

Visualización de la información - Tarea 7

Andrés Urbano Guillermo Gerardo

22 de Marzo del 2022

Introducción

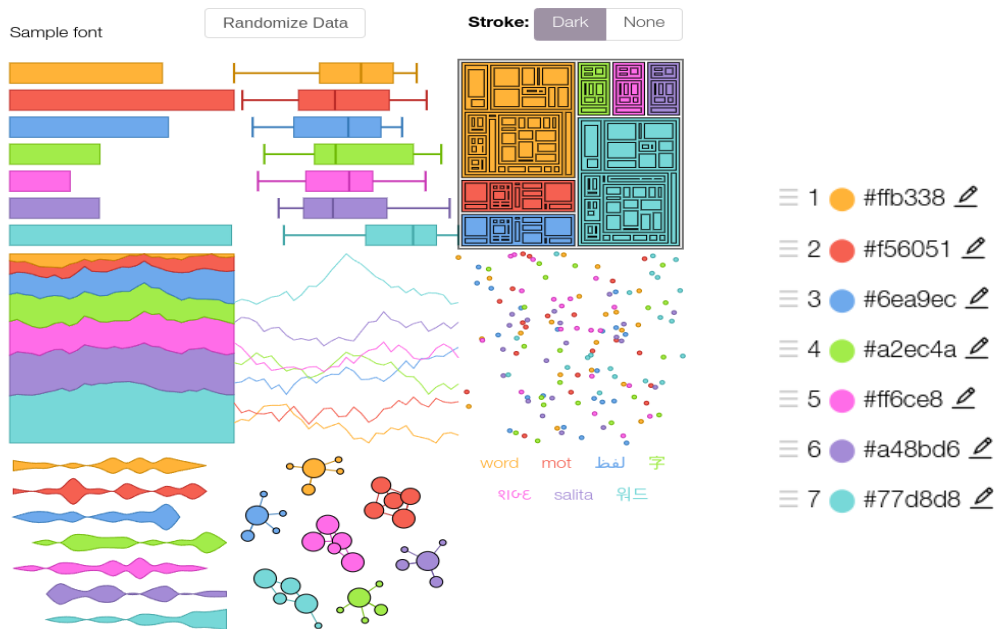
El uso correcto de colores en una visualización es fundamental para un mejor entendimiento de nuestros datos. Ayuda a encontrar patrones que a simple vista tardaríamos más, por lo que es importante buscar herramientas para probar diferentes paletas de colores con distintas visualizaciones. Viz Palette es una herramienta para ayudar a los diseñadores de visualización de datos a evaluar y mejorar sus paletas.

Descripción del problema

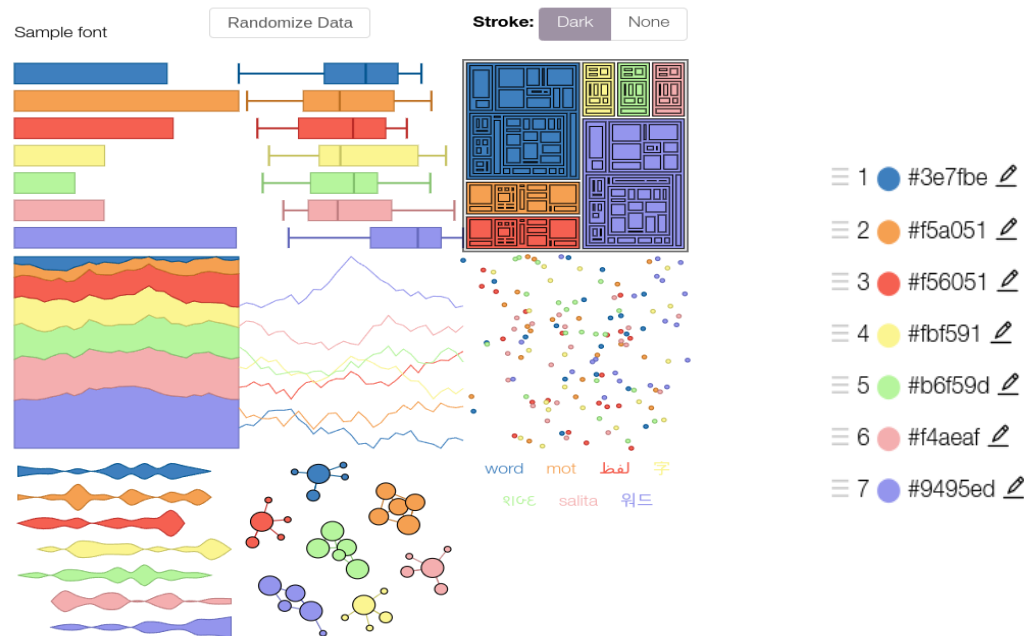
Encontrar la mejor paleta que ayude a visualizar diferentes gráficos.

