

PROJECT-BASED TRAINING IN STATISTICS USING MS EXCEL IN SECONDARY EDUCATION

Assoc. Prof. Desislava Georgieva, PhD

Department of Mathematics

Faculty of Mathematics and Informatics

St. Cyril and St. Methodius University of Veliko Tarnovo, Bulgaria

Phone: +359 887 244 498

E-mail: d.georgieva@ts.uni-vt.bg

Abstract: *A study of mathematics education of 12th grade students of a secondary general education school was conducted on the topic of Statistics. The problems in studying this section were outlined. A review of the scientific literature abroad and in our country was conducted. A model for project-based learning using MS Excel was proposed and tested. Mathematical and digital competence were integrated skilfully. The conducted survey reported increased motivation of the students.*

Keywords: *Statistics training, secondary education, general education, project-based learning, MS Excel.*

INTRODUCTION

Humanity gains experience through repeated observation and establishment of certain patterns. The statistical approach uses registration, processing and analysis of many individual cases to establish various trends. *“The main directions for applying the results of statistical study are: structural analysis of processes and phenomena of reality; forecasting – normative and in the future; making management decisions”* (Arkadiev, 2015). Modern man “immersed in the ocean of information” must possess statistical literacy in order to avoid manipulation. Knowledge in the field of statistics is a starting point in assessing situations of uncertainty and serves to make optimal decisions. *“In their daily lives, all people encounter situations that require statistical thinking and assessment – for example, in terms of weather forecasts, crop development, when discussing issues of income and living standards, when managing the household budget, and many others. For all this, statistical literacy is important for modern people and needs to be developed from school years”* (Sugareva & Murgova, 2021). Statistical competence is necessary for politicians, economists, sociologists, journalists, and many other professions.

Statistics is the science that deals with the collection, organization, analysis, interpretation, and presentation of data. It distinguishes itself from most branches of mathematics by its applied focus.

It is appropriate for students to extract data from official sites, such as Google Data Search or Kaggle, or to collect it through online surveys (Karakoleva, Koleva, 2024).

EXPOSITION

1. Research Parameters

The object of the research is the teaching of mathematics to 12th grade students on the topic of Statistics.

The subject of the research is the pedagogical and technological characteristics of teaching *Statistics* to 12th grade students.

Research aim is to study theoretically the state of the problem of teaching on the topic of *Statistics* and conducting a project-based training that will increase efficiency and contribute to the development of statistical competence.

For achievement of the main goal, the following specific tasks have been completed in advance: basic concepts of the science of Statistics have been clarified; a historical review has

been carried out and the problems in studying this topic in the school mathematics course have been specified; the state educational standards at the high school level have been systematized; a modern technological model of teaching a Statistics lesson in the 12th grade has been developed. Due to the limited volume of this publication, they will not be included in it.

2. Status of teaching on the topic of Statistics

The problems in studying the Statistics section in the 12th grade arise from:

1. The lack of interest among students. This discipline often perceived as too theoretical, contributes to the negative attitude towards studying it.
2. The untargeted preparatory instruction of the studied content.
3. The large number of new abstract concepts.
4. A very limited number of instructional hours, in which the knowledge remains misunderstood and inapplicable in situations other than those given in textbooks.
5. A limited number of teaching resources with which to prepare students for state matriculation exams.
6. Insufficiently good preparation of teachers on the topic, due to its limited inclusion in both school curricula and the university discipline School Course in Mathematics.
7. Lack of appropriate methodological literature.

The conducted NVE and DZI at the high school stage show unsatisfactory results of students on this type of tasks. Therefore, our theoretical research explores contemporary teaching and assessment methods integrating ICT in Bulgaria and abroad through an analysis of documents and educational literature on the topic of Statistics.

3. Literature review

David Moore postulated that statistics education could improve by “strengthening the synergistic relationships between content, pedagogy, and technology, using computers as a tool for learning statistics, not just for statistical processing” (Moore, 1997).

