**PROCESAMIENTO, ANÁLISIS Y VISUALIZACIÓN DE DATA**

**CASO BANCO PERU**

**Descripción de variables**

Las variables de los empleados del BANCO PERÚ son:

* id: es el código del empleado en el Banco.
* genero: es el género del empleado.
* nacimiento: es la fecha de nacimiento del empleado.
* educacion: es el nivel educativo del empleado en años.
* catempleado: categoría de empleado.
* salario\_act: es el salario actual del empleado.
* salario\_ini: es el salario del empleado al momento de ingresar al Banco.
* tiempo: son los meses que tiene el empleado en el Banco desde la firma de su contrato.
* experiencia: es la experiencia previa del empleado en meses al momento de ingresar al Banco
* edad: es la edad del empleado en años.

1. Realizar un análisis de correlación entre las variables cuantitativas y presentar las conclusiones. Haciendo uso de:
   1. Visualización mediante gráfico de dispersión
   2. Prueba de correlación
   3. Grado de correlación
2. Presentar la matriz de correlación haciendo uso de las siguientes visualizaciones:
   1. Corrplot, car, corrgram, psych en R
   2. Heatmap, corrmat en Python
3. Comentar los resultados encontrados haciendo uso de hist, density y box en Python
4. Hacer un análisis comparativo entre las categorías de empleados ¿Cuáles son los insights que se obtienen?, 1: Secretario, 2: Seguridad, 3: Administrativo
5. Hacer un análisis comparativo entre las personas de género masculino y femenino ¿Cuáles son los insights que se obtienen? (Utilizar variables cuantitativas como salario actual vs categoria de empleado) los insights resultantes serian por ejemplo: el empleado secretario gana tanto de dinero.

Algunas variables que podemos crear

* Variacion del salario = (salario actual – salario inicial) / salario inicial
* Salario actual / educacion
* Salario inicial / experiencia1

Fecha de entrega: 22/02/2020

Correo: [rffv.uni@gmail.com](mailto:rffv.uni@gmail.com)

Asunto: Trabajo Final