

Ambiente de aprendizaje para la enseñanza de Estadística II: un enfoque en teoría, prácticas y
autoevaluación

Jorge Eduardo Suárez Cortés

Daniel Alejandro Sánchez Rodríguez

Trabajo de Grado para optar al título de Ingeniero de Sistemas

Director

Andrés Leonardo González Gómez

PhD (c). Ciencias de la Computación

Universidad Industrial de Santander

Facultad de Ingenierías Fisicomecánicas

Escuela de Ingeniería de Sistemas e Informática

Ingeniería de Sistemas

Bucaramanga

2025

Tabla de Contenido

Introducción	10
1. Objetivos	11
1.1. Objetivo General	11
1.2. Objetivos Especificos	11
2. Marco Teorico	13
2.1. Estadística	13
2.1.1. Estadística Inferencial	14
2.2. Ambientes de aprendizaje	15
2.2.1. Políticas TIC en educación	15
2.2.2. Ambiente de aprendizaje	16
2.2.3. Ambientes virtuales de aprendizaje	17
2.3. Tecnologías implementadas en el proyecto	18
2.3.1. Node.js y Express.js	18
2.3.2. React.js	18
2.3.3. Bases de datos relacionales (MariaDB)	19
2.3.4. Virtualización y contenedores (Docker)	19
2.3.5. Moodle como LMS	19
2.3.6. Python y R	20
2.3.7. Google Colab	20

AMBIENTE DE APRENDIZAJE PARA LA ENSEÑANZA DE ESTADÍSTICA II	3
2.3.8. Graders y autoevaluación automatizada	21
2.3.9. Playit.gg y tunelización de servicios	21
3. Estado del Arte	22
3.1. Introducción	22
3.2. Antecedentes	22
3.2.1. La evolución de los LMS	22
3.2.2. Plataformas de Aprendizaje	24
3.2.3. Sistemas de Autoevaluación y Retroalimentación Automatizada	26
3.2.4. Arquitectura tecnológica para la educación	27
3.2.5. Antecedentes nacionales y la UIS	30
4. Metodología	33
4.1. Enfoque metodológico	33
4.2. Roles	33
4.3. Estructura de los sprints	34
4.4. Sprints	34
4.4.1. Sprint 0	34
4.4.2. Sprint 1	36
Referencias Bibliográficas	40

Lista de Figuras

Figura 1.	Estado del arte. Línea de tiempo con base en diversas fuentes históricas..	24
Figura 2.	Estado del arte. Plataformas de aprendizaje.	25
Figura 3.	Estado del arte. Arquitectura tecnológica para la educación.	28
Figura 4.	Estado del arte. Arquitectura de microservicios.	29
Figura 5.	Sprint 1. Esquema de los contenedores Docker.	38

Lista de Tablas

Glosario

Ambiente Virtual de Aprendizaje (AVA): Entorno digital que integra recursos, actividades y herramientas interactivas para apoyar los procesos de enseñanza y aprendizaje en línea o semipresenciales.

Autoevaluación: Estrategia pedagógica que permite a los estudiantes valorar su propio proceso de aprendizaje mediante actividades de retroalimentación automática.

Base de Datos: Conjunto estructurado de información almacenada de manera organizada que permite el acceso, la gestión y la actualización de datos académicos y de evaluación.

Docker: Plataforma de virtualización ligera que permite empaquetar aplicaciones y servicios en contenedores portables, facilitando su despliegue en diferentes entornos.

Estadística Inferencial: Rama de la estadística que utiliza datos muestrales para realizar inferencias, estimaciones o pruebas sobre parámetros poblacionales.

Google Colab: Herramienta en la nube que permite ejecutar código en Python y R desde un navegador, facilitando la realización de prácticas y análisis de datos sin necesidad de instalación local.

Grader: Sistema automatizado de evaluación que corrige ejercicios prácticos, especialmente de programación y estadística, generando resultados inmediatos.

LMS (Learning Management System): Sistema de gestión del aprendizaje que permite planificar, implementar y evaluar procesos educativos en entornos virtuales. Facilita la creación

de cursos, administración de contenidos, seguimiento del progreso de los estudiantes y la comunicación entre docentes y alumnos.

Moodle: Plataforma de gestión de aprendizaje (LMS, por sus siglas en inglés) de código abierto, utilizada para crear cursos en línea, administrar contenidos y facilitar la comunicación entre docentes y estudiantes.

Node.js: Entorno de ejecución de JavaScript orientado al desarrollo de aplicaciones web y servidores, utilizado para gestionar la comunicación entre sistemas del entorno virtual.

Playit.gg: Servicio que permite crear túneles seguros para acceder a aplicaciones locales desde internet, sin necesidad de configuraciones avanzadas en redes.

Resumen

Título: Ambiente de aprendizaje para la enseñanza de Estadística II: un enfoque en teoría, prácticas y autoevaluación.

*

Autores: Jorge Eduardo Suárez Cortés y Daniel Alejandro Sánchez Rodríguez. **

Palabras Clave: Ambiente virtual de aprendizaje, Moodle, Estadística II, Autoevaluación.

Descripción: El presente trabajo de grado describe el diseño, desarrollo e implementación de un ambiente de aprendizaje virtual para la asignatura Estadística II, dirigida a estudiantes de la Escuela de Ingeniería de Sistemas e Informática de la Universidad Industrial de Santander. El ambiente fue construido sobre la plataforma Moodle, con el objetivo de ofrecer una experiencia de aprendizaje más activa, significativa y continua mediante la integración sistemática de contenidos teóricos, ejercicios prácticos y mecanismos de evaluación formativa, tales como autoevaluaciones y retroalimentación inmediata.

El enfoque pedagógico del entorno se basa en el aprendizaje activo y la construcción progresiva del conocimiento, promoviendo la participación continua del estudiante y el desarrollo de competencias relacionadas con el análisis estadístico y la interpretación de datos. Las actividades propuestas están diseñadas para favorecer la autonomía del aprendizaje. Los resultados obtenidos sugieren que el uso del ambiente virtual contribuye a mejorar la comprensión de los contenidos estadísticos, así como a incrementar la motivación y el compromiso de los estudiantes con su proceso de formación. Esta propuesta representa una alternativa metodológica y tecnológica que responde a las necesidades actuales de la educación superior en el ámbito de la ingeniería de sistemas, especialmente en contextos donde la articulación entre teoría y práctica resulta esencial para el desarrollo de competencias profesionales.

* Trabajo de grado

** Facultad de Ingenierías Fisicomecánicas. Escuela de Ingeniería de Sistemas e Informática.
Director: Andrés Leonardo González Gómez

Abstract

Title: Learning Environment for the Teaching of Statistics II: A Focus on Theory, Practice, and Self-Assessment. *

Authors: Jorge Eduardo Suárez Cortés y Daniel Alejandro Sánchez Rodríguez. **

Key words: Virtual learning environment, Moodle, Statistics II, Self-assessment.

Description: This thesis describes the design, development, and implementation of a virtual learning environment for the Statistics II course, aimed at students of the School of Systems Engineering and Computer Science at the Industrial University of Santander. The environment was built on the Moodle platform, with the aim of offering a more active, meaningful, and continuous learning experience through the systematic integration of theoretical content, practical exercises, and formative assessment mechanisms, such as self-assessments and immediate feedback.

The pedagogical approach of the environment is based on active learning and the progressive construction of knowledge, promoting continuous student participation and the development of skills related to statistical analysis and data interpretation. The proposed activities are designed to encourage autonomous learning. The results obtained suggest that the use of the virtual environment contributes to improving the understanding of statistical content, as well as increasing student motivation and commitment to their training process. This proposal represents a methodological and technological alternative that responds to the current needs of higher education in the field of systems engineering, especially in contexts where the articulation between theory and practice is essential for the development of professional skills.

