

PEC 3 - TIPOLOGÍA DE DATOS - ANÁLISIS Y LIMPIEZA DE DATOS

Javier Botija y Nelson Salinas

25/05/2021

Desarrollo de la actividad

Tema principal: Análisis de datos en pacientes con/sin diabetes en referencia a la pandemia actual COVID 19

Introducción

Se han confirmado estudios clínicos que los pacientes con enfermedades crónicas que tienen infección por algún tipo de coronavirus, son más propensos en desarrollar otras enfermedades y recuperarse e incluso en responder a un tratamiento adecuado. En relación a la pandemia actual COVID 19, esta enfermedad se transmite por gotas expiradas por personas contagiadas o por contacto directo, esto hace que se vuelva mayormente transmitible, por lo que los cuidados deben ser más exigibles para evitar el contagio. Los pacientes con diabetes tienen más probabilidades a tener este tipo de enfermedad, según el Profesor Andrew Boulton presidente de la Federación Internacional de Diabetes explica en una entrevista, las personas con diabetes son más propensas en enfermarse y tratar debido a las fluctuaciones de glucosa en la sangre por lo que incluso pueden llegar a morir sino llegan a tomar las precauciones necesarias. Entrevista Andrew Boulton.

Este trabajo será útil para poder analizar la cantidad de pacientes que se detectan con diabetes y que relación tienen con otras variables, por lo que se puede crear un plan de acción para generar recomendaciones a estos tipos de pacientes por la parte médica correspondiente.

Descripción

El desarrollo de esta actividad abarcará la resolución de los puntos determinados para cada actividad, se tomará en referencia una base de datos de pacientes que se han realizado exámenes de glucosa y tienen como referencia si tienen o no diabetes, esto permitirá resolver cual es la relación que hay entre ciertas variables de análisis y los cuidados respectivos en cuanto a las infecciones por COVID 19.

Objetivos de la actividad

- Aplicar los conocimientos obtenidos para poder tratar un Dataset de información.
- Poder analizar las variables y su relación.
- Concluir y resolver el problema planteado.

Problemática

La problemática que se desea resolver es verificar cual es la incidencia de diabeticos entre los dos sexos participantes, y a ello considerar cuales son los pacientes que deben ser mayormente vigilados en relación a la actual pandemia que estamos atrevasando, para la resolución del problema se tomará en cuenta la relación que tienen las variables y con ello poder concluir con un resultado final.

Dataset

El dataset fue obtenido desde la base de datos del Hospital Clínica San Francisco. ubicado en la ciudad de Guayaquil país Ecuador, el mismo que ha otorgado dicha información para poder realizar las actividades para esta tarea. Cabe destacar que por ser información privada no se referencia la identificación de pacientes, es decir no existen los datos como nombres o DNI, tambien se considera el respeto por la legislación ecuatoriana en fundamento a la Ley de Protección de Datos indicada en el siguiente link [Ministerio de Salud](#). . Las características del archivo de datos son:

Nombre del dataset: DIABETES_HCSF

Formato: XLS

Cantidad de registros: 77582 registros

Variables: 16

- SEXO: Masculino o Femenino.
- ESTADOCIVIL: Hace referencia al estado civil del paciente.
- FECHANACIMIENTO: Fecha de nacimiento del paciente.
- CODPACIENTE: Código interno del paciente para identificación del exámen.
- INGRESO: Número de ingreso del paciente para registro interno del Hospital.
- FECHAINGRESO: Fecha de ingreso que tuvo el paciente al ser hospitalizado.
- FECHASALIDA: Fecha en la que salio el paciente o se le dio de alta.
- CODTIPOINGRESO: Código interno por el tipo de ingreso.
- TIPOINGRESO: Tipo de ingreso que tuvo el paciente cuando fue admisionado en el Hospital.
- CODTIPOALTA: Código interno para identificar el tipo de alta.
- TIPOALTA: Tipo de alta que tuvo el paciente al momento de irse del Hospital.
- CODAREA: Código interno para identificar en el area que estuvo el paciente.
- HABITACION: Nombre de la habitación donde estuvo el paciente.
- DIAGNOSTICOINGRESO: Diagnóstico por el cual el paciente fue admisionado.
- GLUCOSA: Valor de la glucosa en formato numérico
- REFERENCIADIABETES: Hace referencia si un paciente tiene o no diabetes, en caso de tener * es positivo.