Problem 1

(a)

Source	DF	Sum of Squares	Mean Square	F Value	$Pr > F$
Model	1	2624.670184	2624.670184	137.906162	≈ 0
Error	96	1827.099916	19.0322908		
Corrected Total	97	4451.7701			

$$MST = \frac{SSH}{DFH} \quad MSE = \frac{SSE}{DFE} \quad F = \frac{MSH}{MSE}$$
$$Pr > F = 1 - pf(137.906162, 1, 96) = 0 \quad \text{(using R)}$$

(b) The model assumptions are:
Homogeneity of variance- every element of ϵ (error terms) has the same variance
Independence- each element of ϵ is independent of all others
Linearity- expected values of WGHT are linear function of the parameters. $E(y) = X\beta$
Existence - ϵ_i has finite first and second moments.
Gaussian errors- error terms are normally distributed. $\epsilon_i \sim N(0, \sigma_i^2)$

(c) $H_0 : \beta_1 = 0$
Since the p-value is approximately 0, reject the null hypothesis and conclude that average daily exercise time is a significant predictor of weight loss.

(d) The analysis suggests that neither variable is significant since both of the p-values are greater than $\alpha = .05$
This occurs in this added-last test because, after adjusting for running mileage, exercise time does not provide any additional useful information. Running mileage and exercise time appear highly correlated.

Problem 3 b

(i)

(ii) $H_0 : Residualsarenormallydistributed$
$H_1 : Residualsarenotnormallydistributed$
$\alpha = .05$
p-value=.242
Since p-value$>\alpha$ fail to reject H_0 thus there is not sufficient evidence to conclude residuals are not normally distributed. Therefore we have met the Gaussian assumption

(iii) Based on the histogram, the studentized residuals appear normally distributed, centered around 0.

(iv) Based on the plot of the studentized residuals vs the predicted values, there appears to be homoscedasticity, and the mean of the residuals appears constant.