* Degree Work

** Faculty of Physicomechanical Engineering. School of Systems and Computer Engineering.
Director: Andrés Leonardo González Gómez

Introducción

1. Objetivos

1.1. Objetivo General

- Diseñar y desarrollar un entorno virtual para la asignatura de Estadística II, de la Escuela de Ingeniería de Sistemas e Informática de la UIS, que facilite la ejecución de un plan de aula propuesto para la práctica, el aprendizaje y la autoevaluación del contenido de la asignatura, mediante herramientas de Python y R.

1.2. Objetivos Específicos

- Diseñar un plan de aula modular para la asignatura Estadística II que organice los contenidos teóricos y prácticos en unidades interactivas basadas en Colab Notebooks, Python y R, e incluya calificadores automáticos y ejercicios con retroalimentación instantánea.
- Implementar un ambiente de aprendizaje virtual en la plataforma Moodle que integre las tecnologías definidas (Colab Notebooks, Python y R) y organice los contenidos teóricos y prácticos junto con materiales didácticos de apoyo, consolidando una experiencia formativa práctica y accesible para la asignatura Estadística II.
- Establecer instrumentos de medición para las actividades prácticas y autoevaluativas en Moodle (desarrolladas con Python y R), que permitan medir el logro de los resultados de aprendizaje durante el desarrollo del proyecto en el curso de Estadística II.
- Pilotear el ambiente virtual de aprendizaje diseñado, con sus contenidos interactivos e instrumentos de evaluación, en por lo menos dos grupos de Estadística II de la Escuela de Inge-

nería de Sistemas e Informática de la UIS, con el fin de validar su usabilidad, satisfacción y efectividad en el logro de los resultados de aprendizaje.

2. Marco Teorico

2.1. Estadística

La estadística constituye una herramienta fundamental en la ciencia, la ingeniería y la administración, al proporcionar métodos rigurosos para recopilar, organizar, presentar, analizar e interpretar datos con el fin de apoyar la toma de decisiones. En la actualidad, desempeña un papel crucial en el análisis de sistemas complejos, la mejora de procesos y el control de calidad en distintos entornos académicos, industriales y científicos.

De acuerdo con Montgomery y Hines (1996), la estadística moderna se divide en dos grandes áreas: estadística descriptiva y estadística inferencial. La primera se ocupa de técnicas que permiten resumir y describir de manera gráfica o numérica los datos disponibles, mientras que la segunda se orienta hacia la formulación de generalizaciones o la toma de decisiones sobre una población a partir de la información contenida en una muestra.

En el contexto de este proyecto, la atención se centra en la estadística inferencial, dado que la asignatura Estadística II aborda conceptos como la estimación puntual, la construcción de intervalos de confianza, las pruebas de hipótesis y el estudio de distribuciones muestrales. Dichos contenidos resultan esenciales para obtener conclusiones válidas a partir de datos incompletos o parciales. Al tratarse de una asignatura que combina teoría matemática con aplicaciones prácticas en problemas reales, se requiere no solo comprender los fundamentos conceptuales, sino también desarrollar competencias aplicadas. Por ello, este proyecto plantea el diseño de un entorno virtual interactivo que permita a los estudiantes ejercitar dichos conceptos con datos reales, actividades

prácticas y retroalimentación inmediata.

2.1.1. Estadística Inferencial

La estadística inferencial es considerada una de las ramas más potentes y aplicadas de la disciplina. Su propósito central es extraer conclusiones sobre una población a partir de la información proporcionada por una muestra. Según Montgomery y Hines (1996), la mayoría de las aplicaciones estadísticas en ciencia, ingeniería y administración incorporan procedimientos de inferencia y toma de decisiones.

Entre los conceptos más relevantes que sustentan esta disciplina se encuentran:

- **Distribuciones muestrales:** distribuciones de probabilidad de estadísticas como la media, la proporción o la varianza, que permiten describir la variabilidad esperada de los estimadores a partir de distintas muestras.
- **Estimación de parámetros:** procedimientos mediante los cuales se obtienen valores aproximados de parámetros poblacionales utilizando la información muestral.
- **Pruebas de hipótesis:** métodos formales que permiten aceptar o rechazar proposiciones acerca de parámetros poblacionales, basándose en la evidencia empírica extraída de una muestra.

2.2. Ambientes de aprendizaje

2.2.1. Políticas TIC en educación

El diseño de ambientes de aprendizaje en la educación superior no solo responde a fundamentos pedagógicos y tecnológicos, sino también a marcos normativos e institucionales que orientan su desarrollo. En este sentido, la Política Institucional de Tecnologías de la Información y la Comunicación (TIC) de la Universidad Industrial de Santander, establecida mediante el Acuerdo del Consejo Superior No. 051 de 2009, constituye un referente clave para la integración de recursos digitales en los procesos académicos.

Dicha política busca garantizar el acceso equitativo a la información, fomentar la innovación pedagógica y promover la incorporación de tecnologías abiertas e interoperables en la formación universitaria. Asimismo, establece como principios fundamentales el fortalecimiento de los procesos de enseñanza-aprendizaje, la inclusión de toda la comunidad académica en el uso de las TIC y la consolidación de una cultura institucional que aproveche estas herramientas para el mejoramiento continuo de la calidad educativa (Universidad Industrial de Santander, 2009).

De esta forma, el presente proyecto se enmarca en los lineamientos de la Política TIC, ya que aprovecha plataformas como Moodle y Google Colab para potenciar la enseñanza de la estadística mediante recursos digitales interactivos, autoevaluaciones y retroalimentación automática. Esto permite articular los objetivos pedagógicos con los principios institucionales de innovación, inclusión y acceso abierto al conocimiento.

2.2.2. Ambiente de aprendizaje

El concepto de ambiente de aprendizaje se refiere al conjunto de condiciones, espacios, interacciones y recursos que permiten y favorecen el desarrollo de procesos educativos efectivos. Este término alude a un escenario dinámico en el que los individuos desarrollan capacidades, competencias, habilidades y valores. Es decir, no se trata de un espacio fijo ni neutral, sino de un entorno que debe transformarse en función de las necesidades de los estudiantes y de las innovaciones educativas (Castro, 2019).

En este sentido, un ambiente de aprendizaje no se limita al aula física, sino que constituye una construcción pedagógica en la que intervienen aspectos sociales, culturales, tecnológicos y metodológicos. De acuerdo con Castro (2019), un ambiente de aprendizaje debe ser:

- Flexible y adaptable a diferentes contextos y tecnologías.
- Fomentar la interacción social y la construcción colectiva del conocimiento.
- Proporcionar recursos adecuados, incluidos los tecnológicos, que potencien las competencias estudiantiles.
- Promover el rol activo del docente como diseñador, facilitador y mediador del aprendizaje.

Desde esta perspectiva, los ambientes de aprendizaje se conciben como espacios dinámicos e integrales, que van más allá de la transmisión de información y buscan estimular la participación activa, la autonomía y la construcción colaborativa del conocimiento.

2.2.3. Ambientes virtuales de aprendizaje

El avance de las tecnologías digitales ha posibilitado la transición de los entornos tradicionales hacia los Ambientes Virtuales de Aprendizaje (AVA). Estos se definen como plataformas tecnológicas que integran recursos didácticos, actividades de aprendizaje, espacios de interacción y mecanismos de evaluación para facilitar procesos formativos en entornos digitales (Salinas, 2004).

En coherencia con las características planteadas por Castro (2019), los AVA son flexibles y adaptables a diferentes contextos, fomentan la interacción social y colaborativa, proporcionan recursos tecnológicos que amplían las competencias estudiantiles y posicionan al docente como un mediador y facilitador del aprendizaje.

Uno de los sistemas más representativos en este campo es Moodle, un Learning Management System (LMS) de código abierto que ha sido adoptado globalmente en educación superior debido a su enfoque pedagógico constructivista, su escalabilidad y su facilidad de integración con otras herramientas (Dougiamas & Taylor, 2003). Moodle permite gestionar cursos, estudiantes, actividades y evaluaciones, convirtiéndose en un eje articulador del proceso formativo.

