**Problème 1**



Dans ce cas, les clients qui achètent le produit ne sont pas influencé par les clients, alors est vrai. Les 2 variables sont indépendantes.

1. Si

On sait que N est le nombre de clients qui vont se présenter durant la saison. En moyenne, il y a donc 24 clients qui vont se présenter, car

En multipliant le nombre de client en moyenne par la probabilité que le client va acheter

Donc, pour maximiser le profit moyen il faut que n=13 unités

**Problème 2**





Dans l’histogramme, on voit la distribution de la variable consommation de l’échantillon ne prend pas vraiment la forme d’une cloche comme la distribution normale.

La ligne du milieu foncée est la médiane qui est de 24, le minimum est la ligne la plus à gauche qui a une valeur de 11, le maximum est la ligne la plus à droite qui a une valeur 44.6. Le premier quartile Q1=17 donc du minimum au premier quartile se situe 25% des données, le deuxième quartile Q2=24 donc du minimum au deuxième quartile se situe 50% des données et le troisième quartile Q3=30.625 donc du minimum au troisième quartile se situe 75% des données.



Graphical user interface, text, application

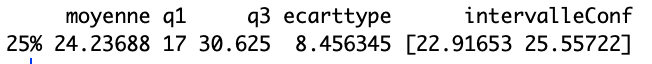
Description automatically generated

Test d’hypothèse effectué par R:

**H0: la distribution suit une loi normale**

**H1: la distribution ne suit pas une loi normale**

En posant α = 0,05, L’hypothèse que la distribution suit une loi normale H0 n’est pas plausible, car la p-value = qui est beaucoup plus petit que la valeur de alpha donc on rejette H0.



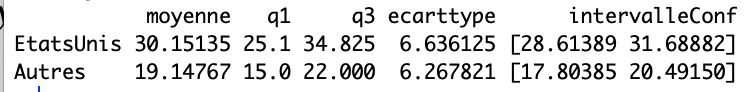
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Moyenne | Premier quartile (Q1) | Deuxième quartile (Q2) | Troisième quartile (Q3) | Écart-type | Intervalle de confiance |
| 24.23688 | 17 | 24 | 30.625 | 8.456345 | [22.91653,25.55722 |





Text

Description automatically generated



Pour vérifier si la variance des 2 échantillons sont les mêmes

**H0 : La variance de la consommation pour les États-Unis et pour les autres pays sont égales**

**H1 : La variance de la consommation pour les États-Unis et pour les autres pays ne sont pas égales**

n1 et n2 sont grands n1 >40, n2 >40, pour ce test, on veut un niveau de confiance de 95%

A picture containing logo

Description automatically generated

Avec les données : ne=74, na=86, Se=6.636125, Sa=6.267821

Table

Description automatically generatedEn utilisant la formule :

A picture containing application

Description automatically generated

Donc

Pour calculer , on utilise la formule :

A picture containing text

Description automatically generated

A picture containing graphical user interface

Description automatically generated

Donc

Alors, on rejette H0, les 2 variances ne sont pas égales.

Pour vérifier si la moyenne des 2 échantillons sont les mêmes

**H0 : La moyenne de la consommation pour les États-Unis et pour les autres pays sont égales**

**H1 : La moyenne de la consommation pour les États-Unis et pour les autres pays ne sont pas égales**

σ12 et σ2 sont inconnues et n1, n2 sont grands (n1 ≥30, n2 ≥30)

A picture containing logo

Description automatically generated

Avec les données : ne=74, na=86, Se=6.636125, Se=6.267821, x̅e=30.15135, x̅a=19.14767

Pour calculer , on utilise la formule :

Diagram

Description automatically generated

A picture containing graphical user interface

Description automatically generated

Donc =10.72869

Alors, on rejette H0, les 2 moyennes ne sont pas égales.

1. Text, letter

   Description automatically generated**Modèle 1 :**

Text

Description automatically generated

L’équation est donc : pour le modèle 1

**Signification** :

En utilisant le **test de** **Fisher**

**H0:**

**H1:**

Comme le seuil critique est de 5% et comme on obtient une p-value < , Donc on rejette H0, a une influence sur . Quand augmente, diminue.

Diagram, schematic

Description automatically generated

On peut voir dans le graph **Residuals vs Fitted**, les points ne sont pas distribués aléatoirement le long de la ligne 0. Cela veut dire qu’il y a Hétéroscédasticité. Dans le graph **Scale-Location** on voit que la distribution des points n’est pas n’est pas dispersée également sur la ligne rouge, il y a des places plus concentrés en point, donc la variance des résidus n’est pas constante. Dans le graph **Normal Q-Q**, les points sont majoritairement sur la droite, donc on peut assumer que les résidus sont distribués selon une normale. Dans le graph **Residuals vs Leverage**, on voit qu’il y a un des points atypiques qui peut avoir une grande influence sur les résultats, le point qui dépasse la ligne **Cook’s distance**.

