92/125

228.371 ASSIGNMENT 1

ID 00001

Wang, Yue
12027710

1.

```
> t.test(gas$HeatRate,mu=9750, alternative='greater')

        One Sample t-test

data:  gas$HeatRate
t = 3.8569, df = 31, p-value = 0.0002714
alternative hypothesis: true mean is greater than 9750
95 percent confidence interval:
 10412.64      Inf
sample estimates:
mean of x
 10932.44
```

Explain

Answer : Sample is not representative.

-2

2.

```
> t.test(gas$InletTemp, gas$ExhTemp, paired=TRUE)


        Paired t-test


data:  gas$InletTemp and gas$ExhTemp
t = 35.3892, df = 31, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 621.139 697.111
sample estimates:
mean of the differences
              659.125
```

Answer : P value is much smaller than 0.05, and then NULL hypothesis is rejected. The default value for NULL hypothesis is zero. So the mean reduction between InletTemp and ExhTemp exists, which value is between 621.139 and 697.111.

3.

```
> t.test(gas$Power[gas$Engine == "Advanced"],gas$Power[gas$Engine == "Traditional"],
alternative="greater")

	Welch Two Sample t-test

data:  gas$Power[gas$Engine == "Advanced"] and gas$Power[gas$Engine == "Traditional"]
t = 2.598, df = 21.509, p-value = 0.008297
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 29931.71      Inf
sample estimates:
mean of x mean of y
153469.62  65017.33
```

What assumptions are required if your findings are to be generalised?

-2

Answer:

on average

As shown above, p-value = 0.008297, alternative hypothesis: true difference in means is greater than 0. Hence the advanced engine have more power than the other.

✓

4

```
> t.test(InletTemp[Engine == 'Aeroderiv'],ExhTemp[Engine == 'Aeroderiv'],paired=TRUE)

	Paired t-test

data:  InletTemp[Engine == "Aeroderiv"] and ExhTemp[Engine == "Aeroderiv"]
t = 9.6462, df = 6, p-value = 7.11e-05
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 438.0989 735.9011
sample estimates:
```

```
mean of the differences
              587

> t.test(InletTemp[Engine == 'Advanced'],ExhTemp[Engine == 'Advanced'],paired=TRUE)

	Paired t-test

data:  InletTemp[Engine == "Advanced"] and ExhTemp[Engine == "Advanced"]
t = 66.5674, df = 12, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 716.9696 765.4919
sample estimates:
mean of the differences
           741.2308

> t.test(InletTemp[Engine == 'Traditional'],ExhTemp[Engine == 'Traditional'],paired=TRUE)

	Paired t-test

data:  InletTemp[Engine == "Traditional"] and ExhTemp[Engine == "Traditional"]
t = 51.2638, df = 11, p-value = 1.914e-14
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 585.9634 638.5366
sample estimates:
mean of the differences
            612.25
```

