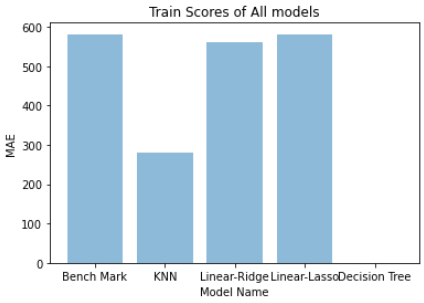
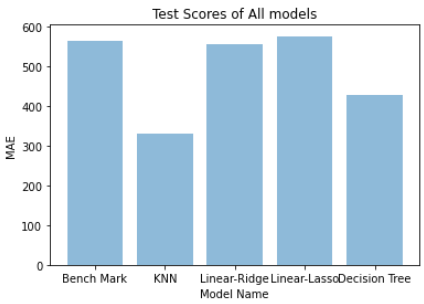
In [1]:
```
import matplotlib.pyplot as plt

#data
x = ['Bench Mark', 'KNN', 'Linear-Ridge','Linear-Lasso', 'Decision Tree']
train_score = [581.9665110196446, 281.32469243164087, 560.9328633555573,580.6874771869304, 0.0026259746405877566]

#bar plot
plt.bar(x, height=train_score, alpha=0.5)
plt.title('Train Scores of All models')
plt.xlabel('Model Name')
plt.ylabel('MAE')
plt.show()
```



In [2]:
```
#data
x = ['Bench Mark', 'KNN', 'Linear-Ridge','Linear-Lasso', 'Decision Tree']
train_score = [565.4238077668508, 329.4823997424956, 556.8012764347889,576.2106723643577, 428.37537008928314]

#bar plot
plt.bar(x, height=train_score, alpha=0.5)
plt.title('Test Scores of All models')
plt.xlabel('Model Name')
plt.ylabel('MAE')
plt.show()
```



it looks like the KNN model has lowest error and the linear lasso has the highest error

Data doesnt satisfy the linear regression assumptions like multi collinearity(>50% of variables have variance inflation factor (VIF) more than 5), this could be the reason for the poor performance

In [ ]: