Ejercicio Individual AE5

| Plan Formativo: Ciencia de Datos | Nivel de Dificultad |
| --- | --- |
| Módulo 3: Análisis exploratorio y programación estadística. | Bajo / medio |
| Tema: Librería Seaborn. | |  | | --- | |
| Intención del aprendizaje o aprendizaje esperado: | |
| | * Presenta información agregada utilizando funciones de agrupación, agregación y pivoteo para el análisis de un set de datos. | | --- | | |
| Ejercicios planteados | |
| Para este ejercicio, primero debe obtener los datos de **gapminder** 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.  Para eso use el siguiente código:  gapminder\_csv\_url ='<http://bit.ly/2cLzoxH>'  gapminder = pandas.read\_csv(gapminder\_csv\_url)  print(gapminder.head(3))  A partir de esto, realice lo siguiente:   1. Realice un resumen estadístico de sus datos. 2. Agrupe por la variable “año”. Comente respecto al resultado al usar groupby. 3. Ahora obtenga la suma (la cantidad), la media y la mediana para cada grupo, es decir, por año. 4. ¿Sabías que en 1967 en Chile se inaugura el Aeropuerto Internacional Arturo Merino Benítez? Usando estos datos se puede especificar un año 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> | |