Yahya Alhinai

EE5373

October 1st, 2019

One Factor Linear Regression Model

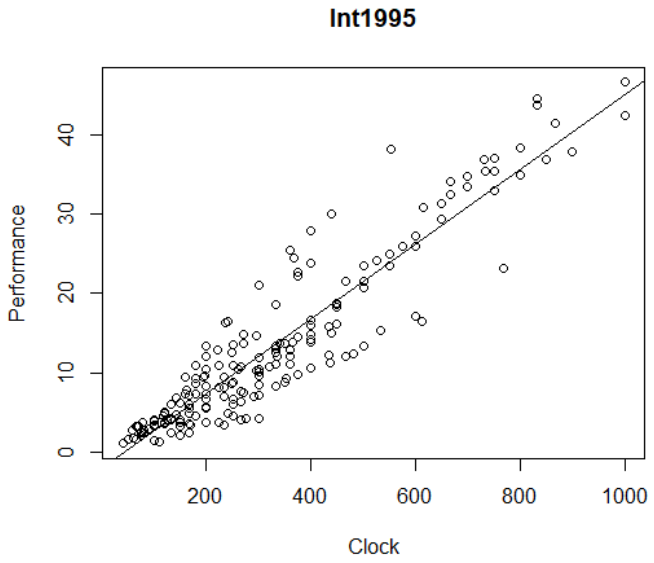
INT1995:

- **Coefficients for the regression model:**

$$perf = -2.0404 + 0.0471 * clock$$

	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept	-2.040420	0.536003	-3.807	0.000192
Clock	0.047096	0.001372	34.316	< 2e-16

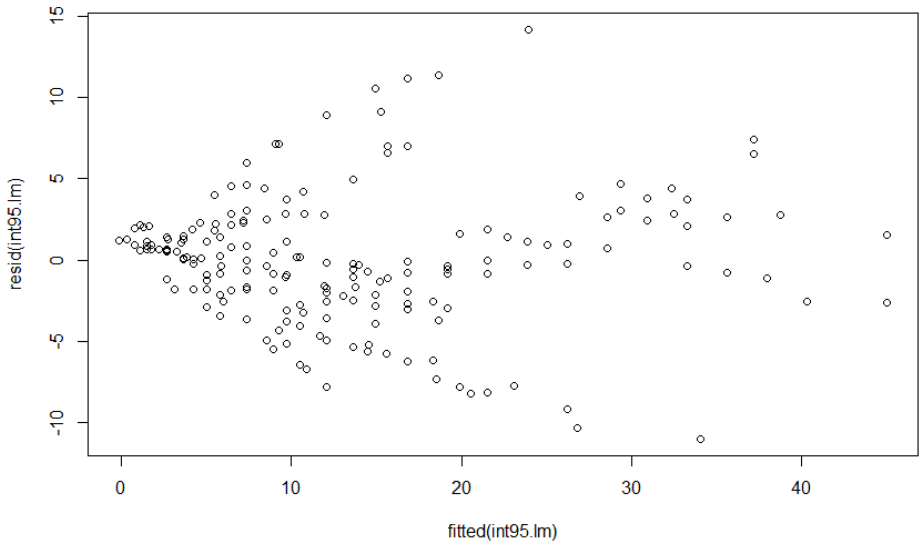
- **Scatter plot with regression line for the model:**



- **Explanation of quality analysis:**
  - **P-values:**
    - This value indicates that the probability that the intercept is not relevant (0.000192) which is a tiny value. Meaning the chance that this specific intercept value is not relevant to the model is 0. 0192%.
    - The p-value of the coefficient is also tiny (almost zero). Which gives a percentage of almost 100% chance that this specific is relevant to the model.
  - **Residual:**

- the residual for this regression model has a good fit because the fact that it is normally distributed around a mean of zero following Gaussian distribution. Also, its median has a value of 0.0299 which is that is roughly balanced around the mean of zero.

- **standard errors:**
  - The standard error is for the Intercept is 3 times smaller which is not ideal. On the other hand, the standard error for the clock is 34.3 times smaller which can be descried as a good model. This means that there is relatively little variability in the slope estimated.
  - A good Residual standard error should be about 1.5 times the standard error which this model does not fall into that description. Meaning that the residuals are not distributed normally.
- **R^2 values:**
  - Values of R^2 that are closer to one indicate a better-fitting model. Our model has a value of 0.8655 which means that the model explains 86.55% of the data's variation.
- **residual analysis**
  - residual analysis shows if the residual values represents a quality model. Residual values represent how good our model in compared to the actual measured value: the closer to zero the better. The figure below says that the higher the clock rate, the farther it deviates from zero which tells us the prediction the model become less accurate as clock rate increases.



