Ashly Lam, Amanda Poon
4181603610, 9451754209
ashlylam@usc.edu
ajpoon@usc.edu

The program implements linear regression to predict the students' grades using data on the students' age, and years of the parents' education.

Since the database has substantial information, we trimmed it to consider the attributes of the students' grades (G1, G2, G3), age (age), and years of the parents' education (Medu and Fedu). We aimed to predict their final grade (G3). We then split the data into testing and training data.

We trained and scored our model with a high accuracy near 80% indicating that our model was highly accurate.

```
linear = linear_model.LinearRegres
linear.fit(x_train, y_train)
acc = linear.score(x_test, y_test)
print(acc)
```

```
0.7992455709230147
```

To check that our model worked on students, we printed the test data, actual final grade, and model-predicted grade.

```
predictions = linear.predict(x_test)

for x in range(len(predictions)):
    print(predictions[x], x_test[x], y_test[x])
```

```
16.39325821466177 [15 16 16  4  2] 16
7.586509009590373 [ 8  8 17  4  1] 8
11.965720924420621 [14 12 19  2  1] 13
18.30275239285161 [16 18 16  3  3] 18
11.757408372815975 [12 12 18  2  1] 13
7.290640165679686 [ 8  8 17  1  1] 9
4.005378332525179 [ 7  5 19  2  3] 0
7.218898209964385 [ 8  8 18  2  2] 0
12.664490286043042 [10 13 15  3  3] 13
6.1022204761566465 [ 7  7 18  2  2] 0
8.406091902416735 [ 8  9 16  3  3] 10
18.47311729653714 [16 18 15  3  2] 19
11.28002754572088 [14 11 16  4  4] 13
10.595697859831857 [11 11 17  2  3] 10
12.791105679528219 [11 13 15  2  2] 14
14.01630429180502 [13 14 16  4  4] 14
```

Instructions

We ran the code in Google Colab, by running each section of the code. We imported our dataset from https://archive.ics.uci.edu/dataset/320/student+performance.