

Proyecto ABP

1. **NOMBRE DEL PROYECTO:** Análisis predictivo de factores críticos en ataques cardíacos utilizando datos clínicos recolectados entre enero y mayo de 2019 en el Hospital Zheen (Erbil, Iraq)
2. **TIPO DE PROYECTO:** Tecnológico y de investigación. El proyecto integra el uso de herramientas tecnológicas para el procesamiento y análisis de datos reales en el ámbito de la salud, con un enfoque exploratorio para interpretar factores de riesgo en ataques cardíacos.
3. **ESPACIO CURRICULAR O ESPACIOS PARTICIPANTES EN EL MÓDULO:**
Procesamiento de datos - Estadística y Exploración de Datos 1
4. **EJES TEMÁTICOS/RED DE CONCEPTOS:** Los ejes temáticos que se trabajan son:
 - Análisis exploratorio de datos (EDA)
 - Limpieza y validación de datos
 - Estadística descriptiva (media, mediana, desviación estándar, valores atípicos)
 - Identificación de valores nulos y registros erróneos
 - Análisis de variables categóricas y numéricas
 - Interpretación de resultados para la toma de decisiones
 - Aplicación de herramientas computacionales (Python, Pandas, Jupyter)
5. **PROBLEMÁTICAS/NECESIDADES:** En la actualidad, los ataques cardíacos representan una de las principales causas de muerte a nivel mundial. El acceso a datos médicos permite realizar análisis que podrían ayudar a comprender mejor los factores críticos asociados. Sin embargo, muchas veces estos datos no son aprovechados por falta de habilidades técnicas. Esta situación plantea la necesidad de formar perfiles capaces de realizar análisis de datos en contextos reales y sensibles, como el de la salud.

6. FUNDAMENTACIÓN: Este proyecto fue elegido por su potencial para fortalecer competencias técnicas en ciencia de datos aplicadas al ámbito de la salud, utilizando un dataset real recopilado en el Zheen Hospital de Erbil, Iraq, durante los meses de enero a mayo de 2019. La base de datos contiene registros biomédicos de pacientes con sospecha de ataque cardíaco, lo que permite trabajar con información representativa de un problema de salud pública global.

El análisis estadístico y exploratorio busca identificar patrones, relaciones entre variables y validar hipótesis que podrían orientar futuras decisiones clínicas o investigaciones. Esta tarea permite a los estudiantes aplicar conocimientos técnicos (como la limpieza, validación y análisis de datos), éticos (como el manejo responsable de datos sensibles) y científicos (como la interpretación crítica de indicadores biomédicos).

Además, este tipo de análisis puede ser de gran utilidad para el personal sanitario, en especial para médicos cardiólogos, ya que puede servir como una herramienta complementaria para el seguimiento de pacientes y la detección temprana de signos de riesgo, contribuyendo a intervenciones preventivas más efectivas.

El proyecto no solo apunta al desarrollo académico y técnico, sino también promueve la toma de conciencia sobre problemáticas de salud pública y la integración de la tecnología en favor del bienestar social.

7. VISIÓN DEL PROYECTO: Construir una base sólida de análisis de datos médicos mediante la utilización de herramientas estadísticas y tecnológicas, con el fin de interpretar correctamente variables biomédicas y generar informes útiles que promuevan una toma de decisiones informada en el ámbito de la salud.

DISEÑO DE LOS OBJETIVOS

- **Objetivo general:** Aplicar técnicas de análisis de datos utilizando Python, Pandas y herramientas estadísticas para identificar patrones relevantes en un dataset médico sobre ataques cardíacos recolectado entre enero y mayo de 2019 en el hospital Zheen de Erbil, Iraq, desarrollando el trabajo durante el primer semestre de 2025.
- **Objetivos específicos:**
 1. Importar, limpiar y preparar un dataset real para análisis estadístico.
 2. Analizar las variables numéricas y categóricas en términos de distribución, media, valores faltantes y posibles errores.
 3. Interpretar los resultados obtenidos y redactar un informe con fundamentos técnicos y científicos.
 4. Crear un entorno reproducible que permita compartir el análisis con la comunidad académica o profesional.
- **Metas:** El proyecto tendrá como meta la presentación de un informe final en formato PDF que refleje un análisis técnico y estadístico riguroso del dataset médico proveniente del Zheen Hospital (Iraq, 2019).

