**INSTITUTO TECNOLÓGICO DE SAN LUIS POTOSÍ**



“**Análisis de información referente a casos de Morbilidad asociados a COVID-19, obtenida de los Anuarios Estadísticos de Morbilidad 2015-2017 de la Dirección General de Epidemiología**”

### Maestría en Ingeniería Electrónica

**Alumno:** Juan Antonio Alvarado Cano

**Profra:** Dra. Luz Roxana de León Lomelí

**Materia:** Programación avanzada

**Análisis de información referente a casos de Morbilidad asociados a COVID-19, obtenida de los Anuarios Estadísticos de Morbilidad 2015-2017 de la Dirección General de Epidemiología**

Se descargaron los datos publicados por la Dirección General de Epidemiología, que pone a disposición de la población en general, la información contenida en los Anuarios Estadísticos de Morbilidad 2015-2017, así como la información referente a los casos asociados a COVID-19.

Se tomaron un total de 10000 registros con datos de cuatro factores, además del resultado de la prueba, que se presentaban en personas que se realizaron la prueba de antígeno para detección de casos asociados a COVID 19.

**Factores para analizar**

|  |  |
| --- | --- |
| 1 | Entidad en la que se realizó la prueba |
| 2 | Sexo |
| 3 | Edad |
| 4 | Tabaquismo |
| 5 | Resultado de la prueba de antígeno |

**Catálogo de datos**

1. Entidades en las que se llevó a cabo la prueba del antígeno

|  |  |  |
| --- | --- | --- |
| **CLAVE\_ENTIDAD** | **ENTIDAD\_FEDERATIVA** | **ABREVIATURA** |
| 01 | AGUASCALIENTES | AS |
| 02 | BAJA CALIFORNIA | BC |
| 03 | BAJA CALIFORNIA SUR | BS |
| 04 | CAMPECHE | CC |
| 05 | COAHUILA DE ZARAGOZA | CL |
| 06 | COLIMA | CM |
| 07 | CHIAPAS | CS |
| 08 | CHIHUAHUA | CH |
| 09 | CIUDAD DE MÉXICO | DF |
| 10 | DURANGO | DG |
| 11 | GUANAJUATO | GT |
| 12 | GUERRERO | GR |
| 13 | HIDALGO | HG |
| 14 | JALISCO | JC |
| 15 | MÉXICO | MC |
| 16 | MICHOACÁN DE OCAMPO | MN |
| 17 | MORELOS | MS |
| 18 | NAYARIT | NT |
| 19 | NUEVO LEÓN | NL |
| 20 | OAXACA | OC |
| 21 | PUEBLA | PL |
| 22 | QUERÉTARO | QT |
| 23 | QUINTANA ROO | QR |
| 24 | SAN LUIS POTOSÍ | SP |
| 25 | SINALOA | SL |
| 26 | SONORA | SR |
| 27 | TABASCO | TC |
| 28 | TAMAULIPAS | TS |
| 29 | TLAXCALA | TL |
| 30 | VERACRUZ DE IGNACIO DE LA LLAVE | VZ |
| 31 | YUCATÁN | YN |
| 32 | ZACATECAS | ZS |
| 36 | ESTADOS UNIDOS MEXICANOS | EUM |
| 97 | NO APLICA | NA |
| 98 | SE IGNORA | SI |
| 99 | NO ESPECIFICADO | NE |

1. El género (sexo), de las personas a las que se realizó la prueba de antígeno para detección de COVID 19

|  |  |
| --- | --- |
| **CLAVE** | **DESCRIPCIÓN** |
| 1 | MUJER |
| 2 | HOMBRE |
| 99 | NO ESPECIFICADO |

1. Edades de las personas que se realizaron la prueba de antígeno para detección de COVID 19, edades que van desde los 0 y hasta los 99 años.
2. Si la persona fuma.

|  |  |
| --- | --- |
| **CLAVE** | **DESCRIPCIÓN** |
| 1 | SI |
| 2 | NO |
| 97 | NO APLICA |
| 98 | SE IGNORA |
| 99 | NO ESPECIFICADO |

1. Finalmente, el resultado de la prueba del antígeno

|  |  |
| --- | --- |
| **CLAVE** | **DESCRIPCIÓN** |
| 1 | POSITIVO A SARS-COV-2 |
| 2 | NEGATIVO A SARS-COV-2 |
| 97 | NO APLICA (CASO SIN MUESTRA) |

Con base a los datos seleccionados, se destaca la siguiente información:

1. Con una muestra de 10000 registros se obtuvo que la mayoría de las personas a las que se les realizó la prueba son de la Ciudad de México (4422 44.22%) y la entidad que menos registro tuvo de pruebas es Tlaxcala (17 0.17%)

|  |
| --- |
| >> tabulate(entidad\_um)  Value Count Percent  1 83 0.83%  2 138 1.38%  3 422 4.22%  4 33 0.33%  5 371 3.71%  6 123 1.23%  7 75 0.75%  8 81 0.81%  9 4422 44.22%  10 189 1.89%  11 448 4.48%  12 58 0.58%  13 80 0.80%  14 266 2.66%  15 426 4.26%  16 85 0.85%  17 39 0.39%  18 45 0.45%  19 576 5.76%  20 80 0.80%  21 143 1.43%  22 190 1.90%  23 157 1.57%  24 196 1.96%  25 90 0.90%  26 112 1.12%  27 434 4.34%  28 99 0.99%  29 17 0.17%  30 104 1.04%  31 294 2.94%  32 123 1.23% |

1. Con respecto al género de las personas que se realizaron la prueba de antígeno para detección de COVID19 se encontró que fueron:

