****

**INSTITUTO SUPERIOR POLITÉCNICO DE CÓRDOBA**

**TECNICATURA SUPERIOR EN CIENCIA DE DATOS E INTELIGENCIA ARTIFICIAL**

**PROYECTO PARA EL MÓDULO**

**ANALISTA DE DATOS**

**Primer cuatrimestre – Año 2025**

**Docentes:**

* Marcos Ugarte
* Nahuel Pratta

**TEMÁTICA DEL PROYECTO**

Partiendo de un conjunto de datos sin depurar, llegar a obtener conclusiones de los mismos mediante el uso de gráficos, medidas de posición y dispersión.

**FUNDAMENTACIÓN**

El módulo Analista de Datos se compone de los EC Procesamiento de Datos y Estadística y Exploración de Datos 1. Lo que se busca con este ABP es la articulación de lo visto en los mismos para llegar a realizar un Proyecto completo que abarque los conocimientos abordados durante el primer cuatrimestre.

**OBJETIVOS DEL PROYECTO**

Los grupos deberán considerar los objetivos para esta primera etapa, teniendo en cuenta los contenidos de los espacios curriculares y la posterior elección de los datos.

**Competencias a desarrollar en el proyecto integrador**

* Capacidad para aplicar análisis gráficos y descriptivos.
* Capacidad de relacionar y/o comparar variables.
* Capacidad de organizar y planificar.
* Capacidad de aplicar los conocimientos e identificar problemas relevantes del contexto profesional.
* Capacidad de comunicar (argumentación y presentación de la información) de forma oral y escrita.
* Razonamiento crítico.
* Desarrollo de actitudes y valores (precisión, revisión, tolerancia), compromiso ético con la igualdad de oportunidades, con la no discriminación por razones de sexo, raza o religión y con la atención a la diversidad.
* Capacidad para la toma de decisiones.
* Capacidad de trabajar en equipo y en entornos multidisciplinares.
* **Estructura (secciones) del proyecto**:

**VER ARCHIVO: ESTRUCTURA PARA EL DISEÑO DE PROYECTOS ABP.**

* **Entregas**: habrá tres instancias de entregas parciales del proyecto:
* Primera entrega: **entre el 21 y 28 de abril**. Esta instancia conformará la nota de la **Evidencia 1**.

Para esta evidencia se solicita lo siguiente:

1. Crear grupos de entre 6 y 8 estudiantes. Poner un nombre. Subir un documento en pdf en la plataforma, en el cual figuren los nombres de los participantes y el link al repositorio que se detalla en el siguiente punto. **IMPORTANTE**, cada estudiante debe hacerlo (si no se considera como Ausente la entrega). También se debe detallar el usuario de GitHub de cada integrante.

2 - Crear en el repositorio una carpeta llamada “Datos” y subir un dataset que cumpla las siguientes características:

* Debe estar en formato .CSV
* El dataset debe contar al menos con al menos 5 variables de interés y 400 registros (ejemplo: 400 estudiantes de un Instituto a los cuales se les miden distintas variables: cualitativas y cuantitativas).
* Segunda entrega: **entre el 12 y 16 de mayo**. Esta instancia conformará la nota de la **Evidencia 2**.

Para esta evidencia se solicita lo siguiente:

1 - Crear una carpeta en el repositorio llamada “Documentación” la cual contendrá una descripción de los datos en formato PDF:

* Origen de los mismos.
* Descripción de cada columna que tipo de dato almacenan carácter, numérico (tipo) o fecha.
* Describir los valores posibles que puede tomar cada columna.
* Describir si los datos están en RAW o no.

2 - Crear un notebook y subirlo a la raíz del repositorio:

* Crear con Markdown una descripción del proyecto y los integrantes del equipo de trabajo (similar documentación)
* Importar los datos.
* Imprimir los primeros diez registros y describir el dataset.
* Realizar el siguiente análisis por cada columna:
* En las columnas numéricas analizar la cantidad de valores válidos sobre el total, media y mediana, desviación estándar, valores ausentes y mal registrados o erróneos.
* En las variables categóricas analizar la cantidad de registros para cada categoría, registros válidos sobre el total, ausentes y mal registrados.
* En otras variables que contienen registros válidos sobre el total, ausentes y mal registrados.
* Tercera entrega: **del 2 al 6 de junio**. Deben presentar el proyecto completo, además las sugerencias o revisiones previamente hechas por los docentes (esta instancia conformará la nota de la **Evidencia 3**).
* Entrega final de proyecto escrito: **del 9 al 13 de junio (junto a la defensa).** Al evaluar el proyecto completo, los docentes haremos (si es necesario) sugerencias o correcciones finales, que deberán incluirlas en la defensa oral.

* **Defensa y exposición oral:** se realizará de forma grupal durante un encuentro sincrónico, el **9 y 20 de junio** y deberán utilizar como soporte para compartir pantalla durante la reunión: un cuadro sinóptico que se adjuntará más adelante.

**CRITERIOS DE EVALUACIÓN**

* Desarrollo de todas las secciones del trabajo planteadas por los docentes y que en cada una esté desarrollada la información solicitada
* Coherencia entre todas las secciones del proyecto, es decir, que haya un hilo conductor.
* Claridad conceptual progresiva a lo largo del desarrollo del proyecto
* Uso correcto del vocabulario específico según cada EC
* En producciones escritas en entornos virtuales: redacción, organización y cohesión del texto, claridad. Caligrafía y ortografía.
* En producciones orales en entornos virtuales: destrezas comunicativas de hablar y escuchar generales en contextos formales de uso, expresiones orales, además se valorará la originalidad, creatividad y el interés del estudiante en prepararse sobre el tema que va a exponer
* Aplicación de los conceptos abordados durante el módulo.
* Cumplimiento de las entregas parciales y final del proyecto, respetando la estructura y formato propuestos, y los plazos de entrega.
* Incorporación de las revisiones y sugerencias realizadas por las docentes en las versiones parciales y final del proyecto.
* Diálogo e intercambio respetuoso entre pares y hacia la docente
* Predisposición y participación hacia el trabajo colaborativo
* Capacidad de síntesis e integración de la información, y su clara comunicación