

ROMAC

DIPLOMADO DE INTELIGENCIA ARTIFICIAL

Defunción por Diabetes Mellitus

Liliana Becerril Tapia

1/diciembre/2024

Abstract

A dual energy radiography method using basis decomposition was developed, the process to do it is shown and it is compared against an alternate more direct method of analyzing the data using the logarithm of the original data, concluding that this second method does work but it is not better than basis decomposition.

1 Introducción

A partir de datos de obtenidos de el INEGI,

2 Marco Teorico

2.1 Diabetes

La diabetes mellitus, también conocida como diabetes, es una enfermedad metabólica que se caracteriza por niveles altos de glucosa en la sangre. Esto se debe a que el cuerpo no produce suficiente insulina o no la utiliza de forma adecuada. La glucosa es la principal fuente de energía del cuerpo, por lo que cuando sus niveles son altos, no llega a las células y puede provocar problemas de salud graves.

3 Resultados

3.1 Análisis exploratorio

Realiza un análisis sobre el conjunto de datos limpio intentando contestar las preguntas de negocio. En esta parte tienes libertad de mostrar tus hallazgos y puedes utilizar los métodos y gráficos que consideres adecuados.

La evaluación de esta sección es un tercio del valor del examen y se evaluará según justifiques cada pregunta.

Importante: Es importante que comentes tu código y escribas paso a paso lo que estás haciendo. Este notebook debe ser un reporte con código y debe poder entenderse claramente tanto procedimiento como explicación (apóyate de las celdas de texto).

Todo lo que hayas usado para explorar que no sirva para responder las preguntas, debes eliminarlo.

Lee el conjunto de datos del INEGI sobre defunciones en México en 2015. Analiza los registros utilizando los métodos y procedimientos vistos en clase.

El análisis y el reporte que hagas debe contestar las siguientes preguntas.

¿Hay similitud o diferencia entre las características de los casos de defunción por diabetes mellitus entre hombres y mujeres? ¿Cuáles son las características principales que observas para las personas que fallecieron por causa de la diabetes mellitus?

¿En qué meses y en qué estados hay mayor abundancia de defunciones por la enfermedad?

3.1.1 Limpieza de datos

1. Con lambda filtre todas las filas que comenzaran con "Diabetes mellitus" que estuvieran en la columna "CAUSA_DEF" y las renombre a "Diabetes mellitus" para que fuera facil graficarlas.

2. Despues seleccione todas las filas que tuvieran "Diabetes mellitus"
3. En la parte de defuncion por mes y estado, elimine los valores para mes de 99 ya que solo hay 12 meses, en la parte de estado elimine todos aquellos que hablaran de manera general o de paises en vez de estados de México.
4. Procedi a graficar.

Primero se identificaron los nombres que se repetían para el caso de La diabetes, de acuerdo a la gráfica de la Figura 4 de barras, es la principal causa de muerte en México, de ahí le sigue "Infarto agudo del miocardio, sin otra especificación", "Neumonía no especificada", etc.

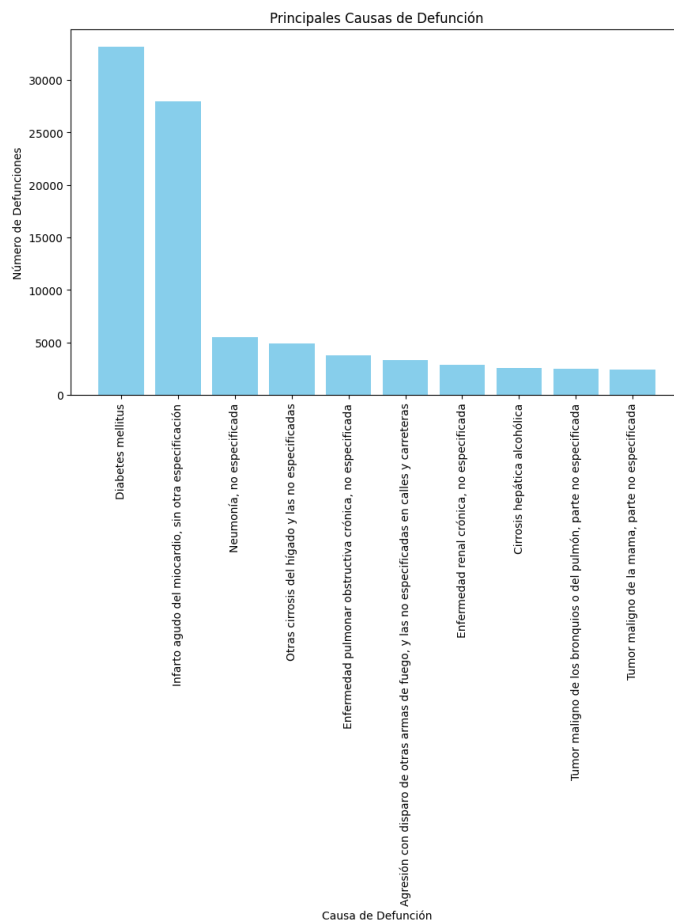


Figure 1: Principales causas de defunción en México.