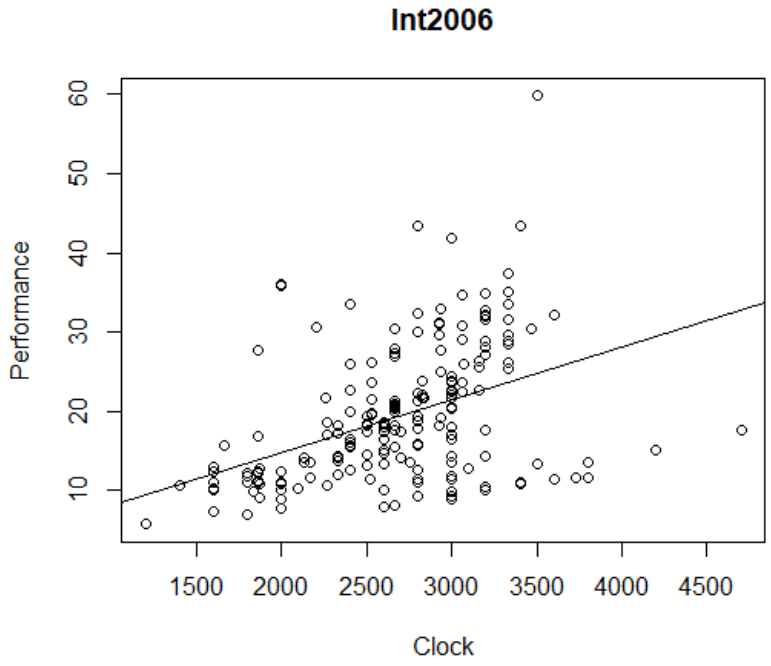
INT2006:

- **Coefficients for the regression model:**

$$perf = 1.514316 + 0.006648 * clock$$

	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept	1.5143157	2.7126572	0.558	0.577
Clock	0.0066481	0.0009971	6.668	2.43e-10

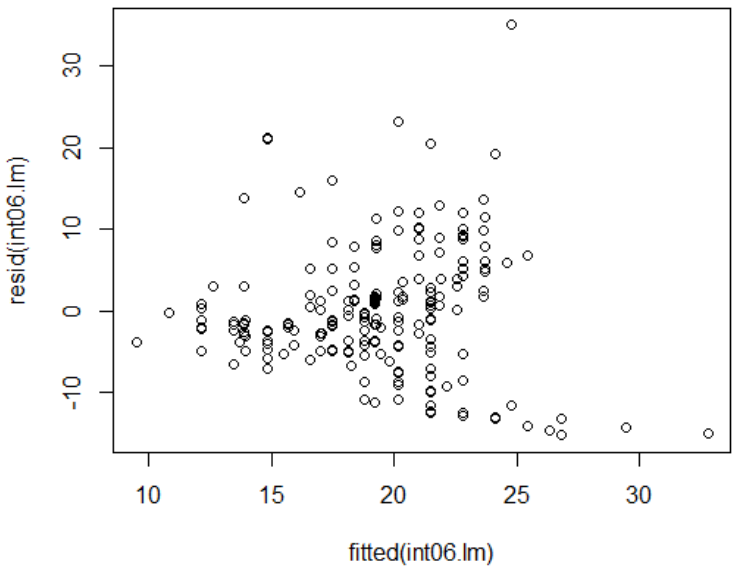
- **Scatter plot with regression line for the model:**



- **Explanation of quality analysis:**
  - **P-values:**
    - The probability that the intercept is not relevant is 0.577 which is a extremally high. Meaning the chance that this specific intercept value is not relevant to the model is more than 57.7%.
    - The p-value of the coefficient is tiny (2.43e-10) which gives a percentage of almost 0% chance that this specific is not relevant to the model.

  - **Residual:**
    - the residual for this regression model is decently good as the value of the median is shifted to the negative value (-1.088). The ideal median would be centered at zero as it follows Gaussian distribution.

- **standard errors:**
  - The standard error is for the Intercept is 0.558 times smaller and the standard error is for the clock is 6.6 times smaller which can be descried as a decently good model as it has to be 5 to 10 time smaller to be descried as good.
  - A good Residual standard error should be about 1.5 times this standard error which is close enough in this model. This means that the residuals are distributed normally.
- **R^2 values:**
  - Values of R^2 that are closer to one indicate a better-fitting model. Our model has a value of 0.1804 which means that the model explains 18.04% of the data's variation which is almost 1 out of 5 data points are expressed by this model.
- **residual analysis**
  - The figure below says that no matter the clock rate is, residual values has almost the same deviation from zero which tells us the prediction the model has almost nothing to so with the clock rate. Except after the clock rate 25, it shows clear leaning towards negative values which indicate that the regression model predicted a value that was too large.



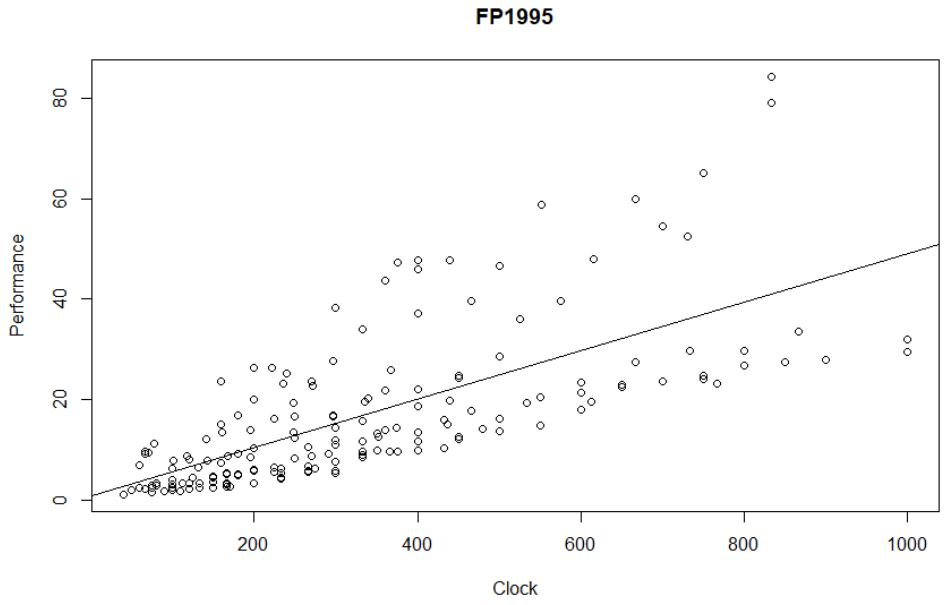
FP1995:

- **Coefficients for the regression model:**

$$perf = 0.79788 + 0.04821 * clock$$

	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept	0.797878	1.516272	0.526	0.599
Clock	0.048208	0.003803	12.677	<2e-16

