



**ESCUELA SUPERIOR POLITÉCNICA DE LITORAL**

**FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICAS**

**ESTADÍSTICA MULTIVARIANTE**

**PERIODO: 1T – 2020**

**VISIÓN ESTADÍSTICA DE LOS FACTORES  
OCURRENTES DURANTE EL TRANCURSO  
DE LA PANDEMIA PRODUCIDO POR EL  
COVID-19**

**Elaborado por:**

Figueroa Quezada Cristhian Marcelo (crimafig@espol.edu.ec).  
Franco Monroy Fernando Alfonso (fafranco@espol.edu.ec).  
Gavilanes Gavilanes Rubén Dario (rubdgavi@espol.edu.ec).  
Parrales Gómez Nathaly Lissette (naliparr@espol.edu.ec).  
Rojas Campuzano Mauricio Javier (maujroja@espol.edu.ec).

**Dirigido por:**

Johnny Pambabay Calero, Ph.D.

Guayaquil- Ecuador



## Visión Estadísticas de los factores Ocurrentes durante el transcurso de la Pandemia producida por el Covid-19

### Índice General

<b>1. Introducción .....</b>	<b>II</b>
<b>2. Objetivos.....</b>	<b>III</b>
<b>2.1. Objetivo general.....</b>	<b>III</b>
<b>2.2. Objetivos específicos .....</b>	<b>III</b>
<b>3. Fuente y metodología .....</b>	<b>IV</b>
<b>4. Análisis descriptivo de datos e implícitos bivariados .....</b>	<b>V</b>
<b>5. Estadística inferencial .....</b>	<b>XVI</b>
5.1. Bondad de Ajuste y Prueba de hipótesis e intervalos de confianza (media, proporción y varianza) .....	XVI
5.2. Análisis de contingencia .....	XXIII
<b>6. Estadística multivariante .....</b>	<b>XXV</b>
6.1. Análisis de componentes principales .....	XXV
6.2. Análisis factorial.....	XXIX
6.3. Análisis de correspondencia.....	XXXIV
<b>7. Conclusiones .....</b>	<b>XLVIII</b>
<b>8. Referencias bibliográficas y electrónicas .....</b>	<b>XLIX</b>
Bibliografía.....	XLIX
<b>9. Anexos .....</b>	<b>L</b>

## **1. Introducción**

El brote de enfermedad por el coronavirus (COVID- 19) en Ecuador detectado desde el primer caso que se registró en el país el 9 de febrero del 2020 hasta ahora, se detalla un estudio estadístico de los casos confirmados, fallecidos, recuperados y defunciones por todas las causas que han ocurrido durante el estado de emergencia Sanitaria en el Sistema Nacional de Salud Pública.

Este estudio ampliara el conocimiento de la situación del país de acuerdo a los reportes obtenidos por provincia en la vigilancia epidemiológica y control médico por parte del Ministerio de Salud Pública, con PCR y pruebas rápidas, presentando un análisis descriptivo de los datos, análisis estadísticos bivariados, estadística inferencial y estadística multivariante.

El Servicio Nacional de Gestión de Riesgos y Emergencias, facilita la información consolidada del Ministerio de Salud Pública (MSP) sobre la emergencia nacional por COVID-19 iniciada en marzo del año en curso, hasta la actualidad, la cual se requiere para el estudio, los Informes e Infografías generados de manera periódica.

Así mismo los reportes del ECU-911, que realizan la atención de emergencia sanitaria de los diferentes tipos de sucesos registrados desde el inicio en la pandemia hasta la actualidad, colaborando activamente con los informes de datos registrados y que han sido dispuestos por el Comité de Operaciones de Emergencia.

## **2. Objetivos**

### **2.1.Objetivo general**

Analizar los resultados estadísticos obtenidos para fines informativos que faciliten el conocimiento y control, basado en los datos obtenidos del ECU-911, Ministerio de Salud Pública y Registro Civil en el transcurso de la emergencia sanitaria en el Ecuador como protocolo de fuente informativa actualizado del estado del país.

### **2.2.Objetivos específicos**

- Establecer un flujo continuo de información del estado actual de contagios y casos generados en el transcurso de la pandemia por los datos obtenidos del ministerio de salud pública.
- Facilitar el conocimiento y control de los resultados estadísticos durante la emergencia sanitaria a la ciudadanía del país.
- Analizar de manera veraz y eficaz los resultados con evidencias estadísticas exactas.

### **3. Fuente y metodología**

Los datos se obtuvieron de los informes de situación e infografías consolidadas del Ministerio de Salud Pública referente a los reportes obtenidos durante el transcurso de la pandemia, así mismo los registros anexados por el ECU-911 y el Registro Civil.

Las variables que se tomaron para el estudio del caso son: el número de casos confirmados, fallecidos, recuperados, llamadas por alertas al 911 y defunciones. Estas variables son tipo cuantitativas.

El método que se utilizó para la aplicación de los resultados es el lenguaje de programación R y el entorno de desarrollo del mismo RStudio, adicionalmente, se aplicaron los conocimientos estadísticos adquiridos en el curso de multivariante de los factores cuantitativos y cualitativos que determinan y explican la cantidad de casos reportados por el COVID-19 y las alertas de emergencia ECU- 911, midiéndose a través de índice relativo y de eficiencia.

Las técnicas aplicadas para los análisis estadísticos son:

- Análisis Descriptivo de Datos
- Análisis Estadístico Bivariado
- Estadística Inferencial
  - Bondad de Ajuste
  - Pruebas de Hipótesis e Intervalos de Confianza (media, proporción, varianza)
  - Análisis de Contingencia
- Estadística Multivariante
  - Análisis de Componentes Principales
  - Análisis Factorial
  - Análisis de Correspondencia

## 4. Análisis descriptivo de datos e implícitos bivariados

**Cuadro 1**

**VISIÓN ESTADÍSTICAS DE LOS FACTORES OCURRENTES DURANTE EL TRANCURSO DE LA PANDEMIA PRODUCIDA POR EL COVID-19.**

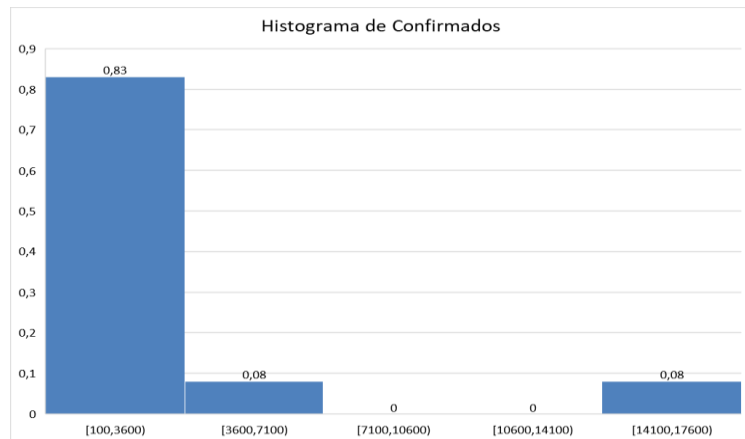
*Estadística Descriptiva de la Variable Fallecidos*

**Tabla de Frecuencias**  
"Confirmados"

Confirmados	Proporción
[100,3600)	0,83
[3600,7100)	0,08
[7100,10600)	0,00
[10600,14100)	0,00
[14100,17600)	0,08
<b>Total</b>	<b>1.000</b>

*Tabla 1 Tabla de Frecuencias*  
"Confirmados"

**Histograma de Frecuencias**  
"Confirmados"



*Ilustración 1 Histograma "Confirmados"*

**Diagrama de Caja**  
"Confirmados"

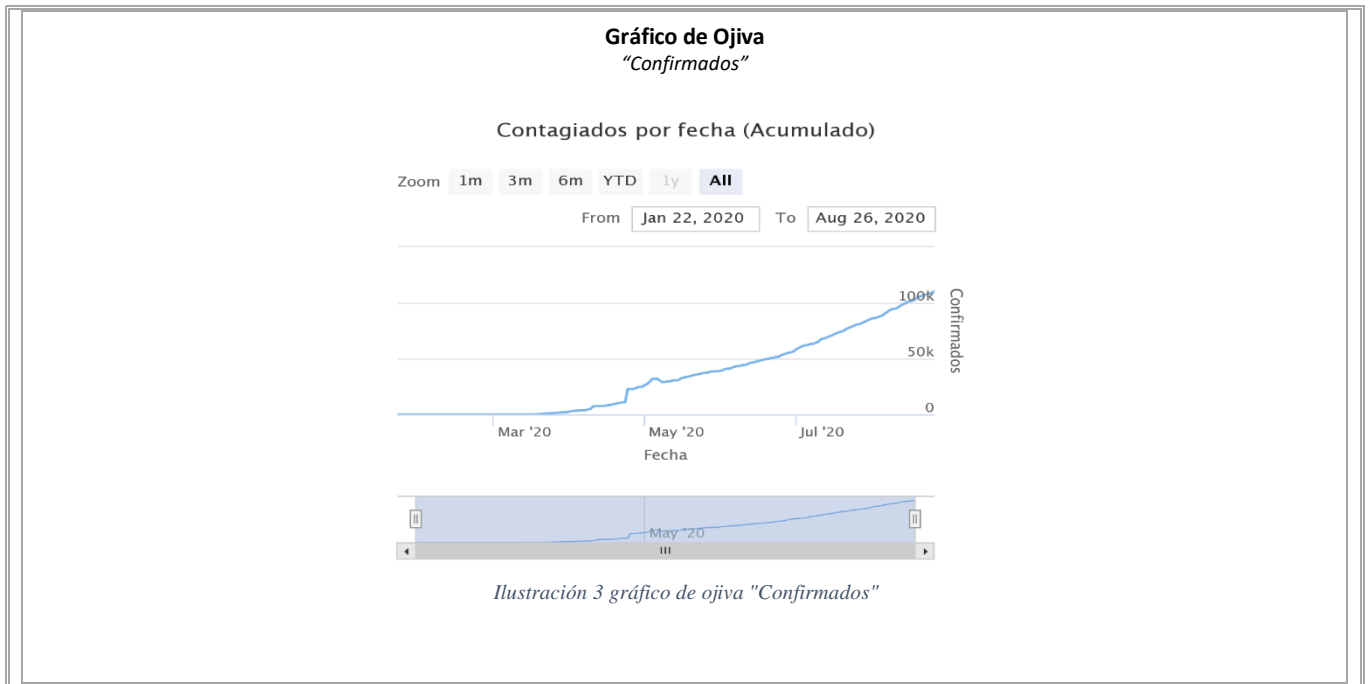


*Ilustración 2 Diagrama de caja "Confirmados"*

**Estadísticas Descriptivas**  
"Confirmados"

Media	Moda	Desviación Estándar	Sesgo	Mínimo	Máximo	Cuartil 1 (Q <sub>1</sub> )	Mediana (Q <sub>2</sub> )	Cuartil 3 (Q <sub>3</sub> )
3183,16	1148	4233,77	2,76	103	17517	1145,0	1743,0	3212,5

*Tabla 2 Estadísticas Descriptivas "Confirmados"*



## ANÁLISIS:

El histograma de Confirmados presenta una gran asimetría positiva, característica compartida por todos los demás histogramas de las variables a analizar más adelante. Esto indica que la mayoría de los casos, un 91%, de confirmados de COVID-19 por provincia se encuentran concentrados en valores que van desde 100 hasta 7100, siendo 2 notables excepciones las provincias del Guayas y Pichincha con 17517 y 14813 casos respectivamente hasta la fecha de realización de este análisis.

**Cuadro 2**  
**VISIÓN ESTADÍSTICAS DE LOS FACTORES OCURRENTES DURANTE EL TRANCURSO DE LA PANDEMIA PRODUCIDA POR EL COVID-19.**

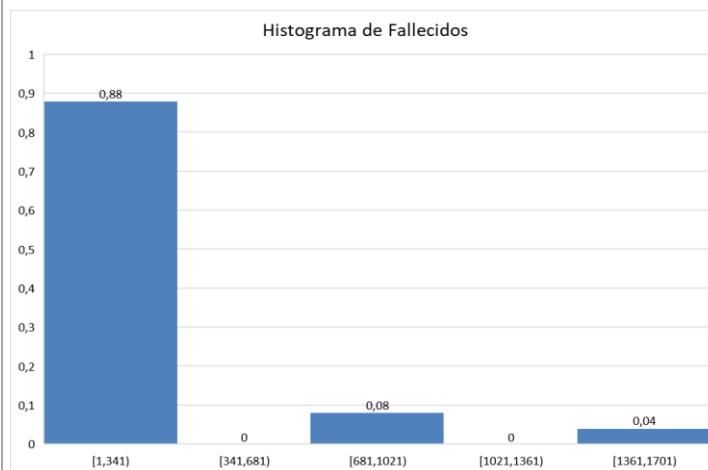
*Estadística Descriptiva de la Variable Fallecidos*

**Tabla de Frecuencias**  
 "Fallecidos"

Fallecidos	Proporción
[1,341)	0,88
[341,681)	0,00
[681,1021)	0,08
[1021,1361)	0,00
[1361,1701)	0,04
<b>Total</b>	<b>1.000</b>

*Tabla 3 Tabla de frecuencias "Fallecidos"*

**Histograma de Frecuencias**  
 "Fallecidos"



*Ilustración 4 Histograma de Frecuencias*

**Diagrama de Caja**  
 "Fallecidos"



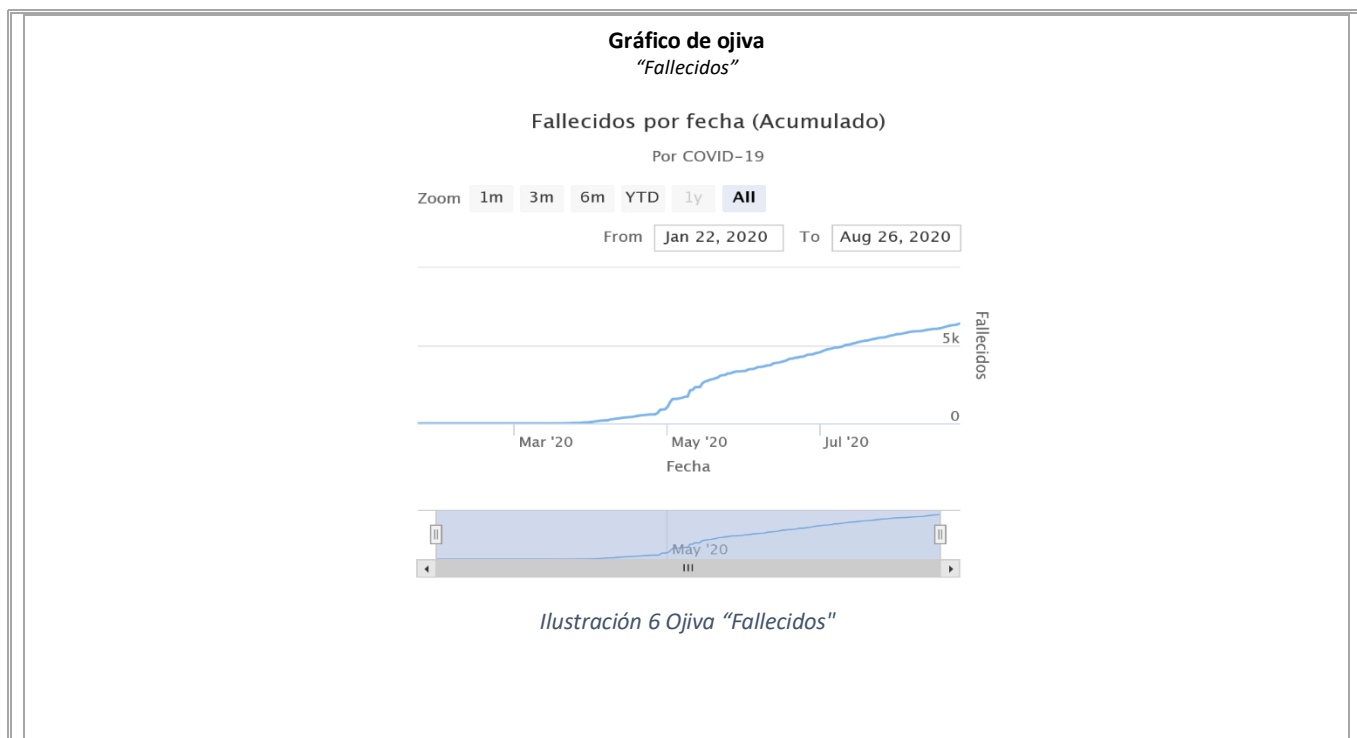
*Ilustración 5 Diagrama de caja "Fallecidos"*

**Estadísticas Descriptivas**  
 "Fallecidos"

Media	Moda	Desviación Estándar	Sesgo	Mínimo	Máximo	Cuartil 1 (Q <sub>1</sub> )	Mediana (Q <sub>2</sub> )	Cuartil 3 (Q <sub>3</sub> )
237,58	51	365,47	2,94	1	1656	45,3	76,5	298,5

*Tabla 4 Estadísticas Descriptivas "Fallecidos"*





## ANÁLISIS:

El histograma de fallecidos presente asimetría positiva en el que el 88% de los datos de fallecidos por provincia se sitúan entre los valores de 1 y 341. Una notable excepción es la provincia del Guayas, que posee el número más alto de fallecidos por Covid-19, contabilizando 1656 casos seguida por Manabí que presenta 787 casos hasta la fecha del estudio.

