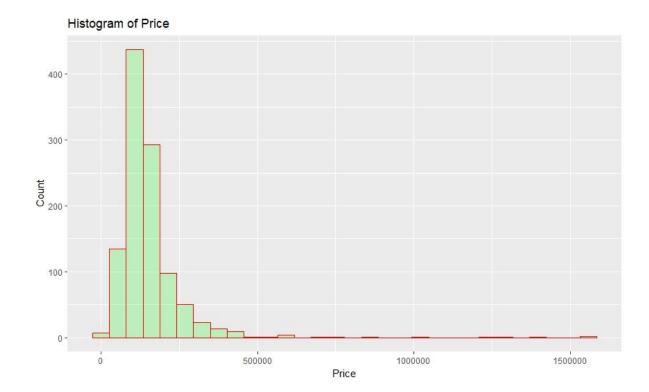
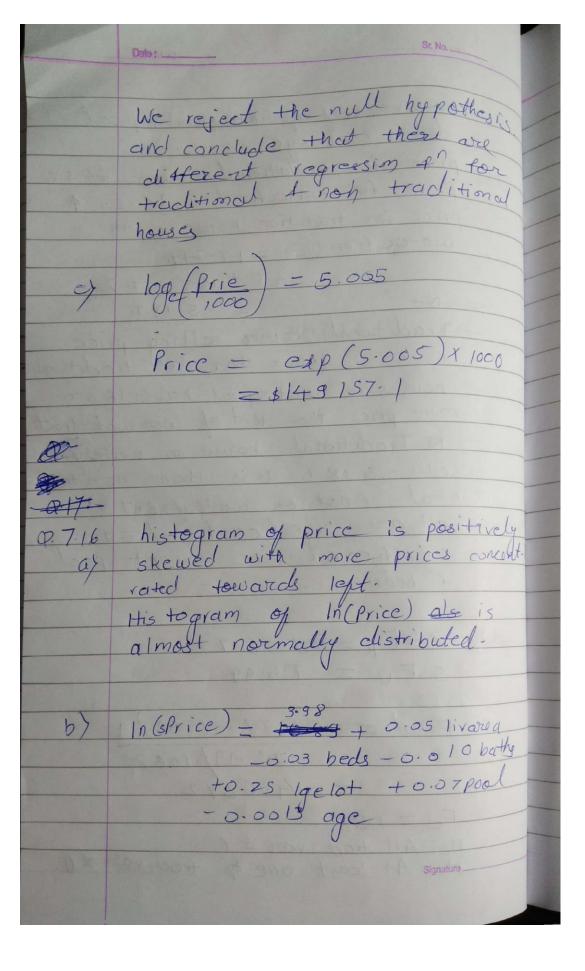


