

ANOVA/Regression I
Solutions to Session 2 : Multiple Regression

1.

Model	Variables Used	SSR	d.f.	SSE	d.f.	Overall F	p	Adj. R ²
1	HGT	588.9225	1	299.3275	10	19.67	0.0013	0.6293
2	AGE	526.3929	1	361.8571	10	14.55	0.0034	0.5519
3	HGT, AGE	692.8226	2	195.4274	9	15.95	0.0011	0.7311
4	HGT, AGE, (AGE) ²	693.0605	3	195.1895	8	9.47	0.0052	0.6978

2.

$$F(\text{AGE} \mid \text{HGT}) = (692.8226 - 588.9225) / (195.4274 / 9) \\ = (299.3275 - 195.4274) / (195.4274 / 9) = 4.78$$

If you look at the **anova, sequential** output of **model 3** and look at the *age* row you will see the F-test = 4.78 and a corresponding p-value = 0.0565. Notice that the t-test p-value for β_2 (the regression coefficient associated with age, is 0.056, and $T^2 = (2.187)^2 = 4.78 = F$. Since the p-value of $F(\text{AGE} \mid \text{HGT})$ is 0.056 which is greater than $\alpha = 0.05$, the addition of AGE doesn't significantly contribute to the model.

3. To answer the same question about $(\text{AGE})^2$ after controlling both height and age, the *partial Type I F-test* $F((\text{AGE})^2 \mid \text{HGT}, \text{AGE}) = 0.01$ with corresponding p-value = 0.924 which is not significant. Thus, even though $(\text{AGE})^2$ might be significant as a single predictor of weight, it is not significant after controlling for height and age. So, a quadratic relationship between weight and age is probably not born out by the data.

4.

$$\begin{aligned} \text{SS}((\text{AGE})^2 \mid \text{HGT}, \text{AGE}) &= 0.24 && \text{from model 4} \\ \text{SS}(\text{AGE} \mid \text{HGT}, \{(\text{AGE})^2\}) &= 3.41 && \text{from model 5} \\ \text{SS}(\text{HGT} \mid \text{AGE}, (\text{AGE})^2) &= 166.58 && \text{from model 6} \end{aligned}$$

5.

$$\begin{aligned} F((\text{AGE})^2 \mid \text{HGT}, \text{AGE}) &= 0.24 / (195.19 / 8) = 0.01 && \text{p-value} = 0.924 \\ F(\text{AGE} \mid \text{HGT}, (\text{AGE})^2) &= 3.41 / (195.19 / 8) = 0.14 && \text{p-value} = 0.718 \\ F(\text{HGT} \mid \text{AGE}, (\text{AGE})^2) &= 166.58 / (195.19 / 8) = 6.83 && \text{p-value} = 0.031 \end{aligned}$$

6. According the above results, *HGT* seems to be the most important predictor of weight since even after controlling for both *AGE* and $(\text{AGE})^2$ the *partial Type III F-test* is significant ($0.031 < \alpha = 0.05$).