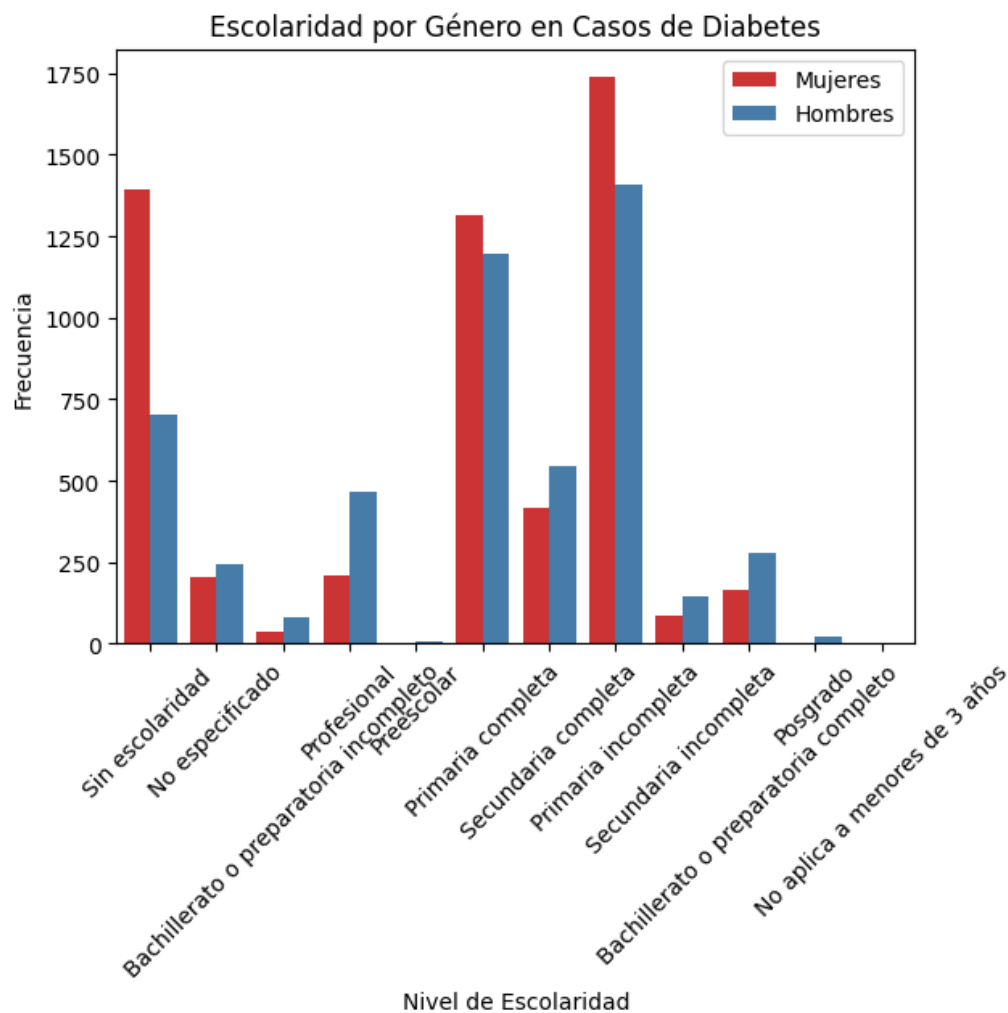


Figure 2: Escolaridad por Género en Casos de Diabetes.

El gráfico de la figura 3 podría sugerir que:

La edad promedio de fallecimiento por diabetes mellitus es muy similar entre hombres y mujeres. Los casos extremos (outliers) podrían indicar situaciones particulares como diabetes de inicio temprano, complicaciones graves, o casos en edades avanzadas. Este análisis puede ser útil para diseñar intervenciones dirigidas a personas de edad avanzada, quienes son las más afectadas.

