ANOVA Part 2

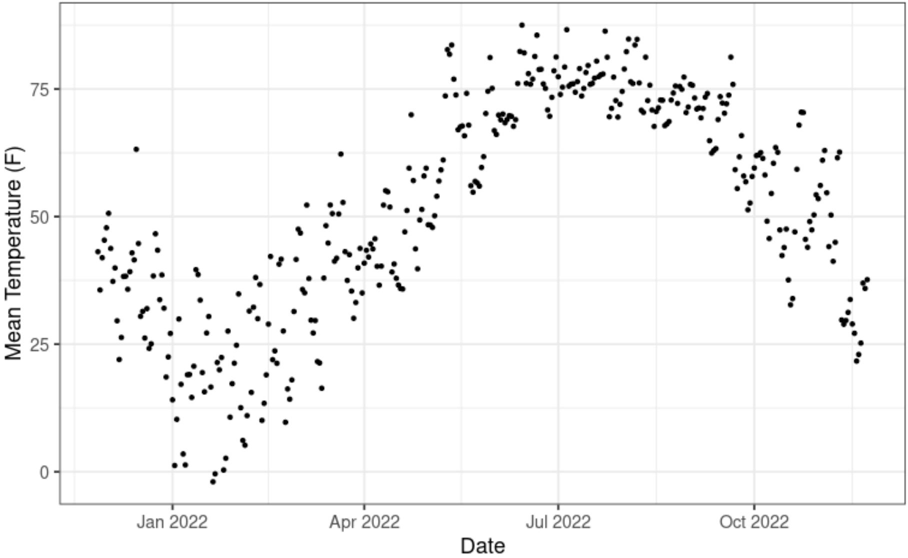
Grinnell College

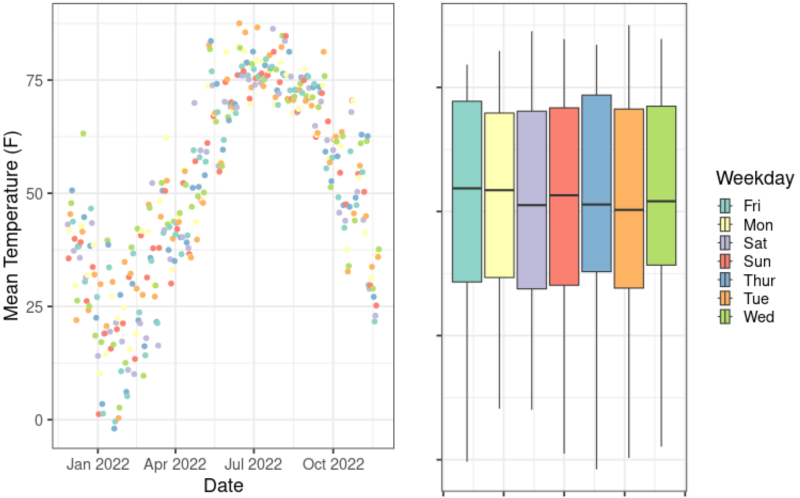
April 27, 2025

Review

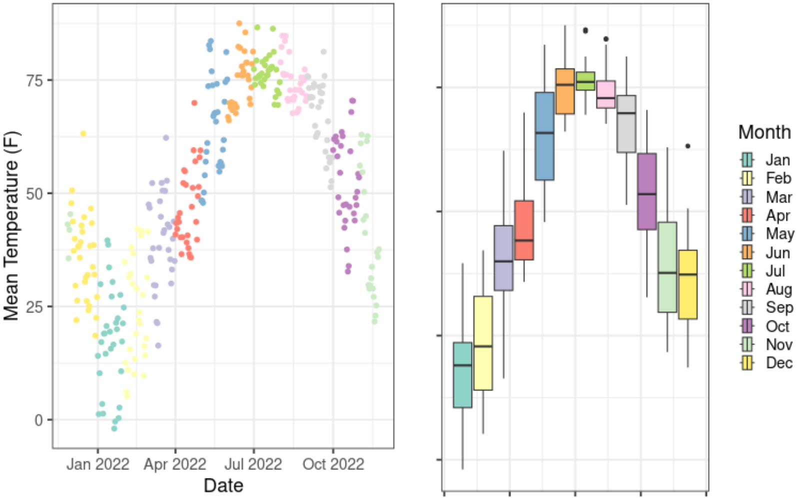
Suppose we had daily average temperatures for Grinnell for the period of one calendar year and we are allowed to choose one categorical variable to help us predict it.

- What makes day of the week a poor candidate for predicting mean values?
- What makes month of the year a good candidate for predicting mean values?
- What kind of attributes make a categorical variable a good or bad candidate? (general conversation question)


Mean Temperature (F)
75
50
25
0
Jan 2022
Apr 2022
Jul 2022
Oct 2022
Date



	Df	Sum Sq	Mean Sq	F value	*p*-value
Weekday	6	342.71	57.12	0.12	0.9939
Residuals	355	168524.83	474.72		



	Df	Sum Sq	Mean Sq	F value	*p*-value
Month	11	138048.06	12549.82	142.52	<0.0001
Residuals	350	30819.48	88.06		

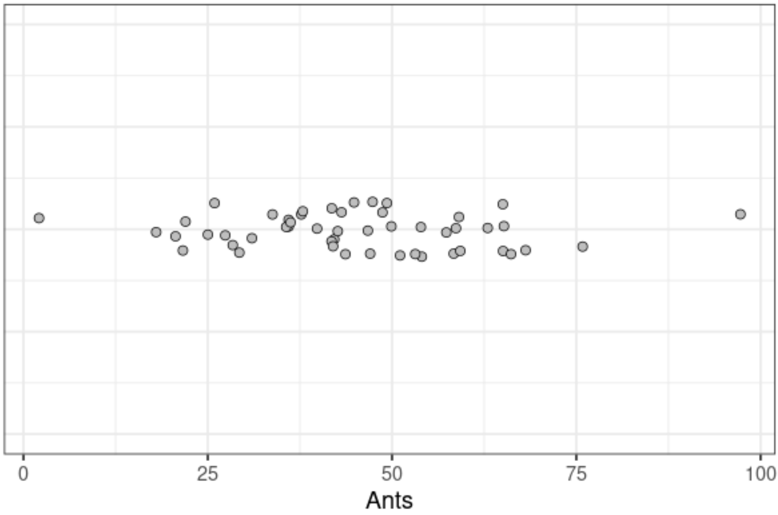
Formulas

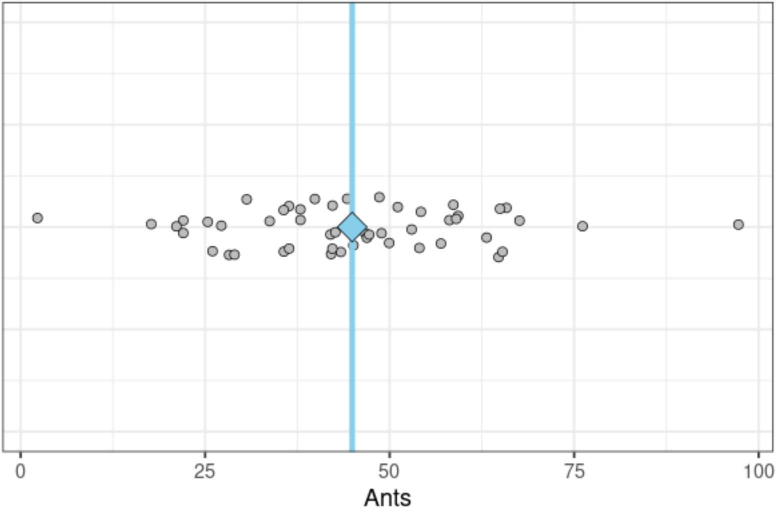
$$\underbrace{\sum_{i}^{n}(x_{ij}-\overline{x})^2}_{\text{SST}} = \underbrace{\sum_{i=1}^{n}(x_{ij}-\overline{x}_j)^2}_{\text{SSE}} + \underbrace{\sum_{j=1}^{k} n_j(\overline{x}_j-\overline{x})^2}_{\text{SSG}}$$

