancia entre ellos ha de ser pequeña. Cuando se elige una medida de similaridad (por ejemplo, el coeficiente de correlación) los grupos formados contendrán individuos con una similaridad alta entre ellos (Fuente, 2016).

Entre las medidas de similitud frecuentemente usadas están:

1. Para datos de intervalo:

   - Distancia euclídea:
   $$d = \sqrt{\sum_{i=1}(x_i - x_j)^2}, \tag{2.6}$$
   donde x_i es la coordenada x en la posición i y x_j es la coordenada x en la posición j.
   - Distancia euclídea al cuadrado:
   $$d = \sum_{i=1}(x_i - x_j)^2. \tag{2.7}$$

2. Para datos de frecuencia:

   - Medida chi cuadrado:
   $$x^2 = \sqrt{\sum_i \frac{(X_i - E(X_i))^2}{E(X_i)} + \sum_i \frac{(Y_i - E(Y_i))^2}{E(Y_i)}}, \tag{2.8}$$
   donde X_i es la frecuencia observada con respecto a la variable X, $E(X_i)$ es la frecuencia esperada de X, Y_i es la frecuencia observada con respecto a la variable Y y $E(Y_i)$ es la frecuencia esperada de Y.

3. Para datos binarios:

   - Distancia (índice) euclídea:
   $$d = \sqrt{b + c}, \tag{2.9}$$
   donde b y c representan las casillas diagonales correspondientes a los casos presentes en un elemento pero ausentes en el otro.

   - Distancia (índice) euclídea al cuadrado:
   $$d = b + c. \tag{2.10}$$

En la siguiente sección se presentá las técnicas de *cluster* espacial.

2.5. *Cluster* espacial

El *cluster* espacial corresponde a un método de agrupamiento, es una extensión del *cluster* tradicional o convencional, teniendo presente la ubicación geográfica de los objetos (como por ejemplo de personas), considerando variables tales como latitud y longitud. El objetivo del *cluster* espacial es identificar patrones espaciales con el fin de estudiar y presenciar aumentos o disminuciones con respecto a la variable de interés en ubicaciones geográficas especifícas.

- Identificación exploratoria de las características de la agrupación global.
  Una variedad de métodos estadísticos espaciales comprueban la agrupación global cuando la hipótesis nula es la aleatoriedad espacial en la distribución de los valores de los datos. Las pruebas se dividen en: basadas en la distancia, cuando se dispone de datos a nivel individual y basadas en el área, las cuales cuentan el número de eventos en subáreas definidas.
  Existen métodos estadísticos espaciales para comprobar la correlación espacial en datos geográficos, entre estos, el *Moran test* y *Geary*, el estadístico G de *Getis-Ord* (Haining, 2003). El término "agrupación" depende de la escala, las pruebas se basan en la estimación de una función que describe la dependencia espacial en la distribución de objetos puntuales en un rango de escalas de distancia. La ubicación de cada individuo en la población es fija, por ende el objetivo del análisis es comprobar si el subconjunto fijado está agrupado espacialmente dentro de la población (Haining, 2003).

- Agrupación en datos geográficos.
  El recuento de uniones para los datos de la variable en estudio se basa en el recuento del número de relaciones de determinados colores. Supongamos que cada zona está coloreada en negro (N) o en blanco (B). Si existe una tendencia a encontrar valores similares juntos (autocorrelación espacial positiva), el número de uniones en las que las coloraciones adyacentes son iguales (NN o BB) será"grande"(o coloración alta) , y el número de uniones en las que las coloraciones adyacentes son diferentes (NB) será "pequeño"(o coloración baja) en relación con los recuentos para un mapa geográfico aleatorio. La agrupación puede generalizarse con más de dos colores y para definiciones generales de diferencia (en términos de ponderación u orden de vecindad) (Haining, 2003).

- Agrupación en un patrón de puntos fijos.
  El objetivo es determinar si algún atributo está agrupado en una población dada su distribución espacial. Cada individuo tiene o no tiene un atributo en particular. Utilizando la terminología epidemiológica, un individuo con un atributo se llama caso (c) y un individuo sin el atributo se llama control (k). (Cuzick y Edwards, 1990) proporcionan una prueba que tiene en cuenta la homogeneidad de la distribución de la población, esta prueba cuenta para cada caso, el número de sus vecinos más cercanos (v) los cuales también se consideran casos. Este recuento se compara con el conteo que se espera bajo la hipótesis nula de una asignación aleatoria de casos y controles. Cuando los casos están agrupados espacialmente, los vecinos más cercanos a un caso tenderán a ser otros casos, por lo que el test de Cuzick y Edward será grande. Un problema de esta prueba es que las distancias de los vecinos más cercanos pueden ser diferentes en todo el mapa (Haining, 2003).

- Datos geográficos.
  Dado que las datos geográficos tienen tamaños de población diferentes, no es apropiado definir los *clusters* en función de las zonas que tienen más casos. Una de las primeras pruebas de

(Choynowski, 1959) calcula la probabilidad Poisson de obtener x o más casos observados bajo la hipótesis nula de una distribución aleatoria de casos en toda el área de estudio. La prueba puede adaptarse fácilmente a diferentes categorías de riesgo diferenciadas, por ejemplo, por características como el sexo, la edad, la ubicación rural-urbana siempre y cuando los datos estén disponibles (Haining, 2003).

En el análisis *cluster* de datos espaciales las matrices de pesos espaciales (o también conocida como matriz de distancias) se utilizan para determinar las interacciones de los lugares estudiadas en el análisis (Aral y Bakir, 2001), está está definida por:

$$W_{ij} = \begin{pmatrix} w_{1,1} & \cdots & w_{1,j} \\ w_{2,1} & \cdots & w_{2,j} \\ \vdots & \vdots \ddots & \vdots \\ w_{i,1} & \cdots & w_{i,j} \end{pmatrix}, \tag{2.11}$$

donde $w_{i,m}$ representa la relación entre las zonas i y j, con $i, j = 1, 2, \ldots, m$.

2.5.1. Índice de Morán

La autocorrelación espacial proporciona la correlación entre las observaciones en un lugar determinado y está asociada a la proximidad geográfica de estas observaciones. Las medidas de autocorrelación espacial se dividen en dos categorías, como escalas globales y escalas locales, según el ámbito o la escala del análisis (Kisilevich, Mansmann, Nanni, y Rinzivillo, 2010). Para verificar la existencia o ausencia de autocorrelación espacial, se pueden utilizar diferentes índices, siendo el I de Morán el más conocido y utilizado en la práctica. Todos los índices tratan de contrastar la hipótesis nula de ausencia de autocorrelación espacial (esto es, aleatoriedad espacial) frente a la hipótesis alternativa de existencia de autocorrelación espacial (positiva o negativa) (Sanchéz, 2008).

Índice de Morán Global

La medida I de Morán se utiliza habitualmente en las mediciones de autocorrelación espacial global, y está dada por:

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij}(x_i - \bar{x})}{\sum_{i=1}^{i=n} \sum_{j=1}^{j=n} W_{ij} \sum_{i=1}^{i=n} (x_i - \bar{x})^2}, \tag{2.12}$$

donde n representa el número de áreas de la muestra; i y j, dos de las unidades de área; W_{ij}, la matriz de peso espacial (la similitud de las posiciones de i y j); x_i el valor de la variable en la posición i; (x_i -$\bar{x}$)(x_j - $\bar{x}$), la similitud de la variable en las posiciones i y j (Aral y Bakir, 2001).

El valor I de Morán varía en el rango de $[-1, +1]$, de forma similar al coeficiente de correlación. Un valor I de Morán de 0 indica que no hay autocorrelación espacial, en este caso, todas las provincias o zonas incluidas en el análisis tienen una distribución aleatoria, lo que significa que no hay *clusters* relevantes en el espacio. Si el I de Morán muestra un valor de 1, indica una correlación perfectamente positiva en la que valores similares se agrupan o interactúan con sus vecinos. En este caso, las provincias o zonas de alto valor o de bajo valor se agrupan juntas. Si el I de Morán muestra un valor de -1, indica una correlación negativa perfecta en la que las observaciones diferentes se agrupan, en este caso las observaciones de valor alto y bajo se agrupan. Cuanto mayor sea el valor absoluto del I de Morán, más fuerte será la autocorrelación espacial (Aral y Bakir, 2001).

Gráfico de Morán (global) (Diagrama de dispersión de Morán)

El "Diagrama de dispersión de Morán" (Morán plot (global)), consiste en la división de cuatro cuadrantes, el superior derecho y el inferior izquierdo son indicativos de autocorrelación positiva, lo que significa que el comportamiento de una variable en algunas zonas está siendo favorecido por su comportamiento en zonas próximas. Por otro lado, el superior izquierdo y el inferior derecho refleja la existencia de autocorrelación espacial negativa, o sea que el comportamiento de una variable está produciendo efectos contrarios en las zonas próximas (Barros y Aroca, 2014).

Por otro lado, el primer cuadrante superior derecho, indica la existencia de valores altos de la variable de estudio rodeados de valores altos, y en el caso del cuadrante inferior izquierdo indica la existencia de valores bajos rodeados de valores bajos, estos dos cuadrantes son indicativos de agrupación espacial de valores altos y bajos. En el caso del cuadrante superior izquierdo, indica la presencia de valores bajos de la variable estudiada rodeado de valores altos, estos se denominan valores atípicos bajos espaciales, y en el caso del cuadrante inferior derecho indica la presencia de valores altos de la variable rodeadas de valores bajos, estos se hacen llamar valores atípicos altos espaciales (Barros y Aroca, 2014). En la figura 2.6 se presenta un ejemplo de un Morán plot global de la variable tasa de incidencia del COVID-19.

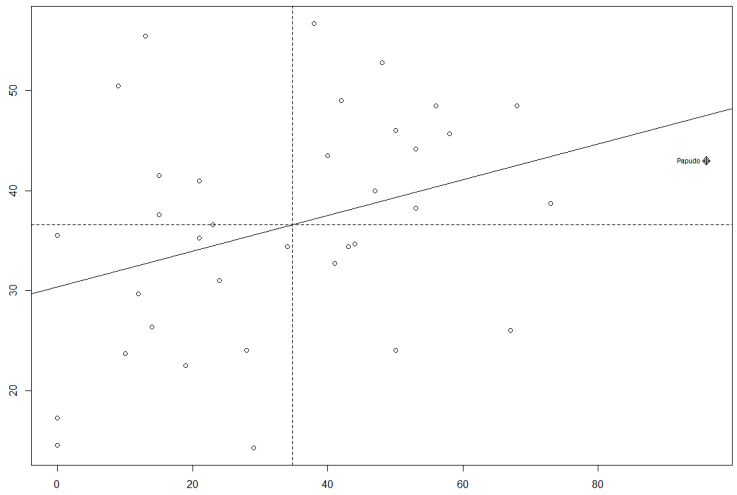

Figura 2.6: Gráfico de Morán global de la tasa de incidencia de COVID-19 para las comunas de la Región de Valparaíso.

Fuente: elaboración propia.

Índice de Moran local

Los indicadores locales de asociación espacial (LISA) se emplean para determinar las agrupaciones espaciales de valores similares en torno a la observación de una variable, indicando la presencia de puntos calientes y fríos con agrupaciones locales de valores de observación altos o agrupaciones locales de valores de observación bajos. Junto con esta clasificación se dan cuatro tipos de asociaciones, esto demuestra la existencia de clusters locales (alto-alto o bajo-bajo) o valores atípicos espaciales locales (alto-bajo o bajo-alto) (Aral y Bakir, 2001). El I de Moran local, está dado por:

$$I_i = (x_i - \bar{x}) * \sum_{j=1}^{j=n} W_{ij}(x_j - \bar{x}), \tag{2.13}$$

donde x_i es un atributo para la característica i, $W_{i,j}$ es el peso espacial entre la característica i y la j y $\bar{x}$ la media de las observaciones vecinas.

Lisa *cluster map*

Para detectar la autocorrelación espacial a nivel local, el gráfico más utilizado es el denominado mapa LISA, correspondiente a los indicadores de asociación local espacial, este mapa utiliza los índices de Moran locales calculados para todas las unidades espaciales analizadas. El mapa LISA representa las localizaciones con valores significativos en estadísticos locales, por lo que se puede detectar la existencia de puntos o valores atípicos espaciales y la formación de *cluster*. En la figura 2.7 se puede observar un LISA *cluster map* de la variable tasa de incidencia del COVID-19, este indica las categorías de asociación espacial y las respectivas comunas que fonrman los *clusters*.

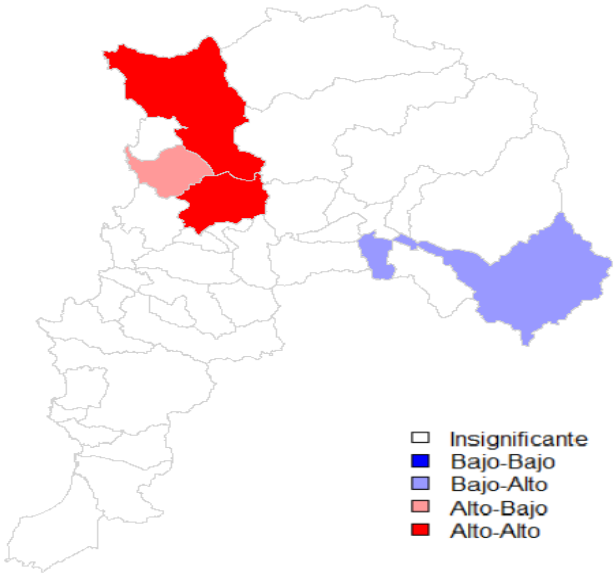

Figura 2.7: LISA *cluster map* de la tasa de incidencia de COVID-19 para las comunas de la Región de Valparaíso.

Fuente: elaboración propia.

Capítulo 3

Resultados

En la siguiente sección se presentan los resultados obtenidos en los análisis desarrollados.

3.1. Base de datos

El conjunto de datos denominado "Datos-COVID1", está disponible en la página web del Ministerio de la Ciencia, Tecnología, Conocimiento e Innovación (MinCiencia), está subdividido en 93 productos (Data products), cada uno de los productos contiene archivos de alguna variable de interés (MinCiencia, 2020). El período de estudio considerado para este trabajo corresponde a Junio del 2020 hasta Junio del 2021, este período se determinó por el cambio de Ministro de Salud y la frecuencia en el registro de los datos COVID-19 y la vigencia del primer "Plan paso a paso" establecido en Chile. La base de datos registra para las 16 regiones de Chile variables de interés epidemiológico, concernientes al COVID-19. Para este estudio se considera el registro de las 16 regiones de Chile continental, dado que es un estudio en donde se tiene en cuenta la ubicación y cercanía de las respectivas regiones y comunas, por ende los registros de las islas Juan Fernández y de Pascua no se consideran en el conjunto de datos ya que interviene el límite natural (oceano Pacífico) y además la escasez de registros de casos dentro de las islas, ya que se determinó cerrar el ingreso y salida de personas al inicio de la pandemia.

Las variables consideradas para este estudio son :

- Casos diarios confirmados con COVID-19: correspondiente al producto 18 del conjunto de datos del Ministerio de la Ciencia, Tecnología, Conocimiento e Innovación (MinCiencia, 2020), este da cuenta de los casos confirmados y casos probables (desde Junio 2020) notificados en cada una de las comunas de Chile, según residencia. Cada uno de los archivos corresponde a un informe publicado por el Ministerio de Salud del país.
- Fallecidos por COVID-19: correspondiente al producto 14 del conjunto de datos del Ministerio de la Ciencia, Tecnología, Conocimiento e Innovación (MinCiencia, 2020), presenta un set de 2 archivos que dan cuenta del número de personas con diagnóstico COVID-19 fallecidas por día, según región de residencia.
- Hospitalizados por grupo de edad y sexo: correspondiente al producto 22 del conjunto de datos del Ministerio de la Ciencia, Tecnología, Conocimiento e Innovación (MinCiencia, 2020), este producto contiene dos archivos, 1) el número acumulado del total de pacientes hospitalizados

con diagnóstico COVID-19 por rango de edad y género y 2) el número acumulado de pacientes internados con diagnóstico COVID-19 en la Unidad de Cuidados Intensivos (UCI) por rango de edad.

- Casos nuevos con síntomas COVID-19: correspondiente al producto 26 del conjunto de datos del Ministerio de la Ciencia, Tecnología, Conocimiento e Innovación (MinCiencia, 2020), presenta un set de 3 archivos que dan cuenta del número de casos nuevos por día según resultado del diagnóstico que presentan síntomas, por región de residencia, reportados por el Ministerio de Salud.

- Casos nuevos sin síntomas COVID-19: correspondiente al producto 27 del conjunto de datos del Ministerio de la Ciencia, Tecnología, Conocimiento e Innovación (MinCiencia, 2020), presenta un set de 3 archivos que dan cuenta del número de casos nuevos por día según resultado del diagnóstico que son asintomáticos, por región de residencia, reportados por el Ministerio de Salud.

- Pacientes en Unidad de Cuidados Intensivos (UCI): correspondiente al producto 8 del conjunto de datos del Ministerio de la Ciencia, Tecnología, Conocimiento e Innovación (MinCiencia, 2020), presenta un set de 2 archivos que dan cuenta del número de pacientes en UCI y que son casos confirmados por COVID-19, por región reportados diariamente por el Ministerio de Salud.

3.2. Análisis descriptivo

3.2.1. Análisis descriptivo clásico

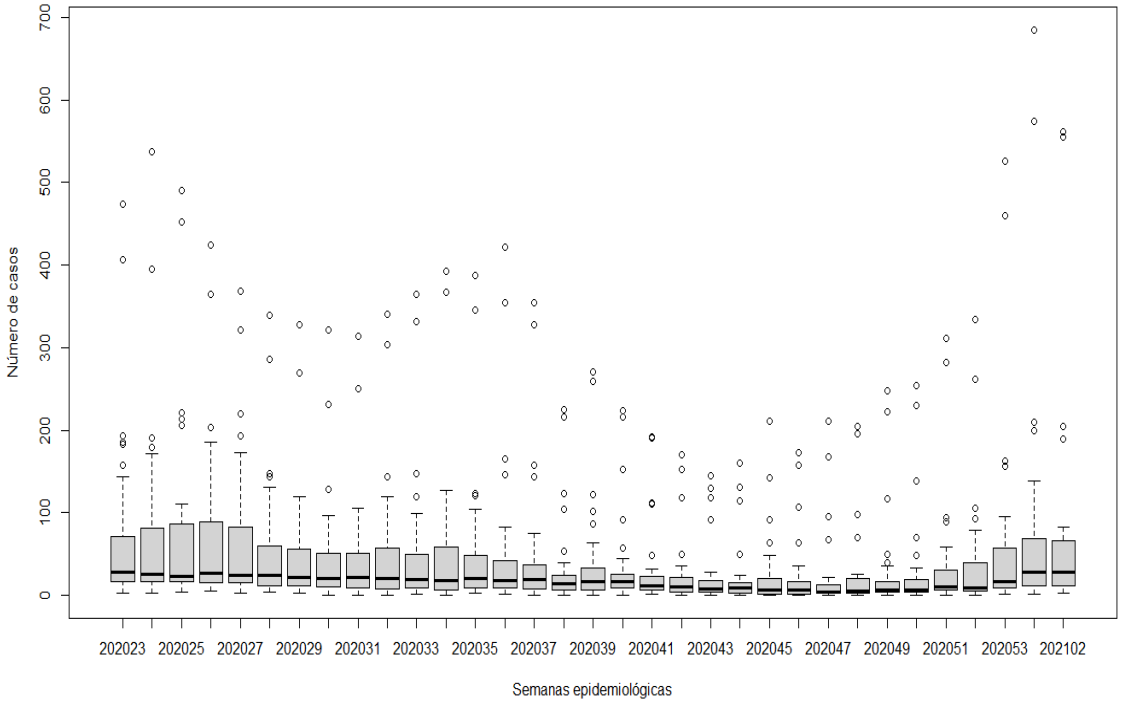

Figura 3.1: Boxplot de la variable Casos Nuevos Confirmados con COVID-19 por semana epidemiológica.

Fuente: elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

En la figura 3.1 se puede observar que en la semana epidemiológica (SE) 01 del 2021 (202101) se presenta el mayor número de Casos Nuevos Confirmados con COVID-19, la cifra se acerca a los 700 casos. Luego de que las cifras disminuyeran considerablemente a finales del 2020, se observa un patrón de crecimiento, superando incluso el período del inicio del estudio.

Región	Min.	1er. cuartil	Mediana	Media	3er. cuartil	Max.
Arica y Parinacota	8	50	150	135	207	254
Tarapacá	22	168	227	208	258	354
Antofagasta	32	252	493	432	562	654
Atacama	0	12	67	62	106	128
Coquimbo	2	74	213	179	270	342
Valparaíso	70	507	882	813	1151	1397
Metropolitana	824	7599	9262	8463	10356	11360
O'Higgins	18	238	422	379	544	660
Maule	17	178	327	328	491	704
Ñuble	18	62	123	122	178	265
Biobío	11	120	338	391	610	1065
La Araucanía	31	78	117	154	229	405
Los Ríos	7	14	20	39	55	161
Los Lagos	12	41	96	125	203	409
Aysén	0	1	1	6	14	20
Magallanes	14	28	83	109	195	273

Cuadro 3.1: Análisis descriptivo para la variable Fallecidos por COVID-19 por región. Fuente: elaboración propia.

Como se observa en el Cuadro 3.1, la región que mayor resalta con respecto a la variable Fallecidos por COVID-19 es la Región Metropolitana, pues es la que presentó mayor número de fallecidos, con 11.360 muertes, este comportamiento es de esperar ya que la región en particular es la que presenta mayor densidad poblacional (sobre los 5 millones de habitantes). Por otro lado, la región la Región de Valparaíso es la segunda con mayor número de fallecidos registrando 1.397 muertes, cabe destacar que es una de las más pobladas después de la Región Metropolitana, además tiene mayor flujo de translado de personas desde y hacia la Región Metropolitana. La Región de Aysén es una de las que presenta menor densidad poblacional en Chile, esta muestra un bajo número de fallecidos por COVID-19 (20) en comparación con el resto de las regiones. La segunda región que registro el menor número de fallecidos es la Región de Atacama, este comportamiento llama bastante la atención, ya que sus regiones vecinas (Antofagasta y Coquimbo) superan los 300 casos.