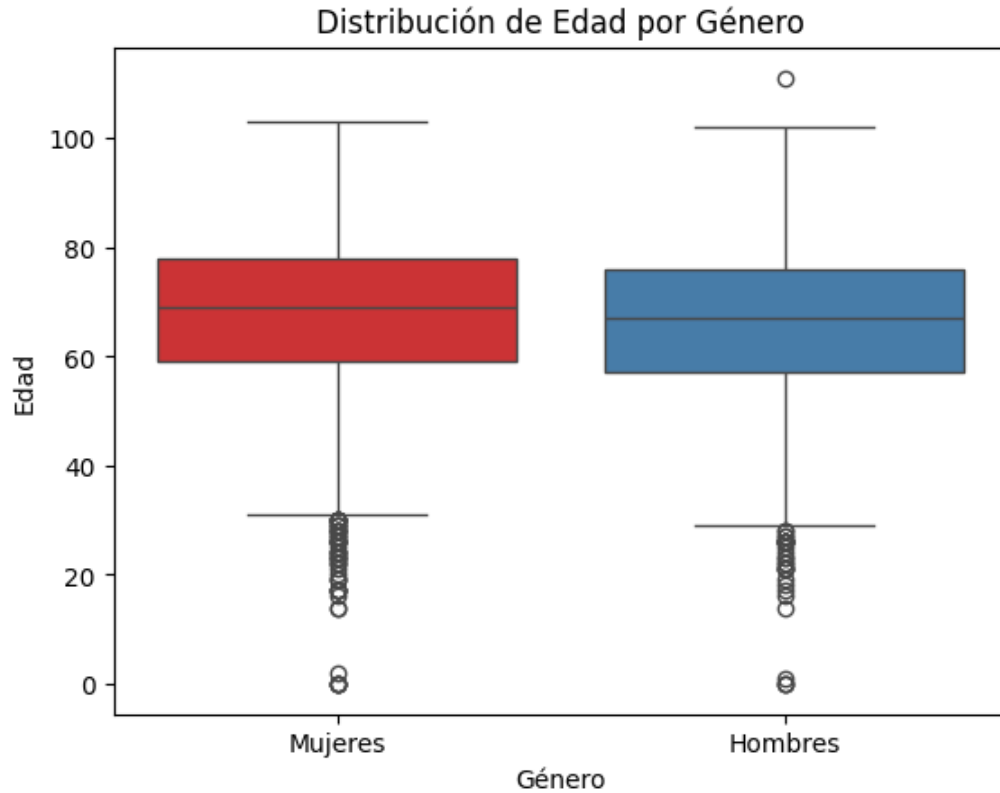


Figure 3: Este gráfico es un diagrama de caja (boxplot) que muestra la distribución de edades de las personas fallecidas por diabetes mellitus, desglosadas por género (mujeres y hombres).

Análisis de la matriz:

Valores de la diagonal:

Cada variable tiene una correlación perfecta consigo misma, por lo que la diagonal está llena de valores 1.

Correlación entre EDAD y otras variables:

- EDAD vs ESCOLARIDAD: Valor de 0.15, lo que indica una correlación positiva débil. Esto podría significar que, en promedio, las personas mayores tienden a tener niveles de escolaridad ligeramente más altos, pero no es una relación fuerte.
- EDAD vs OCUPACIÓN: Valor de 0.014, indicando una correlación casi inexistente entre la edad y la ocupación de las personas.
- EDAD vs ASIST_MEDI: Valor de 0.01, lo que sugiere que no hay una relación significativa entre la edad y la asistencia médica recibida.

ESCOLARIDAD y su relación con otras variables:

- **ESCOLARIDAD vs OCUPACIÓN:** Correlación débil (0.073). Podría sugerir que ciertos niveles de escolaridad están asociados con ciertas ocupaciones, pero no de manera muy marcada.
- **ESCOLARIDAD vs ASIST_MEDI:** Valor de 0.042, también débil, lo que indica poca relación entre el nivel de escolaridad y la asistencia médica recibida.

OCUPACIÓN y ASIST_MEDI:

- Correlación baja (0.032). No hay evidencia de que la ocupación influya significativamente en la probabilidad de recibir asistencia médica.

