



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FÍSICA BIOMÉDICA

SEMESTRE: 2025-2

---

Predicción de diabetes para 2024 y 2025

---

**Autor:**

Cuevas García Ana Cristina.

**Profesores:**

David Alexis García Espinosa.

Luis Eduardo Flores Luna.

**Ayudante:**

Derek Saúl Morán Pérez.

**Introducción a la ciencia de Datos Aplicada a escenarios médico-biológicos**

13 de Junio del 2025

## Resumen

El presente documento presenta un análisis detallado de los datos relacionados con la diabetes en México. Se aplicaron técnicas de análisis exploratorio de datos para detectar patrones e información relevante, además de proponer un modelo que nos ayude a predecir la alza o baja de la enfermedad en el país para el año 2024 y 2025.

## 1 Introducción

La diabetes mellitus es una de las enfermedades más prevalentes y preocupantes a nivel mundial. Su carácter crónico y degenerativo representa un serio problema de salud pública, especialmente en países como México. Esta enfermedad afecta a múltiples sistemas del cuerpo y su crecimiento ha sido alarmante en las últimas décadas.

La diabetes es una enfermedad crónica caracterizada por mantener elevados los niveles de azúcar (glucosa) en sangre, esto, asociado a la deficiencia de insulina, afecta al corazón, ojos, riñones y sistema nervioso.

- **Causas:** La causa principal de la diabetes son los malos hábitos alimenticios llenos de azúcares refinadas, grasas saturadas y sodio, le sumamos a esto el sedentarismo, la obesidad y el factor genético.

### 1.1 Tipos de diabetes

**Existen tres tipos de diabetes:**

- **Diabetes Tipo 1.** Suele aparecer con mayor frecuencia en la juventud, afectando directamente al páncreas al producir poca o nada de insulina.

- **Diabetes Tipo 2.** Es el tipo de diabetes más común, sucede cuando el cuerpo es incapaz de producir insulina y se acumula la glucosa en la sangre; representa la mayoría de los casos y se manifiesta generalmente en adultos, muchas veces con obesidad o hipertensión.

- **Diabetes gestacional.** Esta se presenta durante el embarazo a causa de los cambios que sufre el cuerpo propios en ese estado y suele darse en una etapa avanzada de la gestación, aunque normalmente desaparece tras dar a luz.

- **Síntomas:** Los síntomas son diferentes dependiendo el tipo de la diabetes, pero cuando los niveles de azúcar son altos, se presenta:

una sensación de mucha hambre y sed incluso llegar a perder peso necesidad de orinar muy a menudo sentir cansancio. Por otro lado, en las personas con diabetes tipo 2 es común no presentar síntomas al inicio, incluso es posible que no los tengan durante muchos años. En estos casos, la detección de la diabetes suele darse mediante un análisis de sangre, pero se puede reconocer la enfermedad ante síntomas como disfunción eréctil, visión borrosa y dolor o entumecimiento en los pies o las manos

- **Prevención:** Las personas con alto riesgo de desarrollar diabetes tipo 2, pueden disminuirlo significativamente, realizando cambios en su estilo de vida y mantenerlos a largo plazo, como:

Adopción de una alimentación saludable. Incremento de la actividad física moderada. Mantener un peso saludable Abandono del tabaquismo.

Con el auge de las técnicas de ciencia de datos, es posible no solo analizar tendencias pasadas, sino también proyectar la carga futura de la enfermedad. Este estudio busca predecir la prevalencia de diabetes en México para los años 2024 y 2025 mediante modelos estadísticos de series de tiempo.

## 2 Antecedentes

La diabetes tipo 2 representa la mayoría de los casos en adultos. Según la OMS, más del 10 por ciento de los adultos en México padecen esta enfermedad. Los

factores de riesgo más relevantes incluyen obesidad, sedentarismo y edad. Modelos utilizados en este proyecto como ARIMA y Prophet han sido útiles para predecir la evolución de enfermedades crónicas y apoyar la planeación en salud pública.

### 3 Objetivo

#### General:

- Predecir el número de casos de diabetes en México para los años 2024 y 2025.