**Modèle 2 :**

Après la linéarisation :

Text

Description automatically generated

Text, letter

Description automatically generated

L’équation est donc : ln( pour le modèle 2

**Signification** :

En utilisant le **test de** **Fisher**

**H0:**

**H1:**

Comme le seuil critique est de 5% et comme on obtient une p-value < , Donc on rejette H0, a une influence sur . Quand augmente, diminue.

Diagram

Description automatically generated

On peut voir dans le graph **Residuals vs Fitted**, les points sont distribués assez aléatoirement autour de la ligne 0. Cela veut dire qu’assumer que la relation est linéaire est raisonnable. Dans le graph **Scale-Location** on voit que la distribution des points est dispersée également sur la ligne rouge, donc la variance des résidus est constante. Dans le graph **Normal Q-Q**, les points sont majoritairement sur la droite, il y a peu de points aberrants, donc on peut assumer que les résidus sont distribués selon une normale. Dans le graph **Residuals vs Leverage**, on voit qu’il y a un point en particulier qui dépasse la ligne **Cook’s distance**, donc ce point aurait une grande influence sur les résultats.

**Modèle 3 :**

Après la linéarisation :

Text, letter

Description automatically generated

Text, letter

Description automatically generated

L’équation est donc : pour le modèle 3

**Signification** :

En utilisant le **test de** **Fisher**

**H0:**

**H1:**

Comme le seuil critique est de 5% et comme on obtient une p-value < , , Donc on rejette H0, a une influence sur . Quand augmente, diminue.

Diagram, schematic

Description automatically generated

On peut voir dans le graph **Residuals vs Fitted**, les points ne sont pas distribués aléatoirement le long de la ligne 0. Cela veut dire qu’il y a Hétéroscédasticité. Dans le graph **Scale-Location** on voit que la distribution des points n’est pas n’est pas dispersée également sur la ligne rouge plus de points vers la droite, donc la variance des résidus n’est pas constante. Dans le graph **Normal Q-Q**, les points sont majoritairement sur la droite, il y a peu de points aberrants, donc on peut assumer que les résidus sont distribués selon une normale. Dans le graph **Residuals vs Leverage**, on voit qu’il y a un point en particulier qui dépasse la ligne **Cook’s distance**, donc ce point aurait une grande influence sur les résultats.

**Modèle 4 :** Y =β0 +β1 X2 +ε;

**Text, letter

Description automatically generated**

Text, letter

Description automatically generated

L’équation est donc :

**Signification** :

En utilisant le **test de** **Fisher**

**H0:**

**H1:**

Comme le seuil critique est de 5% et comme on obtient une p-value < , alors , Donc on rejette H0, a une influence sur . Quand augmente diminue.

Diagram

Description automatically generated

On peut voir dans le graph **Residuals vs Fitted**, les points sont distribués aléatoirement le long de la ligne 0. Cela veut dire qu’assumer que la relation est linéaire est raisonnable. Dans le graph **Scale-Location** on voit que la distribution des points n’est pas dispersée également sur la ligne rouge, donc la variance des résidus n’est pas constante. Dans le graph **Normal Q-Q**, les points sont majoritairement sur la droite, il y a peu de points aberrants, donc on peut assumer que les résidus sont distribués selon une normale. Dans le graph **Residuals vs Leverage**, on voit qu’il y a plusieurs points qui dépasse la ligne **Cook’s distance**, donc ces points auront une grande influence sur les résultats.

**Modèle 5 :**

Après la linéarisation :

Text, letter

Description automatically generated

Text, letter

Description automatically generated

L’équation est donc :

**Signification** :

En utilisant le **test de** **Fisher**

**H0:**

**H1:**

Comme le seuil critique est de 5% et comme on obtient une p-value < , Donc on rejette H0, a une influence sur .

**Diagram, schematic

Description automatically generated**

On peut voir dans le graph **Residuals vs Fitted**, les points sont distribués aléatoirement le long de la ligne 0. Cela veut dire qu’assumer que la relation est linéaire est raisonnable. Dans le graph **Scale-Location** on voit que la distribution des points est dispersée également sur la ligne rouge, donc la variance des résidus est constante. Dans le graph **Normal Q-Q**, les points sont majoritairement sur la droite, il y a peu de points aberrants, donc on peut assumer que les résidus sont distribués selon une normale. Dans le graph **Residuals vs Leverage**, on voit qu’il y a un point en particulier qui dépasse la ligne **Cook’s distance**, donc ce point aurait une grande influence sur les résultats.

**Modèle 6**:

Après la linéarisation :

Text, letter

Description automatically generated

Text, letter

Description automatically generated

L’équation est donc :

**Signification** :

En utilisant le **test de** **Fisher**

**H0:**

**H1:**

Comme le seuil critique est de 5% et comme on obtient une p-value < , Donc on rejette H0, a une influence sur .