Signature ____



St. No. didn't change. price for tradational house = -4.5%. price for waterfront = 16.54. 1 price for tradition house with water front = -17-214 -4.5 + 16.5 - 17.21 Traditional houses selling price is 4:57. lower than now traditional houses on waterfront have 16:51. more pric that that of non waterfront A Traditional house on waterfront sells 5-18-1, less than nontradiotal 4 not on waterfront all and significant at 51. significance leve except pool. & bedroom. d) SSER= # 78.77 SSE() = 75.99 Fstat = 78-77-75.99)9 75.99/1062 = 4.6272 Fc = 1.88 Ho: All trad vac = 0 Hi: At least one of tradsignature = 10



```
> library(ggplot2)
> ggplot(br2, aes(br2$price)) +
+ geom_histogram(col="red",fill="green", alpha = .2) +
+ labs(title = "Histogram of Price") +
+ labs(x = "Price", y = "Count")
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
> model <- lm(log(price/1000) ~ I(sqft/100) + bedrooms + baths + age + owner + pool + traditional + fireplace +
waterfront, data = br2)
> summary(model)
Call:
Im(formula = log(price/1000) ~ I(sqft/100) + bedrooms + baths +
  age + owner + pool + traditional + fireplace + waterfront,
  data = br2
Residuals:
  Min
          1Q Median
                         3Q
                               Max
-1.13459 -0.12758 0.00656 0.14785 1.06650
Coefficients:
       Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.9808326 0.0458947 86.738 < 2e-16 ***
I(sqft/100) 0.0299011 0.0014059 21.269 < 2e-16 ***
bedrooms -0.0315060 0.0166109 -1.897 0.058135.
         0.1901190 0.0205579 9.248 < 2e-16 ***
baths
        -0.0062145 0.0005179 -11.999 < 2e-16 ***
age
          0.0674654 0.0177460 3.802 0.000152 ***
owner
pool
        -0.0042748 0.0315812 -0.135 0.892353
traditional -0.0560926 0.0170267 -3.294 0.001019 **
fireplace 0.0842748 0.0190150 4.432 1.03e-05 ***
waterfront 0.1099700 0.0333550 3.297 0.001010 **
```

```
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '' 1
Residual standard error: 0.27 on 1070 degrees of freedom
Multiple R-squared: 0.7373,
                             Adjusted R-squared: 0.7351
F-statistic: 333.7 on 9 and 1070 DF, p-value: < 2.2e-16
> 100*(exp(.1099)-1)
[1] 11.61664
> model <- lm(log(price/1000) ~ I(sqft/100) + bedrooms + baths + age + owner + pool + traditional + fireplace +
        waterfront + I(waterfront * traditional), data = br2)
> summary(model)
Call:
Im(formula = log(price/1000) ~ I(sqft/100) + bedrooms + baths +
  age + owner + pool + traditional + fireplace + waterfront +
  I(waterfront * traditional), data = br2)
Residuals:
  Min
         1Q Median
                        3Q
                              Max
-1.13891 -0.12591 0.00672 0.14693 1.05734
Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)
                   3.9711130 0.0459460 86.430 < 2e-16 ***
I(sqft/100)
                   0.0300308 0.0014034 21.399 < 2e-16 ***
bedrooms
                   -0.0313330 0.0165702 -1.891 0.05890.
baths
                 0.1882577  0.0205208  9.174  < 2e-16 ***
                age
                  0.0683701 0.0177061 3.861 0.00012 ***
owner
                -0.0023939 0.0315125 -0.076 0.93946
pool
                  -0.0449127 0.0175612 -2.557 0.01068 *
traditional
```

```
fireplace
               waterfront
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.2693 on 1069 degrees of freedom
Multiple R-squared: 0.7389,
                         Adjusted R-squared: 0.7364
F-statistic: 302.5 on 10 and 1069 DF, p-value: < 2.2e-16
> Resstricted <- Im(log(price/1000) ~ I(sqft/100) + bedrooms + baths + age + owner + pool + fireplace +
       waterfront, data = br2)
> summary(Resstricted)
Call:
Im(formula = log(price/1000) ~ I(sqft/100) + bedrooms + baths +
 age + owner + pool + fireplace + waterfront, data = br2)
Residuals:
        1Q Median
  Min
                     3Q
                         Max
-1.15673 -0.12355 -0.00287 0.14356 1.03816
Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.9701078 0.0459892 86.327 < 2e-16 ***
I(sqft/100) 0.0301592 0.0014101 21.387 < 2e-16 ***
bedrooms -0.0405182 0.0164592 -2.462 0.013984 *
       0.1894469 0.0206512 9.174 < 2e-16 ***
baths
      age
        0.0650077 0.0178117 3.650 0.000275 ***
owner
       0.0007741 0.0316887 0.024 0.980516
pool
```

fireplace 0.0911987 0.0189852 4.804 1.78e-06 ***

```
waterfront 0.1225762 0.0332869 3.682 0.000243 ***
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.2712 on 1071 degrees of freedom
Multiple R-squared: 0.7347, Adjusted R-squared: 0.7327
F-statistic: 370.7 on 8 and 1071 DF, p-value: < 2.2e-16
> anova(Resstricted)
Analysis of Variance Table
Response: log(price/1000)
       Df Sum Sq Mean Sq F value Pr(>F)
I(sqft/100) 1 185.472 185.472 2521.7197 < 2.2e-16 ***
             1 0.010 0.010 0.1387 0.7096829
bedrooms
          1 15.622 15.622 212.3958 < 2.2e-16 ***
baths
         1 12.521 12.521 170.2406 < 2.2e-16 ***
age
          1 1.749 1.749 23.7851 1.240e-06 ***
owner
          1 0.002 0.002 0.0286 0.8658445
pool
fireplace 1 1.726 1.726 23.4737 1.453e-06 ***
waterfront 1 0.997 0.997 13.5602 0.0002426 ***
Residuals 1071 78.772 0.074
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> Unristricted <- Im(log(price/1000) ~ I(sqft/100) + bedrooms + baths + age + owner + pool + traditional + fireplace +
         waterfront + I((sqft/100) * traditional) + I(bedrooms * traditional)+
          I(baths * traditional) + I(age * traditional) + I(owner * traditional) + I(pool * traditional)+
          I(fireplace * traditional) + I(waterfront * traditional), data = br2)
> summary(Unristricted)
```

