* * * * * * * * * * * * * * * * * Assignment no 5 * * * * * * * * * * * * * * * * *

Name - Kiran Shete

Roll no - 23577

Importing all Libaries

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

plt.style.use("seaborn-whitegrid")
plt.rc("figure", autolayout=True)
plt.rc(
    "axes",
    labelweight="bold",
    labelsize="large",
    titleweight="bold",
    titlesize=14,
    titlepad=10,
)

df = pd.read_csv("housing.csv")
X = df.loc[:, ["MedInc", "Latitude", "Longitude"]]
X.head()
```

Out[1]:

	MedInc	Latitude	Longitude
0	8.3252	37.88	-122.23
1	8.3014	37.86	-122.22
2	7.2574	37.85	-122.24
3	5.6431	37.85	-122.25
4	3.8462	37.85	-122.25

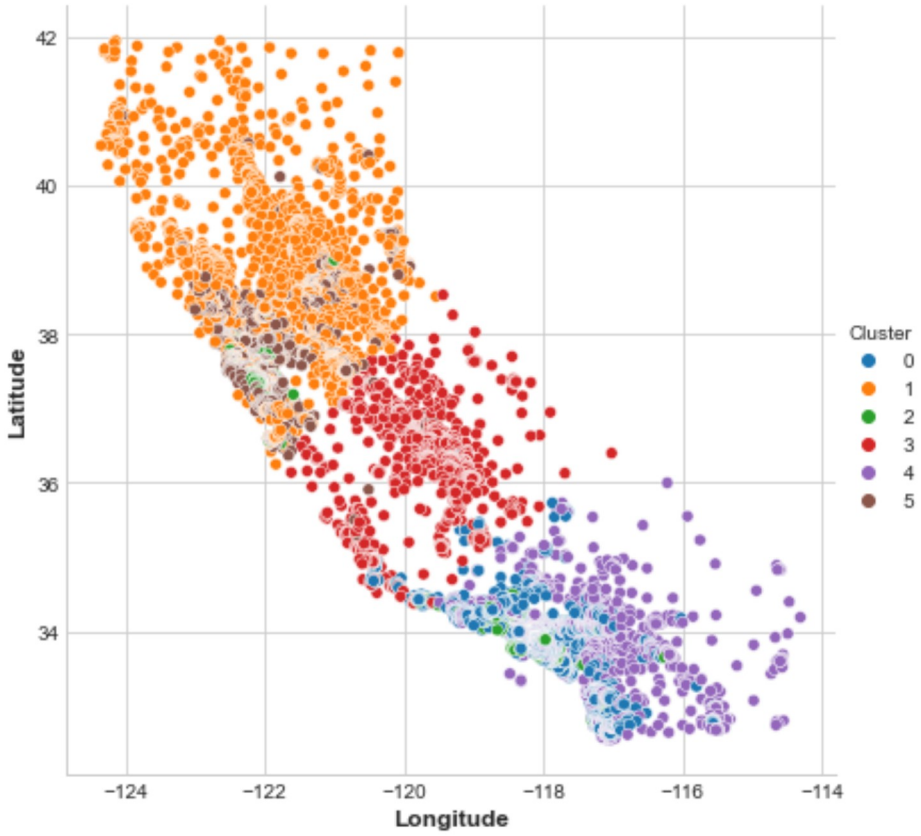
In [2]:
```
# Create cluster feature
kmeans = KMeans(n_clusters=6)
X["Cluster"] = kmeans.fit_predict(X)
X["Cluster"] = X["Cluster"].astype("category")

