**“PROGRAMA ESPECIALIZADO CIENCIA DE DATOS CON PYTHON”**

**PROYECTO Nº 02**

****

## **Ejercicio Nº 1**

Fichero **autos.csv**

Generar una función y utilizar esta para recodificar la variable “**país de origen**” (1, 2, 3) en el dataframe a través de la función **apply.**

## **Ejercicio Nº 2**

Generar un Notebook Colab con código y resultados, con el siguiente fichero **autos.csv** que contiene información de autos de un periodo de tiempo. Resolver los siguientes:

* Limpieza y tratamiento de las variables (Según corresponda)
* Generar 5 nuevas variables aplicando los siguientes criterios con las variables numéricas:
* Log10(var1/var2)
* Sqrt(var1)exp(var2)/200
* Si var 1 > var 2 → 5, caso contrario 3
* 1/logn(var1/var2)\*100
* var2\*\*2/var1

Generar 3 variables cualitativas ordinales con los siguientes criterios:

* Si es mayor a la “media + 1\*desviación estándar” → Alto
* Si está entre “media - 1\*desviación estándar” y “media + 1\*desviación estándar” → Medio
* Si es menor a la “media - 1\*desviación estándar” → Bajo

Nota: Mostrar con las nuevas variables generadas.

Elegir con libertad las variables a trabajar.

* Crear una función que muestre los siguientes gráficos para las variables numéricas según una variable cualitativa:
* Cajas y bigotes
* Barras de la media
* Histograma con Curva
* Crear una función que devuelva un mapa de calor con la correlación de todas las variables numéricas (incluye las generadas), para cada país.

***Proyecto Nº 02****: Se adjunta Data:*

* *Entregable: generar un informe en Python Notebook, documentado paso a paso.*
* *Subir al GitHub.*
* *Enviar al correo:* ***yane.data@gmail.com***
* *Asunto: pro26.abr2022\_proyecto02\_nombredelalumno*

*\*. La tarea es personal.*

*Fecha de entrega máxima: 06 de junio del 2022*

Instructora: Msc (e) Ing. Est. E Inf. Yaneth Ccahuantico Mendoza