#### Específicos:

- Realizar un análisis exploratorio de los datos.
- Evaluar los modelos ARIMA y Prophet. Interpretar los resultados desde una perspectiva médico-social

Actualmente en nuestro país, la diabetes es la tercera causa de muerte; es posible que más de cien mil personas mueran al año por esta enfermedad, dentro de estas cifras, las personas mayores de 65 años ocupan el primer lugar en muertes. Los niños no están exentos, e inclusive pueden nacer con esta enfermedad a causa de la famosa diabetes gestacional que suele darle a la madre durante el embarazo.

el código base que podemos encontrar en [éste link de de la página del IMSS](#).

### 4 Metodología

- Datos: Casos anuales de diabetes en México (2000–2023).
- Herramientas: Python, Pandas, Matplotlib, Statsmodels, Prophet.

1. Limpieza y preprocesamiento.

2. Visualización y análisis exploratorio.
3. Visualización y análisis exploratorio.
4. Modelado con ARIMA y Prophet.
5. Evaluación con MRSE y  $R^2$ .
6. Elección de modelo final.

### 5 Desarrollo

#### 5.1 Descripción de los Datos

La base de datos consta de registros anuales de casos detectados de diabetes en México, desglosados por estados que parte desde el año 2000 hasta el 2023, la primera columna contiene el nombre de la entidad federativa y el total, el resto, el número de casos detectados por año. Iniciamos con la base de datos "Istapla43 – 2023", posteriormente se realizó edición de la base de datos con la que iniciamos trabajando llamada "diabetes2000-2023".

#### 5.2 Análisis Exploratorio de Datos

Antes de aplicar modelos de predicción, se realizó un análisis exploratorio detallado de los datos proporcionados, los cuales contienen el número de casos de diabetes por estado en México desde el año 2000 hasta 2023.

Se analizaron 33 registros correspondientes a las 32 entidades federativas y el total nacional. Cada registro incluye el número absoluto de casos por año, así como la población total del estado en 2020. Esta estructura permitió visualizar el crecimiento tanto a nivel nacional como estatal.

Algunas observaciones generales:

- En el año 2000, el total nacional fue de aproximadamente 5.18 millones de casos.

- Para 2023, esta cifra se elevó a más de 15.8 millones, lo que representa un incremento superior al 200 por ciento en poco más de dos décadas.
- El crecimiento ha sido particularmente fuerte entre 2010 y 2020.
- El estado con mayor número de casos en 2023 fue el Estado de México, seguido de la Ciudad de México y Veracruz.

Estas observaciones se complementan con gráficos que detallan la evolución temporal, la descomposición estacional y la predicción futura, presentados a continuación.

En seguida, se presentan algunas visualizaciones clave del comportamiento histórico de los casos de diabetes en México, con sus respectivos análisis:

La incidencia de diabetes muestra una tendencia creciente. Gráficos lineales revelan un crecimiento sostenido, especialmente entre 2010 y 2020. Se evaluó la estacionalidad, y aunque leve, algunos modelos identificaron ciclos en los datos. Antes de visualizar la gráfica sobre los casos de diabetes en el país, veremos los primeros 10 países con más casos en total:

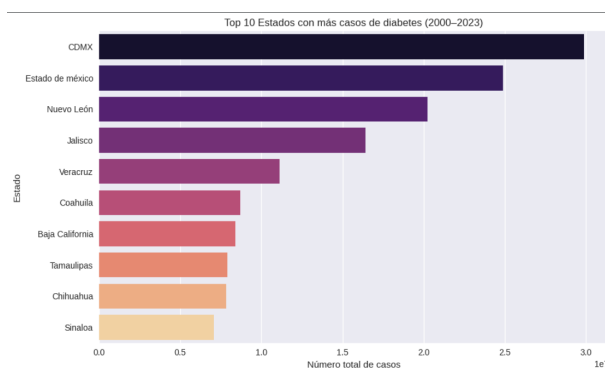


Figure 1: Top 10 estados con más casos.

Observamos que en este gráfico, la CDMX prevalece, enseguida el Estado de México, seguido de Nuevo León, Jalisco, Veracruz,

Coahuila, Baja California, Tamaulipas, Chihuahua y finalmente Sinaloa.

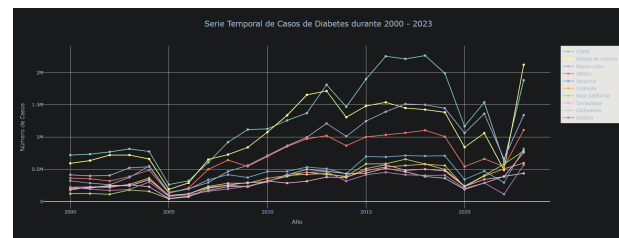


Figure 2: Tendencia histórica de casos de diabetes en México (2000–2023).