Años	Min.	1er. cuartil	Mediana	Media	3er. cuartil	Max.
Hombres						
0-4 y menos	100	303	384	350	414	458
5-17	72	262	350	319	387	447
18-49	1495	3826	7017	6176	7972	9304
50-59	1031	3131	5886	5152	6735	7932
60-69	1036	3271	6220	5483	7289	8731
70-79	791	2670	4654	4171	5503	6722
80 y más	459	1615	2727	2485	3299	3875
Mujeres						
0-4 y menos	102	285	369	342	415	441
5-17	75	273	384	346	431	485
18-49	1325	3433	6142	5462	7099	8319
50-59	629	1841	3427	3038	3966	4815
60-69	666	2223	4125	3629	4796	5766
70-79	609	2047	3553	3181	4202	5113
80 y más	485	1743	2969	2676	3553	4301

Cuadro 3.2: Análisis descriptivo para la variable Hospitalizados por COVID-19 por grupo de edad y sexo.

Fuente: elaboración propia.

Tanto para los hombres como para las mujeres el mínimo de hospitalizados por COVID-19 se sitúa en el intervalo de edad entre los 5 a 17 años (Ver Cuadro 3.2), es importante resaltar el gran número de hospitalizados de hombres y mujeres entre los 18 y los 49 años, 9.304 y 8.319 respectivamente. Finalmente el grupo de mujeres presenta un número mayor de hospitalizados con respecto a las personas adultas (en los 80 años y más) con 4.301 hospitalizados, preocupa bastante el gran número de hospitalizados tanto para los hombres y como para mujeres, en el rango de edad de 70 y 79 años, e inclusive más, puesto que las personas de tercera a edad tienden a padecer enfermedades crónicas que perjudican el estado de salud mientras están hospitalizados por COVID-19. Llama la atención las cifras de hospitalizados, tanto para los hombres y mujeres, en el intervalo de edad entre menos de uno y 4 años de edad, puesto que a muchos pacientes que han padecido de COVID-19 han presentado al menos una secuela a lo largo del tiempo.

Región	Min.	1er. cuartil	Mediana	Media	3er. cuartil	Max.
Arica y Parinacota	0	16	29	34	46	132
Tarapacá	8	18	34	48	68	202
Antofagasta	3	25	48	83	127	375
Atacama	0	9	19	25	34	196
Coquimbo	2	17	45	49	75	149
Valparaíso	26	67	113	122	163	332
Metropolitana	86	241	338	738	570	5264
O'Higgins	11	39	66	77	93	263
Maule	16	57	82	95	117	323
Ñuble	0	22	34	37	49	103
Biobío	12	126	167	188	226	476
La Araucanía	1	28	68	85	126	279
Los Ríos	1	6	21	33	46	155
Los Lagos	0	43	73	102	117	440
Aysén	0	0	3	7	8	44
Magallanes	0	12	52	55	76	253

Cuadro 3.3: Análisis descriptivo para la variable Casos Nuevos con Síntomas COVID-19 por región. Fuente: elaboración propia.

En el Cuadro 3.3 se observa que las Regiones Centrales, entre ellas, Coquimbo, Valparaíso y Metropolitana presentan un comportamiento bastante diferente en el promedio de Casos Nuevos con Síntomas, obviamente la Región Metropolitana registró el mayor número de casos debido a su alto número de población y al hecho de que es la entrada aérea principal del país. Por otro lado, el promedio de casos de las regiones Ñuble y Tarapacá fueron bastaste similares, esto llama bastante la atención ya que ambas regiones son lejanas una de la otra. Es importante resaltar el comportamiento presentado en la Región de Aysén, esta presenta el menor número de casos, esto se debe a una baja densidad poblacional y por ser una región poco turística.

Series de tiempo: Casos confirmados con síntomas de COVID-19

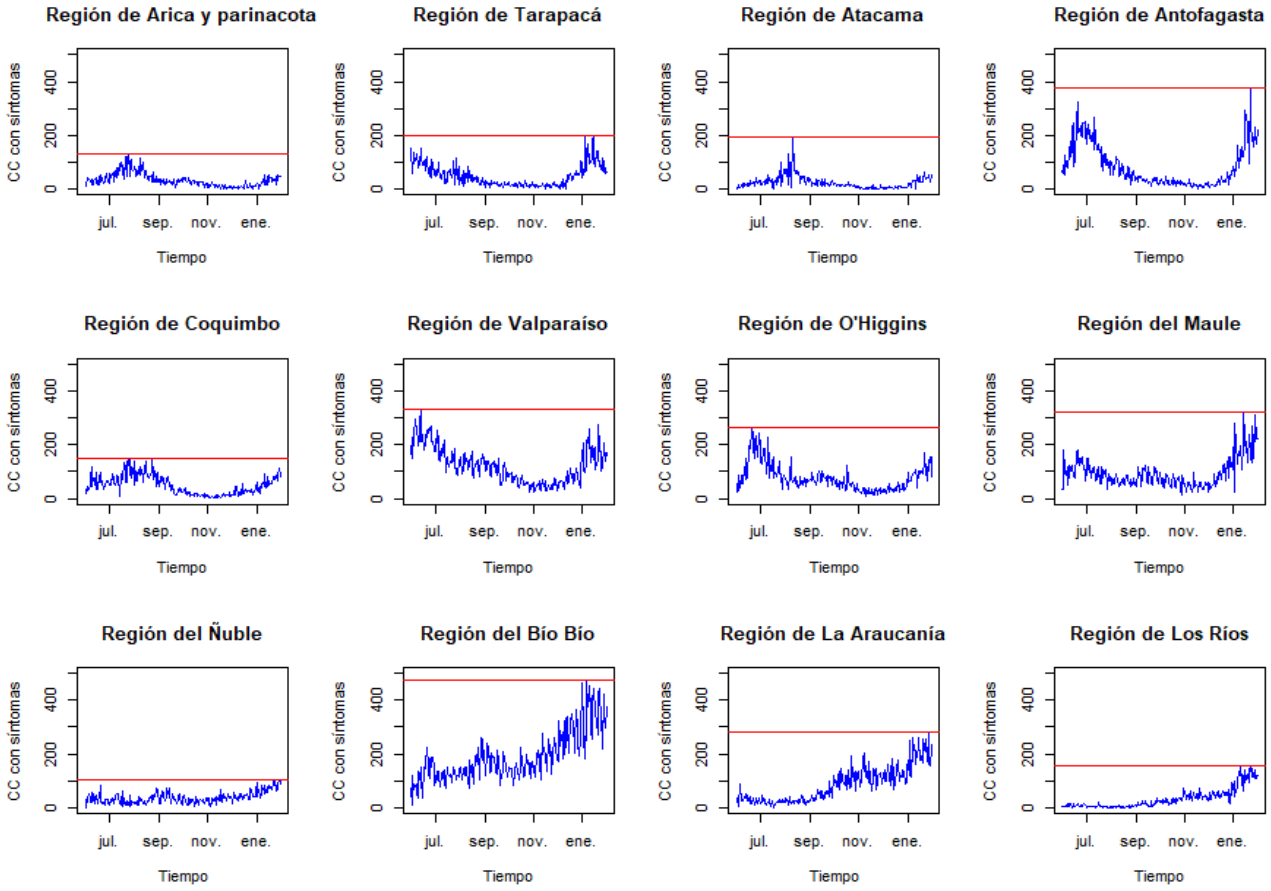

Figura 3.2: Serie temporal del Número de Casos Confirmados (CC) con Síntomas (CS) COVID-19.
Fuente: elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

Las regiones que presentan menor número de Casos Confirmados con Síntomas durante el período de estudio, no superando los 210 casos confirmados, fueron las regiones de Arica y Parinacota, Tarapacá, Atacama, Coquimbo y Los Ríos, siendo la Región de Tarapacá la que presentó el pick mas alto en comparación con las comunas mencionadas anteriormente, este fue de 202 casos en Enero 2021 (Ver Figura 3.2), mientras que el segundo pick mas alto fue registrado por la región vecina la Región de Atacama, con 196 casos en el período de Agosto del 2020. Como se observa en la Figura 3.2. Las regiones de Arica y Parinacota, Tarapacá, Atacama, Coquimbo, Ñuble y Los Ríos, no superan los 200 Casos Confirmados con síntomas a lo largo del período de estudio, mientras que las demás regiones entre ellas la región de Antofagasta, Valparaíso, O'Higgins, El Maule, El Bio Bío y La Araucanía, superan los 220 casos confirmados, entre las regiones sobre los 200 casos. La región que presentó un pick mal alto en comparación a las demás regiones fue la Región del Bio Bío registrando, de 476 en Enero 2021, mientras que la segunda región con el segundo pick mas alto fue la Región de Antofagasta registrando alrededor de 375 casos confirmados con síntomas COVID-19. La Región del Bio Bío llama bastante la atención puesto que el comportamiento de sus casos mantiene una tendencia al alza no como en las demás regiones que presentan aumentos y caídas a lo largo del tiempo. Se esperaría que la Región del Maule, tuviese un comportamiento similar a la región del

Bio Bío, ya que el Maule fue la primera región en registrar los primeros casos de COVID-19 en Chile.

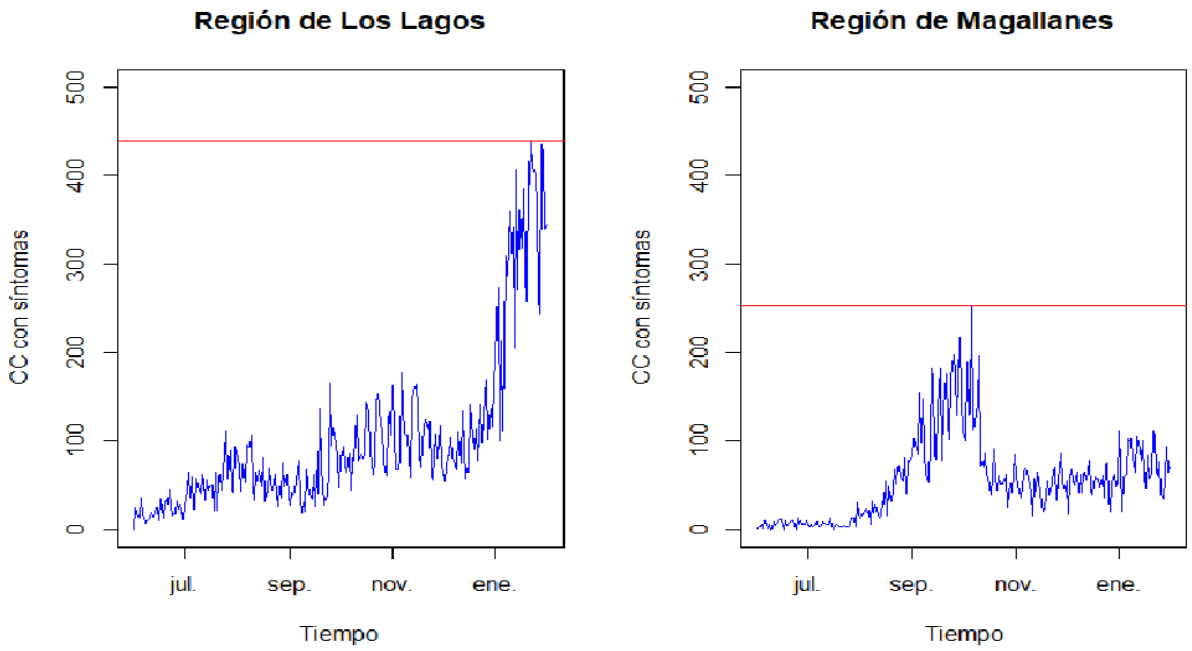

Figura 3.3: Serie temporal del número de Casos Confirmados (CC) con Síntomas COVID-19.
Fuente: elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

La Región de Los Lagos en comparación con la Región de Magallanes presentó mayor número de CC con síntomas superando los 400 casos, mientras que Magallanes registró alrededor de 250 casos confirmados (Ver Figura 3.3), esto se debe a que la última región de Chile (Magallanes) contiene una densidad poblacional considerablemente menor que en el resto de las regiones del país. La Región de los Lagos presentó una tendencia alza, con un pick de 440. Por otro lado, la Región de Magallanes, desde Junio 2020 hasta Agosto 2020, no presentó grandes oscilaciones con respecto al número de caso casos, a finales del Agosto 2020 los casos aumentaron repentinamente llegando a un pick de 253 casos, para luego dismunuir considerablemente, posiblemente debido a las medidas de intervención implementadas ya estaban implementadas a esa fecha.

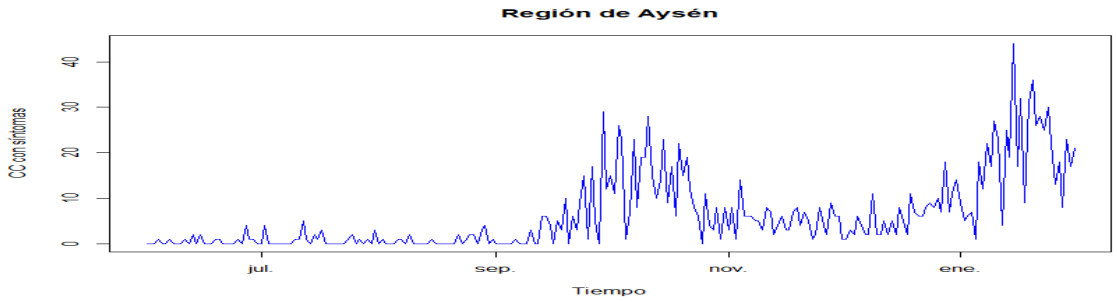

Figura 3.4: Serie temporal del número de Casos Confirmados (CC) con Síntomas COVID-19.
Fuente: elaboración propia.
Datos descargados: Ministerio de la Ciencia.

La Región de Aysén fue la región que presentó el menor número de CC con Síntomas a lo largo del período de estudio, registrando como pick 44 casos en febrero 2021, este comportamiento se debe a que la región no supera los 102.320 habitantes, una densidad poblacional bastante baja en comparación a las demás regiones de Chile. El comportamiento de los casos desde Junio 2020 a Septiembre 2020 no presentó muchas oscilaciones, además no se superaron los 10 contagios. A partir de Agosto 2020 hubo un aumento significativo, llegando a registrar alrededor de 30 casos confirmados (Ver Figura 3.4).

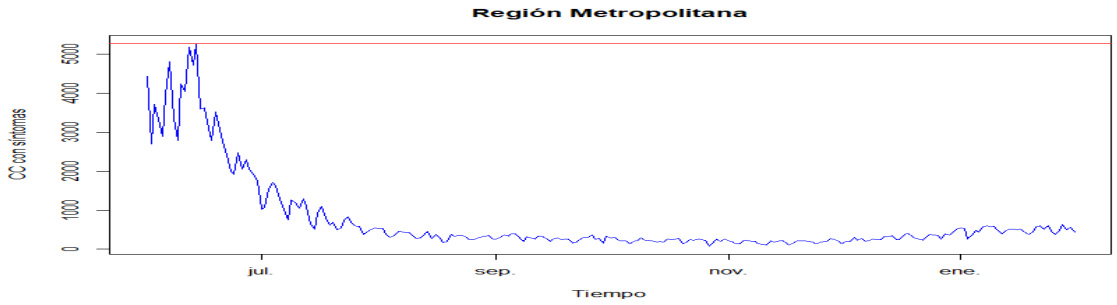

Figura 3.5: Serie temporal del número de Casos Confirmados (CC) con Síntomas COVID-19.
Fuente: elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

En comparación con el resto de las regiones de Chile, la Región Metropolitana fue la región que presento el mayor número de CC con síntomas durante el período de estudio, esto se debe a tu alta densidad poblacional que se sitúa alrededor de los 7.037.000 habitantes según el Censo del 2017. Con respecto al comportamiento de CC, en la Figura 3.5 se observa que en Junio del 2020 los casos se registraban sobre los 5.000 casos, mes donde registro el pick con 5.264 casos. Cando las medidas de intervención se implementaron en Julio 2020, fue de gran de ayuda para controlar el alto contagio de los casos entre residentes de la Región Metropolitana, pues los recintos de salud pública ya estaban colapsados. A principio de Julio 2020 los casos fueron disminuyendo significativamente registrando alrededor de 2000, luego los casos se mantuvieron con pequeños aumentos y caídas, no superando los 1000 y 500 casos respectivamente.

***Cluster* convencional.**

La implementación del método *cluster* se realiza con el propósito de explorar las agrupaciones de las regiones de Chile con respecto a ciertas variables consideradas en este estudio entre ellas: Casos Nuevos con Síntomas, Casos Nuevos sin síntomas, Fallecidos por COVID-19, Pacientes en UCI y Exámenes PCR. Los siguientes *cliusters* convencionales, es decir, sin considerar información espacial, se implemantan con el método de *linkage* (enlace) *average*, éste consiste en la distancia media, y se calcula la distancia entre todos los posibles pares formados por una observación de un *cluster* A y un *cluster* B. El valor promedio de todas ellas se selecciona como la distancia entre los dos *clusters* (mean intercluster dissimilarity).

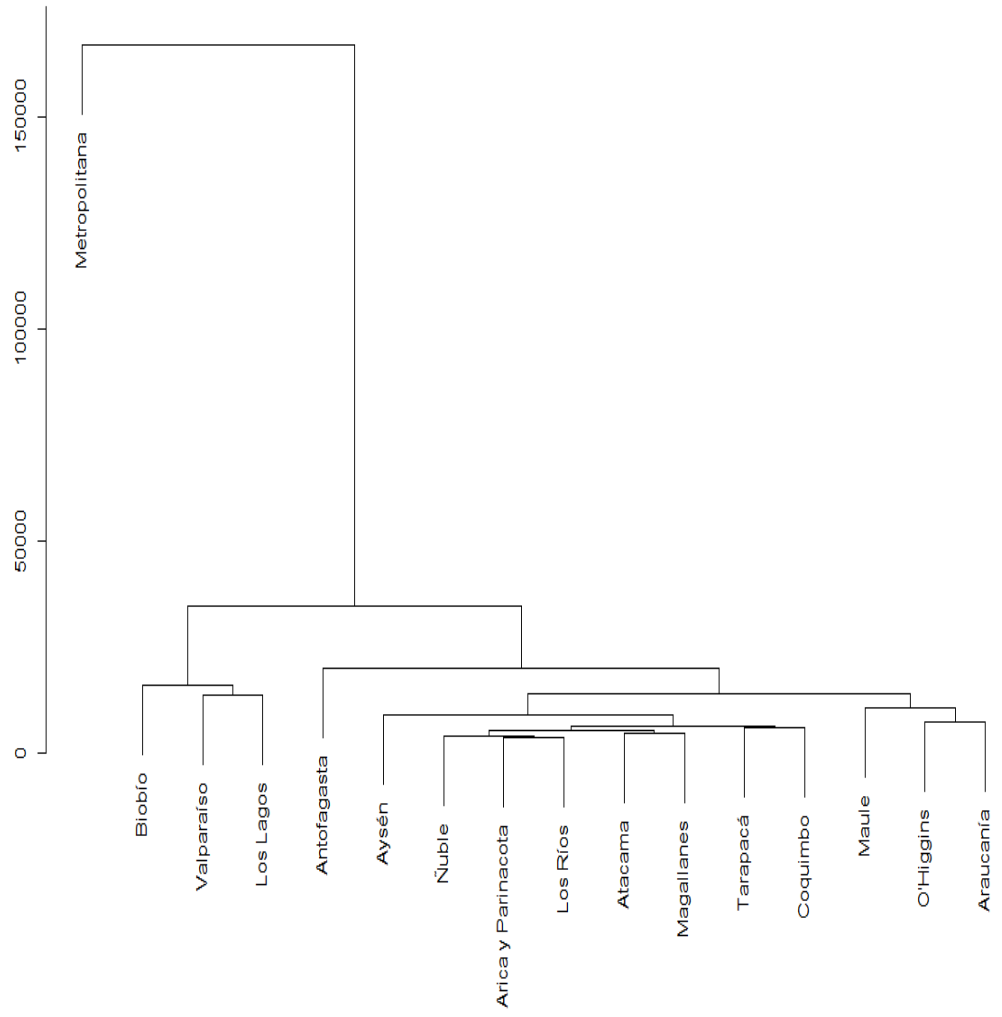

Figura 3.6: *Cluster* de la variable Exámenes PCR por región.
Fuente: elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

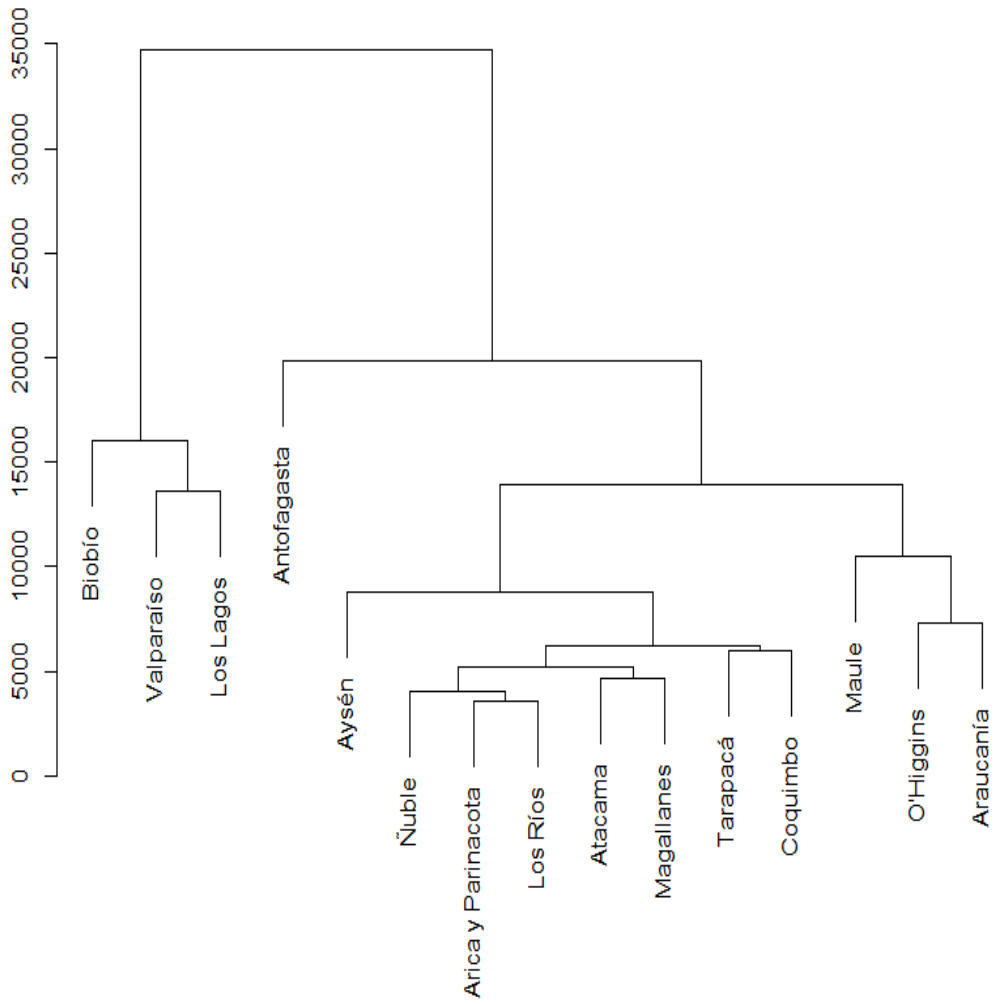

Figura 3.7: *Cluster* de la variable Exámenes PCR por región sin considerar la Región Metropolitana.
Fuente: elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

Las regiones de Chile se han agrupado con respecto a su similitud de la Cantidad de Exámenes PCR hechos durante el período de estudio (Ver Figura 3.6). Se puede observar que a simple vista la Región Metropolitana se ha comportado totalmente distinta a la demás regiones de Chile, esto se debe a la alta población que presenta la región en particular, sin embargo no se pueden ver los *cluster* que contienen a las demás de regiones. En la Figura 3.7 se puede observar el comportamiento de la misma variable de las regiones de Chile sin conciderar la Región Metropolitana con el fin de observar mucho mejor el comportamiendo de las agrupaciones de las regiones con respecto a la variable de estudio. La figura muestra que se han formado al menos cuatro *cluster* entre ellos el *cluster* que contiene a las Regiones Biobío, Valparaíso y Los Lagos, por ende podemos interpretar que las regiones nombradas anteriormente tuvieron un comportamiento similar con respecto al número de Exámenes PCR hechos.

