

## Programa Estining de R-Studio para estudiantes universitarios de Ingeniería

Luisa Graciela Urure Tejada<sup>1\*</sup>, Matilde Peña Casa<sup>1</sup>, Danny Dominguez Pillaca<sup>1</sup>

<sup>1</sup> Escuela de Posgrado. Universidad César Vallejo. Perú.

\* Autor para correspondencia: Luisa Graciela Urure Tejada, luisagracie@gmail.com

(Recibido: 24-12-2023. Publicado: 31-12-2023.)

DOI: 10.59427/rcli/2023/v23cs.3931-3939

### Resumen

*El manuscrito (artículo) se centra en el software R-Studio para la enseñanza de la estadística de los estudiantes universitarios de Arequipa, destacando su importancia en la educación superior. Se discute cómo los estudiantes universitarios comienzan a desarrollar habilidades fundamentales en la competencia analítica de la estadística, el software R-Studio ha transformado la enseñanza de la estadística, ya que, al ser un software libre resulta ser novedoso y de mucha ventaja para los estudiantes de nivel superior, sobre todo en la ciencia de la ingeniería, ya que permitirá que aprendan estadística inferencial en forma más práctica, sencilla y resolver problemas. El artículo subraya la importancia del desarrollo de habilidades para el éxito del análisis estadístico. Por otra parte, muchos estudiantes tienen dificultades para comprender los conceptos estadísticos complejos que se brindan en clase también en la estadística inferencial que es parte de la estadística debido a su abstracción ello, demanda un nivel elevado de comprensión y habilidades de razonamiento lógico. El enfoque de la investigación fue cuantitativo de tipo aplicada el diseño cuasiexperimental y se ubica en el nivel explicativo causal. Se realizó un estudio de investigación científica y revisión de forma sistemática descriptiva de todo tipo de artículos de revistas que se encuentran indexadas a partir del año 2018 hasta el presente año 2023, en las siguientes plataformas de búsqueda: Scopus, SciElo, Redalyc, Latindex, ProQuest, etc. para ello, se utilizó los descriptores: "Programa R" "análisis inferencial" "estadística inferencial", "Pensamiento complejo", "transformación de la enseñanza para el desarrollo de habilidades estadísticas", "Desarrollo del pensamiento crítico reflexivo", "Desarrollo del pensamiento matemático" combinándolos entre ellos con los booleanos AND y OR. Como conclusión se determinó que, la importancia la experiencia exitosa en la instalación, elección y promoción del software, ha resaltado la versatilidad, y el potencial de esta herramienta digital de programaciones en disciplinas vinculadas con la Ingeniería. Esto demuestra que se puede aplicar en otras áreas de la carrera, aprovechando paquetes disponibles en la web de forma gratuita.*

**Palabras claves:** Programa R, análisis inferencial, estadística inferencial, Pensamiento complejo, transformación de la enseñanza para el desarrollo de habilidades estadísticas, Desarrollo del pensamiento crítico reflexivo, Desarrollo del pensamiento matemático.

### Abstract

*The manuscript (article) focuses on the R-Studio software for teaching statistics to university students in Arequipa, highlighting its importance in higher education. It is discussed how university students begin to develop fundamental skills in the analytical competence of statistics. The R-Studio software has transformed the teaching of statistics, since, being free software, it turns out to be novel and very advantageous for students. higher level, especially in engineering science, since it will allow them to learn inferential statistics in a more practical and simple way and solve problems. The article highlights the importance of skill development for the success of statistical analysis. On the other hand, many students have difficulties understanding the complex statistical concepts that are provided in class, also in inferential statistics, which is part of statistics, due to its abstraction, which demands a high level of understanding and logical reasoning skills. The research approach was quantitative, quasi-experimental design applied and is located at the causal explanatory level. A scientific research study and systematic descriptive review of all types of journal articles that are indexed from 2018 to the current year 2023 was carried out on the following search platforms: Scopus, SciElo, Redalyc, Latindex, ProQuest, etc. For this purpose, the descriptors were used: "R Program," "inferential analysis," "inferential statistics," "Complex thinking," "transformation of teaching for*

*the development of statistical skills,” “Development of critical reflective thinking,” “Development of mathematical thinking” by combining them with each other with the booleans AND and OR. In conclusion, it was determined that the importance of the successful experience in the selection, installation and promotion of the software, together with the preparation of classes focused on statistics, has highlighted the potential and versatility of this programming tool in the disciplines related to Engineering. This shows that we can apply it in any area of the career, taking advantage of packages available on the web for free.*

**Keywords:** *R program, inferential analysis, inferential statistics, Complex thinking, transformation of teaching for the development of statistical skills, Development of reflective critical thinking, Development of mathematical thinking.*

## 1. Introducción

En la actualidad, la enseñanza de la estadística ha experimentado una transformación significativa gracias al software R-Studio, que, al ser de código abierto, resulta innovador para nuestros estudiantes de nivel superior, especialmente en el ámbito de las Ingenierías. Este enfoque permite a los estudiantes aprender estadística inferencial de manera más práctica y accesible. Dado que muchos alumnos enfrentan dificultades para comprender los conceptos estadísticos complejos impartidos en clase, la estadística inferencial, siendo una parte abstracta de la estadística, demanda un elevado nivel de comprensión y habilidades de razonamiento lógico (Rubio, 2021). Inicialmente, los estudiantes pueden carecer de los recursos necesarios para familiarizarse con la estadística inferencial utilizando R-Studio, ya que algunos docentes aún no han recibido la capacitación correspondiente. Esta falta de habilidad puede afectar la eficacia de la enseñanza, haciendo que el lenguaje de programación utilizado, quizás, no resulte tan atractivo para los estudiantes. Así mismo Marzouk, (2021) refiere tanto los estudiantes como los docentes que participan en el curso de estadística inferencial, utilizando el Programa Estining de R Studio, se encuentran inmersos en cambios que abarcan desde la comprensión de conceptos hasta la adquisición de habilidades técnicas, la interpretación de datos y la toma de decisiones basadas en los resultados. El objetivo final es lograr resultados óptimos y satisfactorios tanto para el estudiante como para el docente (Abdollahi et al., 2022). Estudios recientes resaltan la importancia de capacitar a profesionales en ingeniería informática y establecer conexiones efectivas entre la universidad y el sector empresarial (González et al., 2020). La estadística desempeña un papel crucial en esta formación al contribuir con la inferencia de resultados y la toma de decisiones óptimas enfocadas en el dominio de ingentes cantidades de base de datos. A pesar de ello, algunos docentes aún se oponen a adoptar esta nueva metodología de enseñanza de estadística, persistiendo en la preferencia por fórmulas y métodos tradicionales en lugar de promover actitudes y aprendizaje integrado que preparen integralmente a los estudiantes para su futuro profesional.

