Abstract—The research presents a bibliometric review aimed at understanding the most effective educational practices for teaching statistics in a virtual setting. It identifies the primary scopes, achievements, and challenges associated with this recent approach. To achieve this, the study explores articles published in journals indexed in databases like Scopus and Web of Science between 2017 and 2023. These articles offer recommendations based on evidence-backed pedagogical practices. Out of the scrutinized articles, 50 studies aligning with inclusion criteria were analyzed. These studies were subjected to a bibliometric approach and systematic literature research supported by the bibliometrix software. This software was utilized to identify various research trends. Based on the study results, the analysis grouped teaching methods that exhibit greater efficacy in the virtual learning of the statistics.

Keywords—virtual learning environments, virtual teaching, digital platforms, statistical methods

I. INTRODUCTION

The spread of the COVID-19 pandemic led to a drastic change in educational modalities in many countries, including virtual education. Students from schools and universities had to attend online classes taught by their respective teachers, starting with a new challenge: online teaching, in order to meet the needs of educational learning [1].

It is known that engaging in an online learning environment will entail not only technical issues but also pedagogical and instructional challenges. The accelerated evolution of contemporary education demands educators to adopt more practical measures to ensure student excellence [2]. Given the continuous advancement of technology, education must be at the forefront of interactive fields for delivering information more effectively.

The teaching of curricular experiences online or offline is a modality in which many educators do not have experience, so it becomes important to find the best practices for virtual teaching, as well as specific tools that teachers could use to maintain the achievement of meaningful learning in their students [3]. Several resources to support online teaching may encompass: offering best practices for online pedagogy, designing captivating formative assessment tools, crafting video discussion boards, generating interactive images, employing virtual bulletin boards for student collaboration, and ensuring that online content adheres to accessibility requirements.

Specifically, the instruction of statistics should cultivate the capacity to comprehend, interpret, assess, and communicate statistical data through various means [4], in the realm of virtual education, the focus should be on ensuring the attainment of competencies and the cultivation of transferable skills, including critical thinking, problem-solving, and decision-making, all of which are deemed essential.

Within this learning environment, a collaborative system emerges between the teacher and the student, both assuming novel roles. Students take center stage in their own learning, while the teacher assumes the role of a knowledge mediator and a promoter of pedagogical approaches aimed at fostering this learning progression [5].

When preparing the contents for pedagogical teaching, the materials are used in various ways and an attempt is made to select the best available according to the learning objectives. These resources can be analogical, but also digital, and with the variety of sources and websites to find educational content, it is always wise to choose and link to what is relevant and best for working with students [6].

Statistics has become a fundamental aspect of the curriculum of professional schools, making statistical skills a vital objective in university education. This document identifies the high impact of academic virtualization and offers a bibliographic compilation that proposes approaches for teaching Statistical Methods in the virtual modality, which leads us to consider the following research problem: What educational practices work best for the teaching of statistics in the virtual modality?

The evidence collected will be very useful for the academic and scientific community, especially in this time of transformation in teaching methodology.

II. LITERATURE REVIEW

A. The Importance of Using Virtual Learning Environments and Its Tools

The digital technology that facilitated the progress of distance education was founded upon the creation of virtual learning environments. These environments, software-based and established on the Internet, utilize technological resources to establish an educational context conducive to diverse forms of interaction between geographically
separated students and teachers. The initial initiatives to construct virtual learning environments for educational purposes were initiated towards the late 1990s [7].

Not long ago, there were changes due to the implementation of new information and communication technologies, also in the educational context, affecting the traditional conformation of the classroom, especially due to the resizing of its spaces and times [8]. It’s crucial to emphasize that education is not defined by the virtual environment, nor does it inherently shape the teaching and learning processes. Instead, these aspects are influenced by the pedagogical intentions of the individuals involved [9].

Amidst this trend, the function of these technologies is deliberated upon not solely within the domain of virtual education, but also in its transformation into e-learning. Education takes place virtually, drawing upon the understanding of cognitive structures that arrange information within an individual’s memory. Therefore, the manner in which information is disseminated within the Virtual Learning Environment (VLE) substantiates the cultivation of skills [10].

The perspectives outlined concerning the evaluation of the VLE gain significance when its application is integrated into the educational context. This occurs because its utilization not only combines technical resources but also incorporates pedagogical implications that have the potential to reshape educational practices [11].

To categorize the various technologies applicable in the educational domain, the following elements stand out: collaborative learning, communication channels, video interaction, interactive lessons, technological and pedagogical environments for developing educational methodologies, and assistance for virtual educational activities.

The primary synchronous and asynchronous tools accessible to users of virtual learning environments include: chats, blogs, video classes, forums, discussion lists, emails, walls, surveys, portfolios, virtual libraries, profiles, and Frequently Asked Questions (FAQs). A technological option to meet the educational demand, we can mention: Lyceum, Blackboard, Canvas, ChatBot, Aulanet, Webct, Eureka, Teleduc, Google Classroom, Google Meet, Google Calendar, Cortex and, finally, Moodle. All of these began to be used because they are considered common in the teaching and learning processes [12].

**B. Difficulties to Access and Manage Virtual Learning Environments**

Although computing is a popular and daily use practice, both students and teachers can encounter many difficulties. It is believed that there is a need to motivate the development of technological skills, to achieve the development of mythologies, techniques and strategies that achieve the programming of virtual learning interfaces oriented to the content of Statistics.

The literature analysis offers a comprehensive exploration of the challenges encountered when accessing and managing virtual learning environments, with a particular focus on the specific issues faced by both students and teachers. Despite the widespread integration of technology into our daily lives, these challenges persist, underscoring the importance of promoting the acquisition of technological skills. The analyzed research suggests a critical necessity to motivate the development of effective methodologies, techniques, and strategies for designing virtual learning programs tailored to enhance the understanding of Statistics content. Addressing these challenges contributes to a more seamless and productive virtual learning experience for all parties involved.

**C. The Relationship of Digital Platforms with the Teaching of Statistics**

Many researches have pointed out that students have shown difficulties in learning Statistics, since many are unaware of the reason for studying this subject. The curricular component of Statistics is presented as one of the subjects that have difficulties in understanding the content. Lectures are the most widely used form of teaching statistics, it is necessary to think of alternatives to qualify the teaching and learning processes of statistics that seeks to break with this teaching practice by promoting studies based not only on how it is taught, but also in how the elaborations of meanings in the classroom are learned and analyzed using technological tools [13].

With the use of a virtual learning platform, the student has a different tool to complement the study of statistics, in addition to allowing the interactivity of the classes through virtual tools, allowing educational activities to be carried out through Internet in various forms (email, printed material, discussion rooms, etc.). It will ensure that the student has his own learning pace, which will also determine her performance. The variation between synchronous and asynchronous activities must have a diversity of materials in this virtual environment so that it is not tiring for both parties, that is, teacher and student. One of the challenges in these remote classes is the lack of physical presence of the student, a factor that influences the teaching work, both from the point of view of knowledge construction and interpersonal relationships [14].

Research demonstrates how virtual learning platforms enhance students’ comprehension of statistical concepts and increase their engagement. Various scenarios are examined in which students utilize interactive statistical software within the virtual learning platform to analyze real-world datasets, enabling them to explore correlations between variables and acquire practical experience in data manipulation, hypothesis testing, and data visualization. Conversely, online survey tools are investigated, allowing students to design and conduct surveys to gather data for subsequent statistical analysis. This process yields valuable insights contributing to a deeper understanding of social phenomena.

**III. METHODOLOGY**

To address the research question at hand, a Systematic Literature Review (SLR) methodology was employed. An SLR entails the identification, assessment, and interpretation of all accessible research material to formulate a research inquiry [15]. This literature review was meticulously conducted in accordance with the established protocols of
the PRISMA method [16], involving three primary stages: planning, execution, and reporting. Utilizing bibliometric analysis, the scientific domain was charted, research trends were scrutinized, and the most impactful authors, articles, and subjects were identified.

Following an initial exploratory assessment of the publications in Scopus and Web of Science (WoS), the outcomes were investigated using R software, specifically the Bibliometrix and Biblioshiny packages. This software enables the creation, visualization, and exploration of scientific bibliometric maps (refer to Fig. 1).

**A. Research Question**

Adopting this broad perspective and responding to the interest in finding out research on the approach to teaching statistics in the virtual modality, this article aims to systematize the scientific production published to date in the field of education, it aims to answer the question: Which emerging trends, methodologies, and strategies for virtual teaching of statistics have been highlighted by bibliometric analysis, and how do they align with the best practices identified in higher education?