**Cuadro 3**  
**VISIÓN ESTADÍSTICAS DE LOS FACTORES OCURRENTES DURANTE EL TRANCURSO DE LA PANDEMIA PRODUCIDA POR EL COVID-19.**

*Estadística Descriptiva de la Variable Aglomeraciones*

**Tabla de Frecuencias**

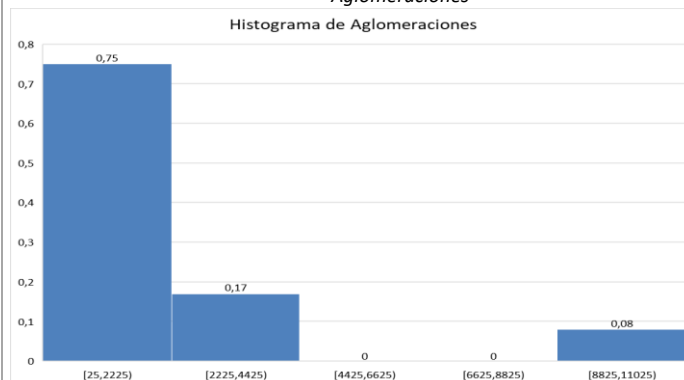
*"Aglomeraciones"*

Aglomeraciones	Proporción
[25,2225)	0,75
[2225,4425)	0,17
[4425,6625)	0,00
[6625,8825)	0,00
[8825,11025)	0,08
<b>Total</b>	<b>1.000</b>

*Tabla 5 Tabla de frecuencia "Aglomeraciones"*

**Histograma de Frecuencias**

*"Aglomeraciones"*



*Ilustración 7 Histograma "Aglomeraciones"*

**Diagrama de Caja**

*"Aglomeraciones"*



*Ilustración 8 Diagrama de cajas "Aglomeraciones"*

**Estadísticas Descriptivas**

*"Aglomeraciones"*

Media	Moda	Desviación Estándar	Sesgo	Mínimo	Máximo	Cuartil 1 (Q <sub>1</sub> )	Mediana (Q <sub>2</sub> )	Cuartil 3 (Q <sub>3</sub> )
1898,38	#N/D	2747,32	2,45	25,00	11008,00	325,00	1063,50	2201,75

*Tabla 6 Análisis descriptivo "Aglomeraciones"*

## ANÁLISIS:

Los datos de aglomeraciones presentan asimetría positiva, pero no presentan datos tan extremos como los de las variables analizadas anteriormente. Las 3 ciudades con más casos reportados de aglomeraciones son Pichincha, Guayas y Manabí en ese orden. Cabe recalcar que los casos de aglomeraciones reportados podrían estar en función de la cantidad poblacional de cada provincia.

**Cuadro 4**  
**VISIÓN ESTADÍSTICAS DE LOS FACTORES OCURRENTES DURANTE EL TRANSURSO DE LA PANDEMIA PRODUCIDA  
 POR EL COVID-19.**

*Estadística Descriptiva de la Variable Cerco Epidemiológico*

**Tabla de Frecuencias**

*"Cerco Epidemiológico"*

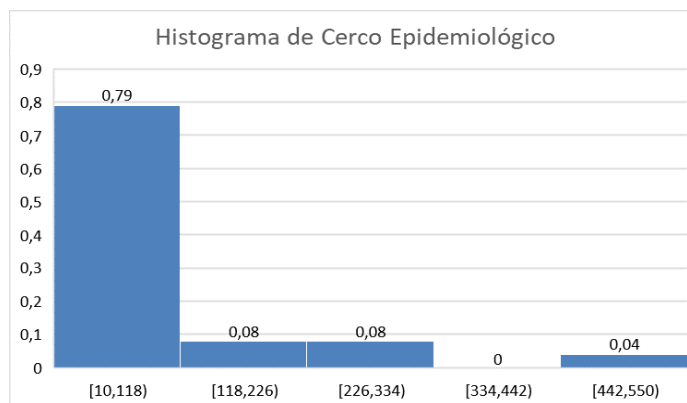
Cerco Epidemiológico	Proporción
[10,118)	0,79
[118,226)	0,08
[226,334)	0,08
[334,442)	0,00
[442,550)	0,04
<b>Total</b>	<b>1.000</b>

*Tabla 7 tabla de frecuencia*

*"Cerco Epidemiológico"*

**Histograma de Frecuencias**

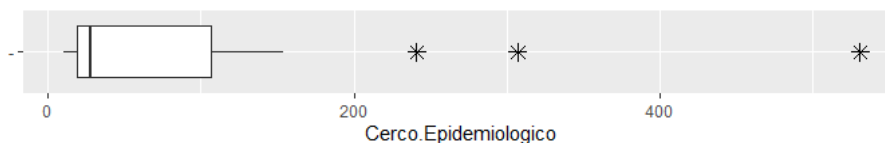
*"Cerco Epidemiológico"*



*Ilustración 9 Histograma "Cerco Epidemiológico"*

**Diagrama de Caja**

*"Cerco Epidemiológico"*



*Ilustración 10 Diagrama de caja "Cerco Epidemiológico"*

**Estadísticas Descriptivas**

*"Cerco Epidemiológico"*

Media	Moda	Desviación Estándar	Sesgo	Mínimo	Máximo	Cuartil 1 (Q <sub>1</sub> )	Mediana (Q <sub>2</sub> )	Cuartil 3 (Q <sub>3</sub> )
86,13	107,00	121,88	2,58	10,00	531,00	15,00	28,00	107,00

*Tabla 8 Análisis descriptivo "Cerco Epidemiológico"*

## ANÁLISIS:

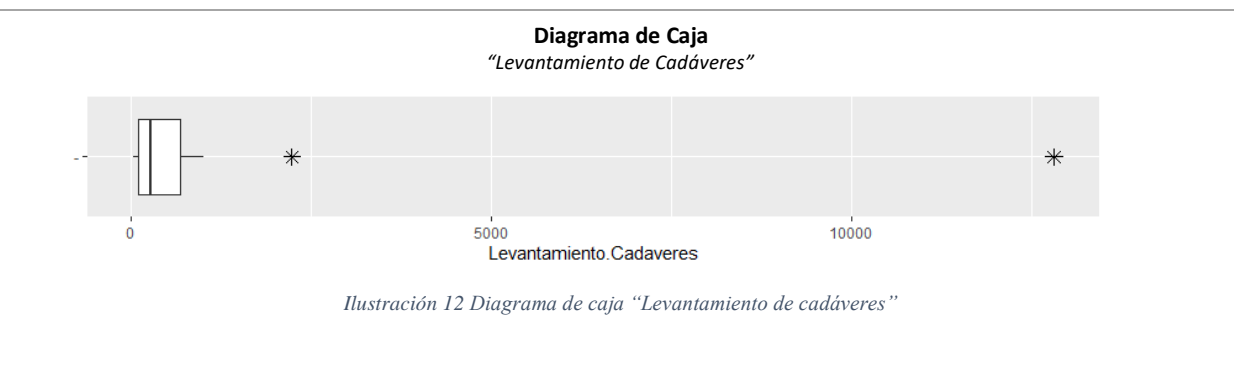
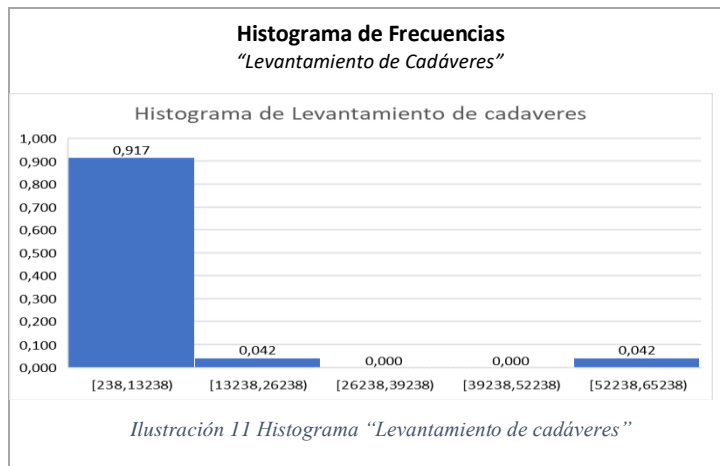
Los datos de reportes por cerco epidemiológico si bien se concentran en los valores más bajos de la distribución, no presentan casos tan extremos como las anteriores variables analizadas.

**Cuadro 5**  
**VISIÓN ESTADÍSTICAS DE LOS FACTORES OCURRENTES DURANTE EL TRANSURSO DE LA PANDEMIA PRODUCIDA POR EL COVID-19.**

*Estadística Descriptiva de la Variable Levantamiento de Cadáveres*

Tabla de Frecuencias "Levantamiento de Cadáveres"	
Levantamiento de cadáveres	Proporción
[41,2613)	0,96
[2613,5185)	0,00
[5185,7757)	0,00
[7757,10329)	0,00
[10329,12901)	0,04
<b>Total</b>	<b>1.000</b>

*Tabla 10 Tabla de frecuencia*  
 "Levantamiento de Cadáveres"



Estadísticas Descriptivas "Levantamiento de Cadáveres"								
Media	Moda	Desviación Estándar	Sesgo	Mínimo	Máximo	Cuartil 1 (Q <sub>1</sub> )	Mediana (Q <sub>2</sub> )	Cuartil 3 (Q <sub>3</sub> )
952,08	107,00	2570,76	4,63	41,00	12805,00	107,00	277,00	700,50

*Tabla 9 Estadísticas Descriptivas "Levantamientos de Cadáveres"*

## ANÁLISIS:

Dentro de los datos de levantamiento de cadáveres se presenta un dato extremadamente atípico perteneciente a la provincia del Guayas, con un número de levantamiento de cadáveres de 12805, sobrepasando por mucho a las demás provincias, siendo Manabí la que ocupa el segundo puesto con 2234 reportes.

**Cuadro 6**  
**VISIÓN ESTADÍSTICAS DE LOS FACTORES OCURRENTES DURANTE EL TRANCURSO DE LA PANDEMIA PRODUCIDA POR EL COVID-19.**

*Estadística Descriptiva de la Variable Edad*

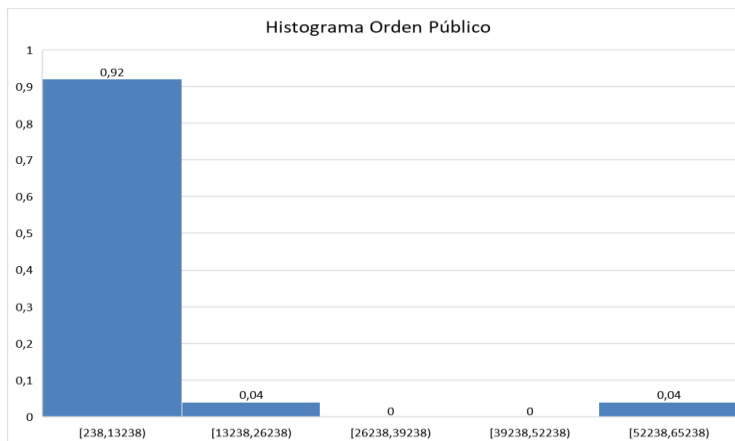
**Tabla de Frecuencias**

*"Orden Público"*

Orden Público	Proporción
[238,13238)	0,92
[13238,26238)	0,04
[26238,39238)	0,00
[39238,52238)	0,00
[52238,65238)	0,04
<b>Total</b>	<b>1.000</b>

**Histograma de Frecuencias**

*"Orden Público"*



*Ilustración 13 Histograma "Orden público"*

**Diagrama de Caja**

*"Orden Público"*



*Ilustración 14 Diagrama de caja "Orden público"*

**Estadísticas Descriptivas**

*"Orden Público"*

Media	Moda	Desviación Estándar	Sesgo	Mínimo	Máximo	Cuartil 1 (Q <sub>1</sub> )	Mediana (Q <sub>2</sub> )	Cuartil 3 (Q <sub>3</sub> )
5327,13	#N/D	13418,44	4,17	238,00	64660,00	335,75	1603,00	3536,50

*Tabla 12 Estadísticas Descriptivas "Orden Público"*

## ANÀLISIS:

Los datos de reportes por Orden público vienen dominados por dos grandes valores en comparación a los de las demás provincias, siendo Guayas y Pichincha las provincias que más casos presentan con 64660 y 22179 respectivamente hasta la fecha del estudio.

**Cuadro 7**  
**VISIÓN ESTADÍSTICAS DE LOS FACTORES OCURRENTES DURANTE EL TRANCURSO DE LA PANDEMIA PRODUCIDA  
 POR EL COVID-19.**

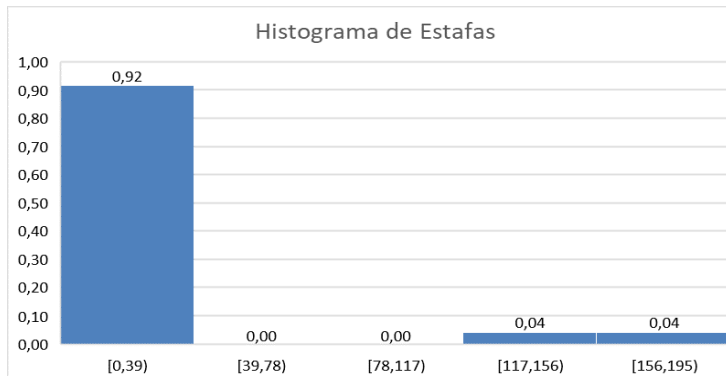
*Estadística Descriptiva de la Variable Estafas*

**Tabla de Frecuencias**  
**"Estafas"**

Estafas	Proporción
[0,39)	0,92
[39,78)	0,00
[78,117)	0,00
[117,156)	0,04
[156,195)	0,04
<b>Total</b>	<b>1.000</b>

*Tabla 13 Tabla de frecuencias "Estafas"*

**Histograma de Frecuencias**  
**"Estafas"**



*Ilustración 15 Histograma "Estafas"*

**Diagrama de Caja**  
**"Estafas"**



*Ilustración 16 Diagrama de cajas "Estafas"*

**Estadísticas Descriptivas**  
**"Estafas"**

*Tabla 14 Análisis estadísticos "Estafas"*

Media	Moda	Desviación Estándar	Sesgo	Mínimo	Máximo	Cuartil 1 (Q <sub>1</sub> )	Mediana (Q <sub>2</sub> )	Cuartil 3 (Q <sub>3</sub> )
23,92	0,00	47,59	3,06	0,00	194,00	2,00	9,00	17,00

## ANÁLISIS:

Pichincha y Guayas lideran una vez más los datos de estafas reportadas con valores de 194 y 153 respectivamente, en comparación a las demás provincias que presentan valores en las decenas como mucho.

