Chemical reagents have a limited shelf-life. To determine the effect of light on a reagent's stability, a freshly prepared solution is stored for one hour under three different light conditions: total dark, subdued light, and full light. At the end of one hour, each solution was analyzed three times, yielding the following percent recoveries; a recovery of 100% means that the measured concentration is the same as the actual concentration.

↓ trial/condition →	A (total dark)	B (subdued light)	C (full light)
1	101	97	90
2	101	95	92
3	104	99	94

Calculating a One-Way Analysis of Variance

1. Treat the data as one large data set and calculate its mean and its variance, which we call the global mean, $\bar{\bar{x}}$, and the global variance, $\bar{\bar{s}}^2$.

$$\bar{\bar{x}} = \frac{\sum_{i=1}^{h}\sum_{j=1}^{n_i} x_{ij}}{N}$$

$$\bar{\bar{s}}^2 = \frac{\sum_{i=1}^{h}\sum_{j=1}^{n_i} (x_{ij} - \bar{\bar{x}})^2}{N-1}$$

   where h is the number of treatments, n_i is the number of replicates for the i^{th} treatment, and N is the total number of measurements.

2. Calculate the within-sample variance, s_w^2, using the mean for each treatment, $\bar{x}_i$, and the replicates for that treatment.

$$s_w^2 = \frac{\sum_{i=1}^{h}\sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2}{N-h}$$

3. Calculate the between-sample variance, s_b^2, using the means for each treatment and the global mean

$$s_b^2 = \frac{\sum_{i=1}^{h}\sum_{j=1}^{n_i} (\bar{x}_i - \bar{\bar{x}})^2}{h-1} = \frac{\sum_{i=1}^{h} n_i(\bar{x}_i - \bar{\bar{x}})^2}{h-1}$$

4. If there is a significant difference between the treatments, then s_b^2 should be significantly greater than s_w^2, which we evaluate using a one-tailed F–test where

   - $H_0 : s_b^2 = s_w^2$
   - $H_A : s_b^2 > s_w^2$

5. If there is a significant difference, then we estimate σ^2_{rand} and $\sigma^2_{systematic}$ as
   - $s^2_w \approx \sigma^2_{rand}$
   - $s^2_b \approx \sigma^2_{rand} + \bar{n}\sigma^2_{systematic}$

   where $\bar{n}$ is the average number of replicates per treatment.

This seems like a lot of work, but we can simplify the calculations by noting that

$$SS_{total} = \sum_{i=1}^{h}\sum_{j=1}^{n_i}(x_{ij} - \bar{\bar{x}})^2 = \bar{\bar{s}}^2(N-1)$$

$$SS_w = \sum_{i=1}^{h}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2$$

$$SS_b = \sum_{i=1}^{h} n_i(\bar{x}_i - \bar{\bar{x}})^2$$

$$SS_{total} = SS_w + SS_b$$

and that SS_{total} and SS_b are relatively easy to calculate; thus

source of variance	sum-of-squares	degrees of freedom	variance
between	$\sum_{i=1}^{h} n_i(\bar{x}_i - \bar{\bar{x}})^2$	$h-1$	$s^2_b = \frac{SS_b}{h-1}$
within	$SS_w = SS_{total} - SS_b$	$N-h$	$s^2_w = \frac{SS_w}{N-h}$
total	$SS_{total} = \bar{\bar{s}}^2(N-1)$		