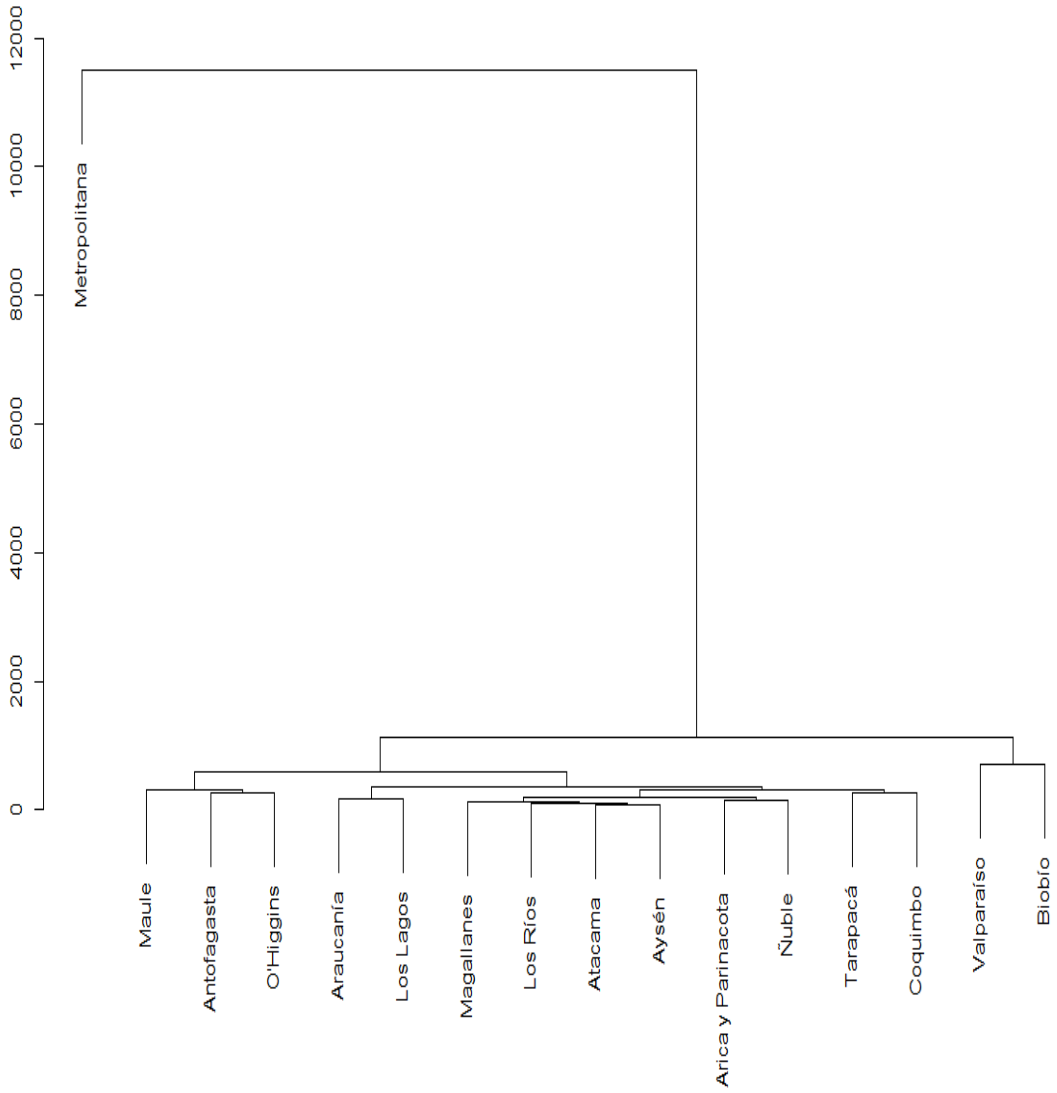

Figura 3.8: *Cluster* de la variable Pacientes en UCI.
Fuente: elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

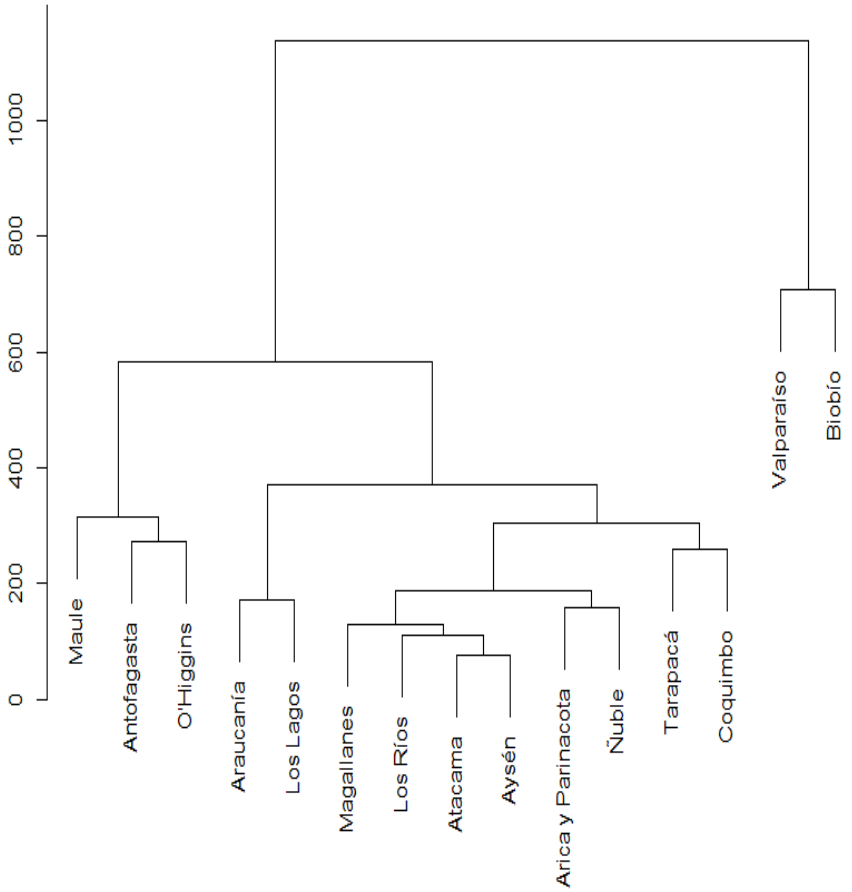

Figura 3.9: *Cluster* de la variable Pacientes en UCI sin considerar la Región Metropolitana.
Fuente: elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

En la Figura 3.8 se puede observar el agrupamiento de las regiones con respecto a la variable Número de Pacientes en UCI. Nuevamente la Región Metropolitana no se encuentra en ningún grupo sino alejada con un comportamiento totalmente diferente. A simple vista no se pueden observar la cantidad de *cluster* que se han formado con respecto a la variable en estudio. Para una mejor interpretación la Figura 3.9 indica el comportamiento de los *cluster* con respecto a la misma variable, sin considerar la Región Metropolitana, observándose con mejor precisión que existen al menos tres grupos, entre ellos el grupo que contiene a las Regiones El Maule, Antofagasta y O'Higgins, registrando alrededor de 300 Pacientes en UCI, es decir que las regiones antes mencionadas presentaron un comportamiento similar con respecto a la variable de estudio. Lo que llama bastante la atención puesto que son regiones con densidades poblacionales muy disímiles, además se encuentran muy lejanas de una de la otra.

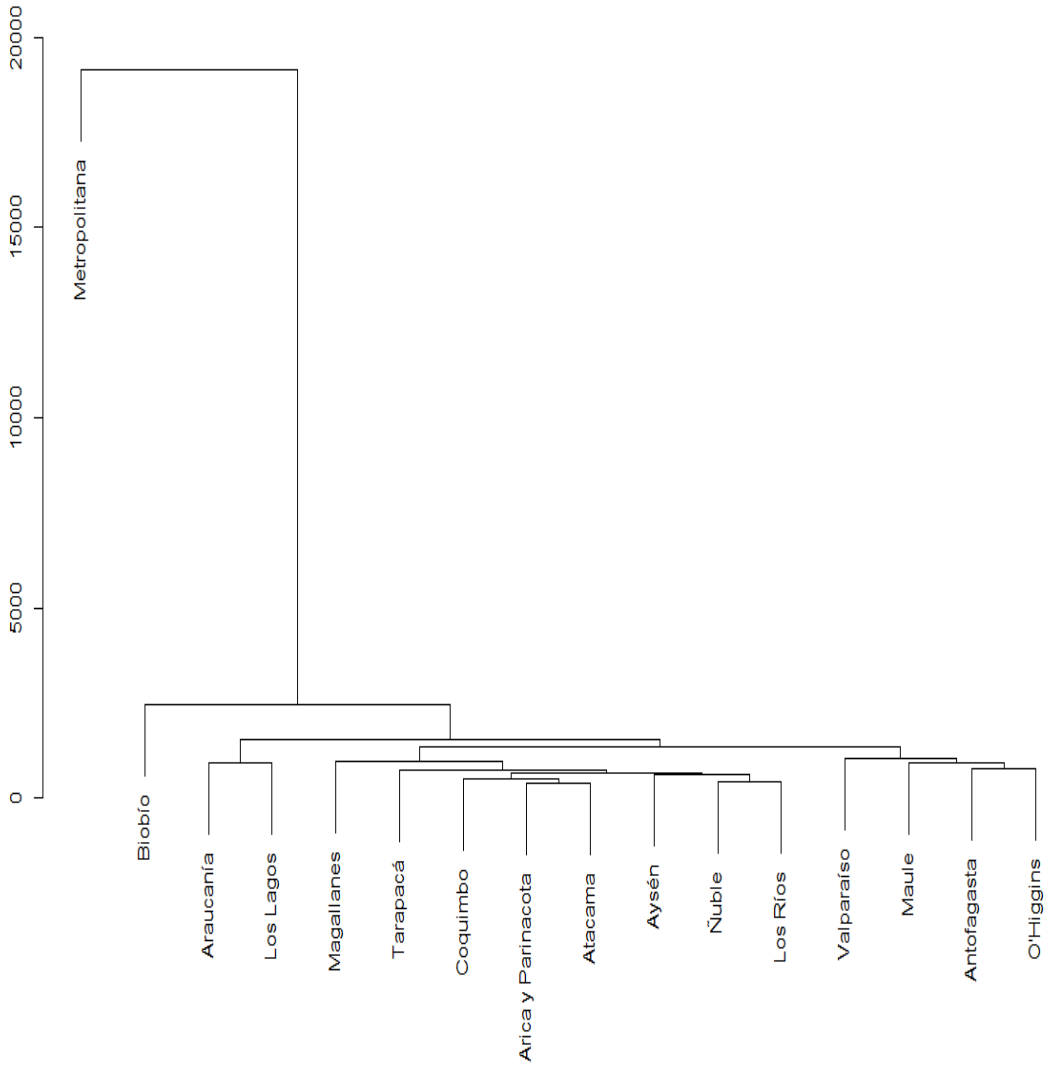

Figura 3.10: *Cluster* de la variable Pacientes Nuevos con Síntomas de COVID-19.
Fuente : elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

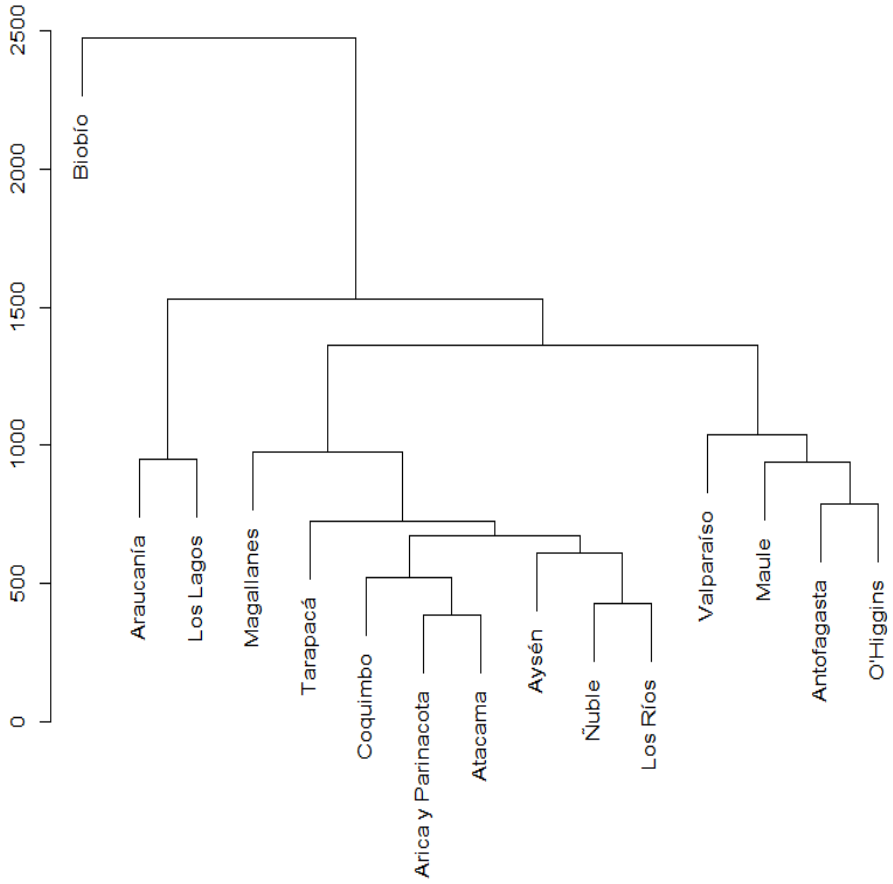

Figura 3.11: *Cluster* de la variable Pacientes Nuevos con Síntomas de COVID-19 sin considerar la Región Metropolitana.

Fuente : elaboración propia.

Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

En la Figura 3.10 nuevamente se observa el comportamiento distintivo de la Región Metropolitana con respecto al resto de las regiones de Chile, registrandose alrededor de 20.000 Pacientes Nuevos con COVID-10 con Síntomas. La región presenta un comportamiento tan distinto a las demás regiones que no se logra ver los demás *cluster*. La Figura 3.11 presenta una mejor clasificación, sin el efecto de la Región Metropolitana. Se observan al menos cuatro grupos entre ellos el *cluster* que contiene a las Regiones de Valparaíso, El Maule, Antofagasta y O'Higgins, registrando alrededor de 1.000 Casos Nuevos Confirmados con Síntomas de COVID-19. Nuevamente las comunas se encuentran bastante lejanas de una a la otra, a excepción de las Regiones de Valparaíso y O'Higgins quienes son regiones vecinas. Por otro lado, la Región del Biobío, se comportó muy distinta a las demás regiones de Chile registrando alrededor de 2.500 Casos Nuevos con Síntomas COVID-19, apartándose significativamente de los grupos formados.

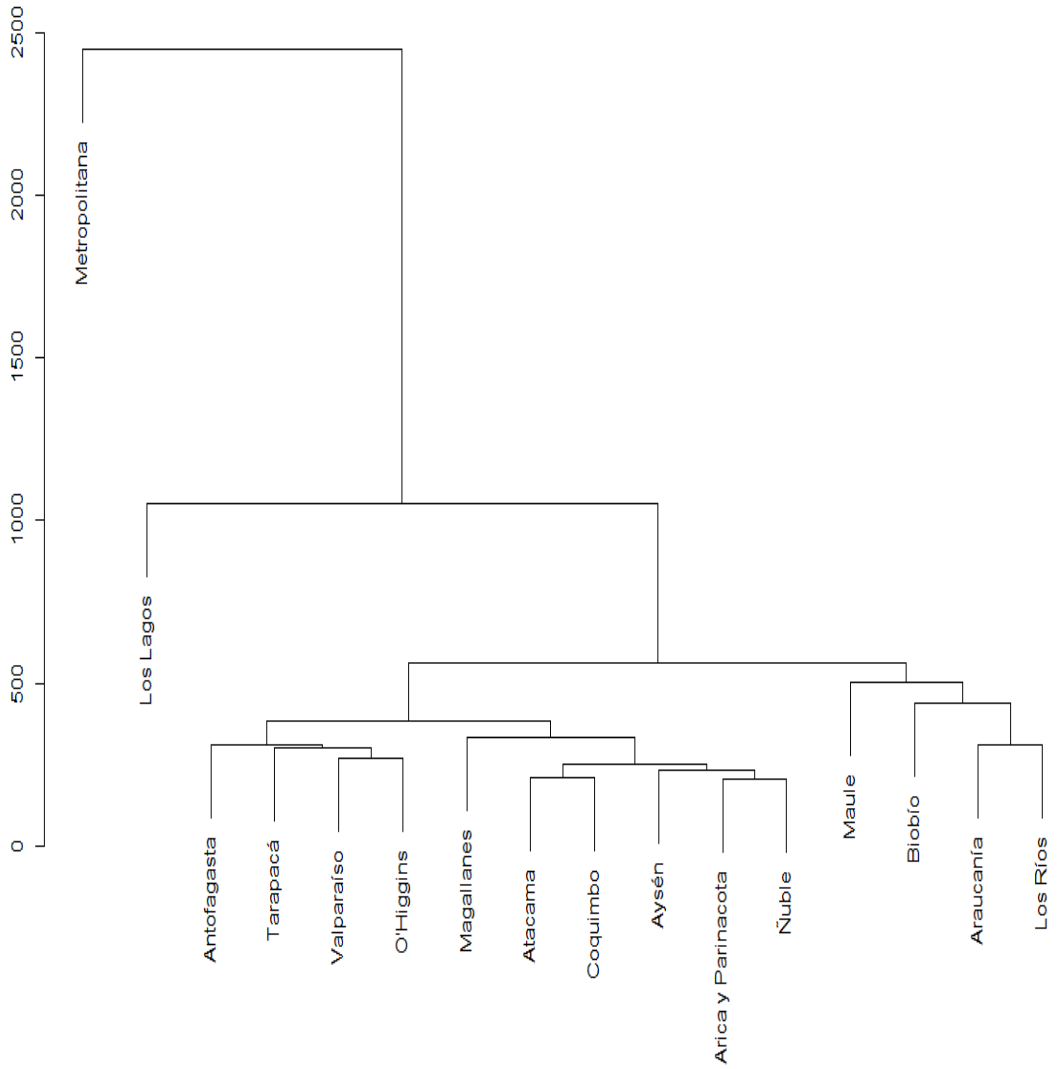

Figura 3.12: *Cluster* de la variable Pacientes Nuevos sin Síntomas de COVID-19.
Fuente : elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

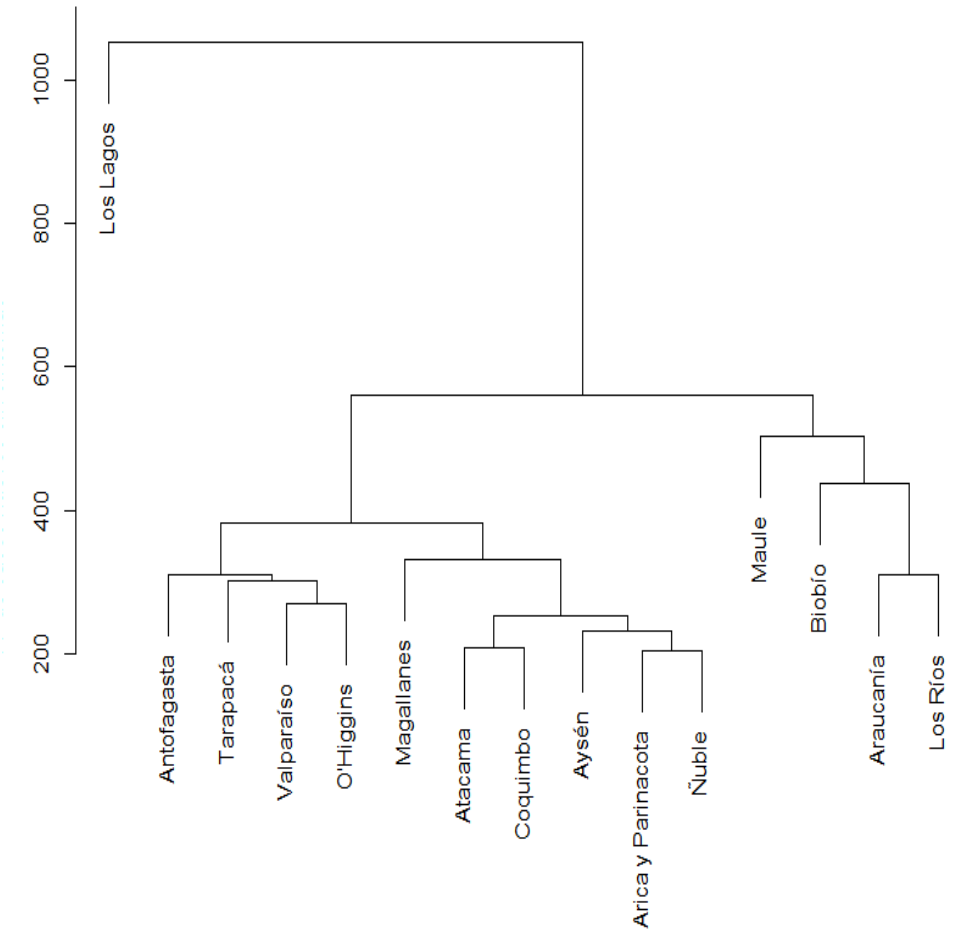

Figura 3.13: *Cluster* de la variable Pacientes Nuevos sin Síntomas de COVID-19 sin considerar la Región Metropolitana.

