

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline

import seaborn as sns
sns.set(style="whitegrid", color_codes=True)
```

```
In [2]: metric = pd.read_csv('decision_model_evaluation.csv')
```

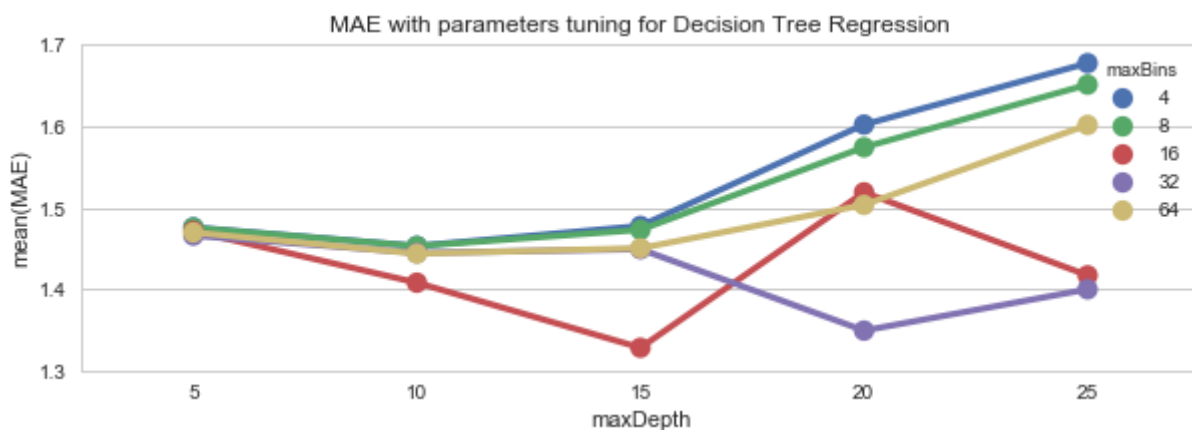
```
In [3]: metric.loc[123, 'RMSE'] = 1.487954
```

```
In [4]: print(metric[metric['RMSE'] == metric['RMSE'].min()])
```

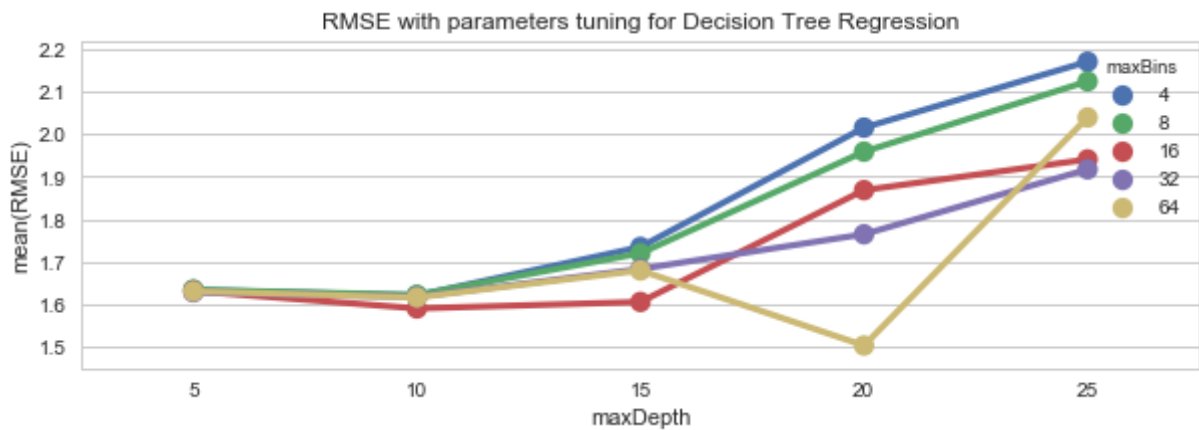
	SampleData	maxDepth	maxBins	MAE	RMSE	TIME
127	400	15	4	1.232904	1.475254	NaN