Answers: Three different types of turbines have three different estimated mean temperature reductions. The values are 587, 741.2308 and 612.25.

```
> anova(lm(InletTemp-ExhTemp ~ Engine))
Analysis of Variance Table

Response: InletTemp - ExhTemp
          Df Sum Sq Mean Sq F value    Pr(>F)
Engine     2 150419   75209   11.26 0.0002405 ***
Residuals 29 193699    6679
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

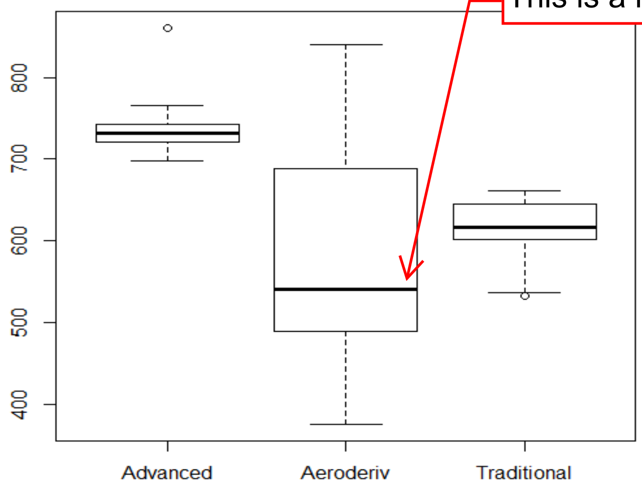
There are at least one pair of engine types that have different power.

Comment

−3

boxplot((InletTemp - ExhTemp) ~ Engine)

This is a median



The graph shows the mean of the heat difference between inlet and outlet is different significantly.

5.

> cor.test(gas$Power, gas$Airflow)

Pearson's product-moment correlation

data: gas$Power and gas$Airflow

t = 36.0711, df = 30, p-value < 2.2e-16

alternative hypothesis: true correlation is not equal to 0

95 percent confidence interval:

 0.9766740 0.9945111

sample estimates:

cor

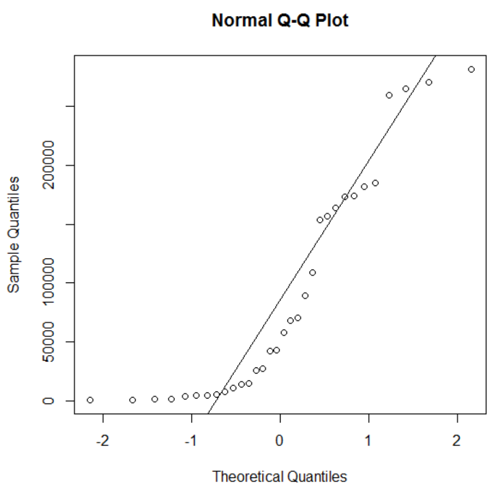
0.9886671

Answer: P value is less than 2e-`6, Null hypothesis is rejected, hence this linear model is suitable for the relationship between Airflow and Power.

6

> qqnorm(Power)

> qqline(Power)

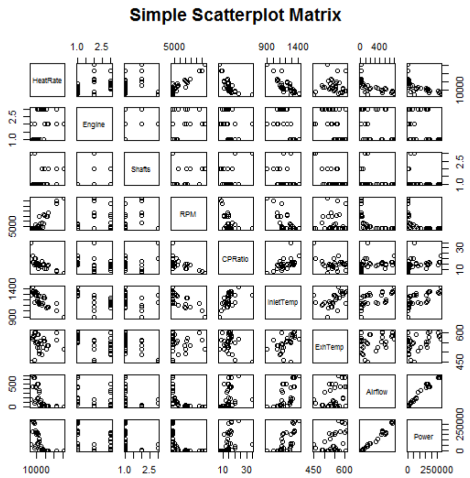


```
> t.test(Power)

        One Sample t-test

data:  Power
t = 5.3919, df = 31, p-value = 6.981e-06
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
  55776.87 123644.19
sample estimates:
mean of x
 89710.53

> shapiro.test(Power)

        Shapiro-Wilk normality test

data:  Power
W = 0.8383, p-value = 0.0002344
```

Answer: As shown above the trend does not follow qqline. P value is much smaller than 0.05, so NULL hypothesis has been rejected. Null hypothesis stands for Power is normally distributed, hence this is not a normal distribution of the power of the turbines.

7

pairs(HeatRate~Engine+Shafts+RPM+CPRatio+InletTemp+ExhTemp+Airflow+Power,data=gas, main="Simple Scatterplot Matrix")



Answer: RPM and InletTemp are the numeric variables are useful as linear regression predictor.

8.

> summary(lm(HeatRate ~ RPM))

Call:

lm(formula = HeatRate ~ RPM)

Residuals:

Min	1Q	Median	3Q	Max
-1424.64	-404.31	-36.18	351.38	1575.73

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	9.238e+03	1.689e+02	54.70	< 2e-16	***
RPM	1.901e-01	1.384e-02	13.74	1.77e-14	***

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

Residual standard error: 652.9 on 30 degrees of freedom

Multiple R-squared: 0.8628,   Adjusted R-squared: 0.8583

F-statistic: 188.7 on 1 and 30 DF, p-value: 1.772e-14

Answer : p value is 1.777e-14 and Multiple R-squared: 0.8628 which is close to 1, hence these two data sets fit to a linear model.

9.