Se realizó el siguiente análisis:

- **Tendencia ascendente general (2005–2018):**

La mayoría de los estados muestran un aumento constante en los casos de diabetes desde 2005 hasta aproximadamente 2018.

- **Caída en 2020:**

Hay una disminución abrupta en los casos en casi todos los estados durante 2020. Esto probablemente se deba a la pandemia de COVID-19, que afectó el registro y diagnóstico de enfermedades crónicas, ya que los servicios de salud estaban centrados en la emergencia sanitaria.

- **Aumento repentino en 2022-2023:**

A partir de 2022, hay un repunte drástico en varios estados, notablemente en Estado de México y CDMX, lo que podría reflejar un efecto de "rebote" en la atención médica pospandemia o una mejor recuperación del sistema de vigilancia epidemiológica.

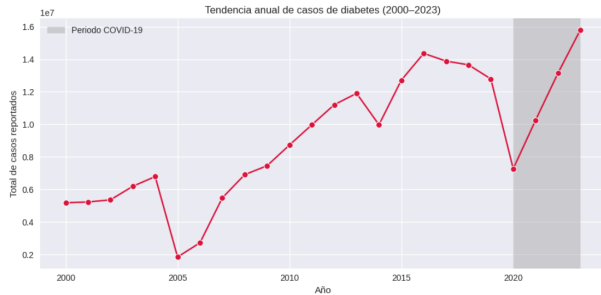


Figure 3: Tendencia anual total de casos de diabetes reportados en México entre 2000 y 2023 (con un enfoque claro en el impacto del periodo de la pandemia por COVID-19 (sombreado en gris, aproximadamente 2020–2023).)

### Lo interesante de esta gráfica es el Impacto COVID-19 (2020–2022):

En 2020 hay una caída muy fuerte (a menos de 8 millones de casos).

Este descenso no refleja una mejora en la salud, sino menor detección y atención médica, como resultado de la saturación del sistema de salud y el confinamiento.

En 2021 y 2022 se observa una recuperación gradual.

### Máximo histórico en 2023:

Se alcanza el pico más alto de todo el periodo (casi 16 millones de casos).

Esto puede ser consecuencia de:

- Reanudación de diagnósticos post-pandemia.
- Mayor acceso a servicios médicos.
- Posible aumento real de la prevalencia.

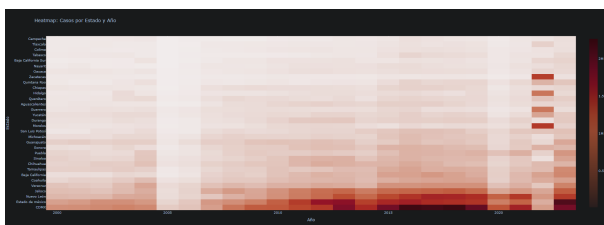


Figure 4: heatmap de casos reportados de diabetes por estado en México entre 2000 y 2023.

En la presente gráfica podemos observar:

**Distribución geográfica:** Estados con mayor concentración de casos (zona inferior del gráfico):

- CDMX, Estado de México, Nuevo León, Jalisco, Veracruz, Coahuila y Baja California muestran una mayor carga de casos (colores más oscuros).

Estos estados concentran mayor población urbana y envejecida, lo que podría explicar la alta prevalencia.

**Estados con menor carga (zona superior del gráfico):**

- Campeche, Tlaxcala, Colima, Baja California Sur, Nayarit y Zacatecas presentan colores más claros de forma constante.

En general, son entidades con menor población y menor densidad urbana.