Call:

```
Im(formula = log(price/1000) ~ I(sqft/100) + bedrooms + baths +
    age + owner + pool + traditional + fireplace + waterfront +
    I((sqft/100) * traditional) + I(bedrooms * traditional) +
    I(baths * traditional) + I(age * traditional) + I(owner *
    traditional) + I(pool * traditional) + I(fireplace * traditional) +
    I(waterfront * traditional), data = br2)
```

Residuals:

Min 1Q Median 3Q Max
-1.1376 -0.1248 0.0045 0.1462 1.0578

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 4.0672688 0.0576957 70.495 < 2e-16 ***

I(sqft/100) 0.0324010 0.0018412 17.598 < 2e-16 ***

bedrooms -0.0713737 0.0236571 -3.017 0.002614 **

baths 0.1831139 0.0288594 6.345 3.29e-10 ***

age -0.0054674 0.0007292 -7.498 1.37e-13 ***

owner 0.0388479 0.0258967 1.500 0.133884

pool 0.0021253 0.0419397 0.051 0.959594

traditional -0.3350839 0.0944926 -3.546 0.000408 ***

fireplace 0.0578017 0.0296703 1.948 0.051662.

waterfront 0.1729789 0.0406915 4.251 2.32e-05 ***

I((sqft/100) * traditional) -0.0052974 0.0028196 -1.879 0.060549.

I(bedrooms * traditional) 0.0989064 0.0335594 2.947 0.003277 **

I(baths * traditional) 0.0310767 0.0412135 0.754 0.450991

I(age * traditional) -0.0012916 0.0010325 -1.251 0.211211

I(owner * traditional) 0.0586870 0.0353000 1.663 0.096703.

I(pool * traditional) -0.0237596 0.0630941 -0.377 0.706566

I(fireplace * traditional) 0.0650471 0.0386865 1.681 0.092982.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2672 on 1062 degrees of freedom

Multiple R-squared: 0.7447, Adjusted R-squared: 0.7406

F-statistic: 182.2 on 17 and 1062 DF, p-value: < 2.2e-16

> anova(Unristricted)

Analysis of Variance Table

Response: log(price/1000)

Df Sum Sq Mean Sq F value Pr(>F)

I(sqft/100) 1 185.472 185.472 2598.5844 < 2.2e-16 ***

bedrooms 1 0.010 0.010 0.1429 0.7054960

baths 1 15.622 15.622 218.8698 < 2.2e-16 ***

age 1 12.521 12.521 175.4297 < 2.2e-16 ***

owner 1 1.749 1.749 24.5101 8.593e-07 ***

pool 1 0.002 0.002 0.0294 0.8638352

traditional 1 1.295 1.295 18.1388 2.236e-05 ***

fireplace 1 1.428 1.428 20.0066 8.548e-06 ***

waterfront 1 0.792 0.792 11.0991 0.0008934 ***

I((sqft/100) * traditional) 1 0.061 0.061 0.8608 0.3537384

I(bedrooms * traditional) 1 0.778 0.778 10.9046 0.0009912 ***

I(baths * traditional) 1 0.106 0.106 1.4898 0.2225246

I(age * traditional) 1 0.093 0.093 1.2968 0.2550634

I(owner * traditional) 1 0.324 0.324 4.5370 0.0333982 *

I(pool * traditional) 1 0.017 0.017 0.2353 0.6276871

I(fireplace * traditional) 1 0.197 0.197 2.7545 0.0972747.

I(waterfront * traditional) 1 0.606 0.606 8.4846 0.0036566 **

Residuals 1062 75.799 0.071

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '' 1

> qf(.95,9,1062)

[1] 1.88868

> d = data.frame(sqft = 2500, bedrooms = 3, baths = 2, age = 20, owner = 1, pool = 0, traditional = 1, fireplace =1, waterfront =0)

>

> predict(Unristricted, d)

