

Requirements for “Report” in *Elementary Statistics***Assumptions**

Data obtained from instructor represent results of survey conducted in 2 primary schools located in different regions of a country, rural and urban. The goal of the survey is to **detect and describe determinants of pupil’s absence in each type of school and make identify similarities and differences between schools.**

Smaller data set (20 cases) comes from rural area, bigger (40 cases) - from urban area.

Each data set comprises of the following variables:

<i>variable name</i>	<i>survey question</i>	<i>value set</i>
id	identification number of a pupil (in school)	(1-20 and 1-40)
gender	gender of a pupil	(0 - female, 1 male)
age	age of a pupil	(10 -15)
games	Do you often play computer games	(0 - No, 1 - Yes)
sport	In how many sport activities do you participate systematically	{0, 1, 2}
winterab	How many schooldays did you omit during winter semester ?	
summerab	How many schooldays did you omit during summer semester ?	

Structure of the “Report” - task description**I. Comparison of single distributions**

Compare two of schools with regard to:

- social structure of pupil’s population (use information on gender and age of pupils),
- type of leisure time usage of pupils (sports and computing activities),
- extent of pupil’s school absence (winter, summer and the whole year)

Statistical and computational tools used in the chapter:

univariate distributions and univariate statistics – central tendency, dispersion; COMPUTE, FREQUENCIES, CROSSTABS, VAR LAB, VAL LAB, SPLIT

Content of the chapter:

1. 6 charts representing distributions of all variables (**gender** to **summerab**) in both schools,
2. chart representing distribution of new variable “**total absence**” during year in both schools
3. description of differences and similarities of distributions between two schools in regard to their shape, central tendency and dispersion (see “GUIDE”).

II. Relationships between pupil’s absence and it’s social determinants in two schools

Compare dependencies of three kinds of absences (dependent variables) on pupils’ characteristics (independent variables) in both schools in regard to:

- shape and the strength of dependency in the regression of means,
- direction and the strength of dependency in the linear regression,
- similarity of the 1st type (means) and linear regression functions (degree of linearity).

Take into account associations of each kind of absence with each of pupil’s characteristics:

- location in the demographic structure: **gender, age,**
- profile of the leisure time usage: **games, sport**

Statistical and computational tools used in the chapter:

Regression of means, linear regression, coefficients of statistical dependency;

MEANS, REGRESSION.

Content of the chapter:

1. 12 charts (3 kinds of absence, 4 independent variables: **gender, age, games sports**) representing dependencies in both schools,
2. tables containing parameters of statistical associations for all bivariate distributions,
3. description of differences and similarities of statistical dependencies between two schools (see “GUIDE”)

III. Linear model of absence in school

Compare multivariate linear models of school absences in both schools for each dependent variable (**winterab, summerab** and “**total absence**”) in regard to:

- goodness of fit of a linear model,
- direction of each determinant (independent variable) association with an absence within a model,
- relative strength of determinants’ impact on the dependent variable prediction within a model,
- explanatory power of each determinant in linear prediction of pupil’s absence.

Statistical and computational tools used in chapter:

Multiple regression, multiple correlation coefficient, standardized regression coefficient; REGRESSION.

Content of the chapter

1. 3 tables containing parameters of each linear model for both schools,
2. 3 charts representing relative strength and direction of determinants impact on dependent variable prediction (within a model),
3. 3 charts representing relative explanatory power of each determinants in a model,
4. description of differences and similarities of models identified for both schools (see “GUIDE”)