A	B	C
68	84	72
61	67	72
84	67	51
78	75	82
93	85	86
76	62	79
92	62	62
68	74	79
79	71	83
76	81	79
81	69	80
87	87	74

Suppose you give 3 exams over 3 years to 30 different people. Do these exams differ in difficulty over the years?

You could compare two means at a time with independent measures t-tests. 3 tests.

A vs. B, A vs. C, B vs. C. Let $\alpha = .05$

Problem: the probability of 1 or more type I errors is greater than α if we run multiple tests.

suppose we have m tests

$p(1 \text{ or more Type I}) = 1 - (1-\alpha)^m$ if $m=3$, $\alpha = .05$ $p = 0.14 > .05$

"Familywise error"

Means 81 72 75

We really just want to test:

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

H_A: H_0 is false

How different are these means from each other? Variance!

variance of 81, 72 and 75 is $\underline{22.5}$

what do we compare the variance of the means to?

Suppose σ^2 is the population variance for all 3 years

$s_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ ← central limit theorem

$s_{\bar{x}}^2 = \frac{\sigma^2}{n} \Rightarrow \sigma^2 = n \cdot s_{\bar{x}}^2$

variance of population, sample size, variance of means

n x variance of means is an estimate of σ^2

$n \cdot s_{\bar{x}}^2 = (10)(22.5) = \underline{225}$

Another way of estimating σ^2 from the data:

calculate the variance for each sample.

A	B	C
68	84	72
61	67	72
84	67	51
78	75	82
93	85	86
76	62	79
92	62	62
68	74	79
79	71	83
76	81	79
81	69	80
87	87	74

variances 61 94 55

σ^2 can be estimated by taking the mean of the variances.

mean of 61, 94 and 55 is 70

we have two estimates of σ^2: 1) n×variance of the means: 225

2) mean of the variances: 70

If H_0 is true, then these #s estimate the same thing (σ^2)

If H_0 is false then the variance of the means will increase but the mean of the variances stays the same

define $F = \dfrac{n \times \text{variance of means}}{\text{mean of variances}}$

H_0 is true, then $F \sim 1$

$F = \dfrac{225}{70} = 3.23$

	1	2	3
27	**4.21**	**3.35**	**2.96**
	7.68	5.49	4.60
28	4.20	3.34	2.95

F table has two degrees of freedom

one for numerator: k = # of groups, df = k − 1, 3 − 1 = 2

one for denominator: n = sample size for each group, $k(n-1) =$

$kn - k =$

$N - k = 30 - 3 = 27$

total sample size

F_{crit} for df 2,27 = 3.35 < 3.23

Fail to reject H_0.