Actividad 1

Generar dos paletas de color que resalten mejor los datos.



Elegí esta paleta de colores porque considero que para cada gráfica resaltan los valores sin llegar confundir a que clase pertenece si fueran datos. Por otro lado los colores son vividos para una mayor claridad para cualquier gráfico.



Para esta paleta de colores, seguí con la misma idea anterior, pero ahora generando colores más tenues, con mayor transparencia para lograr un estilo de colores pastel. Vemos que el resultado es correcto para muchas gráficas, es posible que para otras se vean no tan nítido por los pocos gráficos que se manejan para la gráfica, pero en su mayoría todo sea adecuada perfectamente.

Actividad-2

Se requiere mostrar una Curva de densidad con histograma, que relacione el kilometraje en ciudad (cty) con la clase de vehículo (class) en particular los tipos compacto, SUV y minivan.

```
[28]: df = pd.read_csv("data/mpg_cars.csv")
df.head()
```

```
[28]:  manufacturer model  displ  year  cyl    trans  drv  cty  hwy  fl  class
0         audi    a4    1.8  1999    4  auto(15)   f   18   29   p  compact
1         audi    a4    1.8  1999    4  manual(m5)  f   21   29   p  compact
2         audi    a4    2.0  2008    4  manual(m6)  f   20   31   p  compact
3         audi    a4    2.0  2008    4  auto(av)    f   21   30   p  compact
4         audi    a4    2.8  1999    6  auto(15)   f   16   26   p  compact
```

Para ello, considere lo siguiente:

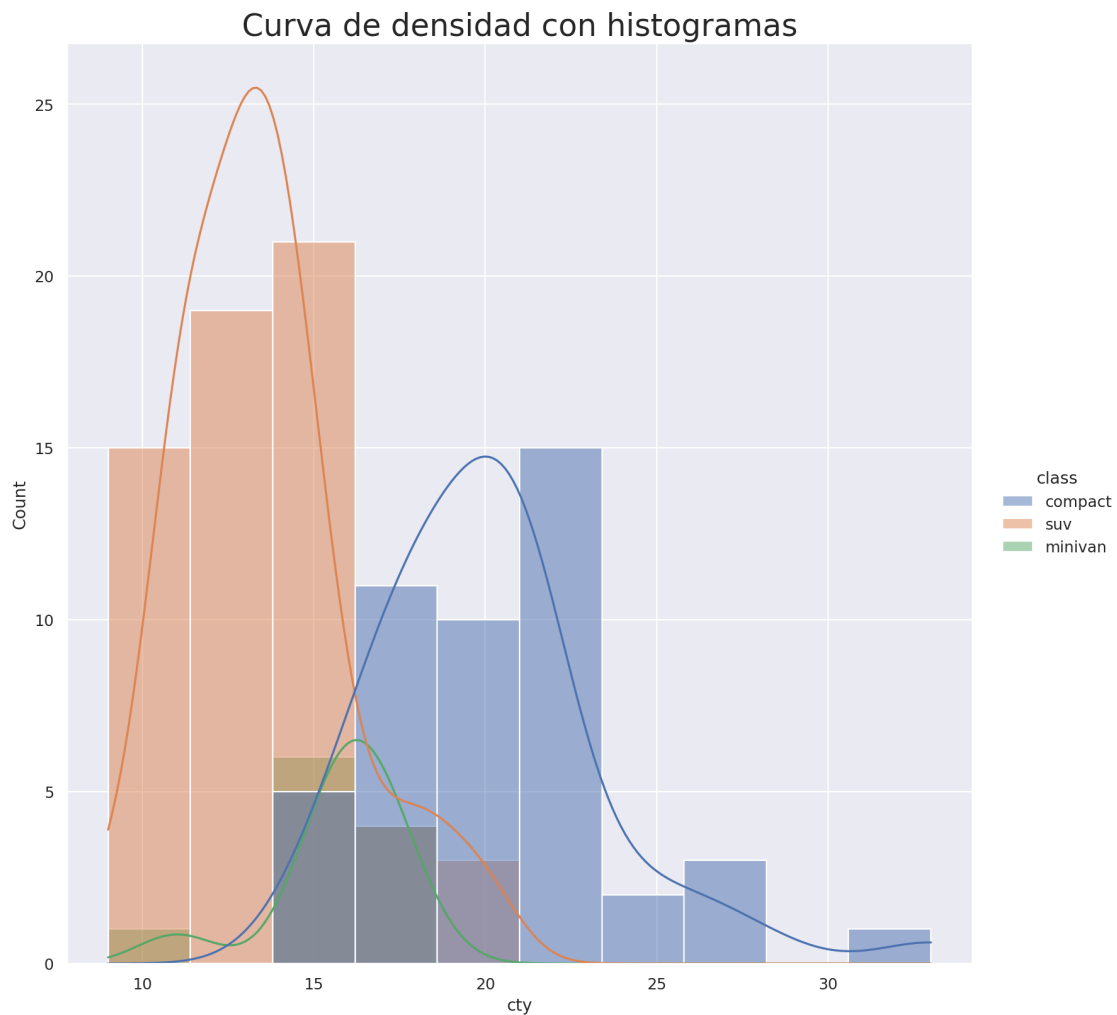
- Para crear el tamaño de la Curva de densidad con histograma que representan los datos, considere una dimensión de la imagen de 1040 x 800 pulgadas y 80 puntos (pixel) por pulgada (dpi).
- Establezca la Curva de densidad con histograma (línea e histograma en el mismo gráfico) con un color diferente para cada clase.
- Para el caso del histograma ajuste la transparencia de la grafica a 0.7 y para la gráfica de línea, considere un ancho de línea de 3 puntos.
- Escriba un título para la Curva de densidad con histograma el cual deberá tener un tamaño de fuente de 22 puntos.
- De acuerdo con la gráfica obtenida describa brevemente sus conclusiones.

