

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL



**FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICAS
DEPARTAMENTO DE POSTGRADO**

PROYECTO DE TITULACIÓN

PREVIO A LA OBTENCIÓN DEL TÍTULO DE:

MAESTRÍA EN ESTADÍSTICA APLICADA

TEMA:

**IDENTIFICACIÓN DE FACTORES DE RIESGO EN LOS INGRESOS
HOSPITALARIOS POR COVID-19 A TRAVÉS DE UN ESTUDIO MULTIVARIANTE DE LA
CIUDAD DE MILAGRO.**

AUTOR:

GUISELLA ISABEL VILLAMAR VASQUEZ

Guayaquil - Ecuador

2022

RESUMEN

De los diversos datos registrados por el distrito de salud zonal 5, se propuso determinar factores de riesgo en los ingresos hospitalarios de pacientes positivos con Covid-19 en la ciudad de Milagro para el año 2021. De los resultados se obtuvieron que la cantidad de hombres y mujeres es proporcional, el nivel de atención más utilizado fue servicio auxiliar que corresponde a hospitales y clínicas privadas; quienes estuvieron dentro de las edades 20 a 49 y 50 a 64 años, fueron los más vulnerables, ya que son consideradas personas en edad activa o productiva, sin embargo, de acuerdo al estado civil la mayoría de grupos vulnerables son los solteros y casados. Los factores más significativos fueron el nivel de atención, los síntomas, lugar de atención, aquellas personas que fueron atendidos de manera ambulatoria no presentaron descensos, las contribuciones de los diversos factores se corroboró con un ACM observando que el primer hiperplano explica una variabilidad considerable, estos distintivos ayudaron a desarrollar un modelo logit multinomial, con una tasa de clasificación correcta del 91,38%. El modelo logró explicar una variabilidad de 87,02% valor determinado mediante el R^2 de Mc-Fadden.

Palabras claves: Factores de riesgo, Análisis multivariante, Análisis de Correspondencia Múltiple, Covid-19, Contagio, Síntomas de Covid-19.

ABSTRACT

From the various data recorded by the zonal health district 5, it was proposed to determine risk factors in hospital admissions of Covid-19 positive patients in the city of Milagro for the year 2021. The results showed that the number of men and women is proportional, the most used level of care was auxiliary service that corresponds to hospitals and private clinics; those who were within the ages 20 to 49 and 50 to 64 years, were the most vulnerable, since they are considered people in active or productive age, however, according to marital status most vulnerable groups are single and married. The most significant factors were the level of care, symptoms, place of care, those persons who were attended on an outpatient basis did not present declines, the contributions of the various factors were corroborated with an MCA observing that the first hyperplane explains considerable variability, these distinctions helped to develop a multinomial logit model, with a correct classification rate of 91.38%. The model was able to explain a variability of 87.02% value determined using Mc-Fadden's R^2 .

Key words: Risk factors, Multivariate analysis, Multiple Correspondence Analysis, Covid-19, Contagion, Covid-19 symptoms.

DEDICATORIA

Dedico este trabajo a Dios, por mantenerme con salud, trabajo y con todas las fuerzas para seguir preparándome en el área profesional, a mi Esposo, mis hermosos hijos Daniel y Néstor, por ustedes es mi lucha constante, para que aprendan el significado de que la capacitación constante es lo mejor, a mi Mami Nelly sin su apoyo nada de esto fuera posible, la amo mamita, a mi mamita Chabela, mi guerrera y valiente y mi Ñaña Jane, por estar siempre conmigo, mis hermanos, mi familia y amigos.

Papi Néstor, ñaña Gina y todos los que han partido y sé que desde el lugar que se encuentren siempre deseaban lo mejor para mí, de la tierra al cielo se los dedico.

AGRADECIMIENTO

Agradezco infinitamente a Dios por todas las bendiciones, por la salud y por la fortaleza que me da día a día para cumplir con mis metas, a mi Amada viejita mi Mami Chabela gracias porque sé que en tus oraciones siempre estoy presente, a mi Mami Nelly que es mi ejemplo y mi apoyo incondicional para mi educación, a mi Ñaña Jane gracias negrita por tanto y por todo, a mi Esposo por su apoyo incondicional, a mis hijos Daniel y Néstor, que entiendan que todo sacrificio tiene su recompensa que mami los ama mucho, a mis hermanos, familiares y amigos por quererme y darme siempre su respaldo, a mi suegra Sra. Narcisa a usted por toda su ayuda y consejos, A mi Paty, que hago sin ti, gracias.

A mi tutora, sin sus sabias enseñanzas, correcciones y toda la paciencia que me ha tenido le agradezco de todo corazón, sin usted nada de esto sería posible mi Estimada Mariela fue un placer ser su estudiante.

Mis queridos y estimados amigos que hice en esta maestría, Andrés y Jonathan, mi grupo 5, a parte de lo aprendido fueron lo mejor de esta travesía, gracias por permitirme ser su amiga, los quiero chicos, gracias por tus enseñanzas Andrés eres el mejor.

Gracias Kike, por aterrizarme en tiempo y espacio eres un crack en lo que haces y en lo que sabes, agradezco tu ayuda, apoyo y tu tiempo.

Finalizo con mis ángeles del cielo, los recuerdo siempre con cariño, sé que siempre querían lo mejor para mí, sé que en el lugar en el que se encuentren siguen cuidando de mí y que les hubiese encantado compartir conmigo estos momentos, los llevo en mi corazón siempre.

DECLARACIÓN EXPRESA

La responsabilidad por los hechos y doctrinas expuestas en este Proyecto de Titulación, me corresponde exclusivamente y ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de mi total autoría. El patrimonio intelectual del mismo, corresponde exclusivamente a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance del Trabajo de Titulación referido.

Guisella Isabel Villamar Vásquez

TRIBUNAL DE GRADUACIÓN

OMAR HONORIO RUIZ BARZOLA PhD
PRESIDENTE

MARIELA GONZALEZ NARVAEZ PhD
TUTOR

FRANCISCO ANTONIO MOREIRA VILLEGAS MSc.
DOCENTE EVALUADOR

ABREVIATURAS O SIGLAS

ACM	Análisis de Correspondencia Múltiple
ANOVA	Analysis of Variance
Covid-19	COronaVirus Disease
IESS	Instituto Ecuatoriano de Seguridad Social
INSPI	Instituto Nacional de Investigación en Salud Pública
MSP	Ministerio de Salud Publica
OC	Organ Culture
OMS	Organización Mundial de la Salud
PIB	Producto Interno Bruto
RNA	Ácido Ribonucleico
RT-PCR	Real Time - Polymerase Chain Reaction
SARS-CoV-2	Severe acute respiratory syndrome-related coronavirus type 2

TABLA DE CONTENIDO

CAPÍTULO 1	1
1. INTRODUCCIÓN	1
1.1. Antecedentes.....	1
1.2. Descripción del problema.....	6
1.3. Objetivos.....	7
1.3.1. Objetivo general	7
1.3.2. Objetivos específicos	7
1.4. Hipótesis.....	8
1.5. Alcance.....	8
CAPÍTULO 2	9
2. MARCO TEÓRICO Y ESTADO DEL ARTE.....	9
2.1. Estado del arte	9
2.1.1. SARS-CoV-2	9
2.1.1.1. Origen.....	9
2.1.2. Taxonomía del SARS-CoV-2	11
2.1.3. Transmisión	12
2.1.4. Variantes detectadas.....	13
2.1.5. Parámetros clínicos.....	15
2.1.5.1. Historias clínicas	15
2.1.6. Factores de riesgo.....	15
2.1.7. Análisis Multivariante en casos Covid-19	18
2.2. Métodos Estadísticos Multivariantes.....	23
2.2.1. Análisis de Conglomerados	23
2.2.1.1. Partición de los datos.....	24
2.2.1.2. Construcción de jerarquías	24
2.2.1.3. Clasificación de variables.....	24
2.2.2. Métodos Clásicos De Partición	25

2.2.2.1.	Fundamentos del algoritmo de k-medias.....	25
2.2.3.	Análisis de correspondencia Múltiple	26
2.2.3.1.	Matriz binaria	26
2.2.3.2.	Matriz de Burt	27
2.2.3.3.	Identificación del número correcto de dimensiones	27
2.2.4.	Modelo logit multinomial.....	28
CAPÍTULO 3		30
3.	METODOLOGÍA	30
3.1.	Tipo de enfoque	30
3.2.	Diseño.....	30
3.2.1.	Investigación Exploratoria	30
3.2.2.	Investigación Descriptiva	31
3.3.	Alcance	31
3.4.	Población.....	31
3.5.	Muestra.....	31
3.6.	Definición de variables	32
3.7.	Métodos o instrumentos de recolección de información.....	33
3.8.	Procedimiento o fases de la investigación.....	34
3.8.1.	Pre-tratamiento de la información.....	34
3.8.2.	Análisis Descriptivo	34
3.8.3.	Análisis Relacional	35
3.8.4.	Análisis Multivariante.....	36
CAPÍTULO 4		37
4.	RESULTADOS.....	37
4.1.	Procesamiento de la información.....	37
4.2.	Estadísticos descriptivos.....	38
4.2.1.	Distribución sociodemográfica	38
4.2.1.1.	Distribución biológica	38
4.2.1.2.	Distribución Social.....	39
4.2.2.	Distribución según el nivel de atención en hospitales y laboratorios.....	41
4.2.3.	Distribución según la condición y diagnóstico del paciente	42
4.2.4.	Ingreso de paciente en función del tiempo	45
4.3.	Análisis relacional.....	45
4.4.	Análisis multivariante.....	47
4.4.1.	Análisis de correspondencia múltiple	47
4.4.2.	Modelo logit Multinomial.....	52
CAPÍTULO 5		59

5.	CONCLUSIONES Y RECOMENDACIONES.....	59
6.	Referencias.....	1
7.	Apéndices y anexos.....	5

LISTADO DE FIGURAS

Figura 2.1 <i>Estructura de un coronavirus</i>	10
Figura 2.2 <i>Taxonomía del SARS-CoV-2.</i>	12
Figura 2.3 <i>Mecanismos de transmisión del SARS-CoV-2</i>	13
Figura 2.4 <i>Protocolo de valoración de la vulnerabilidad del trabajador ante la infección del Covid-19.</i>	17
Figura 4.5 <i>Distribución de la información en la data</i>	37
Figura 4.6 <i>Distribución biológica de pacientes según</i>	39
Figura 4.7. <i>Distribución social del paciente en la ciudad de Milagro</i>	41
Figura 4.8 <i>Distribución según el nivel de atención en hospitales y laboratorios</i>	42
Figura 4.9 <i>Distribución de la condición del paciente</i>	44
Figura 4.10 <i>Pacientes que ingresan al centro de salud por problemas de Covid-19</i>	
Figura 4.11 <i>Variabilidad explicada por cada dimensión</i>	48
Figura 4.12 <i>Plano factorial en base a las variables</i>	49
Figura 4.13 <i>Conglomerados de los individuos</i>	50
Figura 4.14 <i>Análisis de correspondencia múltiple de: A) con las variables significativas y B) óptimas del modelo según V de Cramer</i>	51

LISTADO DE TABLAS

Tabla 2.1	<i>Mutaciones y variantes del SARS-CoV-2</i>	14
Tabla 2.2	<i>Variantes de Covid-19 detectadas en Ecuador</i>	18
Tabla 3.3	<i>Detalle de variables utilizadas para el análisis</i>	32
Tabla 3.4	<i>Valores del índice V de Cramer</i>	36
Tabla 4.5	<i>Dependencia del diagnóstico final con las variables</i>	46
Tabla 4.6	<i>Variables significativas según el criterio V de Cramer</i>	47
Tabla 4.7	<i>Variables consideradas para estimar el modelo logit multinomial</i>	52
Tabla 4.8	<i>Tabla cruzada de datos originales y predichos</i>	53
Tabla 4.9	<i>Aporte de variables al modelo logit multinomial</i>	54
Tabla 4.10	<i>Tabla cruzada sin la variable "muestra" y "confirmado por"</i>	55

CAPÍTULO 1

1. INTRODUCCIÓN

1.1. Antecedentes

En diciembre del año 2019 se reporta la aparición de una nueva enfermedad conocida como, SARS-CoV-2, determinado como el causante del Covid-19. Este coronavirus pertenece a una amplia familia de virus y una de sus capacidades es que pueden ser transmitidos entre especies, en este caso de animales a humanos. Se sabe que los primeros casos se dieron en la localidad de Hubei, ciudad Wuhan, China (Chauca, 2021).

La Organización Mundial de la Salud (OMS) confirmó 4351 casos positivos de coronavirus en 49 países, esto fue hasta febrero del 2020, por lo que la expansión de este padecimiento fue considerada de riesgo muy alto. Esta disposición fue tomada por la rápida infección y por las diversas dificultades causadas en diferentes países, ya sea desarrollado o subdesarrollado. Sin embargo, la OMS mencionó que el tema no debe causar pánico más bien señalaron que la parte administrativa debe estar preparada para enfrentarse a esta enfermedad, de modo que se pueda detener la propagación y salvaguardar la vida de sus pobladores (OMS, 2020).

Cabe mencionar que los diversos países no tomaron acciones a pesar de las advertencias y las alertas emitidas, por ende, no se pudo contener la propagación del virus alrededor del mundo. La peor parte de la pandemia fue para las familias que vivían en las

zonas urbanas donde existe una mayor proporción de habitantes y áreas de alta aglomeración para las distintas actividades económicas.

Fleitas, et al. (2021) desarrollaron un estudio con el objetivo de puntualizar los principales síntomas asociados al Covid-19, así como sus características diagnósticas. Realizaron un análisis de la información recolectada de todos los pacientes diagnosticados por RT-PCR para SARS-CoV-2 entre abril y mayo de 2020 en Argentina. La data estaba compuesta de información clínica y demográfica de todos los pacientes que ingresaron por posible caso de coronavirus (n=67318, donde el 12% fue positivo para SARS-CoV-2), para una descripción agrupada dividieron la información en cuatro rangos de edad: pediátrica (0-17 años), adultos jóvenes (18-44 años), adultos (45-64 años) y ancianos (65-103 años) esto permitió conocer a las personas más vulnerables. Aplicaron la regresión logística multivariante para estudiar la asociación de todos los síntomas y crear un modelo de diagnóstico basado en las sintomatologías. Sus resultados mostraron que la anosmia, disgeusia, cefalea, fiebre baja, odinofagia y malestar general fueron los síntomas asociados. Sin embargo, estos síntomas fueron diferentes entre los distintos grupos de edad. Los indicios asociados a problemas respiratorios (dolor torácico, dolor abdominal y disnea) tuvieron una asociación negativa con el Covid-19 no presentaron relevancia estadística. Así mismo el modelo basado en 16 síntomas relevantes, la edad y el sexo, presentó una sensibilidad del 80% y una especificidad del 46%, este modelo puede ser utilizado como método de diagnóstico para casos sospechosos de esta enfermedad.

En Bahía Blanca, Argentina Viego, et al. (2020) desarrollaron un estudio que permitió conocer y estimar el periodo de incubación del Covid-19. Obteniendo como resultados un intervalo temporal aproximado de 4,6 a 11,1 días con un promedio de

incubación de 7,9 días. Los autores aclaran que la transmisión de asintomáticos y presintomáticos no deben ser despreciados.

La OMS (2020) explica que:

En la mayoría de los casos, el Covid-19 produce síntomas leves como tos seca, cansancio y fiebre, aunque la fiebre puede no ser un síntoma en algunas personas mayores. Otros síntomas son: dolor, congestión nasal, rinorrea, dolor de garganta o diarrea. Algunas personas que se infectan no presentan ningún síntoma ni se sienten mal (p. 2).

Fisiopatológicamente no se ha esclarecido a ciencia cierta cómo responde el sistema inmunológico del ser humano al contagiarse con Covid-19, en algunas personas la presencia del virus es asintomática. En otras se manifiesta como un pequeño resfriado, otros presentan una neumonía y en casos más graves y severos fallan todos los organismos de control y defensa del cuerpo, entre los sistemas que se ven afectados están el sistema nervioso, respiratorio, digestivo produciendo así la muerte del individuo.

Al primer mes del inicio de la pandemia en Madrid, España Arribas, *et al.* (2020) realizaron un estudio con el fin de describir el avance de esta enfermedad y representar la realidad que vive una unidad hospitalaria en la sección de hemodiálisis. Los resultados de un total de 90 personas que se sometieron a las pruebas PCR donde el 41,1% arrojaron positivos y de estas el 59,4 eran sintomáticos y el resto asintomáticos, de los pacientes infectados el 43,2% fueron ingresados y el 16,2% fallecieron, se concluye que existe una gran prevalencia del Covid-19 y se notó síntomas en el personal de salud dentro del área.

Un estudio realizado por Chauca (2021) afirmó que el Ecuador es uno de los países sudamericanos con el mayor número de contagios ya que su propagación fue de manera exponencial. Por tanto, se planteó realizar un análisis descriptivo que exponga el avance de esta enfermedad, donde los resultados expresaron un alto grado de contagio en la ciudad de Quito y Guayaquil con una tasa de mortalidad de 2,9 personas contagiados por cada 100 mil habitantes y comentan acerca del ajuste de políticas que ayuden y protejan a la población.

