```
In [1]:  import matplotlib.pyplot as plt
         import seaborn as sns
         from sklearn.cluster import KMeans
         import pandas as pd
         import numpy as np
         from PIL import Image
```

```
In [2]:  def rgb_to_hex(rgb):
             return '#%02x%02x%02x' % (int(rgb[0]), int(rgb[1]), int(rgb[2]))
```

```
In [4]:  # Convert PIL image to array
         img = Image.open('dogs.jpeg')
         np_array = np.array(img)

         print("Array shape")
         print(np_array.shape)
```

```
Array shape
(185, 272, 3)
```

```
In [5]:  print("\nDimensions of the image")
         print("Height : "+str(img.height))
         print('Width : '+str(img.width))
```

```
Dimensions of the image
Height : 185
Width : 272
```

```
In [8]:  #Reshape the array in 2 dimensions
         flat_array = np_array.ravel()
         new_array = flat_array.reshape(np_array.shape[0]*np_array.shape[1], np_array.s
         hape[2])
         new_array
```

```
Out[8]:  array([[240, 240, 240],
                [255, 255, 255],
                [255, 255, 255],
                ...,
                [254, 254, 254],
                [254, 254, 254],
                [254, 254, 254]], dtype=uint8)
```

```
In [10]:  #Initializing the Kmeans
          kmeans = KMeans(n_clusters=3)
          kmeans.fit(new_array)

          print("\nCluster_Centroid")
          print(kmeans.cluster_centers_)
```

```
Cluster_Centroid
[[210.72326684 174.95593525 143.85791367]
 [ 29.52558775  21.59168945  20.48217605]
 [251.51161734 247.74204465 243.25012628]]
```

```python
#Count elements of each clusters
print("Elements of each cluster ")
unique, counts = np.unique(kmeans.labels_, return_counts = True)
print(dict(zip(unique, counts)))
```

```
Elements of each cluster
{0: 12228, 1: 18293, 2: 19799}
```

```python
df = pd.DataFrame(new_array, columns=["col1", "col2", "col3"])
df["cluster"] = kmeans.labels_
print(df.head())
```
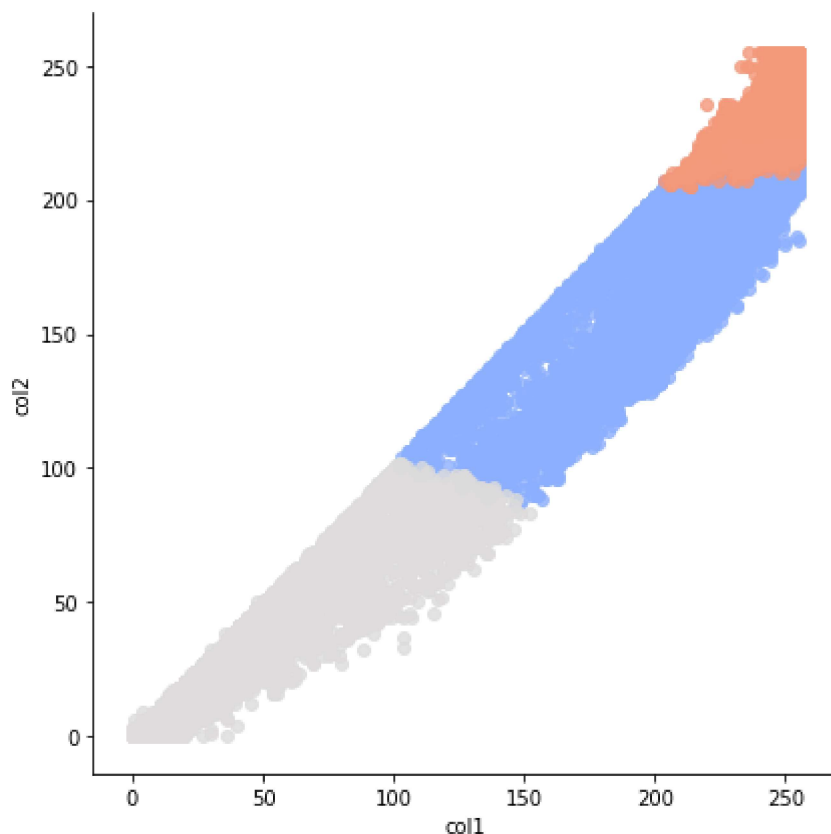
```
     col1  col2  col3  cluster
0    240   240   240        2
1    255   255   255        2
2    255   255   255        2
3    251   251   251        2
4    255   255   255        2
```

```python
#plot cluster points
sns.lmplot('col1', 'col2', data=df, hue="cluster", palette='coolwarm', size=6,
aspect=1, fit_reg=False)
```

```
C:\Users\hp\anaconda3\lib\site-packages\seaborn\regression.py:574: UserWarnin
g: The `size` parameter has been renamed to `height`; please update your cod
e.
  warnings.warn(msg, UserWarning)
```

<seaborn.axisgrid.FacetGrid at 0x25bbcd27448>

```
In [20]: X_compressed = kmeans.cluster_centers_[kmeans.labels_]
         X_compressed = np.clip(X_compressed.astype('uint8'), 0, 255)
         X_compressed = X_compressed.reshape(np_array.shape[0], np_array.shape[1], np_a
         rray.shape[2])

         fig, ax = plt.subplots(1,2, figsize=(20,8))
         ax[0].imshow(img)
         ax[0].set_title('Original Image')
         ax[1].imshow(X_compressed)
         ax[1].set_title('Compressed Image with 3 colors')
         for ax in fig.axes:
             ax.axis('off')
         plt.tight_layout()
```


Original Image — Compressed Image with 3 colors

```
In [21]: # Find Dominant color
         centroid_colors = kmeans.cluster_centers_.astype(int)
         colours = [rgb_to_hex(color) for color in centroid_colors]

         sns.palplot(colours)
         plt.show()
```



```
In [ ]:
```