

# Análisis exploratorio de Datos: Proyecto III

## Clustering Analysis

24 de junio de 2025

### 1. Objetivo

Analizar bases de datos de diferentes áreas del conocimiento aplicando los conceptos de clustering.

### 2. Instrucciones

En equipos de máximo 3 personas, realice las siguientes tareas:

1. **Seleccione una base de datos:** Investigue y determine las características de las instancias y atributos que la componen.
2. **Describa la base de datos:** Detalle cada uno de sus atributos, instancias y el contexto donde los datos fueron obtenidos o creados.
3. **Pre-procese los datos:** Identifique datos faltantes y atípicos. Establezca, describa y justifique la metodología utilizada para evitar que estos datos afecten el desempeño de los algoritmos que posteriormente se aplicarán.
4. **Identifique los grupos de instancia:** Realice el análisis de clustering correspondiente siguiendo la siguiente metodología:
  - a) Determine la tendencia de los datos a formar clusters.
  - b) Encuentre el número óptimo de grupos.
  - c) Utilice alguna técnica de reducción de la dimensión para representar el dataset.
  - d) Aplique al menos 4 algoritmos de clustering y compárelos.
5. **Asignación de clusters:** Para cada una de las instancias de la base de datos, determine la asignación al cluster correspondiente.
6. **Análisis de resultados:** Cada equipo deberá establecer las características de los grupos encontrados de acuerdo al contexto de la base de datos. Presentar tablas y gráficas que permitan mejorar su interpretación.

7. **Conclusiones:** Una vez que los clusters ya se encuentran asignados, deberán precisar cómo esta información puede resultar útil dentro de la toma de decisiones para los interesados. Elabore una propuesta de estrategia de aplicación.

**Formato:** IEEE Short paper. No exceder el máximo de páginas.