Pérez *et al.* (2021) manifiestan que, una vez que las autoridades chinas revelaron la caracterización de la nueva cepa mortal de coronavirus llamado SARS-CoV-2, el estudio científico sobre los factores de riesgo de transmisión del Covid-19 ha sido intenso y la necesidad de este tipo de análisis se ha vuelto importante. Así como la relación entre el Covid-19 con las condiciones ambientales, empleando diferentes análisis estadísticos multivariantes. Los resultados de las componentes principales y la reducción de los factores de riesgo fueron tres componentes que explican el 71% de la variabilidad original, mientras que el análisis de conglomerados fue utilizado para delimitar el territorio de Madrid de acuerdo a las componentes de riesgo. Mediante un test ANOVA se reveló diferentes tasas de incidencia entre los territorios delimitados por los componentes y finalmente, se aplicó un conjunto de modelos lineales para demostrar cómo los factores ambientales presentan una mayor influencia en las infecciones por Covid-19 que las dimensiones socioeconómicas.

Llanos y Bermúdez (2020) plantean que la vulnerabilidad de las personas frente al Covid-19 se encuentra relacionada con las condiciones de vida, y estudian la propagación de esta enfermedad, respecto a las características socioeconómicas en el Municipio de Pitalito Huila. Los resultados mostraron una correlación alta con el número de habitantes

que viven en condición de arrendatarios y el número de cuartos por vivienda esto dentro de la parte urbana.

Devkota (2021) realizó un trabajo con el fin de comprender la dinámica de la propagación del Covid-19 en Nepal. Esto se realiza aplicando técnicas estadísticas multivariantes, como la regresión logística multinomial, la razón de momios, los modelos lineales de efectos mixtos y el análisis de componentes principales. Consideran que este tipo de estudios es de suma importancia para países con datos limitados y escasos. Utilizando datos diarios multivariantes, se analizó minuciosamente el impacto de la restricción gubernamental en forma de cierre estricto 1, cierre parcialmente relajado 1, cierre completamente relajado 1 y cierre estricto 2. Estos controles se ejercieron para frenar la propagación de la pandemia. También se estudia aquí el papel de estos controles para frenar la propagación de la pandemia. Los resultados obtenidos en este estudio pueden aplicarse a otros países del sur de Asia y de África ya que muestra una aproximación ideal a la información recolecta y a la realidad que vivió Nepal.

Para el Ecuador se planteó un confinamiento con limitación humana este empezó el 17 de marzo del 2021 y la normalidad fue de manera progresiva y por ciudades, la ciudad de Guayaquil fue una de las más golpeadas a inicios de pandemia por lo que los controles fueron exhaustivos y la movilidad humana fue controlada a grandes escalas.

Actualmente dentro del país no existe un estudio centrado en los diversos factores de riesgo de contagio de la enfermedad y de las actividades realizadas por las personas antes de contraer el virus. Esto ayudará a comprender las principales causas de contraer dicho padecimiento y permitirá prevenir un rebrote de la pandemia, y garantizará una mejora en la calidad de vida de los pobladores.

La ciudad de Milagro se localiza en la provincia de Guayas, es la cabecera Cantonal y se considera uno de los poblados más grandes de su jurisdicción. Ubicada al sur de la región litoral, se considera una extensa llanura la cual es atravesada por el río Milagro. Según el censo del año 2010 se constató un total de 133,508 habitantes.

En la siguiente investigación el área de estudio corresponde a la parte urbana de la ciudad de Milagro. Sus principales parroquias urbanas son Camilo Andrade Manrique, Chirijos, Coronel Enrique Valdez C. y Ernesto Seminario Hans, donde se pretende realizar un análisis exploratorio de sus pacientes con problemas de Covid-19.

1.2. Descripción del problema

En la actualidad, la pandemia por Covid-19 se ha convertido en una de las barreras que el mundo no ha logrado superar, afectando a los diversos sectores de los países. Ecuador no es la excepción, recibe uno de los golpes más fuertes tanto en la parte económica como en lo sanitario. Según organismos de la Salud Pública del país, el Instituto Nacional de Investigación en Salud Pública (INSPI), menciona que han ingresado 2'470.170 muestras para RT-PCR Covid-19, de los cuales 732.038 son casos confirmados con pruebas PCR, dicha cifra se actualiza de manera diaria (MSP, 2022).

Los centros de salud se vieron colapsados a causa de la propagación de la enfermedad, por lo que es primordial efectuar un análisis estadístico descriptivo y exploratorio con el fin de describir la distribución de la afectación del virus en la población, así como también, el identificar posibles patrones de comportamiento, semejanzas y principales variables que ayuden a explicar de manera óptima las causas antes de un posible contagio en la Ciudad de Milagro. Para ello se pretende realizar un estudio con

información recolectada durante el año 2021, la demografía de Milagro se encuentra segmentada demográficamente en área urbana y rural, la ciudad en su totalidad cuenta con 199835 habitantes, (INEC, 2013) y está ubicada en el centro-sur de la región litoral del Ecuador.

En Ecuador no se ha desarrollado un estudio del comportamiento de pacientes antes de contraer la enfermedad del Covid-19, por lo tanto, es importante analizar las diversas actividades que realiza una persona previamente a ser hospitalizado por el contagio, esto ayudaría a mejorar las medidas de prevención tanto en lo personal como a nivel nacional.

1.3. Objetivos

1.3.1. Objetivo general

Determinar los factores de riesgo de ingresos hospitalarios de los habitantes de la ciudad de Milagro en el 2021, mediante la aplicación de técnicas estadísticas multivariantes, para la identificación de variables relacionadas con el evento de contagio por Covid-19.

1.3.2. Objetivos específicos

- Describir las condiciones sociodemográficas de pacientes que fueron ingresados con síntomas de Covid-19, mediante la aplicación de técnicas de clasificación, segmentación y distribución de los casos confirmados, para el reconocimiento de grupos vulnerables.

- Analizar relaciones / asociaciones entre variables, mediante el uso de métodos estadísticos de correlaciones y/o pruebas de independencia, para la identificación de posibles factores de riesgo.
- Identificar las principales variables relacionadas con el evento de contagio por Covid-19 mediante la aplicación de técnicas multivariantes, para la representación de posibles casos de ocurrencia.

1.4. Hipótesis

En el siguiente estudio no se plantea hipótesis.

1.5. Alcance

Para la elaboración de este análisis se utilizarán los datos proporcionados por el sistema de salud de la Coordinación de Salud Zona 5, que contiene el registro de los habitantes en la ciudad de Milagro que fueron ingresados a un establecimiento hospitalario por presentar síntomas de Covid-19 en el año 2021.

El estudio permitirá establecer políticas que mejoren la toma de decisiones acorde a la situación, por parte de las autoridades de la ciudad de Milagro. Esto ayudará a mejorar la calidad de vida con medidas de cuidado personal, familiar, laboral, etc., minimizando y controlando la propagación de esta enfermedad; la investigación podrá ser replicada en otras ciudades aledañas del país.

CAPÍTULO 2

2. MARCO TEÓRICO Y ESTADO DEL ARTE

2.1. Estado del arte

2.1.1. SARS-CoV-2

2.1.1.1. Origen

Desde el descubrimiento o aparición del coronavirus SARS-CoV-2, en diversas partes del mundo científicos han entrado en debate sobre su origen, especulando que es un producto manipulado en laboratorios, sin embargo, científicos aclaran que no se puede sostener esta hipótesis ya que no es un virus derivado de otro antecedente. Se sabe que es similar al SARS de los murciélagos y este podría ser huésped (Ciotti et al., 2020).

Torres (2020) menciona que el origen del Covid-19 es un tema de múltiples debates en redes sociales, además indica que la comunidad científica trabaja arduamente para poder tener bases sólidas y explicar de manera clara el surgimiento de esta enfermedad. Asimismo, Oliva (2020) menciona que es una enfermedad respiratoria con una gran variabilidad de manifestación que pasa de leve a grave con gran rapidez e indica que su aparición es presuntamente en un mercado en la ciudad de Wuhan.

Según Ortega, *et al.* (2021) indican que el coronavirus es una enfermedad que afecta a mamíferos causando enfermedades gastrointestinales y respiratorias. Este virus forma partículas pleomorfas o tipo esféricas con un diámetro que varía desde los 100 hasta

160 nm. Las variantes HCoV-OC43, HKU1, HCoV-NL63 y HCoV-229E son causantes de problemas respiratorios leves, pero los más peligrosos son MERS-CoV y SARS-CoV, estas son causantes de síndromes respiratorios severos, las características del coronavirus concuerdan con datos recolectados por Maguiña, Gastelo y Tequen (2020).

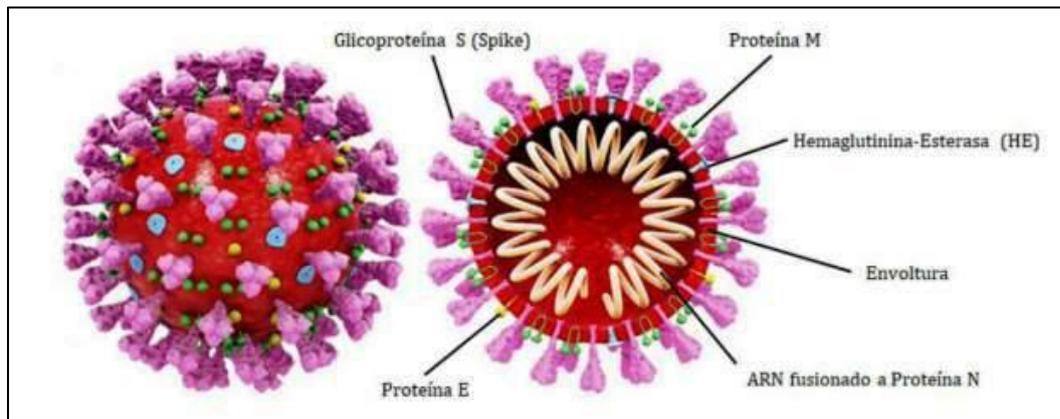


Figura 2.1 Estructura de un coronavirus

Fuente: Maguiña, Gastelo y Tequen, (2020)

A pesar de los diversos debates en los diferentes medios de comunicación, una búsqueda exhaustiva realizada por Ruiz y Jiménez (2020) hallaron que los primeros coronavirus detectados en humanos fueron identificados por la década de los 60, esto fue gracias a la investigación establecida para analizar virus respiratorios que se realizó en Inglaterra en el año de 1960. Constataron una muestra denominada B814, esta era capaz de conservar su infectividad y el resultado se obtuvo de cultivos en tráquea humana realizados en diversos voluntarios, además se constató que su replicación no fue detectable en cultivos de fibroblastos y riñones humanos y de primates, distinguidos en aquella época como virus de gripe, adenovirus, rinovirus y entre otros. En el mismo periodo, en Chicago se aislaron 5 clases de virus diferentes a los que causan afecciones al sistema respiratorio. De estos antecedentes en los siguientes años cepas similares fueron descubiertas, a las

cuales se las clasificó como *Organ Culture* (OC), ya que su aislamiento no es posible en cultivos celulares, a causa de esto se vieron obligados a realizar cultivos de órganos.

2.1.2. Taxonomía del SARS-CoV-2

Helmy, *et al.* (2020) en su investigación mencionan que la enfermedad de los coronavirus (CoV) es uno de los más grandes grupos de virus que pertenecen al orden Nidovirales, suborden Coronidovirineae y a la familia Coronaviridae, mientras que, Coronaviridae se divide en dos subfamilias, a saber, Letovirinae y Orthocoronavirinae. Letovirinae incluye el género Alphaletovirus, mientras que Orthocoronaviridae, sobre la base del análisis filogenético y la estructura del genoma, se enlista cuatro géneros: Alphacoronavirus, Betacoronavirus, Gammacoronavirus y Deltacoronavirus.

La creación de un super reino se dio a conocer en el año 2017 a la cual se denominó Riboviria, realizado con el fin de incluir cada uno de los taxones de virus que tengan RNA-polimerasa y además que dependa de RNA (Ruiz & Jiménez, 2020), su distribución taxonómica se puede observar en la Figura 2.2.

SUPERREINO <i>Riboviria</i>			
ORDEN <i>Nidovirales</i>			
SUBORDEN <i>Cornidovirineae</i>			
FAMILIA <i>Coronaviridae</i>			
SUBFAMILIA <i>Orthocoronavirinae</i>	GÉNERO <i>Betacoronavirus</i>	SUBGÉNERO <i>Embecovirus</i>	ESPECIE <i>Human coronavirus 229E</i> <i>Human coronavirus HKU1</i>
		<i>Hibecovirus</i>	
		<i>Merbecovirus</i>	<i>Middle East respiratory syndrome-related coronavirus (MERS-CoV)</i>
		<i>Nobecovirus</i>	
		<i>Sarbecovirus</i>	<i>Severe acute respiratory syndrome-related coronavirus (SARS-CoV, SARS-CoV-2)</i>

Figura 2.2 *Taxonomía del SARS-CoV-2.*

Fuente: Ruiz y Jiménez, (2020)

2.1.3. Transmisión

El SARS-CoV-2 es un virus que afecta principalmente a las vías respiratorias, por lo que su transmisión es más efectiva por este medio. Con un período de incubación de 5,7-5,6 días y un promedio de 6,4 días del mismo modo, Lauer *et al.* (2020) estima valores aproximados con un tiempo de 5,1 días en un rango de 4,5- 5,8 días. Por tanto, su monitoreo es recomendable durante 14 días.

La OMS (2020) presentó una reseña donde se informa más acerca del panorama de contagio por Covid-19, señalando que la transmisión puede ser por contacto directo o indirecto de personas infectadas. Esto a través de saliva, secreciones o gotículas respiratorias expulsadas al momento de toser, estornudar, entre otras. Además, aclaran que la orina y heces de algunos de los pacientes había rastros de esta enfermedad, sin embargo, no se confirma un informe relativo de la transmisión.

Estudios realizados revelan que la actual transmisión y propagación del virus es por el contacto directo de una persona infectada incluida en aglomeraciones. La importancia de analizar los diferentes mecanismos de contraer el Covid-19, fue lo principal que ayudó a referir tips para protegerse y cuidar la salud de las familias y así frenar la propagación de esta enfermedad (Vargas et al., 2020).

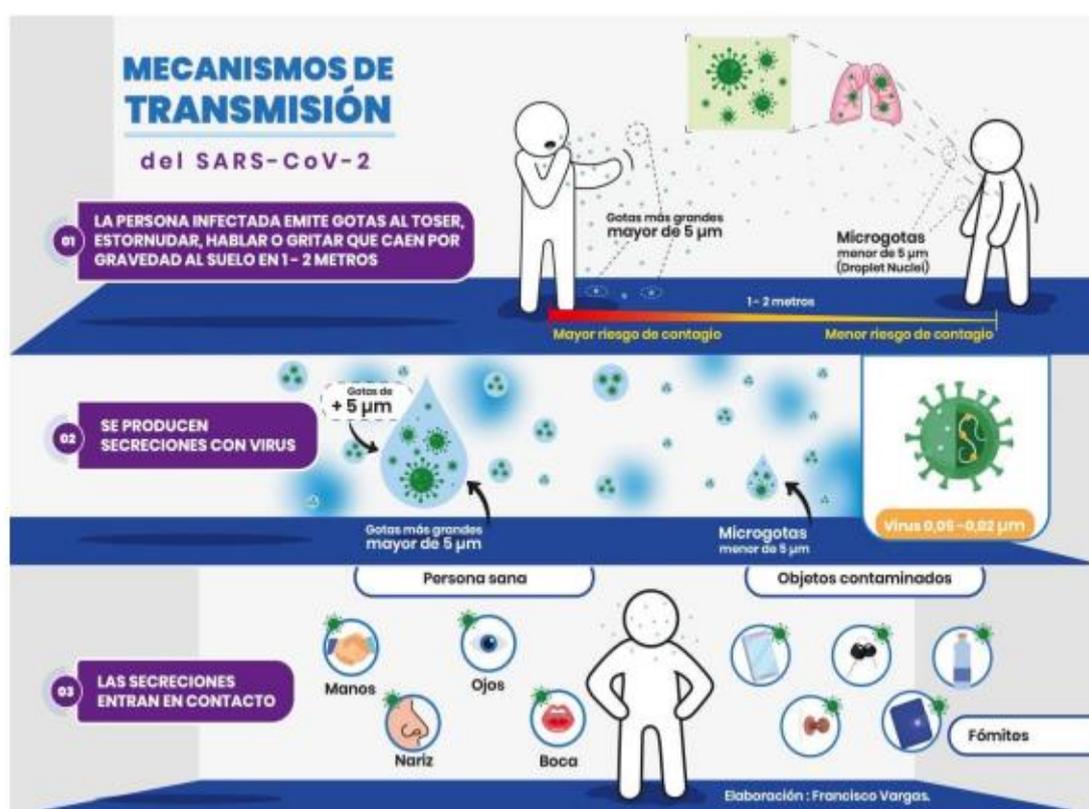


Figura 2.3 Mecanismos de transmisión del SARS-CoV-2

Fuente: Vargas, et al., (2020)

2.1.4. Variantes detectadas

El virus SARS-CoV-2 desde su aparición en el último mes del 2019, y la llegada a los diferentes puntos del mundo ha permitido que mute y genere diversas variantes las

cuales son más adaptadas al ambiente donde se encuentre, además son consideradas las más peligrosas, la caracterización de las diversas clases presentadas se observa en la Tabla 1 (Bedoya et al., 2021, p. 445).