**Cuadro 8**  
**VISIÓN ESTADÍSTICAS DE LOS FACTORES OCURRENTES DURANTE EL TRANSURSO DE LA PANDEMIA PRODUCIDA POR EL COVID-19.**

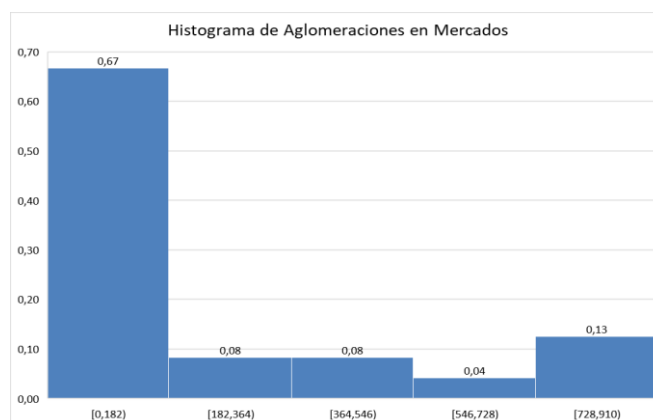
*Estadística Descriptiva de la Variable Aglomeraciones Mercados*

**Tabla de Frecuencias**  
 "Aglomeraciones Mercados"

Aglomeraciones Mercados	Proporción
[0,182)	0,67
[182,364)	0,08
[364,546)	0,08
[546,728)	0,04
[728,910)	0,13
<b>Total</b>	<b>1.000</b>

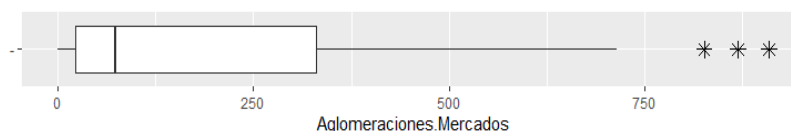
*Tabla 15 Tabla de frecuencias*  
 "Aglomeraciones Mercados"

**Histograma de Frecuencias**  
 "Aglomeraciones Mercados"



*Ilustración 17 Histograma de Frecuencia de Aglomeraciones mercados*

**Diagrama de Caja**  
 "Aglomeraciones Mercados"



*Ilustración 18 Diagrama de cajas "Aglomeraciones Mercados"*

**Estadísticas Descriptivas**  
 "Aglomeraciones Mercados"

*Tabla 16 Estadísticas Descriptivas "Aglomeraciones Mercados"*

Media	Moda	Desviación Estándar	Sesgo	Mínimo	Máximo	Cuartil 1 (Q <sub>1</sub> )	Mediana (Q <sub>2</sub> )	Cuartil 3 (Q <sub>3</sub> )
226,38	0,00	303,48	1,38	0,00	909,00	15,75	73,00	370,75

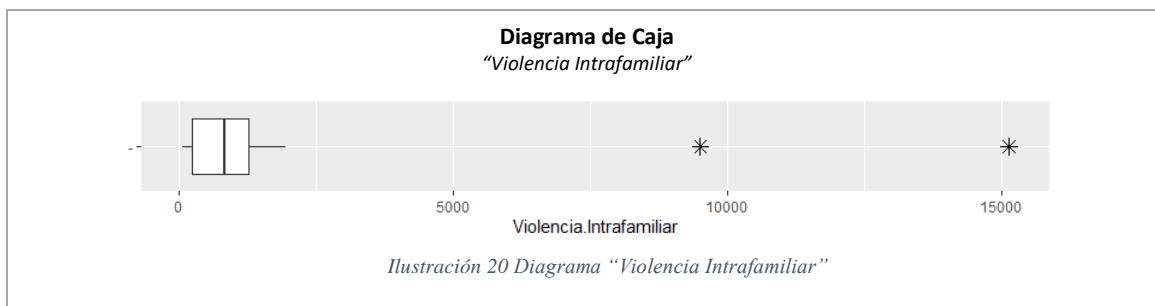
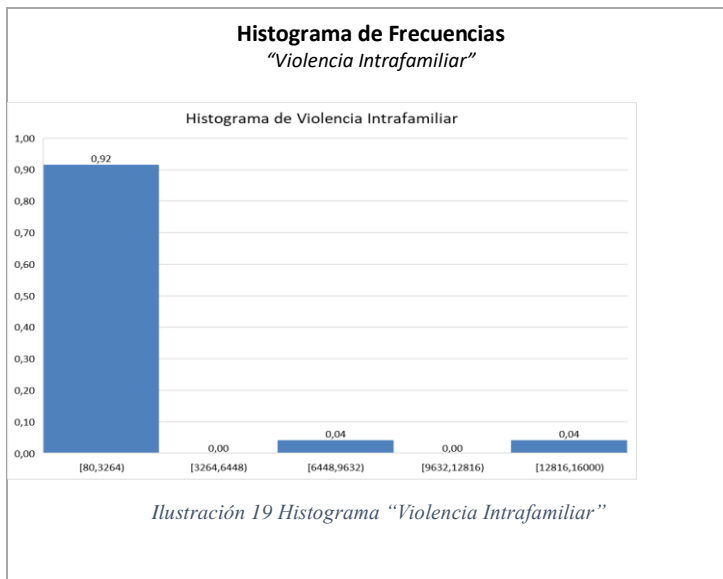
## ANÁLISIS:

Los datos de Aglomeraciones en mercados presentan una distribución un poco más uniforme en comparación a las variables estudiadas anteriormente y alcanzan un máximo de 909 reportes pertenecientes a la provincia de Los Ríos.

**Cuadro 9**  
**VISIÓN ESTADÍSTICAS DE LOS FACTORES OCURRENTES DURANTE EL TRANCURSO DE LA PANDEMIA**  
**PRODUCIDA POR EL COVID-19.**  
*Estadística Descriptiva de la Variable Violencia Intrafamiliar"*

Tabla de Frecuencias "Violencia Intrafamiliar"	
Violencia Intrafamiliar	Proporción
[80,3264)	0,92
[3264,6448)	0,00
[6448,9632)	0,04
[9632,12816)	0,00
[12816,16000)	0,04
<b>Total</b>	<b>1.000</b>

Tabla 17 Tabla de frecuencia "Violencia Intrafamiliar"



Estadísticas Descriptivas "Violencia Intrafamiliar"								
Media	Moda	Desviación Estándar	Sesgo	Mínimo	Máximo	Cuartil 1 (Q <sub>1</sub> )	Mediana (Q <sub>2</sub> )	Cuartil 3 (Q <sub>3</sub> )
1728,79	#N/D	3408,48	3,40	80,00	15129,00	237,75	847,50	1317,25

Tabla 18 Estadísticas Descriptivas "Violencia Intrafamiliar"

## ANÁLISIS:

Intrafamiliar presenta un máximo de 15129 reportes, hasta la fecha del estudio, pertenecientes a la provincia del Guayas. Esta cantidad de casos sobrepasa con creces a el de las demás provincias, siendo Pichincha la provincia que ocupa el segundo puesto en número de casos de Violencia Intrafamiliar con un total de 9504 reportes.



## 5. Estadística inferencial

### 5.1. Bondad de Ajuste y Prueba de hipótesis e intervalos de confianza (media, proporción y varianza)

El estadístico Ji-cuadrada es indispensable para inferir sobre la homogeneidad presente en los contrastes de hipótesis de que K parámetros binomiales poseen el mismo valor.

#### CONFIRMADOS+PROBABLE POR GENERO Y GRUPO ETARIO

#### DEL GRUPO ETARIO: 20-49 AÑOS

Cuadro 10 VISIÓN ESTADÍSTICAS DE LOS FACTORES OCURRENTES DURANTE EL TRANSCURSO DE LA PANDEMIA PRODUCIDA POR EL COVID-19. Prueba de Hipótesis para una Media para la variable “ <b>confirmados+ probable por género y grupo etario</b> ”			
Contraste para proporciones $H_0: p_1 = p_2$ $H_a: p_1 \neq p_2$			
SUPUESTOS	ESTADÍSTICOS DE PRUEBA		VALOR P
Los datos provienen cualquier tipo de distribución.	$x_1 = \text{Hombres}$ $x_2 = \text{Mujeres}$ $\hat{p}_1 = \frac{3685}{6407} = 0.575$ $\hat{p}_2 = \frac{2000}{3214} = 0.622$	$X^2 = \sum_{j=1}^n \frac{(n_j - \text{Esperado}_j)^2}{\text{Esperado}_j}$ $X^2_{0.05, k-t-1}$ $X^2 = 19.468$	5.114e-06
INTERVALO DE CONFIANZA			
$(\hat{p}_1 - \hat{p}_2) - Z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} < p_1 - p_2 < \hat{p}_1 + Z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$			[-1.00,-0.03]

#### ANÁLISIS:

Previamente al análisis se observó que dentro todos los niveles de Grupo etario, las personas entre 20-40 años presenta una tasa de contagio más alta, adicional a esto en el constrste de hipotesis observar que la proporción entre Grupos Etarios de las personas de casos confirmados con COVID de acuerdo al Género Masculino o Femenino son diferentes, debido a que el análisis nos confirma con un valor p menor a 0.05 que existen diferencias marcadas en este nivel a diferencia de otros niveles.

## FALLECIDOS POR GENERO Y GRUPO ETARIO

### DEL GRUPO ETARIO: 20-49 AÑOS

<b>Cuadro 11</b> <b>VISIÓN ESTADÍSTICAS DE LOS FACTORES OCURRENTES DURANTE EL TRANSCURSO DE LA PANDEMIA</b> <b>PRODUCIDA POR EL COVID-19.</b> <i>Prueba de Hipótesis para una Media para la variable "fallecidos + probable por género y grupo etario"</i>			
Contraste para proporciones $H_0: p_1 = p_2$ $H_a: p_1 \neq p_2$			
SUPUESTOS	ESTADÍSTICOS DE PRUEBA		VALOR P
Los datos provienen cualquier tipo de distribución.	$x_1 = \text{Hombres}$ $x_2 = \text{Mujeres}$ $\hat{p}_1 = \frac{28709}{48904} = 0.587$ $\hat{p}_2 = \frac{25476}{42148} = 0.604$	$X^2 = \sum_{j=1}^n \frac{(n_j - \text{Esperado}_j)^2}{\text{Esperado}_j}$ $X^2_{0.05, k-t-1}$ $X^2 = 28.35$	5.061e-08
INTERVALO DE CONFIANZA			
$(\hat{p}_1 - \hat{p}_2) - Z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} < p_1 - p_2 < \hat{p}_1 + Z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$			[-1.00 , -0.012]

### ANÁLISIS:

Podemos observar que la proporción entre Grupos Etarios de las personas entre 20-40 años de casos Fallecidos con COVID de acuerdo al Género Masculino o Femenino son diferentes, debido a que el análisis nos confirma con un valor p menor a 0.05 que existen diferencias marcadas en este nivel a diferencia de otros niveles.

## CONFRIMADOS-FALLECIDOS

<b>Cuadro 12</b> <b>VISIÓN ESTADÍSTICAS DE LOS FACTORES OCURRENTES DURANTE EL TRANSURSO DE LA PANDEMIA</b> <b>PRODUCIDA POR EL COVID-19.</b> <i>Prueba de Hipótesis para una Media para la variable "confirmados + fallecidos"</i>			
Contraste para proporciones  $H_0: \sigma_1^2 = \sigma_2^2$ $H_a: \sigma_1^2 \neq \sigma_2^2$			
SUPUESTOS	ESTADÍSTICOS DE PRUEBA		VALOR P
$X_1, X_2$ cualquier población $\sigma_1^2, \sigma_2^2$ desconocidos e iguales Muestras grandes $n_1 \geq 30; n_2 \geq 30$	$F = \frac{s_M^2}{s_m^2}(n_1 - 1, n_2 - 1)$	$F = 134.19$	$2.2e-16$
INTERVALO DE CONFIANZA			
$F_{n-1, m-1, 1-\frac{\alpha}{2}} > \frac{s_1^2/\sigma_1^2}{s_1^2/\sigma_1^2} < F_{n-1, m-1, 1-\frac{\alpha}{2}}$			$[58.05 .310.21]$

## ANÁLISIS:

Dentro del análisis inferencial se propone determinar la homogeneidad de la varianza entre casos Confirmados con COVID y casos de los Fallecidos, a fin de concebir si existe una relación, pero al aplicar el estadístico de Fisher nos indica que se rechaza la hipótesis nula al obtener un valor-p menor a 0.05 por lo tanto se concluye las varianzas son diferentes.

## FALLECIDOS-LEVANTAMIENTO DE CADAVERES

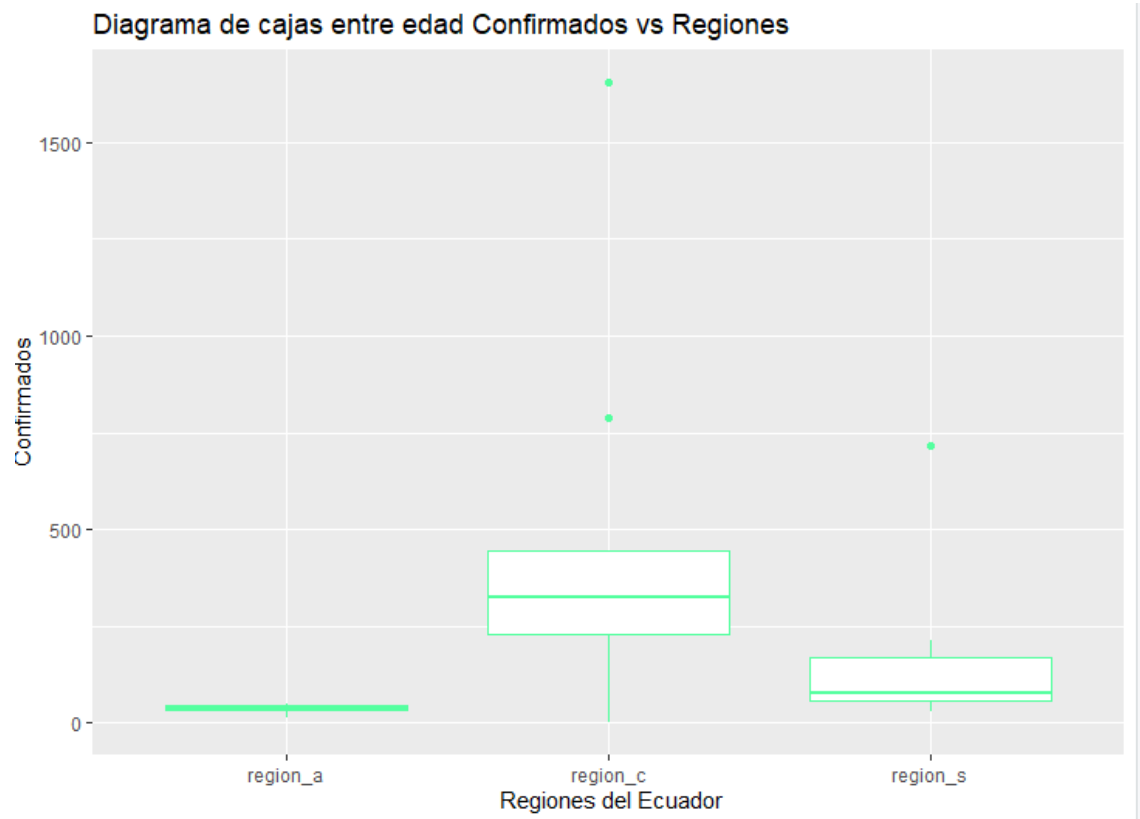
<b>Cuadro 13</b> <b>VISIÓN ESTADÍSTICAS DE LOS FACTORES OCURRENTES DURANTE EL TRANSURSO DE LA PANDEMIA</b> <b>PRODUCIDA POR EL COVID-19.</b> <i>Prueba de Hipótesis para una Media para la variable “fallecidos – levantamiento de cadáveres”</i>			
Contraste para proporciones  $H_0: \sigma_1^2 = \sigma_2^2$ $H_a: \sigma_1^2 \neq \sigma_2^2$			
SUPUESTOS	ESTADÍSTICOS DE PRUEBA		VALOR P
$X_1, X_2$ cualquier población $\sigma_1^2, \sigma_2^2$ desconocidos e iguales Muestras grandes $n_1 \geq 30; n_2 \geq 30$	$F = \frac{s_M^2}{s_m^2} (n_1 - 1, n_2 - 1)$	$F = 134.19$	$2.2e-16$
INTERVALO DE CONFIANZA			
$F_{n-1, m-1, 1-\frac{\alpha}{2}} > \frac{s_1^2/\sigma_1^2}{s_1^2/\sigma_1^2} < F_{n-1, m-1, 1-\frac{\alpha}{2}}$			$[58.05 .310.21]$

### ANÁLISIS:

Dentro del análisis inferencial se propone determinar la homogeneidad de la varianza ente los casos de personas Fallecidas por COVID y el Levantamiento de cadáveres, a fin de concebir si existe una relación, pero al aplicar el estadístico de Fisher nos indica que se rechaza la hipótesis nula al obtener un valor-p menor a 0.05 por lo tanto se concluye las varianzas son diferentes.

