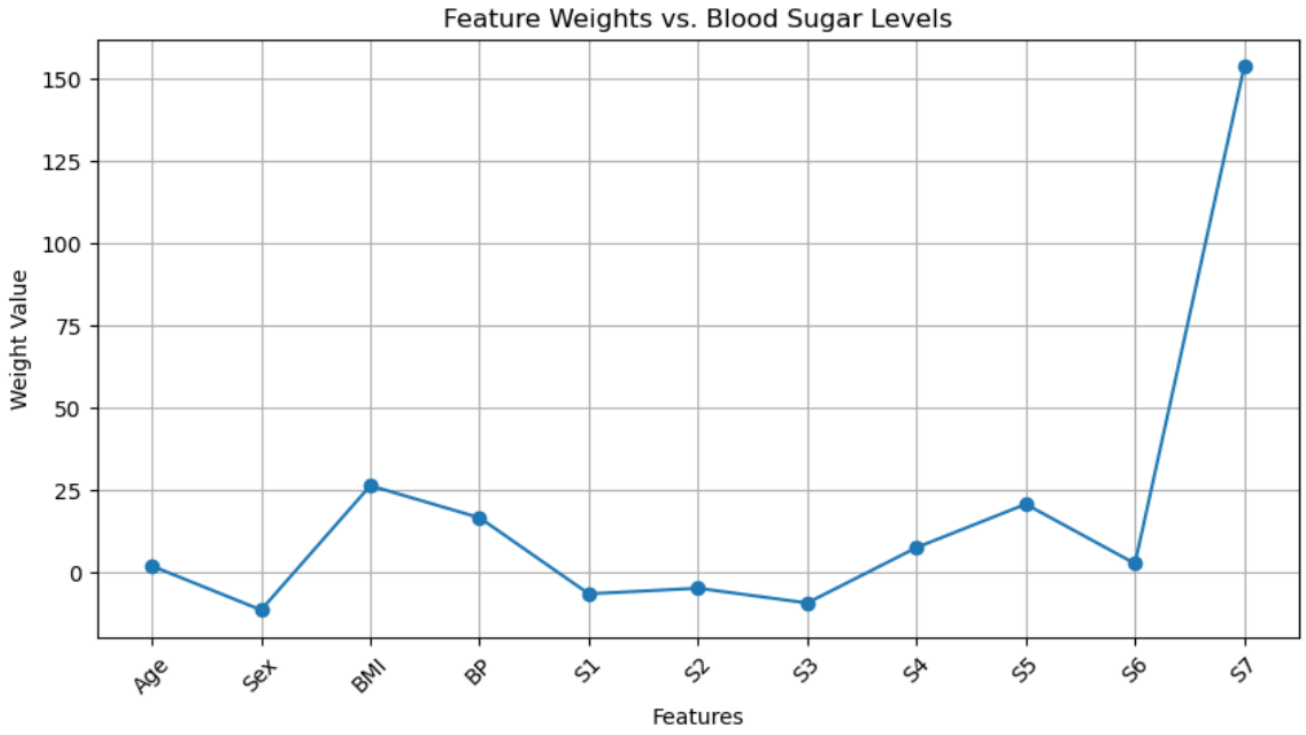
Q5: What conclusion if any can be drawn from the weight values? How does gender and BMI affect blood sugar levels?

Age, Sex, BMI, blood pressure, and some blood serum measurements have a more influence on the diabeti c's blood sugar levels. In blood Serum measurements, S7 has the biggest Impact overall, with a weight of ar ound 153.7 and the highest weight of all.

A positive weight for Age suggests that, as the Age increases the blood sugar level also increases. Blood pre ssure has a positive weight indicating more impact on sugar levels but less than BMI and some blood serum measurements. S1, S2 and S3 have negative weights suggesting that increase in these features will have les s impact on the sugar levels. Similarly, S4, S5, S6 have the opposite effect on the overall prediction compare d to S1, S2 and S3 sugar levels.

As per the weights, the negative weights indicate a difference in predicted blood sugar levels between females and males. So being male is associated with higher blood sugar level compared to female (Considering male as 0 and female as 1). On the other hand, BMI also has significant impact on sugar levels, so if your BMI is more, you likely have high sugar levels.



Model weights:
Age: 1.9484, Sex: -11.4389, BMI: 26.3563, BP: 16.5998, S1: -6.5293, S2: -4.8259, S3: -9.3081, S4: 7.5340, S5: 20.7292, S6: 2.6660, S7: 153.7325

Now estimate the error on the test set. Is the error on the test set comparable to that of the train set? What can be said about the fit of the model? When does a model over/under fit?