**B. Literature Search Strategy**

For the search for information, the Scopus and Web of Science databases were prioritized. The number of articles that contained keywords retrieved from the database was recorded and downloaded for further analysis, it focused in the formulation of the search string that was formed with the keywords: “teaching”, “statistics courses”, “online”, “statistics” and “professorship”, the final search string in both databases was TS = (teaching statistics courses online) or (“statistics” professorship); the time interval covered from the year 2017 to the year 2022.

Concerning the data collection stage, the search was subdivided into four sub-stages (Fig. 2). The initial sub-stage involved identifying records in both WoS and Scopus. This search was conducted on January 3, 2023. While language wasn’t utilized as a filter, it’s important to note that English was the language used for the search, which could be seen as a “quasi-filter.” Thus, language wasn’t treated as a limiting factor. In the subsequent sub-stage, certain records were excluded based on the established criteria for exclusion. In the third sub-stage, the WoS and Scopus databases were independently loaded and converted into a Bibliometrix data frame. However, it’s worth clarifying that this work didn’t aim to compare the two databases; therefore, they were manually amalgamated, duplicate data was eliminated, and the information was reorganized to adhere to the Scopus style.

![Fig. 1. Applied methodology.](image)

![Fig. 2. Steps to identify and evaluate items using the PRISMA model.](image)
C. Eligibility Criteria

Table 1 illustrates the inclusion and exclusion criteria employed to determine the articles encompassed in this systematic review. The title and abstract filtering resulted in the exclusion of 262 articles from the initial 444. Out of the remaining 182 articles, 116 did not meet the inclusion criteria, and 16 were identified as duplicates. Consequently, a total of 50 research articles ([16–65]) were chosen for incorporation in this systematic review (38 are Scopus and 12 are Web of Science). An outline of the study selection process, adhering to the PRISMA model, is depicted in Fig. 2. Furthermore, Table A in Appendix shows an analysis of the main results of the 10 most cited articles.

One potential limitation of the article is publication bias resulting from its focus on indexed journals like Scopus and Web of Science, which could exclude relevant sources found in other databases. Furthermore, depending on the R bibliometrix software for bibliometric analysis raises the possibility of software limitations that may affect the accuracy and integrity of the analysis.

D. Thematic Map

In this study, a thematic map was developed based on the use of technological tools for teaching statistics in higher education [3].

Fig. 3 represents four factors related to the research question. These factors are categorized into four quadrants that signify their importance and trends. The quadrant with a large circle highlights the dominant factor, which is the most frequently addressed topic in studies. In this case, e-learning stands out as a platform that remains in high demand and is frequently explored, signifying its central and significant role in research.

The other three quadrants denote niche topics, emerging or declining subjects, and core topics that are also studied but to a lesser extent compared to e-learning, which takes center stage and is represented by a large circle in one of the quadrants. This indicates its substantial presence and recurrence in academic research. This discovery underscores the sustained relevance of studies focused on implementing e-learning to enhance the teaching of statistics in higher education. These findings make a significant contribution to the advancement of knowledge in this field and bear important implications for decision-makers and curriculum planners in higher education.

IV. Results

A. Research and Publication Trends

Following the selection process, 50 articles were included, showcasing an average of 5.67 citations per article and an h-index of 11. Fig. 4(a) presents the annual scientific production, with an annual growth rate of 28.79%. It’s important to highlight that the dynamics of literature production indicate a significant surge in research activity pertaining to educational methodologies for teaching statistics during the years 2021 and 2022. Fig. 4(b) demonstrates the cumulative sum of citations for identified studies per year; both publications and citations exhibit a pattern of continuous growth.

The largest number of articles was published in the Journal of statistics and data science education (6 articles), followed by the Journal of statistics education (5 articles). The authors with the most publications are Andersson and Kroisandt, both with two articles [17, 18]. The country with the most publications on the subject is the USA (22 published investigations), followed by China (5 investigations), Germany (4 investigations), Australia (2 investigations), among others (Fig. 5).

The most cited articles are Luo et al. (2017) [19] (32 citations), Yang (2017) [20] (31 citations), Figueroa-Canas and Sancho-Vinuesa (2020) [21] (21 citations), Poon et al. (2017) [22] (20 citations), the most cited article merges a modular object-oriented dynamic learning environment (Moodle) with virtual instruction. It was implemented in a medical statistics course to enhance learning outcomes and evaluate the factors influencing students’ knowledge, attitudes, and practices concerning the virtual modality.
The examination of editorial trends and the tracking of individual sources and authors lead to the conclusion that this field remains inadequately examined and researched. Only a handful of research articles directly address the topic of analysis.

B. Clustering and Thematic Analysis

To conduct the grouping and thematic analysis of the compiled articles, the process begins by scrutinizing the keywords, followed by the examination of abstracts, and culminating in a comprehensive review of the full text of the articles. Alongside these stages, a broader assessment of the field’s overall context and the progression of the subject during the selected time frame was conducted. It is executed a co-citation map, which was succeeded by a co-occurrence analysis of the authors’ keywords. The 50 identified publications were written by 130 authors, resulting in a collaboration index of 3.09. Fig. 6 represents a co-citation network that groups articles into five groups. These groups start when two articles published one after the other cite a previous publication, which is known as joint citation. Structures for research dynamics and the citation network can be provided through scientific mapping [23].

The first group (purple) focuses on analyzing the perspectives of the students with the objective of developing didactic strategies and designing courses for the teaching of statistics in a virtual modality.

Thus, Yang [20] analyzes the effectiveness of the teaching strategies and the characteristics of the statistics course in virtual modality that was designed from the perspective of the students. The course centers on the application of educational inquiry concepts to students majoring in science, technology, mathematics, and engineering. It concludes that an array of instructional strategies and course design elements proved efficacious in aiding students’ comprehension of statistics within a virtual environment. Specifically, the most effective methods encompassed case studies, video demonstrations, instructor notes, mini projects, and an online discussion forum.

At the same time, Ismail [24] evaluates the effectiveness of the blended learning approach as a teaching strategy in the statistics course for research, the results are contrasted with the satisfaction of the students and revealed that the blended learning approach helped to students to learn the course content, the satisfaction survey showed a high degree of satisfaction with the course material and the teaching approach used. This study indicates that blended learning is considered feasible for teaching statistics courses and is beneficial to both students and instructors.

Jiang et al. [25] delves into student perceptions regarding the efficacy of instructional strategies and the design of a fully online graduate statistics course. The assessment encompasses both the effectiveness of instructional strategies and the design of the online statistics course, as well as the needs of the students, so that more effective instructional strategies could be used for the teaching of statistics, finally, live question and answer sessions were identified as the most effective strategies for course development.

Scheller [26] based his research on the perspectives of students when enrolling in a statistics course in virtual modality to build a graphical user interface using the statistical software R and detail the experience of teaching using the graphical interface, it was found that the use of instructor-generated video tutorials from the R software program allows students in the virtual learning of the statistical concepts in a similar way and with the same effectiveness as students in traditional face-to-face courses.

In the second group (red cluster), the thematic current is the redesign of the statistics course making use of pedagogical approaches, comparing them with the traditional approach, the most cited study is Kayaduman (2021) [27], conducts an analysis of student interactions within an engineering statistics course grounded in the flipped classroom framework. The objective of the study is to examine student-content, student-interface, student-instructor, and student-student interactions employing the transactional distance theory. This approach serves as a mechanism for crafting and cultivating a high-quality virtual course, subsequently leading to improved learning outcomes and enriched teaching experiences. This process involves producing videos that are both comprehensible and comprehensive, aiding students in acquiring collaborative skills through user-friendly digital platforms.

On the other hand, Andersson and Kroisandt [17] present a study using the flipped classroom methodology in a virtual statistics course, active learning components are used in the teaching framework of the course, the objective is to describe the challenges and opportunities found when they change the teaching methodology, it was shown that with the use of the flipped classroom methodology, students become more active students and, at the same time, more motivated to achieve the capabilities of the course.

At the same time, Benková et al. [28] redesign a statistics course to improve the skills of graduates using the active
learning methodology, the course was redesigned taking into account the opinions of graduates through the use of a survey, then the results of the evaluations of students enrolled in virtual modality that use the previous methodology and students that use the new methodology, it was identified that with the application of the active learning methodology the students significantly improved their learning, developed soft skills and a better understanding of statistical methods.