Fuente : elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

La Región Metropolitana con respecto a la variable Casos Nuevos sin Síntomas COVID-19 mantuvo el mayor número de casos registrando, alrededor de 2.500 casos asintomáticos junto con la Región de Los Lagos que presentó un comportamiento considerablemente tambien diferente a las demás regiones registrando cerca de los 1000 casos asintomáticos (Ver Figura 3.12). El resto de las regiones no superan los 600 casos. Se pueden observar al menos tres grupos, entre ellos la agrupación que contiene a las Regiones Aysén, Ñuble y Arica y Parinacota, es decir que las regiones antes mencionadas presentaron un comportamiento similar con respecto a la variable en estudio en el período de tiempo analizado (Ver Figura 3.13).

Análisis para la Región de Valparaíso

Las semanas epidemiológicas (SE) consideradas para el análisis de la Región de Valparaíso se observa en el Cuadro 3.4.

height Semana	Correspondiente a la S.E	Primer día	Último día
1	23-2020	31-05-2020	06-06-2020.
2	24-2020	07-06-2020	13-06-2020.
3	25-2020	14-06-2020	20-06-2020.
4	26-2020	21-06-2020	27-06-2020.
5	27-2020	28-06-2020	04-07-2020.
6	28-2020	05-07-2020	11-07-2020.
7	29-2020	12-07-2020	18-07-2020.
8	30-2020	19-07-2020	25-07-2020.
9	31-2020	26-07-2020	01-08-2020.
10	32-2020	02-08-2020	08-08-2020.
11	33-2020	09-08-2020	15-08-2020.
12	34-2020	16-08-2020	22-08-2020.
13	35-2020	23-08-2020	29-08-2020.
14	36-2020	30-08-2020	05-09-2020.
15	37-2020	06-09-2020	12-09-2020.
16	38-2020	13-09-2020	19-09-2020.
17	39-2020	20-09-2020	26-09-2020.
18	40-2020	27-09-2020	03-10-2020.
19	41-2020	04-10-2020	10-10-2020.
20	42-2020	11-10-2020	17-10-2020.
21	43-2020	18-10-2020	24-10-2020.
22	44-2020	25-10-2020	31-10-2020.
23	45-2020	01-11-2020	07-11-2020.
24	46-2020	08-11-2020	14-11-2020.
25	47-2020	15-11-2020	21-11-2020.
26	48-2020	22-11-2020	28-11-2020.
27	49-2020	29-11-2020	05-12-2020.
28	50-2020	06-12-2020	12-12-2020.
29	51-2020	13-12-2020	19-12-2020.
30	52-2020	20-12-2020	26-12-2020.
31	01-2021	27-12-2020	02-01-2021.
32	02-2021	03-01-2021	09-01-2021.
33	03-2021	10-01-2021	16-01-2021.
34	04-2021	17-01-2021	23-01-2021.
35	05-2021	24-01-2021	30-01-2021.
36	06-2021	31-01-2021	06-02-2021.
37	07-2021	07-02-2021	13-02-2021.
38	08-2021	14-02-2021	20-02-2021.
39	09-2021	21-02-2021	27-02-2021.
40	10-2021	28-02-2021	06-03-2021.
41	11-2021	07-03-2021	13-03-2021.
42	12-2021	14-03-2021	20-03-2021.
43	13-2021	21-03-2021	27-03-2021.
44	14-2021	28-03-2021	03-04-2021.
45	15-2021	04-04-2021	10-04-2021.
46	16-2021	11-04-2021	17-04-2021.
47	17-2021	18-04-2021	24-04-2021.
48	18-2021	25-04-2021	01-05-2021.
49	19-2021	02-05-2021	08-08-2021.
50	20-2021	09-05-2021	15-05-2021.
51	21-2021	16-05-2021	22-05-2021.
52	22-2021	23-05-2021	29-05-2021.
53	23-2021	30-05-2021	05-06-2021.
54	24-2021	06-06-2021	12-06-2021.
55	25-2021	13-06-2021	19-06-2021.
56	26-2021	20-06-2021	26-06-2021.
57	27-2021	27-06-2021	03-07-2021.

Cuadro 3.4: Semanas epidemiológicas.
Fuente: elaboración propia.

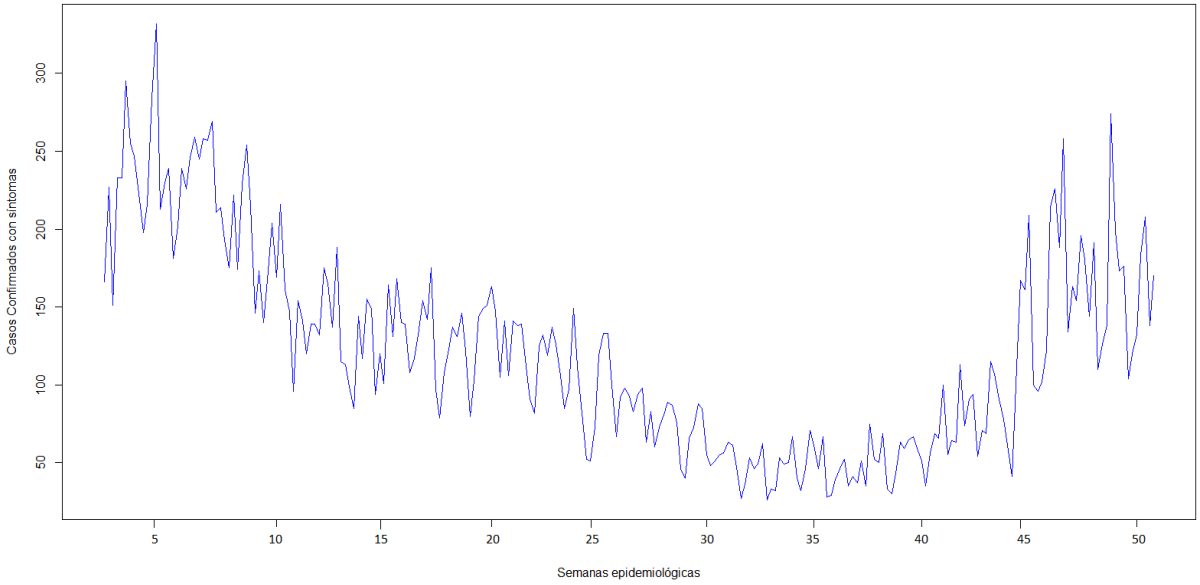

Figura 3.14: Serie de tiempo de los Casos Confirmados con síntomas COVID-19 por SE en la Región de Valparaíso, Chile.

Fuente: elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

En la SE 5, 28-06-2020 al 04-07-2020, la Región de Valparaíso presentó el pick más alto, registrando alrededor de 300 casos confirmados. Ninguna de las comunas de la región de Valparaíso se encontraba con restricción, puesto que para esa fecha en particular aún no se implementaban las fases del conocido "Plan Paso a Paso". En la semana 9, 26-07-2020 al 01-08-2020, exactamente el 28-07-2020 se implementaron las primeras fases en las comunas de la Región de Valparaíso, como se puede observar en la Figura 3.14, en la semana 9 el se presenta una tendencia al decrecimiento, debido a las medidas de intervención. En la semana 31, 27-12-2020 al 02-01-2021, los casos disminuyeron más de lo habitual registrando no más de 100 casos, pues la tendencia duró 10 semanas. En la semana epidemiológica 44 los casos fueron aumentando significativamente, registrándose alrededor de 250 casos confirmados con síntomas COVID-19 en la semana epidemiológica 49.

En el presente análisis se consideraron 36 comunas para la Región de Valparaíso, no se registraron las cifras de las Islas de Pascua y Juan Fernández, debido a la frontera natural (Oceano Pacífico) que las separa del resto de las comunas. Todas las comunas presentaron comportamientos totalmente diferentes con respecto a la variable Casos Nuevos Confirmados con COVID-19 (por semana epidemiológica) como se aprecia en las figuras 3.15, 3.16, 3.17 , 3.18 y 3.19 :

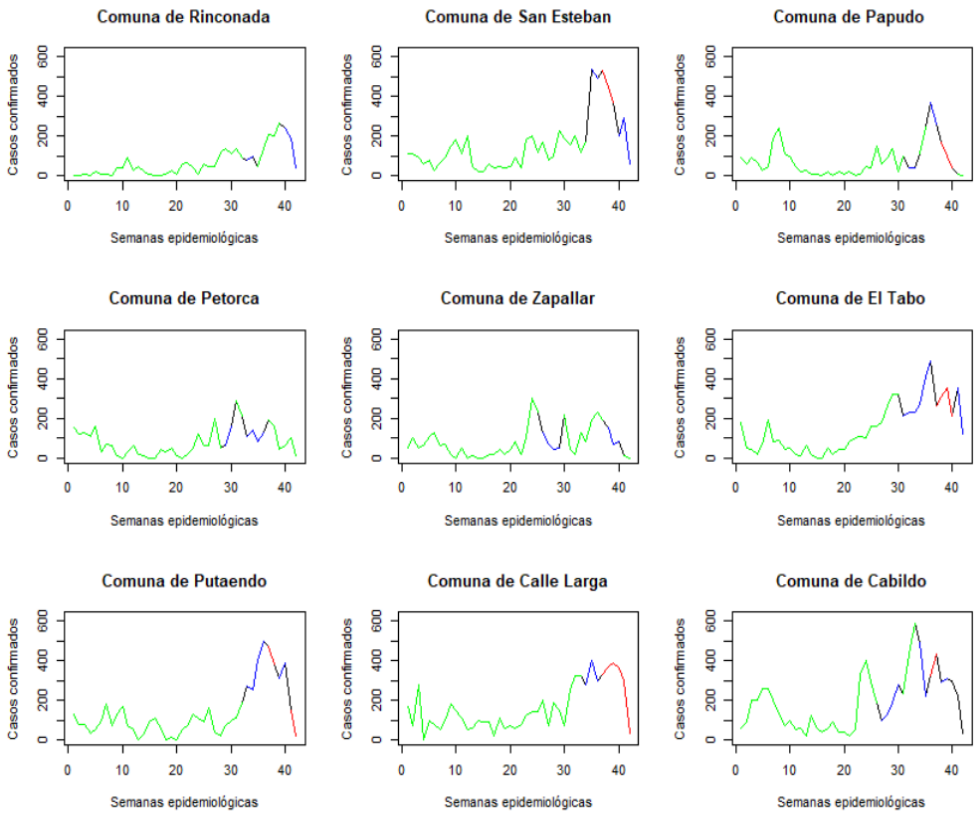

Figura 3.15: Casos Confirmados con COVID-19 por SE en las comunas de la Región de Valparaíso con respecto a la medidas de intervención "Paso a Paso". Fase 1 (Rojo), Fase 2 (azul), Fase 3 (verde).

Fuente: elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

En la Figura 3.15, todas las comunas comienzan con la medida de intervención en Fase 3, es decir que todas ellas al inicio del estudio no se encontraban con ninguna restricción con respecto a las salidas fuera de los hogares de residencia. Las comunas de Rinconada, Petorca y Zapallar fueron las únicas que durante el período de estudio solo estuvieron en dos Fases, 2 y 3, mientras que las demás se encontraban en alguna semana epidemiológica en la Fase 1, además aquellas 3 comunas no superan los 300 Casos Confirmados de COVID-19. Por otro lado la Comuna de Cabildo fue una de las que presentó mayor número de casos confirmados en la SE 33, mientras que la segunda comuna con mayor número de casos es la Comuna de San Esteban, con cerca de 500 casos. Las comunas de El Tabo, Putaendo y Calle Larga entre las SE 1 y la SE 33 se encontraban en Fase 3, como lo casos iban aumentando a excepción de la Comuna de Petorca fue la que registró una caída de sus casos, estas cambiaron a Fase 2, esta no fue una medida efectiva ya que el número de casos aumentó entre las SE 30 y 35, situación que era de esperar, ya que la medida de intervención catalogada como Fase 2 solo corresponde a cuarentena los fines de semana.

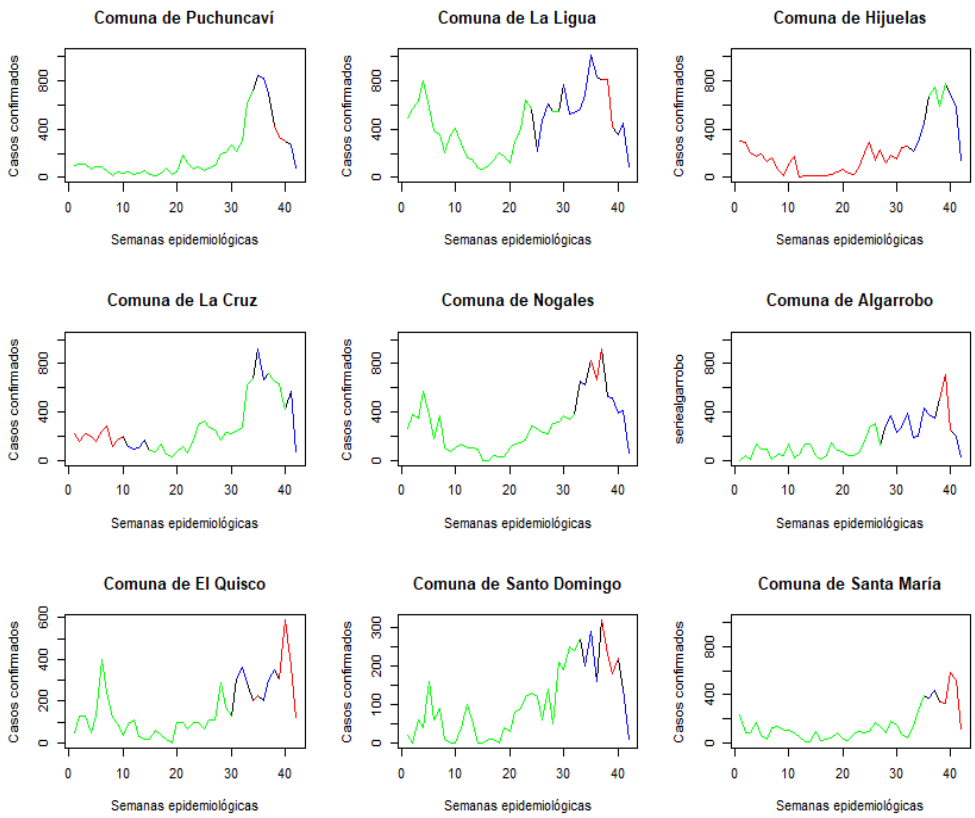

Figura 3.16: Casos Confirmados con COVID-19 por semana epidemiológica en las comunas de la Región de Valparaíso con respecto a la medidas de intervención "Paso a Paso". Fase 1 (Rojo), Fase 2 (azul), Fase 3 (verde).

Fuente: elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

En la Figura 3.16 se puede observar el comportamiento de las comunas de la Región de Valparaíso que presentaron mas de 500 casos, la comuna que está más cercano a los 500 Casos Confirmados es la Comuna de Santa María, esta comuna desde la SE 1 a la 36 se encontraba en Fase 3, como los casos aumentaron desde la SE 31 a la 36, cambió de Fase 3 a Fase 2, los casos aumentaron un par de cifras, sin embargo disminuyeron en SE 38, siendo la última decisión del gobierno de Chile cambiar de Fase 2 a Fase 1 en la SE 39, dado el aumento significato de los casos, llegando a su pick de alrededor de 500 casos, sin embargo desde la SE 40 los casos disminuyeron rápidamente. La mayoría de las comunas que se presentan en la Figura 3.16 comienzan en Fase 3, a excepción de Hijuelas y La Cruz que comenzaron en Fase 2. Sin embargo en el caso de la Comuna de Hijuelas, luego de estar en Fase 1, es decir una cuarentena total, el MINSAL decidió cambiar de Fase 1 a Fase 2, medida de intervención que no fue eficaz, puesto que los casos aumentaron rápidamente en el transcurso de las semanas epidemiológicas, luego la comuna volvió a cambiar de Fase 2 a Fase 3, presentando un pick de alrededor de 600 casos, mucho mayor que cuando esta comuna estaba en Fase 1.

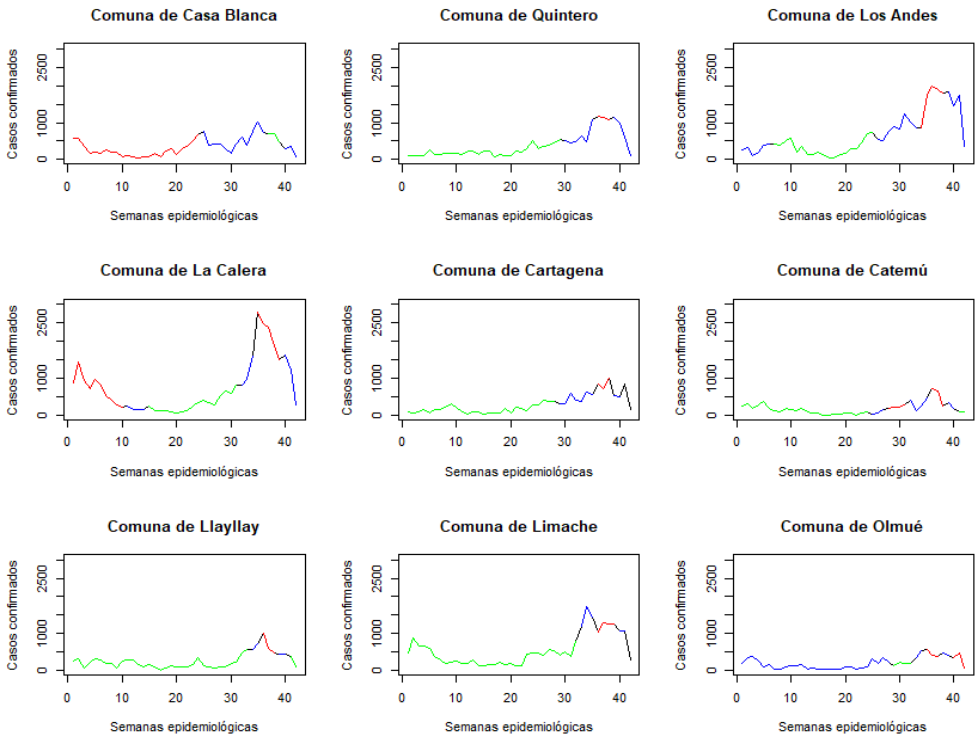

Figura 3.17: Casos Confirmados con COVID-19 por SE en las comunas de la Región de Valparaíso con respecto a la medidas de intervención "Paso a Paso". Fase 1 (Rojo), Fase 2 (azul), Fase 3 (verde).

Fuente: elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

En la Figura 3.17 se presentan las comunas con un número de Casos Confirmados de COVID-19 superior a 500. La Comuna de Olmué llama bastante la atención, pues fue la única que comenzó su comportamiento en Fase 2 teniendo menos de 500 casos, luego cambia a Fase 3 en la SE 29 pero en la SE 31 los casos aumentaron llegando a su pick cerca de los 550 casos, retrocediendo nuevamente a Fase 2, hasta declararse en Fase 1 (al final del periodo de estudio, aunque presenta una tendencia al decreciemiento). Las Comunas de Casa Blanca y La Calera comenzaron en Fase 1, Casa Blanca desde la SE 1 a la 25, mientras que La Calera desde la SE 1 a la 10. Mientras que las Comunas Los Andes, Cartagena, Catemú, Llayllay y Limache comenzaron en Fase 3. La Comuna de La Calera fue la comuna que presentó mayor número de casos, situando un pick sobre los 2.000 casos, esta comuna en particular al inicio del período de estudio comenzó en Fase 1, puesto que presentaba un alto número de casos, cercano a los 1.500, desde la SE 10 a la 15 la comuna cambió a Fase 2, pues los casos se mantuvieron hasta la SE 25, encontrándose en Fase 3, los casos fueron aumentando, este comportamiento era de esperar, ya que en la Fase 3 no hay restricciones, como los casos aumentaron significativamente el MINSAL cambio de Fase 3 a Fase 2, esta medida no sirvió de mucho, los casos aumentaron mucho más, hasta que se decidió el nuevo cambio a la Fase 1, disminuyendo así los casos rápidamente.

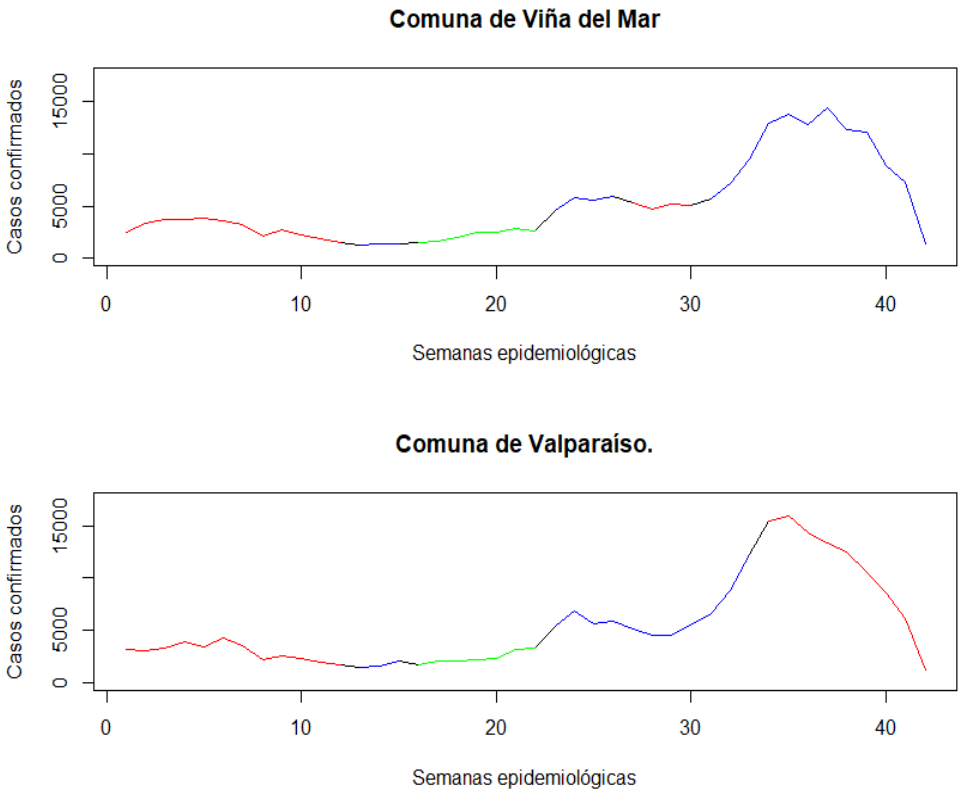

Figura 3.18: Casos nuevos confirmados con COVID-19 por SE en las comunas de la Región de Valparaíso con respecto a la medidas de intervención "Paso a Paso". Fase 1 (Rojo), Fase 2 (azul), Fase 3 (verde).