## ANALISIS DE VARIANZA (ANOVA)

### DIAGRAMA DE CAJAS BIVARIADOS

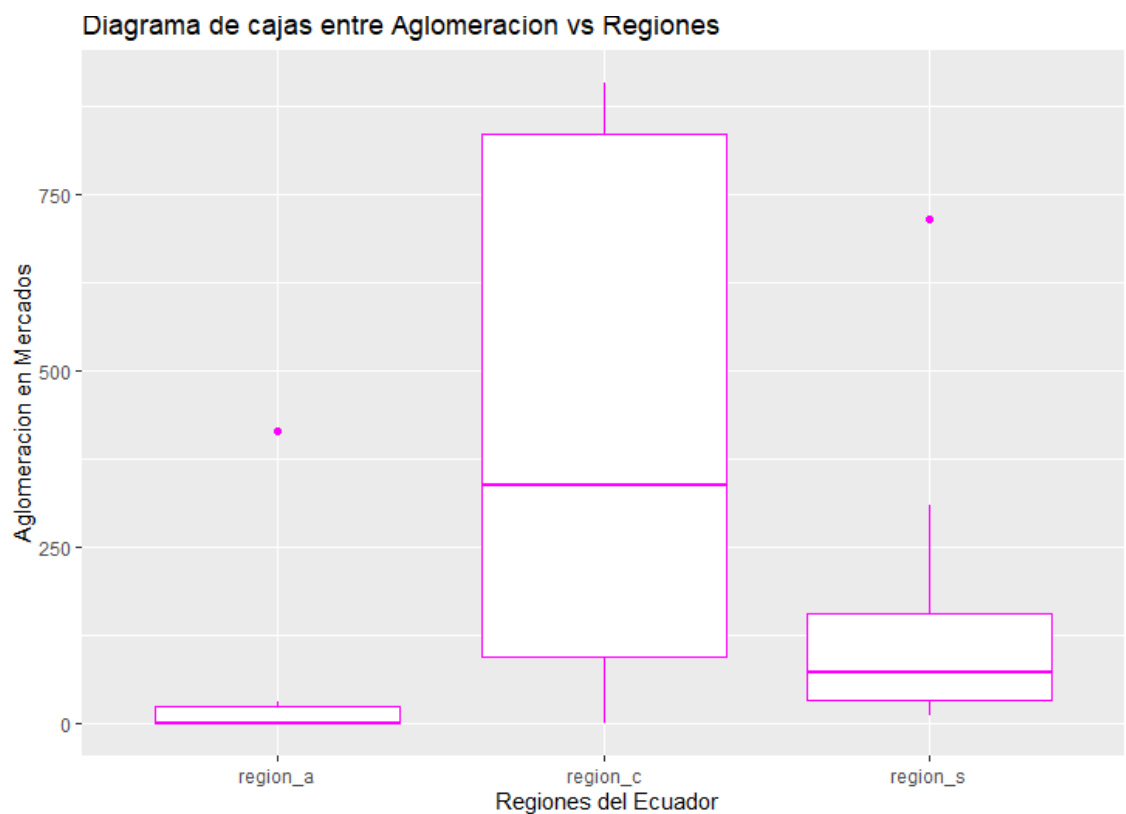


*Ilustración 21 Diagrama de cajas "Edad confirmados vs Regiones"*

### ANALISIS:

En el presente diagrama de cajas y bigotes entre la variable cuantitativa Confirmados y la variable cualitativa Regiones del Ecuador (Regios Costa + Galápagos, Región Sierra, Región Amazónica), podemos observar que el número de casos confirmados con COVID fue mucho más notorio en la Región Costa + Galápagos, pero también podemos observar que existen diferencias marcadas de variabilidad por regiones en cuanto a la media de casos confirmados.

## DIAGRAMA DE CAJAS BIVARIADOS



*Ilustración 22 Diagrama de cajas "Aglomeración vs Regiones"*

### ANÁLISIS:

En el presente diagrama de cajas y bigotes entre la variable cuantitativa Aglomeración en Mercados y la variable cualitativa Regiones del Ecuador (Regios Costa + Galápagos, Región Sierra, Región Amazónica), podemos observar que la Aglomeración en mercados es más fuerte en la región costa + Galápagos, a la vez también podemos observar que existen diferencias marcadas de variabilidad por regiones en cuanto a la media de casos confirmados.

## FALLECIDOS-REGIONES DEL ECUADOR

<b>Cuadro 14</b> <b>VISIÓN ESTADÍSTICAS DE LOS FACTORES OCURRENTES DURANTE EL TRANSURSO DE LA PANDEMIA</b> <b>PRODUCIDA POR EL COVID-19.</b> <i>Prueba de Hipótesis para una Media para la variable "fallecidos – regiones"</i>					
Contraste para proporciones					
$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2$ $H_a: \sigma_1^2 \neq \sigma_2^2 \neq \sigma_3^2$					
ANOVA					
Fuentes Var	DF	SUM	SQ	F VALUE	Pr(>F)
Regiones_ecu	2	775983	387991	3.548	0.047 *
Residuals	21	2296221	109344		

### ANÁLISIS:

Dentro del análisis inferencial se propone determinar la homogeneidad de las varianzas entre los casos de personas Fallecidas por COVID respecto a las Regiones del Ecuador, esto fue observado previamente mediante los diagramas de cajas, pero al realizar una ANOVA bajo el soporte del estadístico de Fisher nos indica que se rechaza la hipótesis nula de homogeneidad de varianzas al obtener un valor-p menor a 0.05 por lo tanto se concluye las varianzas son diferentes.

## AGLOMERACION EN MERCADOS-REGIONES DEL ECUADOR

<b>Cuadro 15</b> <b>VISIÓN ESTADÍSTICAS DE LOS FACTORES OCURRENTES DURANTE EL TRANSURSO DE LA PANDEMIA</b> <b>PRODUCIDA POR EL COVID-19.</b> <i>Prueba de Hipótesis para una Media para la variable "Agglomeraciones en mercados- regiones del Ecuador"</i>					
Contraste para proporciones					
$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2$ $H_a: \sigma_1^2 \neq \sigma_2^2 \neq \sigma_3^2$					
ANOVA					
Fuentes Var	DF	SUM	SQ	F VALUE	Pr(>F)
Regiones_ecu	2	517064	258532	3.391	0.053
Residuals	21	1601211	76248		

## ANÁLISIS:

Dentro del análisis inferencial se propone determinar la homogeneidad de las varianzas entre los casos de Aglomeración de mercados por COVID respecto a las Regiones del Ecuador, esto fue observado previamente mediante los diagramas de cajas, pero al realizar una ANOVA bajo el soporte del estadístico de Fisher nos indica que se rechaza la hipótesis nula de homogeneidad de varianzas al obtener un valor-p menor a 0.05 por lo tanto se concluye las varianzas son diferentes.

### 5.2. Análisis de contingencia

Las Tablas de contingencia conformada por reglones (r) y columnas (c) para variables cualitativas, mediante las cuales obtenemos los totales de reglones y columnas denominadas “Frecuencias Marginales”.

Nos interesa describir la relación entre el Género en cada Grupo Etario propuesto de acuerdo a número de Confirmados y Fallecidos.

#### CONFIRMADOS+PROBABLE POR GENERO Y GRUPO ETARIO

	GÈNERO		
GRUPO ETARIO	MASCULINO	FEMENINO	TOTAL
0-11 meses	100	100	200
1-4 años	209	164	373
5-9	346	337	683
10-14	592	565	1156
15-19	1038	1111	2149
20-49	28709	25476	54185
50-64	10571	8814	19385
>65	7339	5581	12920
TOTAL	48904	42128	91052

Tabla 19 Tabla de contingencia "Confirmados+ probable por género y grupo etario"

<p align="center"><b>Cuadro 16</b>  <b>VISIÓN ESTADÍSTICAS DE LOS FACTORES OCURRENTES DURANTE EL TRANCURSO DE LA PANDEMIA</b>  <b>PRODUCIDA POR EL COVID-19.</b>  <i>Prueba de Hipótesis para una Media para la variable “Confirmados+ probable por género y grupo etario”</i></p>			
<p align="center"><b>INDEPENDENCIA DE DOS VARIABLES CATEGÓRICAS.</b></p> <p align="center"><b>Ho: Grupo Etario y Genero son independientes</b>  <b>Ha: Grupo Etario y Genero son dependientes.</b></p>			
SUPUESTOS	ESTADÍSTICOS DE PRUEBA		VALOR P
Los datos provienen cualquier tipo de distribución.	$\chi^2 = \sum_{j=1}^n \frac{(n_j - \text{Esperado}_j)^2}{\text{Esperado}_j}$	$\chi^2_{0.05, k-t-1}$  $\chi^2 = 45.372$	1.099e-05



## ANÁLISIS:

Al analizar el valor p del estadístico asociado tenemos que es menor a 0.05 por tanto podríamos afirmar que el Grupo Etario de las personas de casos confirmados con COVID depende del Genero

### FALLECIDOS POR GENERO Y GRUPO ETARIO

GRUPO ETARIO	GÉNERO		
	MASCULINO	FEMENINO	TOTAL
0-11 meses	3	8	11
1-4 años	3	4	7
5-9	3	4	7
10-14	1	3	4
15-19	3	8	11
20-49	693	313	1006
50-64	2016	874	2890
>65	3685	2000	5685
TOTAL	6407	3214	9621

Tabla 20 Tabla de contingencia "Fallecidos por género y grupo etario"

<p align="center"><b>Cuadro 17</b>  <b>VISIÓN ESTADÍSTICAS DE LOS FACTORES OCURRENTES DURANTE EL TRANCURSO DE LA PANDEMIA</b>  <b>PRODUCIDA POR EL COVID-19.</b>  <i>Prueba de Hipótesis para una Media para la variable "Fallecidos por género y grupo etario"</i></p>			
<p align="center"><b>INDEPENDENCIA DE DOS VARIABLES CATEGÓRICAS.</b></p> <p align="center"><b>Ho: Grupo Etario y Genero son independientes.</b>  <b>Ha: Grupo Etario y Genero son dependientes.</b></p>			
SUPUESTOS	ESTADÍSTICOS DE PRUEBA		VALOR P
Los datos provienen cualquier tipo de distribución.	$X^2 = \sum_{j=1}^n \frac{(n_j - Esperado_j)^2}{Esperado_j}$	$X^2_{0.05, k-t-1}$  $X^2 = 99.27$	0.0004998

## ANÁLISIS:

Al analizar el valor p del estadístico asociado tenemos que es menor a 0.05 por tanto podríamos afirmar que el Grupo Etario de las personas de las personas fallecidas a causa del COVID depende del Genero

## 6. Estadística multivariante

### 6.1. Análisis de componentes principales

Se presenta primero la matriz de correlaciones de las variables.

La gran mayoría de las correlaciones tienen valores cercanos a 1, siendo la correlación más baja la

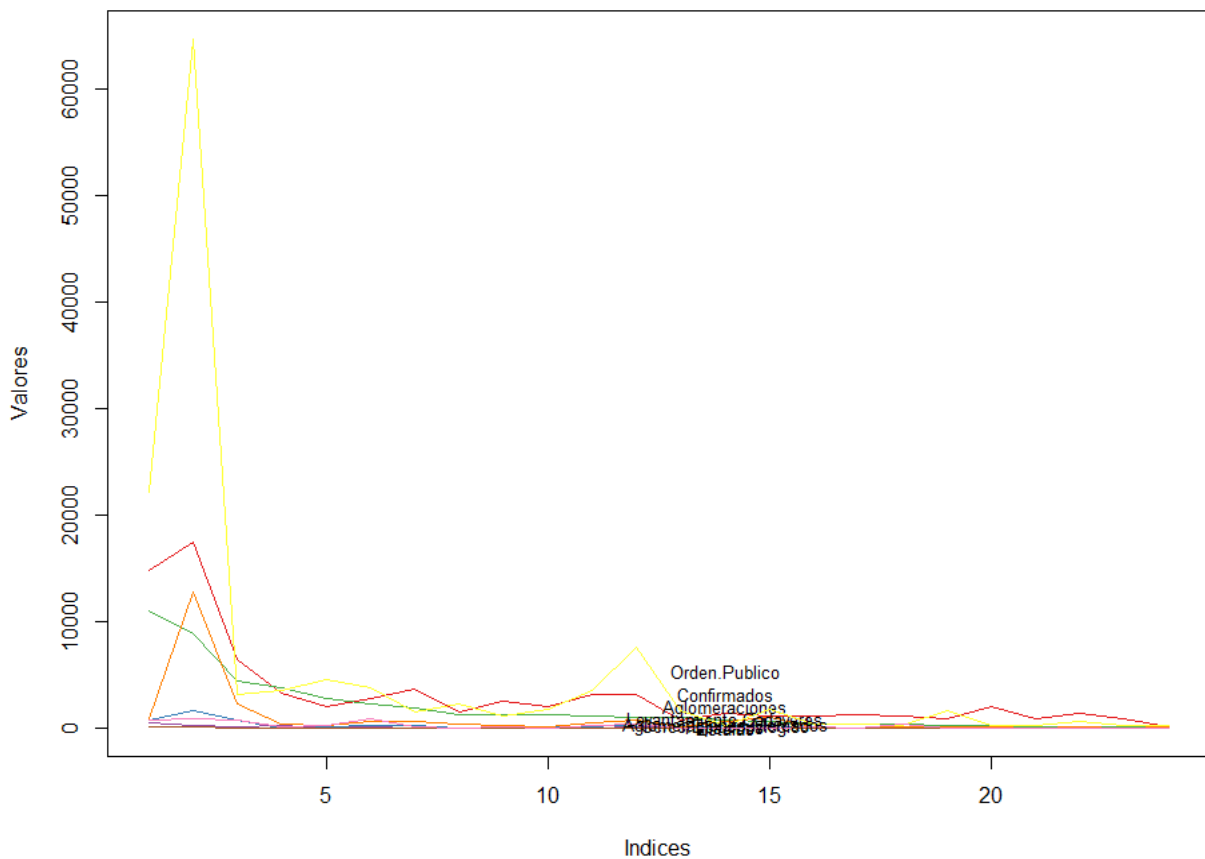
Confirmados	1	0.91	0.95	0.87	0.78	0.89	0.96	0.73	0.96
Fallecidos	0.91	1	0.81	0.66	0.91	0.91	0.78	0.76	0.9
Aglomeraciones	0.95	0.81	1	0.95	0.61	0.77	0.96	0.73	0.88
Cerco.Epidemiologico	0.87	0.66	0.95	1	0.45	0.64	0.93	0.65	0.77
Levantamiento.Cadaveres	0.78	0.91	0.61	0.45	1	0.95	0.62	0.56	0.86
Orden.Publico	0.89	0.91	0.77	0.64	0.95	1	0.8	0.6	0.96
Estafas	0.96	0.78	0.96	0.93	0.62	0.8	1	0.62	0.92
Aglomeraciones.Mercados	0.73	0.76	0.73	0.65	0.56	0.6	0.62	1	0.64
Violencia.Intrafamiliar	0.96	0.9	0.88	0.77	0.86	0.96	0.92	0.64	1
	Confirmados		Cerco.Epidemiologico			Estafas			

*Ilustración 23 Matriz de correlación*

que hay entre las variables “Cerco.Epidemiologico” y “Levantamiento.Cadaveres” con un valor de 0.45.

Luego, se realiza un gráfico comparativo para observar las posibles diferencias en escalas y en varianzas de las variables, con el fin de determinar cómo deberían ser tratados los datos antes de realizar un análisis de componentes principales.

**Gráfico comparativo de las variables**



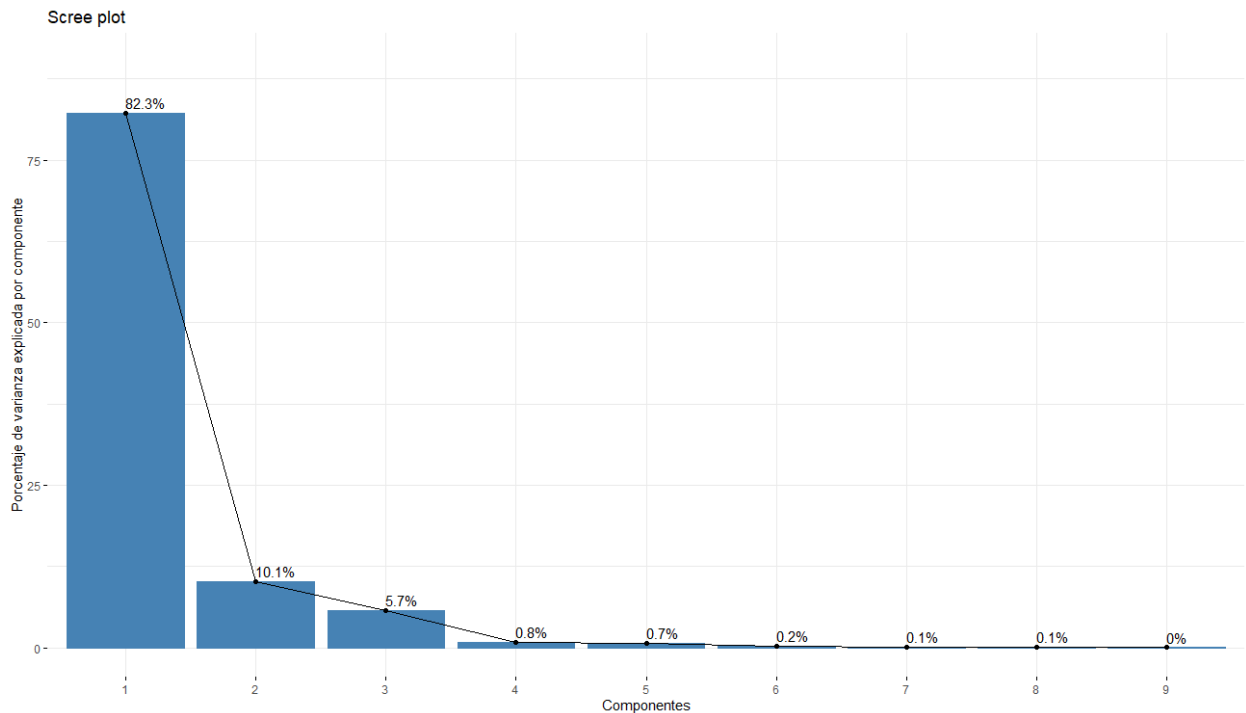
*Ilustración 24 Gráfico comparativo de las variables*

Se observa que existen muy grandes diferencias en las escalas con variables como “Orden Público” alcanzando valores de más de 60000 mientras que otras variables como “Estafas”, apenas sobrepasan el valor de 1000.