In 2005, the American Statistical Association issued a report with recommendations for reforming statistics courses in high school and college. In 2016, it was updated and recommendations for assessment were added. In 2019, a national survey was conducted in the United States on statistics educational practices in colleges and universities (Legacy et al., 2024). The study found that education emphasizes the application of basic statistical methods. Although the lecture method remains dominant, teachers are increasingly integrating active forms of learning and the use of statistical software. The main challenges are the weak mathematical preparation of students, subject anxiety, and difficulties in organizing active learning in large groups. Only a part of teachers adopt innovative teaching methods and technologies (Tintle et al., 2016).

A study showed that combining a flipped classroom and project-based learning significantly improves students’ problem-solving skills, critical thinking, and creativity (Burckhardt et al., 2021).

MS Excel is often used effectively in teaching Statistics. It facilitates experimental and cooperative learning (Harits et al., 2019). The Excel environment is integrated with constructivist teaching approaches to improve learning outcomes (Basturk, 2005; Alismail & McGuire, 2015). Statistics lessons using spreadsheets in secondary education have been studied (Borkulo et al., 2023).

An experimental group participated in hands-on activities using Excel, while a control group relied on traditional methods (Kazika, Malambo, 2025). The study found a significant difference in statistical literacy and technological literacy. Working with real data using MS Excel encourages critical thinking and collaborative problem solving. Students in the experimental

group say that they understand statistics more easily and perceive it as suitable for application in real life.

In Bulgaria, for general education in secondary school, collections containing tasks from Statistics have been published (Zapryanov, Marinkova, 2008; Zapryanov, Marinov, 2018). A digital classroom for 10th grade students on the topic "Statistics and Data Processing" has been developed (Karakoleva, Dimitrova, 2021). Educational videos intended for students are freely available, which are also suitable for students.

Only in 2025 was a methodological guide intended for teachers published, which has an applied nature (Sugareva, M. et al., 2025)

4. Project-Based Learning on the topic of Statistics in 12th grade

In 2024-2025, modern methods were developed and tested in the education of 12th grade students in the Statistics section. The organization of the Statistics section lessons includes the following activities:

the teacher demonstrates the application of MS Excel;

develops a lesson plan for the exercise;

conducts a project-based lesson at the end of the section to consolidate statistical knowledge and understand its applied nature.

The project-based lesson follows all the stages and requirements of the educational project method (choice of topic, teams, project result, presentation, etc.).

First stage: Setting the topic "Place and role of Statistics"

At the beginning of the lesson, the teacher sets the condition of the project task: Using the website of the National Statistical Institute, or other similar sites, select appropriate information in the area of your choice. Perform a statistical analysis of the selected data by applying all the knowledge acquired so far.

Students are familiar with the requirements for statistical analysis, which for the main project tasks are:

1. Grouping the data in an interval statistical order (determining an appropriate number of intervals).
2. To construct histograms and polygons of absolute and relative frequencies.
3. To construct the accumulated (accumulated, cumulative) absolute and relative frequencies.
4. To calculate the arithmetic mean, linear deviation, standard deviation.
5. To construct a confidence interval for estimating the relative share of the studied characteristic at a confidence probability chosen by you.

Second stage: The solutions are described and the project tasks are arranged and considered to the way of presenting the project.

Third stage: Presenting the project to an audience (class).

The evaluation of the student projects is carried out according to a rating scale (Table 1)

Table 1. Project evaluation scale in Statistics

Project tasks	Points	Evaluation
1	2	
2	4	
3	4	
4	6	
5	2	
Presentation	2	
Total score <i>N</i>	20	$2 + 0,2 \cdot N$

In the presented project of a team of two students, waste disposal in several regions of Bulgaria for the period 2011 - 2018 was analysed (Fig. 1).

The students initially enter the data into a table - the row headings are the names of the regions, and the column headings are the years. They write all the data in one column and sort them in ascending order using the Excel option. In an adjacent column, they write their frequencies. They determine the ends of the intervals in column P. The students use the sum formula for part of the cells in column N, they determine the frequencies in the corresponding intervals. Thus, the data are grouped in an interval statistical order. In column Q with the heading x_i the middles of the corresponding intervals are placed in column R with the title f_i are the frequencies of the variables from the corresponding intervals.