Tabla 2.1

Mutaciones y variantes del SARS-CoV-2

Variante	País de Detección	Mutaciones en proteína S	Observaciones del fenotipo
Alfa (B.1.1.7)	Reino Unido Septiembre, 2020	D614G, N501Y, P681H, DEL69/70	30-90% más transmisibilidad 1,64 mayor mortalidad
Beta (B.1.351)	Sudáfrica Agosto, 2020	D614G, N501Y, E484K, K417T	Transmisibilidad y escape del sistema inmune incrementados
Gamma (P.1)	Brasil Diciembre, 2020	D614G, N501Y, E484K, K417T, L18F	1,4-2,2 más transmisibilidad Capaz de evadir la inmunidad protectora en 25-61%
Delta (B.1.617.2)	India Diciembre, 2020	D614G, P618R, L452R, T478K, T19R	Transmisibilidad severa Mayor escape a la inmunidad natural mayor al de la variante alfa
Eta (B.1.525)	Estados Unidos Diciembre, 2020	D614G, E484K, Q677H, DEL69/70	Posiblemente mayor escape al sistema inmune
Iota (B.1.526)	Estados Unidos Diciembre, 2020	D614G, E484K	Posiblemente mayor escape al sistema inmune
Lambda (C.37)	Perú Diciembre, 2020	D614G, L452Q, F490S, T859N	Posiblemente mayor escape al sistema inmune
Mu (B.1.621)	Colombia Enero, 2021	D614G, N501Y, P681H, E484K, R346K	Posiblemente mayor escape al sistema inmune

Fuente: Bedoya, *et al.*, (2021)

2.1.5. Parámetros clínicos

Los parámetros clínicos son utilizados para determinar y diagnosticar debidamente la infección de las enfermedades. En esta circunstancia es necesario contar con parámetros hematológicos, muestras de suero y el análisis sanguíneo, además se verifica la demografía con el fin de estudiar la evolución de esta enfermedad en el paciente (Sánchez et al., 2021).

2.1.5.1. Historias clínicas

Una historia clínica es un documento legal utilizado en el área de salud, donde abarca la información del paciente como residencia, parámetros clínicos, diagnósticos y entre otras. La elaboración de dicho documento debe ser profesional, con objetivo, lícito y una ejecución típica al momento de que el paciente es atendido (Guzmán & Arias, 2012).

La información que se recoge en las historias clínicas se guarda como antecedentes que representan la salud del paciente, diagnósticos, tratamientos, cirugía y entre otros. Además, se detalla que este documento es de uso exclusivo del centro de salud y además debe mantenerse bajo criterios de confidencialidad.

2.1.6. Factores de riesgo

Los factores de riesgo son aquellas enfermedades que ayudan a que la letalidad del coronavirus aumente. Muchas de estas se relacionan a la comorbilidad y enfermedades previas del paciente, es decir los factores de riesgo permiten cuantificar de forma más específica la vulnerabilidad de las personas (Herrero et al., 2020).

Un estudio realizado por Herrero et al., (2020) denotó como mayores factores de riesgo sociodemográficos a la edad (personas mayores a 60 años) y al sexo (mujeres en gestación), mientras que, de acuerdo a los aspectos clínicos fue el padecimiento de: diabetes, hipertensión, obesidad, enfermedades cardiovasculares, entre otros; y que el sector más vulnerable fue el sector sanitario con su personal de limpieza. Del mismo modo, Calixto *et al.* (2021) concuerdan en que la hipertensión y la diabetes fueron los problemas que más asistieron a la muerte temprana del paciente.

Otra de la comorbilidad asociada al riesgo de muerte de Covid-19, fue la hipertensión arterial, con la mayor tasa de letalidad registrada en América del Sur, del mismo modo, la obesidad, hipotiroidismo y dislipidemia. Estas fueron diagnosticadas luego de que los pacientes dieran positivo a coronavirus (Vintimilla, 2021).

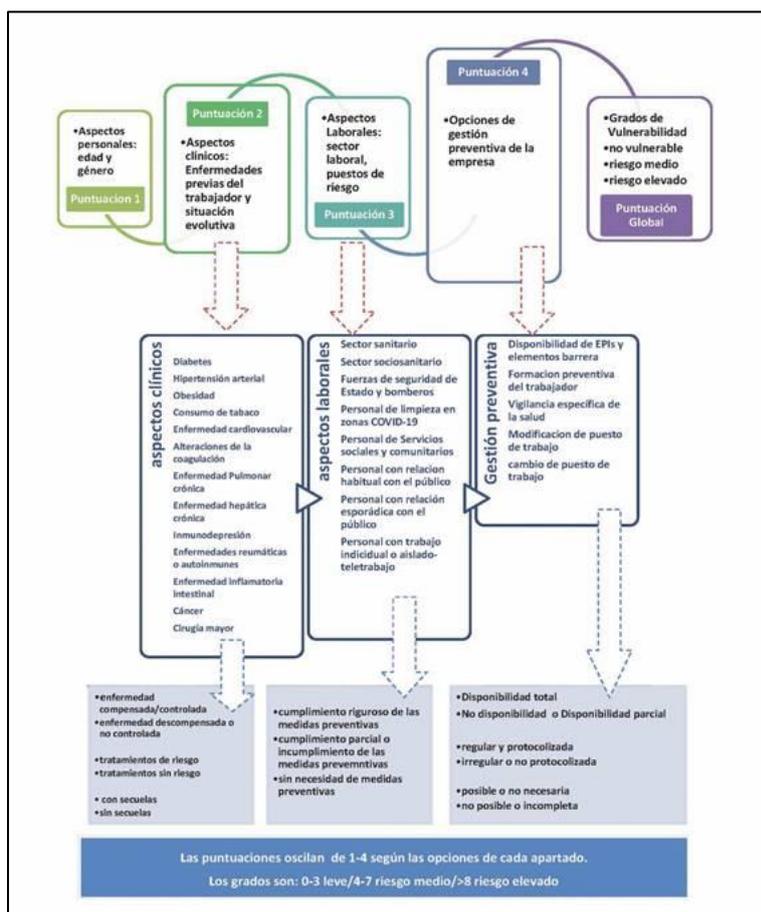


Figura 2.4 Protocolo de valoración de la vulnerabilidad del trabajador ante la infección del Covid-19

Fuente: Herrero et al., (2020)

La Figura 2.4 destaca el protocolo de valoración emitido por la Asociación española, reconocidos especialistas en medicina de trabajo, este protocolo ayudó a estratificar diversos riesgos de contagio bajo aspectos como: sociodemográficos, clínicos y gestión de la empresa. Los cuales toman valores puntuales empezando en 1 y si esta valoración sobrepasa al 8 significa que el riesgo de contagio es elevado.

El Instituto Nacional de Investigación y Salud Pública (INSPI) (2021) identificó un total de 3163 casos de variantes de Covid-19 en Ecuador y el 52,26% corresponden a las mutaciones (Delta, Gamma, Alpha, Ómicron) de mayor preocupación. De estas la predominante es la Delta con un total de 1134 casos positivos registrados, en las provincias de Guayas y Pichincha. Hasta septiembre del 2021, se registró un contagio comunitario en el cantón Milagro registrándose 8 nuevos casos de la variante Delta, aunque los pacientes no mostraron signos de letalidad por tanto no fueron hospitalizados y fueron puestos en cuarenta en sus respectivos hogares.

Tabla 2.2

Variantes de Covid-19 detectadas en Ecuador

Variante	Observaciones de transmisibilidad y mortalidad
Alfa (B.1.1.7)	30-90% más transmisibilidad 1,64 mayor mortalidad
Beta (B.1.351)	Transmisibilidad escape del sistema inmune incrementados
Gamma (P.1)	1,4-2,2 más transmisibilidad Capaz de evadir la inmunidad protectora en 25-61%
Delta (B.1.617.2)	Transmisibilidad severa Mayor escape a la inmunidad natural mayor al de la variante alfa

Fuente: Bedoya, *et al.*, (2021); INSPI, (2021)

2.1.7. Análisis Multivariante en casos Covid-19

Desde el origen de la pandemia en Wuhan, China; y debido a su rápida propagación a nivel mundial, cada uno de los países han empezado a desarrollar investigaciones referentes a factores de riesgo, pronósticos de contagio, letalidad y evolución. Por lo que varios científicos buscan desarrollar una cura que garantice la libre circulación de las personas. Por ello en este apartado se detalla las investigaciones más relevantes realizadas por distintos autores, en diversas localidades.

Reyes *et al.* (2020) disertan que el origen del Covid-19 es de los murciélagos, y reiteran la peligrosidad, transmisibilidad y adaptabilidad que tiene este virus. Por tanto, se propuso desarrollar un análisis de la letalidad de esta nueva enfermedad mediante un estudio descriptivo en una muestra de pacientes con problemas de demencia neurodegenerativas desarrollado en el Instituto de Investigación Biomédica de Málaga. Comparando las características principales entre el grupo control y pacientes con demencia. Tomando la letalidad del virus como variable dependiente se aplica una regresión logística en donde sus resultados muestran que la diabetes mellitus, hipertensión arterial y la demencia actúan de manera independiente, sin embargo, aquellas variables de las cuales depende la letalidad fueron la enfermedad pulmonar, cardiopatía, sexo y la edad del paciente.

Heras *et al.* (2020) dicen que la prioridad es contar con marcadores de gravedad del Covid-19, por tanto, la investigación trata de buscar factores relacionados con la gravedad de esta enfermedad, para ello se tuvo una data con características demográficas, laboratorio y resúmenes clínicos, para empezar, estimaron un modelo de regresión logística binaria con una muestra de 133 pacientes. Los resultados mostraron un aumento en el factor de riesgo de la enfermedad con lactato deshidrogenasa (LDH), aspartato aminotransferasa (AST) y proteína C reactiva (PCR), así mismo la disminución de las cifras de plaquetas, obesidad y personas que tengan problemas de diabetes mellitus tipo 2 son las más propensas a sufrir un fallecimiento. De acuerdo con los resultados ratifican la importancia de contar con datos clínicos, lo que permite predecir de mejor manera los diagnósticos y riesgos de contagio.

García *et al.* (2021) mencionan que la llegada del Covid-19 ha provocado una gran demanda de recursos hospitalarios y el colapso del sistema de salud, por lo que afirman que elaborar pronósticos de propagación, mortalidad y diagnóstico acerca de la enfermedad permitirá frenar la propagación y cuidar la salud de los pobladores. Por tanto, elaboran una investigación que determine los factores de riesgo de mortalidad mediante un estudio transversal descriptivo de pacientes que acudieron entre el 8 de marzo y el 7 de abril del año 2020 al hospital Universitario de Jaén. De modo que para identificar las variables o características que se relacionen con el deceso del paciente empezaron aplicando la regresión de Cox univariada lo que permitió observar la representatividad de la variable respecto al factor riesgo que conlleva a la muerte del paciente. Seguidamente se eligieron aquellas características con un p valor inferior a 0,2. Para los pronósticos se desarrolló un modelo de regresión logística de mortalidad con tiempo límite de 30 días y la validación se realizó con nuevos pacientes, de acuerdo con la metodología TRIPOD sus resultados fueron favorables con una capacidad de pronóstico del 88,6%.

Un estudio realizado en Madrid por Pérez *et al.* (2021) donde realizan un estudio de los factores de riesgo aplicando técnicas multivariadas, para ello seleccionan los factores mediante una investigación bibliográfica, consecutivamente se pretende reducir la dimensionalidad de la data aplicando un análisis de componentes principales. Donde su variabilidad total explicada fue del 71%. El número de componentes fue base para poder aplicar un análisis de conglomerados aplicando el método k-means, lo que permitió agrupar diversas variables en nuevos factores de riesgo. Finalmente, en los resultados de la regresión logística, el factor territorio evidenció una significancia y relación positiva con el número de contagios, mientras que, la componente socioeconómica fue representada de mejor manera por la contaminación y el clima. De los resultados se evidenció un buen

ajuste del modelo y respecto a los factores de riesgo que no fueron significativos, aclararon que pudo ser provocado por la baja potencia del muestreo.

En un estudio realizado por Devkota (2021), recoge datos del Ministerio de Salud y población de la gobernación de Nepal. La data cuenta con información de nuevos casos, muertes diarias, recuperaciones diarias, total de casos, recuperaciones, infectados los cuales son aislados y muertes. La muestra es recolectada respecto a la atención realizada desde el 24 de mayo al 2 de septiembre de 2020, la metodología de procesamiento contó con la descripción, modelos de regresión y componentes principales. Sus resultados importantes muestran que aplicando el modelo ARIMA se tuvo resultados óptimos ya que se adapta de manera clara a la tendencia con una aproximación del 99,99936% para el modelo 1 y 99,99945% para el modelo 2. Al aplicar las componentes principales se obtuvo una reducción de la dimensionalidad con una variabilidad explicada del 74,579%. Lo cual concluye que los resultados son recomendables tanto en los modelos como en las componentes y recalcan que estos resultados pueden variar ya que muchas de las personas no tienen recursos para acercarse a un centro de salud, por lo que consideran a la CP1 como un indicador de datos precisos por las altas puntuaciones factoriales en las variables de nuevos casos, muertes diarias, recuperados e infectados.

Llanos y Bermúdez (2020) narran acerca del gran problema mundial que se vive a causa de la pandemia y reiteran la necesidad de estudiar el efecto, diagnósticos y pronósticos según el avance de este problema. Para ello proponen una investigación descriptiva exploratoria con información recolectada y certificada por Sisbén, boletas epidemiológicas de la secretaría de salud, para determinar factores socioeconómicos y variables con posible relación a la propagación de esta enfermedad en la zona sur de Colombia, en el municipio de Pitalito, en este caso se contó con un total de 135998

registros, de los cuales se estimó una muestra de 2595. Aplicando la metodología de los coeficientes de correlación lineal de Pearson y correlaciones parciales se determinan variables con aporte al modelo planteado, seguidamente estiman un vector de coeficientes β . Mediante la aplicación del análisis de varianza ANOVA, se corroboró la significancia de las variables y su efecto con su respectivo intervalo de confianza. La parte espacial se desarrolló en ArcMap. Sus resultados mostraron que Pitalito presentó mayor cantidad de casos positivos de Covid-19, mientras que Bruselas con mayor población tiene una tasa de contagios baja, la parte urbana de Pitalito representa el 18% de personas en la parte urbana mientras que Bruselas cuenta con el 45% en la parte de la ciudad. Uno de los aspectos a considerar fue que en Pitalito los contagios representan mayor cantidad en la parte rural y esto provoca mayor propagación ya que la gente del campo acude a las ciudades a vender o comprar víveres.

Medina *et al.* (2022) es uno de los estudios más recientes, donde analizan la evolución de la pandemia en países americanos mediante observaciones exploratorias y la búsqueda de asociaciones del Covid-19 con las diversas variables. Los resultados mostraron que Canadá, Brasil y Estados Unidos invierten al menos un 10% de su Producto Interno Bruto (PIB) en salud y tienen un aproximado de 780 profesionales del área médica por cada 100 mil habitantes, del mismo modo el plan de vacunación efectuado en varios países tiene incidencia en la reducción de casos positivos y las muertes acumuladas de esto se pudo apreciar que Ecuador, México y Perú tienen una mayor tasa de letalidad. El análisis clúster determinó que Brasil, Argentina, Chile, Canadá, Colombia, Cuba, Costa Rica, Panamá, Uruguay y Estados Unidos tienen una tasa de vacunación aproximada del 60% de la población total.

En la actualidad se cuenta con investigaciones relevantes a nivel Continental y mundial, sin embargo, la revisión bibliográfica ha constatado la falta de estudios

relacionados a detección de factores de riesgo, variables asociadas a la evolución y avance del Covid-19 en el país, más aún a nivel provincial y cantonal. Vivas y Vivas (2021) realizaron una breve descripción poniendo como herramienta potente el software Power BI y RStudio, con información recabada por el Ministerio de Salud Pública del Ecuador donde se aplicó el modelo de regresión logística y de Gompertz en la cual sus resultados muestran un ajuste significativo para las provincias de Pichincha, Santo Domingo, Manabí y los Ríos. En cambio, para Guayas no se obtuvo un ajuste deseado por tanto no se podrá analizar casos futuros o predicciones para esta provincia, además los ajustes se asemejan en los dos modelos. La observación desarrollada según la tasa de letalidad fue alta para todo el Ecuador, Santa Elena y Chimborazo con 19,76% y 12,15% respectivamente, mostraron una mayor letalidad, mientras que Morona Santiago y Galápagos mostraron letalidad de 0,74% y 0,39%. Por último, los autores recalcan la necesidad de obtener datos accesibles con el fin de garantizar resultados que aproximen de manera más eficiente a la realidad que vive el país.