En el contexto del presente proyecto, Moodle constituye el espacio institucional de referencia en la UIS, donde los estudiantes acceden a materiales de teoría, actividades interactivas y prácticas en Google Colab, además de las autoevaluaciones conectadas con el backend. Su importancia radica en que integra de manera armónica la teoría, la práctica y la retroalimentación automática, consolidando un entorno virtual de aprendizaje efectivo.

2.3. Tecnologías implementadas en el proyecto

2.3.1. *Node.js* y *Express.js*

Node.js es un entorno de ejecución basado en el motor V8 de Google Chrome que permite ejecutar JavaScript en el lado del servidor, diseñado bajo un modelo de I/O no bloqueante. Esto lo convierte en una herramienta adecuada para aplicaciones que requieren manejar múltiples solicitudes concurrentes, como plataformas de aprendizaje en línea (Tilkov & Vinoski, 2010).

Express.js es un framework minimalista para Node.js que facilita la construcción de APIs y aplicaciones web mediante la organización de rutas, middleware y controladores (Brown, 2019). En el proyecto, Node.js y Express.js cumplen el rol de backend principal, gestionando la validación de respuestas de estudiantes, el control de intentos y la comunicación con la base de datos MariaDB.

2.3.2. *React.js*

React.js es una biblioteca de JavaScript desarrollada por Facebook que permite construir interfaces de usuario basadas en componentes reutilizables (Banks & Porcello, 2017). Su arquitectura con virtual DOM optimiza la actualización de vistas, mejorando la eficiencia y la experiencia de usuario.

El frontend, desarrollado en React.js, centraliza la gestión de calificaciones y el control de intentos de los estudiantes. Su principal ventaja es la transparencia: ante cualquier discrepancia con la calificación calculada por el backend (Node.js), un instructor puede revisar tanto el código enviado por el estudiante como la nota obtenida. Esto facilita una verificación rápida y fundamentada,

garantizando la equidad en la evaluación.

2.3.3. Bases de datos relacionales (MariaDB)

MariaDB es un sistema de gestión de bases de datos relacional de código abierto, derivado de MySQL, que organiza los datos en tablas relacionadas bajo el modelo relacional propuesto por Codd (1970).

Su función en el proyecto es ser el repositorio central de información, donde se almacenan los registros de estudiantes, intentos de ejercicios, calificaciones y respuestas. Gracias a sus características de integridad y consistencia, MariaDB asegura la fiabilidad en el almacenamiento de los resultados generados en los procesos de autoevaluación.

2.3.4. Virtualización y contenedores (Docker)

Docker es una plataforma de contenedores ligeros que empaqueta aplicaciones y dependencias en un entorno aislado, lo que garantiza portabilidad y reproducibilidad (Merkel, 2014).

En el proyecto, Docker permite desplegar servicios como el backend, frontend, playit.gg y la base de datos en un entorno controlado, asegurando que las pruebas piloto y los despliegues sean consistentes sin importar el sistema operativo. Su uso facilita la escalabilidad y reduce los problemas de compatibilidad entre entornos de desarrollo y producción.

2.3.5. Moodle como LMS

Moodle es un sistema de gestión de aprendizaje (LMS) de código abierto ampliamente utilizado en la educación superior. Ofrece herramientas para gestionar cursos, usuarios, recursos y

actividades de evaluación (Dougiamas & Taylor, 2003).

En este proyecto, Moodle funciona como el entorno de acceso institucional provisto por la UIS, donde los estudiantes encuentran los enlaces a los notebooks de Google Colab, las actividades prácticas y los recursos teóricos. Moodle sirve como el punto de integración pedagógica que articula los recursos tecnológicos desarrollados.

2.3.6. *Python y R*

Python y R son lenguajes de programación consolidados en el campo del análisis estadístico y la ciencia de datos. Python, gracias a bibliotecas como NumPy, SciPy y Pandas, se ha convertido en una herramienta versátil para la enseñanza y aplicación de métodos estadísticos (McKinney, 2017). R, por su parte, es un lenguaje especializado en estadística y visualización de datos, ampliamente usado en contextos académicos (R Core Team, 2023).

En el proyecto, para el desarrollo de los graders se emplearán ambos lenguajes de programación, lo que permitirá dotarlos de una mayor versatilidad en su uso.

2.3.7. *Google Colab*

Google Colab es una plataforma en la nube que permite ejecutar notebooks de Python sin necesidad de instalar software localmente. Combina teoría, código ejecutable y resultados en un único entorno interactivo (Bisong, 2019).

Su integración en el proyecto permite a los estudiantes resolver ejercicios en tiempo real, ejecutar simulaciones y enviar sus respuestas al backend para validación. Además, Colab facilita el aprendizaje activo y autónomo.

2.3.8. Graders y autoevaluación automatizada

Los graders son sistemas de evaluación automática que comparan respuestas de estudiantes con soluciones esperadas, otorgando calificaciones y retroalimentación inmediata (Kurnia et al., 2001). Herramientas como nbgrader o Otter-Grader han demostrado su efectividad en la enseñanza de programación y estadística.

En este proyecto, los graders están implementados en el backend desarrollado con Node.js y conectados a los notebooks de Colab. De esta forma, los estudiantes reciben retroalimentación instantánea, lo que fomenta el aprendizaje autónomo y reduce la carga de corrección manual para los docentes.

2.3.9. Playit.gg y tunelización de servicios

Playit.gg es una herramienta que permite crear túneles seguros entre una máquina local y la web pública, facilitando la exposición de servicios locales sin necesidad de configuración compleja de red.

Durante el desarrollo y pruebas piloto del proyecto, Playit.gg se utiliza para exponer el backend y el frontend alojados en contenedores Docker hacia los estudiantes y notebooks de Colab. Esta estrategia permite realizar pilotos de manera controlada.

3. Estado del Arte

3.1. Introducción

La enseñanza de la estadística en educación superior presenta desafíos significativos debido a la abstracción de sus conceptos y la necesidad de aplicar teoría a problemas del mundo real. En este contexto, los Ambientes Virtuales de Aprendizaje (AVA) y plataformas digitales han emergido como recursos clave para mediar procesos formativos, permitiendo integrar materiales, actividades prácticas y evaluaciones (S. Al-Haddad, 2024). Sin embargo, aunque las soluciones existentes han avanzado en accesibilidad y organización de contenidos, persisten limitaciones relacionadas con la retroalimentación inmediata, la integración de entornos de análisis de datos y la evaluación automatizada de procesos complejos como la estadística inferencial.

3.2. Antecedentes

3.2.1. La evolución de los LMS

Los sistemas de enseñanza asistida han experimentado una evolución significativa desde sus orígenes en la década de 1960 hasta la consolidación de plataformas como Moodle en la actualidad. Uno de los primeros hitos fue el desarrollo de *PLATO* (Programmed Logic for Automated Teaching Operations) en la Universidad de Illinois en 1960. Este sistema permitía la entrega de contenidos programados, evaluaciones automatizadas, comunicación entre estudiantes y docentes mediante foros y mensajería, además de seguimiento del progreso individual. A pesar de los altos costos por el uso de mainframes, *PLATO* fue implementado en más de 1.000 instituciones educativas y sentó

las bases técnicas y pedagógicas de los actuales LMS (Woolley, 2016).

En 1965, Patrick Suppes y Richard Atkinson, en la Universidad de Stanford, exploraron el uso de computadoras para la enseñanza de matemáticas y lectura en escuelas de Palo Alto. Su sistema ofrecía retroalimentación inmediata y ramificaciones condicionales según el desempeño del estudiante. Los resultados mostraron una mejora del 25 % en la retención de conceptos matemáticos básicos en comparación con métodos tradicionales, lo que evidenció el potencial del aprendizaje asistido por computadora para personalizar la enseñanza (Suppes, 1966).

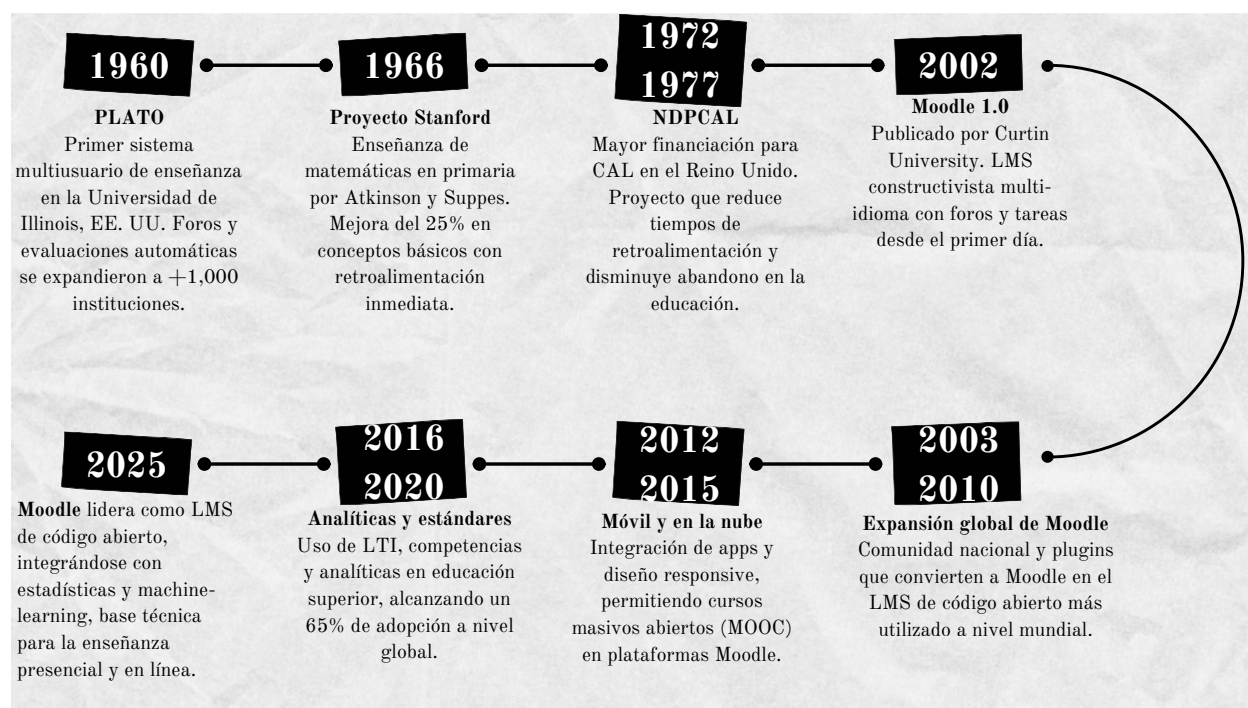
Posteriormente, entre 1973 y 1977, el *National Development Programme in Computer Aided Learning* (NDPCAL) en el Reino Unido financió más de 30 proyectos universitarios que integraban gestión de contenidos digitales, evaluación automatizada y seguimiento del progreso estudiantil. Uno de los proyectos más destacados fue implementado en la Universidad de Leeds, donde se logró reducir el tiempo de retroalimentación en cursos de estadística de semanas a horas, mejorando significativamente la experiencia de aprendizaje (Hooper, 1977).

El punto de inflexión en la historia de los LMS modernos llegó el 20 de agosto de 2002, cuando Martin Dougiamas publicó la primera versión estable de Moodle 1.0. Desarrollado como parte de su tesis doctoral en la Universidad de Curtin (Australia), Moodle fue concebido bajo principios constructivistas y liberado como software de código abierto. Esta primera versión permitía la creación de cursos interactivos, gestión de tareas, foros de discusión y evaluaciones básicas. A pesar de su simplicidad inicial, Moodle fue rápidamente adoptado por comunidades educativas de todo el mundo, traducido a múltiples idiomas y adaptado a diversos contextos pedagógicos (Moodle, 2024).

Hoy en día, Moodle supera los 160.000 sitios registrados y 240 millones de usuarios en más de 240 idiomas. Es utilizado por aproximadamente el 67 % de las instituciones de educación superior a nivel mundial. Su arquitectura abierta y modular permite integrar herramientas especializadas como Jupyter Notebook, nbgrader, R, Python y entornos de evaluación automática, lo que lo ha convertido en una plataforma clave para la enseñanza de la estadística y la ciencia de datos en la educación superior (Goh, 2025; Moodle, 2024).

Figura 1.

Estado del arte. **Linea de tiempo con base en diversas fuentes históricas..**



Fuente: Autores

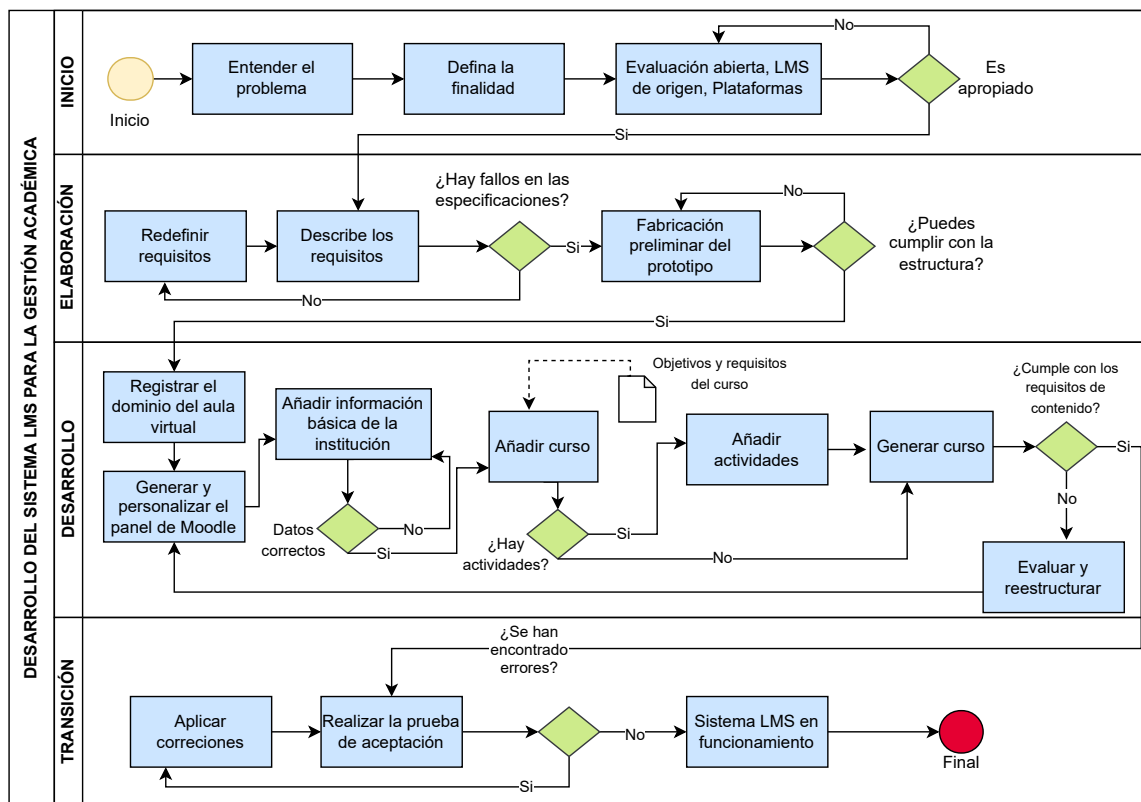
3.2.2. Plataformas de Aprendizaje

Las plataformas de gestión del aprendizaje (LMS) como Moodle han revolucionado la educación digital desde su implementación. El caso pionero de la Universidad de Curtin en Australia

en 2002 demostró que la adopción de Moodle en cursos de estadística descriptiva resultó en una mejora significativa en la participación estudiantil y una reducción del 30% en los índices de reprobación, atribuible a la organización modular de contenidos y el acceso asíncrono a materiales (Pacheco et al., 2025).

Figura 2.

Estado del arte. **Plataformas de aprendizaje.**



Fuente: Adaptada de Pacheco et al. (2025)

Más recientemente, en 2025, la Universidad Nacional de Educación a Distancia (UNED) de España desarrolló un MOOC de estadística en Moodle que alcanzó una tasa de finalización del 68 %, superior al promedio de cursos masivos. Este éxito se atribuyó al uso de evaluaciones

automáticas, foros moderados y contenido multimedia adaptativo (Goh, 2025). En contextos de menores recursos, la Universidad Pública de El Alto en Bolivia implementó Moodle para Estadística II en 2024, logrando mejoras significativas en la claridad de contenidos y el rendimiento académico de los estudiantes (Ndibalema, 2025).

Estos ejemplos demuestran que Moodle no solo cumple funciones administrativas, sino que puede ser una fuente valiosa de información para mejorar la enseñanza de la estadística cuando se complementa con herramientas externas de análisis y práctica computacional. Las analíticas de aprendizaje basadas en el uso del LMS se han convertido en un campo emergente, donde la persistencia y constancia de los estudiantes en el uso de la plataforma se correlaciona directamente con el éxito académico, abriendo la puerta a modelos predictivos de intervención temprana (Goh, 2025).

3.2.3. Sistemas de Autoevaluación y Retroalimentación Automatizada

La retroalimentación inmediata es crucial en la consolidación del aprendizaje estadístico. Mientras que los entornos tradicionales se centran en respuestas cerradas, las herramientas modernas han expandido su alcance hacia problemas complejos. El proyecto *nbgrader*, implementado en la Universidad de Berkeley en 2017, permitió evaluar automáticamente tareas de estadística en Jupyter Notebooks, validando tanto la exactitud numérica como la lógica del código en Python, resultando en una reducción del 40 % en el tiempo de calificación y un aumento en la precisión de las evaluaciones (Blank et al., 2017).

Sin embargo, investigaciones recientes identifican limitaciones persistentes. En 2024, la Universidad de Kuwait implementó un sistema de autoevaluación en estadística inferencial que,

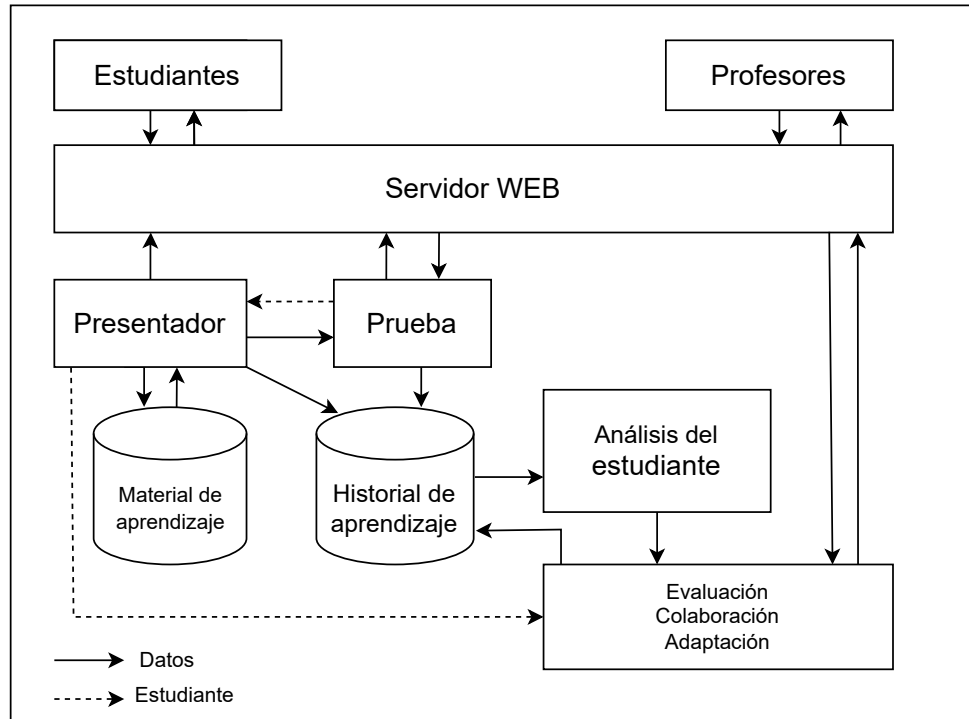
aunque efectivo para respuestas numéricas, no pudo evaluar el razonamiento estadístico ni la justificación de procedimientos, evidenciando la necesidad de graders más avanzados que trasciendan la simple comparación de valores (S. Al-Haddad, 2024).

Estas limitaciones subrayan la importancia de desarrollar sistemas de evaluación más sofisticados. Diversos estudios han sugerido que una alternativa futura es integrar graders más avanzados que almacenen y auditen código, permitiendo no solo verificar respuestas correctas o incorrectas, sino también analizar el proceso de resolución seguido por el estudiante, aumentando la transparencia y equidad en la evaluación.

3.2.4. Arquitectura tecnológica para la educación

En el ámbito de la enseñanza de la estadística, la discusión sobre arquitecturas tecnológicas ha pasado de enfocarse en plataformas de gestión administrativa hacia propuestas que buscan integrar entornos de análisis, retroalimentación inmediata y evaluación automatizada. Moodle continúa siendo una referencia obligada en la educación digital, especialmente por su capacidad para organizar contenidos y actividades. Sin embargo, varios estudios coinciden en que, aunque resulta eficaz para estructurar cursos, no responde por completo a las exigencias de la enseñanza de la estadística inferencial, en la que interesa no solo el resultado final, sino también el razonamiento seguido por el estudiante (Ndibalema, 2025; Pacheco et al., 2025).

Figura 3.

 Estado del arte. **Arquitectura tecnológica para la educación.**


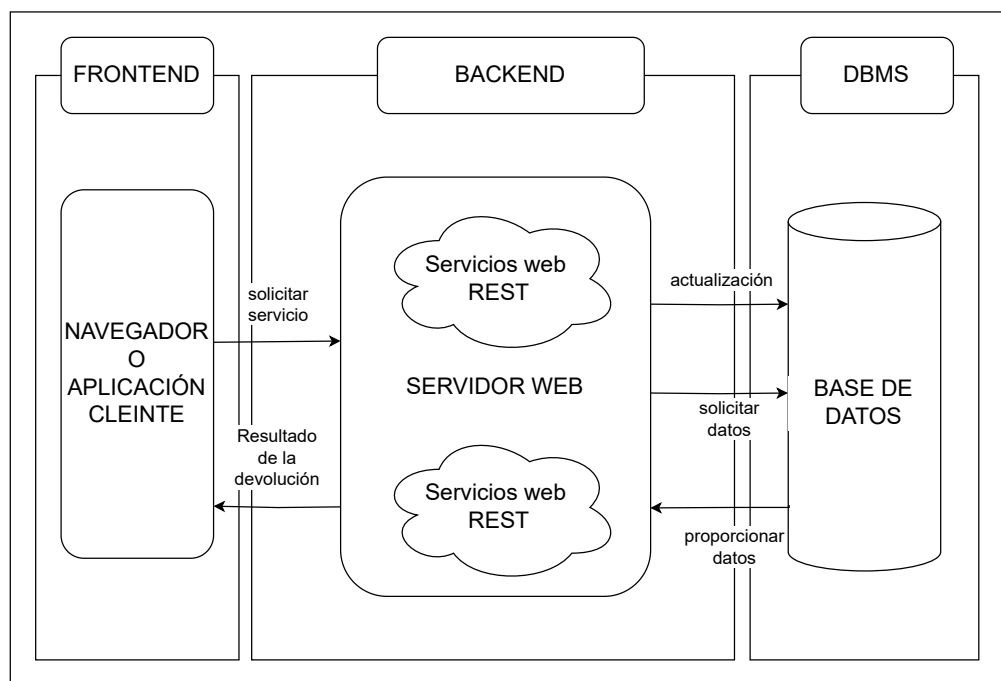
Fuente: Adaptada de Mostefai et al. (2025)

Ante esta limitación, algunas instituciones han apostado por combinar los LMS con entornos de programación y análisis de datos. Un ejemplo temprano es JupyterHub con nbgrader, desarrollado en la Universidad de California, Berkeley, que se consolidó como un referente en cursos masivos de ciencia de datos al permitir la entrega, ejecución y calificación automática de notebooks de Python (UC Berkeley, 2018). Por otro lado, la Universidad de Auckland ha incorporado RStudio Cloud para facilitar la práctica con R en cursos de inferencia y análisis de datos, priorizando la accesibilidad en la web (Rubio-Sánchez et al., 2023). Finalmente, en Carnegie Mellon University se ha implementado Autolab, pensado para la calificación de código en línea con control de intentos y reportes de error detallados (Rubio-Sánchez et al., 2023).