```
> summary(lm(HeatRate ~ InletTemp))

Call:
lm(formula = HeatRate ~ InletTemp)

Residuals:
    Min       1Q   Median       3Q      Max
-1989.23  -847.72   -33.47   547.72  2864.75

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 24798.940   1958.107  12.665 1.43e-13 ***
InletTemp     -11.501      1.616  -7.115 6.50e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1075 on 30 degrees of freedom
Multiple R-squared:  0.6279,    Adjusted R-squared:  0.6155
F-statistic: 50.63 on 1 and 30 DF,  p-value: 6.498e-08
```

Answer : Multiple R-squared: 0.6279, hence InletTemp is ok, but RPM is a better choice.

10.

```
> summary(lm(HeatRate~InletTemp*RPM))

Call:

lm(formula = HeatRate ~ InletTemp * RPM)


Residuals:

    Min      1Q  Median      3Q     Max

-1510.85  -217.73    84.35   208.47  1176.65


Coefficients:

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

(Intercept)    1.438e+04  1.778e+03   8.088 8.33e-09 ***

InletTemp     -3.957e+00  1.421e+00  -2.785   0.0095 **

RPM            1.711e-01  1.037e-01   1.651   0.1100

InletTemp:RPM -2.007e-05  9.584e-05  -0.209   0.8356

---

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


Residual standard error: 563.3 on 28 degrees of freedom

Multiple R-squared:  0.9047,    Adjusted R-squared:  0.8945

F-statistic: 88.62 on 3 and 28 DF,  p-value: 2.103e-14
```

Answer : the interaction term was not required.

Why?

−1

11.

```
> summary(lm(formula = HeatRate ~ Shafts + RPM + CPRatio + InletTemp + ExhTemp + Airflow + Power))

Call:
lm(formula = HeatRate ~ Shafts + RPM + CPRatio + InletTemp +
    ExhTemp + Airflow + Power)

Residuals:
    Min      1Q  Median      3Q     Max
-991.83 -215.47  -62.23  193.12  954.98

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.168e+04  2.196e+03   5.318 1.86e-05 ***
Shafts       4.559e+01  2.146e+02   0.212   0.8335
RPM          1.082e-01  2.282e-02   4.741 8.02e-05 ***
CPRatio     -1.619e+01  4.058e+01  -0.399   0.6935
InletTemp   -6.337e+00  2.283e+00  -2.776   0.0105 *
ExhTemp      1.156e+01  5.436e+00   2.126   0.0440 *
Airflow      9.665e-01  3.260e+00   0.296   0.7694
Power       -4.860e-03  8.253e-03  -0.589   0.5615
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 441.6 on 24 degrees of freedom
Multiple R-squared:  0.9498,    Adjusted R-squared:  0.9351
F-statistic: 64.86 on 7 and 24 DF,  p-value: 4.793e-14
```

Comment on the significance

−2

12.

Answer: According to beta values, when predictor variables of Shaft, CPRatio, InletTemp, Power increase, the response variable decrease with each particular ratio. -2

by how much?

13.

```
> summary(lm(HeatRate~InletTemp+RPM))

Call:
lm(formula = HeatRate ~ InletTemp + RPM)

Residuals:
    Min      1Q  Median      3Q     Max 
-1529.9  -243.1    83.9   208.5  1168.8 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)  1.457e+04  1.504e+03   9.688 1.35e-10 ***
InletTemp   -4.123e+00  1.158e+00  -3.561   0.0013 ** 
RPM          1.497e-01  1.633e-02   9.169 4.55e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 553.9 on 29 degrees of freedom
Multiple R-squared:  0.9046,	Adjusted R-squared:  0.898 
F-statistic: 137.4 on 2 and 29 DF,  p-value: 1.605e-15

> summary(lm(HeatRate~Shafts+RPM+CPRatio+InletTemp+ExhTemp+Airflow+Power))

Call:
lm(formula = HeatRate ~ Shafts + RPM + CPRatio + InletTemp + 
    ExhTemp + Airflow + Power)
```

```
Residuals:
    Min      1Q  Median      3Q     Max
-991.83 -215.47  -62.23  193.12  954.98

