Why One-Way ANOVA and Regression Give the Same F-Statistic

Both one-way ANOVA and regression decompose the total variation in Y into two parts: Explained variation and unexplained variation. They are mathematically equivalent when the explanatory variable represents group membership through dummy variables.

Example:

Groups:
A: 10, 12, 8
B: 20, 22, 18
C: 15, 13, 17

Step 1: Means
Group means:
A = 10
B = 20
C = 15
Overall mean = 15

Step 2: ANOVA decomposition

Total SS (SST) = $\Sigma(Y_i$ - overall mean$)^2$ = 174
Between-group SS (SSB) = $\Sigma n_j(\blacksquare_j$ - overall mean$)^2$ = 150
Within-group SS (SSW) = SST - SSB = 24

F = (SSB / (k - 1)) / (SSW / (N - k)) = (150 / 2) / (24 / 6) = 18.75

Step 3: Regression view

Use dummy variables for groups B and C:

$Y_i = \beta\blacksquare + \beta_B * x_B + \beta_C * x_C + \varepsilon_i$

$\beta\blacksquare$ = 10, β_B = 10, β_C = 5

Regression SS (SSR) = 150
Residual SS (SSE) = 24
F = (SSR / 2) / (SSE / 6) = 18.75

Thus, ANOVA F-statistic = Regression F-statistic.

Why they are the same:
Both models partition total variance identically:
Total = Explained + Unexplained

ANOVA uses group means; regression uses dummy-coded predictors.
Mathematically, they represent the same linear model.