X.head()
```

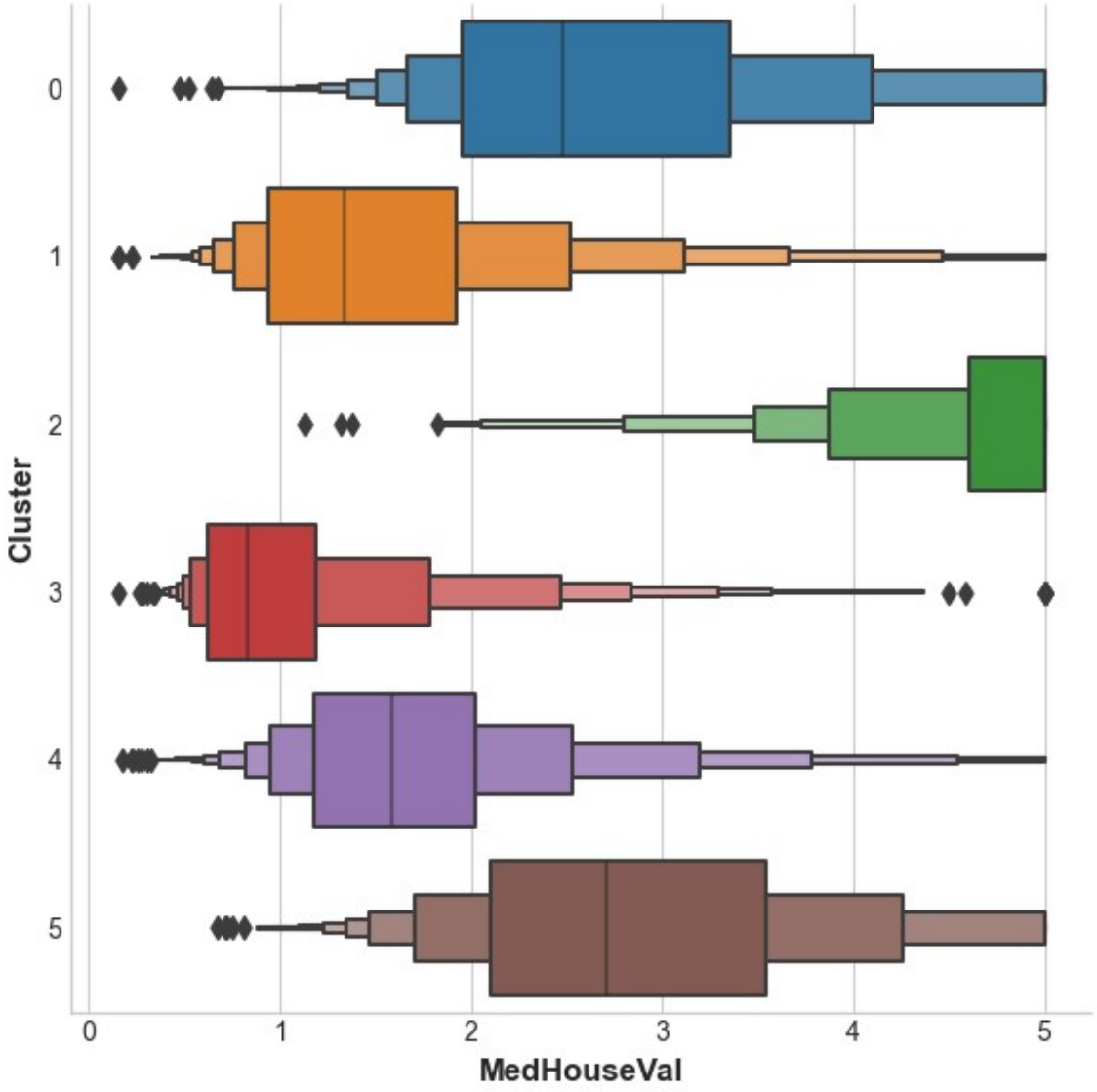
Out[2]:

	MedInc	Latitude	Longitude	Cluster
0	8.3252	37.88	-122.23	5
1	8.3014	37.86	-122.22	5
2	7.2574	37.85	-122.24	5
3	5.6431	37.85	-122.25	5
4	3.8462	37.85	-122.25	1

In [3]:
```
sns.relplot(
    x="Longitude", y="Latitude", hue="Cluster", data=X, height=6,
);
```



In [4]:
```
X["MedHouseVal"] = df["MedHouseVal"]
sns.catplot(x="MedHouseVal", y="Cluster", data=X, kind="boxen", height=6);
```



In [ ]: