model1 =lm(SALES ~ POP)

# lm is for linear model(Y ~ X)

summary(model1)

Call:

lm(formula = SALES ~ POP)

Residuals:

Min 1Q Median 3Q Max

-6046.7 -1460.9 -670.5 485.6 18229.5

Coefficients:

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

(Intercept) 469.70360 702.90619 0.668 0.507

POP 0.64709 0.04881 13.258 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3792 on 48 degrees of freedom

Multiple R-squared: 0.7855, Adjusted R-squared: 0.781

F-statistic: 175.8 on 1 and 48 DF, p-value: < 2.2e-16

# Equation: SALES = 469.7036 + .64709(POP)

# R-sqr =.7855, 78.55% of the variation in SALES IS BEING explained

# by POP THIS is pretty high,which is really good

#

# R-sqr could be low yet good to go

# s= 3792,this is hard to judge without another model to compare

#but this seems high which is bad

plot(POP,SALES)

abline(model1)

lm(SALES ~ POP)

Call:

lm(formula = SALES ~ POP)

Coefficients:

(Intercept) POP

469.7036 0.6471

Above is a short description

Error: unexpected symbol in "Above is"

dim(lotto)

[1] 50 10

# dimension of the lotto dataset,50 rows (observations) and 10 columns(variables)

anova(model1)

Analysis of Variance Table

Response: SALES

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

POP 1 2527165015 2527165015 175.77 < 2.2e-16 \*\*\*

Residuals 48 690116755 14377432

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

# SSE = 690116755,SSR = 2527165015, we r not usually interested in this

# output

summary(model1)

Call:

lm(formula = SALES ~ POP)

Residuals:

Min 1Q Median 3Q Max

-6046.7 -1460.9 -670.5 485.6 18229.5

Coefficients:

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

(Intercept) **469.70360** 702.90619 0.668 0.507

POP **0.64709** 0.04881 13.258 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3792 on 48 degrees of freedom

Multiple R-squared: 0.7855, Adjusted R-squared: 0.781

F-statistic: 175.8 on 1 and 48 DF, p-value: < 2.2e-16

# IS the X that we are using a significant predictor of the Y

#IN THE above -----.04881 is the std.error of .64709

# df n-(k+1)= 50-(1+1)=48

qt(.025, 48)

[1] -2.010635

# quantile t score (area to the left of the t, Df)

df(.975, 48)

Error in df(0.975, 48) : argument "df2" is missing, with no default

qt(.975, 48)

[1] 2.010635

#since out t test stat + 13.258> t critical value=2.0106,we reject

# Ha(not Ho):Beta doesn't 0.this leads us to believe that POP is a significant

# predictor of SALES

# wE CAN ALSO make the same decision with the p-value=<2e-16 =0.

#since the P-value <alpha =0.05,u also would reject Ho

#and belive that ur x is a significant predictor of y

# the t score evaluate each X individually and if they help predict Y

# the F score and its p-value evaluates the model as a whole,how all the

# X's together are doing at predictinf the Y

# SALES =469.703 + .647POP

# interpret bo: The value of Y when x= 0.If there r 0 ppl in the

# population,there d be $ 469.7 in lottory sales

**confint(model1, level=0.95)**

**Find the confidence interval of model1**

# Ho:Beta doesn't 0.this leads us to believe that POP is a significant