- SST = SSE + SSG
- SSE = sum of squares *within groups*
- SSG = sum of squares *between groups*
- $MSG = \frac{SSG}{k-1}$
- $MSE = \frac{SSE}{n-k}$
- $F = \frac{MSG}{MSE} = \left(\frac{n-k}{k-1}\right)\frac{SSG}{SSE}$

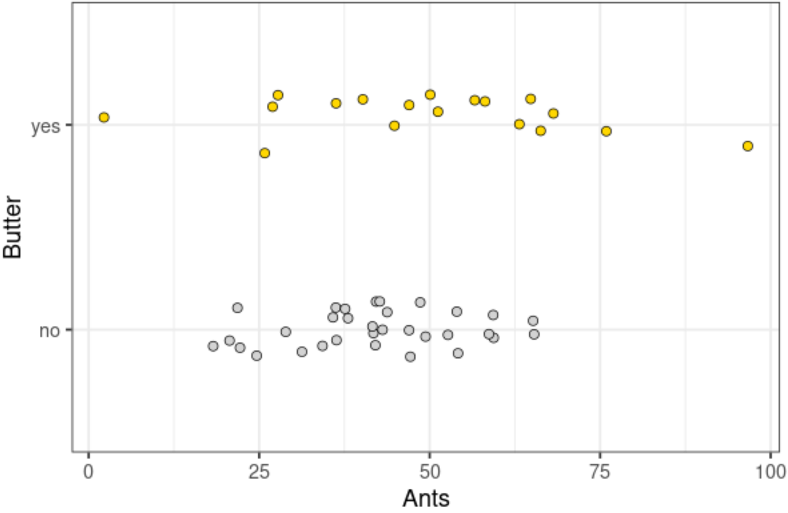
Source	df	Sum Sq	Mean Sq	F value	Pr(>F) / *p*-value
Group	k-1	SSG	$MSG = \frac{SSG}{k-1}$	$F = \frac{MSG}{MSE}$	Upper tail
Error	n-k	SSE	$MSE = \frac{SSE}{n-k}$		
Total	n - 1	SSTotal			