Por lo tanto, la educación superior debe poner un énfasis particular en la aplicación práctica de la estadística en la sociedad, ya que desempeña un papel crucial para garantizar una formación integral en el desarrollo de la autonomía y toma de decisiones para los profesionales. La relevancia de emplear software como herramienta pedagógica en la enseñanza de la estadística en ingeniería, así como los desafíos que enfrentan los estudiantes al aprender estadística, resaltan la necesidad de que los profesores estén debidamente preparados y competentes en el manejo de software (Arriaza et al., 2018). En las últimas décadas, la obtención y transmisión de información han experimentado una transformación significativa, donde las Tecnologías de la Información y la Comunicación (TIC) han jugado un papel fundamental en el ámbito educativo al ofrecer mayor accesibilidad y flexibilidad en los métodos de impartir conocimiento e instrucción (Fernández, 2020). Para el estudio en el Zheng, (2021) señalaron que, es fundamental destacar que las TIC han simplificado la creación y distribución de recursos educativos, fortaleciendo la relación entre estudiantes y profesores y ampliando las oportunidades de adquirir conocimientos más allá de las limitaciones físicas de las aulas tradicionales. Estas tecnologías han impulsado la transición hacia nuevas formas de enseñanza y aprendizaje, fomentando enfoques más interactivos y participativos (Mares y Domínguez, 2022). Por esta razón, Carrera y Villacrés et al., (2019) subrayan la importancia de integrar las TIC en los procesos educativos para aprovechar al máximo sus beneficios y lograr los mejores resultados en el aprendizaje. Los avances de la tecnología a generado la aparición de enfoques educativos diversos y estrategias de enseñanza innovadoras. Modelos pedagógicos como el aprendizaje colaborativo y el aprendizaje en línea han surgido como alternativas que transforman la instrucción educativa.

En este contexto, es crucial que las instituciones educativas se comprometan a adaptarse a las nuevas tecnologías y recursos disponibles, ajustando sus prácticas y modelos educativos. Esta adaptación se vuelve aún más imperativa debido al crecimiento en la demanda de profesionales con habilidades estadísticas para la toma de decisiones a nivel empresarial, según (Jahuey et al., 2022). Es así que, se formula el siguiente problema general ¿Cuáles son los efectos del programa Estining de R-Studio en el aprendizaje de estadística Inferencial en estudiantes universitarios de Ingeniería? Desde una perspectiva teórica, la justificación se basa en el enfoque y metodología utilizado para abordar los fundamentos teóricos del aprendizaje de estadística inferencial en estudiantes universitarios. A través del programa Estining de R-Studio, se busca potenciar el entendimiento y la transmisión de

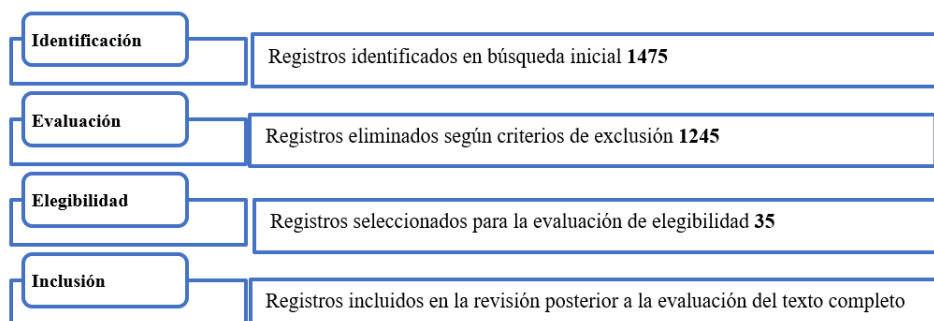
los beneficios que esta herramienta aporta al proceso de enseñanza al mismo tiempo el aprendizaje de los estudiantes de ingeniería. Por su parte Tapia, (2022) en términos metodológicos, la investigación se fundamenta en la definición de los pasos y técnicas necesarios para llevar a cabo el desarrollo del curso de estadística inferencial. Se propone una secuencia de acciones que permitirá la identificación de un problema, su planteamiento y la búsqueda de soluciones utilizando el programa Estining de R-Studio. Además, se abordó la gestión eficiente de herramientas y tiempo, así como la generación de nuevo conocimiento, con el fin último de alcanzar una solución satisfactoria (Adolfsson et al., 2019). Este estudio permitió promover el trabajo colaborativo, exploración de conocimientos previos y la aplicación de destrezas como el análisis, entendimiento y la solución de problemas utilizando el programa Estining de R-Studio. Esto les permite adquirir actitudes en su desarrollo, donde podrán aplicar su ingenio para resolver problemas de su entorno.