Chart, diagram, scatter chart

Description automatically generated

On peut voir dans le graph **Residuals vs Fitted**, les points ne sont pas distribués aléatoirement autour de la ligne 0. Cela veut dire qu’il y a Hétéroscédasticité. Dans le graph **Scale-Location** on voit que la distribution des points n’est pas n’est pas dispersée également sur la ligne rouge, donc la variance des résidus n’est pas constante. Dans le graph **Normal Q-Q**, les points sont majoritairement sur la droite, il y a peu de points aberrants, donc on peut assumer que les résidus sont distribués selon une normale. Dans le graph **Residuals vs Leverage**, on voit qu’il y a un point en particulier qui dépasse la ligne **Cook’s distance**, donc ce point aurait une grande influence sur les résultats.

**Le meilleur modèle parmi les 6 est le modèle avec la plus petite erreur standard résiduelle et la valeur R2 la plus proche de 1 qui est le modèle 5 avec une erreur standard résiduelle de 0.1668 qui est la plus petite erreur, donc le modèle a la plus grande précision et une valeur R2=0.7861 qui signifie que 78% des variations de la variable consommation peut être expliqué. Alors, le modèle 5 est le meilleur modèle.**

1. le meilleur modèle trouvé en c) dépend seulement de X2 donc sur R je vais seulement utiliser X2=2200

Text

Description automatically generated

Donc l’intervalle de confiance obtenue est [3.072505,3.734992], cela veut dire que 95% des valeurs de consommation (mpg) se retrouve dans l’intervalle.

Residual standard error: 5.16 on 158 degrees of freedom

Multiple R-squared: 0.63, Adjusted R-squared: 0.6276

F-statistic: 269 on 1 and 158 DF, p-value: < 2.2e-16

> m2 <- lm(ln(Y)~ln(X1))

Error in ln(Y) : could not find function "ln"

In addition: Warning messages:

1: In doTryCatch(return(expr), name, parentenv, handler) :

display list redraw incomplete

2: In doTryCatch(return(expr), name, parentenv, handler) :

invalid graphics state

3: In doTryCatch(return(expr), name, parentenv, handler) :

invalid graphics state

> m2 <- lm(log(Y)~log(X1))

> summary(m2)

Call:

lm(formula = log(Y) ~ log(X1))

Residuals:

Min 1Q Median 3Q Max

-0.67031 -0.10273 0.02338 0.11278 0.44306

Coefficients:

Estimate Std. Error t value

(Intercept) 7.05100 0.18882 37.34

log(X1) -0.85876 0.04118 -20.85

Pr(>|t|)

(Intercept) <2e-16 \*\*\*

log(X1) <2e-16 \*\*\*

---

Signif. codes:

0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1

‘ ’ 1

Residual standard error: 0.1862 on 158 degrees of freedom

Multiple R-squared: 0.7335, Adjusted R-squared: 0.7318

F-statistic: 434.9 on 1 and 158 DF, p-value: < 2.2e-16

> anova(m2)

Analysis of Variance Table

Response: log(Y)

Df Sum Sq Mean Sq F value

log(X1) 1 15.0735 15.0735 434.88

Residuals 158 5.4765 0.0347

Pr(>F)

log(X1) < 2.2e-16 \*\*\*

Residuals

---

Signif. codes:

0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1

‘ ’ 1

> plot(m2)

> m3 <- lm(log(Y)~log(X1))

> summary(m3)

Call:

lm(formula = log(Y) ~ log(X1))

Residuals:

Min 1Q Median 3Q Max

-0.67031 -0.10273 0.02338 0.11278 0.44306

Coefficients:

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

(Intercept) 7.05100 0.18882 37.34 <2e-16 \*\*\*

log(X1) -0.85876 0.04118 -20.85 <2e-16 \*\*\*

---

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

Residual standard error: 0.1862 on 158 degrees of freedom

Multiple R-squared: 0.7335, Adjusted R-squared: 0.7318

F-statistic: 434.9 on 1 and 158 DF, p-value: < 2.2e-16

> anova(m3)

Analysis of Variance Table

Response: log(Y)

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

log(X1) 1 15.0735 15.0735 434.88 < 2.2e-16 \*\*\*

Residuals 158 5.4765 0.0347

---

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

> m3 <- lm(log(Y)~X1)

> summary(m3)

Call:

lm(formula = log(Y) ~ X1)

Residuals:

Min 1Q Median 3Q Max

-0.65802 -0.11937 0.00555 0.12813 0.45036

Coefficients:

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

(Intercept) 3.9201351 0.0426064 92.01 <2e-16 \*\*\*

X1 -0.0076953 0.0003854 -19.97 <2e-16 \*\*\*

---

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

Residual standard error: 0.1921 on 158 degrees of freedom

Multiple R-squared: 0.7162, Adjusted R-squared: 0.7144

F-statistic: 398.7 on 1 and 158 DF, p-value: < 2.2e-16

> anova(m3)

Analysis of Variance Table

Response: log(Y)