2.2. Métodos Estadísticos Multivariantes

2.2.1. Análisis de Conglomerados

El análisis de conglomerados o también llamado análisis clúster tiene por objeto aglomerar o agrupar elementos en grupos homogéneos en función de las similitudes o similaridades que pueda existir entre la información obtenida. Normalmente se agrupan las observaciones, sin embargo, el análisis de conglomerados puede también aplicarse para agrupar variables o características observadas de la unidad de análisis. Estos métodos son conocidos bajo el enfoque de clasificación automática, no supervisada o de reconocimiento de patrones sin supervisión.

2.2.1.1. Partición de los datos

Al momento de observar la matriz de datos formada disponemos de datos de la cuales dudamos su homogeneidad. Bajo este criterio se propone dividir en un número de grupos prefijado, con el fin de:

- Cada elemento pertenezca a uno y solo uno de los grupos.
- Todo elemento quede clasificado.
- Cada grupo sea internamente homogéneo.

2.2.1.2. Construcción de jerarquías

En la construcción de jerarquías se desea estructurar los elementos de un conjunto de forma ordenada por su similitud. Este tipo de clasificación es muy frecuente en biología, al clasificar animales, plantas etc.

Estos métodos no definen grupos, sino la estructura de asociación en cadena que pueda existir entre los elementos. Sin embargo, como veremos, la jerarquía construida permite obtener también una partición de los datos en grupos.

2.2.1.3. Clasificación de variables

Los métodos de partición utilizan la matriz de datos, pero los algoritmos jerárquicos utilizan la matriz de distancias o similitudes entre elementos. Para agrupar variables se parte de la matriz de relación entre variables: para variables continuas suele ser la matriz de

correlación, y para variables discretas, se construye, como veremos, a partir de la distancia ji-cuadrado. Se representa en primer lugar los métodos de partición.

2.2.2. Métodos Clásicos De Partición

2.2.2.1. Fundamentos del algoritmo de k-medias

Al tener una muestra de n elementos con p variables. El objetivo es dividir esta muestra en un número de grupos prefijado, que se denotará como G . El algoritmo de k-medias requiere el desarrollo de las siguientes etapas:

1. Seleccionar G puntos como centros de los grupos iniciales. Esto puede hacerse.
2. Asignando aleatoriamente los objetos a los grupos y tomando los centros de los grupos así formados.
3. Tomando como centros los G puntos más alejados entre sí.
4. Construyen los grupos con información a priori, o bien seleccionando los centros a priori.
5. Calcular las distancias euclídeas de cada elemento al centro de los G grupos, y asignar cada elemento al grupo más próximo. La asignación se realiza secuencialmente y al introducir un nuevo elemento en un grupo se recalculan las coordenadas de la nueva media de grupo.
6. Definir un criterio de optimalidad y comprobar si reasignando uno a uno cada elemento de un grupo a otro, mejora el criterio.
7. Si no es posible mejorar el criterio de optimalidad, terminar el proceso.

2.2.3. Análisis de correspondencia Múltiple

El análisis de correspondencia múltiple (ACM) es una extensión del análisis de componentes principales, la diferencia radica en que se puede utilizar variables de tipo cualitativo, estas generan particiones de individuos y características, de este modo forma homogeneidad dentro del grupo y heterogeneidad entre cada conjunto.

De acuerdo a lo anterior el objetivo principal del ACM es:

1. Describir asociaciones entre categorías de las variables.
2. Descubrir patrones de individuos, que determine grupos.
3. Sintetizar la información en planos factoriales que visualicen los posibles grupos.
4. Explorar la explicación de la estructura inducida por las categorías.

2.2.3.1. Matriz binaria

En este caso se dispone de K variables categóricas, se tiene una matriz Z de orden $n \times J$, donde J denota el número de categorías de cada una de las variables, n denota el número de casos o registros. La matriz binaria se forma de la unión de k variables binarias, de modo que sus valores son (0,1) y 1 representa la acción de la característica.

$$J = \sum_{k=1}^K J_k \text{ y } Z = (Z_1, Z_2, \dots, Z_Q)$$

El valor del estadístico Chi cuadrado asociado para el test de independencia usando la matriz binaria Z se denota como:

$$X^2 = n(J - Q)$$

2.2.3.2. Matriz de Burt

Uno de las alternativas dentro del análisis de correspondencia es usar la matriz de Burt en reemplazo de la matriz Binaria, esta se conforma de tablas de contingencia N_{qs} , esta resulta de cruzar las variables de par en par con $q \neq s$. Además, esta contiene su diagonal D_q la cual contiene las frecuencias marginales de las diversas categorías cruzadas.

$$B = Z^T Z = (Z_1, Z_2, \dots, Z_Q)^T (Z_1, Z_2, \dots, Z_Q)$$

De mismo modo para el test de independencia Chi Cuadrado sobre la matriz de Burt está definido como:

$$X^2(B) = \sum \sum_{(q \neq s)} X_{qs}^2 + n(J - Q)$$

2.2.3.3. Identificación del número correcto de dimensiones

En este tipo de análisis surge la interrogante de ¿Cuántas dimensiones considerar?, Camiz y Gomez (2012) arguyen que esta incógnita puede ser resuelta empleando la explicación que sugirió Greenacre (1998) y Benzécri (1978), quienes indican que se debe usar los autovalores mayores que su media $1/Q$. Para calcular los autovalores de la matriz de Burt se debe extraer la raíz cuadrada a los autovalores de la matriz binaria y se escogen estas se cumplen.

$$\sqrt{\lambda_k} > \frac{1}{Q}$$

$$S_\lambda^2 = \frac{1}{k} \sum_{k=1}^K \left(\lambda_k - \frac{1}{Q} \right)^2$$

Se estima σ^2 como:

$$\sigma^2 = E[S_\lambda^2] = \frac{1}{nQ^2(J-Q)} \sum_{i \neq j} (J_i - 1)(J_j - 1)$$

De este modo se presenta la expresión $\frac{1}{Q} \pm 2\sigma$, la cual debe contener aproximadamente el 95% de los autovalores.

2.2.4. Modelo logit multinomial

La regresión logística multinomial considera que la variable dependiente tiene tres o más categorías y se asume que esta tiene una distribución multinomial, este tipo de modelo es la extensión de la regresión logit que considera dos categorías en Y.

Se tiene una variable dependiente Y con J categorías cada una con su respectiva probabilidad. Normalmente este modelo se desarrolla planteando una de las categorías como respuesta base. Esta puede ser cualquiera de las clases que esté dentro de J, de acuerdo a esto se plantea el modelo.

$$\log \left(\frac{\pi_j}{\pi_J} \right) = \alpha_j + \sum_{k=1}^K \beta_{jk} X_{jk}$$

Donde $j = 1, 2, \dots, J - 1$, y el modelo tiene $J-1$ ecuaciones, cada uno con su propio parámetro y estas representan el efecto respecto a la categoría base o de referencia y una vez calculado estos parámetros se procede a estimar las probabilidades.

$$\pi_j = \frac{e^{\alpha_j + \sum_{i=1}^n \beta_j X_i}}{1 + \sum_{h=1}^{J-1} e^{\alpha_h + \sum_{i=1}^n \beta_h X_i}}; \quad j = 1, 2, \dots, J$$

CAPÍTULO 3

3. METODOLOGÍA

3.1. Tipo de enfoque

La siguiente investigación se trabajó bajo un enfoque primordialmente cuantitativo, usando como base la aplicación de la Estadística, lo que permitió realizar la descripción de cada una de las variables - características en estudio con precisión y enfocado a la realidad que vivió la ciudad de Milagro durante el año 2021. Además, el estudio pretende utilizar la información a fin de que sea replicable a nivel nacional.

3.2. Diseño

Este trabajo de investigación tiene un diseño de tipo exploratorio y descriptivo, no experimental, dado a que no se manipuló ninguna de las variables - características, en el cual se realizó un análisis detallado, exhaustivo de las variables relacionadas con el tema del Covid-19.

3.2.1. Investigación Exploratoria

El trabajo de investigación será de nivel exploratorio debido a que el Covid-19 es una enfermedad manifestada en Wuhan y que se diagnosticó en Ecuador en febrero del 2020. En la actualidad el caso es poco estudiado en la ciudad de Milagro y ha inducido a realizar investigaciones que ayuden a entender la evolución de este problema, por tanto, el estudio se centra en proporcionar conocimiento para futuras investigaciones.

3.2.2. Investigación Descriptiva

El estudio se considera de tipo descriptivo ya que cada una de las variables serán caracterizadas según su tipo de dato, además permitirá analizar las posibles causas que provocan el rápido avance de la enfermedad dentro de la ciudad de Milagro. Bajo esta perspectiva se recurrirá al uso de gráficos que muestren de manera más precisa la representación de cada variable.

3.3. Alcance

Esta investigación pretende mostrar la relación de las distintas variables con el contagio de Covid-19, mediante la aplicación de técnicas Estadísticas univariantes y multivariantes. Estos resultados podrán ser utilizados en futuras investigaciones y podrá ser replicado a nivel cantonal, provincial o nacional.

3.4. Población

La población de estudio son los habitantes de la ciudad de Milagro que fueron atendidos en las áreas hospitalarias y que fueron contagiados o no con el virus del coronavirus en el año 2021.

3.5. Muestra

El sistema de salud de la Coordinación de Salud Zona 5 proporcionó una data con un total de 11274 registros, la población se caracteriza por los pacientes que acudieron al

centro de salud más cercano. La información es filtrada para aquellos con posibles problemas de Covid-19. De este modo se obtiene un total de 2483 registros y 54 variables, además se considera omitir los valores faltantes pues son referidos a la comorbilidad, y los resultados que se obtienen en laboratorios, ya que si se imputa estos datos se puede generar sesgos en los resultados.

3.6. Definición de variables

Las variables de estudio son de tipo categórico dicotómico y politómico, estas se detallan a continuación.

Tabla 3.3

Detalle de variables utilizadas para el análisis

VARIABLES	DETALLE
Institución	Institución a la que pertenece
Establecimiento de Salud	Nombre del centro de salud
Nivel de atención	Nivel de atención según las instalaciones
Zona Establecimiento	Zona donde está ubicado el centro de salud
Distrito Establecimiento	Zona donde está ubicado el centro de salud
Provincia Establecimiento	Zona donde está ubicado el centro de salud
Condición	Paciente ingresa vivo o muerto
Estado Civil	Estado civil del paciente
País nacionalidad	País de origen del paciente
Sexo	Sexo del paciente
Edad	Edad del paciente
Grupo Edad	Grupo de edad al que pertenece el paciente
Ocupación	Ocupación del paciente
Tiene antecedentes pruebas anteriores	Indica si el paciente se realizó pruebas anteriores por posible Covid-19
Tipo prueba anterior	El tipo de prueba anterior aplicado
Presentó síntomas	Indica si el paciente mostró síntomas

Síntomas que presenta	Indica que tipo de síntomas presentó
Semana Epidemiológica	Semana en la cual fue ingresado el paciente
Diagnóstico Inicial	Diagnostico breve al ingreso
Existe comorbilidad	Indica si el paciente tuvo otros problemas
Comorbilidad	Indica que tipo de problema se presentó
Segunda Comorbilidad	Indica si hubo más problemas de enfermedad
Tercera Comorbilidad	
Antecedente Viaje	Muestra si el paciente tuvo viajes
Lugar de viaje	Indica el lugar de destino
Embarazada	Presencia del embarazo
Semanas Gestación	Semana de gestación de la embarazada
Muestra	Muestra que se toma para las pruebas
Resultado lab 1	Resultado de los exámenes
Adecuada m1	Indica si la muestra tomada es adecuada
Agente 1	Método de extracción de muestra aplicado
Resultado lab 2	
Resultado lab 3	Resultado de los exámenes
Resultado lab 4	
Lugar de atención	Indica donde fue atendido el paciente
Condición final	Indica si el paciente falleció o sigue con vida
Clasificación final del caso	Clasificación según los resultados de exámenes
Diagnostico final	Diagnóstico de las pruebas realizadas si es coronavirus u otra enfermedad
Estado del registro	Muestra registros de personas que no terminaron sus análisis
Tipo de prueba	Tipos de pruebas aplicadas
Ensayo	Ensayos aplicados hasta el diagnostico

3.7. Métodos o instrumentos de recolección de información.

En este estudio se utilizó la información proporcionada por el Sistema de Salud de la Coordinación de Salud Zona 5, organismo público que se encarga de registrar los casos de cada uno de los pacientes que asistieron a la unidad de salud durante el año 2021.

3.8. Procedimiento o fases de la investigación

3.8.1. Pre-tratamiento de la información

Con la información proporcionada por el Sistema de Salud de la Coordinación de Salud Zona 5, se realizó el procesamiento de la misma con el fin de explorar la data y visualizar su estructura, cantidad de variables, individuos, valores perdidos o faltantes y se verificó la correcta tabulación de la data. Antes de empezar con el análisis descriptivo de los datos, primero se verificó la presencia del dato o las características tomadas de cada individuo. La falta de información o vacíos hacen referencia a los valores faltantes que se denotan en R como NA. Para empezar, se propuso seleccionar variables óptimas que ayuden a describir de manera clara y precisa factores relacionados al Covid-19 mediante un pre-procesamiento de datos, lo cual permitió determinar la necesidad de aplicar un relleno de dato o la exclusión del registro de la base de datos, de esta manera se garantiza la confiabilidad de los resultados.

3.8.2. Análisis Descriptivo

Se realizó un análisis descriptivo en el que se exploró el comportamiento de cada una de las variables. Al tener en su gran mayoría variables del tipo cualitativo, se utilizaron las frecuencias absolutas y relativas porcentuales para analizar la distribución de los datos. Se dio soporte a estos resultados con la ayuda de los gráficos de barras y de series.

De acuerdo al pre-procesamiento de la data, se realizó una descripción de las variables cualitativas, de acuerdo a la edad, sexo y otras características.

3.8.3. Análisis Relacional

Para identificar la asociación entre pares de variables se usó el test de independencia Chi cuadrado. El cual plantea en su hipótesis nula (H_0) la no dependencia entre las variables, esta prueba se basa en las tablas cruzadas o de doble entrada con dimensión 2x2 en adelante. Si estas tablas tienen una categoría inferior a 5 se usa la corrección de Yates. Y para cuantificar la asociación se utilizó el coeficiente V de Cramer, el cual considera una relación aceptable con valores superiores de 0,3.

Mediante el cálculo del p valor se determina si se rechaza o no la hipótesis nula. Para ello se aplicó la siguiente decisión, si valor $p < \alpha$ se rechaza H_0 y se asume la dependencia del diagnóstico final con las variables independientes.

Test de independencia Chi cuadrado

H_0 : Las variables X e Y son independientes

H_1 : Las variables X e Y son dependientes

$$\chi_{obs}^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$$
$$e_{ij} = \frac{\text{total filas} * \text{Total columna}}{\text{total}}$$

La decisión final es de acuerdo a la condición: Si valor $p < \alpha = 0,05$ se rechaza H_0 .

Tabla 3.4

Valores del índice V de Cramer

Valores del Índice V de Cramer	Criterio
0-0,10	El grado de efecto es despreciable
0,10-0,30	El efecto es pequeño
0,30-0,50	Considera el efecto moderado
0,50-1	El efecto es grande

Fuente: Betancourt y Caviedes (2018)

3.8.4. Análisis Multivariante

Una vez realizado el estudio de asociación y relaciones entre las distintas características con el diagnóstico final, para la identificación de posibles factores de riesgo se determinaron las principales variables relacionadas con el evento de contagio por Covid-19. Se aplicó un Análisis de Correspondencia Múltiple y un modelo logit multinomial, que permitió validar las asociaciones entre pares de variables con el diagnóstico final y las variables independientes que se consideran representativas.

La validación del modelo propuesto se realizó a través de pruebas de los coeficientes de Wald resultante y mediante el cálculo del valor estadístico (p) con una prueba de significancia de 0.05. La variabilidad explicada del modelo será mostrada mediante el cálculo del Pseudo R^2 de MacFadden, y la tasa de clasificación correcta e incorrecta, de esta manera se pudo determinar los diversos factores de contagio.

CAPÍTULO 4

4. RESULTADOS

4.1. Procesamiento de la información

La base de datos original cuenta con un total de 11 274 registros o historias clínicas y con 195 características o variables que fueron recolectadas durante el año 2021. La distribución de la información, mostró de color celeste a la presencia de dato o información, sin embargo, se observó variables que no tienen información lo cual se pude visualizar de color naranja, por lo que se consideró excluirlas del estudio ya que no cuentan con ninguna información relevante para el análisis (Fig. 4.5).