Fuente: elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

Se puede observar (Ver Figura 3.18) que las comunas hermanas Valparaíso y Viña del Mar presentaron tendencias similares a lo largo de las semanas epidemiológicas. Sin embargo la implementación de las fases fue un poco diferente en algunas semanas epidemiológicas, pues la Comuna de Viña del Mar comenzó con un alto número de casos situándose cerca de los 5.000 un número bastante alto para estar en Fase 1, luego desde la SE 13 a la SE 16, la comuna cambió a Fase 2 y desde la SE 17 a la SE 24 a Fase 3, con un comportamiento fue casi constante hasta la SE 22 cuando los casos aumentaron considerablemente, por esto nuevamente cambió a Fase 2 y luego nuevamente a Fase 1, cuyo periodo fue muy corto hasta el nuevo cambio a Fase 2 con un aumento significativo llegando a su pick de casi 15.000 Casos Confirmados. Por otra parte, en la Comuna de Valparaíso se presentaron las mismas fases que la Comuna de Viña del Mar, a excepción de las SE 23 y 34, donde Valparaíso solo se mantuvo en Fase 2, terminando el período de estudio en Fase 1, los casos disminuyeron bastante, pues la Comuna de Valparaíso registraba un pick sobre los 15.000. Cabe mencionar que estas comunas son las que presentaron el mayor número de Casos Confirmados en la Región de Valparaíso, esto se debe a la alta densidad poblacional y flujo de personas desde por ejemplo, la Región Metropolitana por su cercanía.

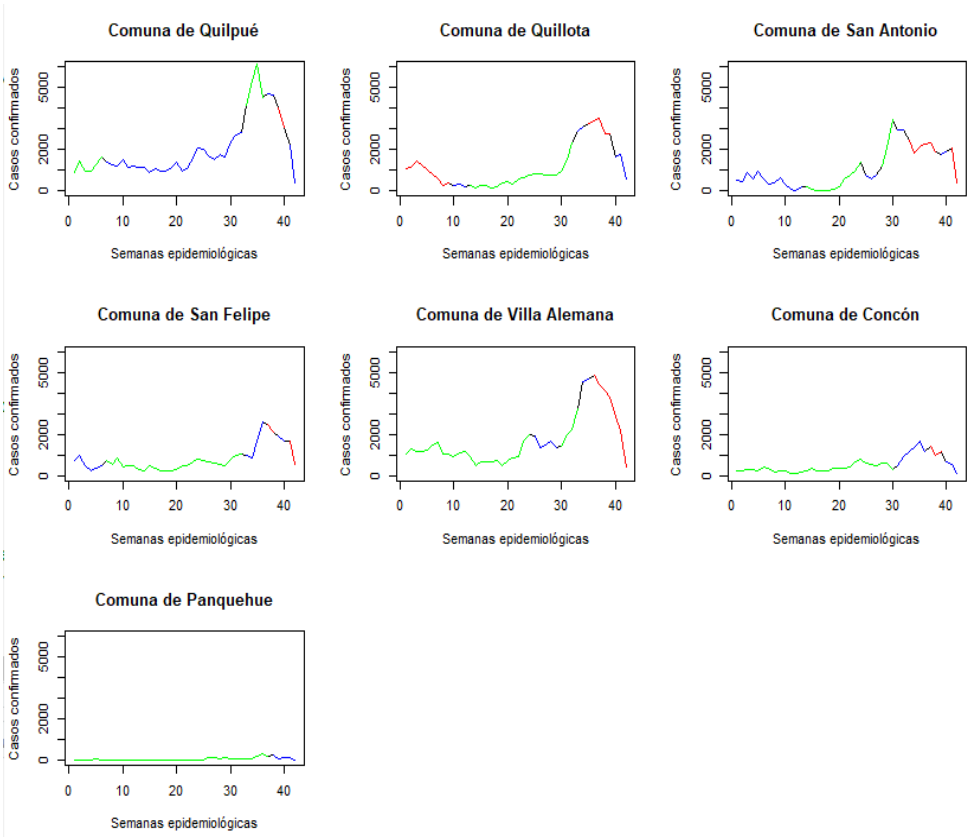

Figura 3.19: Casos nuevos confirmados con COVID-19 por SE en las comunas de la Región de Valparaíso con respecto a la medidas de intervención "Paso a Paso". Fase 1 (Rojo), Fase 2 (azul), Fase 3 (verde).

Fuente: elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

Las comunas presentadas en la Figura 3.19 tienen un número de Casos Confirmados de COVID-19 sobre los 3.000, a excepción de las comunas Concón y Panquehue, pues la Comuna de Panqehue fue la que registró el menor número de casos confirmados en este grupo de comunas de la Región de Valparaíso, situando como máximo cerca de 300 contagios. Por otro lado, la Comuna de Concón registraba en la SE 34 un pick de cerca de 2.000 contagiados encontrándose en Fase 2. La comuna que presentó mayor número de contagiados en este grupo fue la Comuna de Quilpué, pues si se observa con detalle al inicio del período de estudio la comuna estuvo en Fase 3, luego la comuna cambió a Fase 2 donde en las SE 22 y 32 hubo un alza de los casos situándose cerca de los 4.000 casos confirmados con COVID-19, cambiando a Fase 3, decisión que fue inapropiada pues lo casos aumentaron mucho más, llegando a su pick sobre los 6.000 contagiados, cambiando a Fase 2, luego a Fase 1 y finalmente a Fase 2, restricción acertada ya que los casos disminuyeron rápidamente. La segunda comuna que presentó mayor número de casos confirmados fue la Comuna de Villa Alemana, esta es una de las comunas con mayor población situanda alrededor de los 300.000 habitantes, presentando un pick de casi 5.000 contagiados.

3.2.2. Análisis espacial

Las semanas epidemiológicas (SE) consideradas para el análisis espacial se presentan en el Cuadro 3.5, estas semanas fueron determinadas aleatoriamente a través de software y algunas semanas epidemológicas elegidas estratégicamente ya que contienen fechas con festividades, estas corresponden a una semana epidemiológica por mes, desde Junio 2020 a Junio 2021, siendo 12 semanas epidemiológicas en total.

S.E	Abreviación	Primer día	último día
25/2020	2520	14/06/2020	20/06/2020 .
30/2020	3020	19/07/2020	25/07/2020.
33/2020	3320	09/08/2020	15/08/2020.
37/2020	3720	06/09/2020	12/09/2020.
42/2020	4220	11/10/2020	17/10/2020.
48/2020	4820	22/11/2020	28/11/2020.
52/2020	5220	20/12/2020	26/12/2020.
02/2021	0221	10/01/2021	16/01/2021.
06/2021	0621	07/02/2021	13/02/2021.
11/2021	1121	14/03/2021	20/03/2021.
15/2021	1521	11/04/2021	17/04/2021.
20/2021	2021	16/05/2021	22/05/2021.
23/2021	2321	06/06/2021	12/06/2021.

Cuadro 3.5: Semanas epidemiológicas.
Fuente: elaboración propia.

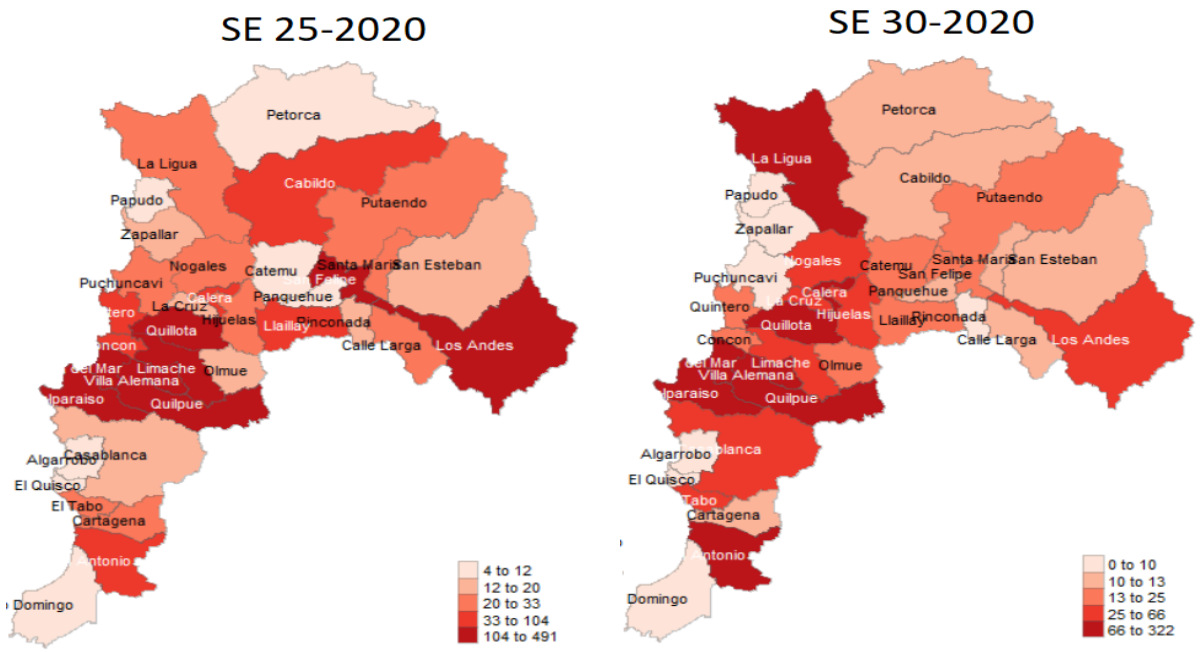

Figura 3.20: Casos Confirmados con COVID-19 por SE en las comunas de la Región de Valparaíso.
Fuente : elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

Al inicio del período de estudio en la SE 23-2020 las comunas de Valparaíso, Viña del Mar, Villa Alemana, Quilpué, Limache y Quillota presentaron un alto número de Casos Confirmados, entre los 104 y 491 casos respectivamente (Ver Figura 3.20), entre estas, Valparaíso y Viña del Mar presentan mayor densidad poblacional, con alrededor de 400.000 habitantes, mientras que las demás comunas presentan una cantidad de habitantes entre los 200.000 y 300.000. En la SE 23-2020 las comunas se encontraban sin ninguna medida de intervención impuesta, puesto que las fases correspondientes al "Plan Paso a Paso" se implementaron en la SE 31-2020, específicamente el 28 de Julio del año 2020. Si se observan con mayor detalle las Comunas de San Felipe y Los Andes, estas también presentan un alto número de casos con una población similar a las comunas de Limache y Quillota, lo que llama la atención es lo lejanas que se encuentran del resto de as comunas que presentan el mismo comportamiento, pues se espera que comunas que compartan alguna frontera o límite presenten comportamientos similares con respecto a la variable en estudio. En la SE 30-2020 las mismas comunas de la SE 23-2020 siguen presentándo un alto número de Casos Confirmados, a excepción de Limache, Los Andes y San Felipe, es decir que durante 7 semanas transcurridas las comunas de Valparaíso, Viña del Mar, Quilpué, Limache y Quillota presentaron un alto número de contagiados, sumándose al mismo grupo las Comunas de San Antonio y La Ligua, registrando entre 66 y 322 casos respectivamente.

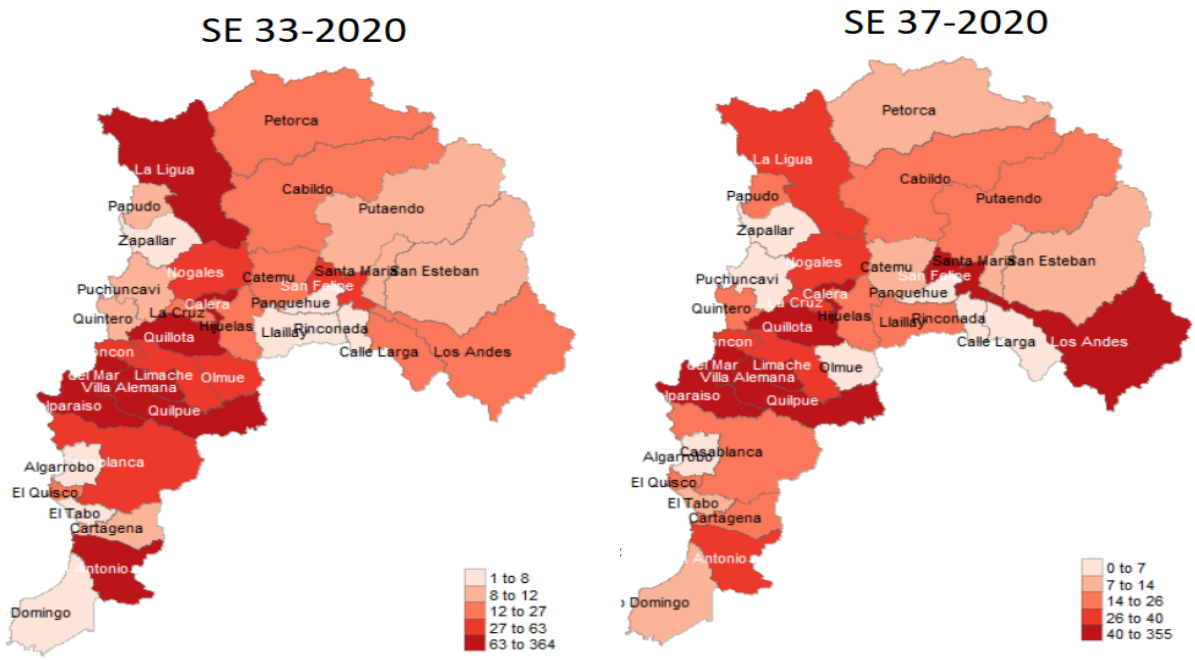

Figura 3.21: Casos Confirmados con COVID-19 por SE en las comunas de la Región de Valparaíso.
Fuente : elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

Luego de transcurrir exactamente tres semanas, las mismas comunas con un alto número de contagiados en la SE 30-2020 , Valparaíso, Viña del Mar, Quilpué, Villa de Alemana, Quillota, La Ligua y San Antonio (Ver Figura 3.20), siguen siendo las mismas comunas con un gran número de casos en la SE 33-2020 (Ver Figura 3.21), es decir que durante 3 semanas epidemiológicas no disminuyeron las cifras, por el contrario aumentaron, situándose entre los 63 y 364 contagiados. Las comunas con mayor cantidad de casos se encontraban en aquel período de tiempo ya con las medidas de intervención implementadas, entre ellas las comunas Quillota, La Ligua, Valparaíso y Viña del Mar estaban en Fase 1, es decir con una cuarentena total, por otro lado las Comunas Quilpué y Villa Alemana se encontraban en Fase 2, solo cuarentena los fin de semanas (sábados y domingos) mientras que la Comuna de San Antonio estaba en Fase 3, sin cuarentena, lo que llama la atención puesto que es una comuna con alta densidad poblacional. La medida implementada para esta comuna no fue la más acertada, puesto que registraron alrededor de 364 contagiados. En la SE 37-2020 la comuna de San Antonio disminuyó al menos 100 casos estando en Fase 2, medida de intervención implementada bastante bien, puesto que durante 4 semanas epidemiológicas no disminuía el número de casos. Las comunas que volvieron a presentar un alto número de casos fueron Los Andes y San Felipe con alrededor de 355 contagiados, esto era de esperar puesto que las comunas en la SE 37-2020 se encontraban en Fase 3, es decir sin cuarentena.

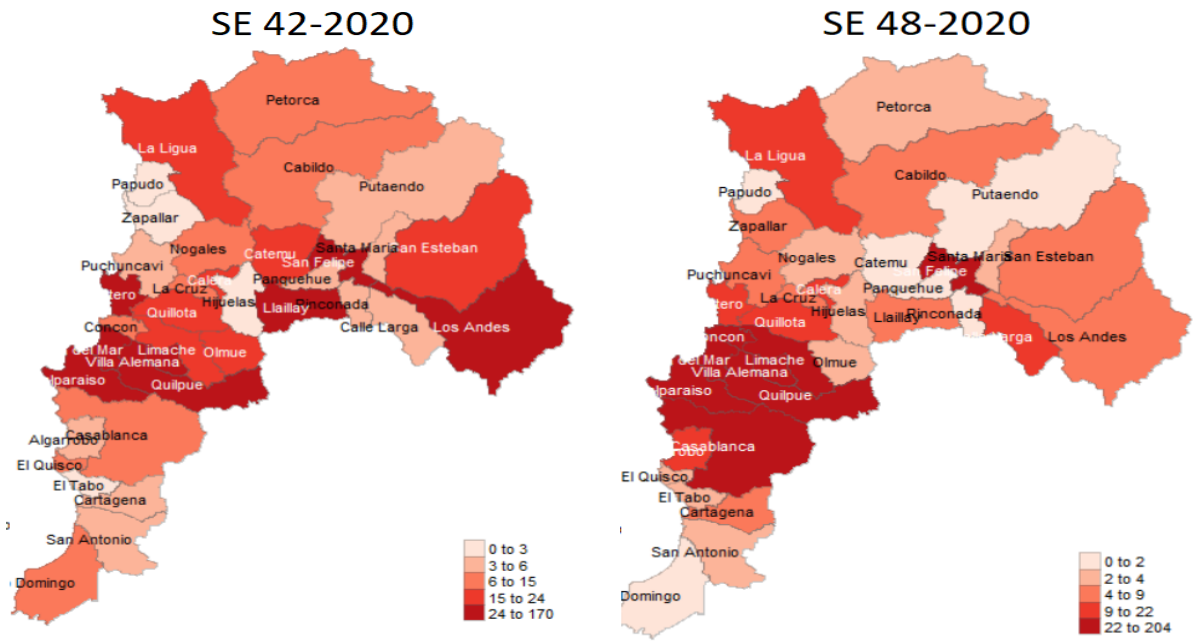

Figura 3.22: Casos Confirmados con COVID-19 por SE en las comunas de la Región de Valparaíso.
Fuente : elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

La Comuna de Quintero junto con la de Llayllay en la SE 42-2020 fueron las únicas que se agregaron al grupo de los casos altos desde la SE 37-2020 (Ver Figura 3.21), ambas comunas presentan una población que no superan los 200.000 habitantes, y en aquella semana epidemiológica se encontraban en Fase 3, es decir no presentaban ninguna restricción para la salida fuera del hogar de los residentes. A través de la medida de intervención implementada las comunas aumentaron desde 14 casos (en la SE 37-2020 Figura 3.21) a 170 contagiados en la SE 42-2020 (Ver Figura 3.22). En la misma semana epidemiológica la comuna de San Antonio disminuyó significativamente sus casos luego de cuatro semanas transcurridas situándose entre 3 a 6 contagiados, encontrándose respectivamente en Fase 2. Si se observa la SE 48-2020 la Comuna de San Felipe se encuentra muy alejada de las demás comunas con un alto número de contagiados entre ellas: Valparaíso, Viña del Mar, Quilpué, Villa Alemana, Limache, Concón y Casa Blanca. Todas las comunas con un alto número de casos confirmados se encontraban en Fase 2, es decir presentaban cuarentena sólo los fines de semanas (sábados y domingos), a excepción de la Comuna de Limache que estaba en Fase 3. Cabe destacar que las comunas con un alto número de casos desde la SE 42-2020 hasta las 48-2020 aumentaron sus casos alrededor de 30 contagiados más.

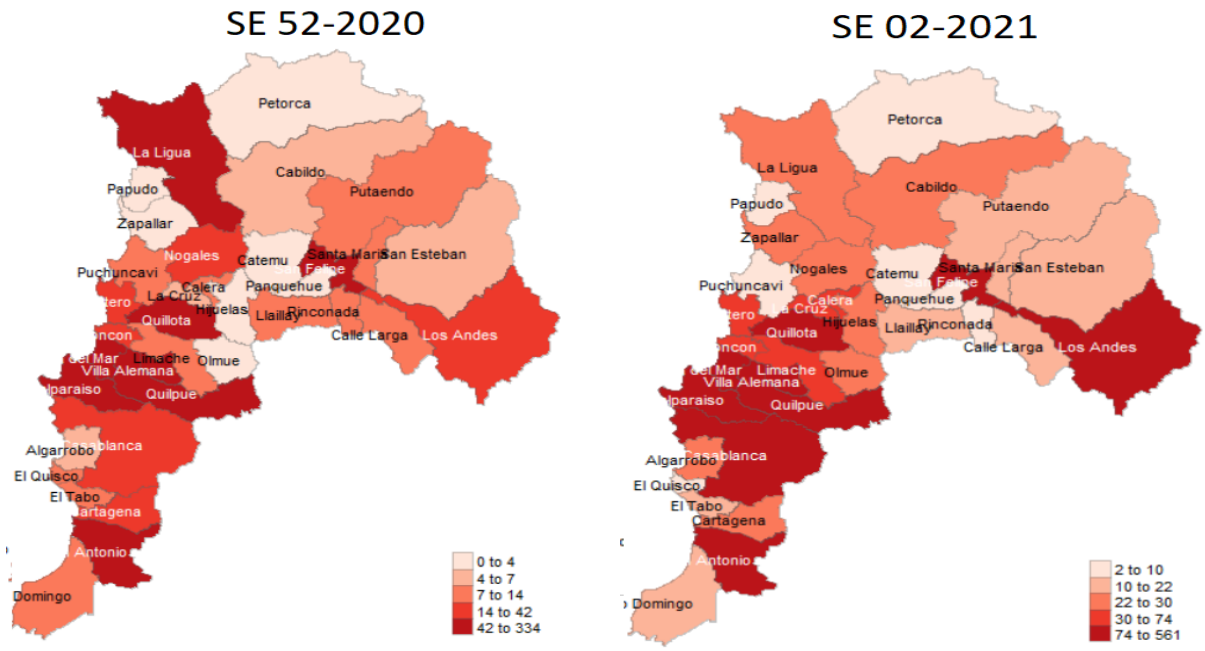

Figura 3.23: Casos Confirmados con COVID-19 por SE en las comunas de la Región de Valparaíso.
Fuente : elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

A diferencia de la SE 48-2020 (ver Figura 3.22), en la SE 52-2020 las comunas de Casa Blanca y Limache dejaron de estar en el grupo de las comunas con un alto número de casos, disminuyendo sus casos entre los 4 y 14 contagiados (Ver Figura 3.23), ambas comunas se encontraban en Fase 3 en esa semana epidemiológica. Mientras que las comunas que se mantuvieron desde la SE 48-2020 hasta la 52-2020 con un alto número de contagiados fueron: Valparaíso, Viña del Mar, Quilpué y Villa Alemana, situándose en las Fases 2, 2, 3 y 3 respectivamente. Las comunas que se sumaron al grupo de alto número de contagios fueron las comunas de San Antonio, Quillota y La Ligua, estas no presentaban casos altos desde la SE 33-2020 y la 37-2020 (Ver Figura 3.21), registrando alrededor de 334 casos confirmados. La Comuna de La Ligua en la SE 02-2021 dejó de presentar un gran número de contagios registrando entre 22 y 30 casos, alrededor de 300 casos menos desde la SE 52-2020, esta se encontraba en Fase 2, es decir que los residentes podían salir fuera de sus residencias de Lunes a Viernes sin ninguna restricción. Por otro lado, en la misma semana epidemiológica la Comuna de Los Andes nuevamente presentaba un gran número de casos registrando entre 74 y 561 casos, las comunas que acompañaba a Los Andes con un alto número de contagiados son: Valparaíso, San Antonio, Viña del Mar, Quilpué, Villa Alemana y Quillota, todas ellas se encontraban en Fase 2 a excepción de la Comuna de Quillota que se encontraba en Fase 3.