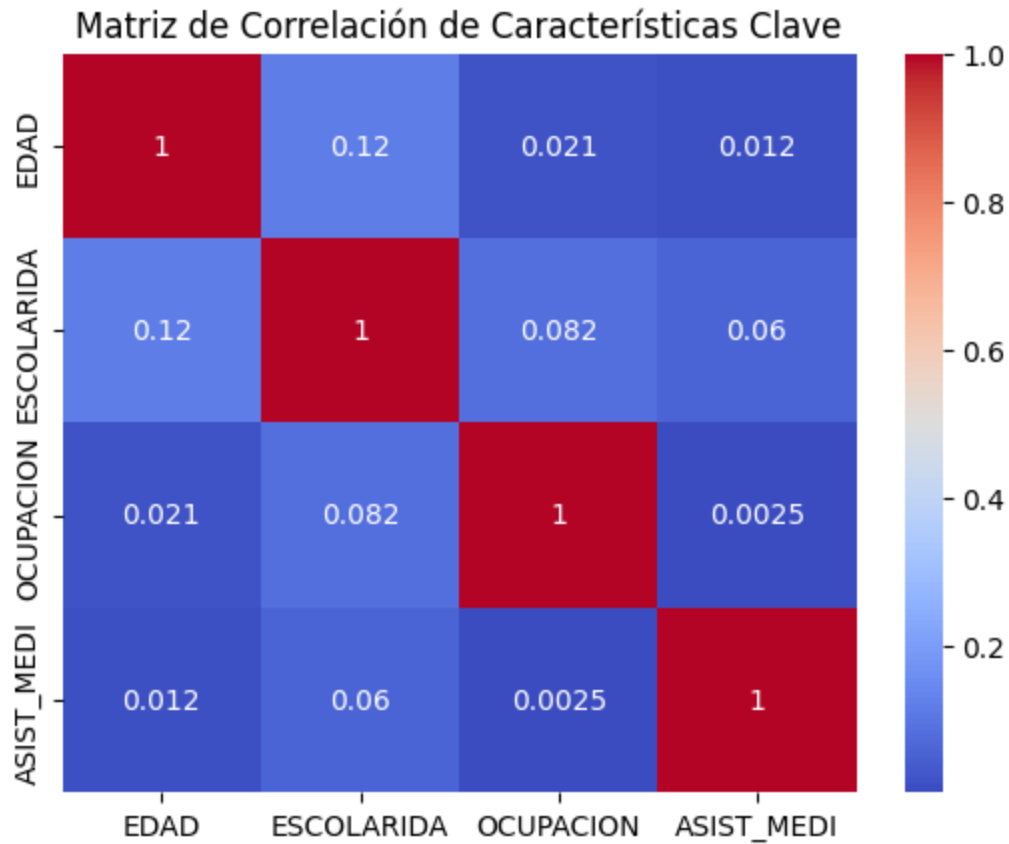


Figure 4: Matriz de Correlación de Características Clave.

En la Figura 5 se muestran las defunciones por mes debido a la "Diabetes Mellitus", enero y diciembre son los meses con un incremento de muertes por esta causa, es posible que el incremento en el índice de muertes por diabetes entre diciembre y enero esté relacionado con las festividades. Aquí hay algunas razones plausibles que podrías considerar al analizar esta tendencia:

1. Cambios en la alimentación

Consumo de alimentos poco saludables: Durante las festividades, es común consumir alimentos ricos en carbohidratos, azúcares y grasas, que pueden ser perjudiciales para las personas con diabetes.

Descontrol en la dieta: Las personas con diabetes pueden relajarse en el control de sus hábitos alimenticios debido a las celebraciones.

2. Menor actividad física

Durante las festividades, muchas personas disminuyen su nivel de actividad física debido al clima (en el hemisferio norte es invierno), las vacaciones o el enfoque en reuniones familiares.

3. Estrés asociado a las festividades

Estrés emocional y financiero: Las festividades pueden generar estrés por las obligaciones sociales, familiares o económicas, lo cual puede tener un impacto negativo en los niveles de glucosa en sangre.

4. Clima invernal

El frío puede agravar ciertas condiciones de salud, como enfermedades cardiovasculares, que están relacionadas con la diabetes.

Además, las infecciones respiratorias son más comunes en invierno y pueden ser más peligrosas para personas con diabetes.

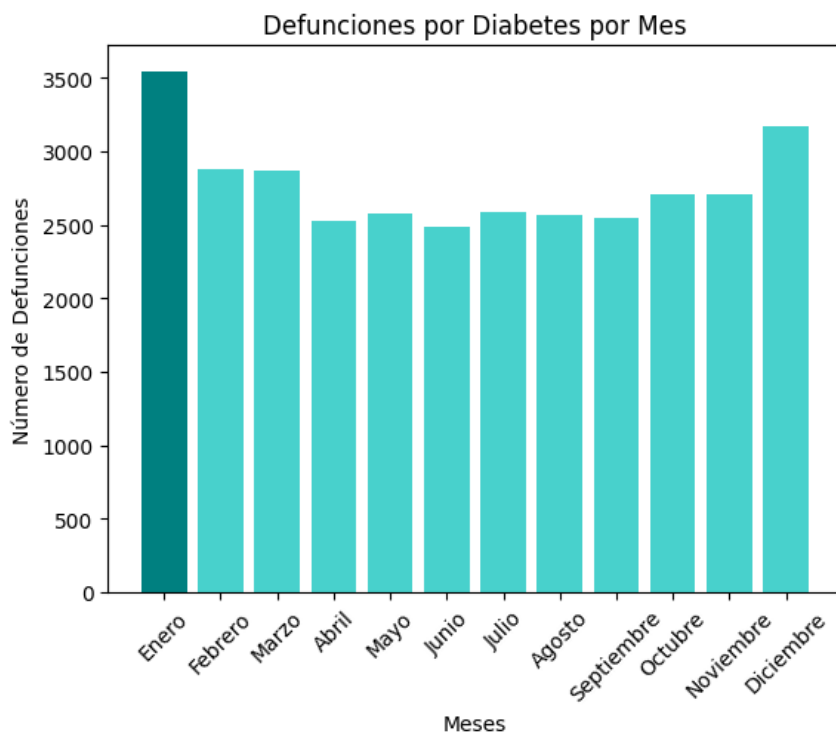


Figure 5: Grafica de barras Defunción por Diabetes por Mes.

Que la Ciudad de México sea el estado con más muertes por diabetes mellitus puede reflejar varios factores relacionados con el contexto demográfico, social, económico y de salud pública. Aquí te dejo algunas interpretaciones posibles:

1. Alta densidad poblacional

Mayor cantidad de habitantes: La Ciudad de México es una de las regiones más pobladas de México, lo que naturalmente aumenta el número absoluto de muertes por cualquier causa, incluyendo la diabetes.

