**PROPUESTA DE PROYECTO FINAL**

**ANÁLISIS AVANZADO DE DATOS**

**2024**

Se espera que cada grupo elija una problemática específica, identifique y recolecte un conjunto de datos apropiado, y luego aplique una combinación de técnicas y modelos analíticos para abordar la cuestión planteada

**Objetivos del proyecto:**

* Identificar una problemática real y relevante que pueda ser analizada mediante técnicas de análisis de datos.
* Seleccionar y preparar un conjunto de datos adecuado para el análisis.
* Elegir y justificar la combinación de al menos dos metodologías o modelos analíticos distintos aprendidos durante el curso.
* Realizar un análisis exhaustivo de los datos y validar los modelos seleccionados.
* Presentar los hallazgos y conclusiones de manera clara y bien fundamentada.

**Etapas del Proyecto:**

1. **Selección del Tema y del Dataset:**

* Definir una problemática de interés.
* Buscar y preparar un conjunto de datos que se ajuste a la problemática elegida.

1. **Metodología:**

* Seleccionar y justificar las metodologías y modelos a utilizar (por ejemplo, regresión, métodos de suavización, modelos generalizados lineales, análisis de datos dependientes, etc.).
* Desarrollar un plan de análisis coherente con los objetivos del proyecto.

1. **Análisis y Validación:**

* Aplicar las metodologías seleccionadas al conjunto de datos.
* Realizar la validación y evaluación de los modelos empleados.

1. **Presentación de Resultados:**

* Preparar un informe detallado que incluya la metodología, el análisis realizado, los resultados obtenidos y las conclusiones.
* Realizar una presentación que resuma los aspectos más importantes del proyecto.

**Evaluación**

la evaluación consta de dos etapas:

**Primera etapa (Entrega):**

En esta fase, se debe realizar la selección del tema y del dataset, así como establecer la metodología a seguir (Etapa 1 y 2 del proyecto). El grupo de estudiantes debe presentar un notebook que contenga el análisis exploratorio y la preparación de los datos para el modelado, junto con un documento que incluya la problemática planteada, la pregunta a resolver, el enlace a la base de datos a utilizar y las metodologías de análisis de datos que se emplearán en la siguiente fase.

Fecha de entrega de la primera etapa: 23 de Mayo

Porcentaje de evaluación:

* Heteroevaluación: 10%
* Autoevaluación grupal: 5%

**Segunda etapa (Entrega final):** Para esta entrega, se requiere presentar el análisis y validación de los modelos seleccionados previamente, así como una presentación de resultados que será utilizada para la coevaluación (Etapa 3 y 4 del proyecto).

Fecha de entrega de la segunda etapa: 30 de Mayo

Porcentaje de evaluación:

* Heteroevaluación: 15%
* Coevaluación: 10%

La coevaluación se llevará a cabo asignando a cada grupo de trabajo un proyecto desarrollado por otro grupo. A partir de una rúbrica establecida, se realizará la coevaluación de la siguiente manera:

* El grupo A coevaluará al grupo B.
* El grupo B coevaluará al grupo A.
* El grupo C coevaluará al grupo D.
* El grupo D coevaluará al grupo C.

La fecha de entrega de la coevaluación es el 1 de Junio.