UNIVERSIDAD METROPOLITANA CASTRO CARAZO

FACULTAD DE INFORMÁTICA

Técnico en análisis de datos

TEMA

Indagación - Seaborn

NOMBRE DEL ESTUDIANTE

Marco Campos García

NUMERO DE CEDULA

114660668

TRABAJO DE INVESTIGACIÓN

CURSO DE ANÁLISIS DE DATOS

SAN JOSÉ, COSTA RICA

Agosto – 2023

Bibliotecas : SEABORN

Python es uno de los tantos lenguajes para programación que existen, este se destaca por su facilidad para codificar, y se puede aplicar en diferentes ámbitos, tales como la tecnología, machine learning y análisis de datos.

Teniendo esto en cuenta, existen varias herramientas que nos ayudan a implementar las diferentes tipos de bibliotecas que existen en el campo del Data Analyst, algunas de ellas podrían ser como :

* **Numpys** : Es una biblioteca de código abierto que se utiliza principalmente para la computación científica, donde sus funciones permiten soportar datos multidimensionales y su rapidez para procesar los datos, mediante sus funciones matemáticas.
* **Pandas**: Es una biblioteca, que permite modelar y poder manipular de manera eficaz, los diferentes tipos de archivos de datos, en sus diferentes tipos de formatos Excel, archivos de texto y CSV, Microsoft, formatos HDF5 y bases de datos SQL.
* **Matplotlib:** Es una biblioteca de código abierto, que permite crear visualizaciones en python fijas, interactivas y animadas, donde incluye varias interfaces de plot tales como Seaborn, HoloViews, ggplot, etc)

Con esto tenemos en cuenta algunas de las principales bibliotecas vistas hasta el día de hoy(en clase), donde buscamos extender el tema de la biblioteca de SEABORN.

BIBLIOTECA : SEABORN

La visualización de los diferentes tipos de datos permite a los científicos de datos poder transformar data “cruda”, en gráficos y diagramas para poder ilustrar los diferentes estudios.

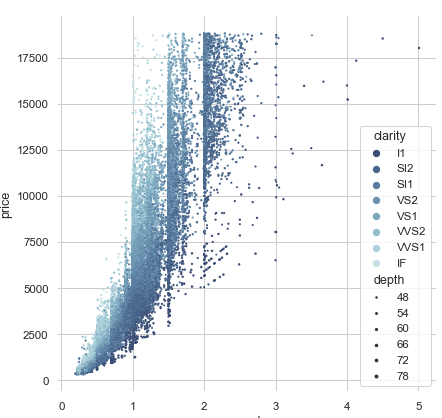
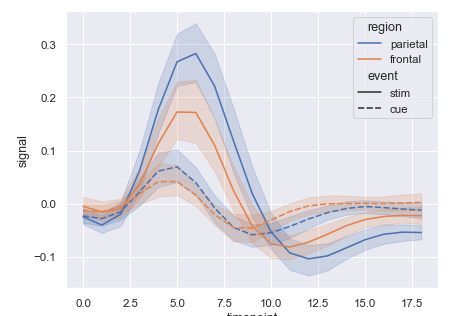
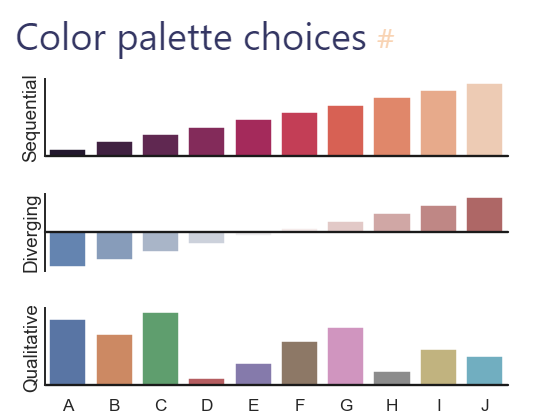
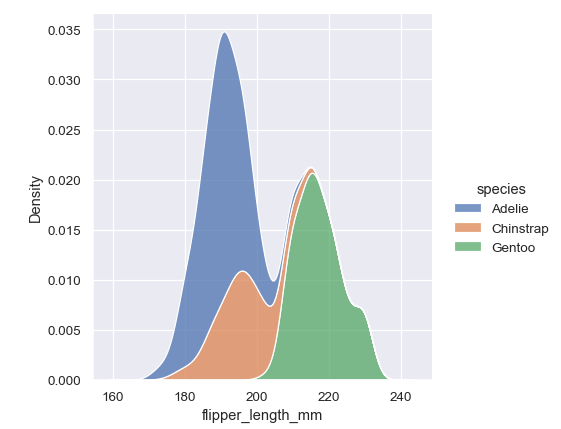
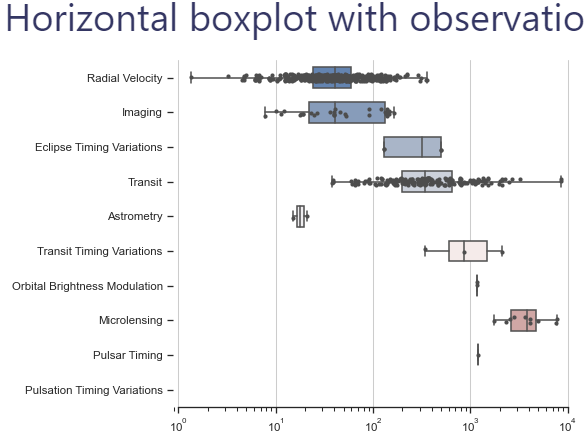
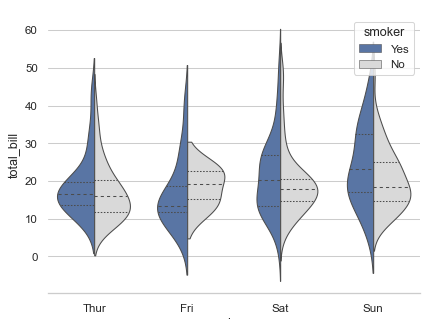
Seaborn, es otra biblioteca de visualización de datos, que está basado en Matplotlib, teniendo una interfaz de alto nivel utilizada para crear visualizaciones estadísticas estéticamente atractivas y valiosas para un análisis, conectada entre los ya mencionados Numpys, y Pandas, para poder ser esencial en el análisis y exploración de los datos.