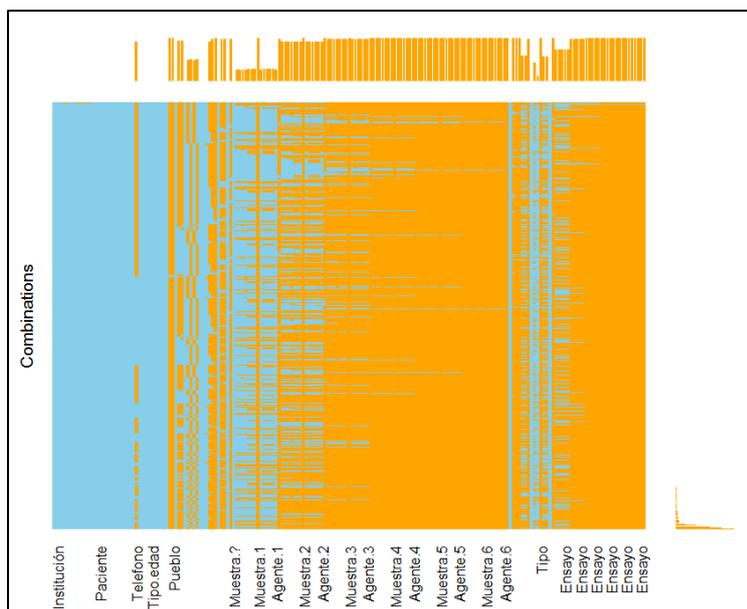


Figura 4.5 Distribución de la información en la data

Realizado por: Villamar Isabel, 2022

Se pudo determinar un total de 7 variables sin información, por tanto, se decidió excluir de la base de datos. Seguidamente se planteó eliminar aquellas variables con información personal del paciente, doctor que lo entendió, dirección, nombres de

laboratorio, entre otros, esto con el fin de resguardar la integridad de pacientes y profesionales. Seguidamente se verificó la existencia de variables que contengan información repetitiva de las características observadas de las personas.

A manera de ejemplo se obtuvieron 4 clases de la variable edad: "Edad" en números, "Tipo edad" medida de la edad en años, "Edad completa" fue medida en años y meses, finalmente "Grupo Edad" que se dividió en 8 clases de diferentes rangos. En este caso se consideró usar la variable "Grupo Edad" y excluir las demás, ya que solamente fue información repetida con diferente escala de medida.

4.2. Estadísticos descriptivos

4.2.1. Distribución sociodemográfica

4.2.1.1. Distribución biológica

Sexo: Se obtuvo una muestra de 2483 personas y los resultados mostraron que, la mayoría de pacientes que acudió a un centro de salud en la ciudad de Milagro eran mujeres que representó el 52,72%. Por otro lado, se pudo apreciar que la proporción de hombres que fueron registrados es el 47,28%. De modo que durante la pandemia en el año 2021 existió más peligro para las personas del sexo femenino (Fig. 4.6 A).

Edad: La distribución de los pacientes según grupos de edad, mostró una clara evidencia de que las personas en edad activa fueron las más propensas al contagio por Covid-19, lo que representó el 54,45%. Quienes estaban entre las edades de 50 hasta los 64 años de edad constituyeron el 21,91%, mientras que, el grupo de la tercera edad incorporó el 13,81%. La proporción de bebés, niños y jóvenes

permanecieron por debajo del 5%, lo que indicó que este grupo de individuos se expuso ante la enfermedad en muy pocas proporciones (Fig. 4.6 B).

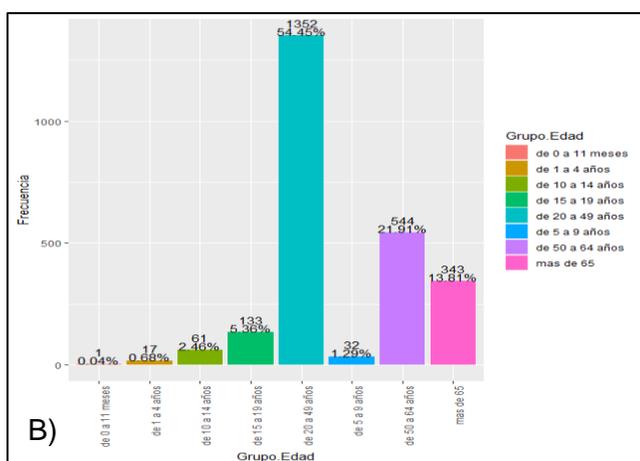
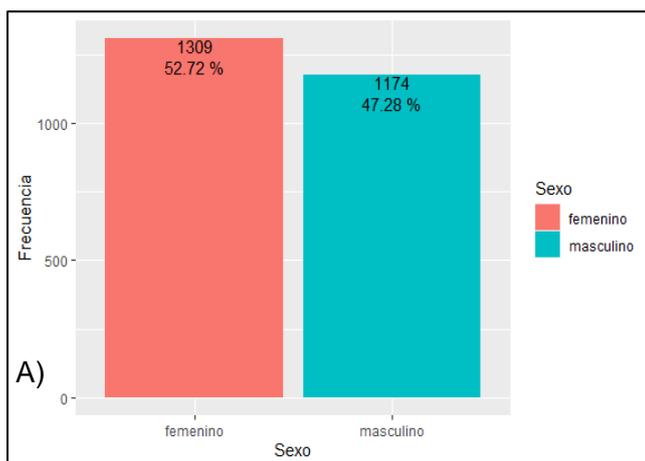


Figura 4.6 Distribución biológica de pacientes según Sexo y Grupo Edad.

Realizado por: Villamar Isabel, 2022

4.2.1.2. Distribución Social

Nacionalidad: El 96,34% de las personas atendidas en el centro de salud de la ciudad de Milagro fueron ecuatorianas, sin embargo, hubo personas extranjeras que estuvieron de viaje o viviendo en la ciudad, el 1,17% eran de Estados Unidos, mientras que, los pacientes nativos de Bolivia, Brasil, Chile, Colombia y otros países estuvieron por debajo del 1%. De acuerdo a estos datos se pudo constatar que

Milagro es uno de los puntos elegidos por individuos de otras nacionalidades, ya sea a manera de turistas o como residentes (Fig. 4.7 A).

Estado civil: El análisis del estado civil mostró que las personas solteras y casadas son las más propensas a ingresar a un centro de salud por problemas de Covid-19 lo que representa un 48,25% y 36,25% respectivamente. Las personas divorciadas, viudas y en unión de hecho fueron atendidos en una baja proporción. De lo observado se pudo concluir que las personas con obligaciones o cargos familiares son las que más han asistido al centro de salud (Fig. 4.7 B).

Autoidentificación: El 96,34% de la muestra se identificaron como mestizos de esto pudimos afirmar que la mayoría de sus pobladores son oriundos de la ciudad local, sin embargo, 2 de ellos indicaron que según su naturaleza no necesitan ser identificados. Por otro lado, el 3,85% (89) señalaron que no saben o simplemente no desearon responder a dicha pregunta (Fig. 4.7 C).

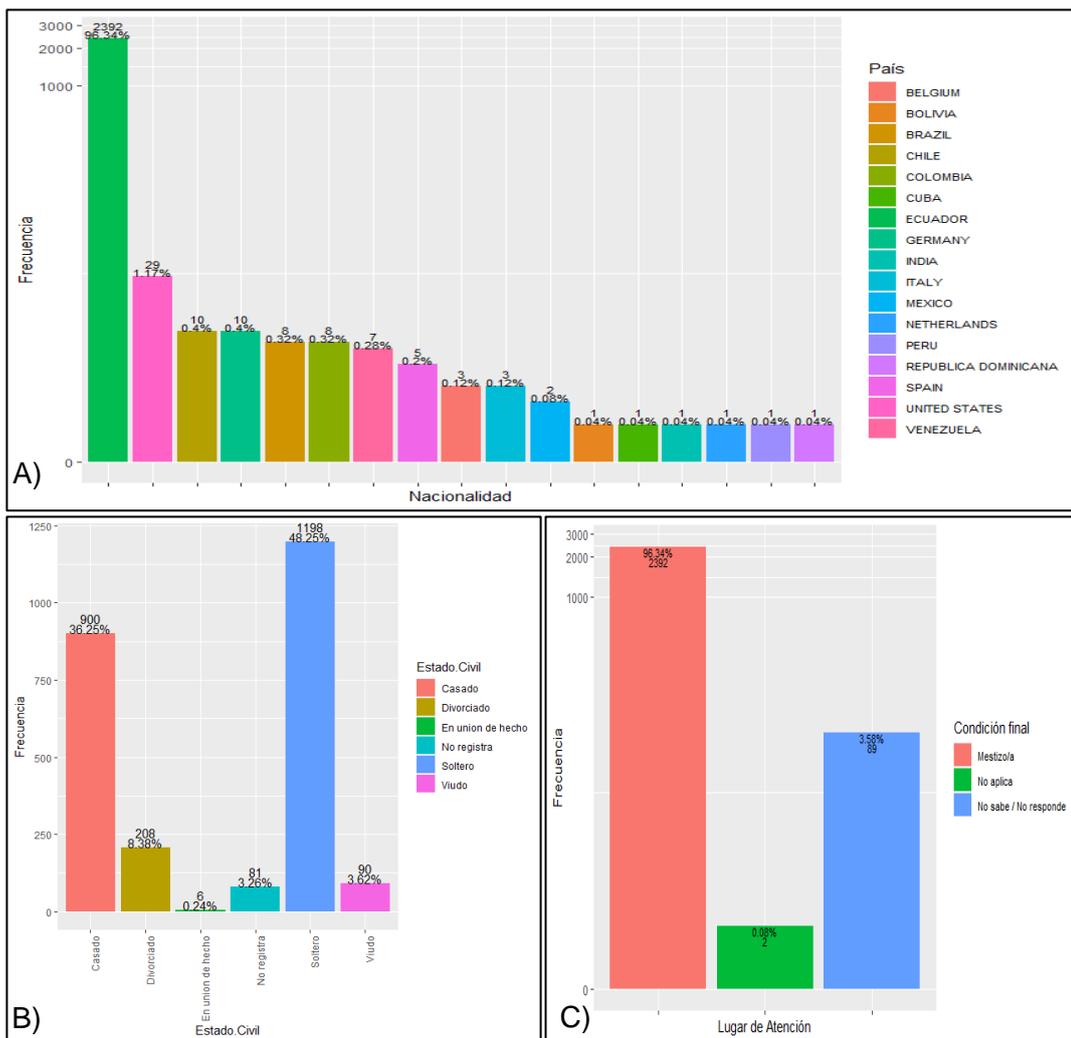


Figura 4.7. Distribución social del paciente en la ciudad de Milagro

Realizado por: Villamar Isabel, 2022

4.2.2. Distribución según el nivel de atención en hospitales y laboratorios.

Institución: se caracterizó la proporción de pacientes atendidos por los distintos centros de salud la cual se divide en diferentes niveles de atención. El sector privado atendió una proporción del 73,98% de personas, lo cual indicó que los enfermos prefirieron ser atendidos por personal médico del sector privado pues en el IESS y MSP tenían las camas llenas. Por otra parte, el IESS y el MSP acogieron una proporción de 15,71% y 10,31% respectivamente (Fig. 4.8 A).

Nombre de centros: Las muestras tomadas de los pacientes en el Laboratorio Clínico INTERLAB fue del 72,94%, en UNILAB el 1,05%. Mientras que, el 15,71% acudió al Hospital General Milagro y el 10,31% al centro Los Monos para su respectiva toma de muestra. De acuerdo a estos datos se observó que INTERLAB fue donde hubo mayor atención de muestras ya que para estas fechas el sector público no contaba con suficientes insumos (Fig. 4.8 B).

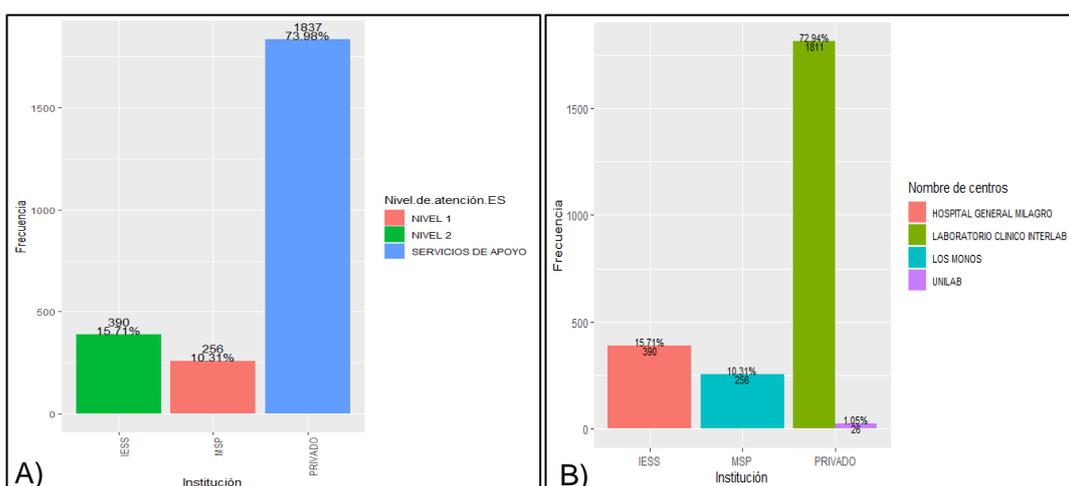


Figura 4.8 Distribución según el nivel de atención en hospitales y laboratorios

Realizado por: Villamar Isabel, 2022

4.2.3. Distribución según la condición y diagnóstico del paciente

Diagnóstico final de paciente: El diagnóstico final de los pacientes que acudieron a los diversos centros de salud indicó que el 54,34% no presentó ningún problema, el 24,89% de personas fueron diagnosticadas con la enfermedad de Covid-19. Por otro lado, quienes presentaron problemas de rinofaringitis aguda y neumonía fue del 14,38% y 7,93% respectivamente. De los resultados se rescató la letalidad en la cuarta parte de los enfermos con pronóstico de coronavirus (Fig. 4.9 A).

Condición final: El estudio tomó en cuenta aquellos pacientes quienes ingresaron por posible contagio de Covid-19, donde se pudo apreciar que el 1,49% registró defunciones con problemas de coronavirus, mientras que, aquellos sin seguimiento el 14,46% estuvieron con sospecha, el 10,63% dio positivo y el 36,09% fueron descartados. Por otro lado, aquellos pacientes que aún seguían vivos, el 12,77% fueron casos confirmados, el 24,33% fueron excluidos y el 0,24% estuvieron con probable caso de la enfermedad (Fig. 4.9 B).

Comorbilidad: Del total de la muestra hubo solamente un paciente presentó comorbilidad de hipertensión esencial (primaria) y se la conoce como hipertensión arterial alta sin causa reconocible, se da cuando supera 140/90 mmHg. Mientras que 2482 personas no mostraron o generaron este tipo de situación (Fig. 4.9 C).

Lugar de atención: El 32,42% fue atendido de manera ambulatoria (no necesitó ingreso hospitalario) ya que la atención que recibió el paciente se realizó en pocas horas. Mientras que el 0,12% fue intervenido en su domicilio y el 5,65% fue hospitalizado el 0,72% ingreso a la unidad de cuidados intensivos (UCI) y el 61,1% no registra atención (Fig. 4.9 D).

Antecedentes de pruebas: Muchos de los pacientes que no tuvieron pruebas anteriores no presentaron síntomas al momento de ingresar al centro de salud y representó el 64,56% del total de la muestra, mientras que, el 34,19% no presentó antecedentes de pruebas anteriores, pero si mostró síntomas. Por otro lado, quienes sí tuvieron pruebas anteriores y no manifestaron síntomas fue solamente el 1,09%, y el 0,16% si presentó síntomas de Covid-19 y tuvo pruebas anteriores. Esto es un

ejemplar de la gravedad del virus ya que existieron muchos casos nuevos tanto en personas sintomáticas y asintomáticas (Fig. 4.9 E).

