

Here is some data for Diamonds

- Run this first to set nice print outs

```
options("scipen"=100, "digits"=4)
```

carat	clarity	price
0.71	3	2554
0.72	5	2691
0.33	4	854
1.52	4	9576
1.23	4	9116
0.54	6	1829
0.32	5	828
1.31	3	8767
1.11	2	5648
1.03	2	4283
0.56	5	1131
1.29	2	5601
0.53	5	1577
1.7	5	11257
0.3	4	615

- create a data `df` using the vectors

```
carat<-c(0.71,0.72,0.33,1.52,1.23,0.54,0.32,1.31,1.11,1.03,0.56,1.29,0.53,1.7, 0.3)
clarity<-c(3,5,4,4,4,6,5,3,2,2,5,2,5,5,4)
price<-c(2554,2691,854,9576,9116,1829,828,8767,5648,4283,1131,5601,1577,11257,615)
df<-data.frame(CODE)
```

- print out info for the dataframe using `str`

```
CODE
```

- print out the first few rows in the dataframe use `head`

```
CODE
```

- create scatterplots of the columns `plot(df, upper.panel=NULL, pch=19)`

CODE

- create correlation matrix for `df` using the code `cor(df)`

CODE

- run a linear regression model using `carat` and `clarity` to predict `price`

```
model1<-lm(CODE,data=df)
summary(model1)
```

Test for overall significant of linear relationship:

overall p-value = ????

Clarity:

p-value for clarity = ???

Carat:

p-value for carat = ???

Eliminate Clarity from the model:

- run a new linear regression model using `carat` to predict `price`

```
model2<-lm(CODE,data=df)
summary(model2)
```

Fill in the ??? below with the right value:

Test for overall significant:

overall p = ???

The regression equation is this:

price = ???(carat) + ???

The standard error is:

std error = ???

Make predictions when:

- carat is 1.3
- carat is 0.6
- carat is 1.8

To do this:

- assign a vector `carat` to hold the predictors above (replace CODE below)
- create a dataframe called `predictors` using the `carat` vector
- print out the `predictors` dataframe

```
carat<-CODE  
predictors <- data.frame(carat)  
predictors
```

Now lets do the predictions:

- make predictions using the `predict` function passing in the `model2` and the `predictors` dataframe you just made (pass 2 arguments where you see CODE)
- print out the `predictions`

```
predictions<-predict(CODE)  
predictions
```

Fill in the correct answers below:

- Prediction when `carat` is 1.3: ???
- Prediction when `carat` is 0.6: ???
- Prediction when `carat` is 1.8: ???

