Agenda

1. two-way ANOVA
2. Interaction plots
3. two-way ANOVA with interaction

	Adam	Brenda	Cathy	Dave	Emily	Mean
Exam 1	62	94	68	86	50	72
Exam 2	87	95	93	97	63	87
Exam 3	74	86	83	70	28	68
Exam 4	77	89	73	79	47	73
Mean	75	91	79	83	47	75

Two-way ANOVA Simple block design has two factors with exactly one data value (observation) in each combination of the factors

Factor A has K levels, Factor B has J levels, so $n = KJ$ values.

$$Y = \mu + \alpha_k + \beta_j + \epsilon$$

To compute the ANOVA, find the mean for each treatment (row means), each block (column means) and grand mean. Partitition the SST into three pieces:

$$\begin{aligned} SST &= SSA + SSB + SSE \\ SSE &= \sum (y_i - \bar{y})^2 = (n-1)s_Y^2 \\ SSA &= \sum J(\bar{y}_k - \bar{y})^2 \\ SSB &= \sum K(\bar{y}_j - \bar{y})^2 \\ SSE &= SST - SSA - SSB \end{aligned}$$

If you get the R output, it will look something like this

	df	SS	MS	F	p-value
Trts/A	K-1	SSA	SSA/(K-1)	MS/MSE	
Block	J-1	SSB	SSB/(J-1)	MSB/MSE	
Error	(K-1)(J-1)	SSE	SSE/(K-1)(J -1)		
Total	n-1	SST			

This tests two hypotheses:

$$\begin{aligned} H_0: &\quad \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0 \\ H_A: &\quad \text{at least one } \alpha_k \neq 0 \end{aligned}$$

$$\begin{aligned} H_0: &\quad \beta_1 = \beta_2 = \cdots = \beta_j = 0 \\ H_A: &\quad \text{at least one } \beta_j \neq 0 \end{aligned}$$

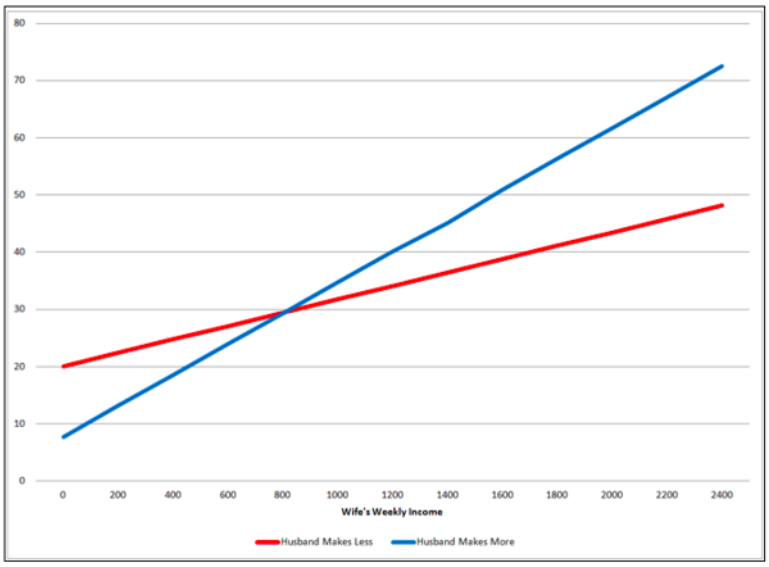


Figure 1: Figure from the Atlantic article, "Emasculated Men Refuse to Do Chores–Except Cooking."

Interaction plots A common way to visualize the interaction between two categorical variables is with an interaction plot.

As an example, consider Figure 1, from `https://www.theatlantic.com/health/archive/2016/10/the-only-chore-men-will-do-is-cook/505067/`

This is an example from regression (so the x-axis has meaning) but we also commonly see them for ANOVA.

