

TECNOLÓGICO DE COSTA RICA
PROGRAMA DE CIENCIA DE DATOS
APRENDIZAJE AUTOMÁTICO (4B)

PROF: MSc. FELIPE MEZA

TUTORIAL PANDAS

Tarea #2 (Grupos de dos integrantes)

Crear un ambiente de trabajo (.ipynb) basado en PYTHON3 y llamarle `Tarea-2.ipynb`, con las librerías que considere oportuno, en especial debe contar con PANDAS.

El objetivo de la Tarea 2 consiste en elegir un dataset, ejecutar las diferentes tareas de pre-procesado y crear un tutorial sobre el uso de PANDAS considerando los siguientes aspectos:

- Debe ser dirigido a una audiencia con poca experiencia en el tema, imagine que bajo su rol de **Científico de Datos**, le han asignado un colaborador junior al que usted le corresponde alinearlos con los procesos de aprendizaje automático, en particular con el uso de PANDAS en labores de pre-procesado.
- Debe enfocarse en especial en aspectos al pre-procesado de datos.
- Se recomienda el uso de un dataset ejemplo, que permita “jugar” con las diferentes técnicas de PANDAS para pre-procesado. **Tome en cuenta que NO solo se trata de un ejercicio de pre-procesado de un dataset**, sino que adicionalmente debe construir un **TUTORIAL** sobre el uso de PANDAS, usando como base el dataset elegido.
- Recuerde seguir las buenas prácticas de comunicación efectiva, como parte de la notebook, utilizando los recursos que considere oportunos.
- Recuerde llevar a cabo el EDA y explotar al máximo la visualización, puede usar matplotlib con estilos varios, pero se recomienda instalar al menos otra librería extra de visualización, como Bokeh o Seaborn. **No olvide que el uso de la librería extra para la visualización, debe ser parte del tutorial.**
- Siéntase en libertad de agregar cualquier “truco” que considere oportuno o feature reciente incorporado a PANDAS!

El entregable consiste en el **url** de la notebook `Tarea-2.ipynb` en *GITHUB*.

Subir al TEC Digital un archivo .txt simple con el url de la notebook (github, collab...que pueda ser accedido) a más tardar el día acordado en clase e indicado en el TEC Digital, envíos pasado el límite de entrega se aplicará la penalización discutida en clase.