## 2. Metodología

Con respecto a la metodología esta investigación se realizó en el marco del método científico por lo que se siguió un proceso basado en la lógica deductiva. Por ende, García et al., (2020) resalta que, el estudio en particular se adhiere al enfoque hipotético-deductivo, en el cual se plantean hipótesis y se someten a verificación mediante inferencias. En relación al enfoque del estudio, se empleó un enfoque cuantitativo. El propósito principal de este enfoque se basa en la descripción, predicción, explicación en consecuencia el control del objetivo de la causa del fenómeno, así como en la predicción de su ocurrencia a partir de la comprensión de dichas causas, estudio que llevó a cabo utilizando un diseño de investigación cuasi experimental (Roa, 2015). Por su parte Fernández y Escobar, (2021) resaltan que, el estudio se centró principalmente en este diseño mediante la aplicación de un experimento con dos grupos: por un lado, el grupo de control y el otro grupo experimental que permitió aplicar el programa para demostrar los resultados de la investigación. Se recopilaron datos antes y después del experimento. no cuenta con la asignación aleatoria de los participantes a los grupos, lo que implica ciertas limitaciones en términos de control de variables.

Es importante destacar que, el aprendizaje en la estadística inferencial es el proceso de adquirir conocimientos y habilidades relacionados con la aplicación de técnicas y métodos estadísticos para realizar inferencias y tomar decisiones sobre una población en base a una muestra de datos de acuerdo con Gallardo, (2017), el proceso de operacionalización se manifestó en un enfoque cuantitativo mediante la observación y medición de variables. En la investigación se hizo una revisión metodológica y sistemática de diversos artículos que se ubicaron en revistas publicadas e indexadas, de los últimos cinco (5) años a partir de 2018 hasta 2023, utilizando plataformas de búsqueda como: Scopus, SciElo, Latindex, Redalyc y ProQuest, utilizando para esto los siguientes descriptores: “Programa R” “análisis inferencial” “estadística inferencial”, “Pensamiento complejo”, “transformación de la enseñanza para el desarrollo de habilidades estadísticas”, “Desarrollo del pensamiento crítico reflexivo”, “Desarrollo del pensamiento matemático”.

Por otra parte, fue necesario aplicar el método PRISMA (Preferred Reporting Items for Systematic Reviews and Meta Analyses), en la que, se escogieron las investigaciones que contenían información sobre el tema tratado, utilizando para esto los criterios de identificación, evaluación, elegibilidad e inclusión (Fernández Bedoya, 2020). Así mismo Fernández Lizana, (2020) confirma en este sentido, como criterios de inclusión, se consideró el título, resumen, metodología, resultados y conclusiones; para ello se tomó criterios de exclusión en la que se consideró no formar parte de artículos científicos (reseñas, ponencias, tesis, publicaciones de revistas o prensa, entre otros), artículos duplicados, así como también no estar vinculado al título o contexto investigado. Durante el proceso, se pudo obtener 1475 artículos de todas las bases de datos que fueron consultadas, descartándose 1245 por estar dentro de los criterios de exclusión, separándose 35 para su análisis, quedando 20 artículos (Figura 1), las cuales fueron seguidamente analizados en el presente estudio (Fernández, 2020).



**Figura 1:** Diagrama de flujo PRISMA.

### 3. Resultados y discusión

Esta sección introduce y examina las investigaciones como componente integral de los procesos de revisión que se ha llevado a cabo. En este sentido, se muestra en la tabla 1 una lista de los estudios de investigación seleccionados de acuerdo con los criterios establecidos para la inclusión en esta revisión. Por su parte Rodríguez, (2019) manifiesta que, después del empoderamiento y dominio del software y elaborar los recursos de soporte, se solicitó a la universidad considerar el rubro de Probabilidad y Estadísticas la creación de ejercicios prácticos para que los estudiantes los resuelvan utilizando la suite R. Se pidió a los profesores que proporcionaran soluciones precisas, acompañadas de comentarios exhaustivos y evaluaciones. Esto se hizo con el propósito de comparar las ejecuciones del software con las soluciones humanas, permitiendo así determinar el nivel de retroalimentación que se espera de los alumnos.