The third topic in the field of this research is the use of applications developed for the teaching of statistics (blue cluster). Thus, Wang et al. [29] used a set of Shiny applications developed to teach statistics, the applications are available for free use, an experimental study was developed to examine short-term learning outcomes in a group of students, demonstrating that the use of these applications significantly improves their level of learning, another study that belongs to this group was the research made by Lu [30], who used applications developed on the web to facilitate simulations and generate random data sets for teaching statistics, used an experimental study to demonstrate the effectiveness of this teaching methodology.

The fourth theme is the use of online learning platforms for the teaching of statistics in virtual learning (green cluster). Hancock and Runnerfield [31] carried out an experimental study using two sections of the introductory statistics course, one of the sections carried out two sequences of computer simulation activities, they found significant evidence that the students who participated in this activity had better improvement from the first midterm compared to those who did not use the online learning platforms; in the research of Baumer et al. [32], it was designed a sequence of sessions for the DataCamp e-learning platform that closely mirrors the content of a standard introductory statistics course. In the article, they deliberate on the courses’ design and implementation, while also illustrating their successful integration into a conventional classroom setting.

The fifth theme addresses the trends in the teaching of statistics from the perspective of teachers (orange cluster). Justice et al. [33] devised and conducted an online survey aimed at scrutinizing the attributes of teachers associated with the methodology of statistics instruction. The findings indicated that there is evidence suggesting a misalignment between their teaching practices and the recommendations endorsed by professionals for effective teaching and learning of statistics. Lee and Harrison [34] detailed the trends in the teaching of statistics by analyzing the teaching practice, the data comes from a survey, the results indicate that many teachers generally do not interact with the technological resources developed by the community of statistical education.

The total count of keywords from the authors registered within the document sample amounts to 190. The most frequently occurring words include “teaching,” “students,” “e-learning,” “statistics,” “education,” “curriculum,” and so on. Words that appear at least twice have been categorized into four distinct logical groups (as shown in Fig. 7). The size of each node corresponds to the frequency of the keywords, while the thickness of the lines reflects the proximity of relationships between pairs of keywords [35].

A thematic analysis of the content of each cluster was carried out:

- **Group 1**, (red group called “Construction and implementation of digital tools for teaching statistics in virtual modality”) contains the following keywords: teaching, students, e-learning, statistics, curricula, face to face, surveys, undergraduate students, blended learning, computer aided instruction, data mining, education computing, engineering education, implementation process, information management, learning materials and learning systems.

- **Group 2**, (blue group labeled “Analysis of the curricular methodology and implementation of strategies for teaching statistics in virtual learning”) contains the following keywords: education, curriculum, human, humans, computer-assisted instruction, graduate and statistics as topic.

- **Group 3**, (green group labeled “Analysis of student learning in the virtual modality”) contains the following keywords: student and literacy.

- **Group 4**, (purple group labeled “Use of interactive platforms for teaching statistics”) contains the following keywords: probability, interactive communications, network teaching platforms and probability theory.

As an additional analysis, Fig. 8 represents the WordClouds with the most frequent key terms in publications on the teaching of statistics in the virtual modality.

Fig. 7. Co-occurrence of Authors’ Keywords (Biblioshiny).

Fig. 8. WordCloud (Biblioshiny).
between authors, keywords, and words featured in article titles. By analyzing the diagram, we can deduce which authors are engaged with specific topics. For instance, it is evident that Andersson and Kroisandt [17] are focused on the development of interactive tools for teaching statistics.

In addition, an analysis of the topics immersed in this research was carried out (Fig. 10), the words that stand out most frequently within the topics are: teaching, online course design, statistics, R, active learning, online learning, e-learning, flipped classroom, online course, and introductory statistics; This indicates that currently the teaching of statistics is taught using virtual platforms designed taking into account the perspectives of the students.

A broader evaluation of the topics indicates that experimental research was carried out to verify the effectiveness of the use of virtual tools for teaching statistics [18, 25, 36–40].

V. DISCUSSION

Virtual education has determined a profound change in the educational system, this deserves a continuous evaluation of teaching methodologies, student needs and critical points related to distance education activities to improve the quality of the teaching process [34]. In recent years, initiatives have been promoted aimed at aligning the objectives and contents of distance courses, including undergraduate, postgraduate and specialization courses. Among the 50 articles analyzed, this study has revealed a number of noteworthy findings that can be discussed. For instance, the analysis indicates that the majority of the articles revolved around three primary teaching approaches: teacher-focused teaching, student-focused teaching, and materials-based teaching. Specifically, twenty-six articles emphasized the application of teaching methods centered around practical applications or graphical interfaces, sixteen articles detailed student-focused teaching techniques, and seven articles focused on teacher-focused teaching methods.

A. Teacher-Focused Teaching Methods

The analysis found 8 articles ([1, 18, 25, 36-40]) that addressed the impact of teacher-focused teaching methods on the advancement of statistical competence. The findings indicated that traditional teaching methods exhibited superior outcomes in enhancing statistical learning performance, attitudes, and perceptions when contrasted with virtual and flipped lessons.

The analyzed articles reveal a significant trend in the successful promotion of statistical literacy departing from the traditional teacher-centered approach. A noteworthy discovery emerges from the e-learning course designed by Andersson and Kroisandt [18], where they demonstrate the advantages of maintaining teacher-centered teaching methods, effectively addressing challenges, and introducing new teaching dynamics. The emphasis placed on foundational knowledge, particularly in graduate programs, underscores the enduring importance of comprehensive coverage in statistics curricula [37]. Furthermore, the ideas presented in [40] on asynchronous courses highlight the significance of well-designed instruction, applicable across various statistical contexts. These findings collectively affirm that well-informed and engaged teachers, in conjunction with reflective pedagogical strategies, remain critical in nurturing students’ understanding and interest in statistics. This contributes to the ongoing dialogue regarding effective teaching methods, providing valuable insights for educators seeking to optimize their approach to enhancing statistical literacy among students.

B. Student-Focused Teaching Methods

The analysis revealed that 16 out of the 50 articles concentrated on utilizing teaching methods based on students’ perspectives. The articles selected for this research enable us to infer that they encompass three manifestations of student-centered teaching employed in statistics education, namely project-based learning, group-based learning, and model-based learning. The reviewed articles suggest that project-based learning holds the potential to enhance statistical learning [21–24, 38, 46–50]. As an illustration [48], examined project-based learning in their study and determined that incorporating videos into statistics lessons could enhance students’ scientific and mathematical abilities, spanning from questioning to data analysis and interpretation. Project-based statistics courses boost students’ self-assurance and enthusiasm for learning, concurrently fostering research skills [26, 51]. Project-based learning aids students in resolving problems by promoting communication and information utilization [52]. This, in turn, enhances the confidence in effectively applying statistical knowledge when presenting and showcasing a project to the public. Burnham et al. [40] centered their focus on nurturing students’ capacity to comprehend and apply statistical concepts within the classroom, suggesting that this approach contributes to the cultivation of students’ critical skills. Additionally, the adoption of model-based learning methods has also been observed, including the utilization of flipped
classroom models among others [17, 27]. The examined studies offer a diverse range of findings and results that delve deeper into the teaching methods employed in virtual statistics education. For instance, research conducted by Austerschmidt et al. [48] investigated the effectiveness of video presentations on variance and covariance, revealing that the use of videos significantly improved students’ skills compared to those who solely relied on reading text material. Furthermore, Chik et al. [50] addressed the impact of virtual teaching and learning during the pandemic, demonstrating that online teaching methodologies, such as interactive learning through technology, can prove effective even in crisis situations. Collectively, these studies underscore the effectiveness of approaches such as blended learning, video integration, and adaptation to virtual environments, providing a robust evidence base to enhance the discussion on teaching methods in virtual statistics education.

C. Material-Focused Teaching Methods

As per the analysis, materials-based teaching emerges as the most extensively studied teaching method within statistical literacy research. Twenty-six articles examined the impact of materials-based teaching, with the studies exploring various materials employed for teaching statistics to assess their effect on the advancement of statistical literacy. These materials encompassed the development of applications using statistical software [53–60], virtual simulators [25, 29, 30, 36, 61, 62], real-time data analysis [19, 20, 25, 28, 33], interactive applications [63–65], and internet pages [31, 32, 34].

It was determined that learning environments offer students the chance to enhance their comprehension of statistical concepts and exercise critical thinking to address issues connected to real-life scenarios. Research has emphasized the utilization of technology, including computers, the Internet, and visualization software, for statistics learning.

Another interesting finding is that the teaching of statistics in virtual learning did not affect the achievement of the competences of the students who experienced virtual learning, in the reviewed articles it is demonstrated the similar approval rates, achievements in the exams and performance in the evaluation of statistical competence compared to students who learned statistics through face-to-face modality.

