Variance: $\dfrac{\sum (x-\bar{x})^2}{n-1} = \dfrac{SS_x}{df}$ ← "mean squared", MS

K = # of groups

j = group # $1 \le j \le K$

i = sample # within a group

X_{ij} = ith sample within group j $X_{3,2}$ 3rd sample in the 2nd group

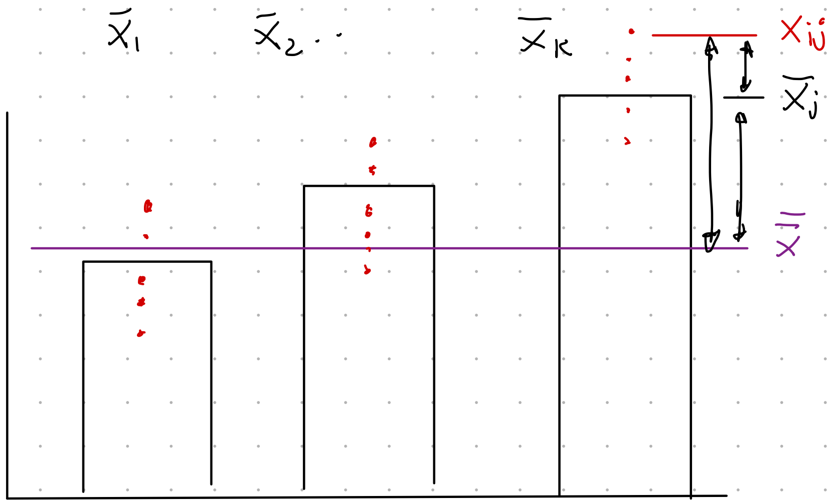
n_j = sample size for group j

$\bar{X}_j$ = mean of group j

$\bar{\bar{X}}$ = grand mean (mean of all samples) N = total sample size $\sum_j n_j$

data:

	X_{11}	X_{12}	....	X_{1K}
2nd sample of 1st group	X_{21}	X_{22}		X_{2K}
	$\vdots$	$\vdots$		$\vdots$
		$X_{n_2 2}$		$X_{n_K, K}$
	$X_{n_1 1}$			



$a - b = (a-c) + (c-b)$

$$(X_{ij} - \bar{\bar{X}}) = (X_{ij} - \bar{X}_j) + (\bar{X}_j - \bar{\bar{X}})$$

$$\sum_j \sum_i (x_{ij} - \bar{\bar{x}})^2 = \sum_j \left[\sum_i (x_{ij} - \bar{x}_j)^2 \right] + \sum_j \left[\sum_i (\bar{x}_j - \bar{\bar{x}})^2 \right]$$

sum of squares from grand mean

Sum of squares <u>within</u> each group

$$\sum_j n_j (\bar{x}_j - \bar{\bar{x}})^2$$

$$SS_{total} + SS_{within} + SS_{between}$$

$$df_{total} = df_{within} \quad df_{between}$$

$$N - 1 = N - k + k - 1$$

$$MS_{total} = \frac{SS_{total}}{df_{total}}, \quad MS_{within} = \frac{SS_{within}}{df_{within}}, \quad MS_{between} = \frac{SS_{between}}{df_{between}}$$

variance within groups

variance between groups

$$F = \frac{MS_{between}}{MS_{within}}$$

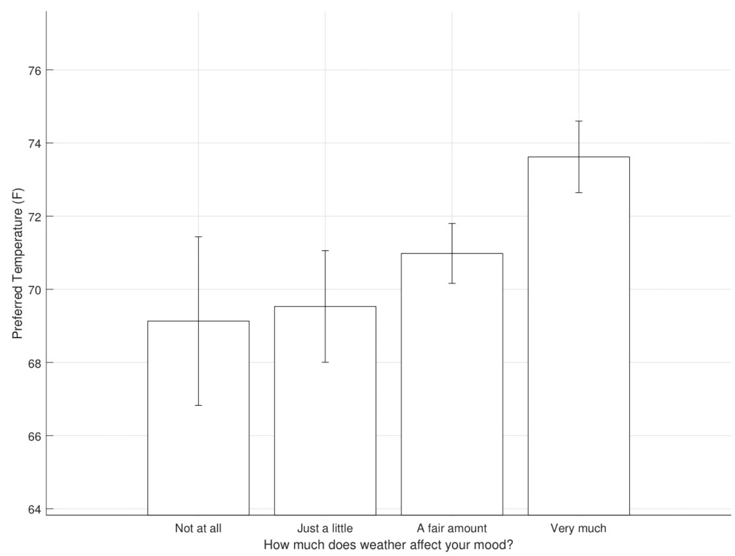
Example 2: Preferred temperature for weather sensitivity

At the beginning of the quarter I surveyed you for your preferred outdoor temperature. I also asked you how much weather affected your mood with the options of Not at all, Just a little, A fair amount and Very much. Let's see if there is a significant difference between the preferred temperatures across these 4 options. We'll us $\alpha = 0.05$ again. Here's a table of statistics:

	Not at all	Just a little	A fair amount	Very much
n	12	35	63	40
mean	69.13	69.53	70.98	73.62
SS	701.5468	2770.8755	2612.9852	1490.838
s S.d.	7.986	9.0275	6.4919	6.1828
sem	2.3054	1.5259	0.8179	0.9776

Totals:	
n	150
grand mean	71.1993
SS_{total}	7961.6499

$\sqrt{\frac{701.5468}{11}} = 7.986$ $\qquad$ $\frac{7.986}{\sqrt{12}} = 2.3054$



Numerator of the F statistic $\quad MS_{between} = \frac{SS_{between}}{df_{between}} \leftarrow k-1 = 4-1 = 3$

$$SS_{between} = \sum n_j(\bar{x}_j - \bar{\bar{x}})^2$$
$$= 9(69.13 - 71.1993)^2 + 35(69.53 - 71.1993)^2 + \dots = 385.4044$$

$$MS_{between} = \frac{385.4044}{3} = 128.7784$$

$\sum(x - \bar{x})^2$

Denominator: $\quad MS_{within} = \frac{SS_{within}}{df_{within}}$

$$SS_{within} = 701.5468 + 2770.8755 + \dots = 7576.2455$$
$$df_{within} = N - k = 150 - 4 = 146$$

$$MS_{within} = \frac{7576.2455}{146} = 51.8921$$

$$F = \frac{MS_{between}}{MS_{within}} = \frac{128.7784}{51.8921} = \underline{2.4817}.$$