Así mismo, se observa una gran diferencia en el rango de los valores de las variables, por lo tanto, una gran diferencia entre sus respectivas varianzas.

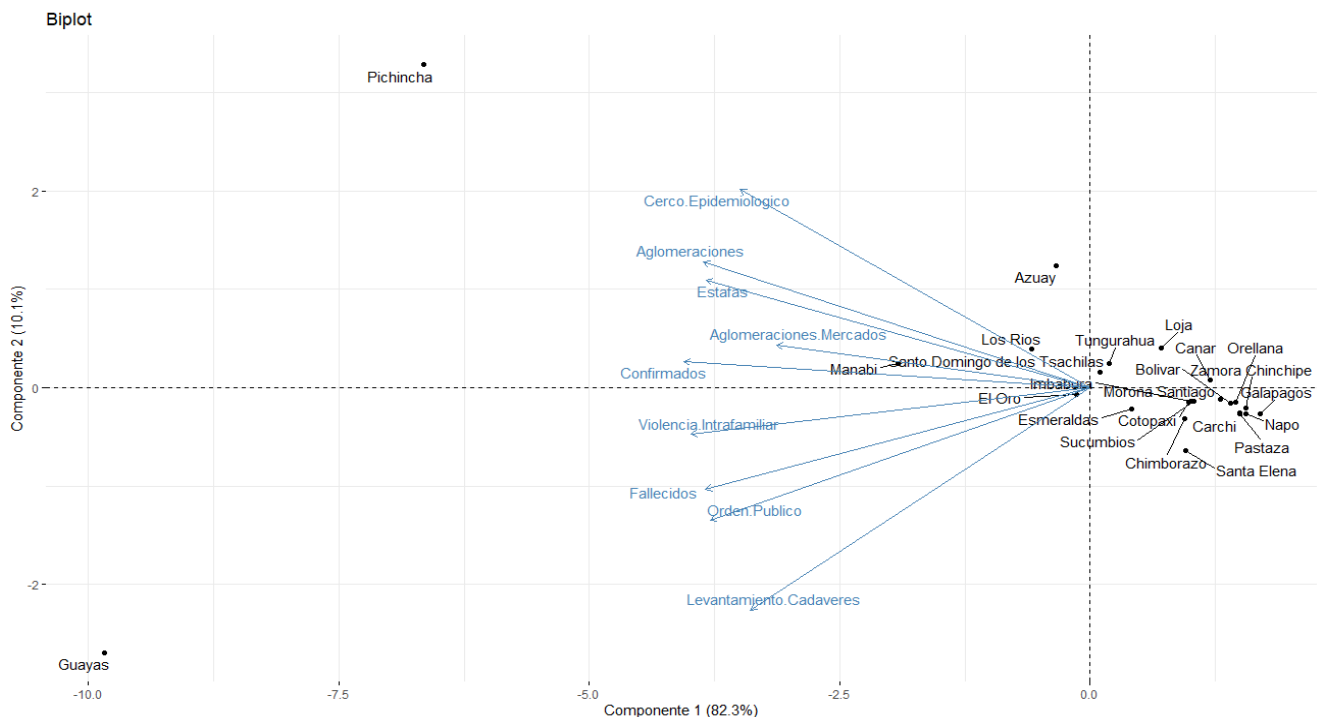
Se decide entonces trabajar con la matriz de correlación para evitar problemas al momento de realizar un análisis de componentes principales y se trabaja primero sin rotación.

Se presenta un gráfico de sedimentación que representa la cantidad de varianza explicada por cada componente.



*Ilustración 25 Grafico de sedimentación*

Se presenta a continuación un biplot sin rotación de las primeras dos componentes principales:



*Ilustración 26 Biplot sin rotación de las primeras dos componentes principales*

Se observa que la primera componente explica el 82.3% de la varianza mientras que la segunda componente explica el 10.1%. Las variables “Confirmados”, “Aglomeraciones.Mercados” y

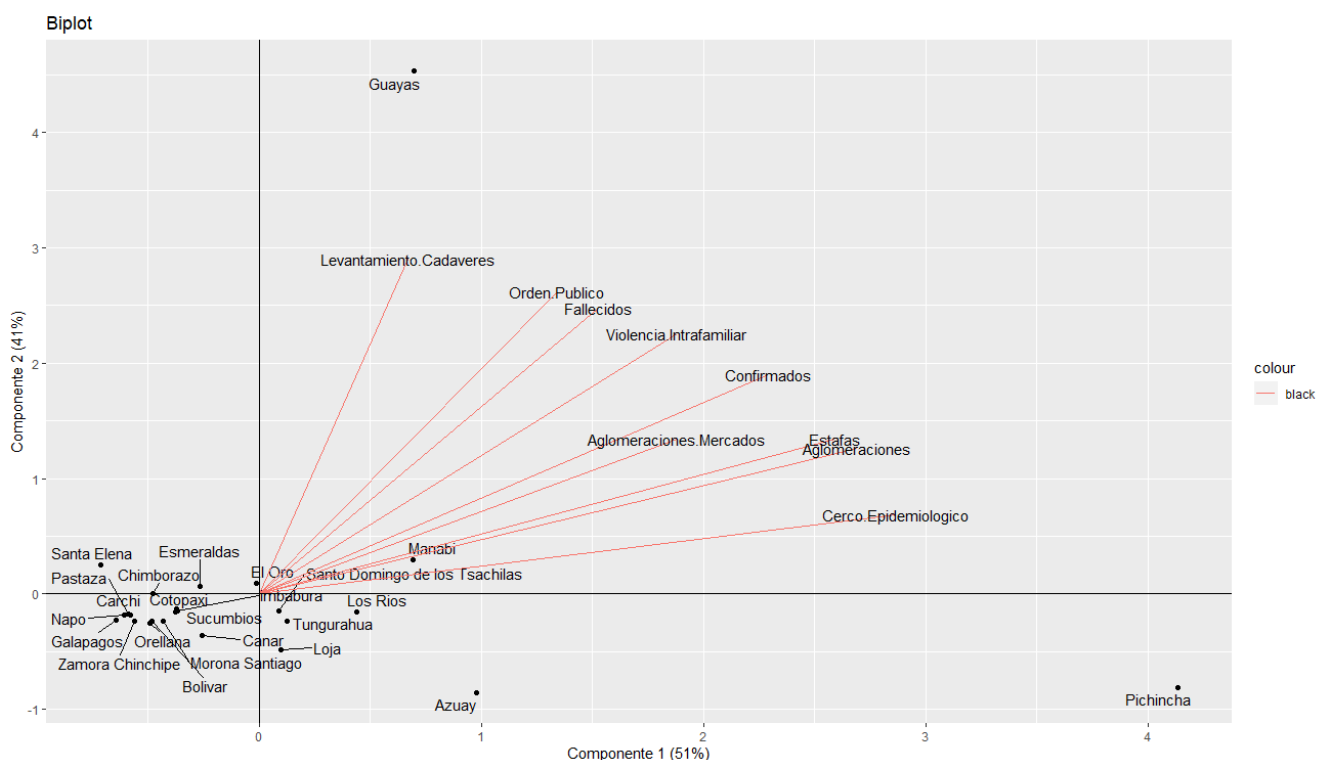
“Violencia.Intrafamiliar” se relacionan fuertemente con la primera componente mientras que el resto de las variables se asocian a ambas componentes, pero con mayor proporción a la primera.

Se podría argumentar que las provincias se agrupan en tres grupos, siendo el primero formado por “Guayas” que se caracteriza por altos valores en “Violencia.Intrafamiliar”, “Fallecidos”, “Orden.Publico” y “Levantamiento.Cadaveres”.

El segundo grupo formado por “Pichincha” que se caracteriza por altos valores en la variable “Cerco.Epidemiologico”, “Aglomeraciones”, “Estafas”, “Aglomeraciones.Mercados”, y “Confirmados”.

El resto de las provincias se caracterizan por tener valores bajos en las variables excepto quizá por Manabí que se encuentra del lado positivo de la primera y segunda componente.

Luego, se procede a realizar de la misma manera un análisis de componentes principales, pero haciendo uso de la rotación “varimax” y se presentan los resultados a continuación:



*Ilustración 27 Biplot del análisis en los componentes principales con rotación "varimax"*

Se observa que ahora cada componente explica una cantidad de varianza parecida 51% y 41% , pero que no se cumple el supuesto de que la primera componente explica la mayor cantidad de varianza, esto debido a la rotación usada.

Se debe tener cuidado al momento de interpretar un biplot correspondiente a un PCA con rotación, debido a que lo que la rotación trata de hacer es dejar a cada componente con pesos altos en algunas variables y bajos en otras, con el fin de obtener una estructura “simple”.

## 6.2. Análisis factorial

Se comienza realizando una validación de la aplicabilidad de usar un modelo factorial para analizar los datos.

Primero se analiza el determinante de la matriz de correlaciones:

$$Det(R) = 3.48 \times 10^{-11}$$

Se obtiene un valor muy cercano a 0, lo que indica alta multicolinealidad entre las variables involucradas en el análisis, condición deseable para un análisis factorial.

Se analiza también la medida de suficiencia o adecuación de la muestra (KMO)

Confirmados	0.90
Fallecidos	0.78
Aglomeraciones	0.92
Cerco.Epidemiologico	0.67
Levantamiento.Cadaveres	0.79
Orden.Publico	0.81
Estafas	0.74
Aglomeraciones.Mercados	0.77
Violencia.Intrafamiliar	0.71

*Tabla 21 Medida de suficiencia o adecuación de la muestra (KMO)*

El menor valor obtenido fue 0.67, perteneciente a la variable “Cerco.Epidemiologico” y el más alto 0.92, para la variable “Aglomeraciones”. Los valores obtenidos son cercanos a 1, lo que indica que es viable realizar un análisis factorial.

Para seleccionar el número de factores a extraer, se escogió la misma cantidad de factores como de componentes principales que se usaron en el análisis anterior.

Primero, se realizó un análisis factorial sin rotación mediante máxima verosimilitud con dos factores y se presentan los resultados a continuación:

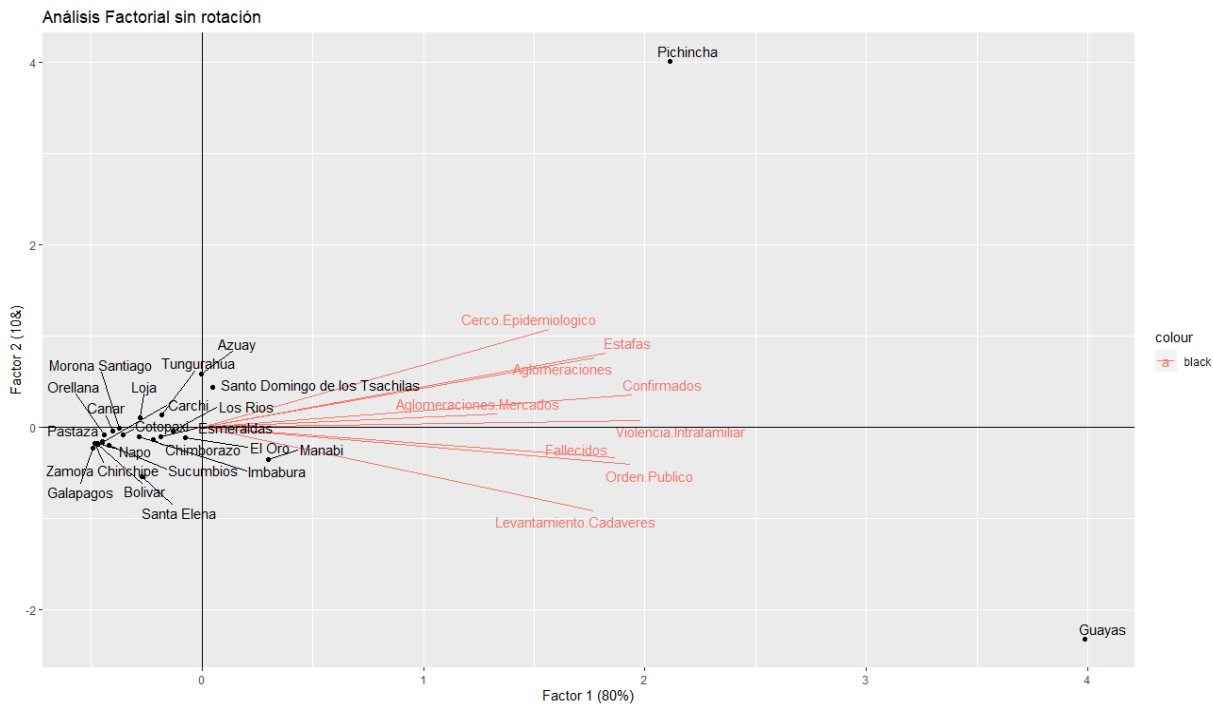


Ilustración 29 Biplot del análisis con dos factores sin rotación

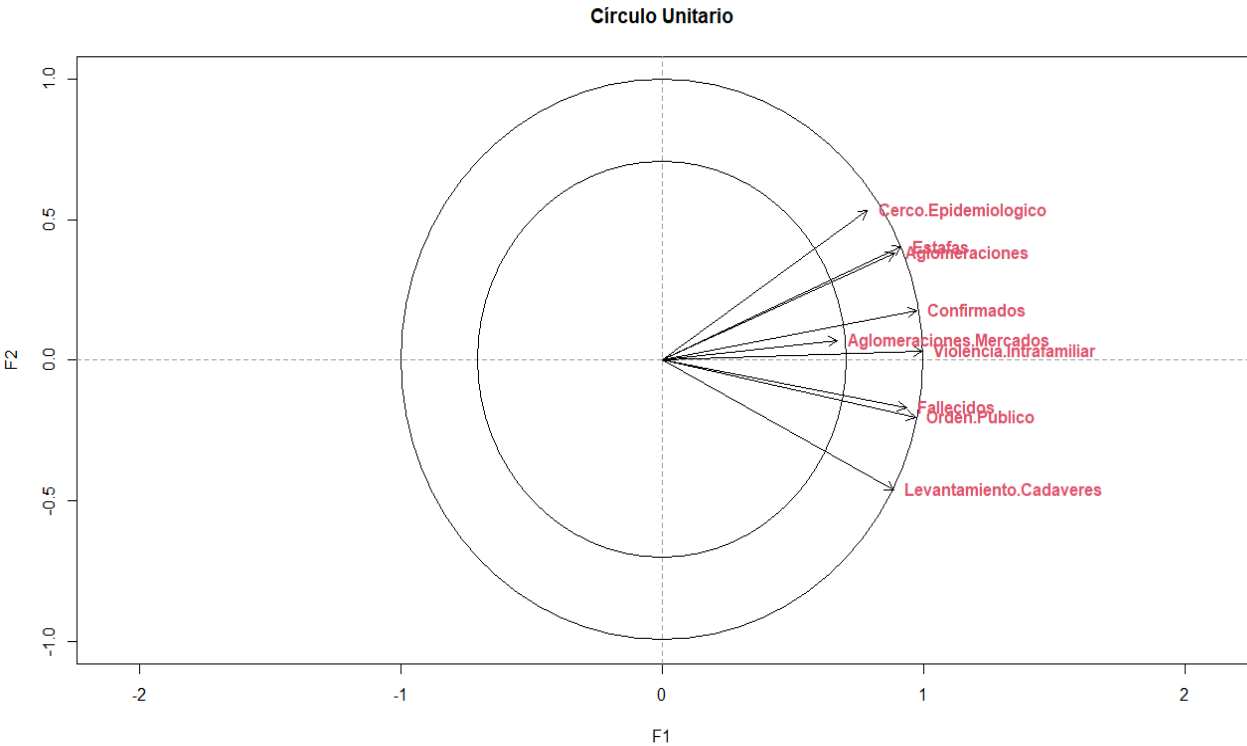
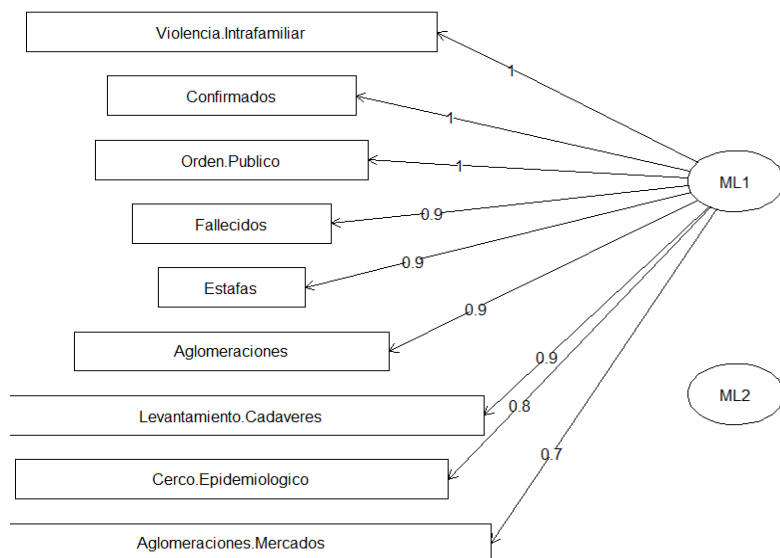


Ilustración 28 Análisis factorial sin rotación mediante máxima verosimilitud con dos factores

### Factor Analysis



*Ilustración 30 Análisis Factorial “con dos factores sin rotación”*

La proporción de varianza explicada por los dos factores es de 80% y de 10% respectivamente para el factor 1 y el factor 2. La cantidad de varianza explicada por los constructos latentes es muy parecida a la cantidad de varianza explicada por las componentes principales del análisis anterior, esto debido a que la variabilidad total de los datos se ve explicada en buena parte por la variabilidad que hay entre los datos.

