Course: Data Analysis

Student’s Name and Surname \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Regression analysis**

Do the multiple regression analysis using the following variables from **data\_games.dta**:

Dependent variable: payment

Independent variables: payment\_type, payment\_method, crystalls\_balance\_before\_buy, crystalls\_bought

*Describe the model:*

1. Specify the regression equation.

2. Assess the goodness-of-fit of the model (R-square, significance of the model);

3. Which independent variables significantly influence the dependent variable (explain your answer)?

4. Describe the relationship between each independent and the dependent variable.

*Do the diagnostics of the model:*

5. Are the residuals normally distributed?

6. Are there any outliers (standardized residuals greater than 3, or less that -3)?

7. Test the multicollinearity and heteroscedasticity.

Please send this MS Word file with answers and the do-file to amelikyan@hse.ru from your personal e-mail.