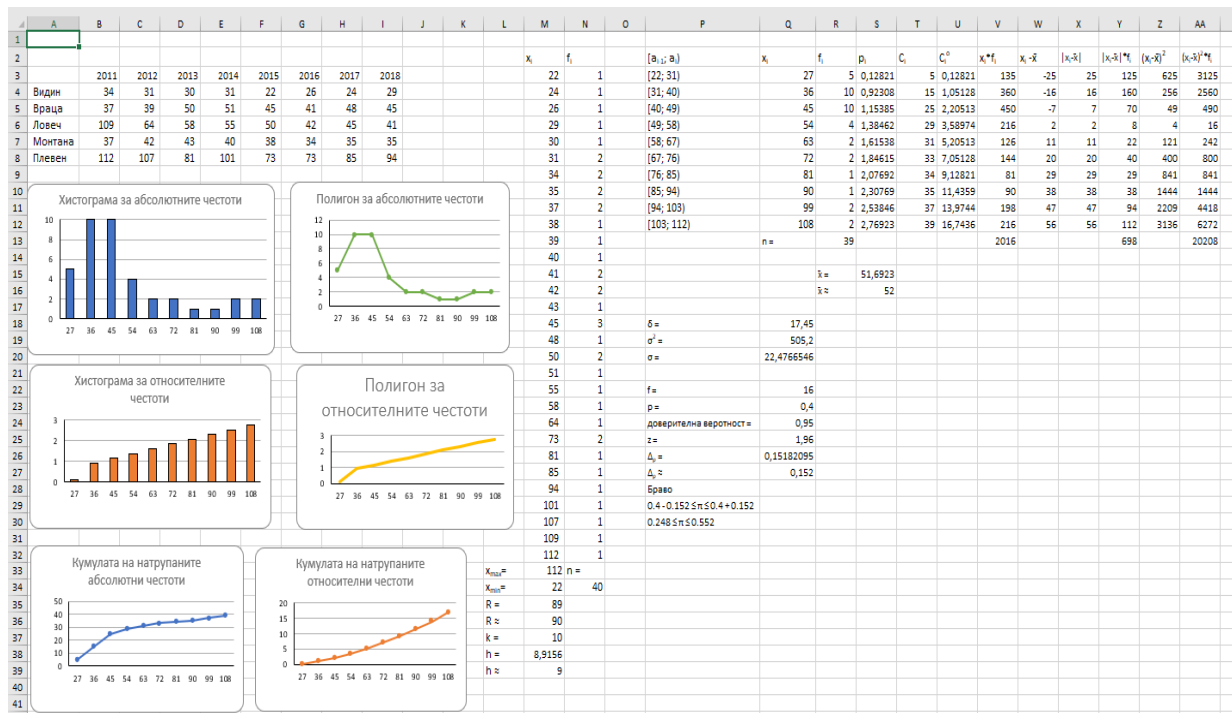


Fig. 1. Completed project of students in Statistics

Using the sum function in MS Excel, sum all the values of f_i , to determine the total number n and use it to find the relative frequency. They apply the mathematical formula for relative frequency $p_i = \frac{f_i}{n}$. In cell S3, enter the formula = R3/\$R\$13, using a relative address in the numerator (changes when copied) and an absolute address in the denominator (does not change). This allows you to copy the formula to the other cells in the column (possibly by dragging). Highlight the columns x_i and f_i and with Excel's chart design option, create histograms and polygons of absolute and relative frequencies.

For the third task, accumulated absolute frequency is needed – C_i and accumulated relative frequency C_i^0 . The accumulated absolute frequency column is created as $C_1 = f_1$, and $C_2 = C_1 + f_2$ etc. Copy the entry for C_2 formula (= R4 + T3) with relative addresses in the remaining cells of the column. Copy this formula into the accumulated relative frequency column. It automatically changes = S4 + U3, to use relative frequency values $C_2^0 = C_1^0 + p_2$.