**Tabla 1:** Artículos seleccionados para la revisión.

Autor(es) / Año/ País	Título	Conclusiones
1 Rodríguez 2019 México	¿Qué puede hacer el software R para resolver tus problemas?	R representa un lenguaje y marco de programación con notables repercusiones en los ámbitos científico, numérico y estadístico, destacándose por sus vastas capacidades. Aunque implica una inversión de esfuerzo aprender este idioma, un fenómeno común al abordar cualquier plataforma de características semejantes, no cabe duda de que se aconseja encarecidamente su aplicación.
2 Escuder (2019) Uruguay	Regionalización de la brecha digital. Desarrollo de la infraestructura de las TIC en Latinoamérica y Uruguay.	La desigualdad digital engendra un ciclo pernicioso en la dinámica de insatisfacción social. Aquellas comunidades más marginadas digitalmente se superponen con la exclusión espacial y socioeconómica de los hogares en las áreas periféricas.
3 Abío et al., (2017) España	El aula invertida y el aprendizaje en equipo: dos metodologías para estimular al estudiante repetidor.	Los resultados revelan la coherencia con la que los estudiantes han adoptado las nuevas metodologías. A pesar de que, en general, no aprecian la necesidad de trabajar de forma individual para preparar las clases, expresan una valoración muy positiva hacia el trabajo en equipo.
4 Fernández (2020) Paraguay	Ventajas de R como herramienta para el Análisis y Visualización de datos en Ciencias Sociales	En cuanto a la pregunta sobre si es necesario ser un "programador," tener conocimientos y experiencia previa en programación para trabajar con R, la respuesta es negativa. No se requiere ser programador ni tener experiencia previa en programación para trabajar con R. Sin duda, un programador experto o alguien con experiencia en programación puede aprender a trabajar con R de manera más rápida, ya que algunos lenguajes de programación comparten ciertas características. No obstante, aquellas personas sin conocimientos previos también pueden aprender a utilizar R de manera efectiva.
5 Morochó et al., (2022) Ecuador	Herramientas tecnológicas para desarrollar la lectura	Las herramientas tecnológicas desempeñan un papel fundamental en la comprensión de las necesidades actuales tanto de docentes como de estudiantes, simplificando el trabajo y garantizando una comunicación constante entre los participantes. Este enfoque dinamiza el proceso de enseñanza y aprendizaje, al mismo tiempo que potencia el desarrollo de habilidades de comprensión lectora, contribuyendo así a proporcionar una educación de alta calidad.
6 Díaz (2021). Perú	Plataforma virtual Moodle para la aplicación del modelo educativo bilingüe-bicultural en la enseñanza de la lengua de señas peruana.	La introducción de un enfoque educativo bilingüe y bicultural mediante la utilización de la plataforma virtual Moodle para la enseñanza de la Lengua de Señas Peruana ha generado un impacto significativo en el proceso de aprendizaje. Las evaluaciones realizadas posteriormente a la aplicación de la prueba muestran mejoras positivas evidentes en el rendimiento académico.
7 Antón (2018) España	Las TIC en la educación bilingüe. Estudio de caso en la enseñanza pública y concertada en un entorno rural	El rol del docente en esta integración es crucial, donde las habilidades informáticas del profesorado constituyen un elemento clave en el proceso de enseñanza-aprendizaje. Por lo tanto, resulta imperativo explorar y profundizar en el marco de competencias TIC de los nuevos docentes para capitalizar los beneficios que pueden derivarse del uso de la tecnología en el aula.
8 Fernández (2020) Perú	Tipos de justificación en la investigación científica	El rol del docente en esta integración es crucial, donde las habilidades informáticas del profesorado constituyen un elemento clave en el proceso de enseñanza-aprendizaje. Por lo tanto, resulta imperativo explorar y profundizar en el marco de competencias TIC de los nuevos docentes para capitalizar los beneficios que pueden derivarse del uso de la tecnología en el aula.
9 Carrera (2020) Ecuador	Manual de r aplicado a la ingeniería del agua	Los estudiantes pueden aprovechar el programa R, ya que se trata de un lenguaje y entorno de programación científica, numérica y estadística con un impacto significativo y vastas potencialidades. Su utilización es ampliamente recomendada.
10 Martínez (2022) México	El programa R: una estrategia inicial para su entendimiento y aprendizaje	Aprender el lenguaje R es equiparable a adquirir una nueva habilidad lingüística. Inicialmente, es probable que surjan errores de código que generen frustración en la mayoría de los usuarios. No obstante, a medida que se practica este nuevo lenguaje, se mejora el rendimiento, la confianza y la agilidad en su utilización. Como resultado, se desarrollan nuevas habilidades para el análisis de datos. Una vez que se alcanza un dominio básico del programa, el usuario de R será capaz de instalar el software y sus diversas bibliotecas, identificar y emplear funciones básicas, construir secuencias lógicas de procedimientos para su ejecución.
11 Mares (2022) México	Evaluación estadística de índices de desempeño para el proceso de división de roles de EVA	El supuesto de normalidad en el cálculo de los PPIs (Índices de Precios al Productor) representa un desafío significativo para el analista, ya que en la realidad es frecuente encontrarse con datos que exhiben un comportamiento no normal. Estos datos son difíciles de transformar o ajustar a otras distribuciones, lo que añade complejidad al proceso de análisis.
12 Bright (2020) Canadá	CMIP6-D&A: un software basado en R con GUI para procesar datos climáticos disponibles en formato de datos común de red	Este software de código abierto está desarrollado en la plataforma R, que es un lenguaje de programación comúnmente utilizado en las comunidades científicas. Debido a su accesibilidad abierta, los usuarios con conocimientos básicos de programación pueden estar en condiciones de modificar el software según sus necesidades.
13 Brownstein (2019) Canadá	Estadísticas descriptivas y visualización de datos del paquete de conjuntos de datos R con implicaciones para la agrupabilidad	La razón subyacente para el uso de la estadística descriptiva es que proporciona un resumen poderoso de los datos y también permite comparaciones sencillas de las variables a estudiar (Brownstein et al., 2019). Para el WGI utilizamos el diagrama de dispersión XY ya que muestra la relación entre el año y el puntaje de cada variable agregada, lo que por lo tanto permite una mejor visualización, observación.
14 Chen 2021	Análisis exploratorio de datos de fallas de LTPP utilizando técnicas estadísticas.	Este estudio investigó cómo los factores relevantes afectan las fallas mediante la realización de un análisis exploratorio de los datos de fallas del LTPP con técnicas estadísticas. Los factores relevantes incluyen los factores en cuatro categorías de repetición del tráfico, información del pavimento, clima local, propiedades de los materiales y tratamientos de mantenimiento. Los datos de fallas del LTPP se dividieron en datos previos y posteriores a la reparación para examinar el impacto de los factores en categorías y la efectividad del tratamiento de mantenimiento.
15 Fernandez (2020) Brasil	La educación estadística a la luz de la educación.	Las calificaciones reveladas se pueden resumir, en primer lugar, en una afinidad relacionada con el reconocimiento de que el desarrollo de la alfabetización, ya sea matemática o estadística, es una formación deseada que debe ser accesible para todos. En consecuencia, el esfuerzo desde ambas perspectivas se orienta cada vez más hacia la creación de oportunidades de acceso y receptividad del conocimiento para un público amplio.
16 Hui (2019) Perú	Aprenda R para estadística aplicada	Obtenga los fundamentos del lenguaje de programación R para realizar estadísticas aplicadas a uso de datos. Este libro cubre temas que van desde conceptos básicos de sintaxis de R, estadísticas descriptivas y visualizaciones de datos hasta estadísticas inferenciales y regresiones. Después de aprender la sintaxis de R, trabajará con visualizaciones de datos como histogramas y diagramas de caja, estadísticas descriptivas y estadísticas inferenciales como prueba t, prueba de chi-cuadrado, ANOVA, prueba no paramétrica y regresiones lineales.
17 Marzouk (2021) USA.	Evaluación del calentamiento global en Al Buraimi, sultanato de Omán, basada en el análisis estadístico de los datos de NASA POWER durante 39 años, y probando la confiabilidad de NASA POWER contra mediciones meteorológicas.	La temperatura del aire a 2 metros muestra un aumento a largo plazo a una tasa promedio de aproximadamente 0.039 °C por año, lo que corresponde a aproximadamente 1.5 °C durante 4 décadas. La desviación estándar de la muestra de la temperatura del aire a 2 metros no muestra un cambio confiable a lo largo del tiempo. No hay evidencia suficiente de que la media y la desviación estándar muestral del rango de temperatura del aire de 2 metros dependan del tiempo.
18 Orellana (2021) Ecuador	Estimación de la temperatura del aire mensual mediante sensores remotos en una región con topografía muy variable y escaso monitoreo en los Andes del sur de Ecuador	La aplicación del kriging de regresión fue limitada ya que menos del 50% de los meses tuvieron autocorrelación espacial en los residuos del modelo de regresión. Sin embargo, en estos meses el kriging de regresión incrementó el desempeño de la estimación. Los resultados de este trabajo de investigación aumentan la comprensión de la distribución espacial de la temperatura mensual del aire en la cuenca del río Paute, lo que mejorará la modelización hidrológica.
19 (2021) China	Aplicación del algoritmo de suavizado óptimo ARIMA-RTS en la predicción de la producción de pozos de gas	Sobre la base del modelo de filtro ARIMA-Kalman con capacidades de análisis eficientes de los factores que influyen en la secuencia temporal, se agrega el algoritmo de suavizado inverso RTS. El modelo ARIMA-RTS establecido puede eliminar la influencia de la asincronía y la histeresis. Cuando la fracturación y otras medidas de mejora de la producción tienen un gran impacto en la producción real, el modelo ARIMA-RTS puede estabilizarse rápidamente para reducir los errores de predicción.
20 Zheng (2022) Australia	Razones, desafíos y algunas herramientas para realizar investigaciones sobre transporte reproducible	Con este fin, la sección 2 resume los principales beneficios de la RR. La Sección 3 revisa la historia de la RR y la práctica de la RR en otras disciplinas. La sección 4 analiza cómo realizar la RR. La Sección 5 propone un procedimiento reproducible para la Investigación del Transporte. Finalmente, la Sección 8 analiza las oportunidades y los desafíos de realizar GR en beneficio de los estudiantes para el uso del programa en el (análisis estadístico).