Concentración urbana: Las áreas urbanas suelen tener mayor incidencia de enfermedades crónicas debido a estilos de vida sedentarios, estrés y cambios en los patrones de alimentación.

2. Acceso limitado a servicios de salud preventiva

Aunque la Ciudad de México tiene un buen acceso a servicios médicos en comparación con otras regiones, la falta de enfoque en prevención puede contribuir a un mal control de enfermedades crónicas como la diabetes.

Las personas pueden no detectar o tratar la diabetes a tiempo, lo que aumenta el riesgo de complicaciones mortales.

3. Estilos de vida urbanos

Dieta poco saludable: En las ciudades, es común el consumo elevado de alimentos procesados, ricos en azúcares y grasas, que incrementan el riesgo de diabetes y complicaciones.

Baja actividad física: El transporte público y el ritmo acelerado de la vida urbana pueden reducir las oportunidades de realizar actividad física regular.

4. Desigualdad socioeconómica

En la Ciudad de México, las desigualdades económicas son muy marcadas. Los sectores con menores ingresos pueden tener menos acceso a alimentos saludables, medicamentos o atención médica regular, lo que incrementa el riesgo de complicaciones.

Educación en salud limitada: La falta de información sobre el manejo de la diabetes puede ser más común en ciertos sectores vulnerables de la población.

5. Envejecimiento de la población

En las ciudades, es frecuente encontrar una mayor proporción de personas mayores, quienes son más propensas a desarrollar diabetes y sus complicaciones.

6. Contaminación ambiental y estrés

Contaminación: Existe evidencia que vincula la contaminación ambiental con un mayor riesgo de enfermedades metabólicas, incluyendo la diabetes.

Estrés crónico: El estrés asociado con la vida en una ciudad grande también puede afectar los niveles de glucosa en la sangre y empeorar el control de la diabetes.

7. Registro más completo

La Ciudad de México tiene una infraestructura de registro más desarrollada que otras regiones del país, lo que puede resultar en un conteo más preciso de muertes por diabetes.

8. Tendencias culturales

En algunas zonas urbanas, las tradiciones culturales relacionadas con la dieta o el cuidado de la salud pueden desempeñar un papel importante en el control de la diabetes.

