**4. Analysis**

**4.1. Statistical test used to test the hypotheses and output**

To test the hypotheses, a **Chi-Squared Test of Independence** was conducted. This test is appropriate as both variables, **income** and **education level**, are categorical. The test evaluates whether there is a significant association between income categories (<=50K and >50K) and education levels.

The test produced the following output:

• **Chi-squared statistic**: 4352.3

• **Degrees of freedom**: 15

• **P-value**: < 2.2e-16

**4.2. The null hypothesis is rejected /not rejected based on the p-value**

Based on the p-value (< 2.2e-16), the null hypothesis is **rejected**. This indicates a statistically significant difference in proportions of income across different education levels. Thus, income distribution is not independent of education level.