El diseño de entornos educativos digitales exige arquitecturas escalables, flexibles y seguras. En 2025, Mostefai propuso una arquitectura de microservicios para plataformas educativas que separa la lógica de negocio, la visualización de resultados y la gestión de datos en componentes independientes. Esta arquitectura permitió escalar cursos de estadística con más de 5.000 estudiantes simultáneos sin pérdida de rendimiento, mejorando significativamente la mantenibilidad y escalabilidad de los sistemas (Mostefai et al., 2025).

Figura 4.

Estado del arte. **Arquitectura de microservicios.**



Fuente: Adaptada de Mostefai et al. (2025)

Estas experiencias muestran un avance real hacia la automatización de la evaluación y la integración con el aprendizaje activo. Sin embargo, también dejan en evidencia limitaciones importantes. Por un lado, dependen de infraestructura institucional robusta o de licencias de pago, lo

que restringe su uso en contextos de menor presupuesto. Además, suelen enfocarse en verificar la exactitud del código o los resultados numéricos, pero no siempre logran capturar la trazabilidad del razonamiento estadístico, que es clave en la formación en inferencia (M. Al-Haddad, 2024).

3.2.5. Antecedentes nacionales y la UIS

En el contexto colombiano, la Universidad Industrial de Santander (UIS) ha sido pionera en el desarrollo de iniciativas que fortalecen la enseñanza y el aprendizaje de la estadística, articulando recursos digitales, programas de formación y unidades de análisis institucional. Desde mediados de la década de 2010, el Grupo SIMON de la UIS diseñó el software Evolución, una herramienta de modelado y simulación con dinámica de sistemas utilizada como apoyo en procesos formativos y de investigación, lo que refleja una apuesta temprana por el uso de simuladores en la enseñanza de disciplinas cuantitativas (Grupo SIMON – UIS, 2016).

De manera complementaria, la Dirección de Nuevas Tecnologías de la Información y las Comunicaciones (NTIC) de la UIS consolidó un repositorio de software y simuladores disponibles para estudiantes y docentes, los cuales abarcan desde áreas de ingeniería hasta aplicaciones estadísticas, contribuyendo al fortalecimiento del aprendizaje autónomo y práctico (NTIC – Universidad Industrial de Santander, s. f.). Esta apuesta tecnológica se ha acompañado de la creación de la Unidad de Información y Análisis Estadístico (UIAES), orientada a centralizar datos institucionales y a promover el uso de la estadística como herramienta de apoyo en la gestión académica y administrativa, lo que evidencia una integración de la estadística no solo en la docencia, sino también en la toma de decisiones estratégicas (Universidad Industrial de Santander, s. f.).

En el ámbito de la formación avanzada, la UIS ha ofrecido desde hace más de dos décadas

la Especialización en Estadística, adscrita a la Escuela de Matemáticas. Este programa de posgrado se ha consolidado como referente nacional, orientando profesionales en el manejo de metodologías estadísticas aplicadas en contextos empresariales, educativos y de investigación (Universidad Industrial de Santander – Escuela de Matemáticas, s. f.). Estas iniciativas evidencian el papel de la UIS como institución líder en la innovación y consolidación de propuestas académicas que integran recursos digitales, analítica institucional y programas de formación especializada.

En la Universidad Industrial de Santander (UIS), el Centro de Experiencias y Tecnologías de Información y Comunicación *EXPERTIC* se ha consolidado como un actor clave en la incorporación de TIC para fortalecer la docencia, la investigación y la extensión universitaria. Este centro promueve la innovación educativa mediante el diseño de recursos digitales, la gestión de plataformas tecnológicas y el acompañamiento a docentes y estudiantes en el uso pedagógico de las herramientas digitales (Universidad Industrial de Santander, 2025).

A nivel nacional, otras universidades también han contribuido a la consolidación de la estadística como disciplina clave para la educación superior y la investigación aplicada. Por ejemplo, la Universidad Católica de Pereira desarrolló un módulo didáctico para el manejo estadístico de datos en laboratorios de Física, destacando la transversalidad de la estadística en diferentes campos del conocimiento (Editorial Universidad Católica de Pereira, 2018). Por su parte, la Fundación Universitaria Los Libertadores ofrece la Especialización en Estadística Aplicada en modalidad virtual, ampliando el acceso a formación especializada y evidenciando el interés creciente por programas flexibles que integren la estadística con entornos digitales (Universidad de La Gran Colombia – Fundación Universitaria Los Libertadores, s. f.).

En este contexto, el presente trabajo de grado se inscribe como un aporte innovador al integrar en la asignatura de *Estadística II* de la UIS un ambiente virtual de aprendizaje que combina *Moodle* con entornos de programación (*Python* y *R*), *Google Colab*, evaluadores automáticos (*graders*) y arquitecturas de despliegue en *Docker*. La propuesta busca responder a las limitaciones señaladas en la literatura, especialmente la ausencia de retroalimentación inmediata. Para ello, se plantea un entorno que incorpora actividades interactivas, mecanismos de autoevaluación y un plan de aula estructurado, implementado en dos cursos. De esta manera, el proyecto no solo fortalece la enseñanza de la estadística inferencial en pregrado, sino que se alinea con las tendencias internacionales de integración entre LMS.

4. Metodología

4.1. Enfoque metodológico

Para el desarrollo del presente proyecto se adoptó la metodología ágil Scrum, debido a que permite organizar el trabajo en iteraciones cortas (sprints), promover la retroalimentación constante y garantizar la entrega continua de productos funcionales. Este enfoque se ajusta a la naturaleza del proyecto, en el cual se requiere diseñar, implementar y validar materiales interactivos y evaluativos para la asignatura Estadística II en un ambiente de aprendizaje

4.2. Roles

El equipo esta conformado por tres integrantes: un docente y dos estudiantes.

Scrum Master (Docente)

- Definir los temas a desarrollar en cada iteración.
- Validar los entregables (material teórico, graders y servidor de calificaciones).
- Proporcionar retroalimentación y ajustes necesarios.

Desarrolladores (2 estudiantes)

Encargados de distribuirse las tareas correspondientes a cada sprint, las cuales incluyen:

- Elaborar el material teórico, diseñado para integrarse en Moodle y con un enfoque interactivo.

- Diseñar e implementar los graders (evaluadores automáticos) para los ejercicios prácticos.
- Programar la lógica del servidor que permite calificar y almacenar los resultados.
- Ejecutar las correcciones solicitadas por el Scrum Master.