En primer lugar, Leines González et al., (2020) indican que, se establece como requisito fundamental para llevar a cabo el desarrollo de resolución de muestras de estadística la programación de datos en formato de vector, permitiendo la opción de realizar esta carga mediante la relevancia del archivo de datos de información. Esto posibilita un procesamiento ulterior de la información con mayor eficacia. La disparidad en el dominio de los estudiantes respecto al procesamiento de datos en archivos se debe a las diferencias en su formación informática, que varía

según las especialidades y los niveles de estudio. Además, los comandos empleados en el procesamiento presentan un formato familiar para aquellos con experiencia en programación, pero puede resultar desconocido para otros estudiantes. En consideración a esto, se hace necesario proporcionar una capacitación y explicación previo a la carga de datos y el empoderamiento de variables de memoria, centrándose en los argumentos específicos de las funciones involucradas (Flores y Ruiz et al., 2017).

En consecuencia, será esencial llevar a cabo una revisión de conocimientos básicos en dos áreas específicas: El dominio de argumentos de las funciones y el almacenamiento de variables en memoria. El almacenamiento y procesamiento de archivos, con el propósito de capacitar a los estudiantes para acceder y manejar eficientemente grandes cantidades de datos durante el procesamiento. Frente a lo expuesto Alina et al., (2019) manifiesta que, se ha constatado que el espacio y tiempo para la capacitación en el uso del programa del software R y la resolución de los trabajos prácticos asignados en clase resulta insuficiente. En consecuencia, se ha llevado a cabo una replanificación de las clases, con la necesidad de tomar decisiones respecto a los conceptos que serán desarrollados en el aula y aquellos que los alumnos deberán abordar de forma autónoma. Como medida para abordar esta situación, se ha fortalecido el apoyo proporcionado a través del aula virtual, incluyendo contenidos diseñados tanto para la resolución en el horario de clase como para verificar en procesos de autoaprendizaje. Esto busca optimizar el aprendizaje de los estudiantes, permitiéndoles acceder a los recursos necesarios y desarrollar habilidades de forma más efectiva dentro de las limitaciones de tiempo existentes (Castro y Domínguez, 2022).

Según Marzouk, (2021) señala que, la colaboración con los catedráticos, se han recuperado los registros originales mostrados en el sitio y las cartillas de trabajos prácticos compartidas con los alumnos. Con base en la actividad valorada, se ha decidido mantener el programa y formato existente sin introducir cambios significativos en la forma de brindar orientación y comunicación. Sin embargo, se llevará a cabo una adaptación en la presentación para alinearla con la plataforma Moodle, asegurando así una transición suave y eficiente. Asimismo, Leines et al., (2020) manifiesta que, la información ha sido reformulada para su posterior publicación y actualización, migrando el entorno virtual al programa R con un formato de navegación mediante Exe-Learning. En respuesta a la solicitud de los catedráticos, el equipo de trabajo asumió la responsabilidad de conducir la clase práctica en el laboratorio. En esta sesión, se presentaron los conceptos iniciales y los primeros pasos para utilizar la suite R, aplicándola a problemas de Probabilidad y Estadísticas. La clase práctica incluyó la resolución efectiva de un trabajo, con entrega por parte de los estudiantes y posterior corrección.

