

Introducción a Python aplicado a Economía y Finanzas

Tarea 2

- I. Copia el siguiente diccionario en una cédula de Colab. A partir de él, crea un dataframe llamado *nfo* utilizando Pandas.

```
data_ciudades = {
    'Ciudad': ['Lima', 'Arequipa', 'Trujillo', 'Cusco', 'Chiclayo', 'Piura', 'Iquitos', 'Huancayo', 'Tacna', 'Pucall'],
    'Habitantes': [1047996, 100169, 92331, 428450, 305717, 484475, 441649, 385098, 294395, 283734],
    'Area_km2': [2672.28, 1545.77, 1487.7, 116.5, 279.89, 6217.26, 368.9, 109.19, 59.4, 483.44],
    'Altitud_m': [154, 2325, 34, 3399, 29, 29, 106, 3271, 562, 156],
    'Densidad_poblacion_por_km2': [3924, 64.8, 62.1, 3673, 1091.6, 77.9, 1196, 3528.6, 4966, 587.2]
}
```

- II. Realiza el mismo procedimiento con el siguiente diccionario. A partir de él, crea un dataframe llamado *nombres*.

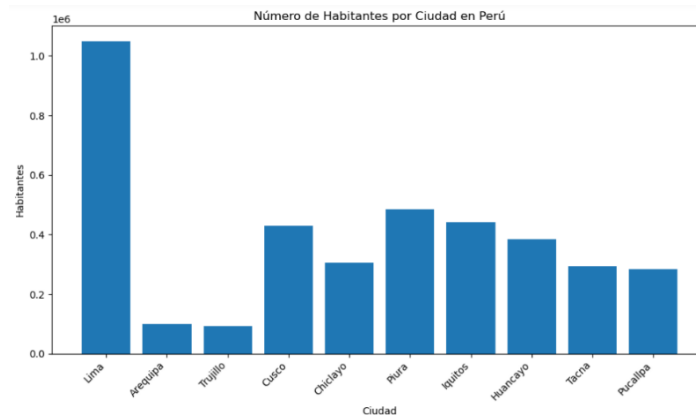
```
data_ciudades_2 = {
    'Ciudad': ['Lima', 'Arequipa', 'Trujillo', 'Cusco', 'Chiclayo', 'Piura', 'Iquitos', 'Huancayo', 'Tacna', 'Pucall'],
    'Gentilicio': ['Limense', 'Arequipeño', 'Trujillano', 'Cusqueño', 'Chiclayano', 'Piurano', 'Iquiteño', 'Huan'],
    'Provincia': ['Lima', 'Arequipa', 'Trujillo', 'Cusco', 'Chiclayo', 'Piura', 'Maynas', 'Huancayo', 'Tacna', 'Cor'],
    'Region': ['Lima', 'Arequipa', 'La Libertad', 'Cusco', 'Lambayeque', 'Piura', 'Loreto', 'Junín', 'Tacna', 'Ucay']
}
```

- III. Con ambos dataframes, realiza un inner join. Guarda el resultado en un nuevo dataframe llamado *cuadro_1*

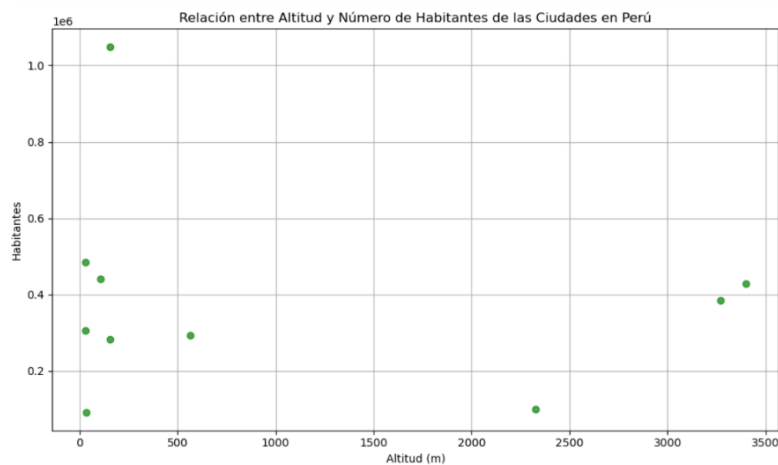
	Ciudad	Habitantes	Area_km2	Altitud_m	Densidad_poblacion_por_km2	Gentilicio	Provincia	Region
0	Lima	1047996	2672.28	154	3924.0	Limense	Lima	Lima
1	Arequipa	100169	1545.77	2325	64.8	Arequipeño	Arequipa	Arequipa
2	Trujillo	92331	1487.70	34	62.1	Trujillano	Trujillo	La Libertad
3	Cusco	428450	116.50	3399	3673.0	Cusqueño	Cusco	Cusco
4	Chiclayo	305717	279.89	29	1091.6	Chiclayano	Chiclayo	Lambayeque
5	Piura	484475	6217.26	29	77.9	Piurano	Piura	Piura
6	Iquitos	441649	368.90	106	1196.0	Iquiteño	Maynas	Loreto
7	Huancayo	385098	109.19	3271	3528.6	Huancaino	Huancayo	Junín
8	Tacna	294395	59.40	562	4966.0	Tacneño	Tacna	Tacna
9	Pucallpa	283734	483.44	156	587.2	Pucallpino	Coronel Portillo	Ucayali

- IV. Resume las estadísticas descriptivas del dataframe y responde ¿Cuál es la mínima densidad poblacional? ¿A cuál ciudad corresponde? ¿Y la máxima?

V. Realizar un gráfico de barras donde se vea la cantidad de habitantes en cada ciudad



VI. Realizar un gráfico de dispersión entre la altura y el número de habitantes en las ciudades del dataframe.



VII. Nombrar el archivo de Google Colab como GRUPO_NUMEROGUPO_Tarea1 y compartirlo al correo a esteban.cabrera@pucp.edu.pe y m.encaladac@pucp.edu.pe dejar el link en el Google Sheets hasta el lunes 19 de agosto a medianoche.

VIII. BIBLIOGRAFÍA

Los scripts desarrollados en clase serán el principal material de referencia el cual será entregado al inicio de cada clase.

Enlace de la carpeta Drive:

<https://drive.google.com/drive/folders/17bgV4HpqzhWbVvAaOHlxjvmhZwrF9Xyr>

Enlace del repositorio:

<https://github.com/qlabpucp/python-basics>

De manera complementaria, se recomienda revisar la siguiente literatura:

- Witten, D., & James, G. (2013). An introduction to statistical learning with applications in Python. Springer publication.