

Data Scientist | Equifax

Información sobre las bases

Tabla1_ejercicio_N3.csv:

Variables:

- *pers_id*: identificador único del cliente
- *sexo*: sexo del cliente
- *edad*: edad del cliente
- *ingreso*: ingreso estimado del cliente
- *situacion_lab*: categoría ocupacional del cliente

Tabla2_ejercicio_N3.csv:

Variables:

- *pers_id*: identificador único del cliente
- *score*: score de riesgo del cliente
- *periodo*: periodo al cual corresponde la información del score de riesgo

Ejercicio

El área de negocios de una empresa necesita información para la venta de un nuevo producto. Entregar análisis y conclusiones de los resultados de los siguientes ítems:

1. Crear una columna nueva (*ejemplo nombre tipo_ocupacion*) e indicar la cantidad de personas con y sin ocupación.
2. Teniendo en cuenta **el periodo más reciente disponible** para cada persona, crear una función que reciba como inputs:

Parámetro 1: datos input.

Parámetro 2: estadístico que se va a calcular.

Parámetro 3: variable a la que se va a aplicar el estadístico.

Parámetro 4: variable con la cual se va a agrupar.

Y calcular:

Score promedio y mediana según categoría ocupacional.

Ingreso y Score máximo y promedio según sexo.

Ejemplo de entorno online python: <https://colab.research.google.com/>