

PROYECTO ABP –

Módulo Analista de Datos 2025.

Grupo Cluster-Fusion .

Repositorio del grupo: [Cluster-Fusion](#)

Integrantes:

- Barbero, Maciel.
- Molina, Jonathan Ariel.
- Molina, Mauricio Leonel.
- Robles, Emilce Lucia Nicole.
- Sosa, Sebastian Cristhian.
- Virinni, Marco.
- Marini, Ian Denis.

1. NOMBRE DEL PROYECTO.

“Análisis Exploratorio de Datos Médicos sobre Enfermedades Cardíacas – Hospital Zheen, 2019”.

2. TIPO DE PROYECTO.

Proyecto de Investigación Tecnológica.

Este proyecto se enfoca en aplicar técnicas estadísticas y computacionales para el análisis de datos médicos reales, con fines exploratorios, utilizando métodos actuales de ciencia de datos.

3. ESPACIO CURRICULAR O ESPACIOS PARTICIPANTES EN EL MÓDULO.

- **Procesamiento de Datos.**
- **Estadística y Exploración de Datos I.**

4. EJES TEMÁTICOS / RED DE CONCEPTOS.

- Limpieza, estructuración y análisis de datos con Python y Pandas.
- Análisis estadístico descriptivo: medidas de tendencia central, dispersión y forma.
- Tratamiento de valores ausentes y errores de carga.

- Clasificación de variables: categóricas y numéricas.
- Representación gráfica de datos médicos.
- Evaluación crítica de la calidad de los datos.

Competencias/Habilidades a fortalecer:

- Pensamiento analítico y crítico.
- Comunicación de hallazgos técnicos.
- Aplicación práctica de herramientas estadísticas.
- Toma de decisiones basada en datos.
- Trabajo colaborativo en entornos digitales.

5. PROBLEMÁTICAS / NECESIDADES.

En la actualidad, las enfermedades cardiovasculares representan una de las principales causas de muerte a nivel mundial. A través de este proyecto, se busca analizar datos reales de pacientes del Hospital Zheen para comprender mejor los factores asociados a ataques cardíacos y evaluar patrones que podrían facilitar diagnósticos preventivos.

6. FUNDAMENTACIÓN.

Elegimos esta temática porque nos permite trabajar con datos reales de salud pública, vinculando nuestras competencias técnicas con un problema de alto impacto global. El análisis estadístico y computacional de este dataset ofrece la posibilidad de identificar patrones de riesgo y mejorar la comprensión de variables asociadas a enfermedades cardíacas. Además, nos permite aplicar herramientas de ciencia de datos y estadística en un contexto médico, alineado con nuestro perfil profesional como analistas de datos.

7. VISIÓN DEL PROYECTO.

Elegimos esta temática porque nos permite trabajar con datos reales de salud pública, vinculando nuestras competencias técnicas con un problema de alto impacto global. El análisis estadístico y computacional de este dataset ofrece la posibilidad de identificar patrones de riesgo y mejorar la comprensión de variables asociadas a enfermedades cardíacas. Además, nos permite aplicar herramientas de ciencia de datos y estadística en un contexto médico, alineado con nuestro perfil profesional como analistas de datos.

8. SELECCIÓN DE ACCIONES.

Objetivo general:

Analizar de manera exploratoria un conjunto de datos médicos del Hospital Zheen, aplicando técnicas de procesamiento y estadística descriptiva, para identificar patrones vinculados a ataques cardíacos durante el primer cuatrimestre de 2025.

Objetivos específicos y Acciones.

Objetivo Específico	Acciones
Importar, depurar y estructurar los datos utilizando Pandas.	<ul style="list-style-type: none"> - Subir dataset al repositorio. - Importar con Pandas. - Verificar tipos de variables. - Eliminar o corregir registros erróneos.
Aplicar medidas estadísticas para describir las variables.	<ul style="list-style-type: none"> - Calcular media, mediana, desviación estándar. - Generar tablas de frecuencia para variables categóricas.
Identificar y graficar relaciones significativas entre variables.	<ul style="list-style-type: none"> - Usar gráficos de dispersión, boxplots y heatmaps. - Analizar correlaciones.
Evaluar la calidad del dataset.	<ul style="list-style-type: none"> - Identificar valores faltantes. - Tratar los datos RAW o incompletos. - Documentar el proceso.
Presentar los hallazgos mediante una exposición oral y escrita.	<ul style="list-style-type: none"> - Elaborar un informe en PDF. - Preparar defensa oral. - Usar visualizaciones claras.

9. CRONOGRAMA.

Objetivos	Mes 1 (Abril)	Mes 2 (Mayo)	Mes 3 (Junio)
Formar grupo, subir dataset, crear repositorio.	x		
Importar datos, revisar estructura.		x	
Calcular estadísticas descriptivas.		x	
Identificar valores		x	

faltantes.			
Documentar en Notebook.		x	
Completar análisis numérico y gráfico.			x
Graficar relaciones entre variables.			x
Armar informe y visualizaciones.			x
Preparar defensa oral y entrega final.			x

11. PRODUCTO FINAL.

El producto final consistirá en:

- Un notebook en Jupyter con todo el análisis de datos realizado.
- Un informe PDF con la documentación técnica de los resultados.
- Un repositorio con todo el trabajo realizado.

Estos productos permitirán presentar un caso real de análisis de datos médicos que puede contribuir a la concientización sobre enfermedades cardíacas. Se articula con ambos espacios curriculares: *Procesamiento de Datos y Estadística y Exploración de Datos I*.

12. BIBLIOGRAFÍA (Normas APA, 7ma edición).

- McKinney, W. (2022). *Python for data analysis: Data wrangling with pandas, NumPy, and Jupyter* (3rd ed.). O'Reilly Media.
- VanderPlas, J. (2017). *Python Data Science Handbook: Essential Tools for Working with Data*. O'Reilly Media.
- Cantatore de Frank, N. (1980). *Manual de Estadística Aplicada*. Tomo II. Editorial Hemisferio Sur.
- Cignacco, G., Craveri, A., & Teran, T. (2008). *Estadística Descriptiva y Probabilidad con Aplicaciones en las Ciencias Veterinarias*. Editorial UNR.

- Cignacco, G., Craveri, A., & Teran, T. (2009). *Inferencia Estadística con Aplicaciones*. Editorial UNR.
- Triola, M. F. (2009). *Estadística* (10ª ed.). Pearson Educación.
- Kaggle. (s.f.). *Heart Attack Dataset*. Recuperado de <https://www.kaggle.com>