

## Ejercicio Individual AE5

Plan Formativo: Ciencia de Datos	Nivel de Dificultad
Módulo: Análisis exploratorio y programación estadística.	Bajo / medio
Tema: Librería Seaborn.	
Intención del aprendizaje o aprendizaje esperado:	
<ul style="list-style-type: none"> <li>● Presenta información agregada utilizando funciones de agrupación, agregación y pivoteo para el análisis de un set de datos</li> </ul>	
Ejercicios planteados	
<p>Para este ejercicio, primero debe obtener los datos de <b>gapminder</b> de la URL de carpintería de software y cargarlos como marco de datos de pandas. Los datos de gapminder tienen información sobre la vida útil, la población y el pib de los países durante varios años.</p> <p>Para eso use el siguiente código:</p> <pre>gapminder_csv_url = 'http://bit.ly/2cLzoxH' gapminder = pandas.read_csv(gapminder_csv_url) print(gapminder.head(3))</pre> <p>A partir de esto, realice lo siguiente:</p> <ol style="list-style-type: none"> <li>1. Realice un resumen estadístico de sus datos.</li> <li>2. Agrupe por la variable “año”. Comente respecto al resultado al usar groupby.</li> <li>3. Ahora obtenga la suma (la cantidad), la media y la mediana para cada grupo, es decir, por año.</li> <li>4. ¿Sabías que en 1967 en Chile se inaugura el Aeropuerto Internacional Arturo Merino Benítez? Usando estos datos se puede especificar un año</li> </ol>	

y obtener un marco de datos más pequeño para el año usando la función `get_group()`. Por ejemplo, podemos obtener el marco de datos agrupados para el año 1967. Comenta tus resultados. Escoge otro año explicando el porqué de la elección y qué resultados obtuviste.

Ahora con los datos anteriores construya tablas pivote según las siguientes instrucciones:

1. Construya una tabla fijando “continent”, con las funciones vistas en clase
2. Realice una tabla de doble entrada, fijando “continent” y “year”. Comente sus resultados.
3. Con las variables anteriores, obtenga el promedio de la variable “pop”.
4. Obtenga un gráfico de barras del punto 3, comente sus resultados.

#### Datos de apoyo al planteamiento

#### Preguntas guía

- Tratamiento de datos

#### Recursos Bibliográficos:

[1] Datos agrupados

<https://www.ibm.com/docs/es/spss-statistics/SaaS?topic=chart-binning-grouping-data-values>

[2] Función separar - aplicar - combinar

<https://medium.com/analytics-vidhya/split-apply-combine-strategy-for-data-mining-4fd6e2a0cc99>

[3] Conceptos estadísticos

<https://docs.python.org/es/3/library/statistics.html>

[4] Tablas pivote

<https://ichi.pro/es/aqui-se-explica-como-construir-una-tabla-dinamica-usando-pandas-en-python-112177747736514>