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.168e+04  2.196e+03   5.318 1.86e-05 ***
Shafts       4.559e+01  2.146e+02   0.212   0.8335
RPM          1.082e-01  2.282e-02   4.741 8.02e-05 ***
CPRatio     -1.619e+01  4.058e+01  -0.399   0.6935
InletTemp   -6.337e+00  2.283e+00  -2.776   0.0105 *
ExhTemp      1.156e+01  5.436e+00   2.126   0.0440 *
Airflow      9.665e-01  3.260e+00   0.296   0.7694
Power       -4.860e-03  8.253e-03  -0.589   0.5615
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 441.6 on 24 degrees of freedom
Multiple R-squared:  0.9498,    Adjusted R-squared:  0.9351
F-statistic: 64.86 on 7 and 24 DF,  p-value: 4.793e-14
```

Quote value

Answer: Applying least squared theory, 7-predictor has smaller Residual standard error. Hence it has better predictive than the 2-predictor model.

14.

```
> summary(lm(HeatRate~Shafts+RPM+CPRatio+InletTemp+ExhTemp+Airflow+Power))

Call:
lm(formula = HeatRate ~ Shafts + RPM + CPRatio + InletTemp +
    ExhTemp + Airflow + Power)

Residuals:
    Min      1Q  Median      3Q     Max
-991.83 -215.47  -62.23  193.12  954.98

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.168e+04  2.196e+03   5.318 1.86e-05 ***
Shafts       4.559e+01  2.146e+02   0.212   0.8335
RPM          1.082e-01  2.282e-02   4.741 8.02e-05 ***
CPRatio     -1.619e+01  4.058e+01  -0.399   0.6935
InletTemp   -6.337e+00  2.283e+00  -2.776   0.0105 *
ExhTemp      1.156e+01  5.436e+00   2.126   0.0440 *
Airflow      9.665e-01  3.260e+00   0.296   0.7694
Power       -4.860e-03  8.253e-03  -0.589   0.5615
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 441.6 on 24 degrees of freedom
Multiple R-squared:  0.9498,    Adjusted R-squared:  0.9351
F-statistic: 64.86 on 7 and 24 DF,  p-value: 4.793e-14
```

Answer: As shown above, it means 94.98% variation.

The model explains 94% of the variation in the HeatRate of the turbines.

-3

This means that some part of the model is useful. It does not say that this model is the best model.

15. -1

P value is 4.793e-14, much less than 0.01, Null hypothesis is rejected. Hence this model is good to predicate HeatRate.

16.

```
> lm1<-lm(HeatRate~Shafts+RPM+CPRatio+InletTemp+ExhTemp+Airflow+Power)
> hatvalues(lm1)
        1         2         3         4         5         6         7         8
0.4284616 0.3252555 0.1641741 0.7507745 0.5133682 0.4849646 0.2036465 0.1885726
        9        10        11        12        13        14        15        16
0.1761790 0.2563977 0.2641164 0.2880030 0.3192141 0.1678053 0.1670899 0.2168007
       17        18        19        20        21        22        23        24
0.2168691 0.1559862 0.3693915 0.1646246 0.1452208 0.1148014 0.4081885 0.1573087
       25        26        27        28        29        30        31        32
0.1635554 0.1737555 0.1835541 0.1591480 0.2287911 0.1181886 0.2243930 0.1013998
> which((hatvalues(lm1)/mean(hatvalues(lm1))>2))
4 5
4 5
```

✔

Answer: two of leverages more than twice the mean leverage. But there is no enough evidence to prove 4,5 are significantly influential.

17.

Applying stepwise algorithm:

```
> m1<-lm(HeatRate~Shafts+RPM+CPRatio+InletTemp+ExhTemp+Airflow+Power)
> m0<-lm(HeatRate~1, data = swiss)
> m2<-step(m0,scope=~Shafts+RPM+CPRatio+InletTemp+ExhTemp+Airflow+Power,
direction="both")
Start:  AIC=478.32
HeatRate ~ 1

          Df Sum of Sq      RSS    AIC
+ RPM      1  80446795 12788995 416.75
```

-1

Start with the multiple regression model and reduce