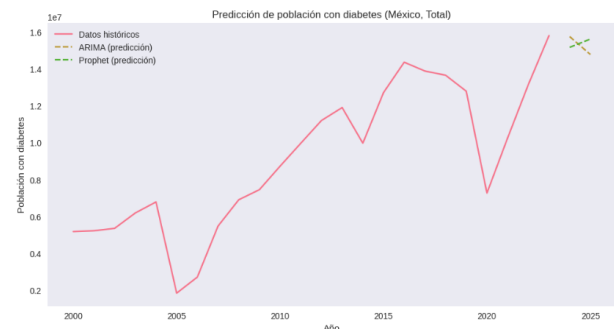


Figure 5: población con diabetes en México a lo largo del tiempo, con una predicción hasta el año 2025

La gráfica presentada nos muestra los datos históricos (Línea continua en color rojo= que ya conocemos, pero tiene el resultado de predicción de dos modelos diferentes, ARIMA (En color naranja) y Prophet (en color café) y podemos analizar lo siguiente:

- **Prophet:**  
Muestra tendencia creciente rumbo a 2025, el modelo es capaz de manejar mejor irregularidades, estacionalidades y tendencias no lineales además de que su predicción parece más optimista en términos de crecimiento, probablemente influenciada por la recuperación post-pandemia y la tendencia general al alza observada en los años previos.
- **ARIMA:**  
Muestra disminución de la población con diabetes para 2025, además de ser un modelo clásico de series temporales que se basa únicamente en los valores históricos, sin factores externos. Su predicción parece conservadora y posiblemente limitada por los valores recientes (como la caída en 2020 dada por pandemia).

Al haber decidido el uso del modelo Prophet dado a las métricas MRSE y  $R^2$  veremos una proyección de casos de diabetes en el país incluyendo los años predichos (2024 y 2025)

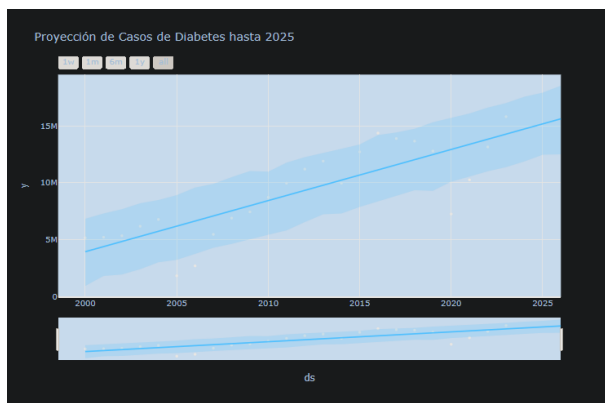


Figure 6: la proyección de casos de diabetes en México hasta el año 2025, utilizando el modelo Prophet.

En esta gráfica, donde **ds** Representa los años, desde el 2000 hasta 2025 e **y** Representa la cantidad de personas con diabetes, la línea azul es la predicción central,

y el área al rededor, de azul claro, representa el intervalo de confianza además de contar con unos puntitos que son nuestros datos históricos. Hablando del intervalo de confianza o incertidumbre del modelo se observa que a medida que avanza el tiempo, esta área se ensancha, lo que indica una mayor incertidumbre en las predicciones más lejanas (como 2024-2025).

Apesar de la existencia de dicha incertidumbre, el intervalo es bastante estrecho, lo que puede indicar que los datos históricos son consistentes y que el modelo confía en su predicción.

También es notorio que los puntos (datos históricos)s se alinean bastante bien con la línea azul (Predicción central), lo que indica un buen ajuste del modelo Prophet a los datos.

## 6 Discusión

El modelo Prophet ha demostrado ser eficaz para capturar la tendencia creciente en los casos de diabetes en México. Algunos puntos clave a discutir:

### Precisión del modelo:

- Prophet logró un buen ajuste a los datos históricos, lo que sugiere que la estructura del modelo es adecuada.
- Sin embargo, al tratarse de un modelo que enfatiza tendencias y estacionalidades, podría mencionar que no capta bien las caídas abruptas (por ejemplo, por subregistro en años de pandemia o cambios en política pública).

### Comparación con ARIMA:

- En comparación con ARIMA (observado en otra gráfica), Prophet proyecta una mayor alza en casos de diabetes, lo que puede ser más consistente con la tendencia histórica.

- Prophet tiende a suavizar las irregularidades y proyectar más en función del crecimiento general.

### Limitaciones:

- No se incorporaron variables externas (como obesidad, consumo de azúcar, políticas de salud, o impacto de la pandemia), lo cual puede limitar la capacidad predictiva real del modelo, me parece que serían buenas variables a considerar.