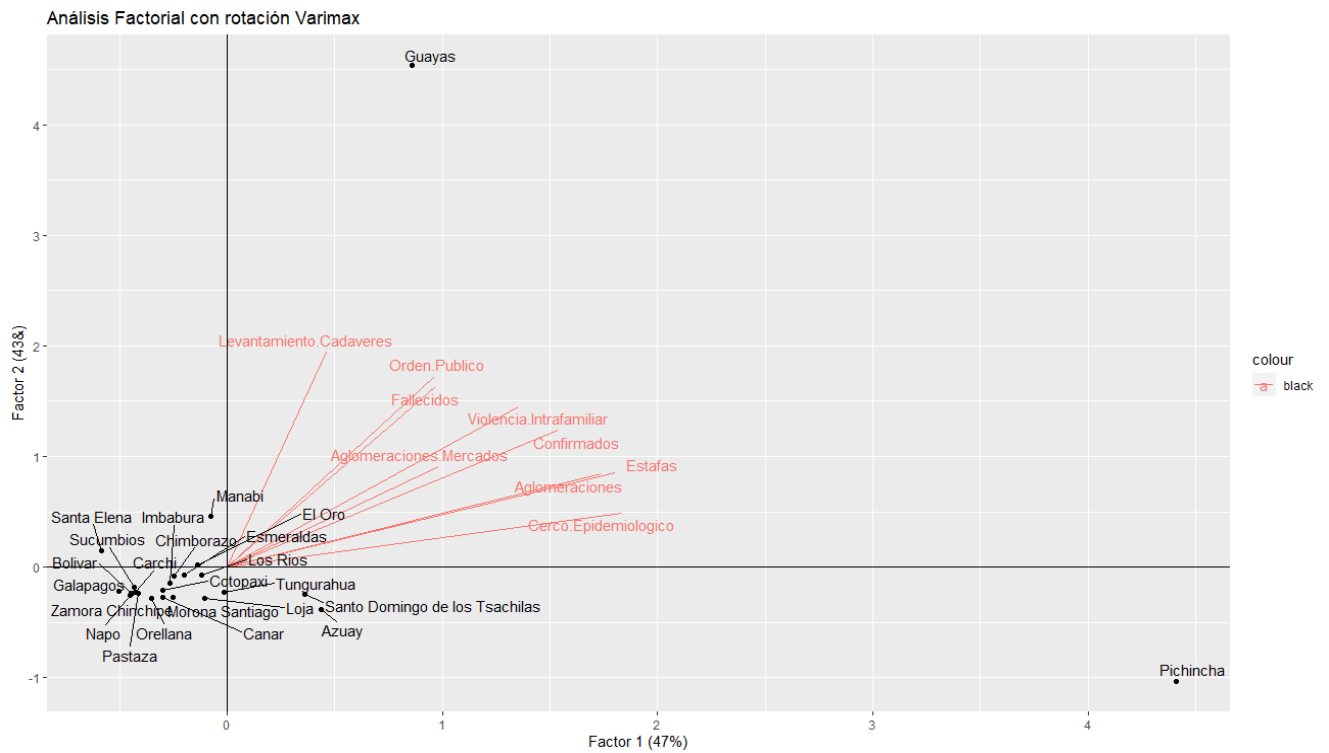
Los gráficos tanto de las cargas como de los “scores”, son muy parecidos a los obtenidos en el análisis de componentes principales.

En el gráfico de Círculo unitario se pueden observar las correlaciones de cada una de las variables con los constructos latentes, se observa que la variable “Violencia.Intrafamiliar” está fuertemente correlacionada con el primer constructo, siendo casi paralelo a este. Las demás variables también se encuentran correlacionadas fuertemente con el primer factor, pero en menor cantidad.

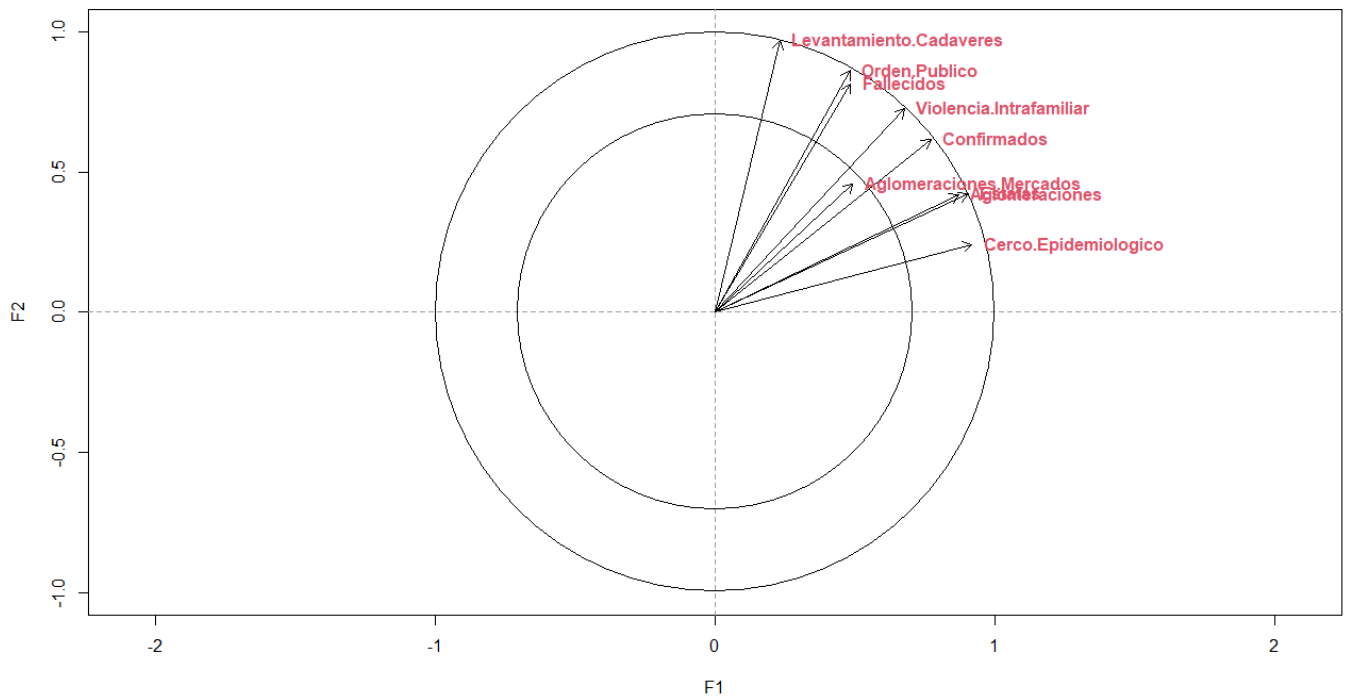
En el esquema de cargas, se observa que todas las variables cargan más en el primer constructo que en el segundo.

Se buscará una estructura más simple donde las variables carguen más en un factor que en otro mediante una rotación “Varimax”.

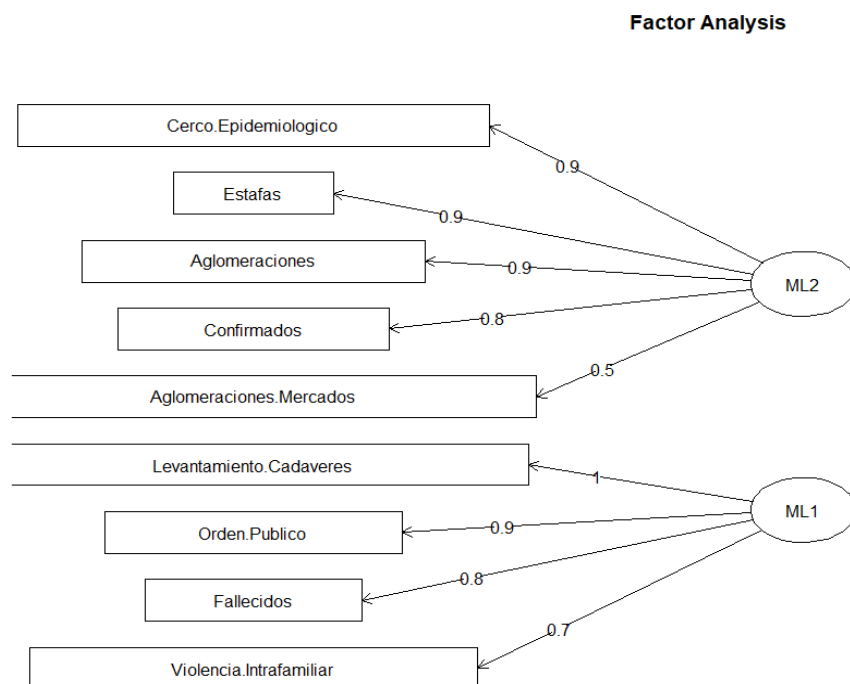




*Ilustración 31 Biplot de Análisis factorial “Todas las variables con rotación Varimax”*



*Ilustración 32 Análisis Factorial “Todas las variables con Rotación Varimax mediante máxima verosimilitud”*



*Ilustración 33 Análisis factorial “todas las variables con rotación”*

Con la rotación “Varimax” se logró llegar a una estructura más “simple”, donde las cargas de las variables se encuentran distribuidas entre los dos constructos facilitando también así la interpretación.

Se observa que el primer factor se relaciona con las variables “Levantamiento.Cadaveres”, “Orden.Publico”, “Fallecidos” y “Violencia.Intrafamiliar”, mientras que el segundo factor se relaciona con “Cerco.Epidemiologico”, “Estafas”, “Aglomeraciones”, “Confirmados”.

Si bien es cierto que una vez realizada la rotación se puede observar de mejor manera los agrupamientos de las variables, las interpretaciones permanecen casi las mismas en ambos casos siendo a su vez parecidas a los resultados obtenidos en el análisis de componentes principales.

Guayas se caracteriza por un valor alto en “Levantamiento.Cadaveres” y “Orden.Publico”, mientras que Pichincha por un valor alto en “Aglomeraciones” y “Cerco.epidemiologico”. Cabe notar que con la rotación “Varimax” Pichincha ya no parece estar tan caracterizada por valores altos en “Confirmados”.

Se podrían interpretar los constructos latentes rotados como las variables con más incidencia en las provincias de Guayas y Pichincha, siendo que en el primer factor cargan más las variables “Fallecidos”, “Orden.Publico”, “Violencia.Intrafamiliar” y “Levantamiento.Cadaveres” que son las que tienes más incidencia en Guayas mientras que las demás variables, cuyos valores altos caracterizaban a Pichincha, cargan más en el segundo factor.

### 6.3. Análisis de correspondencia

#### Fallecidos confirmados y probables

#### Perfil fila de Fallecidos confirmados y probables

Se presenta los perfiles resultantes de la tabla de contingencia con la que se trabajó para el presente análisis. Tenemos las variables, Grupo Etario y Provincia. El Grupo Etario está representado por las columnas y las provincias por las filas.

Obtenemos nuestra primera tabla o matriz de frecuencias relativas, representadas en porcentajes, a condición de estar en la fila provincia.

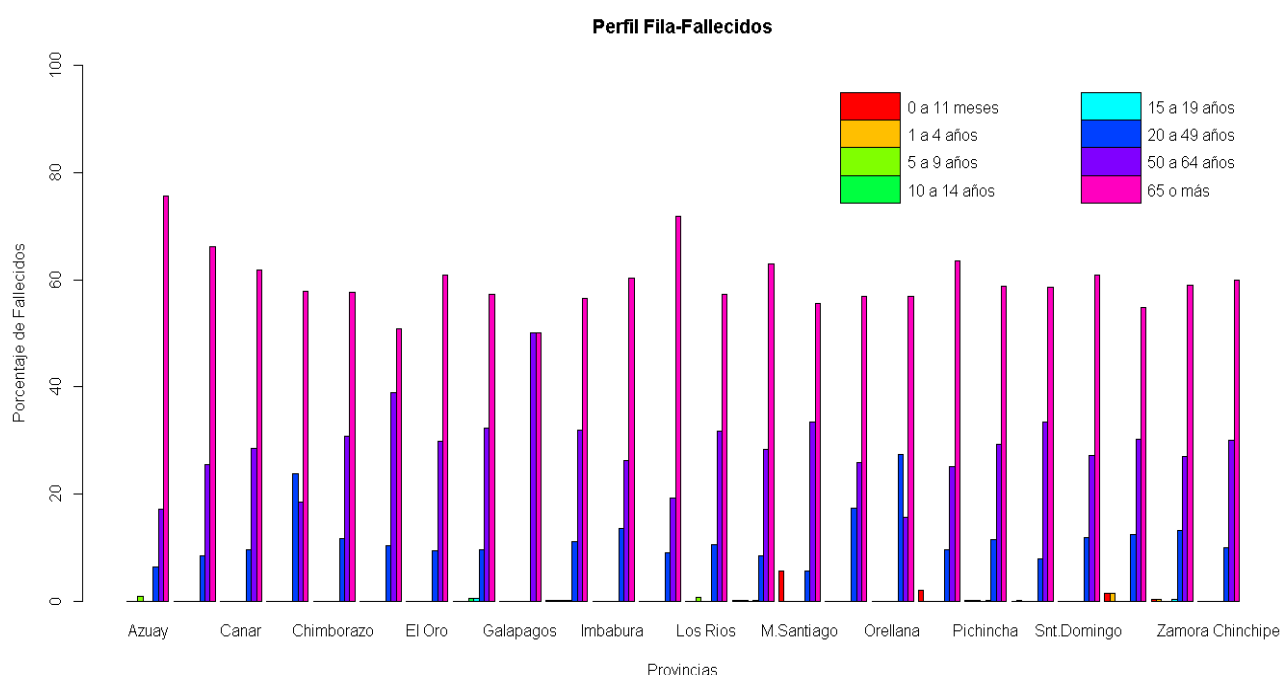
Provincias	0 a 11 meses	1 a 4 años	5 a 9 años	10 a 14 años	15 a 19 años	20 a 49 años	50 a 64 años	65 o más años
<b>Azuay</b>	0.00	0.00	0.90	0.00	0.00	6.31	17.12	75.68
<b>Bolívar</b>	0.00	0.00	0.00	0.00	0.00	8.47	25.42	66.10
<b>Cañar</b>	0.00	0.00	0.00	0.00	0.00	9.52	28.57	61.90
<b>Carchi</b>	0.00	0.00	0.00	0.00	0.00	23.68	18.42	57.89
<b>Chimborazo</b>	0.00	0.00	0.00	0.00	0.00	11.59	30.79	57.62
<b>Cotopaxi</b>	0.00	0.00	0.00	0.00	0.00	10.26	38.89	50.85
<b>El Oro</b>	0.00	0.00	0.00	0.00	0.00	9.35	29.78	60.87
<b>Esmeraldas</b>	0.00	0.00	0.00	0.44	0.44	9.61	32.31	57.21

<b>Galápagos</b>	0.00	0.00	0.00	0.00	0.00	0.00	50.00	50.00
<b>Guayas</b>	0.12	0.06	0.03	0.09	0.18	11.02	31.93	56.57
<b>Imbabura</b>	0.00	0.00	0.00	0.00	0.00	13.64	26.14	60.23
<b>Loja</b>	0.00	0.00	0.00	0.00	0.00	8.98	19.16	71.86
<b>Los Ríos</b>	0.00	0.00	0.59	0.00	0.00	10.50	31.68	57.23
<b>Manabí</b>	0.12	0.06	0.06	0.00	0.06	8.39	28.36	62.95
<b>Morona Santiago</b>	5.56	0.00	0.00	0.00	0.00	5.56	33.33	55.56
<b>Napo</b>	0.00	0.00	0.00	0.00	0.00	17.24	25.86	56.90
<b>Orellana</b>	0.00	0.00	0.00	0.00	0.00	27.45	15.69	56.86
<b>Pastaza</b>	1.92	0.00	0.00	0.00	0.00	9.62	25.00	63.46
<b>Pichincha</b>	0.09	0.09	0.09	0.00	0.19	11.47	29.32	58.74
<b>Santa Elena</b>	0.00	0.16	0.00	0.00	0.00	7.84	33.33	58.66
<b>Santo Domingo de los Tsáchilas</b>	0.00	0.00	0.00	0.00	0.00	11.93	27.16	60.91
<b>Sucumbíos</b>	1.37	1.37	0.00	0.00	0.00	12.33	30.14	54.79

<b>Tungurahua</b>	0.26	0.26	0.00	0.00	0.26	13.14	27.06	59.02
<b>Zamora Chinchipe</b>	0.00	0.00	0.00	0.00	0.00	10.00	30.00	60.00

*Tabla 22 Matriz de frecuencias relativas " Perfil fila de Fallecidos confirmados y probables "*

Consecuentemente se presenta el gráfico del perfil fila. Donde se evidencia, similitud entre la mayoría de las provincias a excepción de las siguientes: Azuay, Carchi, Esmeraldas, Galápagos, Los Ríos, Morona Santiago, Orellana, Pastaza, Sucumbíos y Tungurahua. De los mencionados se puede percatar de cierta similitud entre las provincias: Azuay y Los Ríos, Morona Santiago y Pastaza, Sucumbíos y Tungurahua.