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

X1 1 14.7175 14.7175 398.68 < 2.2e-16 \*\*\*

Residuals 158 5.8326 0.0369

---

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

> m4 <- lm(Y~X2)

> summary(m4)

Call:

lm(formula = Y ~ X2)

Residuals:

Min 1Q Median 3Q Max

-7.738 -3.473 -0.876 2.933 14.400

Coefficients:

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

(Intercept) 48.4319424 1.3159135 36.80 <2e-16 \*\*\*

X2 -0.0082363 0.0004301 -19.15 <2e-16 \*\*\*

---

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

Residual standard error: 4.655 on 158 degrees of freedom

Multiple R-squared: 0.6989, Adjusted R-squared: 0.697

F-statistic: 366.7 on 1 and 158 DF, p-value: < 2.2e-16

> anova(m4)

Analysis of Variance Table

Response: Y

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

X2 1 7946.5 7946.5 366.75 < 2.2e-16 \*\*\*

Residuals 158 3423.5 21.7

---

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

> plot(m3)

> plot(m4)

> m5 <- lm(log(Y)~log(X2))

> summary(m5)

Call:

lm(formula = log(Y) ~ log(X2))

Residuals:

Min 1Q Median 3Q Max

-0.43400 -0.11546 -0.00306 0.11171 0.58714

Coefficients:

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

(Intercept) 12.03177 0.36983 32.53 <2e-16 \*\*\*

log(X2) -1.12107 0.04652 -24.10 <2e-16 \*\*\*

---

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

Residual standard error: 0.1668 on 158 degrees of freedom

Multiple R-squared: 0.7861, Adjusted R-squared: 0.7848

F-statistic: 580.7 on 1 and 158 DF, p-value: < 2.2e-16

> anova(m5)

Analysis of Variance Table

Response: log(Y)

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

log(X2) 1 16.1546 16.1546 580.69 < 2.2e-16 \*\*\*

Residuals 158 4.3955 0.0278

---

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

> plot(m5)

> m6 <- lm(log(Y)~X2)

> summary(m6)

Call:

lm(formula = log(Y) ~ X2)

Residuals:

Min 1Q Median 3Q Max

-0.4588 -0.1203 -0.0126 0.1154 0.5409

Coefficients:

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

(Intercept) 4.212e+00 4.785e-02 88.02 <2e-16 \*\*\*

X2 -3.698e-04 1.564e-05 -23.65 <2e-16 \*\*\*

---

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

Residual standard error: 0.1693 on 158 degrees of freedom

Multiple R-squared: 0.7797, Adjusted R-squared: 0.7783

F-statistic: 559.2 on 1 and 158 DF, p-value: < 2.2e-16

> anova(m6)

Analysis of Variance Table

Response: log(Y)

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

X2 1 16.0229 16.0229 559.22 < 2.2e-16 \*\*\*

Residuals 158 4.5271 0.0287

---

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

> plot(m6)

> length(m6)

[1] 12

> m <- lm(formula=Y~X1,data=Y)

Error in model.frame.default(formula = Y ~ X1, data = Y, drop.unused.levels = TRUE) :

'data' must be a data.frame, environment, or list

> table1 <-[Y,X1]

Error: unexpected '[' in "table1 <-["

> table1 <- c(X1,Y)

