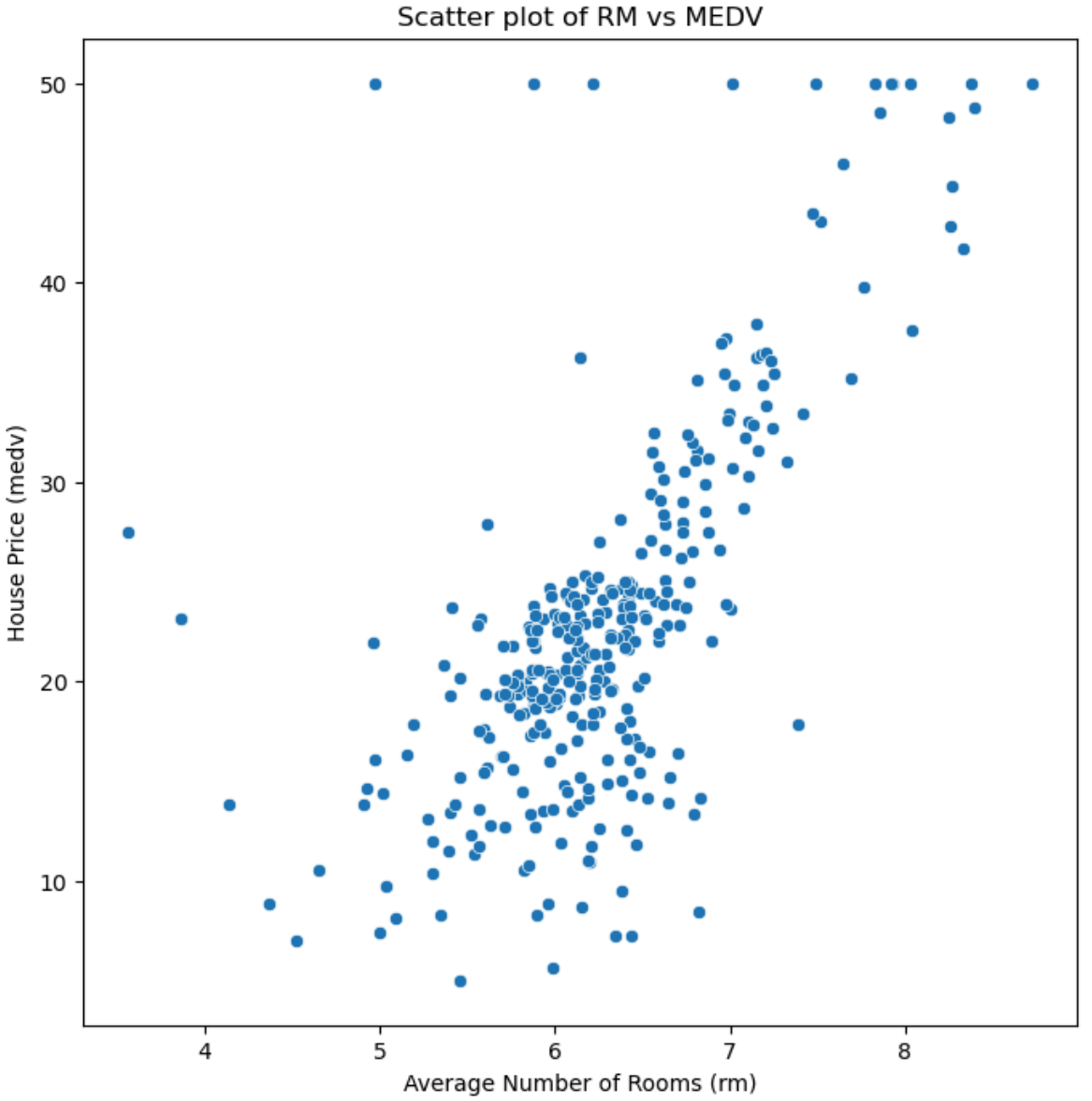
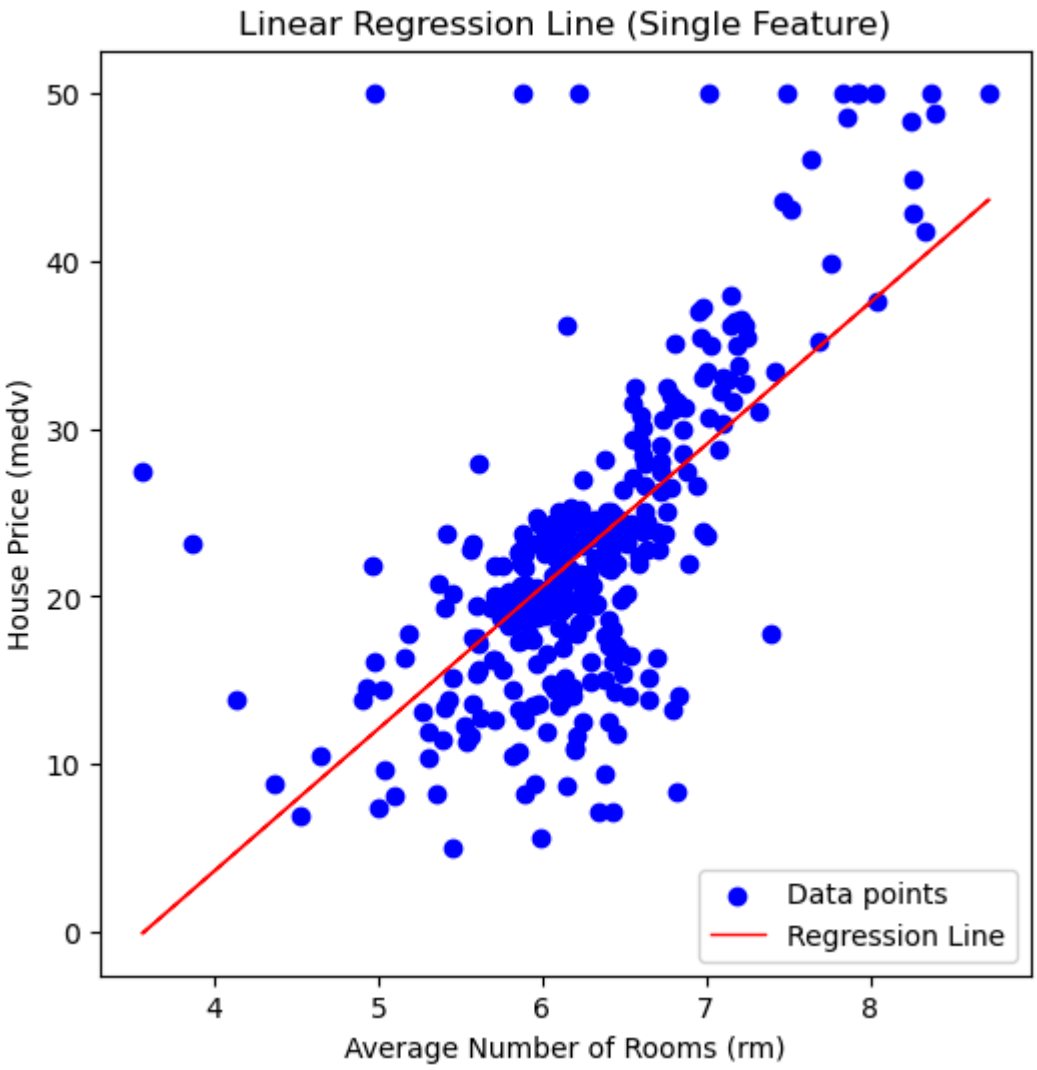
In [ ]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score
```

In [2]:
```
train = pd.read_csv("train.csv")
test = pd.read_csv("test.csv")
submission_example = pd.read_csv("submission.csv")


plt.figure(figsize=(8,8))
sns.scatterplot(x=train["rm"], y=train["medv"])
plt.title("Scatter plot of RM vs MEDV")
plt.xlabel("Average Number of Rooms (rm)")
plt.ylabel("House Price (medv)")
plt.show()
```



In [3]:
```
train.head()
```



In [8]:
```
model.fit(X, y)

test_features = test[["rm"]]
test_predictions = model.predict(test_features)

submission = pd.DataFrame({
    "ID": submission_example["ID"],
    "medv": test_predictions
})

submission.to_csv("submission_single_feature.csv", index=False)
submission.head()
```

Out[8]:

	ID	medv
0	3	31.030718
1	6	24.245969
2	8	21.927472
3	9	17.065818
4	10	20.417753

In [ ]: