1. a c
2. b d
4. a b
5. c d

6. Step 1: Use Simple Linear Regression Model with "Payments made" as the predictor and "APR at Payment" as the response, because we hypothesize that there's linear relationship.
Step 2: Randomly select independent data for the model. Fit data into Simple Linear Regression Model by finding the best estimation for the intercept and the slope. In R, we can use lm() function and read the intercept and read the slope of one variable from output.
Step 3: Use ggplot in r to plote residual plot and normal quantile plot. Use residual vs. fitted-value plot to observe whether "constant spread of error", "zero mean of error" and "linearity" of this model. Use normal-quantile plot to check "normality" of the model.
Based on assumptions above, we can test the model with F-test to understand whether the relationship is significant enough.
Step 4: Use the model for estimating APR. We can compute CI for mean response or Prediction Interval for estimating possible payment range with error.

8 (a) Linearity: Check the assumption of Linearity with residual plot. Provided there're roughly the same amount of data above and below $x=0$. There's no alternative trend observed other than Linear. Therefore Linearity is met.

Zero mean: zero mean assumption is met as we are using the model for least squared error. On the residual plot, we observe that the error is roughly symmetric about zero line and no general shift from zero. Therefore, zero mean is met.

Constant Variance: constant variance assumption is checked with residual plot. The variance of error is generally uniform from left to right. However, we might still get noted that data is more clustered and negative on the left.

Normality: Check the normality assumption with normal quantile plot. The data points are generally along the line, which means that data is roughly normal. However, we still need to get noted that the two data point on two ends are quite off the line, which might imply slight skewness from normal.

We cannot make any justification for independence or randomness from the plot because they're depend on the use of research methods, but we cannot infer from the plots.

8. (b) No. Because change in unit is a linear transformation along x axis (predictor). As predictors are studentized for plotting, there's no change just by scaling the predictor. As there's no change in y, there's no change in residuals. Therefore, there's no change in the two plots.

9. (a) residual standard error 30.84 calories

(c) In this question, we want to know whether the slope is statistically different from 100.

We can use 1-sample t-test for the slope:

$H_0: \beta_1 = 100$ The null hypothesis states that β_1 equals to 100.

$H_A: \beta_1 \neq 100$ The alternative hypothesis states that β_1 does not equal to 100.

Conduct the test by looking at the p-value of $\hat{\beta}_1 = 80$. If the p-value $> \alpha = 0.05$, we might say that the slope is not significantly different from 100 and accept H_0.

Otherwise, $\hat{\beta}_1$ is statistically different from 100 if p-value $< \alpha = 0.05$, there is statistical difference between the slope and 100.

(d) . P-value for test of $\hat{\beta}_1$ is 0.00225, which is very small. This mean that $\hat{\beta}_1$ is significantly different from 0 or no linear relationship. Therefore the linear model is a good fit.

"Multiple R-squared" or R^2 is 0.4313. The coefficient of determination is not close to 0, but it's neither close to 1. This means that the correlation is not very strong and might not be optimal fit.