```
+ InletTemp  1  58543841 34691949 448.68
+ CPRatio    1  45683457 47552333 458.77
+ Airflow    1  39749292 53486498 462.53
+ Power      1  38757766 54478024 463.12
<none>                  93235790 478.32
+ Shafts     1   2105658 91130131 479.59
+ ExhTemp    1     32823 93202967 480.31

Step:  AIC=416.75
HeatRate ~ RPM

            Df Sum of Sq      RSS    AIC
+ CPRatio    1   4995085  7793910 402.90
+ InletTemp  1   3891445  8897550 407.14
+ Shafts     1    950016 11838979 416.28
+ Power      1    853856 11935139 416.54
<none>                  12788995 416.75
+ Airflow    1    392791 12396204 417.75
+ ExhTemp    1      4980 12784015 418.74
- RPM        1  80446795 93235790 478.32

Step:  AIC=402.9
HeatRate ~ RPM + CPRatio

            Df Sum of Sq      RSS    AIC
+ InletTemp  1   1866842  5927068 396.14
+ Power      1   1318310  6475600 398.97
+ Airflow    1   1091714  6702196 400.07
+ ExhTemp    1    835555  6958355 401.27
<none>                   7793910 402.90
+ Shafts     1    234852  7559058 403.92
```

```
- CPRatio    1   4995085 12788995 416.75
- RPM        1  39758423 47552333 458.77

Step:  AIC=396.14
HeatRate ~ RPM + CPRatio + InletTemp

            Df Sum of Sq      RSS    AIC
+ ExhTemp    1    523560  5403508 395.18
<none>                    5927068 396.14
+ Power      1    319598  5607470 396.36
+ Airflow    1    308565  5618503 396.43
+ Shafts     1     33628  5893440 397.96
- InletTemp  1   1866842  7793910 402.90
- CPRatio    1   2970482  8897550 407.14
- RPM        1  19804634 25731702 441.12

Step:  AIC=395.18
HeatRate ~ RPM + CPRatio + InletTemp + ExhTemp

            Df Sum of Sq      RSS    AIC
+ Power      1    699396  4704112 392.74
+ Airflow    1    651517  4751991 393.07
- CPRatio    1     93602  5497110 393.73
<none>                    5403508 395.18
- ExhTemp    1    523560  5927068 396.14
+ Shafts     1      1206  5402302 397.17
- InletTemp  1   1554848  6958355 401.27
- RPM        1  11111263 16514770 428.93

Step:  AIC=392.74
HeatRate ~ RPM + CPRatio + InletTemp + ExhTemp + Power
```

```
          Df Sum of Sq      RSS    AIC
- CPRatio    1     38144  4742256 391.00
<none>                    4704112 392.74
+ Airflow    1     14173  4689939 394.65
+ Shafts     1      5840  4698272 394.70
- Power      1    699396  5403508 395.18
- ExhTemp    1    903358  5607470 396.36
- InletTemp  1   1656870  6360982 400.40
- RPM        1   6797549 11501660 419.35

Step:  AIC=391
HeatRate ~ RPM + InletTemp + ExhTemp + Power

          Df Sum of Sq      RSS    AIC
<none>                    4742256 391.00
+ CPRatio    1     38144  4704112 392.74
+ Airflow    1     27094  4715162 392.82
+ Shafts     1       445  4741811 393.00
- Power      1    754854  5497110 393.73
- ExhTemp    1   4145633  8887889 409.10
- InletTemp  1   6735273 11477529 417.29
- RPM        1   7413766 12156023 419.12
```

Answer: My reduced model is HeaRate vs CPRatio, RPM ,InletTemp.

Fit and comment

−3

18.

```
> full_model <- lm(HeatRate~Shafts+RPM+CPRatio+InletTemp+ExhTemp+Airflow+Power)

> reduced_model <- lm(HeatRate~CPRatio+RPM+InletTemp)

> anova(full_model,reduced_model)

Analysis of Variance Table


Model 1: HeatRate ~ Shafts + RPM + CPRatio + InletTemp + ExhTemp + Airflow +

   Power

Model 2: HeatRate ~ CPRatio + RPM + InletTemp

  Res.Df     RSS Df Sum of Sq     F Pr(>F)