They independently determine which of the studied statistical formulas to use: for arithmetic mean value – $\bar{X} = \frac{1}{n} \sum_{i=1}^k x_i f_i$; for linear deviation – $\delta = \frac{1}{n} \sum_{i=1}^k |x_i - \bar{X}| f_i$; for standard deviation – $\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^k (x_i - \bar{X})^2 f_i}$.

They introduce additional columns $x_i \cdot f_i$; $x_i - \bar{X}$; $|x_i - \bar{X}|$; $|x_i - \bar{X}| \cdot f_i$; $(x_i - \bar{X})^2$, $(x_i - \bar{X})^2 \cdot f_i$. The column $x_i \cdot f_i$ fill in the formula = Q3.R3. They add up the values with the sum button. In a separate cell (S15) they divide this value by the number n to find the average. In a column $x_i - \bar{X}$ introduce = Q3.V\$13/S\$15. For the next column, they use the built-in ABS function for absolute value. They add up the values at the end of the columns. $|x_i - \bar{X}| \cdot f_i$ and $(x_i - \bar{X})^2 \cdot f_i$, to calculate the linear and standard deviation in separate cells.

Students have chosen a confidence probability 0,95 and from a table of values with a normal (Gaussian) distribution they determined the confidence factor $z = 1,96$.

The formula used to calculate the maximum permissible error is $\Delta_p = z \cdot \sqrt{\frac{p(1-p)}{n}}$, which is also calculated with MS Excel.

Then, a confidence interval was constructed using the formulas $-\Delta_p \leq \pi \leq p + \Delta_p$.

The transformation of mathematical formulas into formulas suitable for the MS Excel environment uses the knowledge of students from information technology classes. They must decide which cells to use to find the desired values. They use relative, absolute and mixed addresses to automate the filling of columns. Depending on the level of the students, the teacher assesses whether they can cope without help or if they need it.

The conducted anonymous survey reports increased interest and motivation for studying the topic of Statistics from all students. They realize the applicability of this science in real life. They believe that they have updated, strengthened and expanded their skills in the field of information technology. The mentor teacher observes increased results in project-based learning implemented in the MS Excel environment.

CONCLUSION

ICT supports sustainable teaching strategies by increasing engagement and supporting diverse learning styles. MS Excel allows learners to engage in individual and collaborative problem-solving activities, promoting active learning. Learners develop communication, collaboration, information literacy, technological literacy and statistical literacy skills through hands-on activities.

When processing real data with MS Excel, increased motivation, engagement and activity of learners have been observed. This leads to good results in the students' learning activity in mathematics and information technology. The combination of project-based learning on the topic of Statistics with the use of software builds on mathematical, digital and statistical competence.

ACKNOWLEDGEMENTS

This research is supported by project 2025-FPNO-03, funded by the Scientific Research Fund of the University of Ruse.

I sincerely thank Prof. D. Galabova, PhD for the shared experience and knowledge.

REFERENCES

American Statistical Association (2005). Guidelines for Assessment and Instruction in Statistics Education in Statistics Education (GAISE): College Report 2005.

Alismail, H., & McGuire, P. (2015). 21st Century Standards and Curriculum: Current Research and Practice. *Journal of Education and Practice*, 6(6), 150-154.

Arkadiiev, D. (2015). Between tradition, modernity and practice: some views on teaching statistics to economists. *Statistics*. 2, 71-94. Sofia: National Statistical Institute. (**Оригинално заглавие:** Аркадиев, Д. (2015). Между традицията, модерността и практиката: някои

виждания за преподаване на статистика на икономисти. *Статистика*. 2, 71-94. София: Национален статистически институт.)

URL: https://www.nsi.bg/en/file/16552/sp_STATISTIKA_2_2015.pdf

Basturk, R. (2005). The Effectiveness of Computer-Assisted Instruction in Teaching Introductory Statistics. *Educational Technology & Society*, 8(2), 170-178.