4.3. Estructura de los sprints

El proyecto se organiza en **sprints semanales**, cada uno correspondiente al diseño y validación de una práctica de la asignatura. En total se llevaron a cabo ocho sprints, alineados con las ocho prácticas definidas en el curso de Estadística II. La dinámica de trabajo semanal fue la siguiente:

- **Sábado:** el Scrum Master define los temas correspondientes a la semana.
- **Miércoles (fecha límite):** los estudiantes deben entregar el material teórico, los graders y la versión funcional del servidor de calificaciones.
- **Jueves:** el Scrum Master revisa los entregables y da su aprobación o indica las modificaciones requeridas.
- **Viernes:** se liberan los graders para que los estudiantes (grupo piloto) realicen las prácticas.

4.4. Sprints

4.4.1. *Sprint 0*

Este sprint se definió a partir de la reunión llevada a cabo el 2 de abril de 2025 en la sala de juntas de la Escuela de Ingeniería de Sistemas de la UIS, con la participación de docentes del área

de Estadística II y los dos estudiantes responsables del proyecto (Suárez & Sánchez, 2025).

En dicha reunión se establecieron los lineamientos iniciales del proyecto y se delimitó su alcance, acordando que el enfoque debía orientarse a mejorar los procesos de enseñanza y aprendizaje mediante el uso de herramientas digitales.

Durante las discusiones se resaltó la importancia de analizar las características de los principales lenguajes de programación utilizados en el ámbito estadístico (Python, MATLAB y R). Con este propósito se propuso elaborar un cuadro comparativo que permitiera identificar sus ventajas, desventajas y entornos de desarrollo asociados. Aunque no se seleccionó un lenguaje específico en esta etapa, se dejó abierta la posibilidad de trabajar con más de uno, con el objetivo de garantizar flexibilidad.

Asimismo, se planteó que la herramienta a desarrollar debía ser agnóstica respecto al lenguaje de programación, de modo que el usuario pudiera elegir libremente el entorno de su preferencia. También se discutió la conveniencia de incorporar interfaces visuales y elementos gráficos que fortalecieran la enseñanza para los estudiantes, junto con la organización de los contenidos pedagógicos dentro de Moodle como plataforma de apoyo.

De manera complementaria, se estableció una distribución preliminar de temas en tres unidades (1.3 y 1.4; 2.3 y 2.4; 3.3 y 3.4), con el fin de estructurar las prácticas, realizar pruebas efectivas y evaluar el funcionamiento del proyecto.

Finalmente, en este sprint se tomaron decisiones relacionadas con las tecnologías base que guiarían el desarrollo:

- **Node.js**: para el backend, debido a su rapidez en las respuestas y su estructura escalable.
- **Moodle**: como plataforma para la gestión del contenido teórico.
- **Google Colab**: como entorno de ejecución de los *graders*, aprovechando su accesibilidad y la posibilidad de alternar entre los *kernels* de Python y R.
- **Docker**: para asegurar un despliegue estable y controlado mediante contenedores.
- **Playit.gg**: como servicio de túnel que permite exponer el servidor de calificaciones en internet.

4.4.2. *Sprint 1*

Objetivos del Scrum Master (docente).

- Definir los temas a trabajar durante el sprint: medidas de tendencia central (media, mediana, moda), medidas de dispersión (rango, desviación estándar, varianza), histogramas y diagramas de caja (boxplots).
- Validar la pertinencia y calidad de los contenidos teóricos elaborados por los estudiantes.
- Revisar el funcionamiento de los *graders* y su integración con el servidor de calificaciones.
- Orientar la definición de la arquitectura técnica, verificando que cumpliera con los lineamientos de escalabilidad y modularidad.

Cronograma del Sprint.

- **Inicio del Sprint:** 9 de agosto de 2025 → El Scrum Master definió los temas a desarrollar.
- **Entrega de los estudiantes:** 13 de agosto de 2025 → Se entregaron el material teórico en H5P, los graders y la estructura técnica inicial.
- **Revisión del Scrum Master:** 14 de agosto de 2025 → Se realizaron recomendaciones de mejora sobre los contenidos y ejercicios.
- **Liberación de recursos:** 15 de agosto de 2025 → Los graders y materiales teóricos fueron liberados en Moodle y Colab para ser utilizados por los grupos pilotos.

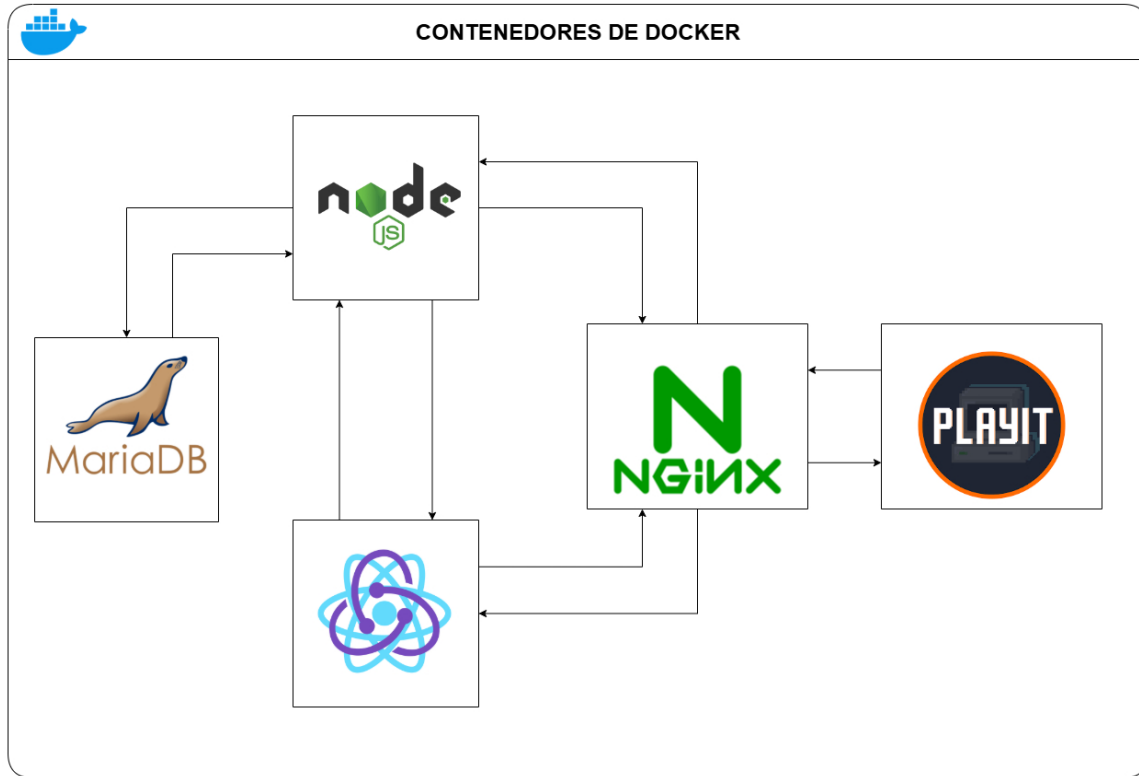
Actividades realizadas por los estudiantes.

1. Estructura técnica del proyecto

- Se definió el esqueleto del sistema, estableciendo la organización de los módulos (frontend, backend y grader).
- Se configuró la exposición de servicios y la comunicación entre módulos.
- Se organizó el sistema en contenedores con Docker, garantizando un entorno estable, replicable y modular.

Figura 5.

Sprint 1. Esquema de los contenedores Docker.



Elaboración propia con base en la estructura de los contenedores usados.

2. Material teórico en Moodle

- Se elaboraron recursos interactivos utilizando H5P dentro de Moodle.
- Los materiales incluyeron infografías sintetizadas, con la información esencial para el aprendizaje de los temas, permitiendo una mayor comprensión visual.
- La construcción de los contenidos se fundamentó en los textos de referencia de Montgomery y Walpole, autores reconocidos en estadística.

3. Graders en Google Colab

- Se desarrollaron notebooks en Google Colab con código en R.
- Estos graders permitieron calcular medidas de tendencia central y dispersión, así como generar representaciones gráficas (histogramas y diagramas de caja).