1     24 4681132

2     28 5927068 -4  -1245936 1.597 0.2075
```

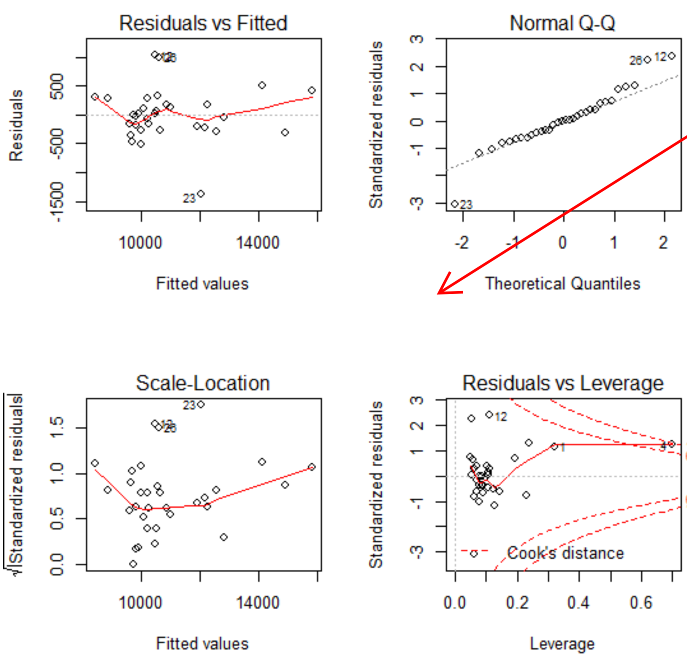
sufficient

Answer: reduced model is better.

19.

```
> par(mfrow=c(2,2))

> plot(reduced_model)
```

The graph generated:



Comment

−3

20.

```
> vif(reduced_model)
  CPRatio       RPM InletTemp 
 1.561515  2.076288  2.124232 
> which(cooks.distance(reduced_model)>1)
named integer(0)
> vif(reduced_model)
  CPRatio       RPM InletTemp 
 1.561515  2.076288  2.124232 
```

Answer: as shown above, cooks distance is less than 1 and vif less than 10.Hence, nothing to worry about.

21.

```
> poly1<-lm(HeatRate~poly(InletTemp,1,raw=T))
> poly2<-lm(HeatRate~poly(InletTemp,2,raw=T))
> poly3<-lm(HeatRate~poly(InletTemp,3,raw=T))
> anova(poly1,poly2)
Analysis of Variance Table

Model 1: HeatRate ~ poly(InletTemp, 1, raw = T)
Model 2: HeatRate ~ poly(InletTemp, 2, raw = T)
  Res.Df      RSS Df Sum of Sq      F  Pr(>F)  
1     30 34691949                              
2     29 28960612  1   5731337 5.7391 0.02327 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
> anova(poly1,poly3)
Analysis of Variance Table

