A manufacturer of medicinal tablets was interested in testing the measurement accuracy of seven different laboratories. A composite was prepared by grinding and mixing together many tablets. Each lab made ten separate measurements on tablet-sized portions of the composite that had nominal dosage levels of 4g per tablet of the active ingredient. The table below shows the mean and standard deviation of these measurements (in grams) of the active ingredient from each lab.

	A	B	C	D	E	F	G
mean	4.059	4.018	4.057	4.121	4.170	4.115	4.102
sd	0.051	0.058	0.027	0.054	0.043	0.029	0.052

A partially completed ANOVA table is below.

1. (8 points) Complete the ANOVA table by filling in the missing values.

   Solution:

   There are a total of $n = 7 \times 10 = 70$ observations, so there should be 69 total degrees of freedom. There are $7 - 1 = 6$ numerator degrees of freedom and $70 - 7 = 63$ denominator degrees of freedom.

   The mean square for residuals will be a weighted average of the sample variances, weighted by their degrees of freedom. As the sample sizes (and thus degrees of freedom) are equal in each of the seven samples, the mean square for residuals in the ANOVA table is simply the mean of the sample variances.

$$MS_{\text{residuals}} = \frac{0.051^2 + 0.058^2 + 0.027^2 + 0.054^2 + 0.043^2 + 0.029^2 + 0.052^2}{7} \doteq 0.002143$$

   Each MS is of the form SS/df and the F statistic is the ratio of the MS factors, which is enough to complete the table.

```
Source       df     SS            MS         F value        Pr(>F)
-------------------------------------------------------------------
lab           6     0.153434      0.025572   11.93          6.975e-09
Residuals    63     0.135036      0.002143
-------------------------------------------------------------------
Total        69     0.288470
```

2. (2 points) State the null and alternative hypotheses for the test associated with the ANOVA table.

   Solution: The null hypothesis is that all seven population group means are equal. $H_0 \colon \beta_A = \cdots = \beta_G$ where β_i is the mean of the ith group, $\mathsf{E}(Y \mid X = i) = \beta_i$. The alternative hypothesis is that there is at least one difference. $H_A \colon$ not H_0

3. (6 points) The p-value in the table depends on several underlying model assumptions. Indicate if each statement is TRUE or FALSE. If FALSE, briefly correct the statement so it is true.

   (a) Each individual observation has a t distribution with 9 degrees of freedom.

   Solution: False. Each individual observation has a normal distribution (with a mean the same as others in the group and all with a common variance).

   (b) The variances of the distributions of individual observations are different in each group.

   Solution: False. The assumption is that the variance is the same for all observations.

   (c) All observations are mutually independent.

   Solution: True.

4. (2 points) The formula for the residual sum of squares is

$$SS_{\text{residuals}} = \sum_{i=1}^{a} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 \ .$$

Recall the formula for the sample standard deviation in a single sample with n observations is

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

and that $\mathsf{E}(S^2) = \sigma^2$ for a random sample where $\mathsf{Var}(X_i) = \sigma^2$. Write $SS_{\text{residuals}}$ in terms of the sample variances from each group and show that $\mathsf{E}(SS_{\text{residuals}}) = (n - a)\sigma^2$.

Solution:

$$\begin{aligned}
\mathsf{E}(SS_{\mathsf{residuals}}) &= \mathsf{E}\left(\sum_{i=1}^{a} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 \right) \\
&= \mathsf{E}\left(\sum_{i=1}^{a} (n_i - 1) S_i^2 \right) \\
&= \sum_{i=1}^{a} (n_i - 1)\sigma^2 \\
&= (n - a)\sigma^2
\end{aligned}$$

5. (2 points) What is a numerical estimate of σ for this data?

Solution: $\hat{\sigma} = \sqrt{0.002143} = 0.046$.