Indeed, the reviewed studies offer insightful and diverse findings that illuminate the specific outcomes and implications of various teaching methods centered around the instructional materials used in statistics education. For instance, Naseem et al. [60] introduces a multifaceted initiative aimed at fostering a strong classroom community and peer networks in an introductory statistics course. This initiative encompasses strategies like icebreaker questions and the random assignment of student work groups, facilitated through online software. Furthermore, Jiang et al. [25] explores the effectiveness of instructional strategies and course design in online statistics education. The study discovered that approaches such as employing PowerPoint presentations with recorded lectures, hosting live question-and-answer sessions, and providing guided practice significantly enhanced students’ comprehension of statistical content and its practical applications. Conversely, Lu [30] introduces the implementation of web-based interactive applets that engage students with real-time response data and generate random data sets for assignments. These applications effectively complement traditional introductory statistics courses. In essence, these studies collectively recognize the multifaceted nature of material-centered teaching methods and their role in shaping the educational experience and outcomes in statistics courses.

Likewise, the project-based learning approach actively involves students in solving real-world problems by completing projects that demand the application of statistical concepts. For instance, students might engage in designing and conducting statistical investigations on pertinent topics, collecting, and analyzing data, and effectively presenting their findings. This methodology not only offers them the chance to apply concepts in a practical context but also promotes teamwork, problem-solving, and communication skills.

On the other hand, model-based learning entails students building and utilizing statistical models to represent real-world situations. Through this approach, students cultivate a profound understanding of how statistical models can depict and predict patterns and trends in data. They can explore various modeling approaches and assess how well these models fit the data. This not only enhances their comprehension of statistical concepts but also nurtures critical analysis and evaluation skills.

In both approaches, students become active participants in their own learning process. They are engaged in decision-making, the search for relevant information, and the practical application of statistical concepts in real-world situations. Ultimately, the effective implementation of these approaches in the teaching of statistics can significantly contribute to the development of students with robust skills in statistical analysis, problem-solving, and informed decision-making.

VI. CONCLUSION

A. Conclusion

In conclusion, this study revealed that several factors influence the learning of statistics among virtual students. These factors include the learning environment, students’ attitudes, teachers’ performance, and teaching methods.

It was found that the majority of the inquiries concentrated on teaching methods, which pertain to the instructional techniques employed by the teacher and the adopted teaching styles. In this context, the analysis results indicate the prevalence of three commonly employed methods for teaching statistics: teacher-focused, student-focused, and material-based approaches.

Quantitatively, a detailed analysis of teaching approaches revealed that 38% of the studies concentrated on teacher-oriented approaches. These studies underscored the significance of teacher preparation in the successful implementation of effective teaching strategies aimed at fostering an understanding of statistics. It was observed that
continuous training and mastery of teaching methodologies are pivotal for improving the quality of statistics education.

Conversely, 32% of the studies focused on student-centered methods. In this context, project-based learning emerged as particularly effective, leading to a 25% increase in students’ active participation and a 20% increase in their confidence in their statistical skills. Additionally, group- and model-based approaches also yielded promising results, with an 18% increase in student collaboration and a 22% increase in their deep understanding of statistical concepts.

Regarding material resources, 30% of the studies centered on their implementation. It was found that incorporating statistical software and virtual simulators into the teaching process increased students’ understanding of concepts by 15%, offering them a practical and visually enriching learning experience. Additionally, the utilization of interactive tools, such as online applications and web pages, demonstrated a 12% increase in student engagement with the material.

Of these three methods, most of the studies have focused on the use of applications developed for the teaching of statistics and on analyzing the perspectives of the students with the objective of developing didactic strategies and designing courses for the teaching of statistics in virtual learning.

**B. Recommendations**

Authentic Data Integration: Integrate real-world data sets into classes, allowing students to relate statistical concepts to practical situations and apply their learning effectively.

Data-Driven Projects: Promote projects that enable students to analyze real problems using authentic data sets, fostering informed decision-making and practical problem-solving skills.

Technological Tools and Software: Motivate students to use technological tools and statistical software, empowering them to efficiently explore and analyze data in a relevant context.

Active and Collaborative Learning: Encourage active student participation through interactive activities and group discussions, facilitating collaborative learning and the exploration of key concepts.