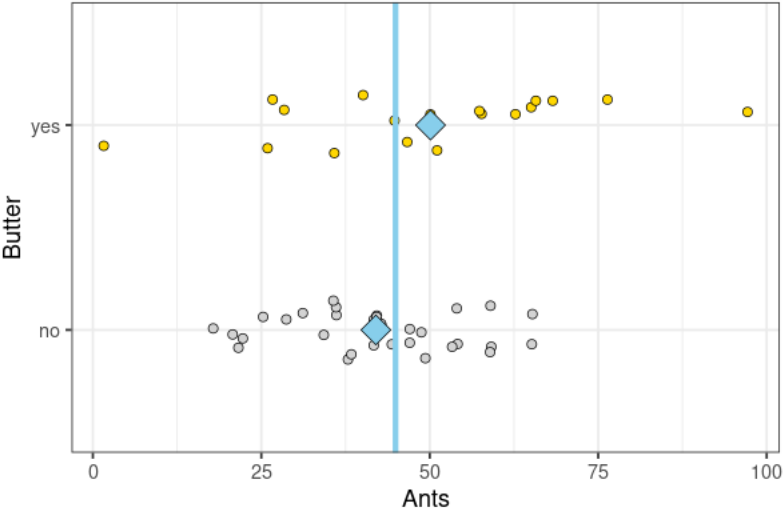
Ants and sandwiches


0
25
50
75
100
Ants

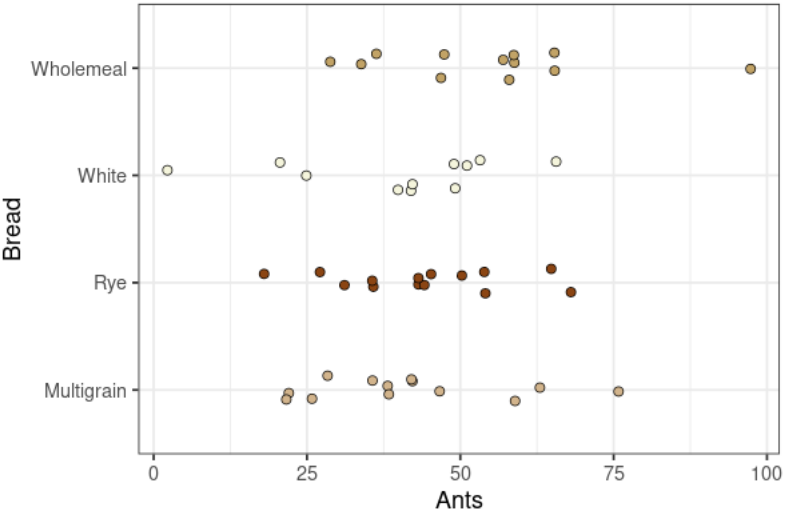


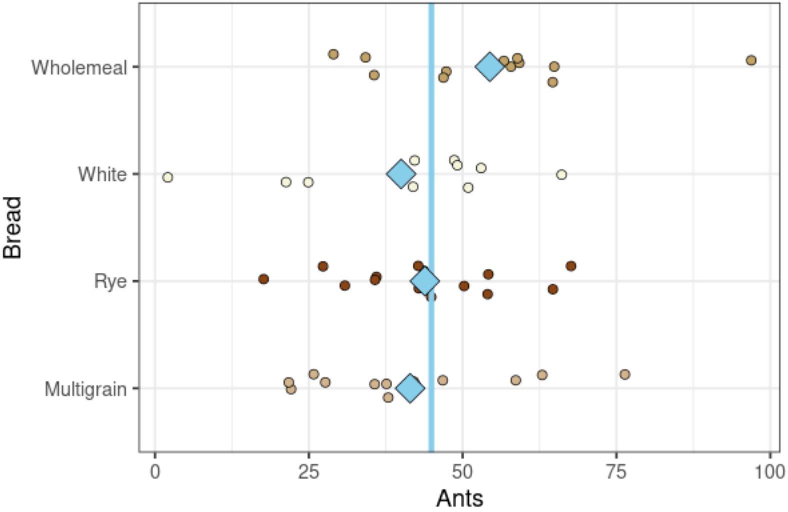
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Residuals	49	14041.68	286.56		


Butter
yes
no
0
25
50
75
100
Ants

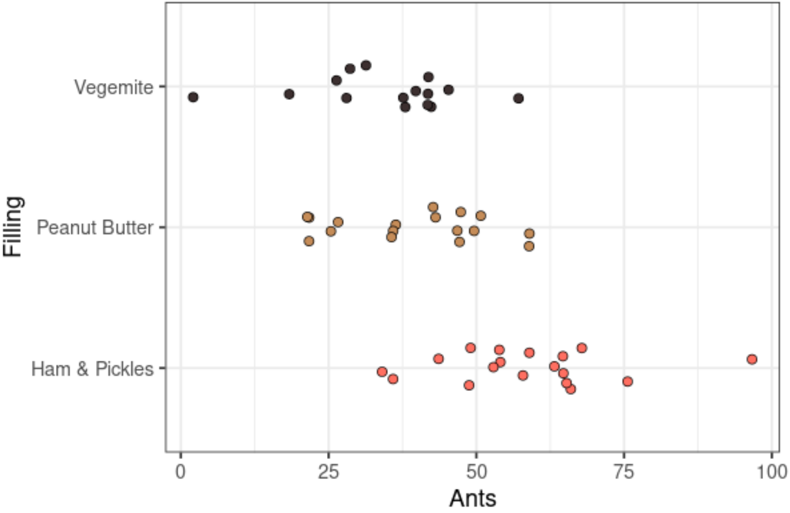


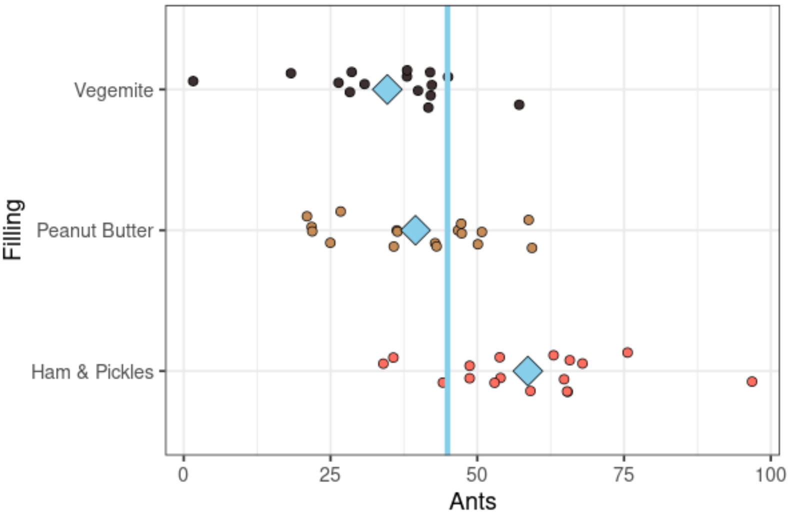
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Butter	1	757.90	757.90	2.74	0.1045
Residuals	48	13283.78	276.75		


Wholemeal
White
Rye
Multigrain
Bread
0
25
50
75
100
Ants

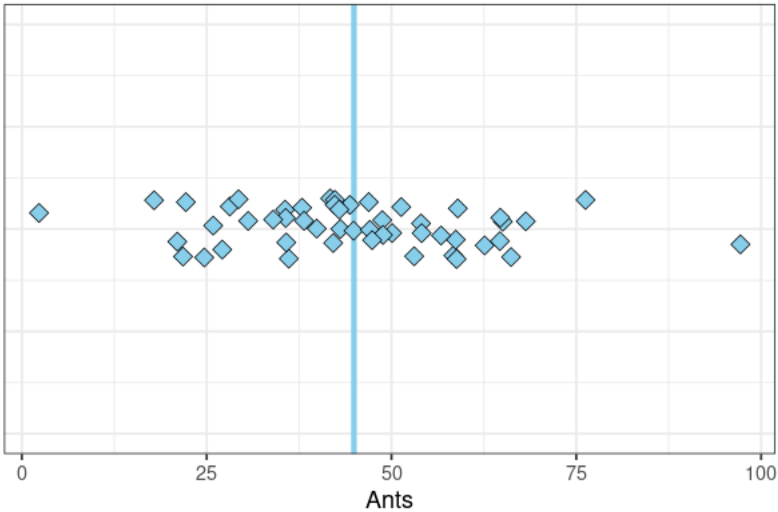


	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Bread	3	1519.82	506.61	1.86	0.1494
Residuals	46	12521.86	272.21		