1

5.005958

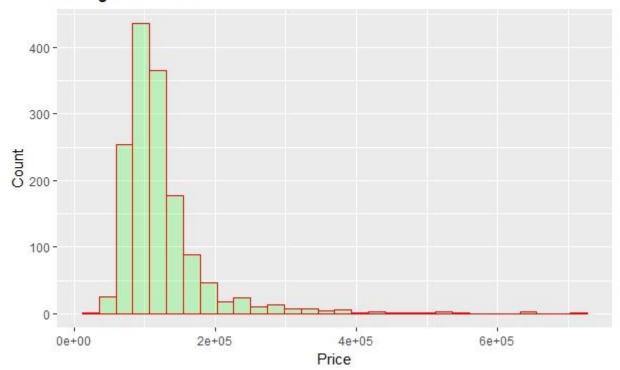
>

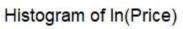
> exp(5.005)*1000

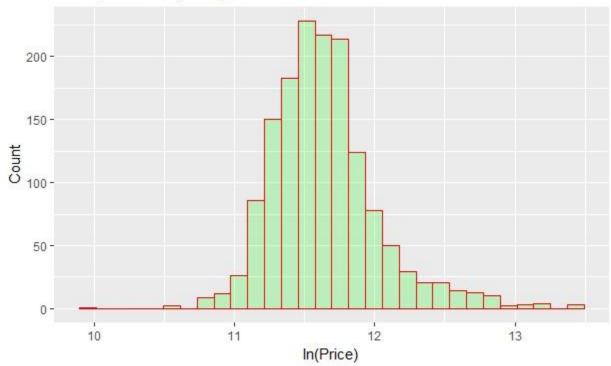
[1] 149157.1

Q 7.16

Histogram of Price







All coefficients except both is 8-tastically significant at 51. 11 vovela house so size grater than 0.5 acres and presence of pool increases the 15 house price. Age, Aumb bedrooms & bothrooms decrease house prices. price of nouses with lot size grater than 0.5 acres is = 100 × (exp(0.2530)-1) = 28.787. grater than that of houses having lot size less than dy The variable is significant at 51. significance revel. houses with larger lotsize than 0.5 acres have reduced selling price by \$ 1.61. It could be the fact that houses having larger lot sixe have larger living asuci. houses Increasing in living and for houses with not size less than 0.5 acre by 100 sq. ft will increase price by 5.8%.

Sr. No. Data: Addition of voriate interaction vou able increased coep of lotsize vou able obcamatically ey Fstat = SSEp-SSEpy J SSEU = 72.063 -65.47/G 65.471 = 1.0986 0.0439 = 24.968 Feritical = 2.10 Ho: toep of lot a variables = 0 we reject the null hypothesis and conclude that ego pricing functi of houses with large lot size & small lot size is not same Signature ____

```
> # -----Question 7.16
> ggplot(stckton4, aes(stckton4$sprice)) +
+ geom_histogram(col="red",fill="green", alpha = .2) +
+ labs(title = "Histogram of Price") +
+ labs(x = "Price", y = "Count")
'stat bin()' using 'bins = 30'. Pick better value with 'binwidth'.
> #With log price
> ggplot(stckton4, aes(log(sprice))) +
+ geom_histogram(col="red",fill="green", alpha = .2) +
+ labs(title = "Histogram of In(Price)") +
+ labs(x = "ln(Price)", y = "Count")
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
> model <- lm(log(sprice/1000) ~ livarea + beds + baths + lgelot + pool + age, data = stckton4)
> summary(model)
Call:
Im(formula = log(sprice/1000) ~ livarea + beds + baths + lgelot +
  pool + age, data = stckton4)
Residuals:
          1Q Median
                              Max
  Min
                        3Q
-1.29751 -0.11979 -0.00427 0.12671 2.00684
Coefficients:
       Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.9859688 0.0373406 106.746 < 2e-16 ***
         livarea
        -0.0382209 0.0113593 -3.365 0.000786 ***
beds
         -0.0102729 0.0165268 -0.622 0.534309
baths
         0.2530908 0.0255382 9.910 < 2e-16 ***
lgelot
         0.0786611 0.0230548 3.412 0.000662 ***
pool
```

```
age
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.2128 on 1493 degrees of freedom
Multiple R-squared: 0.6884,
                           Adjusted R-squared: 0.6871
F-statistic: 549.6 on 6 and 1493 DF, p-value: < 2.2e-16
>
> \exp(.2530)
[1] 1.287883
> # Interaction variable
> unristricted <- model <- Im(log(sprice/1000) ~ livarea + beds + baths + Igelot + pool + age + I(livarea * Igelot), data =
stckton4)
> summary(unristricted)
Call:
Im(formula = log(sprice/1000) ~ livarea + beds + baths + lgelot +
  pool + age + I(livarea * Igelot), data = stckton4)
Residuals:
         1Q Median
  Min
                      3Q
                            Max
-1.17288 -0.12284 -0.00263 0.12812 2.02143
Coefficients:
          Estimate Std. Error t value Pr(>|t|)
(Intercept)
              3.964941  0.037033  107.064  < 2e-16 ***
            livarea
           beds
baths
            -0.020062 0.016398 -1.223 0.221356
```