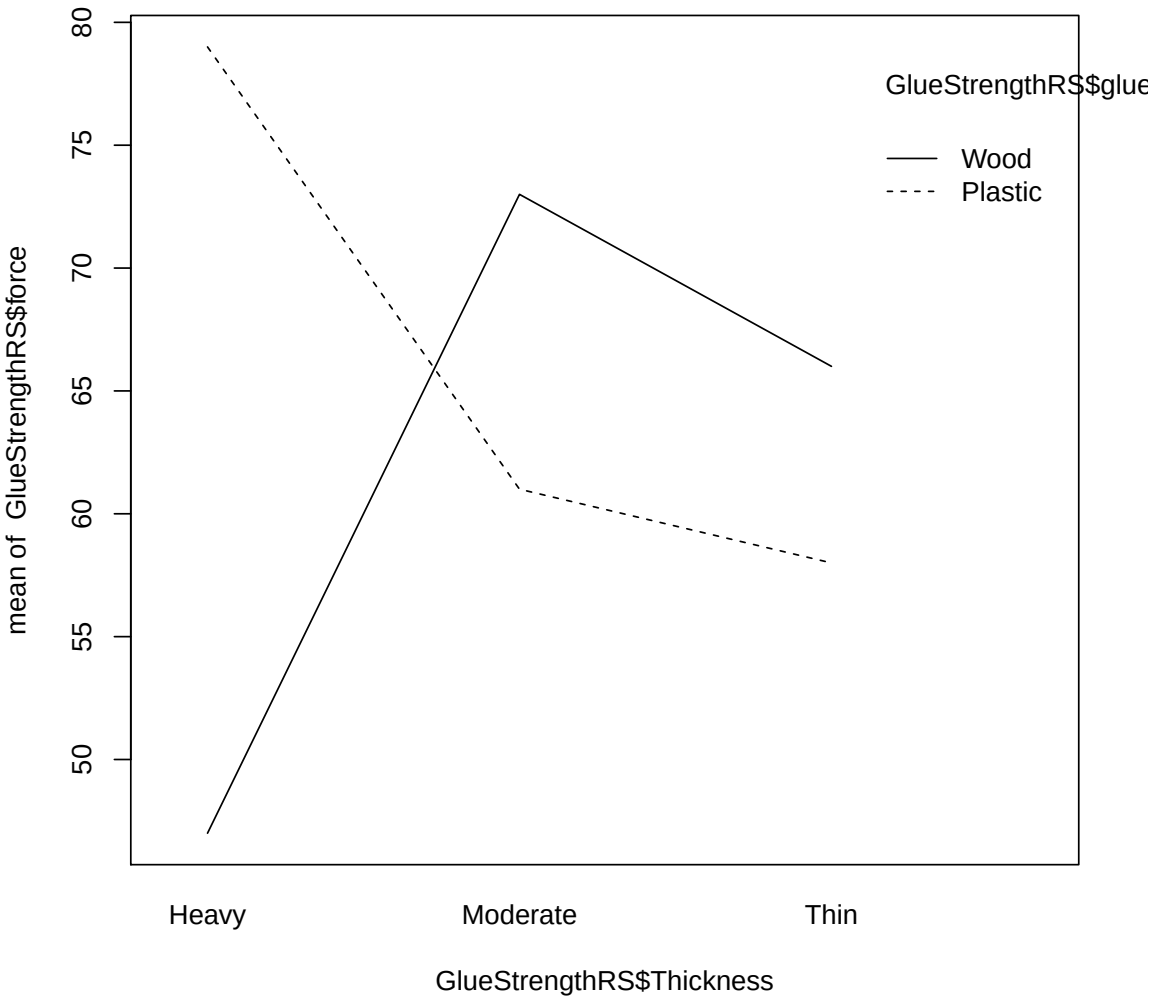
Example: glue strength

```
library(tidyverse)

## -- Attaching packages ---------------------------------------------------------------------
tidyverse 1.2.1 --
## v ggplot2 3.0.0     v purrr   0.2.5
## v tibble  1.4.2     v dplyr   0.7.7
## v tidyr   0.8.1     v stringr 1.3.1
## v readr   1.1.1     v forcats 0.3.0
## -- Conflicts ------------------------------------------------------------------------------
tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

GlueStrength <- data.frame(Plastic = c(52,64,67,55,86,72), Wood = c(72,60,78,68,43,51), Thickness = c(r
GlueStrengthRS <- GlueStrength %>% gather(glue,force,-Thickness)
```