> table1

[1] 85.0 95.0 105.0 170.0 110.0 133.0 46.0 75.0 145.0 86.0 175.0 100.0

[13] 68.0 150.0 150.0 180.0 67.0 78.0 100.0 210.0 112.0 175.0 70.0 150.0

[25] 88.0 90.0 110.0 75.0 75.0 70.0 120.0 77.0 65.0 72.0 105.0 87.0

[37] 48.0 80.0 67.0 90.0 53.0 110.0 72.0 170.0 68.0 155.0 65.0 85.0

[49] 68.0 145.0 103.0 102.0 90.0 129.0 150.0 150.0 78.0 165.0 52.0 97.0

[61] 70.0 72.0 90.0 85.0 90.0 90.0 95.0 48.0 140.0 70.0 140.0 175.0

[73] 208.0 88.0 90.0 150.0 68.0 96.0 67.0 180.0 72.0 160.0 215.0 75.0

[85] 115.0 67.0 79.0 63.0 130.0 76.0 95.0 78.0 150.0 98.0 225.0 79.0

[97] 92.0 165.0 67.0 100.0 67.0 110.0 97.0 65.0 92.0 150.0 90.0 130.0

[109] 88.0 130.0 86.0 90.0 84.0 180.0 150.0 76.0 88.0 67.0 138.0 170.0

[121] 122.0 60.0 69.0 75.0 150.0 78.0 167.0 58.0 54.0 100.0 65.0 100.0

[133] 88.0 70.0 180.0 110.0 74.0 175.0 165.0 95.0 139.0 95.0 75.0 70.0

[145] 92.0 91.0 70.0 49.0 97.0 69.0 78.0 150.0 80.0 75.0 65.0 88.0

[157] 96.0 108.0 87.0 75.0 38.0 25.0 16.0 13.0 21.0 16.2 26.0 36.0

[169] 15.0 28.0 14.0 18.0 37.0 16.0 16.0 12.0 33.8 30.5 16.0 11.0

[181] 18.0 13.0 29.0 18.5 34.0 27.0 18.0 33.7 24.0 36.0 15.5 25.4

[193] 40.8 21.0 26.6 25.0 43.4 28.0 44.6 28.4 33.0 21.5 32.4 15.0

[205] 29.5 16.9 37.2 20.2 31.5 19.2 20.3 20.0 29.8 17.6 15.0 13.0

[217] 34.3 13.0 32.8 24.0 39.4 15.0 21.0 20.8 24.0 24.3 22.0 44.3

[229] 17.0 30.0 19.4 14.0 11.0 36.0 27.0 14.0 30.0 24.0 26.0 11.0

[241] 15.0 14.0 13.0 32.4 25.0 36.4 28.0 30.5 17.0 30.0 23.0 29.0

[253] 15.0 18.5 14.0 26.0 37.0 17.7 38.0 19.0 38.0 17.0 24.0 37.0

[265] 28.0 14.0 19.1 18.0 27.0 13.0 22.0 22.5 26.6 12.0 13.0 41.5

[277] 27.0 32.0 16.5 15.5 20.0 36.1 35.0 31.3 17.0 26.0 12.0 39.1

[289] 23.0 19.0 34.1 18.0 18.0 34.0 16.0 17.5 31.6 13.0 15.0 21.1

[301] 18.1 27.5 32.2 34.2 25.8 20.0 34.5 29.0 27.2 37.3 18.0 15.0

[313] 25.0 26.0 31.0 22.3 32.0 19.0 21.0 29.0

> table1 <- table(mondata$horsepower,mondata$mpg)

> table1

11 12 13 14 15 15.5 16 16.2 16.5 16.9 17 17.5 17.6 17.7 18 18.1 18.5

46 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

48 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

49 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

52 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

53 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

54 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

58 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

60 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

63 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

65 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

67 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

68 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

19 19.1 19.2 19.4 20 20.2 20.3 20.8 21 21.1 21.5 22 22.3 22.5 23 24

46 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

48 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

49 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

52 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

53 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

54 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0

58 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

60 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

63 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

65 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

67 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

68 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

24.3 25 25.4 25.8 26 26.6 27 27.2 27.5 28 28.4 29 29.5 29.8 30 30.5 31

46 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0

48 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

49 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0

52 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

53 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

54 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

58 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

60 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

63 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0

65 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1

67 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0

68 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0

31.3 31.5 31.6 32 32.2 32.4 32.8 33 33.7 33.8 34 34.1 34.2 34.3 34.5

46 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

48 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

49 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

52 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0

53 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0

54 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

58 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

60 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

63 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

65 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0

67 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0

68 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0

35 36 36.1 36.4 37 37.2 37.3 38 39.1 39.4 40.8 41.5 43.4 44.3 44.6

46 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

48 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0

49 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

52 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

53 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

54 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

58 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0

60 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0

63 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

65 0 0 0 0 1 1 0 0 0 0 1 0 0 0 0

67 0 0 0 1 0 0 0 2 0 0 0 0 0 0 1

68 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0

[ reached getOption("max.print") -- omitted 51 rows ]

> table1 <- matrix(c(X1,Y))

> table1

[,1]