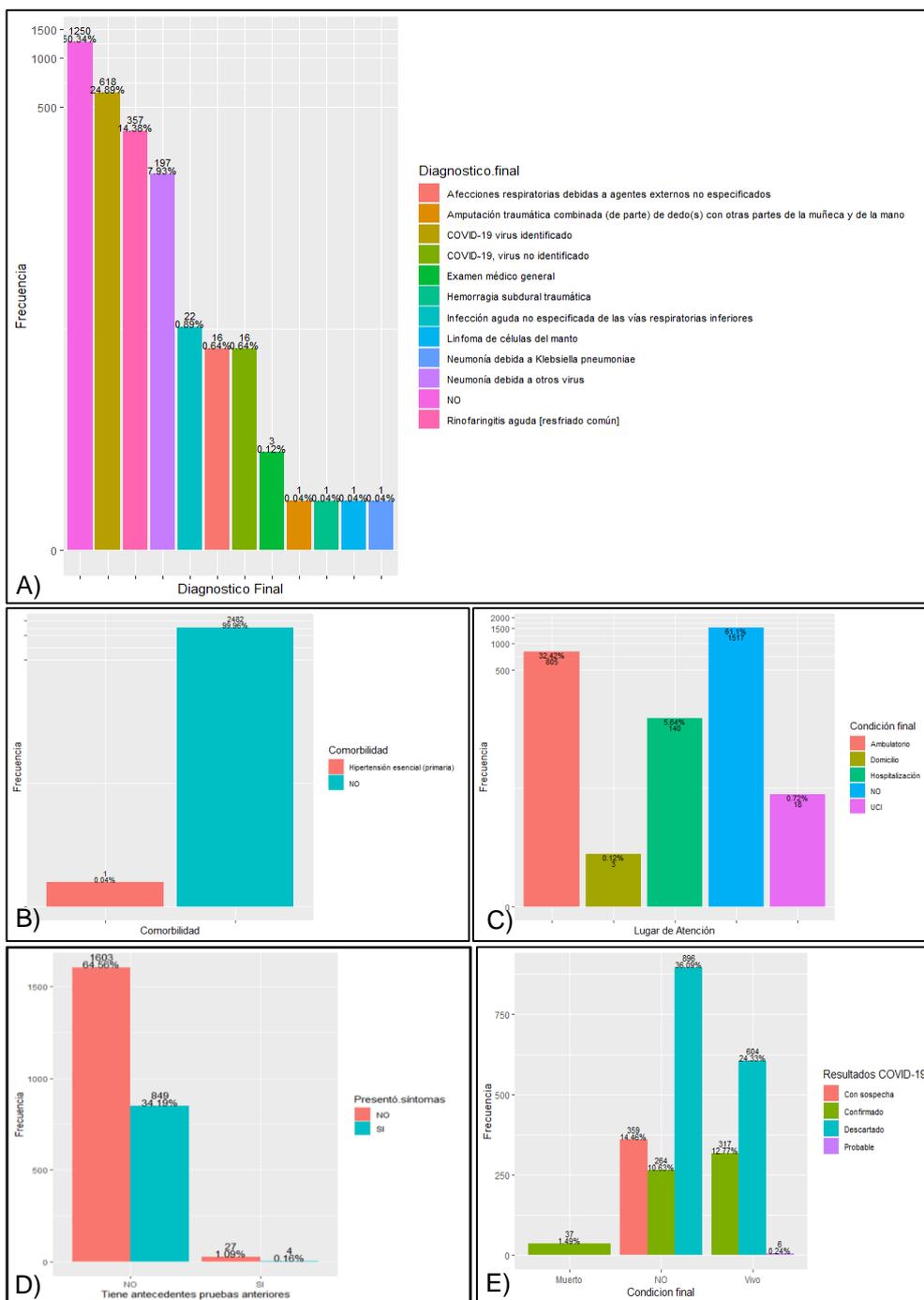


Figura 4.9 Distribución de la condición del paciente
Realizado por: Villamar Isabel, 2022

4.2.4. Ingreso de paciente en función del tiempo

La evolución de pacientes atendidos por centros médicos y laboratorios en la ciudad de Milagro en el 2021, presentó mayor cantidad de personas durante el mes de abril desde entonces las atenciones disminuyeron hasta el mes de septiembre, sin embargo, para octubre, noviembre y diciembre no registran nuevos ingresos, esto mostró con claridad cómo fue el avance hasta la actualidad. El número de casos con síntomas de coronavirus presentes en los hospitales decreció y vale recalcar que las mujeres son la que más acudieron a realizar sus respectivas revisiones y evitar la propagación de este virus (Fig. 4.10).

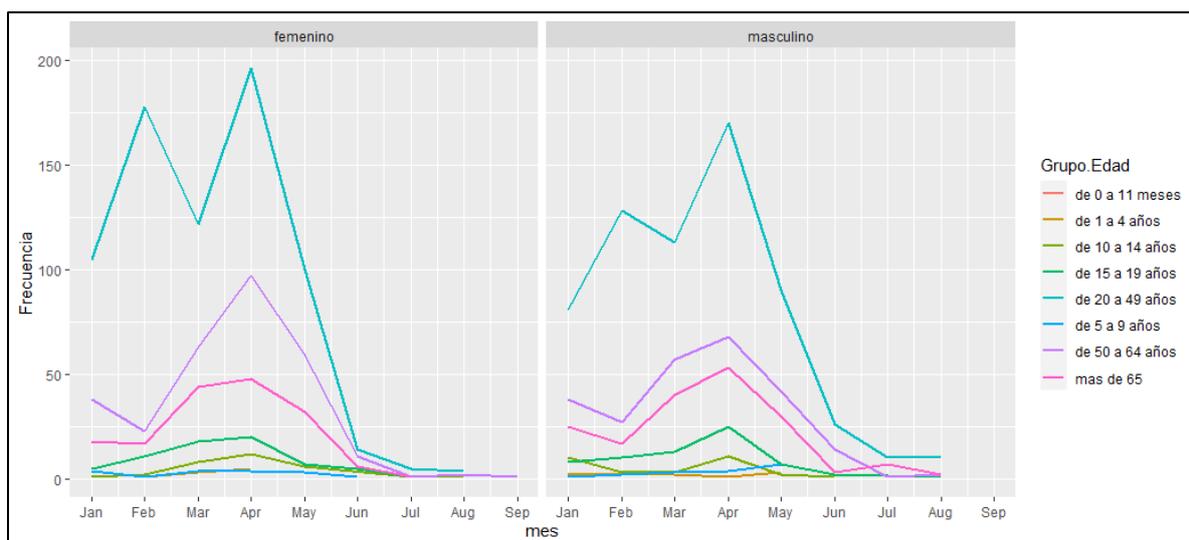


Figura 4.10 Pacientes que ingresan al centro de salud por problemas de Covid-19

Realizado por: Villamar Isabel, 2022

4.3. Análisis relacional

El análisis relacional de las variables se centró en verificar la asociación del diagnóstico final con las variables independientes. Se aplicó el test de independencia Chi y la decisión final se realizó comparando con el p.valor.

Tabla 4.5*Dependencia del diagnóstico final con las variables*

Variables independientes	p.valor	Decisión
Institución	0,0005	Dependiente
Establecimiento.de.Salud	0,0005	Dependiente
Nivel.de.atención.ES	0,0005	Dependiente
Estado.Civil	0,0005	Dependiente
Edad	0,0005	Dependiente
Grupo.Edad	0,0005	Dependiente
Ocupación	0,0005	Dependiente
Autoidentificación	0,0005	Dependiente
Tiene.antecedentes.pruebas.anteriores.	0,0005	Dependiente
Típo.prueba.anterior	0,0005	Dependiente
Presentó.síntomas	0,0005	Dependiente
Síntomas.que.presenta	0,0005	Dependiente
Antecedente.Viaje	0,002999	Dependiente
Lugar.de.viaje	0,002499	Dependiente
Muestra..	0,0005	Dependiente
Resultado.lab.1	0,0005	Dependiente
Adecuada.m1	0,0005	Dependiente
Agente.1	0,0005	Dependiente
Resultado.lab.2	0,0005	Dependiente
Adecuada.m2	0,0005	Dependiente
Agente.2	0,0005	Dependiente
Lugar.de.atención	0,0005	Dependiente
Condición.final	0,0005	Dependiente
Clasificación.final.del.caso	0,0005	Dependiente
Confirmado.por	0,0005	Dependiente
Estado.del.registro	0,0005	Dependiente
Tipo.de.prueba	0,0005	Dependiente
Ensayo	0,0005	Dependiente
Interpretación	0,0005	Dependiente

Realizado por: Villamar Isabel, 2022

Una vez obtenidos los factores que explican el diagnóstico final, se verifica el criterio de asociación según la correlación V de Cramer, la cual muestra la intensidad de correlación. La clasificación del índice como valores de referencia se observa en la siguiente tabla.

De acuerdo al índice de Cramer propuesto por la Facultad de estadística de la universidad de Santo Tomás, se consideró utilizar aquellos valores que estén por encima del 0,30, de modo que se obtenga factores que maximicen la predicción del modelo de clasificación logit multinomial que se presenta más adelante en la sección 4.4.2.

Los resultados del índice V de Cramer se presenta en la Tabla 4.6.

Tabla 4.6

Variables significativas según el criterio V de Cramer

Notación	variable	Cramer V	Criterio
x_1	Institución	0.5114	El efecto es grande
x_2	Establecimiento.de.Salud	0.42	Considera el efecto moderado
x_3	Nivel.de.atención.ES	0.5114	El efecto es grande
x_4	Presentó.síntomas	0.5494	El efecto es grande
x_5	Muestra..	0.3428	El efecto es moderado
x_6	Resultado.lab.1	0.7056	El efecto es grande
x_7	Lugar.de.atención	0.4884	Considera el efecto moderado
x_8	Condición.final	0.6269	El efecto es grande
x_9	Clasificación.final.del.caso	0.7045	El efecto es grande
x_{10}	Confirmado.por	0.447	Considera el efecto moderado
x_{11}	Estado.del.registro	0.8614	El efecto es grande

Realizado por: Villamar Isabel, 2022

4.4. Análisis multivariante

La aplicación del análisis multivariante, se desarrolla de acuerdo al tipo de variables que fueron recolectadas y con el fin de visualizar que las características que puedan estar relacionados con el diagnóstico final. De modo que este trabajo sea motivo de nuevas investigaciones con más propuestas en diferentes líneas de investigación.

4.4.1. Análisis de correspondencia múltiple

La mayoría de las variables utilizadas en este trabajo son de tipo categórica dicotómicas y politómicas, por lo tanto, se consideró aplicar el análisis de correspondencia múltiple. Esta metodología es una extensión del análisis de componentes principales y se

usa para identificar asociaciones entre las categorías de las variables. Para este primer análisis se utilizaron un total de 58 variables.

En la Figura 4.11 se identificó la variabilidad explicada por cada una de las dimensiones, de acuerdo a estos resultados la proporción explicada por el plano factorial es del 17,3% (formado por las Dimensiones 1 y 2).

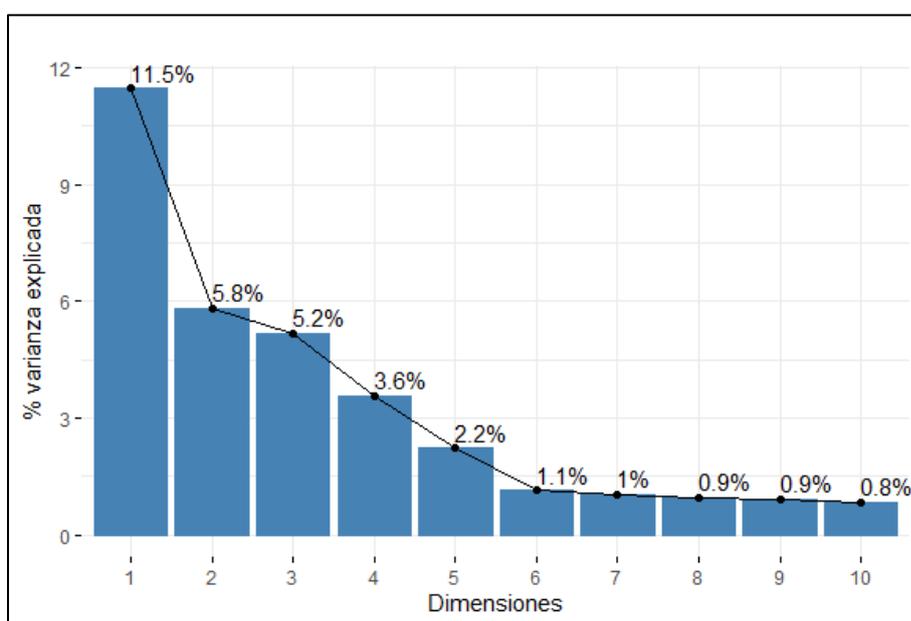


Figura 4.11 Variabilidad explicada por cada dimensión

Realizado por: Villamar Isabel, 2022

En la Figura 4.12 se observó una asociación entre algunas variables. Se determinó que el diagnóstico final puede estar representado por la ocupación u oficio que ejerza el paciente, síntomas que presentó, según los resultados del laboratorio, lugar de atención, entre otros. Por otro lado, las que se encuentran alejadas podrán ser excluidas con el fin de desarrollar un modelo más preciso.

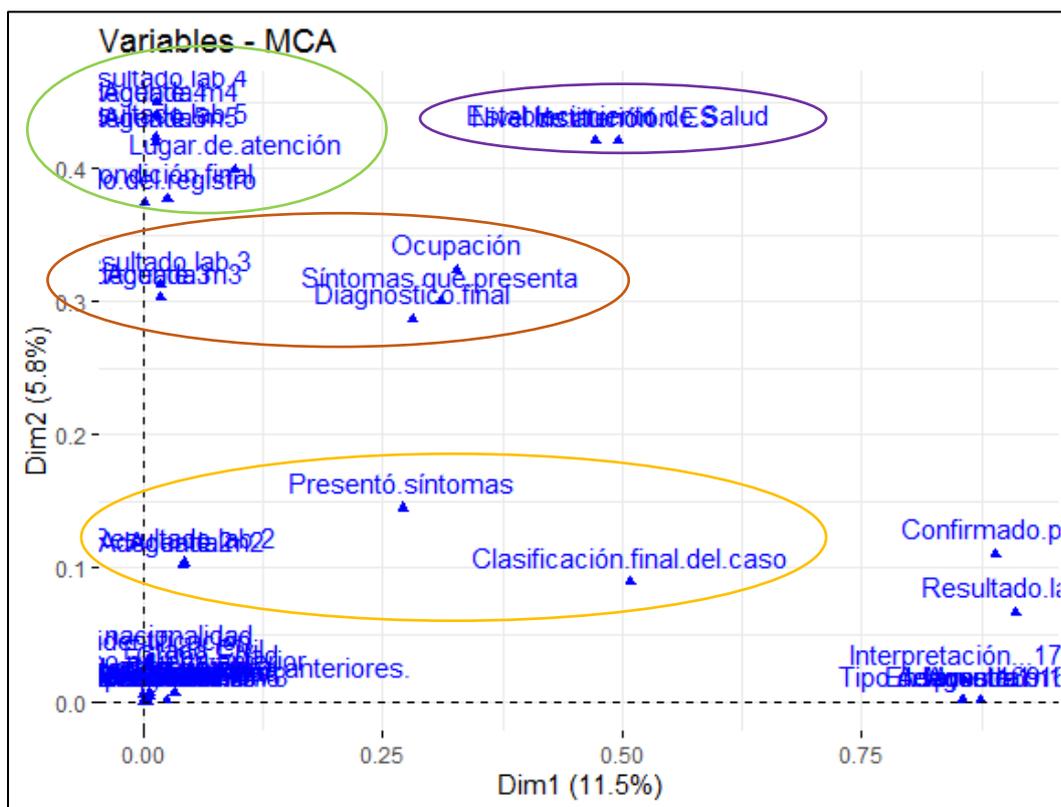


Figura 4.12 Plano factorial en base a las variables

Realizado por: Villamar Isabel, 2022

En la Figura 4.13 se observó el comportamiento de los pacientes divididos en tres grupos, evidenciando un alejamiento por parte del clúster 3 y una cercanía del 1 y 2. De acuerdo a estos resultados se pudo analizar el comportamiento de los individuos en tres conglomerados, quienes fueron atendidos en diversos centros médicos y laboratorios con problemas de coronavirus.

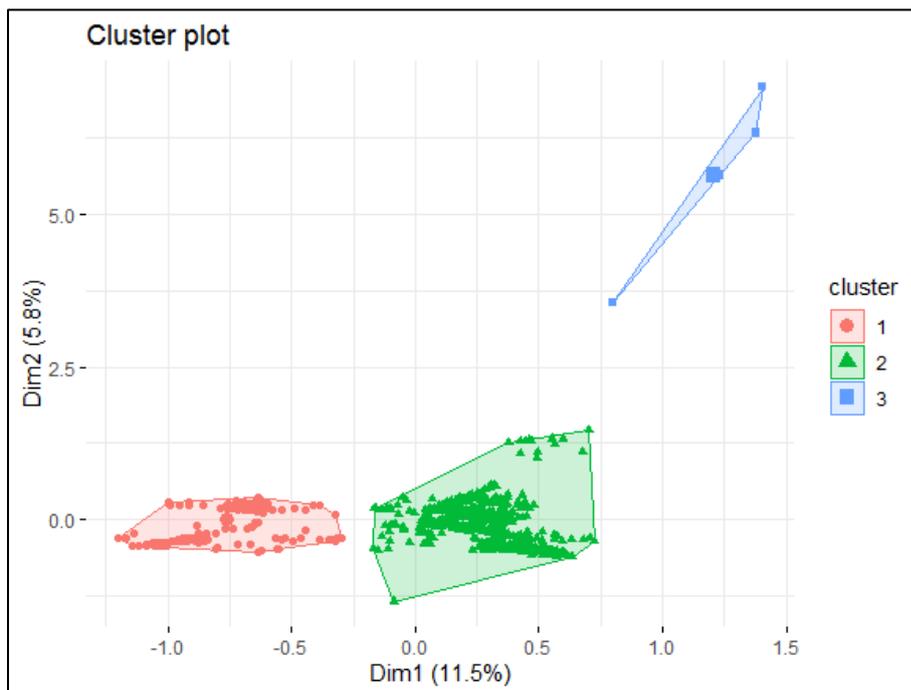


Figura 4.13 Conglomerados de los individuos

Realizado por: Villamar Isabel, 2022

Dentro del clúster 1 se evidenció una agrupación no muy dispersa, donde sus individuos fueron atendidos de manera ambulatoria en el subcentro Los Monos, los de clúster 2 fueron atendidos en el Hospital General Milagro y el Laboratorio Clínico INTERLAB. Entonces la diferencia principal de estos grupos es su punto de atención urbano y rural. Mientras que el clúster 3 tiene registros de pacientes que fueron atendidos como servicios de apoyo por parte del Laboratorio Clínico INTERLAB, además de estos no se registra si están vivos o muertos.