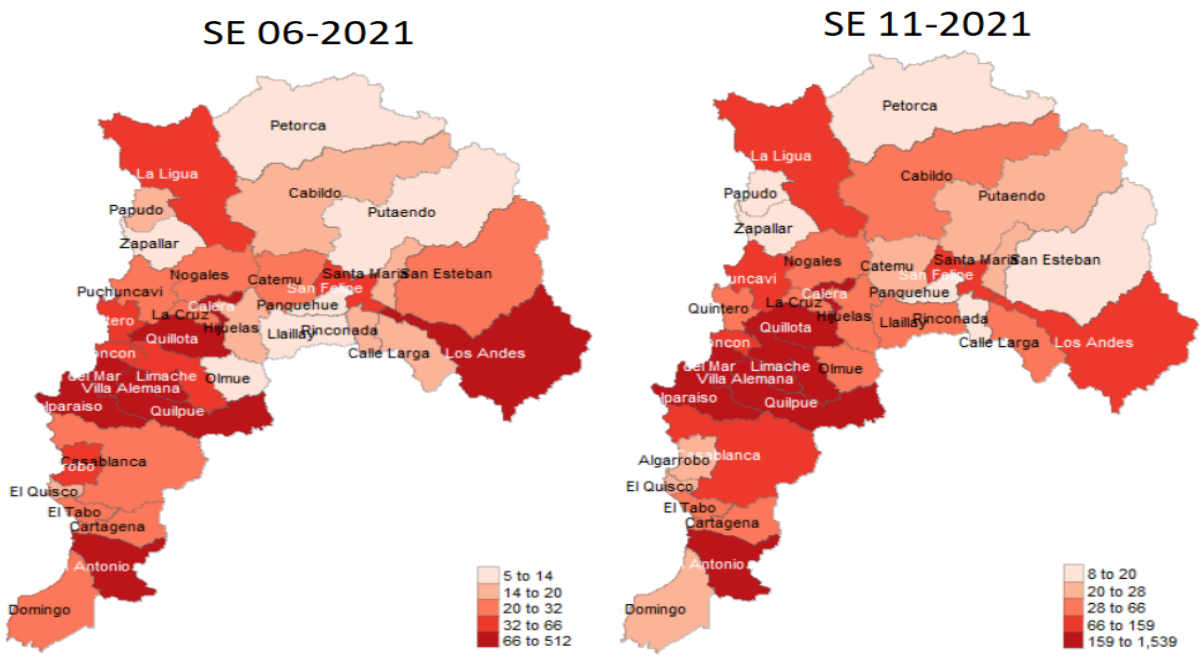

Figura 3.24: Casos Confirmados con COVID-19 por SE en las comunas de la Región de Valparaíso.
Fuente : elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

En la SE 06-2021 las comunas de Valparaíso, Viña del Mar, Quilpué, Villa Alemana, San Antonio, Quillota y Los Andes,disminuyeron sus casos en alrededor de 50 contagios menos desde la SE 02-2021 (Ver Figura 3.23) registrando entre 66 a 512 casos (Ver Figura 3.24), pero se mantiene en el grupo de alto riesgo. Las comunas con mayor número de casos se encontraban en Fase 2 a excepción de la comuna de Quillota, Los Andes y San Antonio quienes se encontraban en Fase 3. La Comuna de San Felipe con cerca de 200.000 habitantes disminuyó sus casos desde la SE 02-2021, registrando entre 32 y 66 contagiados estando en Fase 2, mientras que las comunas que rodean a San Felipe no superan los 31 casos, cabe destacar que las comunas que rodean a San Felipe entre ellas: Catemú, Putaendo, Santa María, Panquehue y Rinconada, todas en Fase 3 al igual que San Felipe, las comunas vecinas a la comuna de San Felipe no superan los 100.000 habitantes. Si se observa la SE 11-2021 la Comuna de Los Andes disminuyó sus casos situándose alrededor de 160 contagios, pues era de esperar puesto que la comuna se encontraba en aquel período de tiempo en Fase 2. No tuvo el mismo comportamiento la Comuna de Casa Blanca, puesto que sus casos aumentaron de 32 a aproximadamente 160, esta situación era de esperarse puesto que Casa Blanca se encontraba en Fase 3 y se considera una comuna relativamente turística.

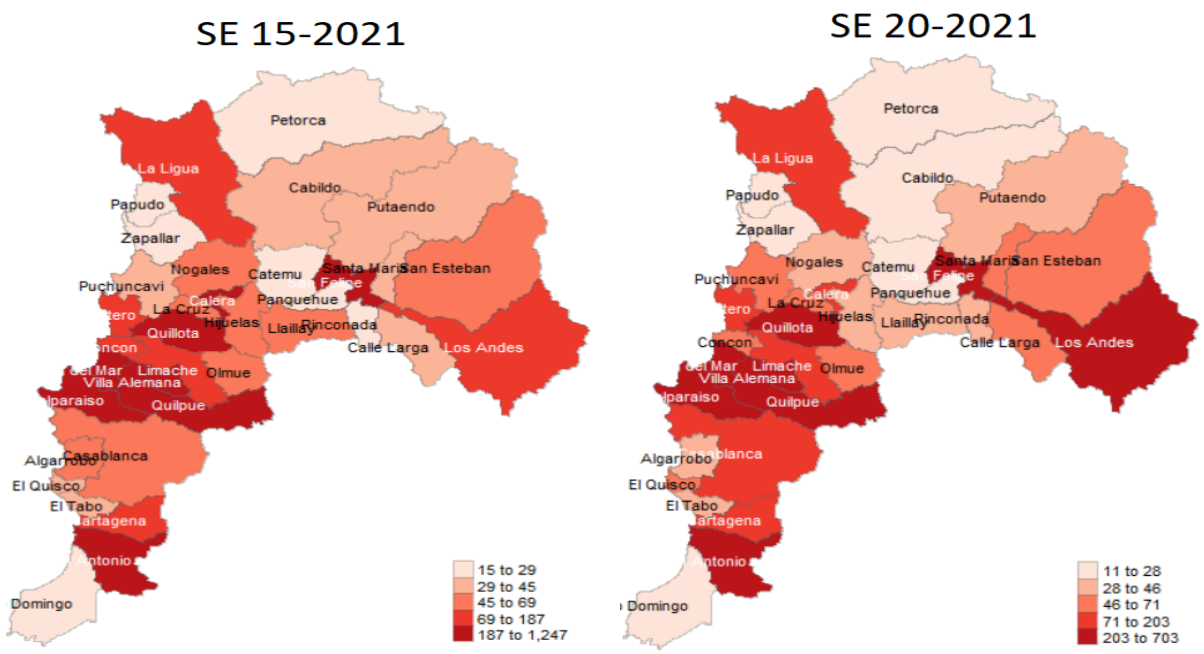

Figura 3.25: Casos Confirmados con COVID-19 por SE en las comunas de la Región de Valparaíso.
Fuente : elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

En la figura 3.25 se observa que nuevamente la Comuna de San Felipe se registra como una de las comunas con mayor número de Casos Confirmados, siendo la SE 15-2021 en la que se presenta mayor número de contagios, entre las comunas con mayor número de casos confirmados además de San Felipe, están Valparaíso, Viña del Mar, Quilpué, Villa Alemana, San Antonio y Quillota, registrando entre 187 y 1.247 casos confirmados, cabe mencionar que en esta semana epidemiológica se terminaba el verano en el territorio nacional. En la SE 20-2021 las mismas comunas de la SE 15-2021 siguen presentándose como comunas con alto número de casos, es decir que durante 5 semanas epidemiológicas consecutivas las comunas con alto número de casos confirmados solo disminuyeron alrededor de 540 casos. Observando con mayor detalle desde la SE 25-2020 (Ver Figura 3.20) en adelante, las comunas Valparaíso, Viña del Mar, Quilpué y Villa Alemana nunca dejaron de presentar alto número de casos, siendo la SE 48-2020 (ver Figura 3.22) el menor registro de casos confirmados dentro del período de estudio analizado. Estas comunas presentan mayor densidad poblacional, registrando entre 300.000 y 400.000 habitantes, además se debe considerar que Valparaíso y Viña del Mar son zonas turísticas, reciben tanto extranjeros como nacionales durante todo el año.

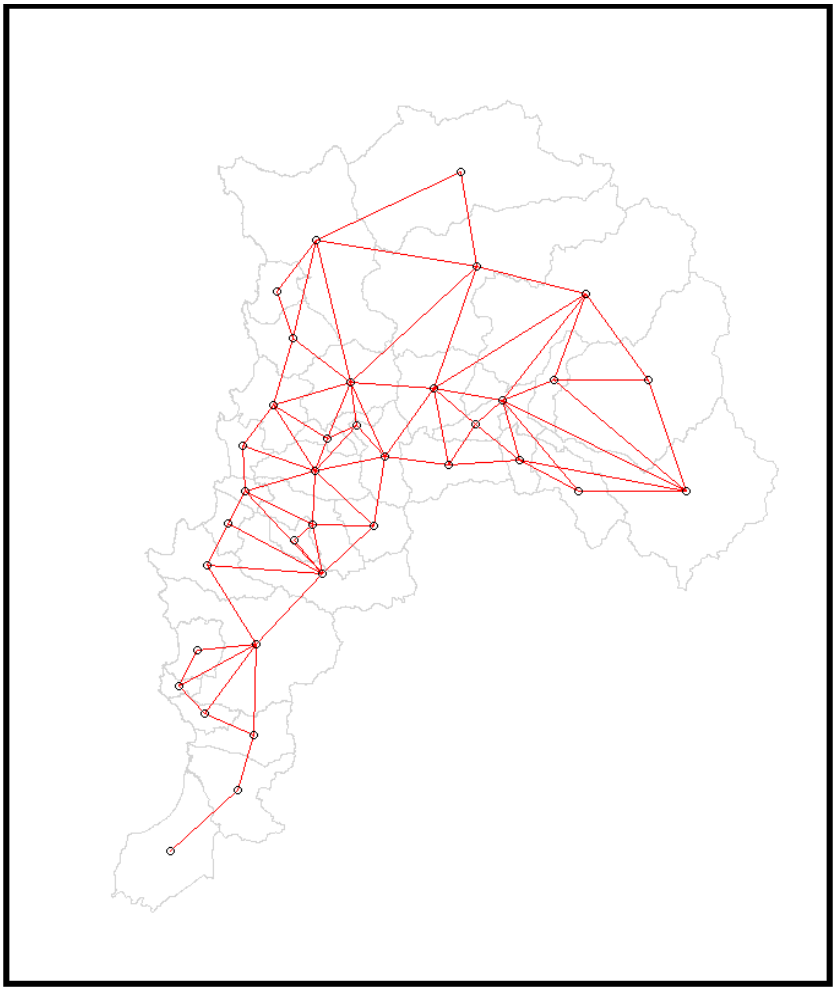

Figura 3.26: Vecinos de las comunas de la Región de Valparaíso.
Fuente: elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

En la Figura 3.26 se pueden observar los vecinos de las comunas de la Región de Valparaíso, en otras palabras, dos comunas serán vecinas si comparten una frontera o borde de longitud no nula. Las líneas rojas corresponden a la red de áreas que trazan la conectividad espacial, es decir que une cada centroide de cada comuna que presenta un borde o frontera en común con otra comuna. A través del cálculo de los vecinos de las comunas de la Región de Valparaíso, se obtiene la matriz de pesos o también conocida como matriz de distancia (Ver sección 2.5), se registraron 148 enlaces y el promedio de estas corresponde a 4,1.

Dependencia espacial: Índice de Morán global

A través del test de Moran global se prueba la dependencia espacial considerando las siguientes hipótesis:

$$H_0 : \text{No existe dependencia espacial.}$$
$$H_1 : \text{Existe dependencia espacial.}$$

El p-valor observado (Ver Cuadro 3.6) es estadísticamente significativo para todas las semanas epidemiológicas, es decir, existe dependencia espacial. Además, a través del estadístico I de Moran se percibe que todos las SE poseen una autocorrelación espacial positiva, en otras palabras, las unidades espaciales vecinas (comunas) presentan valores similares.

SE	I de Morán	p-valor
25-2020	0,40	0,00005
30-2020	0,31	0,0003
33-2020	0,34	0,000005
37-2020	0,43	0,0006
42-2020	0,48	0,00006
48-2020	0,53	0,00018
52-2020	0,41	0,0026
02-2021	0,47	0,00027
06-2021	0,37	0,00025
11-2021	0,45	0,0072
15-2021	0,42	0,0016
20-2021	0,31	0,00006
23-2021	0,33	0,00004

Cuadro 3.6: Test de Morán (global) para los Casos Confirmados con COVID-19 por SE .
Fuente: elaboración propia.

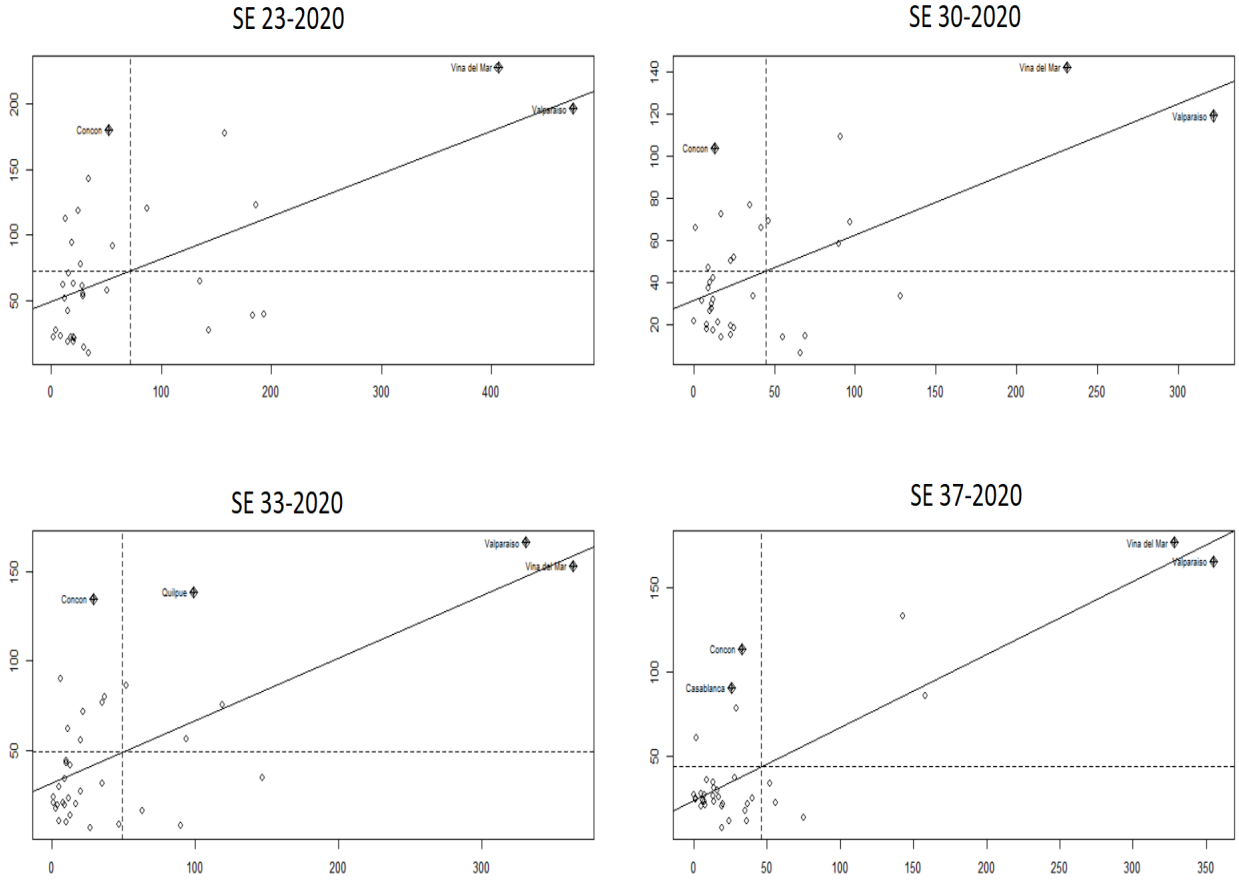

Figura 3.27: Gráfico de Morán (global) de los Casos Confirmados con COVID-19 por SE en las comunas de la Región de Valparaíso.

Fuente: elaboración propia.

Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

Desde la SE 23-2020 a la SE 37-2020 las comunas de Valparaíso, Viña del Mar y Quilpué se mantuvieron en el cuadrante superior derecho en los gráficos de Morán (Ver Figura 3.27), es decir que estas comunas presentaron a lo largo de las semanas epidemiológicas un alto número de Casos Confirmados con COVID-19, teniendo en cuenta que las tres comunas presentan una población entre los 300.000 y 400.000 habitantes. Por otro lado, en las mismas semanas epidemiológicas, pero en el cuadrante superior izquierdo, la Comuna de Concón se mantuvo a lo largo de las semanas epidemiológicas como una comuna de un menor número de Casos Confirmados, rodeada de comunas con alto número de casos, a excepción de la SE 37-2020, quien la acompaña en el mismo cuadrante la Comuna de Casa Blanca. Cabe destacar que entre las SE 23-2020 y la 30-2020, ninguna de las comunas de la Región de Valparaíso se encontraba con alguna medida de intervención, puesto que recién en la SE 31-2020 se implementaron las primeras fases a través del "Plan Paso a Paso". Desde la SE 33-2020 a la 37-2020 las comunas de Quilpué, Concón y Casa Blanca se encontraban en Fase 3, mientras que Valparaíso y Viña del Mar en Fase 1.

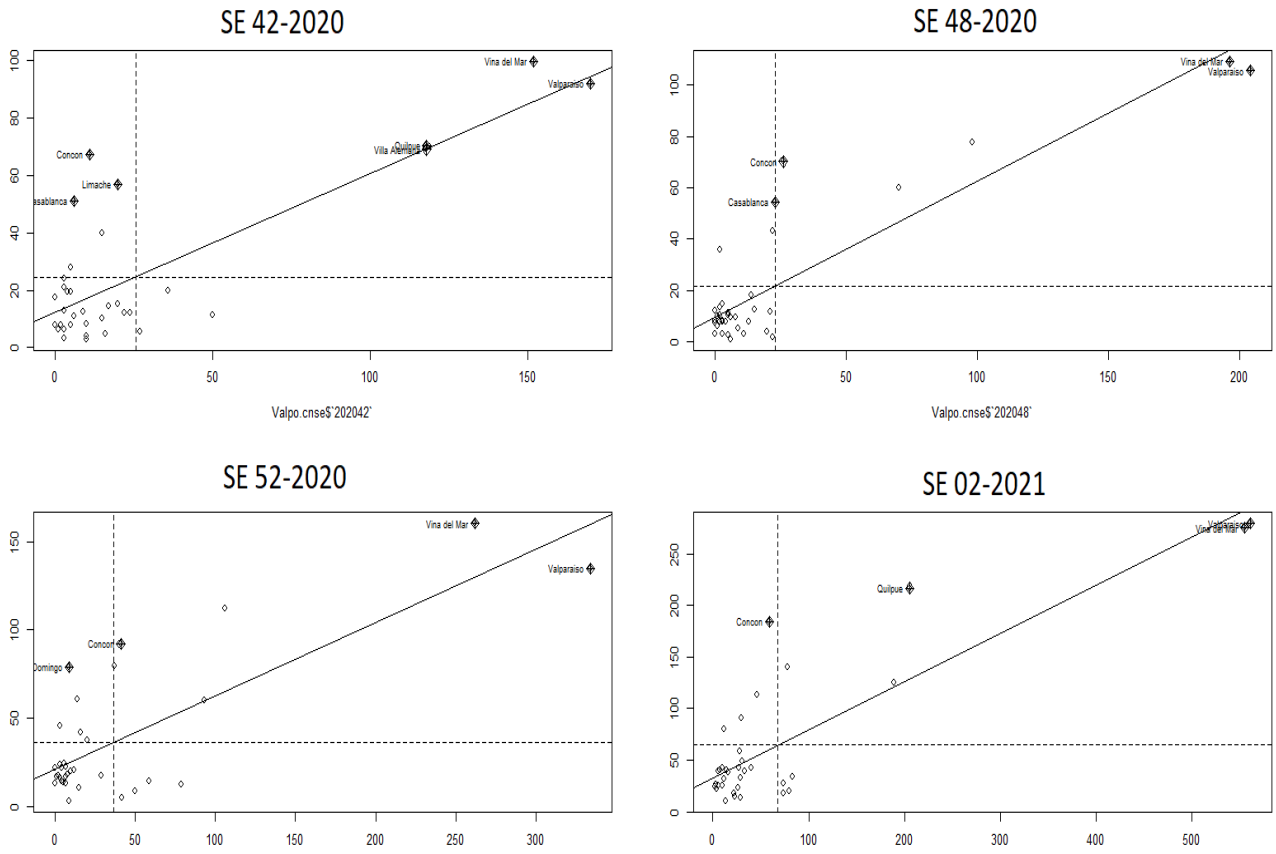

Figura 3.28: Gráfico de Morán (global) de los Casos Confirmados con COVID-19 por SE en las comunas de la Región de Valparaíso.

Fuente: elaboración propia.

Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

En la Figura 3.28 se presenta el Gráfico de Morán para las SE 42-2020, 48-2020, 52-2020 y 02-2021. En la SE 42-2020 en el cuadrante superior derecho las comunas de Viña del Mar, Valparaíso, Quilpué y Villa Alemana se presentaron como comunas con un alto número de Casos Nuevos Confirmados, mientras que en la SE 48-2020 y 52-2020 en el mismo cuadrante se encontraban las mismas comunas a excepción de Quilpué y Villa Alemana. Transcurridas diez semanas la comuna de Quilpué nuevamente se presenta en el cuadrante superior derecho específicamente en la SE 02-2021. En el cuadrante superior izquierdo la Comuna de Concón se mantuvo en aquel cuadrante desde la SE 42-2020 a la 02-2021, mientras que Limache solo se mantuvo en el cuadrante superior izquierdo en la SE 42-2020, Casa Blanca entre las SE 42-2020 y 48-2020, mientras que Santo Domingo solo en la SE 52-2020. Cabe destacar que las Comunas de Valparaíso y Viña del Mar en la SE 42-2020 se encontraban en Fase 2, mientras que desde la SE 48-2020 a la 52-2020 se encontraban en Fase 3 para luego pasar en la SE 02-2021 a Fase 2. Al ser comunas hermanas la mayoría del tiempo se encontraban en las mismas fases, puesto que presentaban tendencias similares a lo largo del tiempo, sin embargo, en algunas semanas algunas comunas se encontraban en fases distintas cuando se presentaba aumento de contgios.

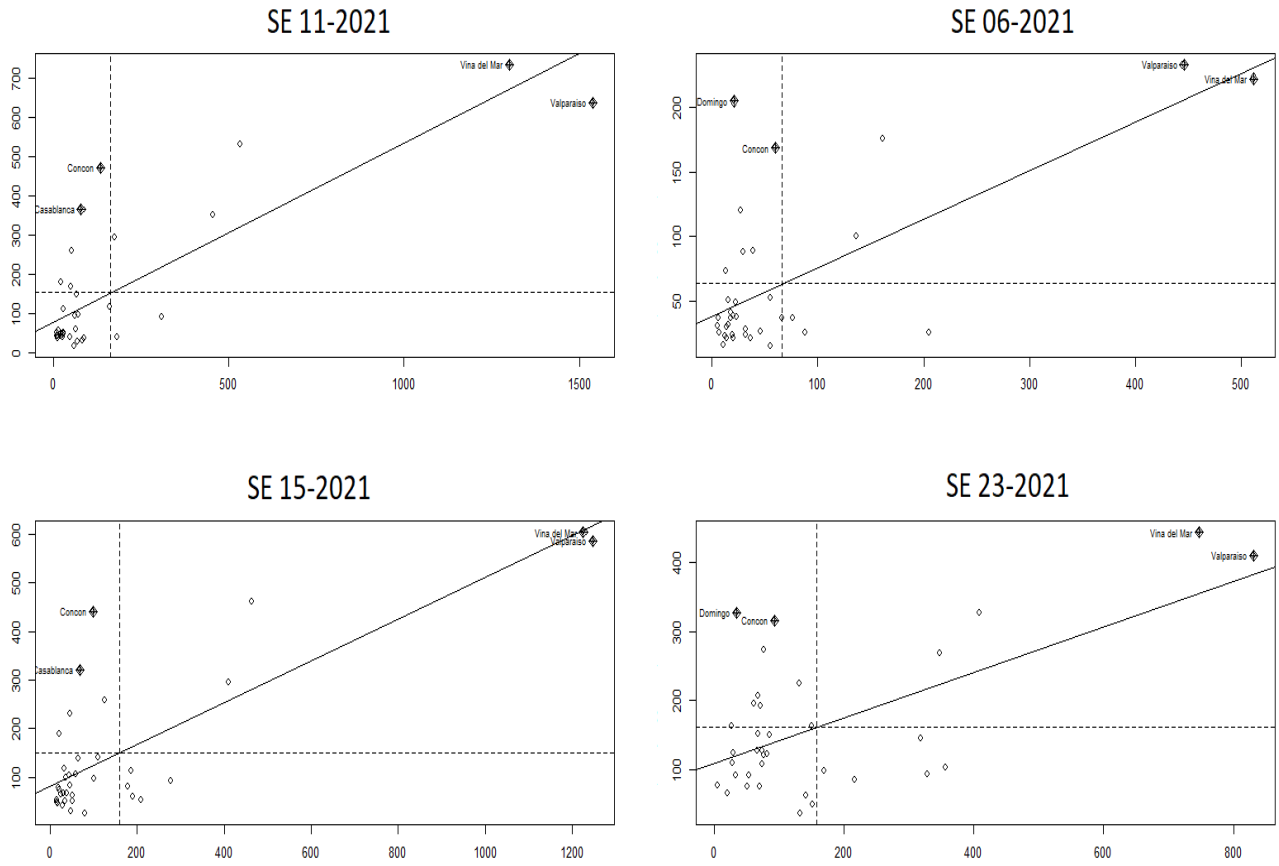

Figura 3.29: Gráfico de Morán (global) de los Casos Confirmados con COVID-19 por SE en las comunas de la Región de Valparaíso.

Fuente: elaboración propia.

Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

En la Figura 3.29 se observa, en el cuadrante superior derecho, que se encuentran las mismas comunas que en las semanas epidemiológicas analizadas anteiormente (Ver Figura 3.28), pero desde la SE 11-2021 a la 19-2021 las Comunas de Valparaíso y Viña del Mar se encontraban en Fase 1, mientras que desde la 20-2021 a la 23-2021 en Fase 2. La Comuna de Concón se mantuvo en el cuadrante superior izquierdo desde la SE 11-2021 a la SE 23-2021, mientras que Casa Blanca solo se mantuvo en el mismo cuadrante en la SE 11-2021 y 15-2021 y Santo Domingo con una población que no superan los 200.000 habitantes se mantuvo en el cuadrante superior izquierdo solo en las SE 06-2021 y 23-2021, por tanto estas comunas, en las respectivas semanas epidemiológicas se presentan como comunas con un menor número de casos nuevos confirmados, rodeadas de otras comunas de la Región de Valparaíso con un alto número de Casos Confirmados, conocidas también como comunas de valores bajos atípicos espaciales.

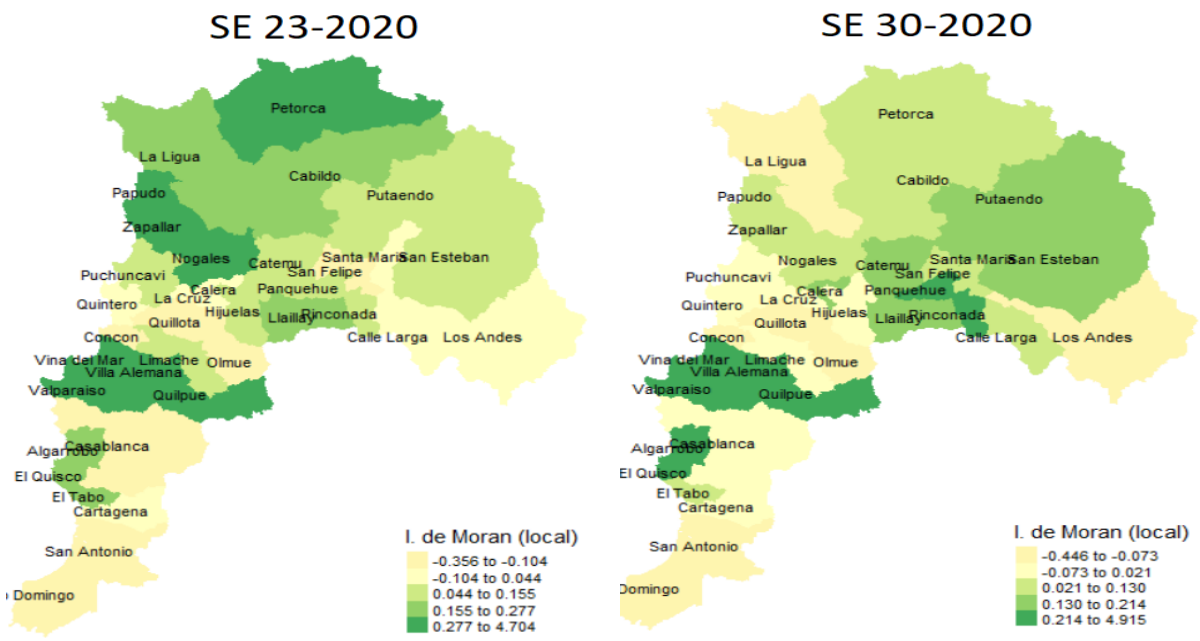

Figura 3.30: *Cluster* espacial de la variable Casos Confirmados con COVID-19 por SE en las comunas de la Región de Valparaíso.

Fuente: elaboración propia.

Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

En la SE 23-2020 el grupo de comunas entre ellas: Valparaíso, Viña del Mar, Villa Alemana y Quilpué junto con Papudo, Zapallar y Nogales se agrupan presentando dependencia espacial alta, pues son comunas que comparten alguna frontera o límite en común, sin embargo la Comuna de Petorca también se encuentra agrupada con las comunas con dependencia espacial alta, pero aquella comuna se encuentra alejada de las que presentan el mismo comportamiento. Por otro lado en la SE 30-2020, Papudo, Zapallar y Nogales se agruparon nuevamente pero en comparación con la SE 23-2020 las comunas presentaron dependencia medianamente alta, mientras que las comunas de El Quisco y Algarrobo junto con Rinconada y Panquehue se agruparon con dependencia espacial alta con las mismas comunas en la SE 23-2020, entre ellas Valparaíso, Viña del Mar, Villa Alemana y Quilpué, estas cuatro comunas presentan una densidad poblacional entre los 300.000 y 400.000 habitantes mientras que El Quisco, Algarrobo, Rinconada y Panquehue no superan los 200.000 habitantes (Ver Figura 3.30).

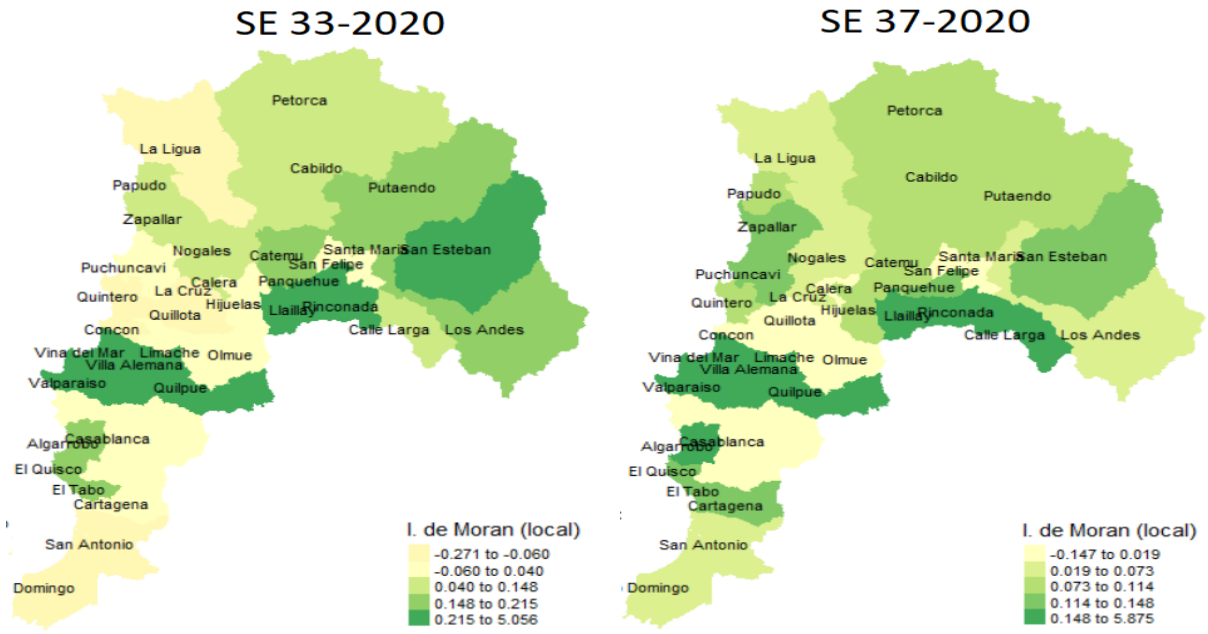

Figura 3.31: *Cluster* espacial de la variable Casos Confirmados con COVID-19 por SE en las comunas de la Región de Valparaíso.

Fuente: elaboración propia.

Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

Las mismas comunas que presentarón una dependencia espacial alta en las SE 23-2020 y 30-2020 (Ver Figura 3.31) nuevamente en la SE 33-2020 mantuvieron el mismo comportamiento, a excepción de las comunas de El Quisco y Algarrobo, sumándose a este mismo comportamiento la Comuna de Llayllay y San Esteban, en esta semana epidemiológica específica ya se encontraban implementadas las medidas de intervención "Plan Paso a Paso". Las Comunas de Viña del Mar y Valparaíso se encontraban en Fase 1, es decir con cuarentena total todos los días de semana, mientras que Quilpué, Villa Alemana, Llayllay, Rinconada, Paquehue y El Quisco se encontraban en Fase 3, todo lo contrario a la descripción de la Fase 1. Transcurridas cuatro semanas epidemiológicas exactamente en la 37-2020 las mismas comunas siguieron con el mismo comportamiento de dependencia espacial alta a excepción de San Esteban. Llama mucho la atención la comuna de San Felipe, pues de encuentra agrupada con las comunas con baja dependencia espacial, pues San Felipe se encuentra rodeada por todas sus fronteras con comunas con dependencia espacial relativamente alta, encontrándose en Fase 3, al igual que las comunas con mayor dependencia espacial. Cabe destacar que las medidas de intervención en las comunas de la Región de Valparaíso no cambiaron entre las SE 33-2020 y 37-2020.

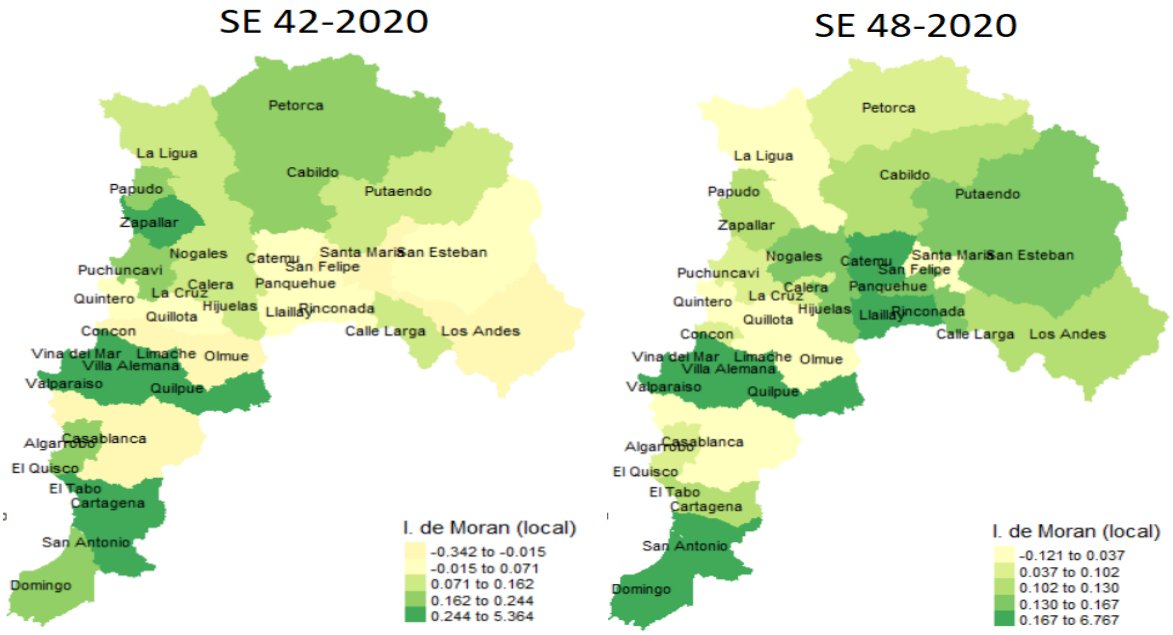

Figura 3.32: *Cluster* espacial de la variable Casos Confirmados con COVID-19 por SE en las comunas de la región de Valparaíso.

Fuente: elaboración propia.

Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

En la SE 42-2020 (Ver Figura 3.32) se puede observar que las comunas de San Antonio, Cartagena, El Tabo y Zapallar se agruparon junto con las comunas mencionadas en las semanas epidemiológicas anteriores, con dependencia espacial alta. Nuevamente llama la atención específicamente la Comuna de Zapallar puesto que alrededor de ella se encuentran comunas con baja dependencia espacial, además Zapallar y las comunas con las que comparte al menos una frontera registran una población similar, que no superan los 100.000 habitantes, además agregar que la Comuna de Zapallar se encontraba en Fase 3 mientras que las demás comunas con dependencia espacial alta en Fase 2, es decir cuarentena total solo sábados y domingos. Por otro lado, en la SE 48-2020 las Comunas Santo Domingo junto con Catemú se unían en el grupo de dependencia espacial alta, Catemú desde la SE 30-2020 no registraba una dependencia espacial relativamente alta, aquella comuna se encontraba en Fase 3 respectivamente al igual que Santo Domingo.

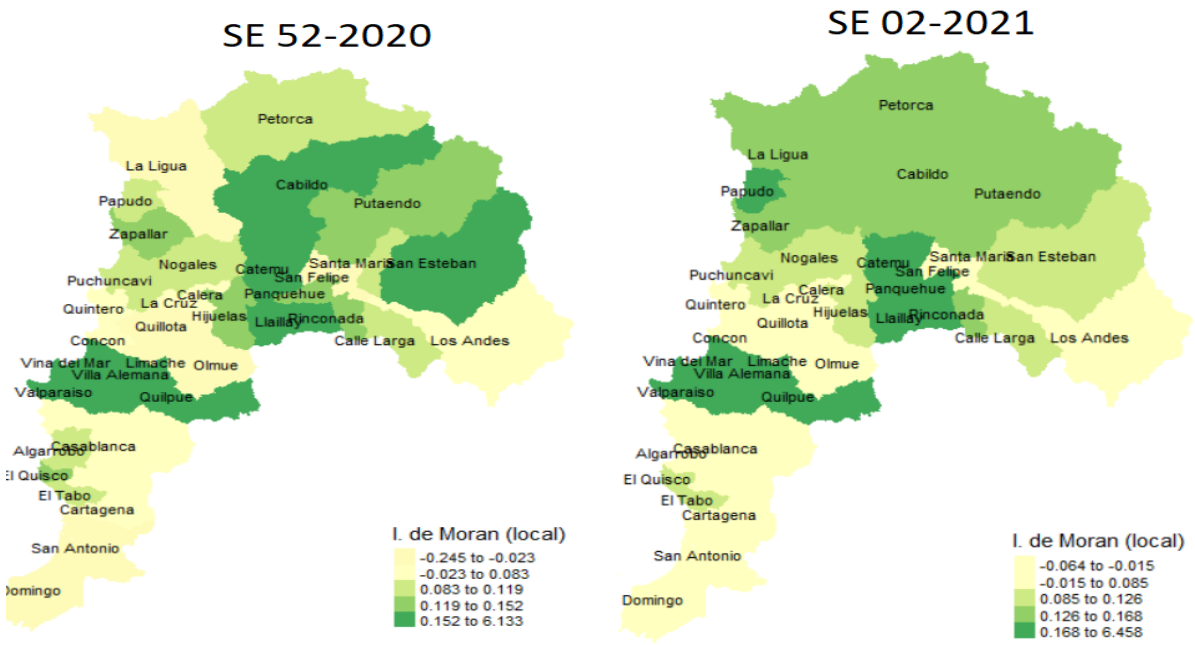

Figura 3.33: *Cluster* espacial de la variable Casos Confirmados con COVID-19 por SE en las comunas de la Región de Valparaíso.

Fuente: elaboración propia.

Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

En la SE 52-2020 se observa (Ver Figura 3.33) el mismo comportamiento descrito anteriormente para el grupo con presencia de dependencia espacial alta, además dos nuevas comunas entran al grupo, estas son la Comuna de San Esteban, que desde la SE 33-2020 no presentaba dependencia espacial alta, y la Comuna de Cabildo. Cabe destacar que todas comunas de la Región de Valparaíso se encontraban en Fase 3, es decir no existía ninguna restricción para la salida de los residentes. Además es importante mencionar que la SE 52-2020 corresponde a la semana de fiestas navideñas y además del inicio de las vacaciones de verano en todo el territorio nacional, por este motivo las comunas turísticas con mayor presencia de población entre ellas Valparaíso y Viña del Mar, siempre presentaron dependencia espacial alta, pues las zonas estaban con exceso de población ya que en los pasos fronterizos no existía una política estable para restringir el acceso de personas.

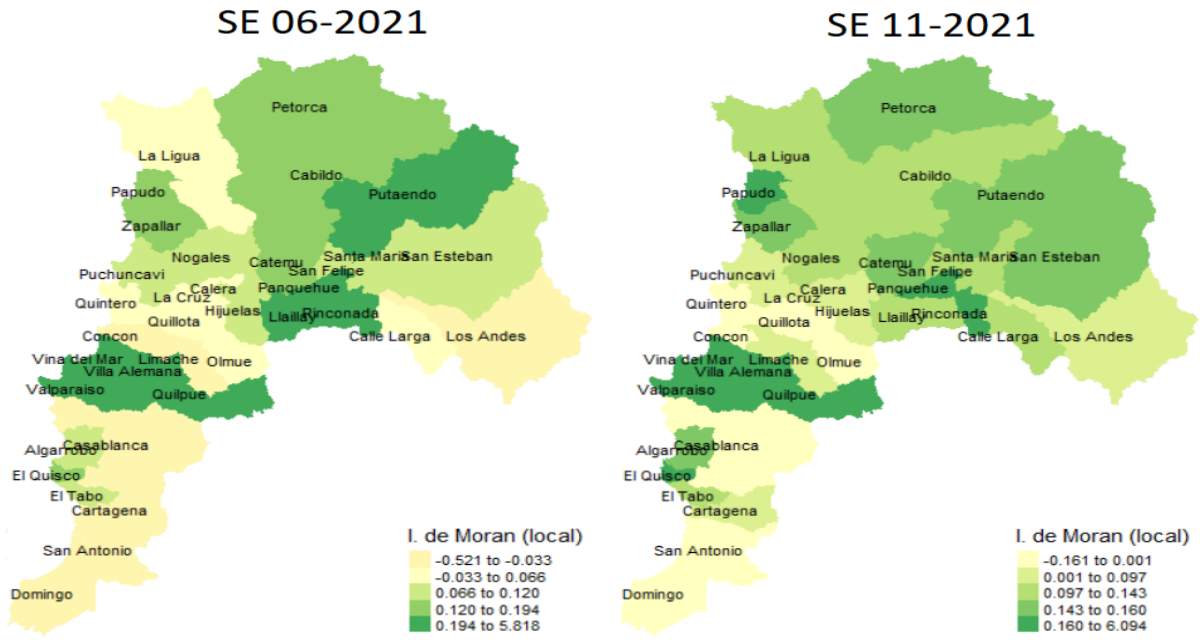

Figura 3.34: *Cluster* espacial de la variable Casos Confirmados con COVID-19 por SE en las comunas de la Región de Valparaíso.