*Ilustración 34 Gráfico "Perfil fila -fallecidos"*

### Perfil columna de Fallecidos confirmados y probables

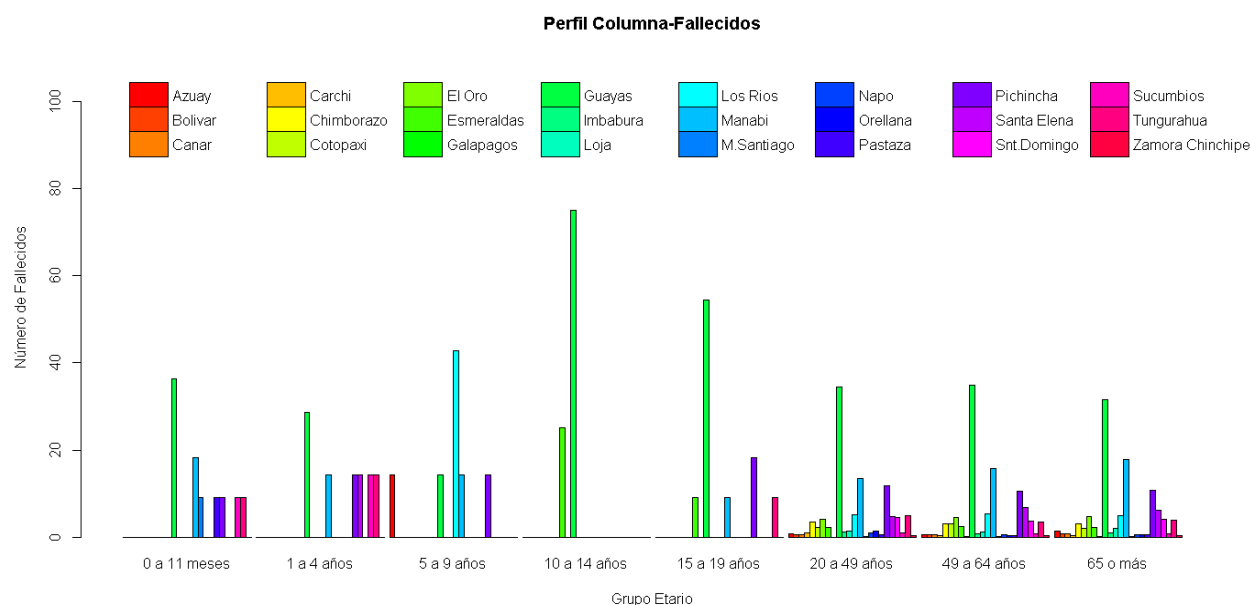
Posteriormente, se obtiene nuestra segunda tabla o matriz de frecuencias relativas, representadas en porcentajes, a condición de estar en la columna grupo etario.

<b>Provincias</b>	<b>0 a 11 meses</b>	<b>1 a 4 años</b>	<b>5 a 9 años</b>	<b>10 a 14 años</b>	<b>15 a 19 años</b>	<b>20 a 49 años</b>	<b>50 a 64 años</b>	<b>65 o más años</b>
<b>Azuay</b>	0.00	0.00	14.29	0	0.00	0.67	0.64	1.43
<b>Bolívar</b>	0.00	0.00	0.00	0	0.00	0.48	0.50	0.66
<b>Cañar</b>	0.00	0.00	0.00	0	0.00	0.58	0.60	0.66
<b>Carchi</b>	0.00	0.00	0.00	0	0.00	0.86	0.23	0.37
<b>Chimborazo</b>	0.00	0.00	0.00	0	0.00	3.36	3.11	2.96
<b>Cotopaxi</b>	0.00	0.00	0.00	0	0.00	2.30	3.04	2.03
<b>El Oro</b>	0.00	0.00	0.00	0	0.00	4.13	4.58	4.77
<b>Esmeraldas</b>	0.00	0.00	0.00	25	9.09	2.11	2.47	2.23
<b>Galápagos</b>	0.00	0.00	0.00	0	0.00	0.00	0.03	0.02
<b>Guayas</b>	36.36	28.57	14.29	75	54.55	34.55	34.88	31.48
<b>Imbabura</b>	0.00	0.00	0.00	0	0.00	1.15	0.77	0.90

<b>Loja</b>	0.00	0.00	0.00	0	0.00	1.44	1.07	2.04
<b>Los Ríos</b>	0.00	0.00	42.86	0	0.00	5.09	5.35	4.92
<b>Manabí</b>	18.18	14.29	14.29	0	9.09	13.44	15.82	17.88
<b>Morona Santiago</b>	9.09	0.00	0.00	0	0.00	0.10	0.20	0.17
<b>Napo</b>	0.00	0.00	0.00	0	0.00	0.96	0.50	0.56
<b>Orellana</b>	0.00	0.00	0.00	0	0.00	1.34	0.27	0.49
<b>Pastaza</b>	9.09	0.00	0.00	0	0.00	0.48	0.43	0.56
<b>Pichincha</b>	9.09	14.29	14.29	0	18.18	11.71	10.43	10.65
<b>Santa Elena</b>	0.00	14.29	0.00	0	0.00	4.61	6.82	6.11
<b>Santo Domingo de los Tsáchilas</b>	0.00	0.00	0.00	0	0.00	4.51	3.58	4.09
<b>Sucumbíos</b>	9.09	14.29	0.00	0	0.00	0.86	0.74	0.68
<b>Tungurahua</b>	9.09	14.29	0.00	0	9.09	4.89	3.51	3.90
<b>Zamora Chinchipe</b>	0.00	0.00	0.00	0	0.00	0.38	0.40	0.41

*Tabla 23 Matriz de frecuencias relativas "Perfil columna de Fallecidos Confirmados y Probables"*

Así, se obtiene el gráfico del perfil columna. En el cual se evidencia cierto grado de similitud entre los Grupo etarios: 20 a 49 años, 50 a 64 años y 65 años o más. Estos dos perfiles visualizados nos dan un indicio de cierta relación entre las modalidades descritas.



*Ilustración 35 Gráfico “Perfil columna- Fallecidos”*

### **Análisis de Correspondencia de Fallecidos confirmados y probables**

A través del gráfico podemos observar que la dimensión 1 absorbe la mayor cantidad de información, con un 31.89%, y la dimensión 2 absorbe una cantidad de inercia del 23.76%. De esta manera, evidenciamos como estas dos dimensiones poseen la mayor cantidad de información extraída de la tabla/matriz original.

Por otro lado, se evidencia una asociación entre la mayoría de provincias y los fallecimientos por covid-19 de los grupos etarios: 20 a 49 años, 50 a 64 años y 65 años o más. Sin embargo, Azuay con 5 a 9 años, Morona Santiago y Pastaza con 0 a 11 meses presentan una asociación entre ellos. A su vez entre: Sucumbíos y 1 a 4 años; y finalmente Galápagos y 15 a 19 años.



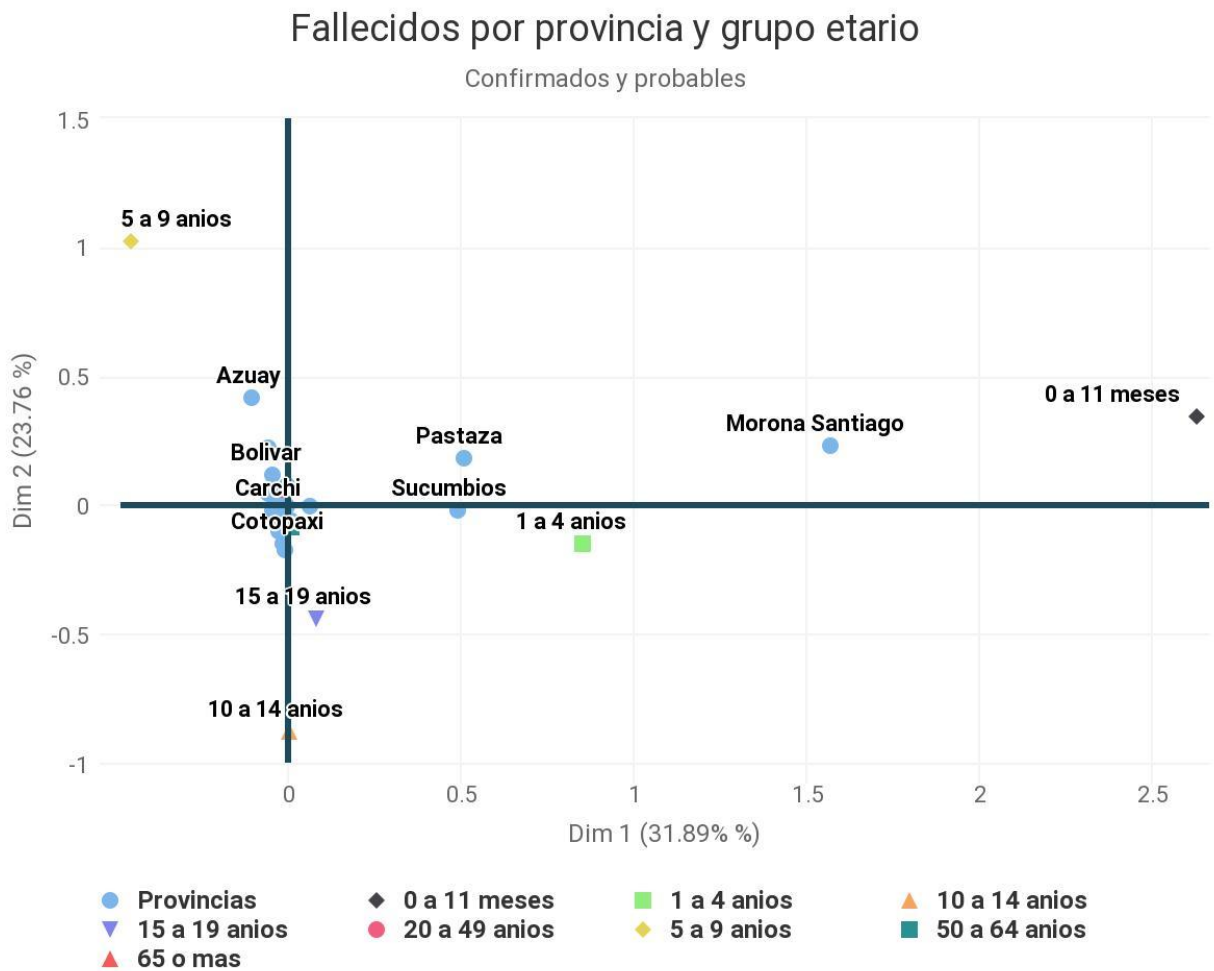


Ilustración 36 Gráfico “Fallecidos por provincia y grupo etario”

### Confirmados

#### Perfil fila de Confirmados

Se presenta los perfiles resultantes de la tabla de contingencia utilizado para el análisis. Tenemos las variables, Grupo Etario y Provincia. El Grupo Etario está representado por las columnas y las provincias por las filas.

Obtenemos la tabla o matriz de frecuencias relativas, representadas en porcentajes, a condición de estar en la fila provincia.

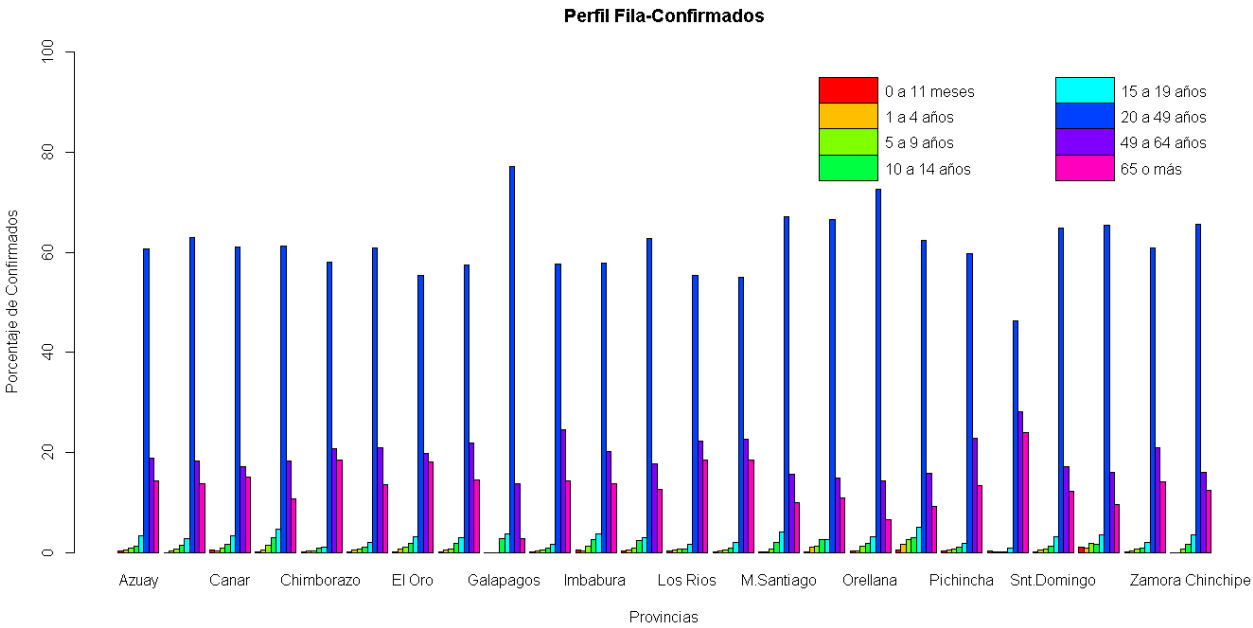
Provincias	0 a 11 meses	1 a 4 años	5 a 9 años	10 a 14 años	15 a 19 años	20 a 49 años	50 a 64 años	65 o más años

<b>Azuay</b>	0.29	0.42	0.95	1.32	3.25	60.65	18.88	14.24
<b>Bolívar</b>	0.00	0.38	0.61	1.37	2.74	62.91	18.20	13.79
<b>Cañar</b>	0.42	0.32	0.95	1.69	3.38	60.99	17.23	15.01
<b>Carchi</b>	0.07	0.54	1.42	2.90	4.72	61.23	18.34	10.79
<b>Chimborazo</b>	0.06	0.28	0.28	0.85	1.13	58.11	20.80	18.49
<b>Cotopaxi</b>	0.15	0.45	0.60	1.12	2.09	60.95	21.00	13.64
<b>El Oro</b>	0.18	0.74	1.08	1.87	3.05	55.33	19.71	18.04
<b>Esmeraldas</b>	0.06	0.44	0.73	1.80	3.03	57.55	21.89	14.50
<b>Galápagos</b>	0.00	0.00	0.00	2.75	3.67	77.06	13.76	2.75
<b>Guayas</b>	0.15	0.30	0.54	0.83	1.63	57.74	24.58	14.23
<b>Imbabura</b>	0.56	0.26	1.29	2.50	3.71	57.78	20.22	13.67
<b>Loja</b>	0.25	0.56	0.85	2.29	2.96	62.74	17.70	12.64
<b>Los Ríos</b>	0.23	0.46	0.72	0.76	1.71	55.39	22.27	18.46
<b>Manabí</b>	0.17	0.27	0.44	0.93	2.02	54.93	22.70	18.54

<b>Morona Santiago</b>	0.14	0.18	0.73	2.04	4.17	67.07	15.65	10.02
<b>Napo</b>	0.19	0.97	1.26	2.52	2.62	66.54	14.94	10.96
<b>Orellana</b>	0.27	0.33	1.20	1.79	3.05	72.58	14.21	6.57
<b>Pastaza</b>	0.40	1.67	2.53	3.04	4.94	62.44	15.85	9.13
<b>Pichincha</b>	0.26	0.40	0.65	1.00	1.81	59.74	22.78	13.36
<b>Santa Elena</b>	0.33	0.17	0.17	0.08	0.83	46.32	28.12	23.99
<b>Santo Domingo de los Tsáchilas</b>	0.14	0.40	0.71	1.33	3.05	64.88	17.17	12.31
<b>Sucumbíos</b>	1.11	0.95	1.85	1.58	3.59	65.35	16.01	9.56
<b>Tungurahua</b>	0.16	0.26	0.67	0.90	1.99	60.94	20.98	14.10
<b>Zamora Chinchipe</b>	0.00	0.00	0.74	1.67	3.53	65.68	16.05	12.34

*Tabla 24 Matriz de frecuencias relativas "Perfil fila de confirmados"*

Consecuentemente se presenta el gráfico del perfil fila. Donde se evidencia, similitud entre la mayoría de las provincias a excepción de las siguientes: Chimborazo, Galápagos, Napo, Santa Elena, Sucumbíos y Zamora Chinchipe.