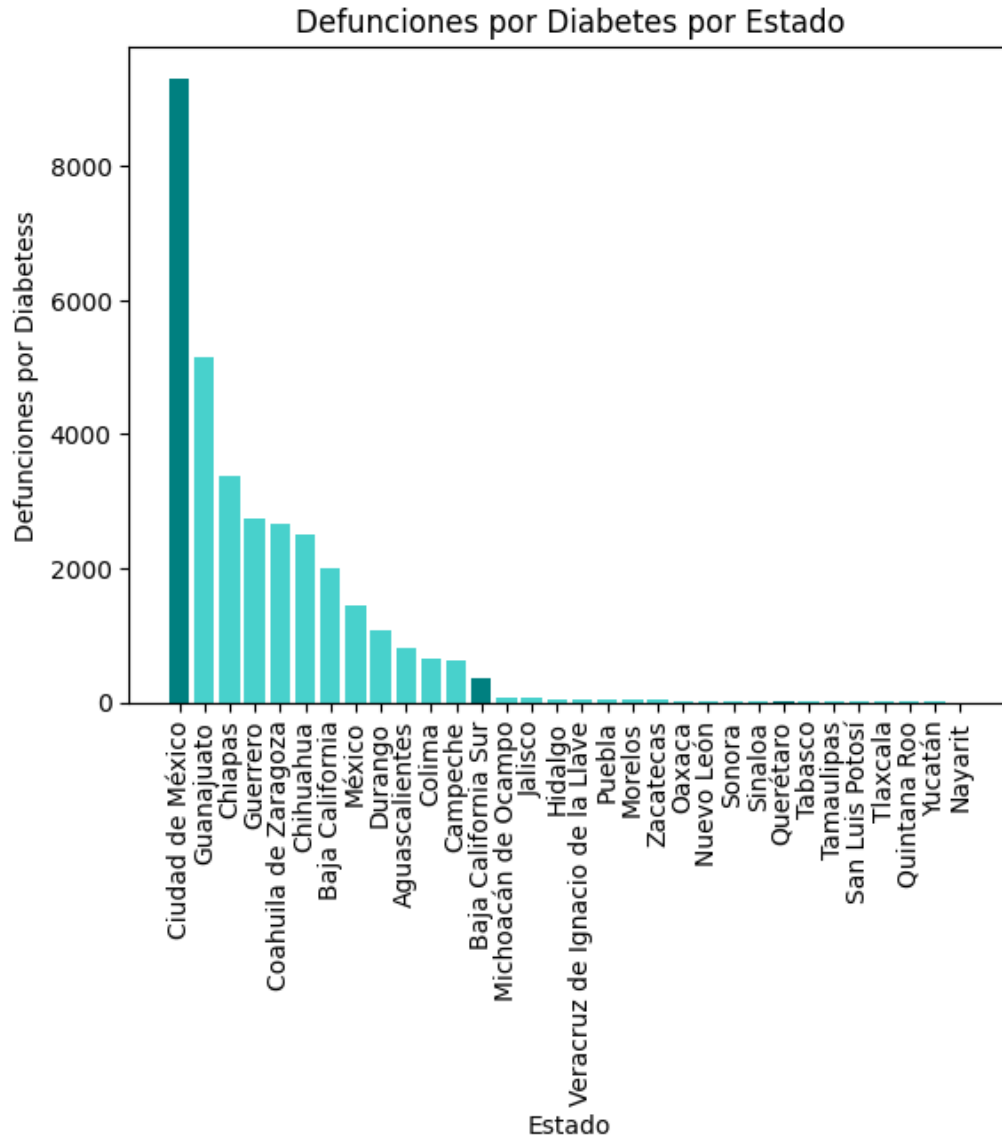


Figure 6: Defunción por Diabetes por Estado.

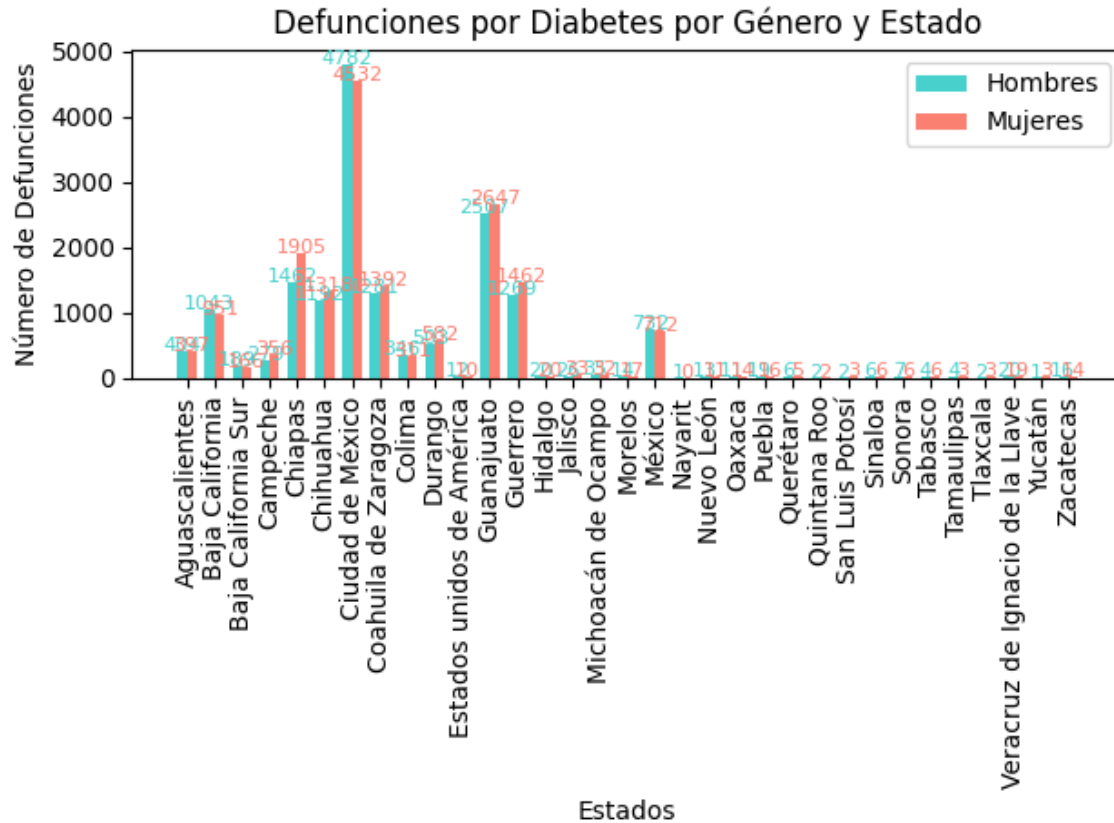


Figure 7: Defunciones por Diabetes por Género y Estado.

3.2 Modelo de clasificación

Utiliza la variable "DEFUN_DIAB" como variable objetivo e implementa un modelo de machine learning (Regresión logística) para clasificar si una persona falleció a causa de esta enfermedad o no. Tienes libertad para utilizar las variables que consideres prudente. Esta parte se califica sobre 3 aspectos

Preprocesamiento de la información (Manipulación de valores nulos o atípicos, separación de conjuntos, imputación, escalamiento, etc.) Implementación del modelo.

Explicación del valor de 1 en la gráfica ROC: Modelo Perfecto: Un valor de 1 en la curva ROC indica que el modelo tiene un área bajo la curva (AUC) de 1. Esto significa que el modelo ha logrado una clasificación perfecta en todos los puntos de corte de probabilidad. Este es un resultado ideal y poco común en la práctica.

