Chapter 10

Study 4: How openness impacts developer’s satisfaction and their contribution

My third empirical study strikingly reveals that openness is important to developer communities and that this is not just a matter of course. Openness of software components is more important than openness of hardware components. Replicability is perceived less important than accessibility and transparency. Highly active developers indicate higher importance ratings than less active developers, while the duration of participation does not influence the indicated importance.

In my last empirical study of this thesis, I aim to develop different statistical models in order to systematically investigate the impact of the degree of openness, of its aspects, and of individual expectations towards it on developer satisfaction. Based on statistical models I attempt to draw conclusions on my remaining sets of research hypotheses \( H_4, H_5, H_6, \) and \( H_7 \). In addition effects on contributed working hours of respondents to their projects shall be investigated to develop an understanding of the broader picture of the examined variables and their effects.

10.1 Ordinary linear regression models on satisfaction

To explore structures in the underlying data, I start with the simplest ordinary linear regression model with one response or dependent variable, satis-
faction (S), and one prediction or independent variable, openness (O):

\[ E(S|O) = \beta_0 + \beta_1 O \text{ or } S = \beta_0 + \beta_1 O + \epsilon \]  

(10.1)

where \( \beta_0 \) and \( \beta_1 \) denote unknown regression parameters, and \( \epsilon \) denotes the residuals. A general linear regression model obtained via QR-decomposition using the R-function `lm` returns:

Residuals:

<table>
<thead>
<tr>
<th>Min</th>
<th>1Q</th>
<th>Median</th>
<th>3Q</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2.66881</td>
<td>-0.35468</td>
<td>0.05025</td>
<td>0.42419</td>
<td>1.15152</td>
</tr>
</tbody>
</table>

Coefficients:

|              | Estimate | Std. Error | t value | Pr(>|t|) |
|--------------|----------|------------|---------|----------|
| (Intercept)  | 2.56511  | 0.22621    | 11.340  | < 2e-16  *** |
| O            | 0.37199  | 0.05738    | 6.483   | 3.58e-10 *** |

Residual standard error: 0.614 on 307 degrees of freedom
Multiple R-squared: 0.1204, Adjusted R-squared: 0.1176
F-statistic: 42.03 on 1 and 307 DF, p-value: 3.584e-10

Figure 10.1: Diagnostic plots of an ordinary linear regression model (Model 10.1)

The estimates for the parameters have been calculated as

\[ \hat{\beta}_0 = 2.57^{***}, \quad \hat{\beta}_1 = 0.37^{***} \]

The model reveals a significant positive impact from openness on satisfaction (p-value < 0.1%). The \( R^2 \) of 11.8% points to an existing degree of explanation, but it also makes obvious that other factors beyond openness impact satisfaction.