*Ilustración 37 Grafico “perfil fila- confirmados”*

**Perfil columna de Confirmados**

Posteriormente, se obtiene la tabla o matriz de frecuencias relativas, representadas en porcentajes, a condición de estar en la columna grupo etario.

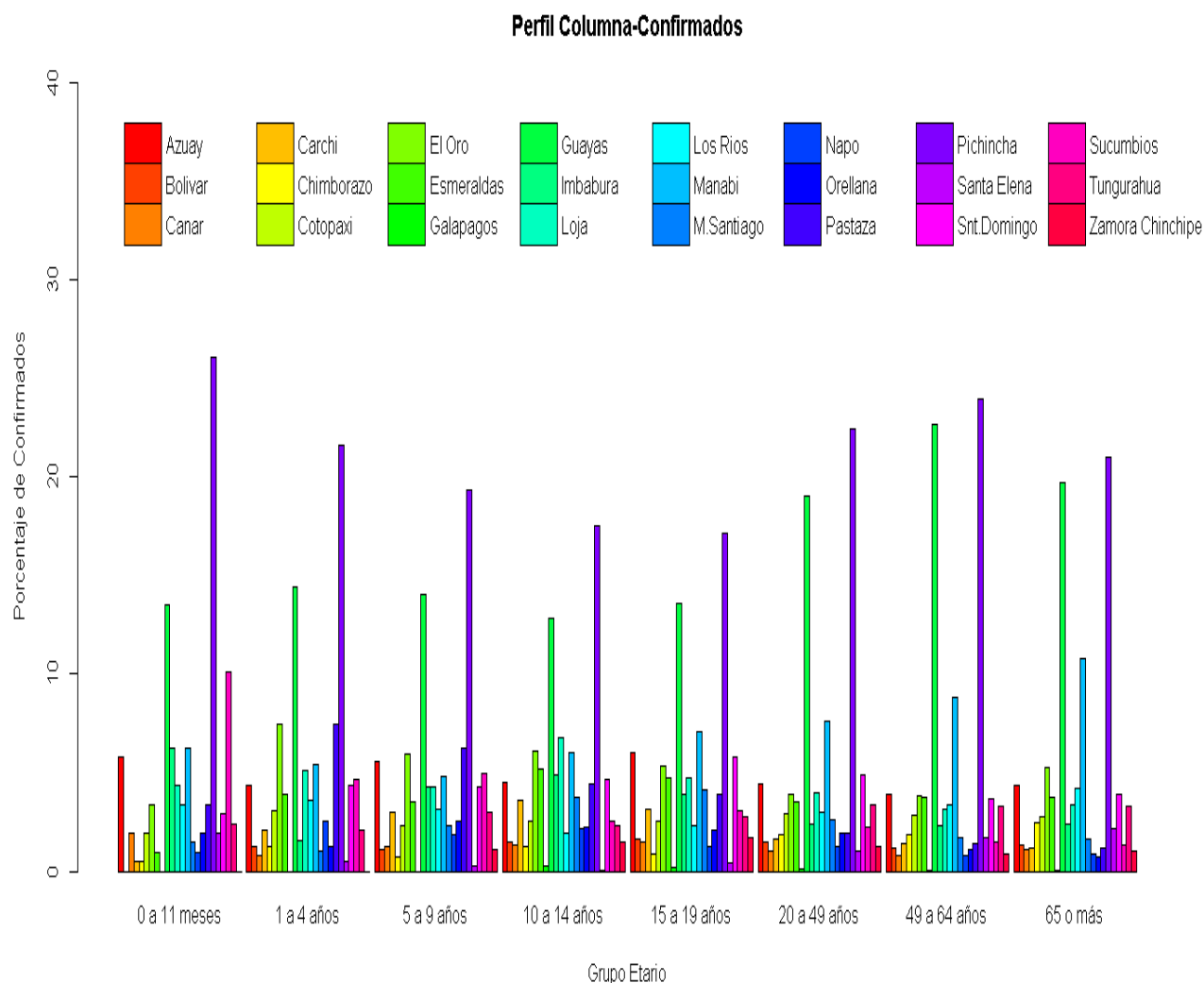
Provincias	0 a 11 meses	1 a 4 años	5 a 9 años	10 a 14 años	15 a 19 años	20 a 49 años	50 a 64 años	65 o más años
Azuay	5.80	4.38	5.55	4.52	6.02	4.45	3.88	4.38
Bolívar	0.00	1.29	1.14	1.51	1.63	1.48	1.20	1.36
Cañar	1.93	0.77	1.28	1.34	1.45	1.03	0.82	1.07
Carchi	0.48	2.06	2.99	3.60	3.17	1.63	1.37	1.20
Chimborazo	0.48	1.29	0.71	1.26	0.90	1.84	1.85	2.46

<b>Cotopaxi</b>	1.93	3.09	2.28	2.51	2.53	2.92	2.82	2.74
<b>El Oro</b>	3.38	7.47	5.97	6.11	5.38	3.87	3.86	5.29
<b>Esmeraldas</b>	0.97	3.87	3.56	5.19	4.70	3.54	3.77	3.74
<b>Galápagos</b>	0.00	0.00	0.00	0.25	0.18	0.15	0.08	0.02
<b>Guayas</b>	13.53	14.43	14.08	12.81	13.57	19.05	22.71	19.70
<b>Imbabura</b>	6.28	1.55	4.27	4.86	3.89	2.40	2.35	2.38
<b>Loja</b>	4.35	5.15	4.27	6.78	4.75	3.98	3.15	3.37
<b>Los Ríos</b>	3.38	3.61	3.13	1.93	2.35	3.02	3.40	4.23
<b>Manabí</b>	6.28	5.41	4.84	6.03	7.06	7.61	8.81	10.78
<b>Morona Santiago</b>	1.45	1.03	2.28	3.77	4.16	2.65	1.73	1.66
<b>Napo</b>	0.97	2.58	1.85	2.18	1.22	1.23	0.77	0.85
<b>Orellana</b>	1.93	1.29	2.56	2.26	2.08	1.96	1.07	0.74
<b>Pastaza</b>	3.38	7.47	6.26	4.44	3.89	1.95	1.39	1.20
<b>Pichincha</b>	26.09	21.65	19.35	17.50	17.19	22.42	23.95	21.05

<b>Santa Elena</b>	1.93	0.52	0.28	0.08	0.45	1.00	1.71	2.18
<b>Santo Domingo de los Tsáchilas</b>	2.90	4.38	4.27	4.69	5.83	4.91	3.64	3.91
<b>Sucumbíos</b>	10.14	4.64	4.98	2.51	3.08	2.22	1.52	1.36
<b>Tungurahua</b>	2.42	2.06	2.99	2.35	2.80	3.40	3.28	3.30
<b>Zamora Chinchipe</b>	0.00	0.00	1.14	1.51	1.72	1.27	0.87	1.00

*Tabla 25 Matriz de frecuencia relativa "Perfil columna de Confirmados"*

Posteriormente, se obtiene el gráfico del perfil columna. En el cual se evidencia cierto grado de similitud entre los Grupo etarios: 1 a 4 años y 5 a 9 años, 10 a 14años y 15 a 19 años; y entre 20 a 49 años, 50 a 64 años y 65 años o más. Estos dos perfiles visualizados nos dan un indicio de cierta relación entre las modalidades descritas.

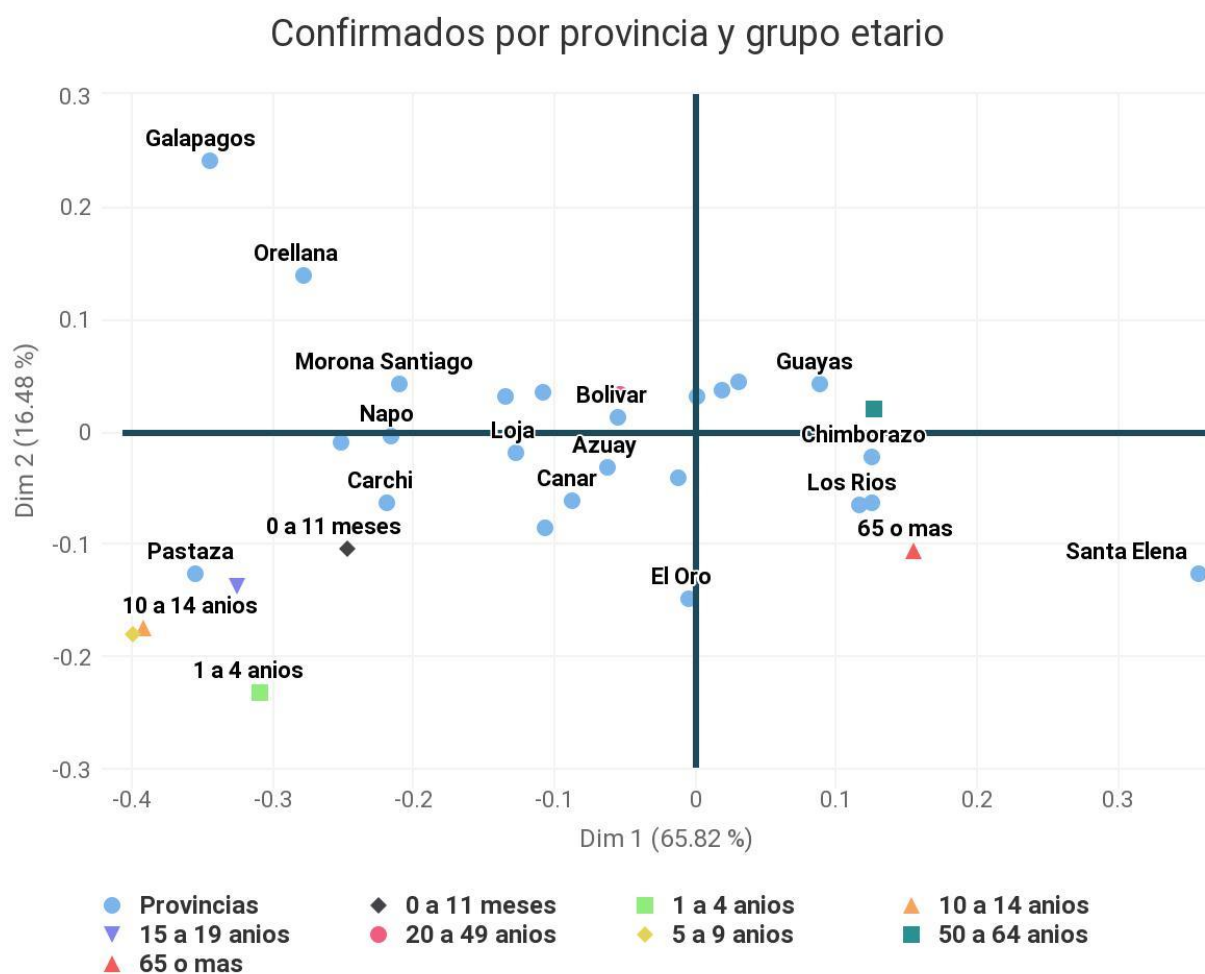


*Ilustración 38 Gráfico “Perfil columna- Confirmados”*

### **Análisis de correspondencia de Confirmados**

Mediante el gráfico se aprecia que la dimensión 1 absorbe la mayor cantidad de información, con un 65.82%, el cual además es el que discrimina los grupos etarios; y la dimensión 2 absorbe una cantidad de inercia del 16.48%. De esta manera, evidenciamos como estas dos dimensiones poseen la mayor cantidad de información extraída de la tabla/matriz original.

Por otro lado, se evidencia una asociación entre Pastaza y los confirmados por covid-19 de los grupos: 1 a 4 años, 5 a 9 años, 10 a 14 años y 15 a 19 años. A su vez, Carchi, Sucumbíos, Napo e Imbabura y 0 a 11 meses. También por parte de las provincias: Bolívar, Cotopaxi, Santo Domingo de los Tsáchilas, Zamora Chinchipe y 20 a 49 años. Finalmente, Guayas con 50 a 64 años; y Los Ríos, Manabí, Santa Elena y Chimborazo muestran asociaciones con el grupo 65 años o más.



*Ilustración 39 Gráfico “Confirmados por provincia y Grupo etario”*



## 7. Conclusiones

- Se puede evidenciar la diferencia entre las provincias de la Costa, con las provincias de las demás regiones, la propagación del virus, al principio, tuvo un comportamiento más “agresivo” en la costa que en la sierra; no obstante, con el paso del tiempo, las personas susceptibles al contagio disminuyeron en la costa, y en la sierra, por la implementación de medidas preventivas, empezaron a contagiarse con una velocidad mayor a la que tenían los primeros días. Siendo Pichincha, la provincia con mayor número de contagios.
- Si centráramos el análisis en las variables tomadas de las alertas del 911, podríamos inferir que en la región Sierra, se tienen más llamadas con fin de prevención, como alertas por posible violación del cerco epidemiológico, aglomeraciones, entre otras, sin embargo en la Costa, las alertas son para anunciar algo ya ha pasado, por ejemplo: Violencia intrafamiliar, levantamiento de cadáveres, etc.
- Hay una asociación bastante marcada entre Pastaza, Sucumbíos, Napo y los niños y adolescentes confirmados por covid-19.
- Los Ríos, Manabí, Santa Elena y Chimborazo son las provincias donde más se evidencia una asociación con los confirmados por covid-19 del grupo etario de 65 años o más.
- Los decesos de niños y adolescentes menores a 15 años no se ven asociados con las provincias de mayor impacto por covid-19. En particular los niños menores de 5 años fallecidos tienen una asociación con provincias como Sucumbíos, Pastaza, Morona Santiago; que pertenecen a la Amazonía Ecuatoriana.

## 8. Referencias bibliográficas y electrónicas

### Bibliografía

- Zurita, G. (2010). Probabilidad y Estadística Fundamentos y Aplicaciones. *Ediciones del Instituto de Ciencias Matemáticas ESPOL, Guayaquil-Ecuador*.
- Pérez, E. R., & Medrano, L. A. (2010). Análisis factorial exploratorio: bases conceptuales y metodológicas. *Revista Argentina de Ciencias del Comportamiento (RACC)*, 2(1), 58-66.
- Freire Álvarez, L. (2016). *Procesamiento de Datos usando análisis de Correspondencia Múltiples* (Bachelor's thesis, Espol).
- Mendieta, I. D. L. L. B. (2020). Incidencias del Covid-19 en Ecuador. *Question/Cuestión*, 1(mayo), e321-e321.
- Ruiz, G. P. I., & León, A. C. I. (2020). Evolución de la enfermedad por coronavirus (COVID-19) en Ecuador. *La Ciencia al Servicio de la Salud*, 11(1), 5-15.

### Anexo 1 Ventana inicial del Dashboard