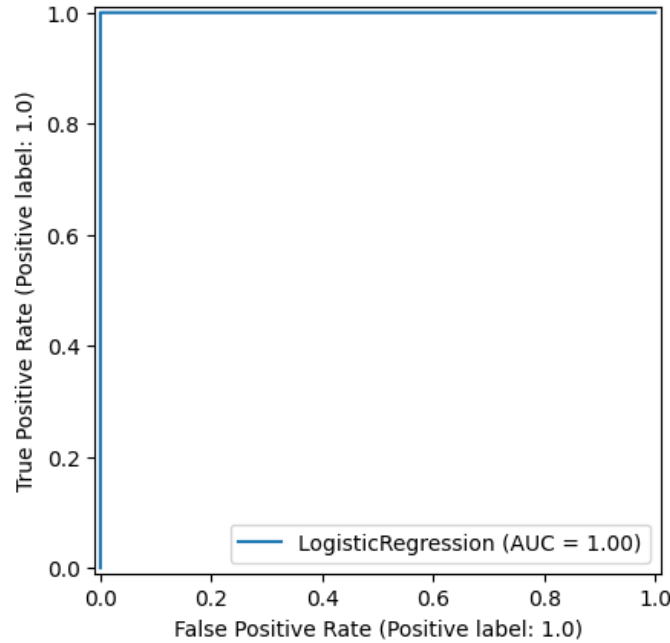


Figure 8: Evaluación del modelo (ROC o métricas sobre la matriz de correlación).

3.3 Forecasting

Utiliza el conjunto de datos "DEFUN_DIAB.TS.csv", que contiene la cantidad de personas fallecidas por causa de la diabetes mellitus por día (desde 01-01-2015 hasta 31-12-2020), para crear un modelo de machine learning tal que puedas predecir el total de muertes por esta causa para los primeros 3 meses de 2021.

La evaluación se clasifica sobre 3 aspectos:

Preprocesamiento de la información (Manipulación de valores nulos o atípicos, separación de conjuntos, imputación, escalamiento, etc.) Implementación del modelo (Regresión lineal, ARMA, Suavización exponencial u otro) Evaluación del modelo (Evaluar el desempeño en el conjunto de prueba usando métricas MSE o R2)

3.3.1 Resultados del Forecasting

Para explicar el incremento de muertes por diabetes mellitus de abril a septiembre de 2020 de acuerdo a las Figuras ?? que muestran las defunciones por diabetes por día, mes y año del 2015 al 2020 se pueden relacionar con la pandemia, puedes considerar los siguientes puntos:

1. Impacto de la COVID-19 en pacientes con diabetes:

La COVID-19 afecta de manera más severa a personas con enfermedades crónicas como la diabetes mellitus debido a su sistema inmunológico debilitado.

Muchas personas con diabetes que contrajeron COVID-19 experimentaron complicaciones graves, como insuficiencia respiratoria y descontrol glucémico, lo que pudo aumentar la mortalidad.

2. Desafíos en el acceso a atención médica:

Durante la pandemia, los sistemas de salud estuvieron sobrecargados, limitando el acceso a consultas regulares y tratamientos para personas con diabetes.

Muchas personas evitaron acudir a hospitales por temor a contagiarse, lo que pudo llevar a un manejo inadecuado de su enfermedad.

3. Restricciones y cambios en el estilo de vida:

Las restricciones de movilidad, el confinamiento y el cierre de gimnasios afectaron la actividad física, un factor clave para el control de la diabetes.

Cambios en los hábitos alimenticios debido al estrés, la ansiedad y el acceso limitado a alimentos saludables también pudieron contribuir a un mal control glucémico.

4. Condiciones socioeconómicas:

La crisis económica generada por la pandemia afectó la capacidad de algunas personas para adquirir medicamentos, insulina y alimentos necesarios para el control de la diabetes.

Las poblaciones más vulnerables pudieron ser las más afectadas, especialmente en zonas urbanas densas como la Ciudad de México.

5. Falta de monitoreo y seguimiento:

Durante la pandemia, muchas clínicas suspendieron consultas regulares y monitoreos de enfermedades crónicas. Esto pudo derivar en un mayor número de complicaciones no detectadas a tiempo.