lgelot

```
0.085349  0.022795  3.744  0.000188 ***
pool
          age
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.2102 on 1492 degrees of freedom
Multiple R-squared: 0.6962, Adjusted R-squared: 0.6948
F-statistic: 488.5 on 7 and 1492 DF, p-value: < 2.2e-16
> # -----Chow Test
> Restricted <- Im(log(sprice/1000) ~ livarea + beds + baths + pool + age, data = stckton4)
> summary(Restricted)
Call:
Im(formula = log(sprice/1000) ~ livarea + beds + baths + pool +
 age, data = stckton4)
Residuals:
        1Q Median
  Min
                      3Q
                           Max
-1.16849 -0.13118 -0.01003 0.12675 2.00675
Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.9794107 0.0385303 103.280 < 2e-16 ***
livarea 0.0606975 0.0016157 37.567 < 2e-16 ***
       beds
       -0.0262415 0.0169748 -1.546 0.1223
baths
       0.0989178 0.0236994 4.174 3.17e-05 ***
pool
       -0.0007805 0.0004716 -1.655 0.0981.
age
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.2196 on 1494 degrees of freedom
Multiple R-squared: 0.6679, Adjusted R-squared: 0.6667
F-statistic: 600.8 on 5 and 1494 DF, p-value: < 2.2e-16
> anova(Restricted)
Analysis of Variance Table
Response: log(sprice/1000)
      Df Sum Sq Mean Sq F value Pr(>F)
         1 142.469 142.469 2953.6328 < 2.2e-16 ***
livarea
         1 1.445 1.445 29.9564 5.175e-08 ***
beds
         1 0.056 0.056 1.1608 0.28147
baths
         1 0.799 0.799 16.5546 4.974e-05 ***
pool
         1 0.132 0.132 2.7389 0.09814.
age
Residuals 1494 72.063 0.048
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
># ----
> Unrestricted <- Im(log(sprice/1000) ~ livarea + beds + baths + pool + age + lgelot + I(lgelot * livarea) +
             I(Igelot * beds ) + I(Igelot * baths)+ I(Igelot * pool) + I(Igelot * age), data = stckton4)
> summary(Unrestricted)
Call:
Im(formula = log(sprice/1000) ~ livarea + beds + baths + pool +
  age + Igelot + I(Igelot * livarea) + I(Igelot * beds) + I(Igelot *
  baths) + I(Igelot * pool) + I(Igelot * age), data = stckton4)
Residuals:
  Min
          1Q Median
                          3Q
                                Max
```

-1.09828 -0.12100 -0.00141 0.12783 2.02787

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) livarea beds -0.033442 0.017394 -1.923 0.054714. baths 0.069685 0.025131 2.773 0.005627 ** pool age lgelot I(lgelot * beds) 0.043412 0.037391 1.161 0.245819 I(lgelot * baths) 0.116104 0.051893 2.237 0.025409 * I(lgelot * pool) 0.056183 0.060423 0.930 0.352616 I(lgelot * age) -0.000219 0.001447 -0.151 0.879738

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2098 on 1488 degrees of freedom

Multiple R-squared: 0.6982, Adjusted R-squared: 0.696

F-statistic: 313 on 11 and 1488 DF, p-value: < 2.2e-16

> anova(Unrestricted)

Analysis of Variance Table

Response: log(sprice/1000)

Df Sum Sq Mean Sq F value Pr(>F)

livarea 1 142.469 142.469 3237.9682 < 2.2e-16 ***

beds 1 1.445 1.445 32.8402 1.209e-08 ***

baths 1 0.056 0.056 1.2726 0.25947

pool 1 0.799 0.799 18.1483 2.172e-05 ***

age 1 0.132 0.132 3.0025 0.08334.

lgelot 1 4.448 4.448 101.0903 < 2.2e-16 ***

```
I(lgelot * livarea) 1 1.708 1.708 38.8091 6.070e-10 ***

I(lgelot * beds) 1 0.117 0.117 2.6499 0.10377

I(lgelot * baths) 1 0.282 0.282 6.4049 0.01148 *

I(lgelot * pool) 1 0.037 0.037 0.8448 0.35818

I(lgelot * age) 1 0.001 0.001 0.0229 0.87974

Residuals 1488 65.471 0.044

---

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '' 1

> qf(.95,6,1488)

[1] 2.104665
```