Con la finalidad de incrementar el porcentaje de explicación de la inercia o variabilidad explicada, se repite el Análisis de Correspondencia Múltiple utilizando las variables que resultaron significativas y óptimas en el análisis del Índice de Cramer.

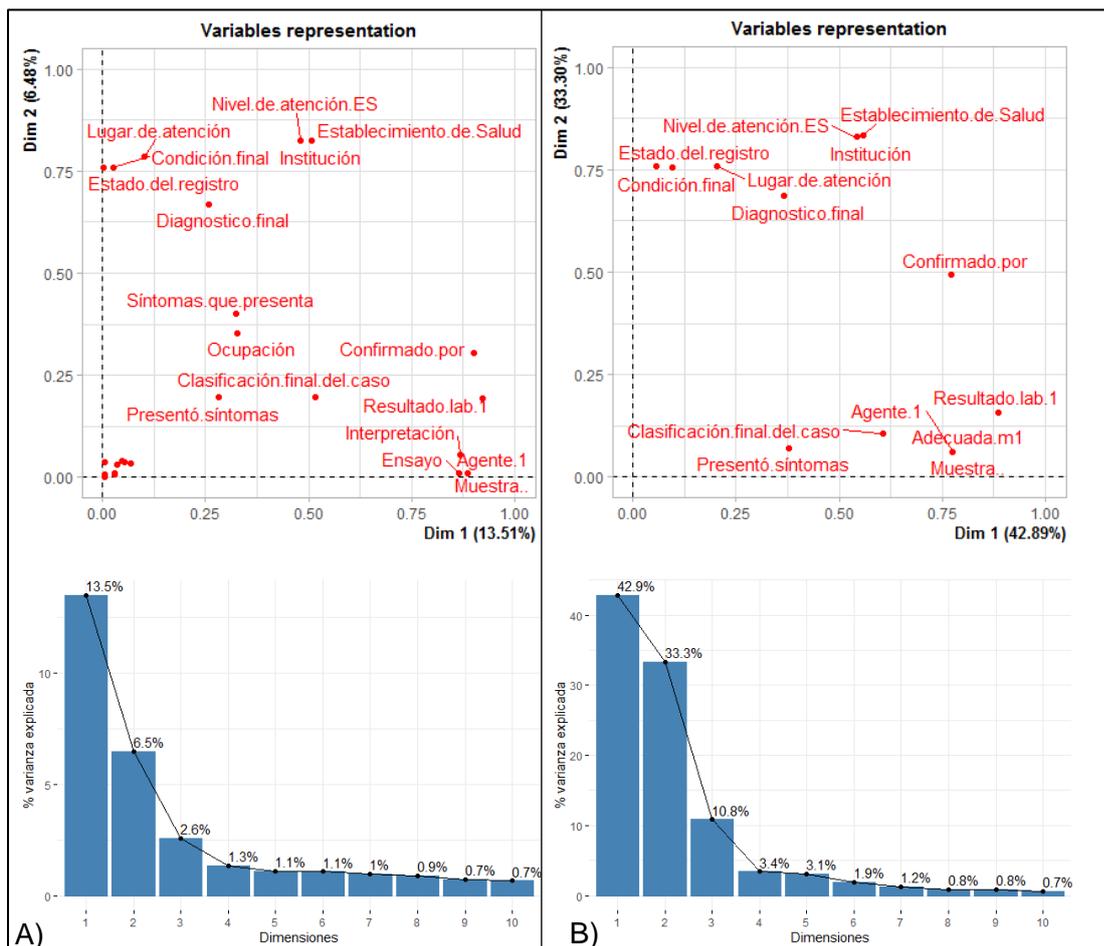


Figura 4.14 Análisis de correspondencia múltiple de: A) con las variables significativas y B) óptimas del modelo según V de Cramer

Realizado por: Villamar Isabel, 2022

En el primer ACM realizado, se tuvo una variabilidad explicada del 17,3%, posteriormente como segunda opción se repite el ACM descartando las variables que resultaron no significativas al realizar el análisis de independencia lo que originó un incremento de inercia en un 20% Gráfico 4.14 A) donde se observa el biplot de las variables y la variabilidad explicada por las dos primeras dimensiones. Sin embargo, como tercera opción se repite el ACM considerando solamente las variables que resultaron óptimas según el cálculo del índice V de Cramer, lo cual permitió tener una mayor representatividad de las

variables que están asociadas con el Diagnostico Final, originando un incremento de la inercia explicada en un 76,2% (Fig. 4.10 B).

De acuerdo a estos resultados obtenidos, se procedió a estimar el modelo logit utilizando las variables óptimas que incrementaron la inercia del ACM.

4.4.2. Modelo logit Multinomial

Se plantea el modelo especificando las variable dependiente e independiente en la Tabla 4.7:

Tabla 4.7

Variables consideradas para estimar el modelo logit multinomial

Variable dependiente		Variable independiente	
Nombre	Notación	Nombre	Notación
Diagnostico.final	Y	Institución	x_1
		Establecimiento.de.Salud	x_2
		Nivel.de.atención.ES	x_3
		Presentó.síntomas	x_4
		Muestra	x_5
		Resultado.lab.1	x_6
		Lugar.de.atención	x_7
		Condición.final	x_8
		Clasificación.final.del.caso	x_9
		Confirmado.por	x_{10}
		Estado.del.registro	x_{11}

Realizado por: Villamar Isabel, 2022

Por tanto, la expresión matemática del modelo logit multinomial se expresa como:

$$\log\left(Y_j = \frac{\pi_j}{\pi_j}\right) = \beta_{1j} + \beta_{2j}X_1 + \beta_{3j}X_2 + \beta_{4j}X_3 + \beta_{5j}X_4 + \dots + \beta_{12j}X_{11}$$

Tabla 4.8*Tabla cruzada de datos originales y predichos*

	COVID-19 virus identificado	COVID-19, virus no identificado	NO	otro	Rinofaringitis aguda [resfriado común]
COVID-19 virus identificado	618	0	0	0	0
COVID-19, virus no identificado	0	10	0	4	2
NO	0	0	1250	0	0
otro	0	0	0	147	95
Rinofaringitis aguda [resfriado común]	0	0	0	105	252

Realizado por: Villamar Isabel, 2022

Los resultados de la estimación del modelo indica que el modelo es un buen predictor para diagnosticar Covid-19, sin embargo, cuando no se detectó la enfermedad se clasificó erróneamente como otro diagnóstico a 4 individuos y 2 como rinofaringitis aguda, 95 casos de otro diagnóstico fueron predichas como personas con problemas de rinofaringitis aguda, mientras que 105 pacientes con rinofaringitis aguda fueron predichas como otro diagnóstico, ver Tabla 4.8.

De acuerdo a estos resultados se tuvo una tasa de clasificación correcta del 91,70% por tanto se considera que el modelo de clasificación es muy bueno ya que solo presenta un error de 8,3%. Seguidamente se estima el Pseudo R² de Mc-Fadden obteniéndose una variabilidad explicada del 87,26%, estos valores ayudaron a validar el modelo, ver Tabla 4.9.

Tabla 4.9*Aporte de variables al modelo logit multinomial*

variable	Clasificación correcta	incorrecta	R ² Mc Fadden
Institución	0,6605	0,3395	0,2127
Establecimiento.de.Salud	0,6605	0,3395	0,2163
Nivel.de.atención.ES	0,6605	0,3395	0,2127
Presentó.síntomas	0,6186	0,3814	0,1227
Muestra	0,5034	0,4966	0,0686
Resultado.lab.1	0,7479	0,2521	0,4412
Lugar.de.atención	0,7060	0,2940	0,4492
Condición.final	0,6617	0,3383	0,4106
Clasificación.final.del.caso	0,7547	0,2453	0,5193
Confirmado.por	0,5683	0,4317	0,2009
Estado.del.registro	0,6532	0,3468	0,4042

Realizado por: Villamar Isabel, 2022

De acuerdo a la Tabla 4.9 se observó el aporte de cada una de las variables, según su tasa de clasificación correcta, incorrecta y el pseudo R^2 de Mc-Fadden. Analizando los resultados, se observó que la variable Muestra tiene una tasa de clasificación correcta del 50,34% y explica una variabilidad de 6,86% lo cual no es considerable, mientras que, la variable Confirmado.por tiene una inercia del 20,09% y un error de clasificación del 43,17%. De acuerdo a estos resultados se consideró eliminar estas tres variables.

De acuerdo al análisis anterior, se plantea el siguiente modelo:

$$\log\left(Y_j = \frac{\pi_j}{\pi_j}\right) = \beta_{1j} + \beta_{2j}X_1 + \beta_{3j}X_2 + \beta_{4j}X_3 + \beta_{5j}X_4 + \beta_{7j}X_6 + \beta_{9j}X_7 + \beta_{9j}X_8 + \beta_{10j}X_9 \\ + \beta_{13j}X_{11}$$

Tabla 4.10

Tabla cruzada sin la variable "muestra" y "confirmado por"

	COVID-19 virus identificado	COVID-19, virus no identificado	NO	Otro	Rinofaringitis aguda [resfriado común]
COVID-19 virus identificado	618	0	0	0	0
COVID-19, virus no identificado	0	8	2	4	2
NO	0	0	125	0	0
Otro	0	0	0	147	95
Rinofaringitis aguda [resfriado común]	0	0	0	105	252

Realizado por: Villamar Isabel, 2022

De acuerdo a los nuevos resultados mostrados en la Tabla 4.10, se tuvo una tasa de clasificación correcta del 91,62% por tanto se considera que el modelo de clasificación es muy bueno ya que solo presenta un error de 8,38%. Seguidamente se estima el Pseudo R^2 de Mc-Fadden obteniéndose una variabilidad explicada del 87,19%. La comparación con el modelo inicial mostro una diferencia del 0,08% en la tasa de clasificación correcta y del 0,10% respecto al R^2 de Mc-Fadden por tanto se concluyó que las variables Muestra y Confirmado.por no aportaban una inercia significativa ya que las diferencias de valores fueron próximos a cero.

Factores de riesgo de contagio

De acuerdo a los resultados se consideró como factor de riesgo principal de contagio si la persona Presentó.síntomas de Covid-19, pues es una de las variables que más aporta en el modelo logit multinomial (Tabla 4.6) y mostró una dependencia alta con el Diagnostico.final, así mismo la atención en los diferentes centros de salud contó como riesgo,

pues una detección temprana del virus permite salvar la vida del paciente o suministrar medicamentos que minimicen los padecimientos del individuo, como se vio en la Tabla 4.5 la ocupación o profesión que ejerce resulta ser dependiente aunque dentro del modelo no resultó ser significativo, esto ocurrió debido a las tantas profesiones que registraron cada paciente que ingresó al centro médico.

Otro de los factores de riesgo resultó ser Antecedente.Viaje que la persona realizó durante el año 2021 ya sea a nivel local o internacional, en este caso se obtuvo una frecuencia muy baja de personas que se trasladaron de un lugar a otro, aunque no resultó ser significativo para el modelo logit multinomial según el índice V de Cramer (menor a 0.3), esto puede ser provocado por la baja proporción de viajeros.

Otro de los factores principales de riesgo es la característica de la Muestra, aunque esta no fue significativa en el modelo logit multinomial, pues la muestra extraída del paciente debe ser guardada en condiciones adecuadas para que sea aceptado como apropiado para los análisis pertinentes por el profesional médico, quien procederá a realizar los exámenes pertinentes hasta detectar la presencia de Covid-19. Se consideró como riesgo ya que si la persona está infectada y la muestra se contamina no se podrá detectar a tiempo la presencia del coronavirus y este individuo no será aislado lo que aumento el riesgo de contagio comunitario.

Discusión

El estudio fue enfocado en analizar la información de todas las personas atendidas y registradas en los distintos centros de Salud y laboratorios en la ciudad de Milagro durante el año 2021, los resultados mostraron que un total de 37 (1,49%) personas murieron con

problemas de coronavirus (Figura 4.9) y el factor de riesgo principal de contagio son Presentó.síntomas, seguidamente por Nivel.de.atención.ES, Condición.final y Lugar.de.atención, del mismo modo el ser hombre o mujer no es un factor de riesgo, esto concuerda con los resultados obtenidos en la investigación de Vintimilla (2021) donde el sexo no resultó ser significativo como factor de riesgo. Así mismo, quienes se encuentran en edades activas de trabajo fueron las más susceptibles a contagiarse del virus, en este caso fue mayor en individuos solteros y casados.

La investigación realizada por Leiva et al. (2021) denotó a la gestación como un riesgo que se debe considerar de prioridad de atención ante posible contagio, sin embargo, en los resultados encontrados la variable semanas de gestación no se consideró significativo ya que existieron 3 mujeres registradas como embarazadas del total de la muestra analizada. Las comorbilidades que fueron consideradas como factor de riesgo fue el asma grave, enfermedad pulmonar, cardiopatía isquémica, entre otros, en nuestro estudio encontramos un paciente que presentó problemas de comorbilidad por lo tanto no es estadísticamente significativo en el modelo logit ni muestra una dependencia con el diagnóstico final, aunque dicho paciente presentó una tercera comorbilidad y viajó fuera del país. De acuerdo a estos hallazgos pudimos afirmar que en la mayoría de casos no hubo problemas de comorbilidad en la ciudad de Milagro.

Se aplicó el modelo logit multinomial donde se encontró aporte significativo de la variable Presentó.síntomas, Condición.final, Nivel.de.atención.ES, Estado.del.registro y Clasificación.final.del.casode Covid-19 (Tabla 4.6), mostrando una dependencia mayor a 0.5 (Efecto grande) alta con el diagnóstico final, la atención en los diferentes centros de salud y laboratorios se consideraron factores de riesgo ya que si en estos lugares no se implementan las medidas sanitarias pertinentes puede el profesional de salud contagiarse y desarrollar una

nueva propagación de la cepa. En este caso la profesión u ocupación del individuo tuvo dependencia con el diagnóstico final del paciente, sin embargo, no aporta mucha relevancia para incluirla en el modelo, sin embargo, Herrero et al. (2020) denoto que las principales labores de mayor riesgo fueron para el sector sanitario y las fuerzas de seguridad del estado.

CAPÍTULO 5

5. CONCLUSIONES Y RECOMENDACIONES

Se establecieron los factores de riesgo de ingresos hospitalarios de los habitantes de la ciudad de Milagro en el 2021, mediante la identificación de variables relacionadas con el evento de contagio por Covid-19, empleando técnicas estadísticas multivariantes.

Se realizó la clasificación, segmentación y distribución de las condiciones sociodemográficas de los casos confirmados, determinando grupos vulnerables.

Se determinaron las relaciones o asociaciones entre variables para la identificación de posibles factores de riesgo y se identificaron las principales variables relacionadas con el evento de contagio por Covid-19.

De todos los pacientes que fueron ingresados por problemas de coronavirus, el 47,28% fueron hombres y el 52,72% mujeres de la ciudad de Milagro, por tanto, al acercarse a los diversos niveles de atención (IESS, MSP y PRIVADO), la atención principalmente era a personas del sexo femenino, lo que indica una mayor proporción en la atención de servicios de ayuda (73,98%).

Por otra parte, quienes estuvieron dentro de las edades 20 a 49 (54,45%) y 50 a 64 (21,91%) años, fueron los más vulnerables al inicio de la pandemia, ya que el primer grupo son consideradas personas en edad activa o productiva ya que son las que mueven la economía del país, sin embargo, de acuerdo al estado civil la mayoría de grupos vulnerables son los solteros quienes representan el 48,25% de la muestra, por otra parte, los casados fueron del 36,25%.

La mayoría de pacientes fueron ecuatorianos, seguidos por estadounidenses, chilenos, brasileños, etc. y la proporción de muertes confirmadas fueron 37 personas a causa del Covid-19. El 10,63% no se conoció su estado y el 12,77% continuaron vivos.

La relación de las distintas variables se desarrolló tomando en cuenta el diagnóstico final como la variable dependiente. La mayoría de variables fueron de tipo cualitativo, debido a eso, se empleó el test de independencia Chi Cuadrado para verificar la dependencia o independencia con el diagnóstico final, entonces, se obtuvieron como factores principales a la Institución o centro de salud a la cual acuden.

El Estado civil mostró que las personas más vulnerables eran los solteros y las personas casadas quienes se ven obligados a laborar por el bienestar de sus allegados, del mismo modo la Edad es un factor que no puede quedar atrás ya que debe haber más prioridad para individuos que mueven la economía del país y están en edad activa de trabajo. Si el paciente tiene antecedentes de pruebas anteriores es de suma importancia analizar este factor ya que puede ser un asintomático. La presencia de síntomas fue otro de los elementos que se tomó en cuenta y los síntomas son de suma importancia, ya que muestran el avance sintomatológico y su gravedad, lo cual es un buen predictor del contagio.

La ocupación laboral no se considera como parte significativa al factor de riesgo, esto puede ser causado por las diversas ramas de trabajo (muchas categorías) por lo que no aporta información específica, así mismo sucede con el antecedente de viaje, esta variable tiene poca cantidad de viajeros por lo que su contribución no se tomó en cuenta.