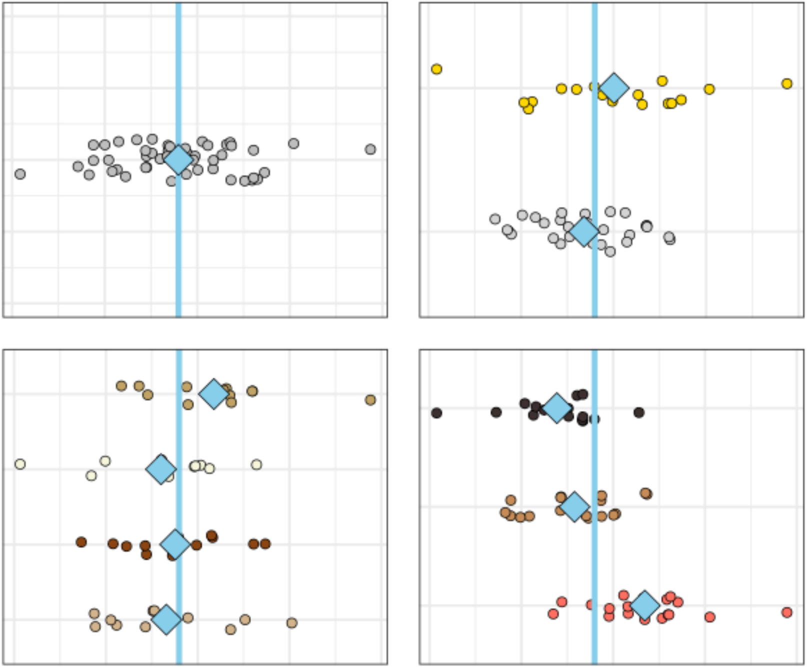

Vegemite
Peanut Butter
Ham & Pickles
Filling
0
25
50
75
100
Ants



	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Filling	2	5455.83	2727.92	14.93	0.0000095
Residuals	47	8585.85	182.68		



	Df	Sum Sq	Mean Sq
Individual	49	14041.68	286.56



	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Residuals	49	14041.68	286.56		

Table 1: No groups

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Butter	1	757.90	757.90	2.74	0.1045
Residuals	48	13283.78	276.75		

Table 2: Butter

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Bread	3	1519.82	506.61	1.86	0.1494
Residuals	46	12521.86	272.21		

Table 3: Bread

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Filling	2	5455.83	2727.92	14.93	0.0000095
Residuals	47	8585.85	182.68		

Table 4: Filling

To t or Not to t

Recall that for ANOVA we are testing the null hypothesis that *all* of our means our equal

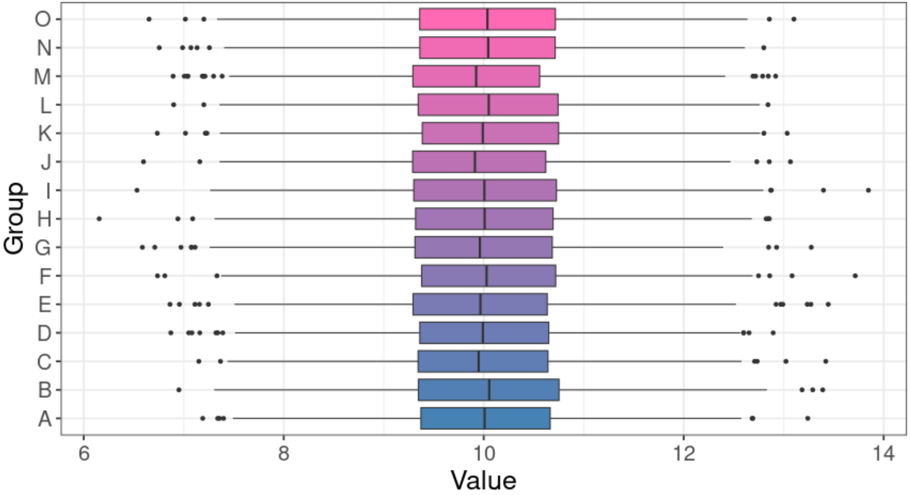
$$H_0 : \mu_A = \mu_B = \mu_C$$

Why not instead just stick with our t-test, doing

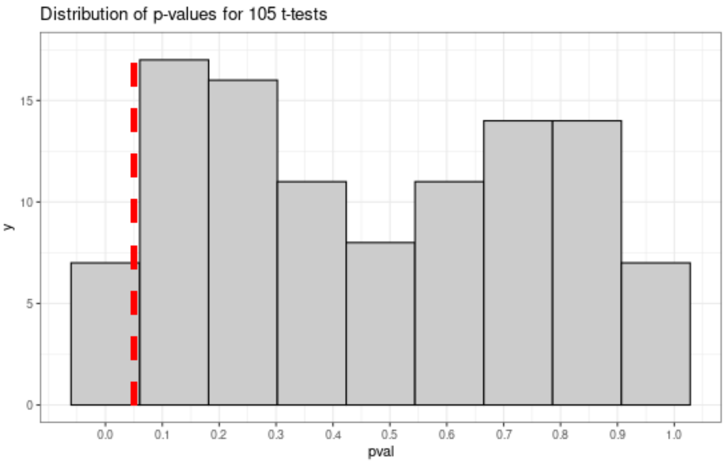
$$H_0 : \mu_A = \mu_B,\ \mu_A = \mu_C,\ \text{and}\ \mu_B = \mu_C$$