Seleccionamos solamente la clase de automoviles que nos interesan:

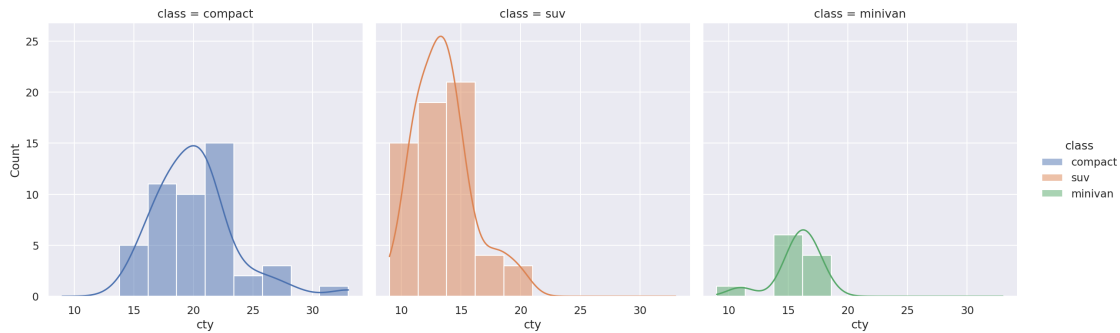
```
[62]: data = df[(df['class'] == 'suv') | (df['class'] == 'minivan') | (df['class'] == 'compact')]
```

Mostramos nuestro datos mediante una curva de densidad:

```
[166]: p = sns.displot(data=data, x="cty", hue="class", kde=True, height=10); # 1040 x 800
p.fig.set_dpi(80)
plt.title('Curva de densidad con histogramas', fontsize=22);
```



```
[167]: sns.displot(data=data, x="cty", hue="class", col="class", kde=True);
```



Vemos que el kilometraje es mayor con los autos de la clase compacto, seguido despues de los suv y por último lo de minivan. Además vemos que siguen una distribucion normal, aunque sea muy pocos datos, vemos que inicia con un valor pequeño y va incrementandose hasta llegar a un maximo para posteriormente ir disminuyendo. Igual podemos observar que en nuestro conjunto de datos hay una mayor cantidad de autos de la clase suv y muy pocos minivan, por lo que podemos inferir que los clientes suelen comprarlo más tal vez por su kilometraje.

Conclusiones

Vemos que la importancia de los colores, puede ayudarnos encontrar mejores patrones y darnos diferentes perspectivas de los datos con los que trabajamos. Además vemos que para diferentes visualizaciones el uso de una misma paleta de colores podría existir algunos gráficos que se no lleven bien del todo, por lo que siempre sera adecuado ir probando para los diferentes gráficos a utilizar. Además vemos que es útil unir diferentes visualizaciones para dar una mejor visión.