[1,] 85.0

[2,] 95.0

[3,] 105.0

[4,] 170.0

[5,] 110.0

[6,] 133.0

[7,] 46.0

[8,] 75.0

[9,] 145.0

[10,] 86.0

[11,] 175.0

[12,] 100.0

[13,] 68.0

[14,] 150.0

[15,] 150.0

[16,] 180.0

[17,] 67.0

[18,] 78.0

[19,] 100.0

[20,] 210.0

[21,] 112.0

[22,] 175.0

[23,] 70.0

[24,] 150.0

[25,] 88.0

[26,] 90.0

[27,] 110.0

[28,] 75.0

[29,] 75.0

[30,] 70.0

[31,] 120.0

[32,] 77.0

[33,] 65.0

[34,] 72.0

[35,] 105.0

[36,] 87.0

[37,] 48.0

[38,] 80.0

[39,] 67.0

[40,] 90.0

[41,] 53.0

[42,] 110.0

[43,] 72.0

[44,] 170.0

[45,] 68.0

[46,] 155.0

[47,] 65.0

[48,] 85.0

[49,] 68.0

[50,] 145.0

[51,] 103.0

[52,] 102.0

[53,] 90.0

[54,] 129.0

[55,] 150.0

[56,] 150.0

[57,] 78.0

[58,] 165.0

[59,] 52.0

[60,] 97.0

[61,] 70.0

[62,] 72.0

[63,] 90.0

[64,] 85.0

[65,] 90.0

[66,] 90.0

[67,] 95.0

[68,] 48.0

[69,] 140.0

[70,] 70.0

[71,] 140.0

[72,] 175.0

[73,] 208.0

[74,] 88.0

[75,] 90.0

[76,] 150.0

[77,] 68.0

[78,] 96.0

[79,] 67.0

[80,] 180.0

[81,] 72.0

[82,] 160.0

[83,] 215.0

[84,] 75.0

[85,] 115.0

[86,] 67.0

[87,] 79.0

[88,] 63.0

[89,] 130.0

[90,] 76.0

[91,] 95.0

[92,] 78.0

[93,] 150.0

[94,] 98.0

[95,] 225.0

[96,] 79.0

[97,] 92.0

[98,] 165.0

[99,] 67.0

[100,] 100.0

[101,] 67.0

[102,] 110.0

[103,] 97.0

[104,] 65.0

[105,] 92.0

[106,] 150.0

[107,] 90.0

[108,] 130.0

[109,] 88.0

[110,] 130.0

[111,] 86.0

[112,] 90.0

[113,] 84.0

[114,] 180.0

[115,] 150.0

[116,] 76.0

[117,] 88.0

[118,] 67.0

[119,] 138.0

[120,] 170.0

[121,] 122.0

[122,] 60.0

[123,] 69.0

[124,] 75.0

[125,] 150.0

[126,] 78.0

[127,] 167.0

[128,] 58.0

[129,] 54.0

[130,] 100.0

[131,] 65.0

[132,] 100.0

[133,] 88.0

[134,] 70.0

[135,] 180.0

[136,] 110.0

[137,] 74.0

[138,] 175.0

[139,] 165.0

[140,] 95.0

[141,] 139.0

[142,] 95.0

[143,] 75.0

[144,] 70.0

[145,] 92.0

[146,] 91.0

[147,] 70.0

[148,] 49.0

[149,] 97.0

[150,] 69.0

[151,] 78.0

[152,] 150.0

[153,] 80.0

[154,] 75.0

[155,] 65.0

[156,] 88.0

[157,] 96.0

[158,] 108.0

[159,] 87.0

[160,] 75.0

[161,] 38.0

[162,] 25.0

[163,] 16.0

[164,] 13.0

[165,] 21.0

[166,] 16.2

[167,] 26.0

[168,] 36.0

[169,] 15.0

[170,] 28.0

[171,] 14.0

[172,] 18.0

[173,] 37.0

[174,] 16.0

[175,] 16.0

[176,] 12.0

[177,] 33.8

[178,] 30.5

[179,] 16.0

[180,] 11.0

[181,] 18.0

[182,] 13.0

[183,] 29.0

[184,] 18.5

[185,] 34.0

[186,] 27.0

[187,] 18.0

[188,] 33.7

[189,] 24.0

[190,] 36.0

[191,] 15.5

[192,] 25.4

[193,] 40.8

[194,] 21.0

[195,] 26.6

[196,] 25.0

[197,] 43.4

[198,] 28.0

[199,] 44.6

[200,] 28.4

[201,] 33.0

[202,] 21.5

[203,] 32.4

[204,] 15.0

[205,] 29.5

[206,] 16.9

[207,] 37.2

[208,] 20.2

[209,] 31.5

[210,] 19.2

[211,] 20.3

[212,] 20.0

[213,] 29.8

[214,] 17.6

[215,] 15.0

[216,] 13.0

[217,] 34.3

[218,] 13.0

[219,] 32.8

[220,] 24.0

[221,] 39.4

[222,] 15.0

[223,] 21.0

[224,] 20.8

[225,] 24.0

[226,] 24.3

[227,] 22.0

[228,] 44.3

[229,] 17.0

[230,] 30.0

[231,] 19.4

[232,] 14.0

[233,] 11.0

[234,] 36.0

[235,] 27.0

[236,] 14.0

[237,] 30.0

[238,] 24.0

[239,] 26.0

[240,] 11.0

[241,] 15.0

[242,] 14.0

[243,] 13.0

[244,] 32.4

[245,] 25.0

[246,] 36.4

[247,] 28.0

[248,] 30.5

[249,] 17.0

[250,] 30.0

[251,] 23.0

[252,] 29.0

[253,] 15.0

[254,] 18.5

[255,] 14.0

[256,] 26.0

[257,] 37.0

[258,] 17.7

[259,] 38.0

[260,] 19.0

[261,] 38.0

[262,] 17.0

