Now we use the above result for checking the goodness of the linear fit. ANOVA for checking the goodness of $Y = X\beta + \epsilon$, or $y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_{i(p-1)} + \epsilon_i$, or equivalently for testing $H_0 : \beta_1 = \cdots = \beta_{p-1} = 0$ is what is needed. Intuitively, if $X_1, \ldots, X_{p-1}$ provide no useful information, then the appropriate model is $y_i = \beta_0 + \epsilon_i$, so $\bar{y}$ is the only quantity that can help in predicting y. Then $\text{RSS}_{H_0} = \sum_{i=1}^{n}(y_i - \bar{y})^2$ is the sum of squares unexplained, and it has $n-1$ d.f. If $X_1, \ldots, X_{p-1}$ are also used in the model, then $(Y - X\hat{\beta})'(Y - X\hat{\beta}) = \text{RSS}$ is the unexplained part with $n - r$ d.f. How much better is RSS compared to RSS_{H_0}? Let SS_{reg} denote the sum of squares due to $X_1, \ldots, X_{p-1}$ and without an intercept. Then,

$$\begin{aligned} \text{RSS}_{H_0} &= \text{RSS} + \text{SS}_{reg} \\ \sum_{i=1}^{n}(y_i - \bar{y})^2 &= \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + \text{SS}_{reg} \end{aligned}$$

In other words,

$$\begin{aligned} Y'Y - \frac{1}{n}Y'1'1Y &= Y'(I-P)Y + \text{SS}_{reg}, \text{ or} \\ Y'Y &= Y'(I-P)Y + \left(\text{SS}_{reg} + \frac{1}{n}Y'1'1Y\right), \text{ or} \\ \text{SSR} &= \hat{\beta}'X'X\hat{\beta} = \hat{\beta}'X'Y = \left(\text{SS}_{reg} + \frac{1}{n}Y'1'1Y\right), \end{aligned}$$

since $Y'Y = Y'(I-P)Y + Y'PY = Y'(I-P)Y + \hat{\beta}'X'X\hat{\beta}$. Now, $\mathbf{1}(\mathbf{1}'\mathbf{1})^{-1}\mathbf{1}' = \frac{1}{n}\mathbf{1}\mathbf{1}' = P_{\mathcal{M}(\mathbf{1})} = P_{\mathcal{M}(X_0)}$, so that $\text{SSR} = n\bar{y}^2 +$ SS_{reg} is the orthogonal decomposition of SSR into components attributed to $\mathcal{M}(\mathbf{1})$ and $\mathcal{M}(X_1, \ldots, X_{p-1})$. Therefore SS_{reg} with $r-1$ d.f. is the quantity to measure the merit of the regressors, $X_1, \ldots, X_{p-1}$.

ANOVA with mean

source of variation	d.f.	sum of squares	mean squares	F-ratio
mean	1	$\text{SSM} = n\bar{y}^2$	$\text{MSM} = \text{SSM}/1$	$F_{mean} = \text{MSM}/\text{MSE}$
regression on $X_1, \ldots, X_{p-1}$	$r-1$	$\text{SS}_{reg} = \hat{\beta}'X'Y - n\bar{y}^2$	$\text{MS}_{reg} = \text{SS}_{reg}/(r-1)$	$F_{reg} = \text{MS}_{reg}/\text{MSE}$
residual error	$n-r$	$\text{SSE} = \text{RSS} = Y'Y - \hat{\beta}'X'Y$	$\text{MSE} = \text{SSE}/(n-r)$	
Total	n	$\text{SST} = Y'Y$		

ANOVA for regression (corrected for mean)

source of variation	d.f.	sum of squares	mean squares	F-ratio
regression (corrected)	$r-1$	$\text{SS}_{reg} = \hat{\beta}'X'Y - n\bar{y}^2$	$\text{MS}_{reg} = \text{SS}_{reg}/(r-1)$	$F_{reg} = \text{MS}_{reg}/\text{MSE}$
residual error	$n-r$	$\text{SSE} = \text{RSS} = Y'Y - \hat{\beta}'X'Y$	$\text{MSE} = \text{SSE}/(n-r)$	
Total (corrected)	$n-1$	$\text{SST(Corrected)} = \sum(y_i - \bar{y})^2$		

How good is the linear fit? There are two things to consider here.
(i) The ANOVA F-test: Under $H_0 : \beta_1 = \cdots = \beta_{p-1} = 0$, the F-ratio, $F_{reg} \sim F_{r-1,n-r}$ and large values of the statistic provide evidence against H_0, or equivalently indicate that the regressors are useful.
(ii) The proportion of variability in y not explained by the actual regressors is: RSS/SST (corrected), so the proportion of variability in y around its mean, explained by the actual regressors is

$$1 - \frac{\text{RSS}}{\text{SST (corrected)}} \equiv R^2 = \text{ Coefficient of determination.}$$

In other words,

$$\begin{aligned}
R^2 &= 1 - \frac{\text{RSS}}{\text{SST (corrected)}} = 1 - \frac{Y'(I-P)Y}{Y'(I - \frac{1}{n}\mathbf{11}')Y} \\
&= \frac{\sum_{i=1}^n (y_i - \bar{y})^2 - Y'(I-P)Y}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{\sum_{i=1}^n y_i^2 - n\bar{y}^2 - Y'(I-P)Y}{\sum_{i=1}^n (y_i - \bar{y})^2} \\
&= \frac{Y'Y - n\bar{y}^2 - Y'Y + Y'PY}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{Y'PY - n\bar{y}^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \\
&= \frac{\text{SSR } - n\bar{y}^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{\text{SS}_{reg}}{\text{SST (corrected)}} \\
&= \text{ proportion of variability explained by regressors.}
\end{aligned}$$

Also,

$$\begin{aligned}
R^2 &= \frac{\text{SS}_{reg}}{\text{SST (corrected)}} = \frac{\text{SS}_{reg}}{\text{RSS} + \text{SS}_{reg}} \\
&= \frac{\text{SS}_{reg}/\text{RSS}}{1 + \text{ SS}_{reg}/\text{RSS}} = \frac{(\frac{r-1}{n-r})F_{reg}}{1 + (\frac{r-1}{n-r})F_{reg}}
\end{aligned}$$

is an increasing function of the F-ratio.

Note that to interpret the F-ratio, normality of ϵ_i is needed. R^2, however, is a percentage with a straightforward interpretation.