

Project-Based Teaching Reform of "Multivariate Statistical Analysis and R Language Modeling" Based on OBE Concept

Shimin Liang*, Yuyi Luo, Junhao Huang

School of Computer Science, Guangdong University of Science and Technology, Dongguan,
Guangdong, China

*Corresponding Author

Abstract: "Multivariate Statistical Analysis and R Language Modeling" is a professional group elective course for students majoring in Data Science and Big Data Technology, designed for those aiming to pursue careers in data analysis and data mining. In traditional teaching, there are issues such as deviations in teaching goal orientation, outdated and lagging teaching methods, and rigid teaching evaluation criteria in the teaching of "Multivariate Statistical Analysis and R Language Modeling". Teaching reform should be carried out under the guidance of the OBE concept and project-based teaching model. Determine teaching goals based on demand, design a dual-module teaching content of "basic accumulation + practical application", adopt a project-based teaching model, and construct a continuously improving assessment and evaluation system, providing a reference for improving the teaching quality of "Multivariate Statistical Analysis and R Language Modeling".

Keywords: OBE Concept; Project-Based Teaching; Multivariate Statistical Analysis and R Language Modeling

1. Issues in the Teaching of "Multivariate Statistical Analysis and R Language Modeling"

"Multivariate Statistical Analysis and R Language Modeling" is a professional group elective course in the Data Science and Big Data Technology major, specifically designed for students planning to pursue careers in data analysis and data mining. The goal is to cultivate data analysts with solid professional skills. This course serves as a bridge between previous and subsequent courses. The prerequisite courses include "Statistical Analysis", "Machine Learning", "Data

Visualization", etc., laying a theoretical and skill foundation for course learning. The subsequent courses include "Multivariate Statistical Analysis and R Language Modeling Course Design", "Graduation Thesis (Design)", "Graduation Internship", etc., serving as a crucial link for students to transform theoretical knowledge into practical abilities and connect with the workplace. As a comprehensive, applied, and practical course, its teaching quality directly affects students' professional competence and employment competitiveness. However, in the traditional teaching mode, there are still many prominent issues in the teaching of this course, as follows:

1.1 Deviation in Teaching Objective Positioning

The core objective of this course is to cultivate data analysts who meet industry needs. However, traditional teaching models overly emphasize knowledge impartation and basic skill training, neglecting the cultivation of process and method objectives. This leads to students finding it difficult to deeply integrate the theoretical knowledge they have learned with practical business scenarios. There are significant shortcomings in practical abilities such as solving practical problems, writing analytical reports, and presenting project reports. The unreasonable teaching objective orientation has caused a disconnect between the ability level of students after graduation and the actual requirements of corporate positions, making it difficult for them to smoothly transition from campus to the workplace and quickly adapt to the job demands of data analysis-related positions.

1.2 Outdated Teaching Methods in Traditional Teaching

Traditional teaching models mainly adopts the outdated teaching mode of "lecture +

demonstration". The teaching process overemphasizes the leading role of teachers and neglects the principal role of students, leading to insufficient classroom interaction and low student participation. This outdated teaching method is difficult to adapt to the learning characteristics of college students in the Internet era, unable to effectively stimulate students' interest in learning and enthusiasm for active exploration. It is not conducive to cultivating students' comprehensive practical abilities and does not match the characteristics of the course, which are "comprehensiveness, applicability, and practicality".

1.3 Rigid Evaluation Criteria

Traditional teaching evaluation criteria overly emphasize final exam scores, while assigning very low weights to students' usual classroom performance, project reports, and project presentations. In terms of evaluation dimensions, they mainly focus on knowledge mastery and basic skill application, with a severe lack of evaluation on students' ability to solve practical problems. Rigid evaluation criteria are difficult to fulfill the guiding, feedback, and improvement functions of teaching evaluation, unable to comprehensively and objectively reflect students' true learning outcomes and professional abilities. It is also not conducive to guiding teaching models towards optimization that emphasizes practice and highlights abilities.

In summary, due to issues such as deviations in teaching goal orientation, outdated and lagging teaching methods, and rigid teaching evaluation criteria, the teaching effectiveness of the "Multivariate Statistical Analysis and R Language Modeling" has failed to truly achieve the core goal of "cultivating data analysts with solid professional abilities". To effectively reverse this situation, overcome teaching difficulties, further improve the quality of teaching, and meet the needs of industry talents, it is urgent to explore new teaching paths that are suitable for the comprehensive, applied, and practical characteristics of this course.

