# A1

## Q1

| Model | Train RMSE | Test RMSE | Num Parameters |
|---|---|---|---|
| 1 | 1.44834 | 1.424781 | 4 |
| 2 | 1.159677 | 1.11978 | 7 |
| 3 | 0.5070891 | 0.5269903 | 13 |
| 4 | 0.5044712 | 0.5332113 | 18 |

Model 3 is best at making predictions because it has the smallest Test RMSE.

## Q2

| Model | Desc | Train RMSE | Test RMSE | Num Parameters |
|---|---|---|---|---|
| `medv ~ .^2` | Baseline | 2.641685 | 2.813286 | 91 |
| `medv ~ .` | Smaller | 4.675465 | 4.767746 | 13 |
| `medv ~ .^2 + I(black^2)` | Larger | 2.636336 | 2.837968 | 92 |

## Q3

| k | Train RMSE | Test RMSE | Fit |
|---|---|---|---|
| 5 | 1.645974 | 2.156945 | Over |
| 10 | 1.703955 | 2.08401 | Over |
| 15 | 1.790904 | 2.047302 | Under |
| 20 | 1.930387 | 2.055042 | Under |
| 25 | 2.023999 | 2.14426 | Under |
| 30 | 2.284619 | 2.360523 | Under |
| 35 | 2.601655 | 2.672844 | Under |
| 40 | 2.957898 | 2.994395 | Under |
| 45 | 3.270317 | 3.287538 | Under |
| 50 | 3.581351 | 3.566584 | Under |

## Q4

| k | Scaled | Test MSE |
|---|---|---|
| 1 | No | 0.4678402 |
| 1 | Yes | 0.520477 |
| 5 | No | 0.2882769 |
| 5 | Yes | 0.298908 |
| 25 | No | 0.2660158 |
| 25 | Yes | 0.258192 |

## Q5

A KNN model that outperforms the given linear model can be constructed as follows:

```
model = FNN::knn.reg(
  train = scale(train_data[, -1]),
  test = scale(train_data[, -1]),
  y = train_data$y,
  k = 6
)
```

This model uses scaled X data and a `k` value of 6. It has a test RSME of `1.182593` which is significantly lower than the linear model, which had a test RSME of `3.495659`.