```
In [5]: param = metric[metric['SampleData'] == 50]
```

```
In [20]: f, ax = plt.subplots(figsize=(10, 3))
plt.title('MAE with parameters tuning for Decision Tree Regression')
sns.pointplot(x=param['maxDepth'], y=param['MAE'], hue=param['maxBins'], data=param)
```

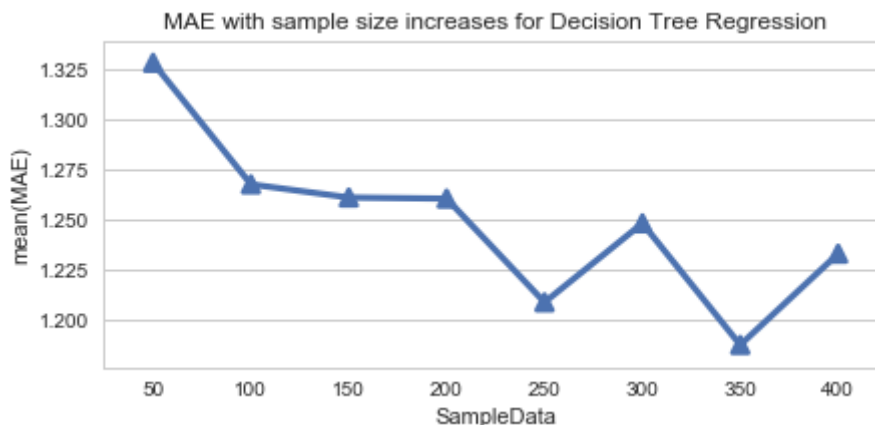


```
In [21]: f, ax = plt.subplots(figsize=(10, 3))
plt.title('RMSE with parameters tuning for Decision Tree Regression')
sns.pointplot(x=param['maxDepth'], y=param['RMSE'], hue=param['maxBins'], data=param)
```

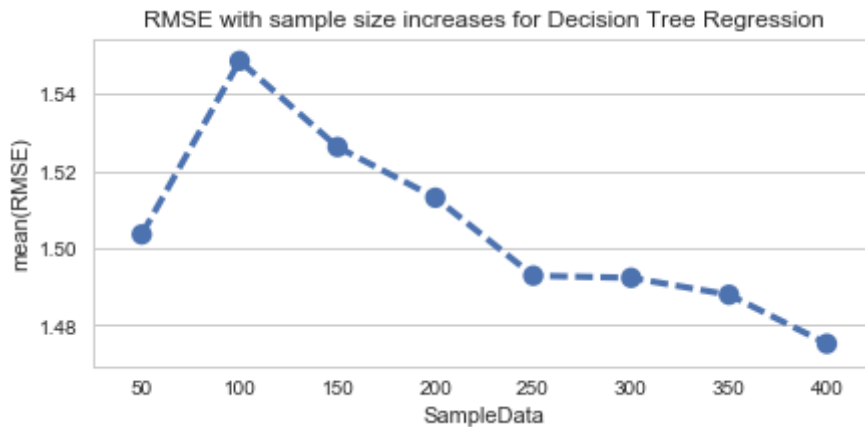


```
In [8]: sample = metric.groupby(metric['SampleData']).min().reindex()
```

```
In [9]: f, ax = plt.subplots(figsize=(7, 3))
plt.title('MAE with sample size increases for Decision Tree Regression')
sns.pointplot(x=sample.index, y=sample['MAE'], data=sample,
               markers=["^"], linestyle=["-"]);
```



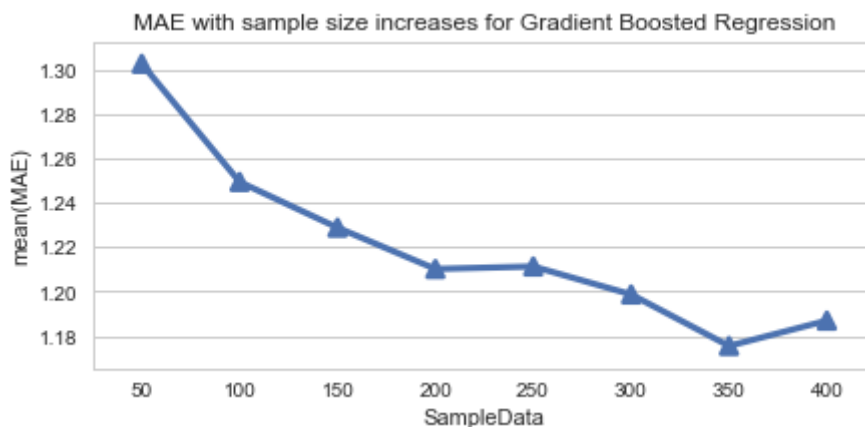
```
In [10]: f, ax = plt.subplots(figsize=(7, 3))
plt.title('RMSE with sample size increases for Decision Tree Regression')
sns.pointplot(x=sample.index, y=sample['RMSE'], data=sample,
               markers=["o"], linestyle=["--"]);
```



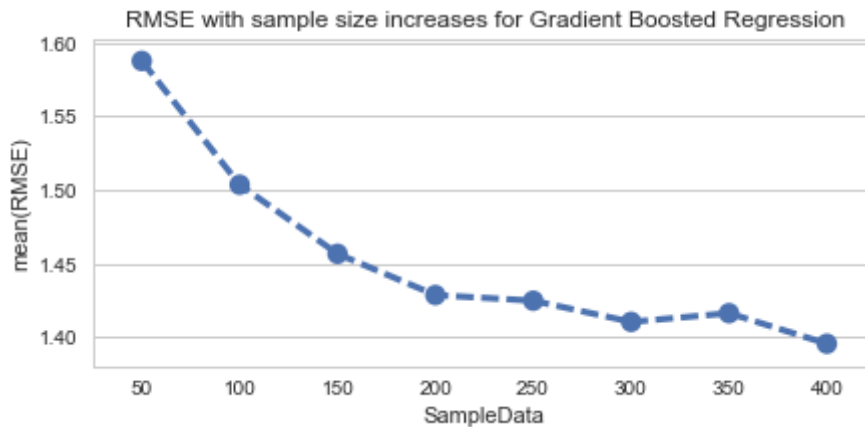
```
In [11]: gbt_metric = pd.read_csv('gbt_model_evaluation.csv')
```