The mean Squared error for the training set is around 2894.8 and mean squared error for the test set is around 2884.9, which is close to the training set, with an error percentage around ~0.34%. So as per the learning rate of 0.1 (alpha) the model is balanced and making accurate predictions for a sample.

Overfitting occurs when model performs well on trained data set but poorly on test data set. The model learns the training data closely. This results, when given a new data it predicts poor and inaccurate results.

Underfitting occurs when the model is not capturing the data patterns accurately. It performs poorly on both train data set and test data set resulting in inaccurate predictions.

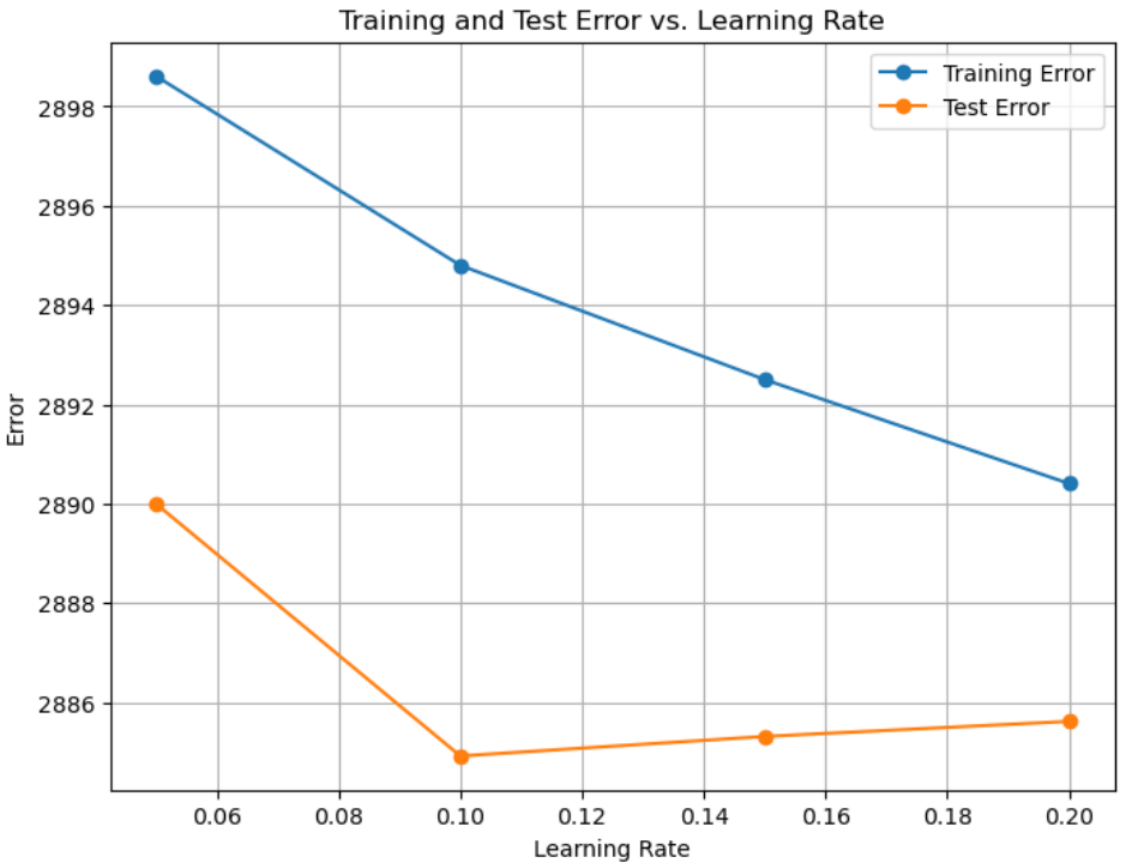
```
Test Set Mean Squared Error: 2884.922607421875
Training Set Mean Squared Error: 2894.800048828125
Error Percentage of training set and test set: 0.3412132527166632
```

Q6: Try the code with a number of learning rates that differ by orders of magnitude and record the error of the training and test sets. What do you observe on the training error? What about the error on the test set?

We considered learning rates ranging from 0.05 to 0.2 and we found that with increase in learning rates, training error is decreasing, indicating the higher the learning rate the model learns from the data more quickly. However, this results in a reduction in training error.

