

```
> summary(model1)
```

Call:

```
lm(formula = House_price ~ Population + Crime + Average_age +  
    Household_income, data = data)
```

Residuals:

Min	1Q	Median	3Q	Max
-26460.7	-6011.9	-386.4	6331.8	24591.6

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	1.807e+05	1.680e+04	10.754	< 2e-16	***
Population	6.577e-01	2.453e-01	2.682	0.00782	**
Crime	-3.358e+02	1.153e+02	-2.913	0.00391	**
Average_age	-8.218e+01	1.186e+02	-0.693	0.48915	
Household_income	-1.957e-02	3.033e-01	-0.065	0.94861	

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9394 on 245 degrees of freedom  
Multiple R-squared: 0.06272, Adjusted R-squared: 0.04741  
F-statistic: 4.098 on 4 and 245 DF, p-value: 0.00311

Here we have our parameter estimates and the t-tests associated with our predictors.

Here are the R-squared and Adjusted R-squared values (which reflects number of predictors in our model).

```
# Notice that Average_age and Household_income do not seem to predict house prices
# Let's drop them in model2
model2 <- lm (House_price ~ Population + Crime, data = data)
anova(model2, model1)
```

```
> anova(model2, model1)
Analysis of Variance Table
```

```
Model 1: House_price ~ Population + Crime
```

```
Model 2: House_price ~ Population + Crime + Average_age + Household_income
```

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	247	2.1666e+10				
2	245	2.1622e+10	2	43401593	0.2459	0.7822

OK, so the models do not differ significantly by this test - we can use another measure of goodness-of-fit - AIC (Aikaike Information Criterion). AIC tells us how much information in our data is not captured by each model - lower values are better - can only be interpreted in a relative sense (i.e., comparing one model to another)...