Multiple tests

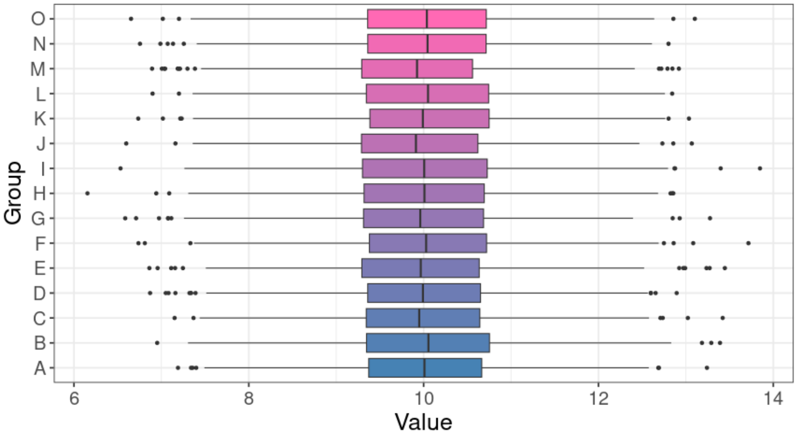
15 groups, all generated with the same mean value:



- 105 pair-wise tests
- 6 with p-value < 0.05



Multiple tests

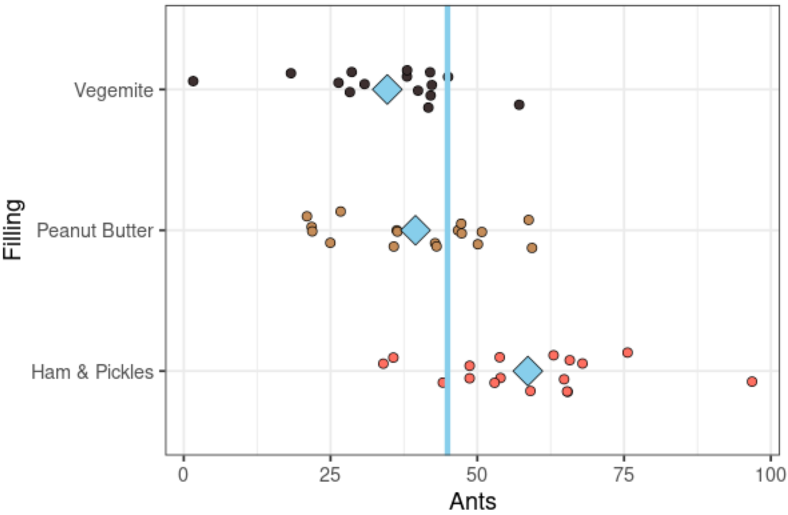


	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Group	14	15.40	1.10	1.10	0.3504
Residuals	14985	14964.85	1.00		

Post-hoc Tests

ANOVA only tells us *that* a difference exists, not where it is or to what degree

If our ANOVA test is such that we rejet the null hypothesis, we can use *post-hoc* testing via the **Tukey Range Test** or the **Tukey Honest Significant Difference Test** to identify any statistically significant pair-wise differences



	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Filling	2	5455.83	2727.92	14.93	0.0000095
Residuals	47	8585.85	182.68		

```
> aov(Ants ~ Filling, sandwich) %>% TukeyHSD()
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Ants ~ Filling, data = sandwich)

$Filling
                                diff     lwr      upr   p adj
Peanut Butter-Ham & Pickles -19.1405 -30.203  -8.0780 0.00036
Vegemite-Ham & Pickles      -23.9444 -35.380 -12.5090 0.00002
Vegemite-Peanut Butter       -4.8039 -16.391   6.7834 0.57845
```

Review

- ANOVA allows us to test equality of many means
  - By comparing ratio of between-group and within-group variances
- Ameliorates problem of multiple testing
- *Post-hoc* testing can be done to determine which groups are different
- Tukey Honest Statistical Difference (TukeyHSD)