8. SELECCIÓN DE ACCIONES

| OBJETIVO ESPECÍFICO | ACCIONES |
|---|---|
| Importar, limpiar y preparar un dataset real para análisis estadístico. | Crear entorno de trabajo en VSCode y/o Colab, importar el CSV, revisar valores nulos y corregir errores |

| | |
|--|---|
| Analizar las variables numéricas y categóricas en términos de distribución, media, valores faltantes y posibles errores. | Usar Pandas y Numpy para describir estadísticamente el dataset y verificar datos erróneos |
| Interpretar los resultados y redactar informe | Generar gráficas, detectar correlaciones posibles, extraer conclusiones claras y relevantes |
| Crear entorno reproducible | Subir el archivo .ipynb al repositorio y generar una presentación o PDF con los hallazgos |

9. CRONOGRAMA: incluir un presupuesto de tiempo estimativo. Pueden ser meses o semanas. Ejemplo:

| CRONOGRAMA | MES O SEMANA 1 | MES O SEMANA 2 | MES O SEMANA 3 | MES O SEMANA 4 |
|--|--|---|---|----------------------------------|
| Objetivo 1: Importar, limpiar y preparar un dataset real para análisis estadístico. | Crear entorno de trabajo en VSCode y/o Colab Subir el dataset CSV original al repositorio | Cargar los datos con Pandas Identificar y tratar valores nulos y registros mal formateados | Verificar los tipos de datos Normalizar o ajustar formatos si es necesario | Documentar la limpieza realizada |

| | | | | |
|--|---|---|--|---|
| <p>Objetivo 2:</p> <p>Analizar las variables numéricas y categóricas en términos de distribución, media, valores faltantes y posibles errores.</p> | <p>Definir las variables numéricas y categóricas del dataset</p> <p>Verificar categorías esperadas (0/1, positive/negative, etc.)</p> | <p>Aplicar análisis estadístico con Pandas y Numpy</p> <p>Obtener media, mediana, desvío estándar y conteos</p> | <p>Detectar posibles outliers o inconsistencias</p> <p>Registrar cantidad de nulos o valores inesperados</p> | <p>Consolidar tabla de resultados con insights por valores.</p> |
| <p>Objetivo 3:</p> <p>Interpretar los resultados y redactar informe.</p> | <p>Comenzar redacción del informe técnico</p> <p>Plantear hipótesis preliminares según patrones visibles</p> | <p>Crear visualizaciones (gráficos de barras, histogramas, etc.)</p> <p>Extraer conclusiones parciales</p> | <p>Redactar versión completa del informe</p> <p>Integrar visualizaciones y referencias</p> | <p>Revisar redacción, formato y ortografía</p> <p>Exportar informe en PDF</p> |
| <p>Objetivo 4: Crear entorno reproducible.</p> | <p>Configurar notebook limpio y comentado</p> | <p>Subir versión final del archivo .ipynb al repositorio</p> | <p>Generar link compartible de Colab o GitHub</p> | <p>Validar acceso, probar desde otro dispositivo</p> |

| | | | | |
|--|--|--|--|--|
| | Probar funcionamiento de todas las celdas | | | |
|--|--|--|--|--|

10. PRODUCTO FINAL: El producto final será un informe de análisis de datos médicos sobre ataques cardíacos, incluyendo estadísticas, gráficas y conclusiones, acompañado de un notebook de Python en Jupyter o Google Colab. Se espera que este producto sea un ejemplo práctico de análisis reproducible, útil tanto en la formación académica como para su posible extensión a proyectos de salud comunitaria.

11. BIBLIOGRAFÍA:

Rashid, T. A., & Hassan, B. (2022). *Heart Attack Dataset*. Mendeley Data, V1.

<https://doi.org/10.17632/wmhctcrt5v.1>

McKinney, W. (2022). *Python for Data Analysis* (3rd ed.). O'Reilly Media.

World Health Organization. (2023). *Cardiovascular Diseases (CVDs)*.

[https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvd-s\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvd-s))

NumPy Developers. (2023). *NumPy Documentation*. <https://numpy.org/doc/>

Python Software Foundation. (2024). *Pandas Documentation*.

<https://pandas.pydata.org/>