A diferencia de Matplotlib, Seaborn es una biblioteca que aporta simplicidad a la hora de usarla y permite explorar rápidamente los datos



Seaborn proporciona diferentes tipos y paletas de colores, para crear gráficos atractivos para poder interpretar de manera más sencilla, tiene diferentes tipos de visualizaciones numéricas, clases, y grupos.

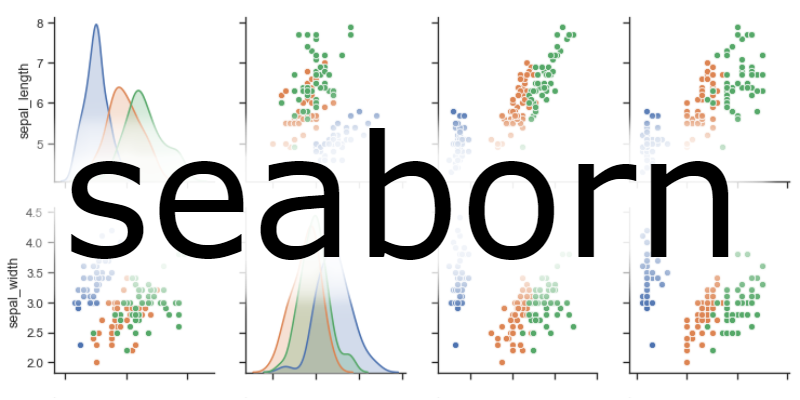
En ellos se pueden lograr tomar algunos de los siguientes gráficos:

* Gráficos de dispersión (scatterplots)
* Gráficos de líneas (line plots)
* Gráficos de barras , apilados , agrupados (bar plots)
* Gráficos de distribución (distribution plots)
* Gráficos de cajas (box plots)
* Gráficos de violín (violin plots) 

Estos son algunos de los diferentes tipos de gráficos que podemos realizar en esta biblioteca donde nos permite, una interfaz más sencilla de utilizar, en comparación a Matplotlib, ya que tiene una amplia variedad de esquemas para las visualizaciones estadísticas.

Parte de las ventajas que tiene esta biblioteca, es su flexibilidad en términos de personalización y algunas veces un rendimiento superior a los otros.

Esto para los analistas de datos permite crecer , ya que es una herramienta en todos los sectores empresariales que se puede permitir usar, para los distintos tipos de análisis.



Conclusión

Como conclusión de esta indagación es importante tener en cuenta como estudiante y/o científico de datos, saber los diferentes tipos de bibliotecas para las diferentes herramientas que vamos a ir aplicando a lo largo del estudio y la aplicación de los conocimientos, en este caso como estamos investigando SEABORN, es importante saber que tipos de datos se pueden visualizar, de diferentes maneras, con los diferentes tipos de gráficos que existen, y cómo aplicarlos, conocer las diferentes funciones que pueden llegar a tener en cada proyecto, esto para explotar a su máximo el potencial que tiene la herramienta.

Es importante seguir indagando este tipo de herramientas para el desenvolvimiento de los proyectos de los clientes en el campo.

Bibliografía

<https://www.maseldata.com/post/python-librerias-data-science>

<https://seaborn.pydata.org/tutorial/distributions.html>

https://seaborn.pydata.org/examples