

## Multicollinearity Example

n = 25 males; height is in inches, Rtfoot and Leftfoot are foot lengths in centimeters

### Correlation matrix:

	Height	LeftFoot	RtFoot
Height	1.0000000	0.5466786	0.5345347
LeftFoot	0.5466786	1.0000000	0.9078141
RtFoot	0.5345347	<b>0.9078141</b>	1.0000000

Note the strong correlation between the feet

---

### Left foot only:

```
lm(formula = Height ~ LeftFoot, data = MaleFeet)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	44.0701	8.8493	4.980	4.9e-05 ***
LeftFoot	0.9986	0.3189	<b>3.131</b>	<b>0.00469</b> **

Clearly, knowing Left Foot helps predict height.

Residual standard error: **2.859** on 23 degrees of freedom

Multiple R-squared: 0.2989, Adjusted R-squared: **0.2684**

F-statistic: 9.804 on 1 and 23 DF, p-value: 0.004689

---

### Right foot only:

```
lm(formula = Height ~ RtFoot, data = MaleFeet)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	46.8408	8.2224	5.697	8.43e-06 ***
RtFoot	0.8964	0.2955	<b>3.033</b>	<b>0.00591</b> **

Clearly, knowing Right Foot helps predict height.

Residual standard error: **2.885** on 23 degrees of freedom

Multiple R-squared: 0.2857, Adjusted R-squared: **0.2547**

F-statistic: 9.201 on 1 and 23 DF, p-value: 0.00591

---

### Both feet:

```
lm(formula = Height ~ LeftFoot + RtFoot, data = MaleFeet)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	43.9334	8.9983	4.882	7.01e-05 ***
LeftFoot	0.6379	0.7730	0.825	<b>0.418</b>
RtFoot	0.3647	0.7096	0.514	<b>0.612</b>

Left Foot is not significant

Right Foot is not significant

---

Residual standard error: **2.905** on 22 degrees of freedom

Multiple R-squared: 0.3072, Adjusted R-squared: **0.2442**

F-statistic: 4.877 on 2 and 22 DF, **p-value: 0.01765** The combination is significant

---

### Anova for model with both, with left foot entered first

Response: Height

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
LeftFoot	1	80.106	80.106	9.4900	0.005468 **
RtFoot	1	2.230	2.230	0.2642	0.612382
Residuals	22	185.704	8.441		

Left foot is significant alone

Right foot not significant, given

left foot is already there.

---

### Both feet, Anova with order of entry reversed:

Response: Height

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
RtFoot	1	76.586	76.586	9.0730	0.006411 **
LeftFoot	1	5.749	5.749	0.6811	0.418060
Residuals	22	185.704	8.441		

Right foot is significant alone

Left foot not significant, given

right foot is already there.

---

### Variance Inflation Factors:

```
> library(car)
> vif(Both)
LeftFoot RtFoot
5.685903 5.685903
```