Model 1: HeatRate ~ poly(InletTemp, 1, raw = T)
Model 2: HeatRate ~ poly(InletTemp, 3, raw = T)
```

Res.Df RSS Df Sum of Sq F Pr(>F)

1 30 34691949

2 28 28358735 2 6333214 3.1265 0.05949 .

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

> anova(poly2,poly3)

Analysis of Variance Table

Model 1: HeatRate ~ poly(InletTemp, 2, raw = T)

Model 2: HeatRate ~ poly(InletTemp, 3, raw = T)

Res.Df RSS Df Sum of Sq F Pr(>F)

Adding x^3 does not significantly improve

-1

1 29 28960612

2 28 28358735 1 601877 0.5943 0.4472

Answer: 0.02327 is smallest one. Hence the model with degree 2 is the best model.

significant

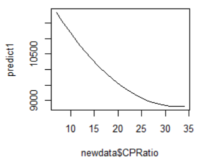
22.

> poly2<-lm(HeatRate~poly(CPRatio,RPM,degree=2,raw=TRUE))

> newdata<-data.frame(RPM=rep(4500,19),CPRatio=seq(7,35,1.5))

> predict1<-predict.lm(poly2,newdata)

> plot(newdata$CPRatio,predict1,type='l')



Answer: When RPM equals to 4900 rpm, and CPRatio is between 7 and 35.

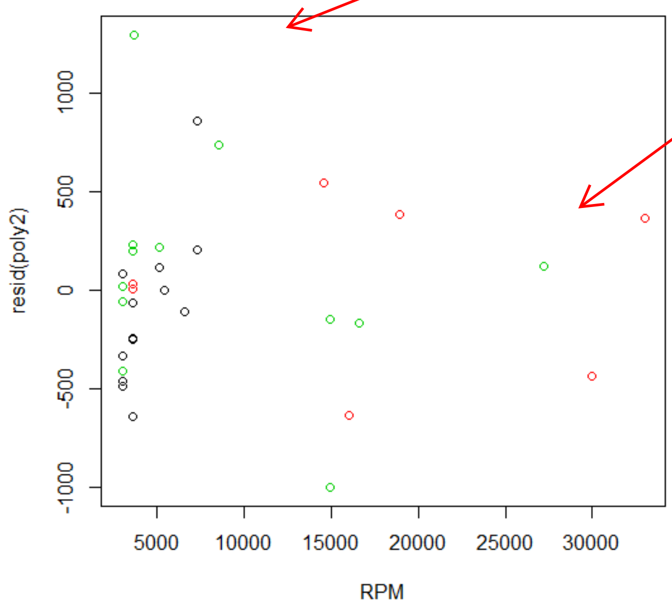
23.

> plot(RPM,resid(poly2),col=Engine)

legend

Comment

−4



24.

> poly_with_engine<-lm(HeatRate~poly(CPRatio,RPM,degree=2,raw=TRUE)+Engine)

> summary(poly_with_engine)

Call:

lm(formula = HeatRate ~ poly(CPRatio, RPM, degree = 2, raw = TRUE) +

Engine)

Residuals:

Min	1Q	Median	3Q	Max
-1048.15	-256.65	3.56	181.36	1138.72

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	1.171e+04	2.242e+03	5.221
poly(CPRatio, RPM, degree = 2, raw = TRUE)1.0	-1.768e+02	1.884e+02	-0.939
poly(CPRatio, RPM, degree = 2, raw = TRUE)2.0	2.213e+00	3.517e+00	0.629

```
poly(CPRatio, RPM, degree = 2, raw = TRUE)0.1  1.851e-01  2.053e-01   0.901
poly(CPRatio, RPM, degree = 2, raw = TRUE)1.1 -3.942e-03  1.015e-02  -0.388
poly(CPRatio, RPM, degree = 2, raw = TRUE)0.2 -1.859e-07  3.413e-06  -0.054
EngineAeroderiv                             2.862e+02  3.371e+02   0.849
EngineTraditional                           3.109e+02  2.578e+02   1.206
                                           Pr(>|t|)
(Intercept)                                2.38e-05 ***
poly(CPRatio, RPM, degree = 2, raw = TRUE)1.0    0.357
poly(CPRatio, RPM, degree = 2, raw = TRUE)2.0    0.535
poly(CPRatio, RPM, degree = 2, raw = TRUE)0.1    0.376
poly(CPRatio, RPM, degree = 2, raw = TRUE)1.1    0.701
poly(CPRatio, RPM, degree = 2, raw = TRUE)0.2    0.957
EngineAeroderiv                                  0.404       Not significant       -2
EngineTraditional                                0.240
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 511.3 on 24 degrees of freedom
Multiple R-squared:  0.9327,    Adjusted R-squared:  0.9131
F-statistic: 47.51 on 7 and 24 DF,  p-value: 1.548e-12
```

Answer: p value: 1.548e-12, hence adding Engine improved this model.

25.

```
> another_poly_with_engine <- lm(HeatRate~poly(CPRatio,RPM,degree=2,raw=TRUE)*Engine)
> anova(another_poly_with_engine,poly_with_engine)
Analysis of Variance Table

Model 1: HeatRate ~ poly(CPRatio, RPM, degree = 2, raw = TRUE) * Engine
Model 2: HeatRate ~ poly(CPRatio, RPM, degree = 2, raw = TRUE) + Engine
  Res.Df     RSS  Df Sum of Sq      F  Pr(>F)
1     14 2218234
```

2   24 6275250 -10 -4057016 2.5605 0.05284 .

--- −3

Comment

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