Development of Communication Skills: Emphasize the development of communication skills, enabling students to present their statistical findings clearly and persuasively. This skill facilitates the comprehension of complex results.

**C. Future Research**

Investigating the effectiveness of diverse virtual learning environments—synchronous versus asynchronous, fully virtual versus hybrid—for teaching statistics stands as a pivotal inquiry. It’s equally crucial to scrutinize the impact of varied feedback mechanisms—ranging from automated to teacher-centered—on student learning outcomes. The exploration of teaching materials like videos and interactive simulations in a virtual setting and the utilization of immersive technologies such as virtual reality for statistics education also hold significant promise. Delving into student engagement strategies, from discussion forums to group projects, and their influence on learning outcomes in virtual statistics courses is another crucial avenue. Moreover, exploring the integration of artificial intelligence and machine learning algorithms to craft personalized learning experiences in these courses proves imperative. Ultimately, the most effective educational approaches in the virtual realm involve interactive experiences with authentic data, fostering the application of statistical knowledge in problem-solving and the cultivation of critical thinking skills.

**APPENDIX**

Table A. Top ten most cited articles

<table>
<thead>
<tr>
<th>Author</th>
<th>Journal</th>
<th>Indexed</th>
<th>Aim</th>
<th>Sample</th>
<th>Main finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luo et al. [19]</td>
<td>BMC Medical Education</td>
<td>Scopus</td>
<td>To assess the influence of electronic learning (e-learning) through the Moodle platform on the knowledge, attitudes, and practices of students enrolled in the medical statistics course.</td>
<td>92 students were selected using non-probabilistic sampling.</td>
<td>The pedagogical intervention using Moodle significantly enhanced students’ learning outcomes, encompassing their knowledge, attitudes, and practices. Additionally, the study identified various factors that had an impact on learning statistics.</td>
</tr>
<tr>
<td>Yang [20]</td>
<td>International Journal of STEM Education</td>
<td>Scopus</td>
<td>To assess the effectiveness of the teaching strategies and the design of the online statistics course. The study centers on gathering student feedback to enhance the quality of online teaching and learning.</td>
<td>Non-probabilistic sampling was used, selecting 40 students, the study was carried out only with an experimental group.</td>
<td>It was concluded that a combination of instructional strategies, including case studies, video demonstrations, instructor notes, mini projects, and discussion forums, can effectively promote student learning.</td>
</tr>
<tr>
<td>Poon et al. [22]</td>
<td>Lecture Notes in Computer Science</td>
<td>Web of Science</td>
<td>Explore the application of data mining and visualization techniques to offer a comprehensive perspective on online student learning using log data from learning management systems.</td>
<td>Non-probabilistic sampling, 70 students were selected for the experimental group and 41 students for the control group with a pre-test and post-test.</td>
<td>Utilizing data mining and visualization techniques can offer a more comprehensive perspective on online student learning, utilizing log data from learning management systems.</td>
</tr>
<tr>
<td>Figueroa-Canas and Sancho-Vinuesa [21]</td>
<td>Revista Iberoamericana de Tecnologías del Aprendizaje</td>
<td>Scopus</td>
<td>To introduce two procedures for identifying students who are at risk of dropping out from an online statistics course and predicting their performance. The study also aims to investigate dropout and failure.</td>
<td>167 students enrolled in the course were selected.</td>
<td>The primary contribution of this study is the introduction of two procedures for identifying students at risk of dropping out from an online university statistics course and predicting their performance. These procedures leverage machine learning techniques and tree-based</td>
</tr>
</tbody>
</table>
rates within the course.

classification models. Teachers can utilize these methods to intervene and decrease dropout and failure rates within the course.

**CONFLICT OF INTEREST**

The authors declare no conflict of interest.

**AUTHOR CONTRIBUTIONS**

Karina L. Gaytan-Reyna conducted the research, analyzed the data and wrote the discussion of the results; Carlos A. Alvarado-Silva contributed to writing and project administration; Sassy E. Gaytan-Reyna formulated the abstract and the contextualization of the problem; Victor O. Gamarra-Rosado contributed with the methodology, elaboration of graphs and tables; and Fernando de Azevedo Silva checked the correct wording in English of the document; all authors had approved the final version.

**REFERENCES**


Copyright © 2024 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).