De acuerdo al análisis desarrollado uno de los factores más significativos fue el nivel de atención que ofrecen los distintos centros de salud, esto se consideró por la estructura

hospitalaria, laboratorios, médicos, etc. lo que permite una atención oportuna y eficiente a los pacientes que ingresaban por problemas de Covid-19.

El lugar de atención es otro de los aspectos a tomar en cuenta pues en algunos casos quienes fueron tratados en consideraciones ambulatorias se mantenían más protegidos y con vida, mientras que, aquellos pacientes que fueron ingresados y que estuvieron en la unidad de cuidados intensivos lamentablemente fallecieron a causas del Covid-19, estas características fueron calculadas aplicando un modelo logit multinomial, con una tasa de clasificación correcta del 91,62%, el modelo logró explicar una variabilidad del 87,16%, valor determinado mediante el R^2 de Mc-Fadden.

Estos tipos de análisis se deben desarrollar dentro del país, anexando más tipos de variables con características comparables ya sean cualitativas o cuantitativas, lo que garantice una mayor veracidad en las predicciones. Del mismo modo, se considera realizar estudios de acuerdo a los síntomas del paciente ya que en su gran mayoría este tipo de factor representó hechos de contagio a no ser que sea asintomático, aunque dichos pacientes sean pocos.

6. Referencias

- Afzal, A., Ansari, Z., Alshahrani, S., Raj, A., Saheer, M., Ahamed, C., & Nisar, K. (2021). Clustering of COVID-19 data for knowledge discovery using c-means and fuzzy c-means. *Results in Physics*, vol. 29. <https://doi.org/10.1016/j.rinp.2021.104639>
- Arribas, P., Bernabé, M., Herrera, E., Martínez, I., Ruiz, S., & Argote-Cervera, P. D. (2020). Estudio descriptivo del primer mes de situación de pandemia por COVID-19 en una unidad de diálisis hospitalaria. *Enferm Nefrol*, vol. 23(n. 2). <https://doi.org/https://doi.org/10.37551/S2254-28842020012>
- Barreto, A. (2012). El progreso de la Estadística y su utilidad en la evaluación del desarrollo. *Papeles de Población*, vol. 18(n. 73), 1-31. <https://www.redalyc.org/pdf/112/11224638010.pdf>
- Bedoya, M., Medina, J., Chau, V., Li, R., Vera, Á., & García, P. (2021). Variantes del SARS-CoV-2: epidemiología, fisiopatología y la importancia de las vacunas. *Revista Peruana de Medicina Experimental y Salud Pública*, vol. 38(n. 3). <https://doi.org/http://dx.doi.org/10.17843/rpmesp.2021.383.8734>
- Benzecri, J. (1978). Sur l'analyse des tableaux binaires associés a une correspondance multiple. *Cahiers de l'Analyse des Données*, 55-71.
- Betancourt, A., & Caviedes, I. (2018). Metodología de correlación estadística de un sistema integrado de gestión de la calidad en el sector salud. *SIGNOS*, vol. 10(n. 02), 119-139. <https://doi.org/https://doi.org/10.15332/s2145-1389.2018.0002.07>
- Calixto, B., Vázquez, M., Martínez, R., Bermeo, J., García, V., Mena, L., . . . López, P. (2021). Pre-existing comorbidity, the highest risk factor for poor prognosis of COVID19 among the Mexican population. <https://doi.org/https://doi.org/10.21640/ns.v13ie.2823>
- Camiz, S., & Gomes, G. (2012). *Multiple and joint correspondence analysis: Testing the true dimension of a study*.
- Castro, M. (2019). Bioestadística aplicada en investigación clínica: conceptos básicos. *Revista Médica Clínica Las Condes*, vol. 30, 50-65. <https://doi.org/https://doi.org/10.1016/j.rmcl.2018.12.002>
- Castro, M., & Ávila, M. (2006). Una introducción a la Imputación de Valores Perdidos. *Terra Nueva Etapa*, vol. 22(n. 31), 127-151. <https://www.redalyc.org/pdf/721/72103106.pdf>
- Chauca, R. (2021). La Covid-19 en Ecuador: fragilidad política y precariedad de la salud pública. *História, Ciências, Saúde – Manguinhos*, vol. 28(n. 22), 587-591.
- Chavent, M., Kuentz, V., Labenne, A., & Saracco, J. (2014). Multivariate analysis of mixed data: The PCAmixdata R package. *ResearchGate*. https://doi.org/https://www.researchgate.net/publication/268525214_Multivariate_analysis_of_mixed_type_data_The_PCAmixdata_R_package
- Ciotti, M., Ciccozzi, M., Terrinoni, A., Jiang, W.-C., Wang, C., & Bernardini, S. (2020). The COVID-19 pandemic. *Critical Reviews in Clinical Laboratory Sciences*, vol. 57(n. 06), 365-388. <https://doi.org/https://doi.org/10.1080/10408363.2020.1783198>
- Craen, S., Commandeur, J., Frank, L., & Heiser, W. (2006). Effects of Group Size and Lack of Sphericity on the Recovery of Clusters in K-means Cluster Analysis. *PUBMED*, vol. 41(n. 2), 127-145. https://doi.org/10.1207/s15327906mbr4102_2
- Cuadras, C. (2014). *NUEVOS MÉTODOS DE ANÁLISIS MULTIVARIANTE*. Barcelona: CMC Editions.
- Devkota, J. (2021). Multivariate Analysis of COVID-19 for Countries with Limited and Scarce Data: Examples from Nepal. *Journal of Environmental and Public Health*. <https://doi.org/https://doi.org/10.1155/2021/8813505>

- Fleitas, P., Paz, J., Simoy, M., Vargas, C., Cimino, R., Krolewiecki, A., & Aparicio, J. (2021). Clinical diagnosis of COVID-19. A multivariate logistic regression analysis of symptoms of COVID-19 at presentation. *Acta Académica*, vol. 11, 221-237. <https://www.aacademica.org/jorge.paz/135>
- García, A., García, S., Martínez, A., Martínez, E., & Coca, R. (2021). UTILIDAD DE LA PRIMERA ANALÍTICA DE URGENCIAS REALIZADA EN PACIENTES CON COVID-19 COMO PREDICTORA DE MORTALIDAD (PMCOVID SCORE). *Revista Española de Salud Pública*, vol. 95(n. 28). https://www.sanidad.gob.es/biblioPublic/publicaciones/recursos_propios/resp/revista_cdrom/VOL95/ORIGINALES/RS95C_202110144.pdf
- Greenacre, M. (1998). *Correspondence analysis of multivariate categorical data by weighted least squares*. Biometrika.
- Guzmán, F., & Arias, C. (2012). La historia clínica: elemento fundamental del acto médico. *Revista Colombiana de Cirugía*, 33-52. <http://www.scielo.org.co/pdf/rcci/v27n1/v27n1a2.pdf>
- Hair, J., Rolph, A., Latham, R., & Black, W. (1999). *Análisis Multivariante*. Madrid: Prentice Hall Iberia. <https://doi.org/8483220350>
- Helmy, Y., Fawzy, M., Elasad, A., Sobieh, A., Kenney, S., & Shehata, A. (2020). The COVID-19 Pandemic: A Comprehensive Review of Taxonomy, Genetics, Epidemiology, Diagnosis, Treatment, and Control. *Journal of Clinical Medicine*, vol. 9(n. 4). <https://doi.org/10.3390/jcm9041225>
- Heras, S., Rodríguez, J., Carretero, M., & Sosa, M. (2020). Utilidad de los parámetros de laboratorio en el pronóstico de los pacientes ingresados por COVID-19. *Revista de Medicina de Laboratorio*, vol. 1(n. 2), 61-68. <https://doi.org/10.20960/revmedlab.00040>
- Herrero, T., Ramírez, V., & Rueda, J. (2020). Criterios de vulnerabilidad frente a infección Covid-19 en trabajadores. *Revista de la Asociación Española de Especialistas en Medicina del Trabajo*, vol. 29(n. 02). https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1132-62552020000200004
- INEC. (2013). *INEC presenta sus proyecciones poblacionales cantonales*. Retrieved 20 de 04 de 2022, from <https://www.ecuadorencifras.gob.ec/inec-presenta-sus-proyecciones-poblacionales-cantonales/>
- INSPI, E. I. (2021). *Cuatro variantes de 'preocupación' predominan en Ecuador*.
- Lauer, S., Grantz, K., Bi, Q., Jones, F., & Zheng, Q. (2020). The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of Internal Medicine*, vol. 172(n. 09), 577-582. <https://doi.org/10.7326/M20-0504>
- Leiva, L., Proaño, R., Quishpe, G., Hernández, E., Villacís, S., Carrero, Y., & Bustillos, A. (2021). Factores de riesgo para la infección por SARS Cov 2 e implicación de las nuevas variantes del Virus, alternativas de prevención en Ambato. *Revista Universitaria con proyección científica, académica y social*, vol. 5(n. 02), 49-60. <https://doi.org/DOI:https://dx.doi.org/10.31243/mdc.uta.v5i2.1085.2021>
- Llanos, E., & Bermúdez, Á. (2020). ANÁLISIS MULTIVARIADO Y ESPACIAL PARA DETERMINAR LA RELACIÓN DE LA PROPAGACIÓN DEL COVID-19 CON CONDICIONES SOCIOECONÓMICAS EN PITALITO-HUILA. *Fundación Universitaria los Libertadores*. <https://repository.libertadores.edu.co/handle/11371/3598>
- Lucio, N. (2010). *Estadística para Ingenieros*. Lima: Macro E.I.R.L. <https://doi.org/978-612-4034-55-8>
- Maguiña, C., Gastelo, R., & Tequen, A. (2020). El nuevo Coronavirus y la pandemia del Covid-19. *Revista Medica Herediana*, vol. 31(n. 2). <https://doi.org/http://dx.doi.org/10.20453/rmh.v31i2.3776>

- Medina, E., Muñiz, J., & Barco, E. (2022). Análisis multidimensional de la evolución de la pandemia de la COVID-19 en países de las Américas. *PAN American Journal of public health*, vol. 46. <https://doi.org/https://doi.org/10.26633/RPSP.2022.49>
- MSP. (2022). *Actualización de casos de coronavirus en Ecuador*. Retrieved 20 de 04 de 2022, from <https://www.salud.gob.ec/actualizacion-de-casos-de-coronavirus-en-ecuador/>
- Oliva, J. (2020). SARS-CoV-2: origen, estructura, replicación y patogénesis. *Departamento de Gobernanza y Gestión del Conocimiento del Instituto Nacional de Salud*, , vol. 3(n. 2), 79-86. <https://doi.org/https://doi.org/10.5377/alerta.v3i2.9619>
- OMS. (2020). *Preguntas y respuestas sobre la COVID-19 para personas mayores*. Retrieved 20 de 04 de 2022, from https://www.who.int/docs/default-source/coronaviruse/covid19-advice-older-adults-qandas-final-spanish.pdf?sfvrsn=721a46f2_2#:~:text=En%20la%20mayor%C3%ADa%20de%20os,dolor%20de%20garganta%20o%20diarrea.
- OMS. (2020). *Transmisión del SARS-CoV-2: repercusiones sobre las precauciones en materia de prevención de infecciones*. https://apps.who.int/iris/bitstream/handle/10665/333390/WHO-2019-nCoV-Sci_Brief-Transmission_modes-2020.3-spa.pdf
- Ortega, C., Rivera, N., Sandoval, X., & Hernández, C. (2021). Análisis de la mutación D614G en secuencia del genoma completo de SARS-CoV-2 en El Salvador. *Revista Científica del Instituto Nacional de Salud*, vol. 4(num. 1). <https://doi.org/https://doi.org/10.5377/alerta.v4i1.10683>
- Pérez, V., Caro, R., & Rua, A. (2021). Multivariate Analysis of Risk Factors of the COVID-19 Pandemic in the Community of Madrid, Spain. *International journal of environmental research and public health*, vol. 18(n. 17). <https://doi.org/10.3390/ijerph18179227>
- Reyes, J., Mena, N., Ojea, T., Gonzales, M., Cabezudo, P., Ciano, N., . . . Serrano, P. (2020). Análisis de letalidad por COVID-19 en pacientes con demencia neurodegenerativa. *Neurología*, vol. 35(n. 9), 639-645. <https://doi.org/https://doi.org/10.1016/j.nrl.2020.07.005>
- Ruiz, A., & Jiménez, M. (2020). SARS-CoV-2 y pandemia de síndrome respiratorio agudo (COVID-19). *Ars Pharmaceutica*, vol. 61(n. 2). <https://doi.org/https://dx.doi.org/10.30827/ars.v61i2.15177>
- Sánchez, N., Saavedra, D., Hidalgo, C., Aguila, M., Abreu, G., Herrera, V., & Rodríguez, I. (2021). Parámetros de laboratorio clínico en pacientes con la COVID-19. *Revista Cubana de Medicina Militar*. <http://www.revmedmilitar.sld.cu/index.php/mil/article/view/1171/910#:~:text=Los%20par%C3%A1metros%20hematol%C3%B3gicos%20analizados%20fueron,eritrosedimentaci%C3%B3n%20y%20el%20d%C3%ADmero%20D>.
- Tenesaca, F., & Amaro, I. (2021). COVID-19 data analysis using HJ-Biplot method: A study case. *Latin American Journal of Biotechnology and Life Sciences*, vol. 6(n. 2), 1778-1784. <https://doi.org/http://dx.doi.org/10.21931/RB/2021.06.02.18>
- Torres, J. (2020). ¿Cuál es el origen del SARS-CoV-2? *Revista Médica Instituto Mexicano del Seguro Social*, vol. 58(n. 01), 1-2. <https://www.medigraphic.com/pdfs/imss/im-2020/im201a.pdf>
- Vargas, F., Ruiz, M., Marín, I., & Moreno, S. (2020). *Transmisión del SARS-CoV-2 por gotas respiratorias, objetos contaminados y aerosoles (vía aérea)*. Asociación Española de Aerobiología. <https://www.sanidadambiental.com/wp-content/uploads/2020/09/Transmisi%C3%B3n-del-SARS-CoV-2-por-gotas-respiratorias-objetos-contaminados-y-aerosoles.pdf>
- Viego, V., Geri, M., Castiglia, J., & Jouglard, E. (2020). Período de incubación e intervalo serial para COVID-19 en una cadena de transmisión en Bahía Blanca (Argentina).

Ciência & Saúde Coletiva, vol. 25(n. 9), 3503-3510. <https://doi.org/DOI:10.1590/1413-81232020259.20852020>

Vintimilla, K. (2021). *Factores de riesgo asociados a mortalidad en pacientes con Covid-19*. UNIVERSIDAD DEL AZUAY, Tesis, Cuenca.

<https://dspace.uazuay.edu.ec/bitstream/datos/11038/1/16580.pdf>

Vivas, A., & Vivas, D. (2021). Estadísticas y análisis del Covid-19 en Ecuador utilizando Microsoft Power BI . *Revista Ingenio, Facultad de Ingeniería y Ciencias Aplicadas*, vol. 4(n.1), 27-39. <https://doi.org/https://doi.org/10.29166/ingenio.v4i1.3068>

7. Apéndices y anexos

ANEXO A: Uso de software R para la implementación del código

En este trabajo se realizó la codificación en el lenguaje R, para desarrollar todos los análisis realizados en la investigación.

```
cov <- read_xlsx('COVID FILTTADO (2).xlsx', sheet = 1, na = c('--', 'NA'))
dim(cov)
names(cov)

cov %>% select("Fecha de Notificación", "Fecha Atención")
cov %>% select(contains(' Establecimiento'))

#Análisis de NA's por variable
nas <- colsums(is.na(cov))

covnatot <- select(cov, -which(nas == nrow(cov)))

covnatot %>% select(contains('Edad'))
covnatot %>% select(contains('Fecha'))
covnatot %>% select(contains('Código'))
covnatot %>% select(contains('Persona'))
covnatot %>% select(contains('Laboratorio'))

covfilt <- covnatot %>% mutate(año = year(`Fecha Atención`), mes = (month(`Fecha
  select(-c(contains('Fecha'), contains(' Establecimiento'), contains("Zona Est
  Telefono, `Celular personal`, `Tipo edad`, `Edad completa`, `Edad comp
#covfilt <- covnatot %>% filter(year(`Fecha Atención`) == 2021, `Parroquia Est
#select(-c(contains('Fecha'), contains('Establecimiento'), contains("Zona Est
#
  Telefono, `Celular personal`, `Tipo edad`, `Edad completa`, `Edad

covfilt$`Tipo prueba anterior` <- ifelse(covfilt$`Tiene antecedentes pruebas ar
covfilt$`síntomas que presenta` <- ifelse(covfilt$`Presentó síntomas` == 'NO',
covfilt$Comorbilidad <- ifelse(covfilt$`Existe comorbilidad` == 'NO', 'NO', cov
covfilt$`segunda comorbilidad` <- ifelse(covfilt$Comorbilidad == 'NO', 'NO', cc
covfilt$`Tercera comorbilidad` <- ifelse(covfilt$`segunda comorbilidad` == 'NO'
covfilt$`Lugar de viaje` <- ifelse(covfilt$`Antecedente viaje` == 'No Aplica',
```