[263,] 24.0

[264,] 37.0

[265,] 28.0

[266,] 14.0

[267,] 19.1

[268,] 18.0

[269,] 27.0

[270,] 13.0

[271,] 22.0

[272,] 22.5

[273,] 26.6

[274,] 12.0

[275,] 13.0

[276,] 41.5

[277,] 27.0

[278,] 32.0

[279,] 16.5

[280,] 15.5

[281,] 20.0

[282,] 36.1

[283,] 35.0

[284,] 31.3

[285,] 17.0

[286,] 26.0

[287,] 12.0

[288,] 39.1

[289,] 23.0

[290,] 19.0

[291,] 34.1

[292,] 18.0

[293,] 18.0

[294,] 34.0

[295,] 16.0

[296,] 17.5

[297,] 31.6

[298,] 13.0

[299,] 15.0

[300,] 21.1

[301,] 18.1

[302,] 27.5

[303,] 32.2

[304,] 34.2

[305,] 25.8

[306,] 20.0

[307,] 34.5

[308,] 29.0

[309,] 27.2

[310,] 37.3

[311,] 18.0

[312,] 15.0

[313,] 25.0

[314,] 26.0

[315,] 31.0

[316,] 22.3

[317,] 32.0

[318,] 19.0

[319,] 21.0

[320,] 29.0

> table1<-matrix(c(X1,Y),ncol = 2)

> table1

[,1] [,2]

[1,] 85 38.0

[2,] 95 25.0

[3,] 105 16.0

[4,] 170 13.0

[5,] 110 21.0

[6,] 133 16.2

[7,] 46 26.0

[8,] 75 36.0

[9,] 145 15.0

[10,] 86 28.0

[11,] 175 14.0

[12,] 100 18.0

[13,] 68 37.0

[14,] 150 16.0

[15,] 150 16.0

[16,] 180 12.0

[17,] 67 33.8

[18,] 78 30.5

[19,] 100 16.0

[20,] 210 11.0

[21,] 112 18.0

[22,] 175 13.0

[23,] 70 29.0

[24,] 150 18.5

[25,] 88 34.0

[26,] 90 27.0

[27,] 110 18.0

[28,] 75 33.7

[29,] 75 24.0

[30,] 70 36.0

[31,] 120 15.5

[32,] 77 25.4

[33,] 65 40.8

[34,] 72 21.0

[35,] 105 26.6

[36,] 87 25.0

[37,] 48 43.4

[38,] 80 28.0

[39,] 67 44.6

[40,] 90 28.4

[41,] 53 33.0

[42,] 110 21.5

[43,] 72 32.4

[44,] 170 15.0

[45,] 68 29.5

[46,] 155 16.9

[47,] 65 37.2

[48,] 85 20.2

[49,] 68 31.5

[50,] 145 19.2

[51,] 103 20.3

[52,] 102 20.0

[53,] 90 29.8

[54,] 129 17.6

[55,] 150 15.0

[56,] 150 13.0

[57,] 78 34.3

[58,] 165 13.0

[59,] 52 32.8

[60,] 97 24.0

[61,] 70 39.4

[62,] 72 15.0

[63,] 90 21.0

[64,] 85 20.8

[65,] 90 24.0

[66,] 90 24.3

[67,] 95 22.0

[68,] 48 44.3

[69,] 140 17.0

[70,] 70 30.0

[71,] 140 19.4

[72,] 175 14.0

[73,] 208 11.0

[74,] 88 36.0

[75,] 90 27.0

[76,] 150 14.0

[77,] 68 30.0

[78,] 96 24.0

[79,] 67 26.0

[80,] 180 11.0

[81,] 72 15.0

[82,] 160 14.0

[83,] 215 13.0

[84,] 75 32.4

[85,] 115 25.0

[86,] 67 36.4

[87,] 79 28.0

[88,] 63 30.5

[89,] 130 17.0

[90,] 76 30.0

[91,] 95 23.0

[92,] 78 29.0

[93,] 150 15.0

[94,] 98 18.5

[95,] 225 14.0

[96,] 79 26.0

[97,] 92 37.0

[98,] 165 17.7

[99,] 67 38.0

[100,] 100 19.0

[101,] 67 38.0

[102,] 110 17.0

[103,] 97 24.0

[104,] 65 37.0

[105,] 92 28.0

[106,] 150 14.0

[107,] 90 19.1

[108,] 130 18.0

[109,] 88 27.0

[110,] 130 13.0

[111,] 86 22.0

[112,] 90 22.5

[113,] 84 26.6

[114,] 180 12.0

[115,] 150 13.0

[116,] 76 41.5

[117,] 88 27.0

[118,] 67 32.0

[119,] 138 16.5

[120,] 170 15.5

[121,] 122 20.0

[122,] 60 36.1

[123,] 69 35.0

[124,] 75 31.3

[125,] 150 17.0

[126,] 78 26.0

[127,] 167 12.0

[128,] 58 39.1

[129,] 54 23.0

[130,] 100 19.0

[131,] 65 34.1

[132,] 100 18.0

[133,] 88 18.0

[134,] 70 34.0

[135,] 180 16.0

[136,] 110 17.5

[137,] 74 31.6

[138,] 175 13.0

[139,] 165 15.0

[140,] 95 21.1

[141,] 139 18.1

[142,] 95 27.5

[143,] 75 32.2

[144,] 70 34.2

[145,] 92 25.8

[146,] 91 20.0

[147,] 70 34.5

[148,] 49 29.0

[149,] 97 27.2

[150,] 69 37.3

[151,] 78 18.0

[152,] 150 15.0

[153,] 80 25.0

[154,] 75 26.0

[155,] 65 31.0

[156,] 88 22.3

[157,] 96 32.0

[158,] 108 19.0

[159,] 87 21.0

[160,] 75 29.0

> m <- lm(formula=Y~X1,data=table1)

