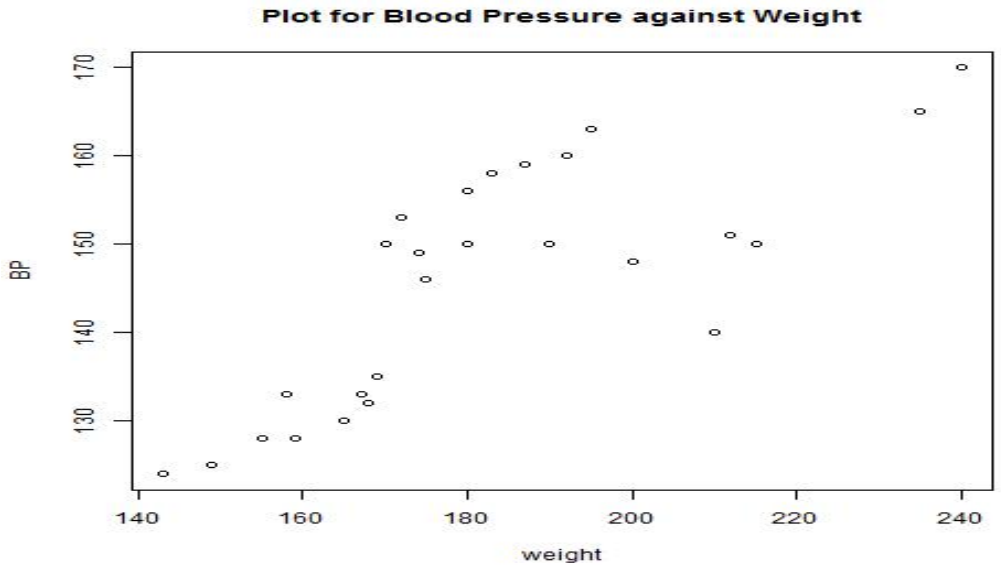
# Stat 6021: Guided Question Set 2 Solutions

1. The estimated linear regression equation is $\hat{BP} = 69.1044 + 0.4194 weight$.

```
Call:
lm(formula = BP ~ weight)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 69.10437   12.91013   5.353 1.71e-05 ***
weight       0.41942    0.07015   5.979 3.59e-06 ***
---
Residual standard error: 8.681 on 24 degrees of freedom
Multiple R-squared:  0.5983,    Adjusted R-squared:  0.5815
F-statistic: 35.74 on 1 and 24 DF,  p-value: 3.591e-06
```




2. Correlation is 0.773. There are two possible answers regarding the reliability and interpretation of this value:

- The interpretation that we have a fairly strong linear relationship between the two variables is reliable if one believes a linear relationship exists, based on the scatterplot.
- The correlation should not be interpreted and is unreliable if one believes there is a curved relationship between the two variables, based on the scatterplot.

Note: Should always use a correlation with a scatterplot to assess if a linear relationship is reasonable. The fact that the correlation is close to 1 is not proof that the correlation is reliable, as non linear relationships can result in high correlations.

3. The predicted blood pressure increases by 0.4194mmHg when weight increases by one pound, for young males. The corresponding 95% confidence interval for the change in blood pressure for a one pound increase in weight is (0.2746, 0.5642).

```
                2.5 %     97.5 %
(Intercept) 42.4591756 95.7495700
weight       0.2746281  0.5642023
```

4. $H_0 : \beta_1 = 0, H_a : \beta_1 \neq 0$. The $t$ statistic is 5.979 with a p-value less than 0.05, so we reject the null hypothesis. Our data support the claim that blood pressure and weight are linearly related for young males. Note: The $F$ statistic of 35.74 is the squared of the $t$ statistic, with the same p-value. This is the case in SLR with 1 predictor.

5. Yes the results are consistent, since we rejected the null hypothesis and 0 lies outside the 95% CI for the slope. Note: A $(1 - \alpha)x100\%$ CI is consistent with a two-sided hypothesis test at significance level $\alpha$.

6. The estimated mean systolic blood pressure for young males who weighs 200 pounds is 152.9874 mmHg. The 95% CI for the mean systolic blood pressure of young males who weigh 200 pounds is (148.6489, 157.3259) mmHg.

```
> newdata <- data.frame(weight=200)
> predict.lm(result, newdata, level=0.95, interval="confidence")
       fit      lwr      upr
1 152.9874 148.6489 157.3259
```

7. We have 95% confidence the blood pressure of a young male who weighs 200 pounds is between 134.5532 and 171.4216 mmHg.

```
> predict.lm(result, newdata, level=0.95, interval="prediction")
       fit      lwr      upr
1 152.9874 134.5532 171.4216
```

8. $H_0 : \beta_1 = 0.35, H_a : \beta_1 > 0.35$.

   The $t$ statistic is $t = \frac{\hat{\beta}_1 - 0.35}{SE_{\hat{\beta}_1}} = \frac{0.4194 - 0.35}{0.0702} = 0.9896$.

   The pvalue is found using 1 - pt(0.9896,24) = 0.1661 which is greater than 0.05. We fail to reject the null hypothesis. The data do not support the claim that the predicted blood pressure increases by more than 0.35 mmHg per pound increase in weight.