Course: Data Analysis (task № 2)

Student’s Name and Surname Iurii Aksenov

1. Open **salary.dta** file.

* 1. Recode **position** variable from string into numeric.

1.2. Calculate a new variable which will show the sum of income taxes **in dollars** which is paid monthly to every person (use today’s exchange rate). Take into account that foreigners pay 30% of income taxes and non-foreigners – 13%.

1.3. Create a new variable which will represent the rank of a person which is calculated on the basis of his/her salary. Person with the highest salary will have rank 1.

1.4. Sort the data of the basis of the salary variable so that the person with the highest salary appears in the first row.

1.5. Save all the changes in **salary.dta** file.

2. Open **data\_games.dta** file.

2.1. Deselect form the further analysis the cases which have the value of “**payment**” variable less than 100 or greater than 14000.

2.2. Divide the remained cases into 4 groups using the following interval width:

1st group: from 100 to 500;

2nd group: from 501 to 1000;

3rd group: from 1001 to 2000;

4th group: greater than 2000.

Indicate the mode group number: **Answer: 1**

3. Open **revenue.dta** file.

Modify the dataset so as it contains three variables:

- year;

- company;

- new variable which represents the mean revenue of each company for each year.

4. Contingency table. Open **data\_games.dta** file.

Create a contingency table between two variables **payment\_method** and **country**. Answer the following questions.

4.1. Which payment method is most frequently used in Russia (RU)?

General (41)

4.2. What is the percentage of payments which were done using a gift card from all the worldwide payments?

58/105507\*100% = **0.05%**

4.3. Players from which country have made the highest number of payments using a mobile payment type?

DE 495

4.4. Create a suitable graph to demonstrate the relationship between these two variables.

We create a bar chart of the countries from Group of Seven (except Japan, it’s not mentioned in data) and their payment method.



We can see that general payment method prevails in all countries over other types of payment. Moreover, US is the first in general and fb\_promotion methods, except mobile where FR has more frequency than other discovered countries. In addition, we see that GE has the lowest position in all payment methods. The top three leaders by frequency are US, GB, FR. However, they have different positions in methods. As we can see in fb\_promotion US, GB, FR; in general US, GB, FR; in mobile FR, US, GB.

Please send to amelikyan@hse.ru from your personal e-mail:

- this MS Word file with answers,

- the do-file with the corresponding commands,

- the modified salary.dta file.