EduDX

Connecting Inquiry and Information (2)

~Series: Practices Toward "Data-Driven Inquiry" as an SSH-Designated School~

Daiki Ito

Oita Maizuru High School

29 August 2024

1. The Necessity of Statistics

As various initiatives toward Society 5.0 progress, the importance of learning statistics and data science in high school subjects "Information I" and "Information II" has been emphasized as central to data-driven learning activities that guide actions and decision-making[1]. In this context, starting in the 2022 academic year, courses related to "Mathematical, Data Science, and AI" were introduced in the first-year curriculum at universities. These courses aim to provide systematic education on knowledge and skills related to these fields to enhance university students' interest in and foundational understanding of mathematical, data science, and AI, while fostering practical abilities to solve problems. Given these trends, learning statistics with an eye on high school-university connections has become increasingly important, and further accumulation of knowledge in this area is necessary. Furthermore, from the perspective of "data collection, analysis, and utilization" in inquiry-based learning and research projects, the position of the high school subject "Information" is considered significant. A review of the Course of Study Guidelines for Information reveals that the content related to data utilization is well-developed, indicating expectations for future information education to address literacy in "statistics and data science."

However, regarding the statistical learning activities within the "Data Utilization" unit of "Information I", there appear to be few practical and universally applicable case studies available. One factor contributing to this is the difficulty of conducting data analysis using individual devices. The primary environment for data processing and statistical analysis is often confined to computer labs, which depend heavily on environmental factors. In individual-device environments, tablets are predominantly used, and limitations in operation and input methods make statistical processing challenging. This has hindered the transition of statistical learning to inquiry-based activities conducted in various classrooms. Additionally, the use of statistical software requires considerable effort to teach operational methods and data

interpretation. It is also heavily influenced by the number and performance of student computers available in computer labs. These challenges have led to concerns among many high school information teachers about teaching statistics. Furthermore, connections to Information II's "Data Science" curriculum and the "Mathematical, Data Science, and AI" courses positioned in university-level education remain areas of concern.

Table 1: Learning Content Related to Data Utilization as Described in the Course of Study Guidelines (Information Edition)

Information I	Information II
・ Qualitative and Quantitative Data ・ Data Organization (e.g., Data Cleaning) ・ Handling Missing Values and Outliers ・ Various Methods of Data Visualization ・ Text Mining ・ Utilization of Open Data ・ Multiple Scatter Plots (Scatter Plot Matrix) ・ Correlation and Confounding Factors ・ Simple Linear Regression Analysis	・ Selection Bias and Information Bias ・ Machine Learning, Training Data, Overfitting ・ Multiple Regression Analysis, Model Adjustment, Polynomial Regression ・ Classification, Conditional Probability, Nearest Neighbor Methods, Tree-Based Predictions ・ Clustering ・ Character Recognition, Image Recognition, etc. ・ Artificial Intelligence

To address these issues, a web application and web manual were developed at an SSH-designated school to enable statistical learning activities and student-driven exploratory statistical analysis and machine learning that are not constrained by temporal, spatial, or technical limitations. Additionally, through trial lessons, specific strategies for curricula focused on "data-driven inquiry activities" and "statistics" were explored. These lessons aimed to examine concrete approaches for fostering information literacy and other related factors through practical teaching methods.
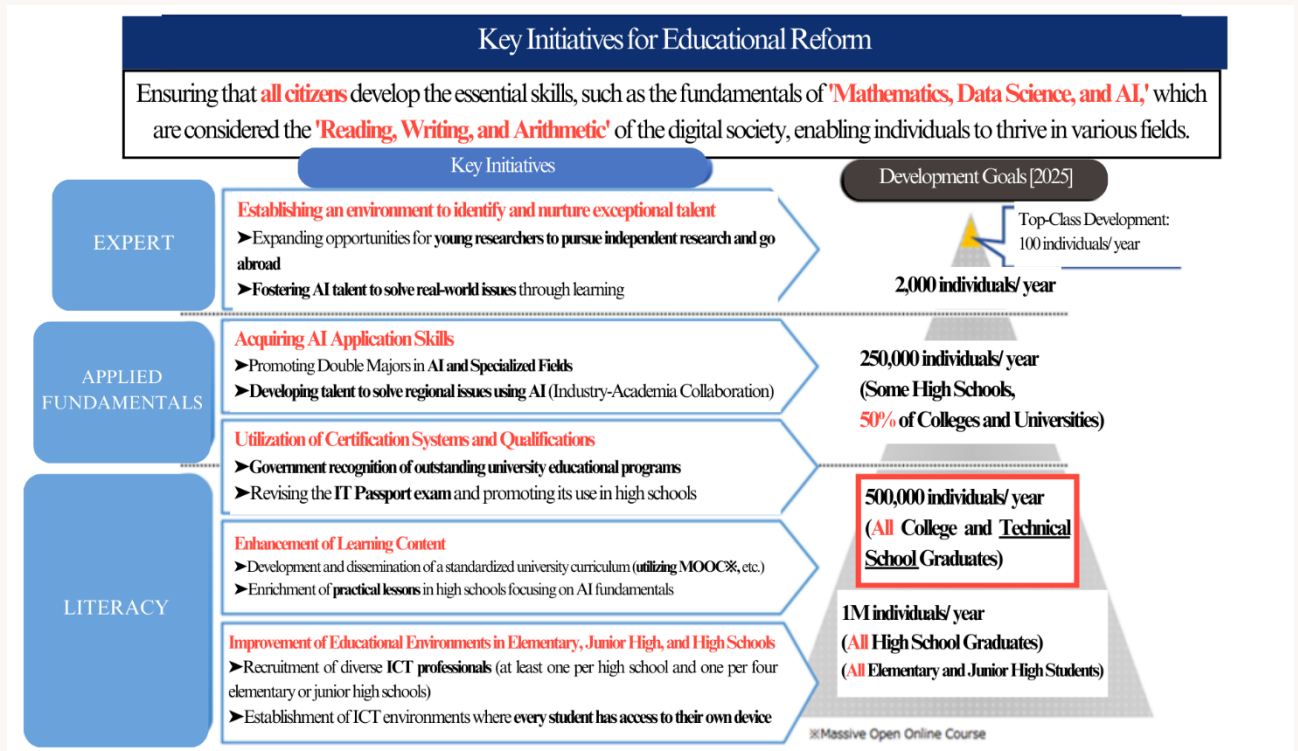
Figure 1: Education on "Mathematics, Data Science, and AI"[3]

In the next section, I will discuss the curriculum and web applications that were developed.

References

1. Shinya Ohashi: "Data Science in the Common High School Subject 'Information'," *Computer & Education*, Vol. 52, No. 1, pp. 3–12 (2022).
2. Yuya Shimoji et al.: "Teachers' Perceptions of the Revised Curriculum Guidelines for the Common High School Subject 'Information'," *Educational Information Research*, Vol. 36, No. 1, pp. 3–12 (2020).
3. Ministry of Education, Culture, Sports, Science and Technology: *"On the Development of AI Human Resources Based on the AI Strategy,"* p. 2 (2019). URL: https://www5.cao.go.jp/keizai-shimon/kaigi/special/reform/wg7/20191101/shiryou2_1.pdf