4. Ejercicios calificables y lógica en el backend

- Se diseñaron tres ejercicios evaluables relacionados con los temas del sprint.
- Se implementó la lógica de corrección en el backend (Node.js) para procesar automáticamente las respuestas.
- El servidor de calificaciones registró y almacenó los resultados, garantizando un flujo de evaluación automatizado.

Bibliografía

- Al-Haddad, M. (2024). Automated assessment systems in higher education: Current challenges and future directions. *Journal of Educational Technology Research*, 17(2), 55-72.
- Al-Haddad, S. (2024). Teaching statistics: A technology-enhanced supportive learning environment. *Teaching Statistics*, 46(1), 1-12. <https://doi.org/10.1080/26939169.2024.2315939>
- Banks, A., & Porcello, E. (2017). *Learning React: A Hands-On Guide to Building Web Applications Using React and Redux*. O'Reilly Media.
- Bisong, E. (2019). *Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners*. Apress.
- Blank, D., Bourgin, D., Brown, A., Bussonnier, M., Frederic, J., Granger, B., & Willing, C. (2017). nbgrader: A tool for creating and grading assignments in the Jupyter Notebook. *Journal of Open Source Education*, 2(11), 32. <https://doi.org/10.21105/jose.00032>
- Brown, E. (2019). *Web Development with Node and Express: Leveraging the JavaScript Stack*. O'Reilly Media.
- Castro, M. C. (2019). Ambientes de aprendizaje. *Sophia*, 15(2), 123-134. <https://revistas.ugca.edu.co/index.php/sophia/article/view/827/1473>
- Codd, E. F. (1970). A relational model of data for large shared data banks. *Communications of the ACM*, 13(6), 377-387. <https://doi.org/10.1145/362384.362685>
- Dougiamas, M., & Taylor, P. (2003). Moodle: Using learning communities to create an open source course management system. *World Conference on Educational Multimedia, Hypermedia and Telecommunications*, 2003(1), 171-178.

- Editorial Universidad Católica de Pereira. (2018). *Módulo didáctico sobre el manejo estadístico de datos para estudiantes de laboratorio de Física I*. Consultado el 17 de septiembre de 2025, desde <https://editorial.ucp.edu.co/omp/index.php/e-books/catalog/download/50/46/1818?inline=1>
- Goh, T.-T. (2025). Learning management system log analytics: The role of persistence and consistency of engagement behaviour on academic success. *Journal of Computers in Education*, 12(1), 45-62. <https://doi.org/10.1007/s40692-025-00358-x>
- Grupo SIMON – UIS. (2016). *Evolución: herramienta software para modelado y simulación con Dinámica de Sistemas*. Consultado el 17 de septiembre de 2025, desde <https://www.researchgate.net/publication/305503842>
- Hooper, R. (1977). *National Development Programme in Computer Aided Learning: The final report of the Director*. Council for Educational Technology.
- Kurnia, A., Lim, A., & Cheang, B. (2001). Online judge. *Computers & Education*, 36(4), 299-315. [https://doi.org/10.1016/S0360-1315\(01\)00013-4](https://doi.org/10.1016/S0360-1315(01)00013-4)
- McKinney, W. (2017). *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*. O'Reilly Media.
- Merkel, D. (2014). Docker: lightweight Linux containers for consistent development and deployment. *Linux Journal*, 2014(239), 2.
- Montgomery, D. C., & Hines, W. W. (1996). *Probabilidad y estadística para ingeniería y administración*. Limusa.
- Moodle. (2024). La historia de Moodle [Consultado el 11 de septiembre de 2025]. <https://moodle.com/es/acerca-de/la-historia-de-moodle/>

- Mostefai, B., Boutefara, T., Bousbia, N., Balla, A., Dhelim, S., & Hammia, A. (2025). Enhancing User Experience in E-Learning Systems: A new User-Centric RESTful Web Services approach. *Computers in Human Behavior Reports*, 100643. <https://doi.org/10.1016/j.chbr.2025.100643>
- Ndibalema, P. (2025). Perspectives on the use of learning management systems in higher learning institutions in Tanzania: The gaps and opportunities. *Social Sciences & Humanities Open*, 11, 101463. <https://doi.org/10.1016/j.ssaho.2025.101463>
- NTIC – Universidad Industrial de Santander. (s. f.). *Software y simuladores*. Consultado el 17 de septiembre de 2025, desde <https://ntic.uis.edu.co/software-simuladores/>
- Pacheco, A., Yupanqui, R., Mogrovejo, D., Garay, J., & Uribe-Hernández, Y. (2025). Impact of digitization on educational management: Results of the introduction of a learning management system in a traditional school context. *Computers in Human Behavior Reports*, 17, 100592. <https://doi.org/10.1016/j.chbr.2025.100592>
- R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.r-project.org/>
- Rubio-Sánchez, C., Soto, M., & Pérez, J. (2023). Automated assessment system for programming courses: A case study for teaching data structures and algorithms. *Educational Technology Research and Development*. <https://doi.org/10.1007/s11423-023-10277-2>
- Salinas, J. (2004). Innovación docente y uso de las TIC en la enseñanza universitaria. *Revista de Universidad y Sociedad del Conocimiento*, 1(1), 1-16. <https://rusc.uoc.edu/rusc/ca/index.php/rusc/article/view/v1n1-salinas.html>
- Suárez, J. E., & Sánchez, D. A. (2025, abril). *Acta de reunión del 2 de abril de 2025* [Acta interna, Escuela de Ingeniería de Sistemas, Universidad Industrial de Santander. <https://docs.>

[google.com/document/d/1smyDmnIUfinXADn5_7-11BeBVsy75D3pmMV-Ch6wmZs/edit?usp=sharing](https://www.google.com/document/d/1smyDmnIUfinXADn5_7-11BeBVsy75D3pmMV-Ch6wmZs/edit?usp=sharing)].

Suppes, P. (1966). The uses of computers in education. *Scientific American*, 215(3), 206-220. <https://doi.org/10.1038/scientificamerican0966-206>

Tilkov, S., & Vinoski, S. (2010). Node.js: Using JavaScript to build high-performance network programs. *IEEE Internet Computing*, 14(6), 80-83. <https://doi.org/10.1109/MIC.2010.145>

UC Berkeley. (2018). *The JupyterHub journey: Starting small and scaling up* [Berkeley Data Science Education Program].

Universidad de La Gran Colombia – Fundación Universitaria Los Libertadores. (s. f.). *Especialización en Estadística Aplicada – Virtual*. Consultado el 17 de septiembre de 2025, desde <https://www.ulibertadores.edu.co/educacion-virtual-distancia/especializacion-estadistica-aplicada-virtual/>

Universidad Industrial de Santander. (2009). Política Institucional de Tecnologías de la Información y la Comunicación [Acuerdo del Consejo Superior No. 051 de 2009. Recuperado de <https://uis.edu.co/wp-content/uploads/2022/06/2009PoliticaTICAcCS051.pdf>]. Consultado el 1 de octubre de 2023, desde <https://uis.edu.co/wp-content/uploads/2022/06/2009PoliticaTICAcCS051.pdf>

Universidad Industrial de Santander. (2025). *EXPERTIC: Centro de Experiencias y Tecnologías de Información y Comunicación*. Consultado el 17 de septiembre de 2025, desde <https://expertic.uis.edu.co/home>

Universidad Industrial de Santander. (s. f.). *Unidad de Información y Análisis Estadístico – UIAES*. Consultado el 17 de septiembre de 2025, desde <https://uis.edu.co/unidad-de-informacion-y-analisis-estadistico-uaes/>

Universidad Industrial de Santander – Escuela de Matemáticas. (s. f.). *Especialización en Estadística*. Consultado el 17 de septiembre de 2025, desde <https://posgrados.uis.edu.co/inicio/especializacion-en-estadistica/>

Woolley, D. R. (2016). PLATO: The emergence of online community. En *The MIT Press eBooks*. MIT Press. <https://doi.org/10.7551/mitpress/9780262034654.003.0005>