Para llevar a cabo esta enseñanza, se diseñó una secuencia didáctica que comprende una fase inicial de recuperación conceptual, seguida por el desarrollo de habilidades (como el uso de funciones y variables y manejo de archivos), y culminando con una evaluación tipo rúbrica. Se adjuntan al final ejemplos de ambas instancias para mayor claridad (Arriaza et al., 2008). En paralelo con esta iniciativa, se ha desarrollado la unidad correspondiente en el aula virtual. Esta unidad incluye recursos destinados tanto para el autoaprendizaje de los estudiantes como para el seguimiento de la clase práctica. Además, se proporcionan materiales específicos para guiar la resolución del trabajo que será evaluado, junto con la correspondiente rúbrica evaluativa. Este enfoque integral tiene como objetivo ofrecer a los estudiantes un conjunto completo de herramientas y orientaciones para abordar la tarea de manera efectiva y facilitar una evaluación clara y constructiva al final del proceso (Fernández, (2020).

Por su parte Martínez et al., (2022) sostiene que, una vez establecida la estructura de la clase, se procedió a informar a los docentes para garantizar el cumplimiento de las condiciones necesarias y asegurar su conocimiento sobre la participación que se esperaba de ellos. Se optó por utilizar la misma secuencia didáctica para todas las especialidades y para todos los docentes involucrados, permitiendo a los alumnos asistir a cualquiera de las clases prácticas, e incluso repetir la asistencia si así lo requerían. Según Rodríguez, (2019) para abordar el desarrollo de los trabajos asignados y el manejo de estadística, se constató la necesidad de poseer conocimientos en áreas adicionales, las cuales fueron revisadas previamente antes de la resolución específica. Estas revisiones abarcaron el uso de herramientas informáticas y el tratamiento de la información, y comprendieron: El manejo de comandos de programación, la transferencia de argumentos en funciones y el almacenamiento de variables en memoria.

Se reconoció que el tiempo necesario para revisar conceptos iniciales debía considerarse cuidadosamente dentro del marco temporal establecido por las cátedras. Por este motivo, se tomó la decisión de abordar algunos de estos conceptos durante el taller, pero se estableció la necesidad de reforzar y consolidar estos conocimientos a través de actividades en el aula virtual, facilitando así procesos de autoaprendizaje. Este enfoque busca proporcionar a los estudiantes las herramientas necesarias para elevar el nivel de sus informes, abordando de manera más completa tanto la resolución de problemas como la expresión escrita de los resultados (Mares y Domínguez, 2022). Dentro del estudio fue necesario revisar las siguientes temáticas:

**Distribuciones de Probabilidad**

La presencia de fenómenos o experimentos no deterministas, en los cuales el conocimiento de las condiciones no garantiza resultados específicos, hace imperativo el empleo de una función que asigne niveles de certeza a cada posible resultado. Aquí es donde entra en juego la teoría de la probabilidad. Los experimentos o fenómenos que exhiben esta característica se llaman aleatorios. Desde una perspectiva intuitiva, al asignar valores numéricos con cierto criterio al fenómeno, se crea lo que se conoce como variable aleatoria. Una comprensión adecuada de estos conceptos resulta esencial para estudiar conjuntos extensos a partir de pequeñas partes llamadas muestras. Esto da lugar a lo que se conoce como inferencia estadística, que permite realizar afirmaciones y conclusiones sobre la población en general basándose en el análisis de muestras representativas.(Alina et al., 2017).

Por otro lado Alonso y Zamora, (2020) indican que, antes de sumergirse en el contenido, se describirán una serie de fenómenos que se pueden asimilar a las distribuciones de probabilidad que se explorarán en detalle en este capítulo, proporcionando así un contexto inicial para comprender la aplicación práctica de estos conceptos. En el caso de un examen tipo test con 10 preguntas, donde cada pregunta tiene 4 opciones y solo una es correcta, el número de aciertos más probable al contestar al azar se encuentra en el centro del rango posible. Dado que cada pregunta tiene 4 opciones, y al responder al azar se tiene una probabilidad del 25% (1/4) de acertar en cada pregunta, el número de aciertos más probable es aproximadamente la mitad de las preguntas, es decir, 5 aciertos. En relación al sorteo de la ONCE, la distribución de las unidades de los premios se basa en la premisa de que cada número entre el 0 y el 9 tiene la misma probabilidad de ser la unidad ganadora. Por lo tanto, las unidades de los premios en el sorteo de la ONCE están uniformemente distribuidas, y cada una tiene la misma probabilidad de salir como ganadora (Contreras, García et al., 2010). Probabilidad de ser elegido al menos uno de los temas preparados:

Probabilidad de esperar menos de 10 minutos para coger el autobús: Dado que hay salidas cada media hora y el individuo llega a la estación de autobuses en un momento desconocido, la probabilidad de esperar menos de 10 minutos es  $10/30=1/3$ . Las probabilidades de funcionamiento o no funcionamiento se pueden calcular multiplicando la probabilidad diaria por el número de días (365).

Estructura de la distribución de los 350 alumnos:

La descripción de la estructura de la distribución de los alumnos requeriría información específica sobre los intervalos o categorías en las que se agrupan las variables (peso, edad, estatura, etc.). Sin esa información, es difícil proporcionar una respuesta detallada. Dependiendo de la naturaleza de las variables, la distribución podría ser uniforme, normal, sesgada a la derecha o a la izquierda, entre otras posibilidades. Para Rodríguez, (2019) las situaciones planteadas implican la realización de experimentos aleatorios, cada uno con resultados de naturaleza diferente. Por ejemplo, extraer el número de las unidades entre diez posibles, o sacar dos temas entre cien, son ejemplos de experimentos aleatorios con resultados distintos.

Estos experimentos llevan a una clasificación fundamental en dos categorías principales: En el primer caso, como el número de aciertos en las 10 preguntas, los resultados son finitos y discreto, variando entre 0 y 10. En el segundo caso, como el tiempo de espera para coger el autobús, los resultados son continuos, ya que pueden tomar infinitos valores dentro del intervalo (0, 30), limitados solo por la precisión de los instrumentos de medición. El paso inicial en la resolución de estos problemas es seleccionar el modelo teórico de probabilidad apropiado para cada situación específica, ya sea un modelo discreto o continuo, lo que proporciona la base para realizar cálculos y análisis estadísticos precisos (Martinez et al., 2022). En la tabla 2 se aprecia la distribución discreta.

