3.
(a)

```
> aismod<-lm(Wt~Sex+Ht+SSF+Bfat, data=ais)
> summary(aismod)

Call:
lm(formula = Wt ~ Sex + Ht + SSF + Bfat, data = ais)

Residuals:
     Min       1Q   Median       3Q      Max
-13.8944  -4.7384  -0.7099   3.5957  31.6139

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -68.693175  11.579901  -5.932 1.32e-08 ***
Sex         -14.017058   2.008283  -6.980 4.41e-11 ***
Ht            0.759674   0.063985  11.873  < 2e-16 ***
SSF           0.199369   0.071631   2.783  0.00591 **
Bfat          0.004401   0.439623   0.010  0.99202
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.858 on 197 degrees of freedom
Multiple R-squared:  0.7623,    Adjusted R-squared:  0.7574
F-statistic: 157.9 on 4 and 197 DF,  p-value: < 2.2e-16
```

(b) least squares estimates:
Intercept: -68.693175
Sex: -14.017058
Ht: 0.759674
SSF: 0.199369
Bfat: 0.004401

(c) The "Multiple R-Squared" value: 0.7623

(d)

```
> which.max((residuals(aismod)))
163
163
> fitted(aismod)[163]
     163
91.58611
```

Observation 163 has the largest residual, and its fitted value is 91.58611.

(e) According to this model, the estimated difference in weight for a male compared to a female is 14.017058.

(f) Variables that have p-values less than 0.05 are statistically significant, so sex, height and sum of skin folds are statistically significant.

(g)

```
> confint(aismod,level=0.95)
                    2.5 %      97.5 %
(Intercept) -91.52965492 -45.8566953
Sex         -17.97755074 -10.0565662
Ht            0.63348925   0.8858580
SSF           0.05810749   0.3406307
Bfat         -0.86257156   0.8713728
```

(h)

```
> newdata<-data.frame(Sex=1, Ht=170, SSF=60, Bfat=0.12)
> predict(aismod, newdata,interval = "prediction")
       fit      lwr      upr
1 58.39696 40.45794 76.33597
```

(i)

```
> aismod2<-lm(Wt~Sex+Ht, data=ais)
> anova(aismod2, aismod)
Analysis of Variance Table

Model 1: Wt ~ Sex + Ht
Model 2: Wt ~ Sex + Ht + SSF + Bfat
  Res.Df     RSS Df Sum of Sq      F    Pr(>F)
1    199 14548.6
2    197  9266.5  2      5282 56.146 < 2.2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Since the p-value is less than 0.05, we reject the null hypothesis. We conclude that adding sum of skin folds and percent body fat improve the model.

4.

(a)

```
> fuelmod<-lm(FuelC~Income+Pop+Tax,data=fuel2001)
> summary(fuelmod)

Call:
lm(formula = FuelC ~ Income + Pop + Tax, data = fuel2001)

Residuals:
     Min       1Q   Median       3Q      Max
-2696785  -113067    -7653   141991  1668076

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.698e+06  5.820e+05   2.918  0.00539 **
Income      -3.083e+01  1.723e+01  -1.789  0.07998 .
Pop          5.638e-01  1.646e-02  34.258  < 2e-16 ***
Tax         -3.373e+04  1.645e+04  -2.051  0.04587 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 522600 on 47 degrees of freedom
Multiple R-squared:  0.9647,    Adjusted R-squared:  0.9625
F-statistic: 428.3 on 3 and 47 DF,  p-value: < 2.2e-16
```

(b)The p-value of income is 0.07998, which is greater than 0.05, so we fail to reject H_0.

(c) The p-value of F-test at the bottom of the output is less than 0.05, so we reject the null hypothesis.

(d)

```
> fuelmod2<-lm(FuelC~Income+Pop+Tax+Drivers,data=fuel2001)
> summary(fuelmod2)

Call:
lm(formula = FuelC ~ Income + Pop + Tax + Drivers, data = fuel2001)

Residuals:
     Min       1Q   Median       3Q      Max
-1697519  -118108   -17214   122668  1849795

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.231e+06  5.231e+05   2.353 0.022954 *
Income      -2.644e+01  1.512e+01  -1.749 0.086930 .
Pop          1.666e-02  1.403e-01   0.119 0.906006
Tax         -2.204e+04  1.470e+04  -1.499 0.140648
Drivers      6.496e-01  1.657e-01   3.920 0.000292 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 457400 on 46 degrees of freedom
Multiple R-squared:  0.9735,    Adjusted R-squared:  0.9712
F-statistic: 423.2 on 4 and 46 DF,  p-value: < 2.2e-16
```

(e)

```
> anova(fuelmod2, fuelmod)
Analysis of Variance Table

Model 1: FuelC ~ Income + Pop + Tax + Drivers
Model 2: FuelC ~ Income + Pop + Tax
  Res.Df        RSS Df  Sum of Sq      F    Pr(>F)
1     46 9.6235e+12
2     47 1.2838e+13 -1 -3.215e+12 15.368 0.0002923 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Since the p-value is less than 0.05, we reject the null hypothesis. So $\beta_{Drivers} \neq 0$.

(f)

The p-value of Drivers in the summary of the linear model is 0.000292, which is the same as the p-value of that in the analysis of variance table.