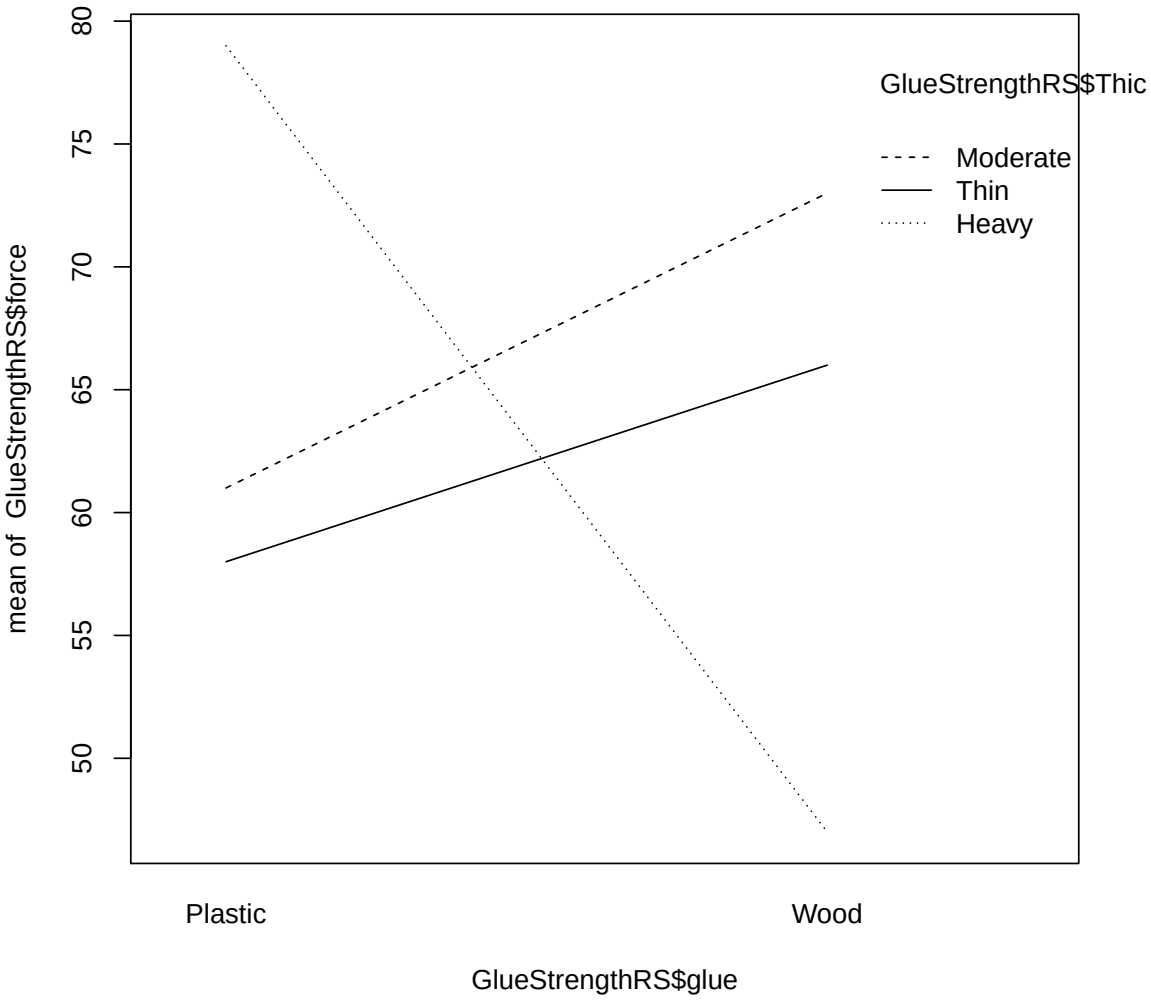
```
GlueStrengthRS %>%
  group_by(glue) %>%
  summarize(mean(force))

## # A tibble: 2 x 2
##   glue    `mean(force)`
```

```
##   <chr>           <dbl>
## 1 Plastic            66
## 2 Wood               62
```

```
interaction.plot(GlueStrengthRS$Thickness, GlueStrengthRS$glue, GlueStrengthRS$force)
```




```
interaction.plot(GlueStrengthRS$glue, GlueStrengthRS$Thickness, GlueStrengthRS$force)
```




```
a1 <- aov(force~Thickness+glue+Thickness*glue, data=GlueStrengthRS)
summary(a1)

##                Df Sum Sq Mean Sq F value Pr(>F)
## Thickness       2     56      28   0.424 0.6725
## glue            1     48      48   0.727 0.4265
## Thickness:glue  2   1184     592   8.970 0.0157 *
## Residuals       6    396      66
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

$$Y = \mu + \alpha_k + \beta_j + \gamma_{kj} + \epsilon$$

$$SST = SSA + SSB + SSAB + SSE$$