**Tabla 2:** Distribución discreta.

DISCRETAS		
Distribución	Parámetros	En Rcmdr
Binomial	$n = size; p = prob$	binom
Binomial negativa	$n = size; p = prob$	nbinom
Geométrica	$p = prob$	geom
Hipergeométrica	$(N, K, n) = (m, n, k)$	hyper
Poisson	$\lambda = lambda$	pois

Según, Fernández y Lizana (2020) la distribución de probabilidad proporciona una descripción sistemática de cómo se distribuyen las probabilidades entre los posibles valores de la variable aleatoria. Esta herramienta es fundamental en la teoría de la probabilidad y la estadística, ya que facilita la interpretación y análisis de los resultados de un experimento aleatorio (Méndez, 2020).

Por su parte Arriaza et al., (2018) menciona que, según la naturaleza de la variable aleatoria, se pueden considerar distribuciones de probabilidad discretas o continuas. Para acceder al programa en R, se dirige a: Distribuciones → Distribuciones continuas, o en Distribuciones → Distribuciones discretas. También se puede registrar en la ventana de instrucciones la denominación de la distribución, precedido por una "D" si se busca la función de densidad, una "P" para la función de distribución, una "Q" para los cuantiles y una "r" para generar una muestra aleatoria de la distribución. Se debe asegurar incluir los argumentos importantes en cada actividad (González et al., 2020).

El script en R para analizar los resultados de la prueba y obtener las impresiones gráficas en un análisis de regresión (figura 2).

```

Usando R
> var.test(residuals(modelo)[x<median(x)],residuals(modelo)[x >
median(x)])

      F test to compare two variances

data: residuals(modelo)[x < median(x)] and residuals(modelo)[x >
median(x)]
F = 1.0277, num df = 4, denom df = 4, p-value = 0.9795
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.1070041 9.8708029
sample estimates:
ratio of variances
      1.027724

```

**Figura 2:** Uso de la R en análisis estadístico.

Los resultados de esta prueba indican (valor  $p = 0.9795$ ) que, al nivel de significancia del 5%, no se dispone de evidencia para afirmar que exista una disparidad en la variabilidad de los residuales entre estas submuestras (figura 3). En otras palabras, no se puede sostener la presencia de heterocedasticidad en los residuales, cumpliendo así con el supuesto. Es relevante notar que el intervalo de confianza para el cociente de varianzas incluye el valor de 1, reafirmando la conclusión mencionada Carrera, Villacrés et al., (2019).

```

Usando R
> var.test(residuals(modelo)[x<median(x)],residuals(modelo)[x >
median(x)])

      F test to compare two variances

data: residuals(modelo)[x < median(x)] and residuals(modelo)[x >
median(x)]
F = 1.0277, num df = 4, denom df = 4, p-value = 0.9795
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.1070041 9.8708029
sample estimates:
ratio of variances
      1.027724

```

**Figura 3:** Ejemplo de la R en análisis estadístico.

Rodríguez, (2019) resalta que, la representación visual superior de la figura 9.2, se evidencian residuales que fluctúan alrededor de cero con valores moderadamente dispersos, sugiriendo homocedasticidad. Así mismo, González, Videgaray et al., (2019) En la sección inferior, los residuales exhiben proximidad a la línea, insinuando una distribución que se asemeja a la normalidad. No obstante, las pruebas de hipótesis diseñadas para evaluar la normalidad y homocedasticidad de los errores son métodos más eficaces y apropiados para explorar estos supuestos en comparación con las interpretaciones basadas en gráficos (figura 4).

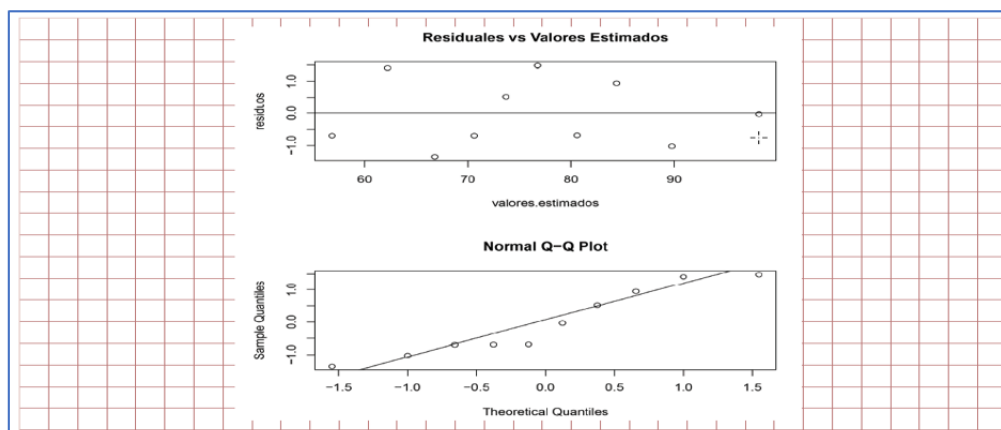


Figura 4: Residuales y pruebas.

## 4. Conclusiones

La experiencia exitosa en la elección, instalación y promoción del Programa R, junto con la programación de clases centradas en estadística, ha resaltado la versatilidad y el potencial de este programa en las disciplinas relacionadas con la Ingeniería. Esto demuestra que podemos aplicarla en cualquier área de la carrera, aprovechando paquetes disponibles en la web de forma gratuita. Esta adopción elimina la necesidad de utilizar herramientas previamente consideradas demasiado complejas, permitiendo a los estudiantes abordar problemas de manera más accesible y eficiente. El trabajo colaborativo e interdisciplinario entre los docentes, estudiantes de ambos grupos no solo facilitó el trabajo en equipo para agilizar las etapas de análisis, aplicación y comprensión, sino que también reforzó en los estudiantes habilidades fundamentales, como el uso del programa R, la capacidad de la motivación y autoaprendizaje, que derivó de resultados exitosos, y la capacidad de abordar situaciones y ejercicios estadísticos. La decisión de contar con un grupo de trabajo con una programación del programa R, junto con la creación de actividades y el intercambio continuo de experiencias, y la interacción entre los docentes, sienta las bases para replicar y generar sostenibilidad de esta experiencia en otras áreas en el futuro. Este enfoque promueve el enriquecimiento mutuo y la expansión de buenas prácticas pedagógicas, generando un impacto positivo en la formación integral de los estudiantes.