Fuente: elaboración propia.

Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

Como se observa en la Figura 3.34 en la parte inferior de la Región de Valparaíso se han agrupado comunas con una dependencia espacial baja, a excepción de El Quisco, El Tabo y Algarrobo, pues estas se agrupan respectivamente con las demás zonas que presentan una dependencia espacial mucho mayor. Además, mencionar que la mayoría de las comunas de la parte baja de la región se encontraban en Fase 3 a excepción de la Comuna de El Quisco, que se encontraba en Fase 2. Si se observa la parte superior de la Región de Valparaíso en la SE 11-2020 aquellas comunas presentan una mayor dependencia espacial que las comunas de la parte inferior de la región, comportamiento que llama la atención ya que las zonas de la parte alta presentan una densidad poblacional que no superan los 100.000 habitantes. Las comunas centrales de la región, entre ellas Valparaíso, Viña del Mar, Quilpué y Villa Alemana siguen estando agrupadas con dependencia espacial alta. Cabe destacar que la mayoría de las comunas de la parte alta de la Región de Valparaíso se encontraba en Fase 3, a excepción de Catemú que se encontraba en Fase 1 y Papudo en Fase 2.

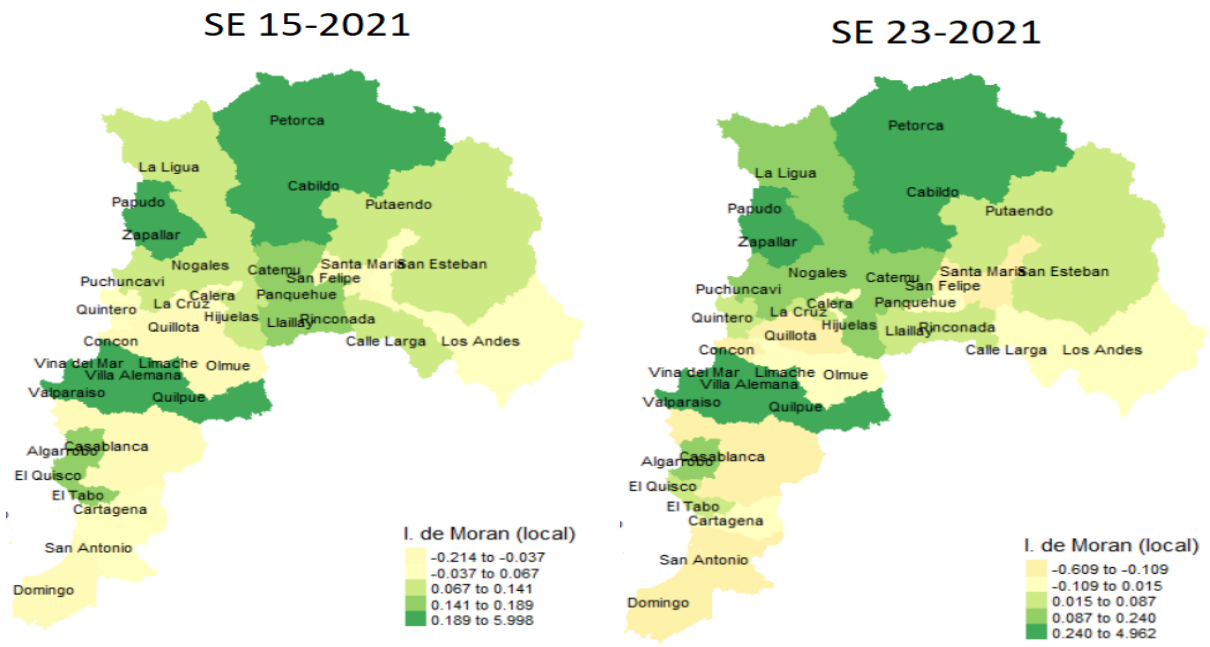

Figura 3.35: *Cluster* espacial de la variable Casos Confirmados con COVID-19 por SE en las comunas de la Región de Valparaíso.

Fuente: elaboración propia.

Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

En la SE 15-2021 las comunas de Petorca y Cabildo se han agrupado con las demás zonas con dependencia espacial alta, desde la SE 23-2020 Petorca no registraba tal comportamiento, mientras que desde la SE 52-2020 Cabildo no registraba una alta dependencia. Nuevamente en la parte inferior de la Región de Valparaíso las comunas se han agrupado con el criterio de dependencia espacial baja, a excepción de El Tabo, El Quisco y Algarrobo, aquellas comunas se han agrupado con las demás zonas con una mayor dependencia espacial. Además se observa en la SE 23-2021 que las mismas comunas se la SE 15-2021 se han agrupado con el mismo comportamiento de dependencia espacial. Cabe destacar que durante esta semana epidemiológica la mayoría de las comunas de la Región de Valparaíso se encontraban en Fase 2, a excepción de las comunas Cabildo, Papudo, Putaendo, Zapallar, San Esteban, Catemú y Concón que se encontraban en Fase 3, mientras que las Comunas de El Tablo, Cartagena y San Antonio en Fase 1 (Ver Figura 3.35).

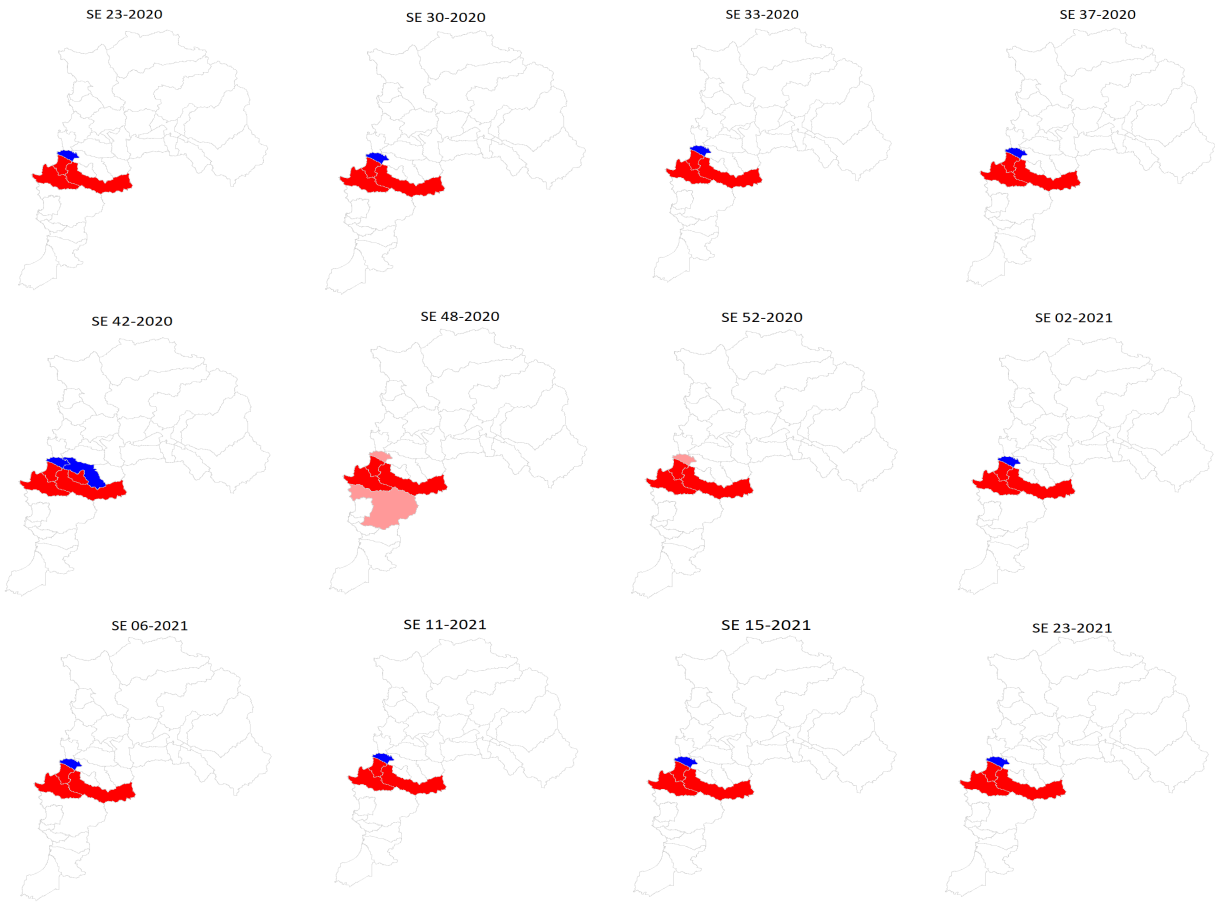

Figura 3.36: LISA *cluster map* de los Casos Confirmados con COVID-19 por SE en las comunas de la Región de Valparaíso. (Rojo) alto-alto , (Rosa) alto-bajo, (Púrpura) bajo-alto, (Azul) bajo-bajo, (Blanco) insignificante

Fuente: elaboración propia.
Datos descargados: Ministerio de Ciencia, Tecnología, Conocimiento e Innovación.

A través del cálculo de los índices de asociación espacial local se pueden obtener los LISA *cluster map* de la variable en estudio, como se observa en la Figura 3.36 en sus respectivas semanas epidemiológicas. Las comunas de Valparaíso, Viña del Mar, Quilpué y Villa Alemana desde la SE 23-2020 a la 23-2021 siempre se mantienen en una categoría de asociación espacial alto-alto con respecto a sus comunas vecinas, mientras que si se observa la comuna de Concón, en la mayoría de las semanas epidemiológicas, presentó una asociación espacial de categoría bajo-bajo con respecto a las comunas con las que comparte al menos una frontera en común, sumándose a la misma categoría en la SE 42-2020 exactamente la Comuna de Limache. Por otro lado en la SE 48-2020 y la 52-2020 las comunas de Concón junto con Casa Blanca presentan una asociación espacial de categoría alto-bajo con respecto a sus vecindades.

Capítulo 4

Conclusión

El presente proyecto de investigación tenía como objetivo explorar la dimámica espacial y temporal del COVID-19 en Chile, obteniendose las siguientes conclusiones:

- A nivel nacional el *cluster* convencional agrupó las regiones con mayor mayor similitud, aunque éstas no fueran regiones vecinas.
- La dinámica espacial del COVID-19 en la Región de Valparaíso en el período de estudio, de Junio 2020 a Junio 2021, presenta asociación de categoría alta-alta para algunas comunas. El índice de Morán local en las comunas de Valparaíso y Viña del Mar junto con Villa Alemana y Quilpué indicó mayor dependencia espacial considerando la variable de Casos Confirmados con COVID-19 entre las semanas epidemiológicas 23-2020 y 23-2021, además de ser las comunas con mayor densidad poblacional. Por otro lado, las mismas comunas se agruparon espacialmente durante todas las semanas epidemiológicas, debido a su dependencia espacial alta y positiva. Cabe señalar que las comunas mayormente pobladas pasaron su mayor tiempo en fases con bastante restricción.
- Con respecto al análisis temporal, las medidas de intervención a lo largo de todo el período de estudio influyeron tanto negativamente como positivamente en algunos períodos de tiempo en las comunas de la Región de Valparaíso, puesto que el comportamiento de los casos confirmados por COVID-19 en algunas zonas era extremadamente elevado, siendo la mejor decisión del Ministerio de Salud determinar fases restrictivas para evitar elevar los casos, no obstante en algunas oportunidades en algunas comunas como por ejemplo la Comuna de San Antonio no se tomó la mejor decisión, puesto que al presentar un alto número de casos la organización decidió implementar fases con mayor facilidad de salir a terreno, esto hacia elevar mucho más los contagios por COVID-19.

Apéndice A

Algoritmos computacionales

Para la obstención de los resultados de este trabajo se realizó una serie de códigos en él software estadístico R, siendo los siguientes:

Análisis espacial

librerías
library(rgdal)
library(rgeos)
library(readxl)
library(sf)
library(sp)
library(raster)
library(tmap)
library(spData)
library(spdep)

rm(list=ls())

sheet para seleccionar el número de hoja

valpocnse = read.xlsx(Ç:/Users/Valentina/correlación espacial/BD COMUNAL.xlsx",sheet =7, col.names = TRUE)
valpocnse = as.data.frame(valpocnse)
head(valpocnse)
names(valpocnse)
dim(valpocnse)

Valparaiso = readOGR(".", ÇOMUNA.C17S")
Valparaiso
head(Valparaiso)
plot(Valparaiso["NOM.COMUNA"])

Valpo.cnse = merge(Valparaiso, valpocnse, by.x=ÇOMUNA", by.y=ÇOD.COMUNA")
head(Valpo.cnse)

tm.shape(Valpo.cnse) + tm.fill("202023", palette = Reds", style = "quantile",title =) + tm.borders(alpha=.4) + tm.symbols(col = "black", size = "Población")

Calculando vecinos
neighbours ¡- poly2nb(Valpo.cnse)
neighbours
plot(Valpo.cnse, border = 'lightgrey')
plot(neighbours, coordinates(Valpo.cnse), add=TRUE, col='red')

Calculate the Rook's case neighbours
neighbours2 ¡- poly2nb(Valpo.cnse, queen = FALSE)
neighbours2

compares different types of neighbours
plot(Valpo.cnse, border = 'lightgrey')
plot(neighbours, coordinates(Valpo.cnse), add=TRUE, col='blue')
plot(neighbours2, coordinates(Valpo.cnse), add=TRUE, col='red')

Convert the neighbour data to a listw object
listw = nb2listw(neighbours2)
listw

global spatial autocorrelation
moran.test(Valpo.cnse‘202023‘, $listw$)

creates a moran plot
$moran < -moran.plot(Valpo.cnse$‘202023‘, listw = nb2listw(neighbours2, style = "W"), labels = as.character (Valpo.cnse$comuna$))

creates a local moran output
$local < -localmoran(x = Valpo.cnse\$202023, listw = nb2listw(neighbours2, style = "W"))$

binds results to our polygon shapefile
moran.map1 ¡- cbind(Valpo.cnse, local)

maps the results

$tm_shape(moran.map12)+tm.fill(col = "Ii", style = "quantile", title = "I.deMoran(local)")+ tm.olygons(labels(Valpo.cnse\$comuna), palette = "Pastel1")$

library(ggrepel)

to create LISA cluster map
quadrant1 ¡- vector(mode="numeric",length=nrow(local1))

centers the variable of interest around its mean
m.qualification1 ¡- Valpo.cnse202025 − $mean(Valpo.cnse!202023)$

```
    centers the local Moran's around the mean
m.local1 ¡- local1[,1] - mean(local1[,1])

    significance threshold
signif ¡- 0.1
builds a data quadrant
quadrant1[m.qualification1 ¿0  m.local1¿0] ¡- 4
quadrant1[m.qualification1 ¡0  m.local1¡0] ¡- 1
quadrant1[m.qualification1 ¡0  m.local1¿0] ¡- 2
quadrant1[m.qualification1 ¿0  m.local1¡0] ¡- 3
quadrant1[local1[,5]¿signif] ¡- 0

    plot in r
par(mfrow=c(4,4))
brks ¡- c(0,1,2,3,4)
colors ¡- c("white","blue",rgb(0,0,1,alpha=0.4),rgb(1,0,0,alpha=0.4),red")
plot(Valpo.cnse,border="lightgray",col=colors[findInterval(quadrant9,brks,all.inside=FALSE)])
box()
legend("bottomleft",legend=c(Ïnsignificante","Bajo-Bajo","Bajo-Alto",."lto-Bajo",."lto-Alto"),
fill=colors,bty="n")
```

Análisis temporal

```
    librerias
library(aTSA)
library(forecast)
library(stats)
library(tseries)
library(uroot)

    library(readxl)
library(xlsx)

    require(mets)
require(data.table)
library(mets)

    install.packages("ggplot2")
library(ggplot2)
library(zoo)
limpiamos
rm(list=ls())

    importamos los datos datos.comunas = read.xlsx(Ç:/Users/Valentina/Desktop/TESIS/Codigo
tesis/datos/Base de datos/comuna/BD COMUNAL.xlsx",
sheet = 7 , col.names = TRUE)
mostramos los datos
```

datos.comunas

Valparaiso =datoscomunasValparaiso
Casablanca =datoscomunasCasablanca
Concon =datos.comunasConcon
Puchuncavi =datoscomunasPuchuncavi

(y así para las 32 comunas más)

par(mfrow=c(3,3))

VALPARAÍSO

valpo = datos.comunasValparaiso
serievalpo¡-ts(valpo)
serievalpo

y ¡- valpo[1:12]
a ¡- valpo[13:15]
b ¡- valpo[16:22]
c ¡- valpo[23:33]
d ¡- valpo[34:42]

seriey¡-ts(y , start = c(1))
seriea¡-ts(a , start = c(13))
serieb¡-ts(b, start = c(16))
seriec¡-ts(c , start = c(23))
seried¡-ts(d , start = c(34))
seriey
seriea
serieb
seriec
seried

serie = plot(serievalpo , main = Çomuna de Valparaíso.", xlab = "Semanas epidemiológicas ",
ylab = Çasos confirmados", ylim =c(0,17500))

lines(seriey, col=(red"))
lines(seriea, col=("blue"))
lines(serieb, col=("green"))
lines(seriec, col=("blue"))
lines(seried, col=(red"))

Acomodar leyenda
legend("topleft", legend = c("Fase 1", "Fase 2", "Fase 3"),
lwd = 3, col = c(red", "blue", "green"), cex = 0.8)

Referencias

Aeroterra. (2021). ¿Qué es SIG?, Sistemas de Información Geográfica [Manual de software informático]. Argentina. Descargado de https://www.aeroterra.com

Alkhamis, M., Youhab, S., Khajahc, M., Haider, N., Alhardan, S., Nabeel, A., ... Al-Sabahb, S. (2020). Spatiotemporal dynamics of the covid-19 pandemic in the state of kuwait. *International Journal of Infectious Diseases*, *98*, 153-160.

Aral, N., y Bakir, H. (2001). Spatiotemporal analysis of covid-19 in turkey. *Sustainable Cities and Society*, *76*. doi: 10.1016/j.scs.2021.103421

Ayala, M. (2014). Introducción a la epidemiologia [Manual de software informático]. Chile. Descargado de https://sintesis.med.uchile.cl/index.php/en/profesionales/informacion-para-profesionales/medicina/condiciones-clinicas2/otorrinolaringologia/1132-7-01-3-013

Barros, O., y Aroca, P. (2014). Econometría espacial y el análisis sociodemográfico. aplicación en la formación de agrupaciones espaciales de envejecimiento en cuba, período 2003-2009. *Scielo*. doi: 40782014000200001

Choynowski, M. (1959). Maps based on probabilities. *ournal of the American Statistical Association*, *54*, 8-385.

Cuartas, D. E., Arango-Londoño, D., Guzmán-Escarria, G., Muñoz, E., Caicedo, D., Ortega, D., y Fandiño-Losada, A. (2020). Análisis espacio-temporal del sars-cov-2 en cali, colombia. *Revista de Salud Pública*, *22*, 1-6.

Cuzick, J., y Edwards, R. (1990). Spatial clustering for inhomogeneous populations. *Journal of the Royal Statistical Society*, *52*, 73–104.

El comercio. (2020). Chile se acerca al colapso por coronavirus con el 97 % de las camas uci ocupadas y los contagios al alza [Manual de software informático]. Lima, Perú. Descargado de https://elcomercio.pe/

Fernández, E. (2014). Para qué sirve la epidemiología. *Cuadernos de las fundación DR. Antonio Esteven*, *32*. doi: 10.1007/s11432-011-4391-8

Fuente, S. (2016). *Análisis cluster*. Madrid: Instrumentos estadísticos avanzados.

Gobierno de Chile . (2020). Plan paso a paso "nos cuidamos todos" [Manual de software informático]. Chile. Descargado de https://www.gob.cl/coronavirus/pasoapaso/

Haining, R. (2003). *Spatial data analysis* (Vol. 2). Cambridge university press.

Hernández, G. (2003). SARS: epidemiología y mecanismos de transmisión. *Medicina intensiva*, *27*(10), 686 - 691. doi: 10.1016/s0210-5691(03)79993-8

ISC III . (2020). Coronavirus: términos epidemiológicos más utilizados [Manual de software informático]. España. Descargado de https://www.isciii.es/InformacionCiudadanos/DivulgacionCulturaCientifica/DivulgacionISCIII/Paginas/Divulgacion/DivulgacionInformeTerminosEpidemiologiaCoronavirus.aspx

Kanox, G. (1964). La detección de interacciones espacio-temporales. *Estadísticas aplicadas*, *13*, 25 - 29.

Kisilevich, S., Mansmann, F., Nanni, M., y Rinzivillo, S. (2010). Agrupación espacio-temporal. *Revista de Estudios Empresariales*, *27*(2), 855-874. doi: 10.1007/978-0-387-09823-4_44

Lawson, A. B. (2006). *Statistical methods in spatial epidemiology* (Vol. 2). CRC Press.

Lawson, A. B., Banerjee, S., Haining, R. P., y Ugarte, M. D. (2016). *A review of gaussian random fields and correlation functions*. Boston, USA: CRC Press.

MinCiencia. (2020). Bases de datos covid-19 [Manual de software informático]. Chile.

OIT. (2020). Chile: efectos de la pandemia generaron consecuencias sin precedentes en el mundo del trabajo [Manual de software informático]. Chile. Descargado de https://www.ilo.org

OMS. (2020). Información básica del covid-19 [Manual de software informático]. Descargado de https://www.who.int/es

Rencher, A. (1934). *Methods of Multivariate Analysis* (Vol. 2). A JOHN WILEY SONS, INC. PUBLICATION.

Sanchéz, M. (2008). Análisis espacial de datos y turísmo; nuevas técnicas para el análisis turísico. Una aplicación al caso extremo. *data mining and knowledge discovery handbook*(2), 48-66. doi: 10.1016/s0210-5691(03)79993-8

Tenisi, B. L. (2021). Coronavirus y síndromes respiratorios agudos (covid-19, mers y sars) [Manual de software informático]. USA.

Vergara-Perucich, J. (2020). Correlación espacial entre la propagación del covid-19 y las zonas urbanas vulnerables en santiago de chile. *critical housing analysis*, *7*, 21-35.

Wong, G., y Hui, D. (2003). Síndrome respiratorio agudo severo (sars): epidemiología, diagnóstico y manejo. *Eurosurveillance*, *8*, 121.