- **Scatter plot with regression line for the model:**

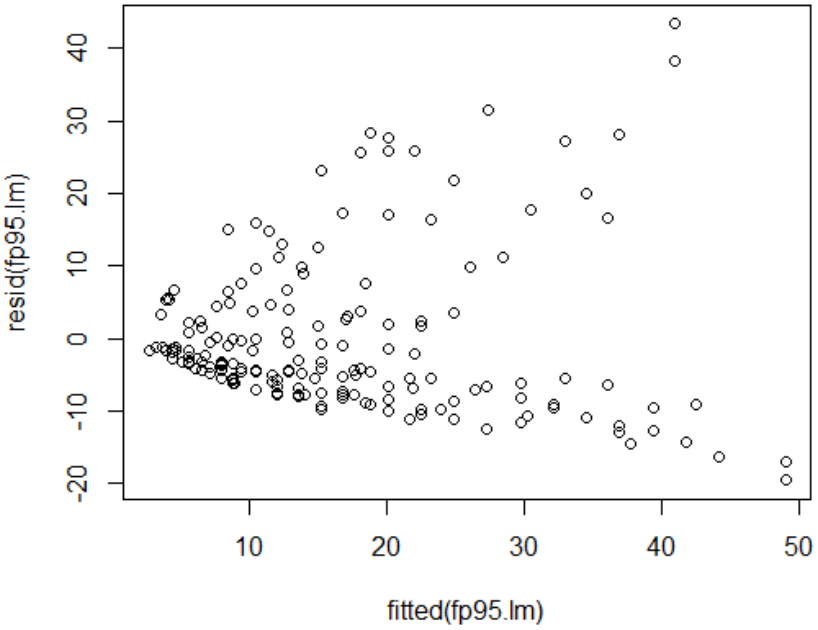


- **Explanation of quality analysis:**
  - **P-values:**
    - The probability that the intercept has a percentage of almost 60%. Meaning the chance that this specific intercept value is not relevant is 6 out 10 data points.
    - The p-value of the coefficient is also almost zero. Which gives a percentage of almost 100% chance that this specific is relevant to the model.
  - **Residual:**
    - the residual for this regression model does not has a good presentation. The normal distribution is not around zero. It is shifted to the negatively (-3.2). It's also not symmetrical which deviate from Gaussian distribution.
  - **standard errors:**
    - The standard error is for the Intercept is 0.57 times smaller. The standard error for the clock is 12.6 times smaller which can be descried as a good

model. This means that there is relatively very little variability in the slope estimated.

- A good Residual standard error should be about 1.5 times the standard error which is good enough the case in this model. It almost has 2 times the standard error meaning that the residuals are almost distributed normally.

- **R^2 values:**
  - Our model has a value of 0.4859 which means that the model explains 48.59% of the data's variation. Meaning, half of the data points are not described by the regression model
- **residual analysis**
  - residual analysis of FP1995 follows the same trend as INT1996 residual analysis. The figure below says that the higher the clock rate, the farther it deviates from zero which tells us the prediction the model become less accurate as clock rate increases. The closer residual to zero the better.



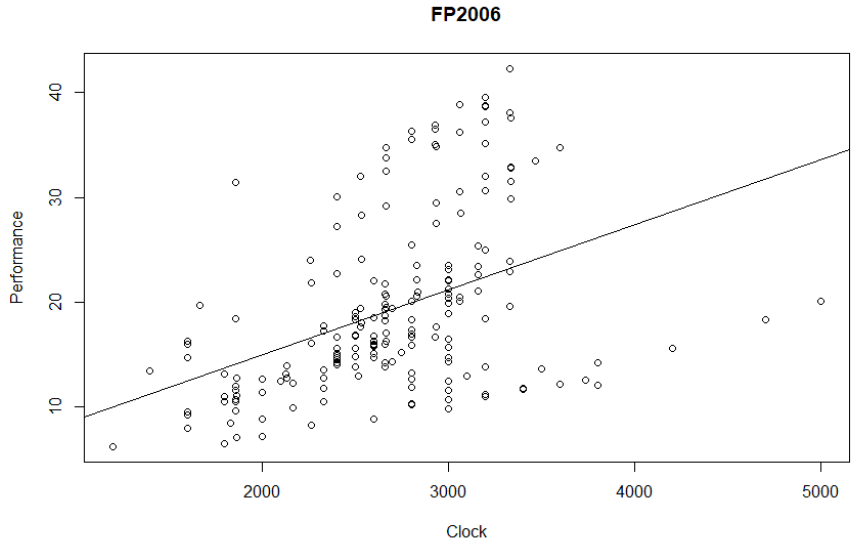
FP2006:

- **Coefficients for the regression model:**

$$perf = 2.491537 + 0.006224 * clock$$

	Estimate	Std. Error	t-value	Pr(>|t|)
Intercept	2.4915371	2.6596790	0.937	0.35
Clock	0.0062239	0.0009696	6.419	1.06e-09

- **Scatter plot with regression line for the model:**



- **Explanation of quality analysis:**
  - **P-values:**
    - This value indicates that the probability that the intercept is not relevant is 0.35. Meaning the chance that this specific intercept value is not relevant to the model is 35% almost one-third of the points.
    - The p-value of the coefficient is almost zero. The relevancy of this model is almost 100%.
  - **Residual:**
    - the residual for this regression model is decently good. The median is slightly shifted from being centered at zero. Not ideally following Gaussian distribution but it has some of its attribute.
  - **standard errors:**
    - The standard error is for the Intercept is 93% the same as the estimated value. The standard error for the clock is 6.4 times smaller which is good enough because it must fall between 5 to 10 times smaller. This means that there is relatively little variability in the slope estimated.

- A good Residual standard error should be about 1.5 times the standard error. Because the model doesn't have a good symmetry, the poison of Q1 and Q3 are not symmetrical which indicates that the residuals are not distributed normally.

- **R^2 values:**
  - Our model has a value of 0.1774 which is very low. It means that this model only explains 17.74% of the data's variation.
- **residual analysis**
  - The figure below shows no or little correlation between the clock rate and the deviation of residual from zero. Except after the clock rate 25, it shows clear leaning towards negative values which indicate that the regression model predicted a value that was too large.