## 5. Referencias bibliográficas

- Abdollahi, M., Farjad, B., Gupta, A., & Hassan, Q. K. (2022). CMIP6-D&A: An R-based software with GUI for processing climate data available in network common data format. *SoftwareX*, 18.
- Adolfsson, A., Ackerman, M., & Brownstein, N. C. (2019). To cluster, or not to cluster: An analysis of clusterability methods. *Pattern Recognition*, 88, 13–26.
- Alina, D., Rodríguez, M., Zamora, D. C. L., Nelsa, D., Sagaró Del Campo, M., Annia, D., & Rodríguez, D. A. (2017). ARTÍCULO ORIGINAL Análisis estadístico implicative para la identificación de factores pronósticos de la mortalidad por cáncer de mama Statistical analysis involving the identification of mortality prediction factors due to breast cancer. In *MEDISAN* (Vol. 21, Issue 4).
- Alonso, D., & Zamora, V. (n.d.). La importancia de la estadística aplicada para la toma de decisiones en Marketing THE IMPORTANCE OF THE STATISTICS APPLIED TO MARKETING DECISIONS.
- Arriaza Gómez, A. J., Fernández, F., López Sánchez, P. M. A., Muñoz Márquez, M., Pérez, S., & Navas, P. A. S. (2008). *Estadística Básica con R y R-Commander*.
- Carrera-Villacrés, D., Bastidas, A., Caza, N., & Chicaiza, M. (n.d.). *MANUAL DE R APLICADO A LA INGENIERÍA DEL AGUA*.
- Contreras García, J. Miguel., Molina Portillo, Elena., & Arteaga Cezón, Pedro. (2010). *Introducción a la programación estadística con R para Profesores*. Grupo de Educación Estadística, Universidad de Granada.



- Fernández Bedoya, V. H. (2020a). Tipos de justificación en la investigación científica. *Espí-ritu Emprendedor TES*, 4(3), 65–76.
- Fernández Bedoya, V. H. (2020b). Tipos de justificación en la investigación científica. *Espí-ritu Emprendedor TES*, 4(3), 65–76.
- Fernández Lizana, M. (2020). Advantages of R as a tool for data Analysis and Visualization in Social Sciences. *Revista Científica de La UCSA*, 7(2), 97–111.
- Fernández-Hernández, F. J., & Andrade-Escobar, L. (2021). Statistical education in the light of critical mathematics education. In *Revista Colombiana de Educacion* (Vol. 1, Issue 83, pp. 1–20). Research Center of Universidad Pedagógica Nacional.
- Flores-Ruiz, E., Guadalupe Miranda-Novales, M., Ángel Villasís-Keever, M., & Ángel Villasís-Keever miguel, M. (2017). Metodología de la investigación. In *Rev Alerg Mex* (Vol. 64, Issue 3).
- Gallardo, E. (2017). Metodología de la Investigación. Manual Autoformativo Interactivo I. *Universidad Continental*, 1, 98.
- García, M. L. S., de Segura, B. I. E., López, J. M. S., & Romero, C. S. (2020). Tablet devices. Dynamic strategy to promote significant learning at University. *Pixel-Bit, Revista de Medios y Educacion*, 59, 97–123.
- González-Videgaray, M., Romero-Ruiz, R., Del, N., & Goris-Mayans, C. (n.d.). “IV Encuentro sobre Didáctica de la Estadística, la Probabilidad y el Análisis de Datos.”
- Jahuey Martínez, F. J., Herrera Ojeda, J. B., & Paredes Sánchez, F. A. (2022). El programa R: una estrategia inicial para su entendimiento y aprendizaje. *Revista Digital Universitaria*, 23(4).
- Leines González, A., Michele, M., Díaz, T., & Carlos, J. (2020a). UNIVERSIDAD TÉCNICA PARTICULAR DE LOJA.
- Leines González, A., Michele, M., Díaz, T., & Carlos, J. (2020b). UNIVERSIDAD TÉCNICA PARTICULAR DE LOJA.
- Mares Castro, A., & Domínguez Domínguez, J. (2022). Evaluación estadística de índices de desempeño para el proceso de división de rollos de EVA. *Ingeniería Investigación y Tecnología*, 23(2), 1–15.
- Marzouk, O. A. (2021). Assessment of global warming in Al Buraimi, sultanate of Oman based on statistical analysis of NASA POWER data over 39 years, and testing the reliability of NASA POWER against meteorological measurements. *Heliyon*, 7(3).
- Méndez Escobar, A. (2020). Educación en tiempos de pandemia (covid-19). *Revista Universidad de La Salle*, 1(85), 51–59.
- Ricardo, M., & Rubio, C. (n.d.). E stadística con aplicaciones en R. Roa, I. (2015). *Metodos Cuantitativos*. 1–52.
- Rodríguez Silva, J. L. Á. (2019). ¿Qué puede hacer el software R para resolver tus problemas? *Revista Digital Universitaria*, 20(3).
- Tapia, F. Z. R. (2022). Estrategias didácticas y competencias digitales en la práctica de la eneseñanza de docentes con trabajo remoto de Lima- 2021. *Ucv*, 1–163.
- Zheng, Z. (2021). Reasons, challenges, and some tools for doing reproducible transportation research. *Communications in Transportation Research*, 1.