Question 1.4: Does the line we found fit the data well? Explain.

Type your answer here, replacing this text.

SOLUTION: Nope

Question 1.6: Using your 95% confidence interval, is it likely that the credit has no effect on the September 2005 bill? Justify your answer.

Type your answer here, replacing this text.

SOLUTION: No, the CI does not contain 0.

Question 1.9: Fill in the function `pred_and_plot` below which models `bill_sep05` based on a column `col`, plots the scatterplot and line of best fit, and computes the RMSE of the model. Then choose a column you think might be related to `bill_sep05` and use your `pred_and_plot` function to determine its prediction RMSE and plot the regression line.

Hint: Your code from Question 1.3 may be helpful here...

```
In [31]: def pred_and_plot(col):
             """Performs single variable OLS to predict bill_sep05 based on col"""
             x = defaults[col]                      # SOLUTION
             y = defaults["bill_sep05"]             # SOLUTION

             beta_1 = slope(x, y)                   # SOLUTION
             beta_0 = intercept(x, y)               # SOLUTION

             y_hat = beta_1 * x + beta_0            # SOLUTION

             model_rmse = rmse(y, y_hat)            # SOLUTION


             ### DO NOT EDIT THE REST OF THIS FUNCTION ###
             print("RMSE: {:.5f}".format(rmse(y, y_hat)))

             plt.scatter(x, y, color="tab:blue", alpha=0.3)
             plt.plot(x, y_hat, color="tab:red")
             plt.title("Predict September bill with {}".format(col))

         """ # BEGIN PROMPT
         ### Provide your column name below ###
         pred_and_plot(...)
         """ # END PROMPT
         pred_and_plot("paid_apr05")  # SOLUTION NO PROMPT)
```
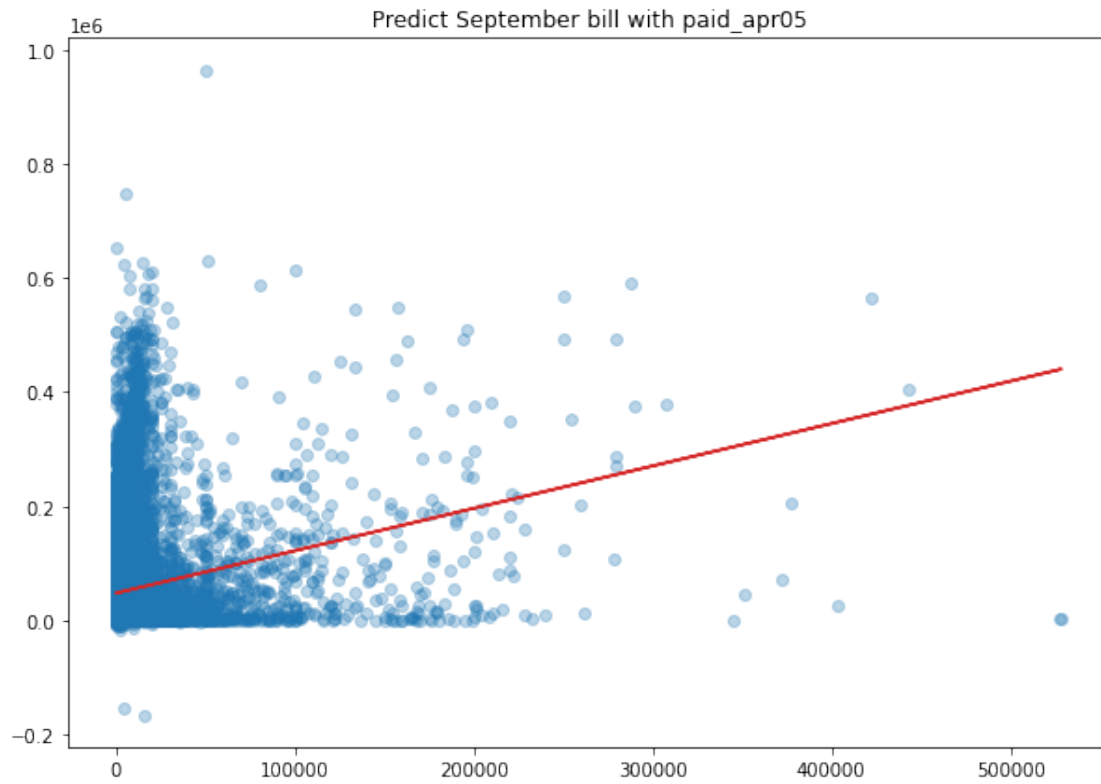
RMSE: 72440.79127

Predict September bill with paid_apr05

Question 1.11: Interpret the value of `default_beta_1`. Basically, what do we expected to happen when `default` changes from 0 to 1? Explain.