The behaviour of the test set based on the learning rate is not linear. The test error with learning rate between 0.05 and 0.1 has reduced the test error from 2890 to 2884.9. However, when the learning rate is increased from 0.1 to 0.2 the learning rate has increased to 2885.61 indicating the potential for overfitting where the model has good on training data but performs poorly on unseen data.

```
Learning Rate: 0.05, Training Error: 2898.60986328125, Test Error: 2890.0009765625
Learning Rate: 0.1, Training Error: 2894.800048828125, Test Error: 2884.92260742187
Learning Rate: 0.15, Training Error: 2892.500244140625, Test Error: 2885.3154296875
Learning Rate: 0.2, Training Error: 2890.406494140625, Test Error: 2885.619140625
```

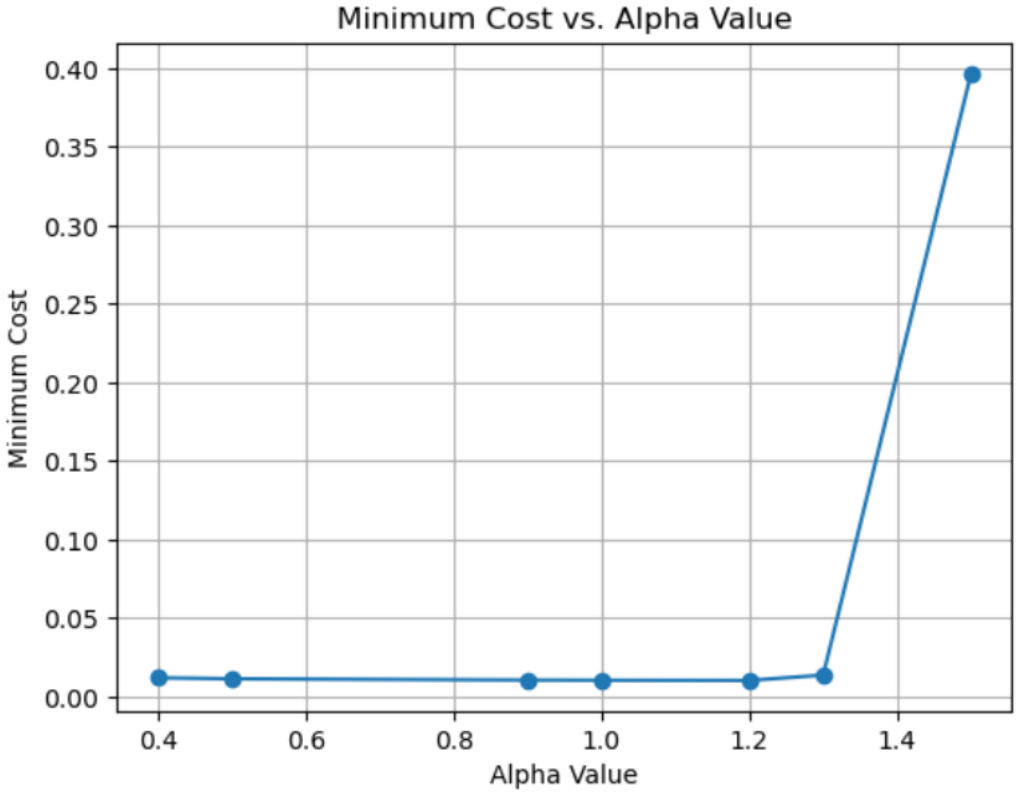


Q8. First of all, find the best value of alpha to use in order to optimize best. Next, experiment with differen t values of λ and see how this affects the shape of the hypothesis.

After testing with different alpha values ranging from .4 to 1.5, the minimum cost remains approximately t he lowest for values till 1.2 around 0.104. Choosing value 1.2 results in the lowest cost and should lead to g ood linear model during the training process.

```
Alpha Value: 0.1 Minimum Cost:0.06648995727300644
Alpha Value: 0.2 Minimum Cost:0.021793313324451447
Alpha Value: 0.3 Minimum Cost:0.014164367690682411
Alpha Value: 0.4 Minimum Cost:0.012157843448221684
Alpha Value: 0.5 Minimum Cost:0.011501406319439411
Alpha Value: 1.2 Minimum Cost:0.010443011298775673
Alpha Value: 1.5 Minimum Cost:0.3964122533798218
```

With the best value for alpha (1.2) we calculated the cost with different values of lambda ranging from .1 to 100 with 100 iterations, but with to the few numbers of data points and ease of fitting due to less noise in the data poi nts, the lambda value doesn't seem to have an huge impact over weight updating.




```
Lambda value: 0.1: Minimum Cost:0.008161894045770168
Lambda value: 0.2: Minimum Cost:0.008161894045770168
Lambda value: 0.5: Minimum Cost:0.008161894045770168
Lambda value: 0.9: Minimum Cost:0.008161894045770168
Lambda value: 1: Minimum Cost:0.008161894045770168
Lambda value: 100: Minimum Cost:0.008161894045770168
```