## 7 Conclusiones

Hemos podido notar un cambio significativo desde el año 2000:

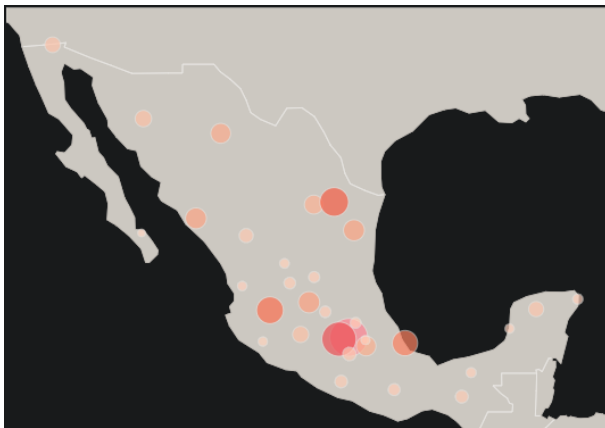


Figure 7: Visualización demográfica de casos de diabetes en el año 2000.

El modelo Prophet predice que la población con diabetes en México seguirá creciendo hasta superar los 15 millones en 2025, lo cual representa un problema grave de salud pública.

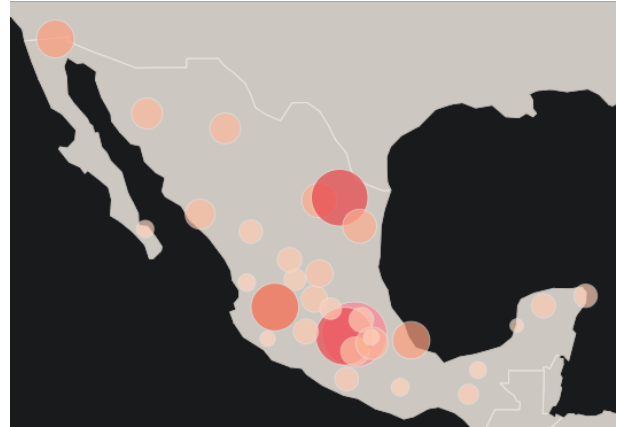


Figure 8: Visualización demográfica de casos predichos de diabetes en el año 2025.

Sería mucho mejor complementar este modelo con análisis que incluyan factores sociales, económicos y de salud pública, para enriquecer la predicción y orientar políticas más efectivas.

Las autoridades sanitarias deben considerar estos escenarios para anticipar necesidades de atención médica, recursos e infraestructura.

**¿Qué recomendaciones podrían tomarse en cuenta para mejorar la situación?**

1. Evitar el desarrollo de la enfermedad:
  - Realizar campañas de educación alimentaria, a temprana edad.
  - Fomentar la actividad física REGULAR en la sociedad, desde escuelas primarias hasta en trabajos además de recomendar mejorar la alimentación.
1. Detección temprana:
  - Ampliar el acceso a la salud, con chequeos médicos preventivos gratuitos.
  - Capacitación médica para diagnóstico temprano.
  - Implementación de clínicas móviles a lo largo del país, principalmente en los estados más afectados, y zonas marginadas, puesto que tenemos información de una base de datos del IMSS con

pacientes registrados, pero seguramente muchas más personas lo padecen sin tener registro o seguimiento médico.

## 8 Referencias

### References

- [1] Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado (ISSSTE). (s.f.). *Diabetes: uno de los principales problemas de salud en México*. Gobierno de México. Recuperado de <https://www.gob.mx/issste/es/articulos/diabetes-uno-de-los-principales-problemas-de-salud-en-mexico?idiom=es>
- [2] Organización Panamericana de la Salud (OPS). (s.f.). *Diabetes*. Recuperado de <https://www.paho.org/es/temas/diabetes>
- [3] Instituto Nacional de Salud Pública (INSP). (s.f.). *Diabetes en México*. Recuperado de <https://www.insp.mx/avisos/3652-diabetes-en-mexico.html>
- [4] Facebook. (s.f.). *Prophet: forecasting at scale*. Recuperado de <https://facebook.github.io/prophet/>
- [5] Instituto Nacional de Estadística y Geografía (INEGI). (2021). *Características de las defunciones registradas en México durante 2019*. Recuperado de [https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2021/EstSociodemo/DefuncionesRegistradas2020\\_Pnles.pdf](https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2021/EstSociodemo/DefuncionesRegistradas2020_Pnles.pdf)