Borkulo, S., Chytas, C., Drijvers, P. et al. (2023). Spreadsheets in Secondary School Statistics Education: Using Authentic Data for Computational Thinking. *Digital Experiences in Mathematics Education* 9, 420–443. <https://doi.org/10.1007/s40751-023-00126-5>

Burckhardt, P., Nugent, R., & Genovese, C. (2021). Teaching Statistical Concepts and Modern Data Analysis with a Computing-Integrated Learning Environment. *Journal of Statistics and Data Science Education*, 29(sup1), S61–S73.

GAISE College Report ASA Revision Committee. (2016). Guidelines for Assessment and Instruction in Statistics Education College Report 2016.

Harits, M., Sujadi, I., & Slamet, I. (2019). Technological, pedagogical, and content knowledge math teachers: to develop 21st century skills students. *Journal of Physics: Conference Series*, 2-7. doi:10.1088/1742-6596/1321/3/032011.

Karakoleva, S., Dimitrova, D. (2021). Use of cloud technologies for training and test control on the topic "Statistics and data processing" in tenth grade. *Proceedings of University of Ruse*. 60 (6.4) 79-84. ISSN 2603-4123 (on-line).

Karakoleva, S., Koleva, M. (2024). Application of cloud technologies in education on the topic "Probabilities and Statistics" in twelfth grade. *Proceedings of University of Ruse*. 63 (6.4) 113-121. ISSN 2603-4123 (on-line).

Kazika, G., Malambo, P. (2025). Effectiveness of Excel-Assisted Instruction in Teaching Statistics for Learners' Acquisition of 21st Century Skills. *Journal of research, innovation and implications in education*, 9(1), 254-264 ISSN 2520-7504 (Online).

Legacy, C. Le, L., Zieffler, A., Fry, E., Vivas Corrales, P. (2024). The Teaching of Introductory Statistics: Results of a National Survey. *Journal of Statistics and Data Science Education*. 32:3, 232-240, DOI: 10.1080/26939169.2024.2333732

Moore, D. (1997), New Pedagogy and New Content: The Case of Statistics. *International Statistical Review*, 65, 123-137.

Sugareva, M. & Murgova, M. (2021). On the need to increase statistical literacy and culture. *Statistics*. 2, 41-54. Sofia: National Statistical Institute. (**Оригинално заглавие:** Сугарева, М., Мургова, М. (2021). За нуждата от повишаване на статистическата грамотност и култура. *Статистика*. 2, 41-54. София: Национален статистически институт.)

Sugareva, M., Borisova-Marinova, K., Moraliyska-Nikolova, S. (2025). *Statistics at school. To help mathematics teachers*. Varna: Slavena. 120 p. (**Оригинално заглавие:** Сугарева, М., Борисова-Маринова, К., Моралийска-Николова, С. (2025). *Статистика в училище. В помощ на учителите по математика*. Варна: Славена. 120 с. ISBN 9786191902705.)

Tintle, N., Chance, B., Cobb, G., Rossman, A., Roy, S., Swanson, T., VanderStoep, J. (2016). *Introduction to Statistical Investigations*. John Wiley & Sons: Hoboken, NJ, USA, ISBN 1119683459.

Zapryanov Z., Marinkova M. (2008). Collection of problems and tests in combinatorics, probability and statistics. Sofia: Trud, 356 p. (**Оригинално заглавие:** Запрянов З., Маринкова М. (2008). Сборник от задачи и тестове по комбинаторика, вероятности и статистика. София: Труд, 356 с. ISBN 9789545288159.)

Zapryanov Z., Marinov M. (2018). Fun Mathematics for 10th, 11th and 12th Grades. Practical Course in Probability Theory, Combinatorics and Statistics. Sofia: Scorpio, 240 p. (Оригинално заглавие: Запрянов З., Маринов М. (2018). Забавна математика за 10., 11. и 12. клас. Практикум по теория на вероятностите, комбинаторика и статистика. София: Скорпио, 240 с. ISBN 9789547928046.)