Error in model.frame.default(formula = Y ~ X1, data = table1, drop.unused.levels = TRUE) :

'data' must be a data.frame, not a matrix or an array

> t1<-data.frame(table1)

> t1

X1 X2

1 85 38.0

2 95 25.0

3 105 16.0

4 170 13.0

5 110 21.0

6 133 16.2

7 46 26.0

8 75 36.0

9 145 15.0

10 86 28.0

11 175 14.0

12 100 18.0

13 68 37.0

14 150 16.0

15 150 16.0

16 180 12.0

17 67 33.8

18 78 30.5

19 100 16.0

20 210 11.0

21 112 18.0

22 175 13.0

23 70 29.0

24 150 18.5

25 88 34.0

26 90 27.0

27 110 18.0

28 75 33.7

29 75 24.0

30 70 36.0

31 120 15.5

32 77 25.4

33 65 40.8

34 72 21.0

35 105 26.6

36 87 25.0

37 48 43.4

38 80 28.0

39 67 44.6

40 90 28.4

41 53 33.0

42 110 21.5

43 72 32.4

44 170 15.0

45 68 29.5

46 155 16.9

47 65 37.2

48 85 20.2

49 68 31.5

50 145 19.2

51 103 20.3

52 102 20.0

53 90 29.8

54 129 17.6

55 150 15.0

56 150 13.0

57 78 34.3

58 165 13.0

59 52 32.8

60 97 24.0

61 70 39.4

62 72 15.0

63 90 21.0

64 85 20.8

65 90 24.0

66 90 24.3

67 95 22.0

68 48 44.3

69 140 17.0

70 70 30.0

71 140 19.4

72 175 14.0

73 208 11.0

74 88 36.0

75 90 27.0

76 150 14.0

77 68 30.0

78 96 24.0

79 67 26.0

80 180 11.0

81 72 15.0

82 160 14.0

83 215 13.0

84 75 32.4

85 115 25.0

86 67 36.4

87 79 28.0

88 63 30.5

89 130 17.0

90 76 30.0

91 95 23.0

92 78 29.0

93 150 15.0

94 98 18.5

95 225 14.0

96 79 26.0

97 92 37.0

98 165 17.7

99 67 38.0

100 100 19.0

101 67 38.0

102 110 17.0

103 97 24.0

104 65 37.0

105 92 28.0

106 150 14.0

107 90 19.1

108 130 18.0

109 88 27.0

110 130 13.0

111 86 22.0

112 90 22.5

113 84 26.6

114 180 12.0

115 150 13.0

116 76 41.5

117 88 27.0

118 67 32.0

119 138 16.5

120 170 15.5

121 122 20.0

122 60 36.1

123 69 35.0

124 75 31.3

125 150 17.0

126 78 26.0

127 167 12.0

128 58 39.1

129 54 23.0

130 100 19.0

131 65 34.1

132 100 18.0

133 88 18.0

134 70 34.0

135 180 16.0

136 110 17.5

137 74 31.6

138 175 13.0

139 165 15.0

140 95 21.1

141 139 18.1

142 95 27.5

143 75 32.2

144 70 34.2

145 92 25.8

146 91 20.0

147 70 34.5

148 49 29.0

149 97 27.2

150 69 37.3

151 78 18.0

152 150 15.0

153 80 25.0

154 75 26.0

155 65 31.0

156 88 22.3

157 96 32.0

158 108 19.0

159 87 21.0

160 75 29.0

> m <- lm(formula=Y~X1,data=t1)

> summary(m)

Call:

lm(formula = Y ~ X1, data = t1)

Residuals:

Min 1Q Median 3Q Max

-14.548 -3.218 -0.202 2.776 14.203

Coefficients:

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

(Intercept) 41.77153 1.14431 36.5 <2e-16 \*\*\*

X1 -0.16977 0.01035 -16.4 <2e-16 \*\*\*

---

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

Residual standard error: 5.16 on 158 degrees of freedom

Multiple R-squared: 0.63, Adjusted R-squared: 0.6276

F-statistic: 269 on 1 and 158 DF, p-value: < 2.2e-16

> plot(m)

> plot(m2)

> plot(m3)

> plot(m4)

> plot(m5)

> plot(m6)

> plot(m4)

> predict(m5,data.frame(X1=120,X2=2200), level=0.95,interval="prediction")

fit lwr upr

1 3.403749 3.072505 3.734992

>