Mujeres: 5611

Hombres: 4388

|  |
| --- |
| >> cantidad=numel(sexo(sexo==1))  cantidad =  5611  >> cantidad=numel(sexo(sexo==2))  cantidad =  4388 |

1. La edad de las personas que se tomaron en cuenta en la muestra es de los 0 a los 99 años, con una edad promedio de 37.73 años. La edad más común entre las personas que se realizaron la prueba fue la de 25 años y la desviación estándar fue de 17.7431

|  |
| --- |
| m\_um=mean(edad)  max\_edad=max(edad)  min\_edad=min(edad)  moda\_edad=mode(edad)  mediana\_edad=median(edad)  desv\_edad=std(edad)  m\_um = 37.7307  max\_edad = 99  min\_edad = 0  moda\_edad = 25  mediana\_edad = 36  desv\_edad = 17.7431 |

1. Con respecto al factor de tabaquismo se encontró que entre los 10000 registros que se seleccionaron se encontró que:

Fuman: 473

No fuman: 9472

Se ignora: 54

|  |
| --- |
| >> cantidad=numel(tabaquismo(tabaquismo==1))  cantidad =  473  >> cantidad=numel(tabaquismo(tabaquismo==2))  cantidad =  9472  >> cantidad=numel(tabaquismo(tabaquismo==97))  cantidad =  0  >> cantidad=numel(tabaquismo(tabaquismo==98))  cantidad =  54  >> cantidad=numel(tabaquismo(tabaquismo==99))  cantidad =  0 |

1. Finalmente, con respecto a lo concerniente al estudio de la prueba de antígeno se encontraron los siguientes datos:

Positivo a SARS-COV-2: 3181

Negativo a SARS-COV-2: 5318

No aplica: 1500

|  |
| --- |
| >> cantidad=numel(r\_antigeno(r\_antigeno==1))  cantidad =  3181  >> cantidad=numel(r\_antigeno(r\_antigeno==2))  cantidad =  5318  >> cantidad=numel(r\_antigeno(r\_antigeno==97))  cantidad =  1500  >> cantidad=numel(r\_antigeno(r\_antigeno==98))  cantidad =  0  >> cantidad=numel(r\_antigeno(r\_antigeno==99))  cantidad =  0 |

**Depuración de datos**

Con el fin de garantizar un análisis preciso, eficiente o significativo. La depuración de datos alude a todos aquellos métodos que se utilizan para buscar, eliminar y reemplazar datos incorrectos o ausentes. La detección de cambios de extremos locales y cambios abruptos puede ayudar a identificar tendencias de datos significativas. El suavizado y la eliminación de tendencia son procesos que se emplean para eliminar el ruido y las tendencias polinómicas de los datos.

Se procedió con la depuración de los datos que así lo necesitaban:

|  |  |
| --- | --- |
| Entidad federativa:  Se depuran los “no aplica”, “se ignora” y “no especificado” | for i=1:10000  if entidad\_um(i)==97;  entidad\_um(i)=0;  end  if entidad\_um(i)==98; %if c(i)~=97;  entidad\_um(i)=0;  end  if entidad\_um(i)==99;  entidad\_um(i)=0;  end  end |
| Sexo:  Se depuran los “no especificado” | for i=1:10000  if sexo(i)==99;  sexo(i)=0;  end  end |
| Tabaquismo:  Se depuran los “no aplica”, “se ignora” y “no especificado” | for i=1:10000  if tabaquismo(i)==97;  tabaquismo(i)=0;  end  if tabaquismo(i)==98; %if c(i)~=97;  tabaquismo(i)=0;  end  if tabaquismo(i)==99;  tabaquismo(i)=0;  end  end |
| Se depuran los “no aplica” (caso sin muestra) | for i=1:10000  if r\_antigeno(i)==97;  r\_antigeno(i)=0;  end  end |

**Correlaciónes**

Una vez aplicada la depuración de los datos se realiza un análisis de correlación entre el resultado de la prueba de antígeno y cada uno de los otros factores

|  |
| --- |
| CR1=corrcoef(entidad\_um,r\_antigeno)  CR2=corrcoef(sexo,r\_antigeno)  CR3=corrcoef(edad,r\_antigeno)  CR4=corrcoef(tabaquismo,r\_antigeno)  figure  subplot(221),heatmap(CR1)  subplot(222),heatmap(CR2)  subplot(223),heatmap(CR3)  subplot(224),heatmap(CR4) |

Interfaz de usuario gráfica, Aplicación, Teams, Gráfico de rectángulos

Descripción generada automáticamente

**Conclusiones**

Después de analizar los resultados de las correlaciones se puede observar que:

1. En el caso de la correlación entre la entidad donde se aplica la prueba y los resultados de la prueba de antígeno, existe una ligera relación y considero que es a partir de que, a mayor población, mayor cantidad de contagios y el resultado es negativo debido a que la relación es por contagios en los distintos lugares que se llevaron a cabo las pruebas

Gráfico

Descripción generada automáticamente

1. La correlación entre el sexo de las personas y la prueba de antígeno se encuentra con muy poca relación y se debe principalmente a que se realizaron la prueba más mujeres que hombres

Imagen que contiene Gráfico

Descripción generada automáticamente

1. En el caso de la correlación de la edad con los resultados de la prueba de antígeno se encontró que no hay una dependencia clara debido a que en todas las edades registradas se detectaron contagios y nuevamente al encontrarse la correlación negativa, tenemos que la relación es más bien de contagios en las distintas edades

Gráfico

Descripción generada automáticamente

1. La última correlación se analizó entre personas que fuman y los resultados de la prueba de antígeno, donde se encontró una baja relación entre las personas que fuman y que resultaron contagiados y puede observarse que es muy poca la población que dijo ser fumadora

Gráfico

Descripción generada automáticamente