## 

**INFORME**

**Integrantes:** Cristian Matus

Cristian Molina

Roderick Zambrano

Ángelo Rodríguez

**Asignatura:** Minería De Datos

# **Índice**

**Contenido**

[Índice 2](#_Toc54200006)

[Introducción 3](#_Toc54200007)

[Reconocimiento del problema 3](#_Toc54200008)

[Contexto de la problemática 3](#_Toc54200009)

[Consecuencias 3](#_Toc54200010)

[Crear estrategia de visión de la calidad de los datos 3](#_Toc54200011)

[Identificar variables, para poder generar conocimiento 3](#_Toc54200012)

[Buscar fuentes de datos 3](#_Toc54200013)

[Preprocesamiento bajo directrices en calidad de los datos 3](#_Toc54200014)

[Directrices de como maneja 3](#_Toc54200015)

[Ausencia de valores en las fuentes de datos 3](#_Toc54200016)

[Inconsistencia de los datos de la fuente de datos 3](#_Toc54200017)

[Duplicidad de los datos de la fuente de datos 3](#_Toc54200018)

[Datos anómalos de la fuente de datos 3](#_Toc54200019)

[Fuente de datos 4](#_Toc54200020)

[Generación de análisis 4](#_Toc54200021)

[Conclusión 4](#_Toc54200022)

[Referencias y/o bibliografía 4](#_Toc54200023)

# **Introducción**

En el presente informe presentaremos un desarrollo relacionado a las herramientas de minería de datos de acuerdo a la problemática de las enfermedades cardiovasculares, las cuales son una de las causantes de muertes más comunes en nuestro país. Realizando un análisis de datos que se extraen a través de distintas fuentes de información en América latina. Arrojando como resultado las variantes a la problemática actual para poder facilitar la toma de decisiones con respecto a cómo resolver la problemática, para obtener los resultados este documento se basará en los siguientes puntos:

* Reconocimiento del problema
* Creación de estrategia de visión de la calidad de los datos
* Preprocesamiento bajo directrices en calidad de los datos

# **Reconocimiento del problema**

Las enfermedades cardiovasculares no son un problema solamente de Chile, sino que también es un problema de salud y un reto global. Pero en nuestro país es una de las principales causas de muertes, representando más del 25% de las defunciones en el país. Estas causantes de muertes esta relacionada por ejemplo a enfermedades hereditarias, mala alimentación, tabaquismo, alcoholismo, vida sedentaria, etc.

## **Contexto de la problemática**

El alcance del informe es abarcar las enfermedades cardiovasculares relacionadas a la obesidad en Chile. Ya que la obesidad cada vez mas se convierte en un problema a nivel país.

## **Consecuencias**

Las consecuencias de no contar con una buena cantidad de datos que nos permita identificar las variables que afectan a la población en cuanto a la previsión de enfermedades cardiovasculares, nos evitara de poder tomar acciones preventivas ante esta situación. Sabemos que no solamente es un problema de salud publica en Chile, sino que también es un problema mundial que año a año se lleva a millones de personas. Por lo anterior es muy importante contar con datos fidedignos para la toma de decisiones y así realizar campañas y educar a la población.

# **Crear estrategia de visión de la calidad de los datos**

Relacionadas a las enfermedades cardiovasculares por la obesidad, la estrategia que se utilizará al momento de ver la calidad de los datos utilizados debe estar dentro de los atributos de exactitud, completitud, integridad, actualización, coherencia y relevancia.

## **Identificar variables, para poder generar conocimiento**

Las fuentes de datos que se tomaran en cuenta para realizar este informe son:

* Identificar regiones o comunas con mayor alto índice de obesidad en el país.
* Edades de víctimas, en tratamiento y recuperados según su género.
* Dato de víctimas en los últimos 12 meses.
* Dato de personas en tratamiento según su género.
* Dato de personas recuperadas en los últimos 12 meses.

## Buscar fuentes de datos

# 

# Preprocesamiento bajo directrices en calidad de los datos

## Directrices de como maneja

### 

### Ausencia de valores en las fuentes de datos

### Inconsistencia de los datos de la fuente de datos

### 

### Duplicidad de los datos de la fuente de datos

### 

### Datos anómalos de la fuente de datos

# Fuente de datos

# 

# Generación de análisis

# Conclusión

# Referencias y/o bibliografía