Type your answer here, replacing this text.

SOLUTION: We expect the bill to go down by approx \$3,485.

Question 2.2: Which bills are likely good predictors of `bill_sep05`? Justify your response.

Type your answer here, replacing this text.

SOLUTION: August, July, and June. These have CIs that don't contain 0, and their t statistics are high.

Question 2.6: Did the RMSE go up or down in Question 2.7 compared to Question 2.4? Why do you think so?

Type your answer here, replacing this text.

SOLUTION: You will get full points as long as you provide a good reason for why you think your RMSE went up or down.

Question 4.1: Explain one choice you made in selecting features while modeling in Part 3 and why you made it. (Your explanation should take at least a few sentences, and should justify your choice mathematically (i.e. with numerical evidence).)
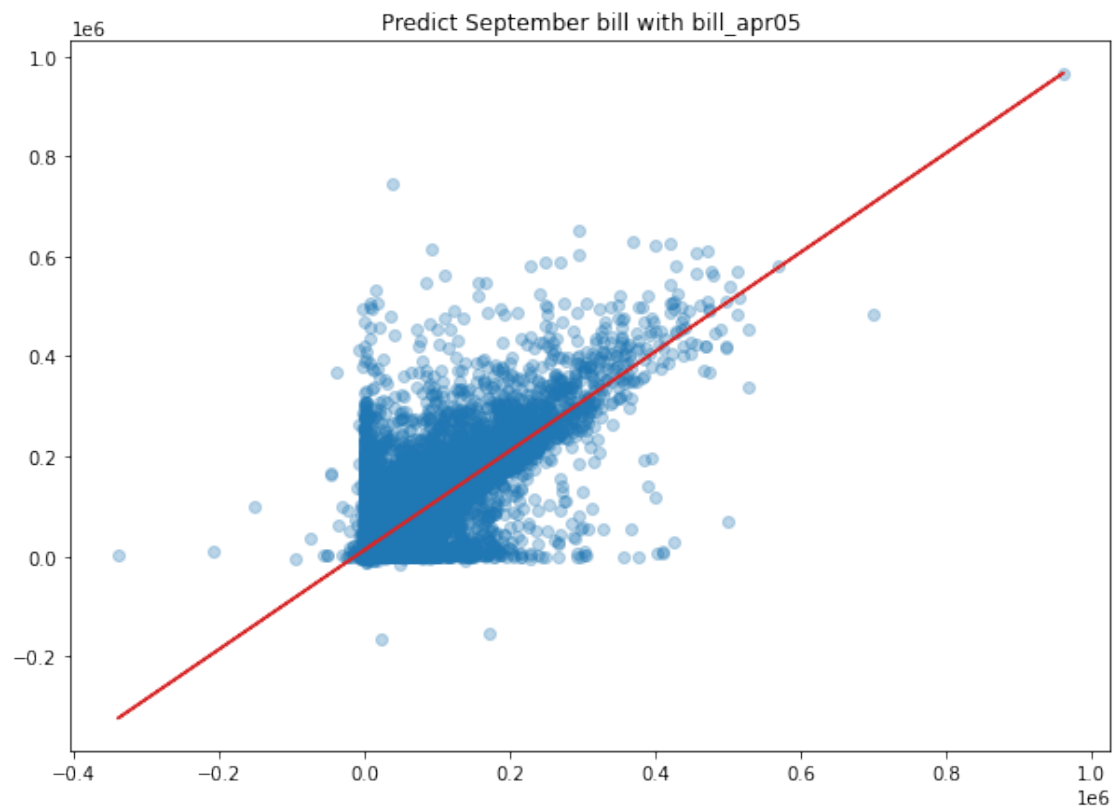
Type your answer here, replacing this text.

SOLUTION: You should describe a choice you made and give mathematical justifications for why you made it. For example, I replace feature A with feature B because A's correlation with y was <a number> but B's was <a number>, and this lowered the RMSE from <a number> to <a number>. Basically, show me the numbers.

Question 4.2: Use your `pred_and_plot` function in the cell below to generate a visualization that helped you choose a feature in Part 3.

```
In [57]: pred_and_plot("bill_apr05") # SOLUTION
```

```
RMSE: 43919.38150
```


Predict September bill with bill_apr05

Question 4.3: Choose a column you regressed on. Report its coefficient, t statistic, and 95% CI. Interpret the coefficient's value. Is the variable likely significant? Explain.

Type your answer here, replacing this text.

SOLUTION: Full points with reporting all values and explanation using t statistic and/or 95% CI.