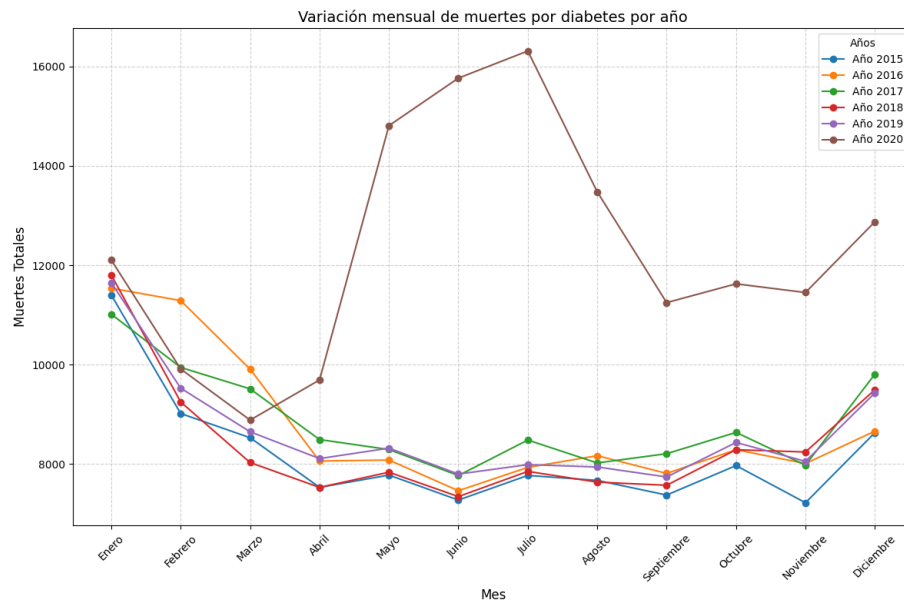


Figure 9: Variación mensual de muertes por diabetes por año.

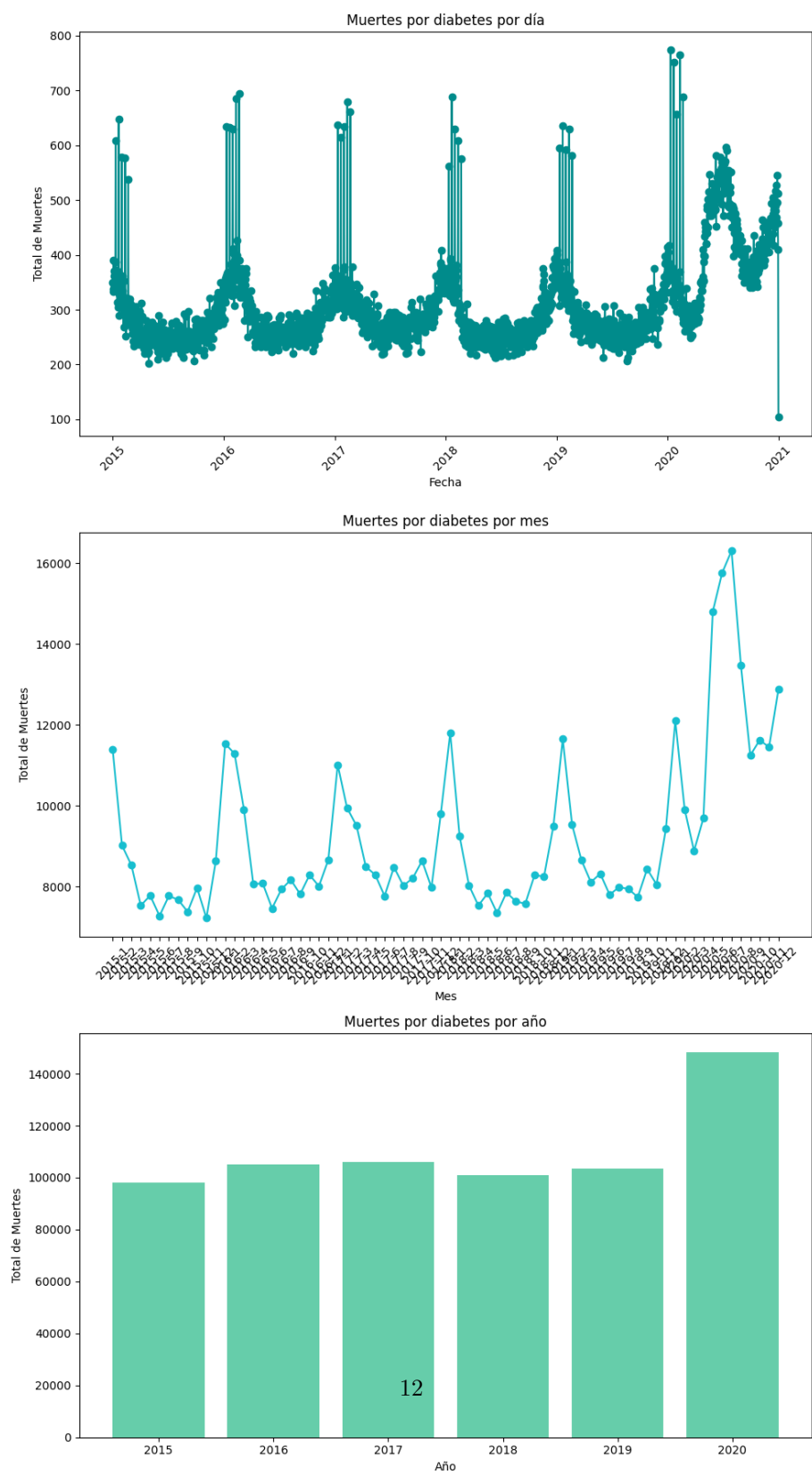


Figure 10: Muertes por diabetes del 2015 al 2020 por día, mes y año.



Figure 11: Hice lo mejor que pude.

4 Open code

https://colab.research.google.com/drive/13bd6E6LzaQ9FGdTvqRHnIR5PwA_e91KJ?usp=sharing