2. Construction of Project-Based Teaching Mode for the "Multivariate Statistical Analysis and R Language Modeling" Based on the OBE Concept

OBE (outcomes based education), as an advanced education concept, its core meaning is

to completely break the traditional "teaching centered" teaching logic, clearly emphasize the student-centered and learning results oriented, and take the ultimate ability improvement of students as the starting point and foothold of the whole education and teaching activities [1]. At the beginning of teaching activities, teachers and students will jointly define the preset learning outcomes [2], that is, after a complete course learning process, students should master the core knowledge of multivariate statistics, the flexible use of R language modeling skills, and the comprehensive qualities of data analysis, logical reasoning, team cooperation and so on. On this basis, the teaching team carries out reverse teaching design around the preset learning results, accurately defines the core objectives of each teaching link, scientifically screens the teaching content, optimizes the teaching methods, plans the implementation path, and closely links each teaching activity with the preset results, effectively breaking the situation of blind learning, fuzzy objectives, and disconnection between knowledge and application in traditional teaching, so as to make the whole teaching process more targeted and effective [3]. At the same time, OBE always insists on taking students as the main body of learning [4], abandoning the assessment mode of relying solely on course examination results and focusing on knowledge memory in traditional teaching, and instead comprehensively evaluating the learning effect through multiple dimensions such as students' actual practice performance, project work quality, problem solving ability and innovative thinking [5], which can more comprehensively and objectively evaluate students' comprehensive ability, and fully meet the core needs of "emphasizing practice, strengthening ability and seeking effectiveness" in the cultivation of applied talents.

Project based teaching is a kind of teaching mode with actual projects as the core carrier and ability training as the fundamental goal [6]. Its core connotation is to break the limitations of knowledge fragmentation and disconnection between theory and practice in traditional teaching [7], and to deeply integrate and organically link the teaching content with the project tasks in the real work scene [8]. This teaching mode takes the complete project implementation process as the main line [9], discards the traditional indoctrination teaching

logic of "teachers' teaching and students' passive acceptance", and guides students to actively participate in the whole process of "demand analysis - data acquisition - data preprocessing - data analysis and modeling - data visualization - analysis report". In the course of multivariate statistical analysis and R-language modeling, project-based teaching adheres to the principle of student-centered and project driven, emphasizing "learning in practice and improving in exploration". Students are required to comprehensively apply the theoretical knowledge of multivariate statistics (such as regression analysis, cluster analysis, classification analysis, etc.) and R-language modeling skills (such as data preprocessing, data visualization, model construction and optimization, etc.), actively think and cooperate to solve all kinds of actual data analysis problems encountered in the process of project implementation, and ultimately realize the effective absorption of knowledge, repeated training of skills and comprehensive improvement of comprehensive quality, so as to achieve the teaching goal of "combination of learning and application, and application of learning".

OBE concept and project-based teaching mode have a high degree of adaptability and complementarity. The organic combination of the two can effectively solve many prominent problems faced by the course of multivariate statistical analysis and R language modeling in teaching practice. Among them, the OBE concept points out the core direction for teaching activities, defines the core issues of "what to do", and ensures that teaching activities are always carried out around the ability training through the preset learning results and the reverse design of teaching links, so as to avoid teaching deviating from the core needs of application-oriented talent training; The project-based teaching mode provides specific methods and paths for teaching design and implementation, and solves the practical problem of "how to do" [10]. Through the real project carrier, the preset learning achievements in the OBE concept take root, and the abstract theoretical knowledge is transformed into specific practical ability. The two support each other and work together. They not only play the goal oriented role of OBE concept to ensure that teaching does not deviate from the core training objectives, but also use the practical

advantages of project-based teaching to enable students to achieve preset learning outcomes in practice, effectively solve the problems of disconnection between theory and practice, students' weak practical ability and single evaluation method in teaching, promote the transformation of teaching from "knowledge imparting" to "ability training", comprehensively improve the quality of teaching and students' comprehensive application ability, and provide strong teaching support for the cultivation of applied talents.

3. Project-Based Teaching Reform Strategy for the "Multivariate Statistical Analysis and R Language Modeling" Based on the OBE Concept

Guided by the concept of OBE and based on the requirements of the core position of data analyst, the research team carried out the project-based teaching reform of multivariate statistical analysis and R language modeling. The course reversely designs teaching objectives based on job requirements, breaks the traditional teaching material framework, constructs a dual module content system of "basic accumulation + practical application", realizes the deep integration of theory and practice by relying on project-based teaching, introduces real industrial projects and competition scenes to strengthen students' comprehensive application ability, and establishes a closed-loop evaluation mechanism of "forward assessment + reverse evaluation", so as to promote the continuous optimization of teaching process, content and evaluation.

3.1 Determining Teaching Objectives Guided by Job Requirements

The research team, based on the actual needs of social positions, combined with the school's educational orientation and the talent cultivation plan for data science and big data technology majors, adopted a reverse design approach to formulate the teaching objectives. The corresponding position for the "Multivariate Statistical Analysis and R Language Modeling" is data analyst. After thorough research and practical verification, the job requirements for data analysts are clearly defined as follows: (1) proficiency in mainstream data analysis tools such as SPSS, R, and Python, with the ability to independently write data analysis reports; (2) proficiency in

the entire process from data collection, cleaning, analysis to visualization; (3) solid grasp of conventional statistical methods and techniques, with basic understanding and application abilities in mathematical modeling and data mining algorithms.

Based on the aforementioned job requirements, the research team has clarified the teaching objectives of "Multivariate Statistical Analysis and R Language Modeling" as follows: (1) Understand and master the core concepts, basic principles, and complete implementation process of multivariate statistical analysis and R language modeling; (2) Proficiently use commonly used libraries such as tidyverse, caret, and plotly in R language, and possess practical abilities in data processing, data analysis and modeling, and data visualization; (3) Be able to comprehensively utilize relevant R language libraries to systematically process, analyze, and model real data in conjunction with actual needs; (4) Master the standardized writing methods of data analysis reports and project reporting skills, and enhance professional expression and results presentation abilities.

3.2 Designing a Dual-Module Teaching Content System of "Basic Accumulation + Practical Application"

To achieve the established teaching objectives, the research team designed a dual-module teaching content system of "basic accumulation + practical application". In the process of refining the teaching content, the inherent limitations of textbook content were broken - textbooks are often compiled following systematic principles, with comprehensive coverage of knowledge points, but for courses like "Multivariate Statistical Analysis and R Language Modeling" that emphasize application and practicality, their theoretical orientation is difficult to fully play a guiding role in teaching. Therefore, the research team adheres to the orientation of job requirements and adopts a reverse design approach to optimize the teaching content, ensuring that all teaching content is centered around the expected learning outcomes and precisely aligns with the teaching objectives.

To effectively enhance students' ability to comprehensively utilize R language-related libraries to process real-world data, the teaching content is organized strictly in accordance with

the actual workflow of data analysis and modeling, namely "demand analysis - data acquisition - data preprocessing - data analysis and modeling - data visualization - analysis report". Emphasis is placed on strengthening the teaching of skills such as data preprocessing and data visualization, precisely tackling key and difficult issues in data analysis and modeling, ensuring that the teaching content is highly aligned with actual job tasks, and facilitating a seamless transition from theoretical learning to practical job operations for students.

3.3 Adopting a Project-Based Teaching Model

Based on a dual-module teaching content system of "basic accumulation + practical application", this course adopts a project-based teaching model to conduct teaching, deeply integrating theoretical knowledge and practical skills into various projects to achieve "combining learning with doing and applying what one has learned". The basic accumulation module uses simple projects as a carrier, integrating the core knowledge and basic skills that students need to master into the entire process of project implementation; the practical application module adopts comprehensive projects covering the complete process of "demand analysis - data acquisition - data preprocessing - data analysis and modeling - data visualization - analysis report", comprehensively honing students' practical operational abilities in their positions.

In the foundational accumulation module, the research team consistently integrates the project of second-hand car data analysis and price prediction, accurately aligning various knowledge points with sub-projects and advancing them simultaneously. Students systematically grasp data preprocessing content such as handling missing values, outliers, and duplicates while completing the sub-project of second-hand car data preprocessing. They solidly learn the usage skills of data visualization libraries such as ggplot2 and plotly while carrying out the sub-project of second-hand car data visualization. In advancing the sub-project of second-hand car price prediction, they gradually master multivariate statistical analysis and modeling methods such as linear regression, decision tree, support vector machine, neural network, and

random forest. In the completion of the second-hand car data visualization and price prediction application development sub project, the theoretical learning and practical ability of shiny application development are improved simultaneously.

The competition is highly aligned with students' post-graduation job scenarios in terms of content setting and implementation format, serving as an important way to strengthen practical abilities. Based on this, in the practical application module, the research team introduced two comprehensive projects from the Kaggle competition: telecom customer churn prediction and housing price prediction. Both projects fully cover the entire process of data analysis and modeling, fitting the actual work requirements of data analysts. During the teaching process, students complete the entire process of project implementation, analysis report writing, and project reporting in groups according to the requirements of the practical training guidebook. They enhance their ability to apply learned knowledge to solve practical problems in practice, while cultivating team collaboration awareness, honing language expression and result presentation skills. In the practical application link, the students' subjective position is fully highlighted, while teachers transform into project mentors and process organizers, providing targeted guidance for students' project implementation throughout the process to ensure teaching effectiveness.

3.4 Building a Sustainably Improving Assessment and Evaluation System

The research team of "Multivariate Statistical Analysis and R Language Modeling" has established a bidirectional assessment and evaluation system consisting of "forward assessment + reverse evaluation" to achieve a closed-loop evaluation and continuous optimization. The forward assessment focuses on students' learning outcomes, integrating both regular and final grades for comprehensive evaluation. Regular grades are assigned based on a combination of classroom performance, project reports, project presentations, and other process-oriented outcomes; final grades are determined through an open-book computer-based test, with practical questions accounting for 50% of the score, emphasizing students' ability to apply theoretical knowledge to solve practical data analysis problems. This

assessment scheme is clearly communicated to students at the beginning of the course, guiding them to shift from rote memorization to deep understanding and flexible application, effectively enhancing learning outcomes and comprehensive qualities. The reverse evaluation component involves conducting multi-dimensional surveys on student satisfaction with teaching and corporate satisfaction with students, extensively incorporating feedback from both students and the industry. This continuously optimizes assessment content, teaching methods, and teaching evaluation standards, promoting the transformation of the course assessment and evaluation system from quality monitoring to sustainable improvement, and facilitating steady improvement in teaching quality.

4. Conclusions

Project-based teaching reform of the "Multivariate Statistical Analysis and R Language Modeling" based on the OBE concept addresses issues such as deviations in teaching goal positioning, outdated and lagging teaching methods, and single and rigid teaching evaluation standards in traditional teaching. It explores project-based teaching reform based on the OBE concept. Guided by the OBE concept and utilizing project-based teaching as a vehicle, the reform determines teaching goals oriented towards job requirements, designs a dual-module teaching content of "basic accumulation + practical application", hones students' practical skills through basic projects and comprehensive competition projects, and constructs a diversified evaluation system of "forward assessment + reverse evaluation". This reform effectively addresses the dilemmas of traditional teaching, achieves a shift from "knowledge impartation" to "ability cultivation", improves the quality of teaching, meets industry employment needs, provides a reference for the cultivation of applied talents in related majors, and will continue to optimize the effectiveness of reform in conjunction with industry development in the future.

Acknowledgments

This paper is supported by Guangdong University of Science and Technology "Quality Engineering" Higher Education Teaching Reform Project (No. GKZLGC2025044).

References

- [1] Killen R. Outcomes-based education: Principles and possibilities. Unpublished manuscript, University of Newcastle, faculty of education, 2000: 1-24.
- [2] Malan S P T. The 'new paradigm' of outcomes-based education in perspective. *Journal of Family Ecology and Consumer Sciences= Tydskrif vir Gesinsekologie en Verbruikerswetenskap*, 2000, 28(1): 22-28.
- [3] Jansen J D. Curriculum reform in South Africa: A critical analysis of outcomes-based education. *Cambridge journal of education*, 1998, 28(3): 321-331.
- [4] Son D, Homma M, Pickering J. Bridging the Gap in Medical Education: A Qualitative Study on the Perspectives of Japanese Medical Students and Patients on Outcomes-Based Education. *Cureus*, 2025, 17(10):e95527-e95527.
- [5] Shi W, Shi X, Zhuang L, et al. Integrating Outcome-Based Education (OBE) and Problem-Based Learning (PBL) in psychiatric clinical practice teaching: an empirical study. *BMC Medical Education*, 2025, 25(1):1610-1610.
- [6] Kokotsaki D, Menzies V, Wiggins A. Project-based learning: A review of the literature. *Improving schools*, 2016, 19(3): 267-277.
- [7] Larmer J, Mergendoller J R. Seven essentials for project-based learning. *Educational leadership*, 2010, 68(1): 34-37.
- [8] Handrianto C, Rahman M A. Project based learning: a review of literature on its outcomes and implementation issues. *LET: Linguistics, Literature and English Teaching Journal*, 2018, 8(2): 110-129.
- [9] Guerra A, Jiang D, Du X, et al. Examining engineering design students' perceptions of agency for sustainability in a problem- and project-based learning (PBL) context—a Q methodology study. *International Journal of Technology and Design Education*, 2024, 35(2): 1-24.
- [10] [10]Ma X, Wu X. Establishing a Seamless Integrated Project-Based Learning Framework Mediated by an Evidence-Based Project-Based Learning System. *Sustainability*, 2025, 17(5): 2325-2325.