```
In [12]: gbt = gbt_metric[gbt_metric['Unnamed: 0'] == 'gbt']
```

```
In [13]: f, ax = plt.subplots(figsize=(7, 3))
plt.title('MAE with sample size increases for Gradient Boosted Regression')
sns.pointplot(x=gbt['SampleData'], y=gbt['MAE'], data=gbt,
               markers=["^"], linestyle=["-"]);
```

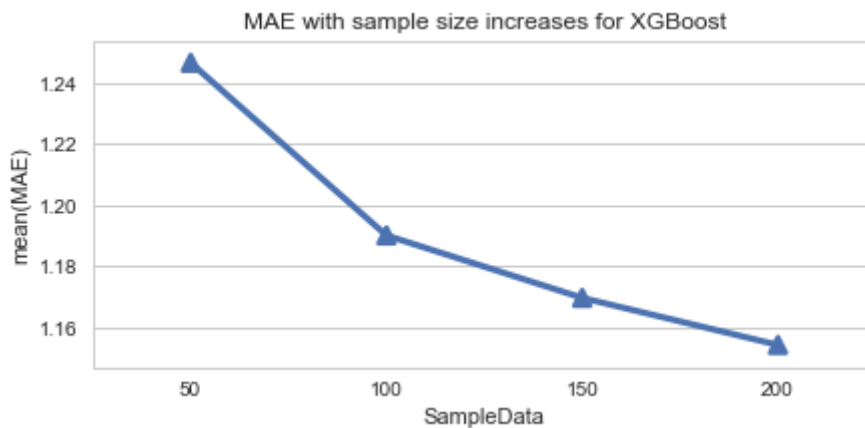


```
In [14]: f, ax = plt.subplots(figsize=(7, 3))
plt.title('RMSE with sample size increases for Gradient Boosted Regression')
sns.pointplot(x=gbt['SampleData'], y=gbt['RMSE'], data=gbt,
              markers=["o"], linestyle=["--"]);
```

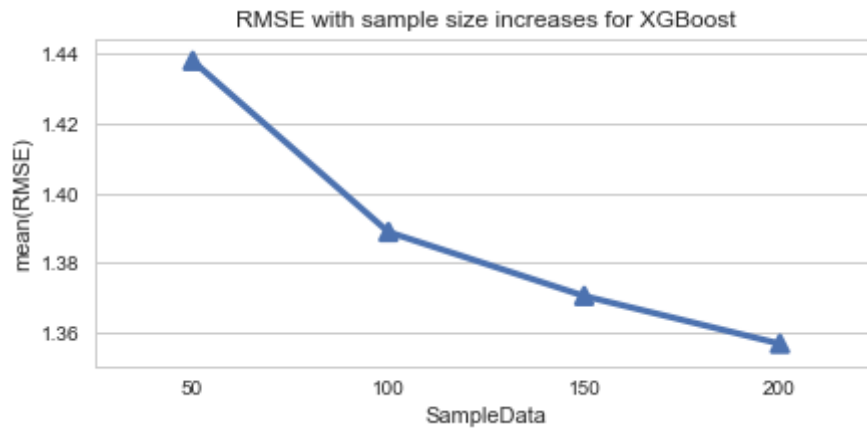


```
In [15]: xgboost = gbt_metric[gbt_metric['Unnamed: 0'] == 'xgboost']
```

```
In [16]: f, ax = plt.subplots(figsize=(7, 3))
plt.title('MAE with sample size increases for XGBoost')
sns.pointplot(x=xgboost['SampleData'], y=xgboost['MAE'], markers=["o"], lines
              markers=["^"], linestyle=["-"]);
```



```
In [17]: f, ax = plt.subplots(figsize=(7, 3))
plt.title('RMSE with sample size increases for XGBoost')
sns.pointplot(x=xgboost['SampleData'], y=xgboost['RMSE'], data=xgboost,
               markers=["o"], linestyle=["--"]);
```



In [ ]: