



Categoría: International Congress on Education, Technology and Science (CIETYC 2024) Quindio - Colombia

ORIGINAL

Data analysis using statistical graphs from an exploratory approach

Análisis de datos usando gráficos estadísticos desde un enfoque exploratorio

Mailen Hoyos¹ ✉, Tulia Esther Rivera Florez¹ ✉, Isabel Cristina Puerta Lopera¹ ✉

¹Universidad Industrial de Santander, Colombia.

Citar como: Hoyos M, Rivera Florez TE, Puerta Lopera IC. Data analysis using statistical graphs from an exploratory approach. SCT Proceedings in Interdisciplinary Insights and Innovations. 2024;2:472. DOI: <https://doi.org/10.56294/piii2024472>

Recibido: 22-04-2023

Revisado: 22-06-2024

Aceptado: 24-07-2024

Publicado: 26-07-2024

Editor: Rafael Romero-Carazas 

ABSTRACT

The degree work entitled "Data analysis using statistical graphics from an exploratory approach" addressed the need to transform the teaching of statistics through a technological and innovative approach. Traditional textbooks were found to be deficient in terms of pedagogical design, which motivated the development of an interactive book that included new pedagogical methods, technology and data analysis.

For the realization of this work, a review of statistics textbooks was carried out and, taking this into account, an interactive book was designed and programmed using the Shiny library of the R studio software. In addition, we conducted a pilot test with a group of students of Statistics I at the Universidad Industrial de Santander and progress was presented at the 32nd Symposium of Statistics, which allowed an evaluation of user interaction with the material, obtaining feedback on the activities proposed.

The results showed the importance of the exploratory approach, as well as interactivity in the teaching of statistics, demonstrating how effective it is to integrate it with contexts and theory. The need to adapt the activities was also mentioned, emphasizing the importance of giving clear instructions, examples and interactive elements for the clear instructions, examples and interactive elements so that students can learn autonomously. The dynamic nature of the book and the commitment to keep it in constant evolution was highlighted, showing commitment to continuous improvement and the ability to adapt to the diverse needs of students.

Keywords: Interactive book, Exploratory approach, R Studio, Shiny, Statistics.

RESUMEN

En el trabajo de grado titulado “Análisis de datos usando gráficos estadísticos desde un enfoque exploratorio” se abordó la necesidad de transformar la enseñanza de la estadística a través de un enfoque tecnológico e innovador. Se detectó que los libros de texto tradicionales tenían deficiencias en cuanto a diseño pedagógico, lo cual motivó al desarrollo de un libro interactivo que incluyera métodos pedagógicos nuevos, tecnología y análisis de datos.

Para la realización de este trabajo, se realizó una revisión a libros de texto de estadística y teniendo en cuenta esto, se llevó a cabo el diseño y programación un libro interactivo haciendo uso de la librería Shiny del software R studio. Además, realizamos una prueba piloto con un grupo de estudiantes de Estadística I en la Universidad Industrial de Santander y se presentaron avances en el 32° Simposio de estadística, lo cual permitió una evaluación de la interacción del usuario con el material obteniendo retroalimentación sobre las actividades planteadas.

Los resultados mostraron la importancia del enfoque exploratorio, así como también de la interactividad en la enseñanza de la estadística, demostrando a su vez, lo efectivo que es integrarlo con contextos y teoría. Se mencionó también la necesidad de adaptar las actividades, haciendo hincapié en la importancia de dar instrucciones claras, ejemplos y elementos interactivos para que los estudiantes puedan aprender de manera autónoma. Se resaltó, la naturaleza dinámica del libro y el compromiso de mantenerlo en constante evolución, mostrando compromiso de una mejora continua y la capacidad de adaptarse a las diversas necesidades de los estudiantes.

Palabras clave: Libro interactivo, Enfoque exploratorio, R Studio, Shiny, Estadística.

INTRODUCCIÓN

La enseñanza de la Estadística se ha posicionado como un tema importante en el sistema educativo gracias a la demanda de profesionales con formación en esta disciplina. Aún así, su enseñanza sigue estando dominada por un enfoque tradicional, el cual es un escenario que no esperamos observar como ciudadanos de este siglo dado el marco conceptual en Didáctica de la Estadística, la variedad de recursos tecnológicos y de softwares especializados existentes.

En cuanto a recursos didácticos, el libro de texto solía tener un uso preponderante en el aula. Hoy día, el acceso a internet ha posicionado otras opciones como páginas web, canales de YouTube y applets. Sin embargo, los recursos tecnológicos que nos ha traído la modernidad nos hacen pensar que es necesario analizar cómo se puede mejorar su uso para que realmente sean una herramienta cognitiva que potencie el aprendizaje del estudiante. Consideramos que es posible hacer trascender el papel del libro de texto en el aula y no limitarlo a ser un simple medio de transmisión de información. Los nuevos formatos pueden hacer posible más formas de interacción, la variedad de recursos que incluyen puede orientarse a: promover un aprendizaje más autónomo, permitir la exploración y análisis de situaciones reales, incentivar la creatividad, potenciar el pensamiento crítico y dar respuesta a diferentes estilos de aprendizaje y necesidades de los estudiantes.

Este trabajo tuvo como principal motivación contribuir al mejoramiento del proceso de enseñanza-aprendizaje en cursos de formación en Estadística a nivel universitario a través de una mediación que se aparta del enfoque tradicional basado en la ejercitación de fórmulas al resolver misceláneas de ejercicios rutinarios. En particular queremos aportar a la comunidad educativa nacional el libro de texto “Análisis de datos desde un enfoque exploratorio”, en este proyecto se presenta un capítulo de éste cuyo tema central es el análisis de datos a través de gráficos estadísticos, para su diseño se combinaron tres elementos: resolución de problemas, énfasis en la exploración dinámica de datos y la incorporación de herramientas tecnológicas.

De acuerdo con las características del proyecto se combinaron conocimientos en tres áreas: Didáctica de la Estadística, Herramientas para diseño web disponibles en R y Elementos básicos en escritura y edición de libros. A continuación, se presentan algunas ideas que soportaron el desarrollo de la propuesta. En cuanto a diseño gráfico, la rama que se enfoca en la presentación de publicaciones, como libros, revistas y periódicos entre otros, se conoce como diseño editorial; en esta se incluyen aspectos visuales, teniendo en cuenta la estética que define la publicación y las condiciones de impresión y recepción (Rosas, 2012)

Como se busca crear armonía entre textos, imágenes y disposición de estos para transmitir mensajes, es importante tener en cuenta que el diseño editorial, debe tener en cuenta algunos objetivos como adaptarse al público, facilitar el proceso creativo, o ser diferente a otras publicaciones.

Ghinaglia (2009) sugiere que para elaborar materiales que incluyan componentes gráficos se deben tener en cuenta elementos como la tipografía, haciendo referencia específicamente a la selección y uso de tipos de letra, así como también el tamaño, y estilo entre otras características ya que esto permite la legibilidad en los textos. Otro elemento a tener en cuenta, son las imágenes, que tienen como objetivo transmitir ideas o mensajes, pudiendo ser ilustraciones o gráficos, por ejemplo. El color es un elemento clave también ya que se puede utilizar para transmitir mensajes, crear jerarquías o resaltar elementos importantes.

Según Iñigo y Makhlof (2014), los autores deben atender a las siguientes ideas en cuanto a diseño editorial:

- Párrafo: es la unidad básica de escritura y consiste en una serie de oraciones interconectadas que se unen para formar una idea o argumento.
- Legibilidad: Se refiere a la facilidad con la que el lector lee y comprende el texto.
- Portada: es la parte exterior de un libro que protege y muestra el contenido
- Maquetación: se refiere a la presentación visual de las páginas del libro, incluida la selección y organización del texto y los elementos gráficos de cada página.
- Retícula: es una estructura invisible en un diseño que se utiliza para dividir una página en columnas y filas.

Serrano y Pugo (2021) citando a Aubia, mencionan que el diseño interactivo es caracterizado por la búsqueda de soluciones en las cuales el lector se relaciona con el contenido a través de herramientas que posibilitan realizar acciones o navegar a través del contenido, por ejemplo, incluyendo recursos como hipervínculos, animaciones, y actividades interactivas.

En relación con los elementos de didáctica de la estadística, Carballo y Domínguez (2010), sugieren que el pensamiento estadístico está basado en tres principios, los cuales son: Todo trabajo ocurre en un sistema de procesos

interconectados, la variación existe en todos los procesos y la clave del éxito se alcanza comprendiendo y reduciendo la variación del proceso. Estrella (2017), por su parte, menciona que este pensamiento tiene cuatro características fundamentales: reconocer la necesidad de datos, tras numeración, percepción de la variación y razonamiento con incertidumbre.

Es por esto que el aprendizaje del pensamiento estadístico no es unidimensional, y como sugieren Behar y Ojeda (2004), entender la estadística va más allá de un enfoque lineal o limitado. Se resalta la importancia de tener una perspectiva amplia, para abarcar la complejidad de la estadística en el contexto de la educación superior.

La enseñanza basada en proyectos beneficia a los cursos de estadística, ya que como menciona Batanero (2004), está relacionada con sus aplicaciones, además permite que los estudiantes apliquen conceptos a situaciones del mundo real y permite desarrollar habilidades de pensamiento crítico. Batanero también señala que trabajar a través de proyectos incrementa la motivación en los estudiantes, pues esto les permite abordar temas que son de su interés usando datos reales.

MATERIALES Y MÉTODOS

El capítulo desarrollado, gira en torno al análisis de datos a través de gráficos estadísticos desde un enfoque exploratorio, integrando el aprendizaje teórico y práctico, ofreciendo a los usuarios un medio dinámico que permite la exploración de conceptos a través de gráficos y simulaciones programadas en el software R studio, el cual es un entorno de desarrollo integrado para el lenguaje de programación R.

Para el desarrollo del capítulo, se usó una serie de herramientas entre las cuales se destacan librerías que permiten la creación de aplicaciones interactivas, para la elaboración de visualización de datos, y para incorporar mapas interactivos entre otras. Además, se añadieron elementos visuales y se añadieron funciones como lectura de texto y adaptación del tamaño del texto mediante botones, haciendo uso de JavaScript, lo cual refleja el compromiso con la mejora de la accesibilidad.

Además de haber sido diseñada para proporcionar conocimientos teóricos, esta unidad didáctica, busca asegurar que el lector pueda aplicar dicha teoría, combinando la exploración de datos, la interactividad y la preparación para la comunicación efectiva de resultados preparándolo así para enfrentar desafíos del mundo real en el ámbito del análisis de datos.

RESULTADOS Y DISCUSION

A continuación, se presenta de manera resumida la estructura del capítulo del libro.

1. Introducción

Esta sección tiene como objetivo que el usuario reconozca la importancia de seleccionar medidas de tendencia central adecuadas. Además, se muestra cómo al usar el promedio de forma indiscriminada, en ocasiones, puede conducir a presentar errores en la descripción del comportamiento de variables, evidenciando la importancia del uso de criterios técnicos para incorporar de forma adecuada las medidas de resumen. Así mismo, esta sección sienta bases para los temas que se abordarán en los siguientes

capítulos, tales como pruebas estadísticas de normalidad, medidas de tendencia central y técnicas más avanzadas de análisis de datos como la detección de valores extremos o ajuste de modelos.

2. Validando un supuesto distribucional

En esta sección se muestra al usuario una guía de cómo validar supuestos distribucionales en diferentes contextos de análisis de datos. Se establece un marco conceptual para entender la importancia de validar la forma de la distribución de una variable, tanto en situaciones donde ya existe un modelo preestablecido como cuando no lo hay. Con relación a contenido, se introduce el concepto de supuestos distribucionales, destacando su importancia en la estadística. Brindando una revisión sobre cómo y cuándo aplicar modelos distribucionales preestablecidos.

También, se aborda la complejidad de trabajar con datos que no tienen un modelo preestablecido, enfatizando la tarea del analista en identificar la distribución que mejor se ajusta a los datos.

Figura 1

Sección 1: Cómo validar un supuesto distribucional



Esta se enriquece con histogramas dinámicos, los cuales permiten que el lector explore a profundidad los conceptos y los aplique a otros conjuntos de datos. Esta sección es importante ya que permite al lector evitar conclusiones erróneas o interpretaciones incorrectas.

Se introduce también, el uso de diversas herramientas gráficas para la validación de supuestos, como histogramas, boxplot y gráficos PP y QQ, entre otros.

3. Pruebas gráficas

El objetivo de esta sección es entregar al usuario las herramientas necesarias para la visualización y la comprensión de conjuntos de datos a través de diferentes gráficos y representaciones. Concretamente, se busca que el lector aprenda a crear e interpretar diferentes diagramas como histogramas, caja y bigotes, CC, PP y QQ plots.

3.1. Diagrama de caja y bigotes.

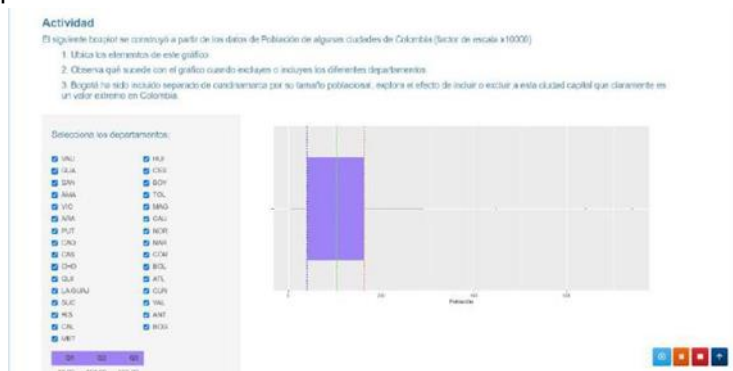
Se inicia con una analogía que permite relacionar el uso de estadísticos de orden para resumir conjuntos de datos lo que puede facilitar al lector la interpretación y uso

práctico de estos recursos. Además, se presentan escenarios interactivos que muestran diferentes diagramas de caja y bigotes como se observa en la figura 2, en el cual se permite modificar el conjunto

inicial de datos, al activar o desactivar datos particulares, mostrando en tiempo real, cómo afecta al diagrama la inclusión o exclusión de los datos.

Figura 2

Sección 2, pruebas gráficas. Subsección, diagrama de caja y bigotes. Boxplot interactivo población por departamento.



Esta actividad tiene como objetivo que el lector identifique los elementos claves de este gráfico a partir de un código de color donde se relacionan las definiciones con la ubicación en el gráfico. Los escenarios planteados, se complementan con preguntas que indagan por comportamientos inusuales en los diagramas de caja y bigote que permiten valorar la presencia de ciertos detalles y elementos en el gráfico, como por ejemplo de dónde se obtienen, qué variaciones pueden presentar, y en general qué características estadísticas de la distribución de los datos deja en evidencia el gráfico.

3.2. Histograma.

En esta sección se proporciona un algoritmo para construir histogramas manualmente, así como también un tutorial de cómo hacerlo utilizando el lenguaje de R. Además, se incluyen diferentes visualizaciones para ilustrar los conceptos estadísticos, y se incluyó una herramienta que ilustra diferentes gráficas de densidad de una distribución y ayuda a identificar las características de variación más comunes.

Así mismo, esta sección se constituye como un puente entre las medidas de tendencia central, y temas más avanzados como pruebas de hipótesis, que serán abordados en secciones posteriores.

3.3. PP y QQ plot.

El propósito de esta sección es mostrar al usuario las instrucciones para la creación e interpretación de métodos gráficos para evaluar normalidad en una muestra. Se introducen 3 tipos de gráficos: C-C plot, gráfico que ha mostrado ser efectivo para la interpretación de los dos próximos, el percentil-percentil o P-P Plot y el cuantil-cuantil o QQ Plot, los cuales permiten determinar si los datos son procedentes de una distribución normal.

3.4. Ejercicios propuestos.

El objetivo de esta sección es profundizar en la comprensión y aplicación de los métodos gráficos estudiados, centrando la atención en que el usuario, haciendo uso de diferentes softwares como "ESTAT" y "R", compare y concilie los gráficos construidos a partir de una misma muestra. Adicionalmente, la sección incluye otras herramientas

interactivas, que permiten al lector interactuar, comparar y aprender de manera más efectiva.

3.5. Comunicando resultados.

El objetivo de esta sección supone un conocimiento previo en probabilidad y algunos elementos de estadística descriptiva, tales como medidas de tendencia central, dispersión y posición. Se incluyen video tutoriales como se aprecia en la figura 3. Con el propósito de guiar la construcción de dashboards con gráficos dinámicos construidos en Excel, la cual se considera una herramienta accesible y muy utilizada en el mundo profesional.

Figura 3

Sección 2, pruebas gráficas. Subsección, Comunicando Resultados. Video tutoriales para realizar Dashboard en Excel



Para el desarrollo de este capítulo se llevó a cabo una prueba piloto dividida en dos sesiones de dos horas, con estudiantes del curso de Estadística I. Se permitió a los estudiantes, interactuar con el recurso a partir de la instrucción de la docente; donde se pidió que probaran las diversas funciones incorporadas. Los estudiantes realizaron comentarios con respecto a la exploración, los cuales fueron recopilados para realizar ajustes y mejorar la interfaz del recurso y la utilidad de algunas de las herramientas incorporadas en el libro. En la segunda sesión, los estudiantes utilizaron la versión actualizada, la cual incluía las mejoras que derivaron del ejercicio previo de exploración.

Esta segunda sesión, permitió la mejora de las instrucciones para que pudieran aprovechar el material en un contexto autodidacta, ya que hicieron uso del recurso de manera independiente.

CONCLUSIONES

Este trabajo ofrece una visión crítica, y resalta la necesidad de reformar los enfoques tradicionales en la enseñanza de la estadística, ya que se identifican carencias que dificultan en aprendizaje, lo que requiere priorizar un aprendizaje autónomo y el uso de herramientas interactivas que permitan un desarrollo académico óptimo. Como respuesta a dichas limitaciones, se propuso un enfoque que reúne elementos exploratorios e interactivos, respaldado por el uso de herramientas como Rshiny, que transforman la manera en que se presentan los conceptos y ofrecen entornos de aprendizaje dinámicos y flexibles.

La retroalimentación valida la efectividad del enfoque propuesto, subrayando la necesidad de proporcionar instrucciones claras y el aporte de actividades interactivas para facilitar el aprendizaje autónomo. Este proyecto es un punto de partida, para desarrollar nuevos capítulos que aborden otros temas estadísticos, reafirmando el compromiso con la evolución y mejora continua, así como también la adaptabilidad de acuerdo con las necesidades tanto de los estudiantes como del campo estadístico en general.

REFERENCIAS

1. Batanero, C., & Díaz, C. (2004). El papel de los proyectos en la enseñanza y aprendizaje de la estadística. Aspectos didácticos de las matemáticas, 125-164.
2. Behar Gutiérrez, R y Ojeda Ramírez, M. (2013). El proceso de aprendizaje de la estadística.
3. Behar Gutiérrez, R. & Grima Cintas, P. (2004). La Estadística en la Educación Superior ¿Formamos Pensamiento Estadístico? Ingeniería y Competitividad, 5(2).
4. Berihuete Macías, Á., Ramos García, C. D., & García Ramos, J. A. (2018). Libro dinámico y colaborativo de Estadística descriptiva. Proyectos de Innovación y Mejora Docente 2017/2018. Universidad de Cádiz.
5. Burbano-Pantoja, V. M. A., Valdivieso-Miranda, M. A., & Aldana-Bermúdez, E. (2017) Conocimiento base para la enseñanza: un marco aplicable en la didáctica de la probabilidad. Rev.investig. desarro.innov, 7(2), 269-285. Doi: 10.19053/20278306.v7. n2.2017.6070
6. Carballo Monsivaia, C.A. & Domínguez Domínguez, J. (2010). Impulsando el Pensamiento Estadístico con Seis Sigma. En Estadística Aplicada: “Didáctica de la Estadística y Métodos Estadísticos en Problemas Socioeconómicos”
7. Cuétara Hernández, Y., Salcedo Estrada, I. M., & Hernández Díaz, M. (2016). La enseñanza de la estadística: antecedentes y actualidad en el contexto internacional y nacional. Atenas, 3(35), 125-140.
8. Devore, J. L. (2019). Introducción a la probabilidad y estadística para ingeniería y ciencias. Cengage Learning.
9. Estrella, Soledad. (2017). Enseñar estadística para alfabetizar estadísticamente y desarrollar el razonamiento estadístico.
10. Garfield, J. et al., 2005. Guidelines for Assessment and Instruction in Statistics Education (GAISE): College report. Alexandria, Virginia: The American Statistical Association.
11. Garfield, J., Hogg, B., Schao, C., y Whittinghill, D. (2002). First courses in statistical science: The status of educational reform efforts. Journal of Statistics Education, 10(2), 1-14.
12. Ghinaglia, D. (2009). Taller de diseño editorial: entre corondeles y tipos. Diseño en Palermo. IV Encuentro Latinoamericano de Diseño.
13. Guerrero Reyes, L. (2016). El Diseño Editorial. Guía para la realización de libros y revistas [Trabajo de fin de máster, Universidad Complutense de Madrid]. Repositorio de la Universidad Complutense de Madrid.
14. Iñigo Dehud, L. S., & Makhlof Akl, A. (2014). Diseño Editorial: Manual de conceptos básicos (1st ed.). Universidad Autónoma del Estado de Morelos.
15. Kapr, A. (1985). 101 reglas para el diseño de libros. Empresa Editoriales de Cultura y Ciencia.
16. Los libros electrónicos: la tercera ola de la revolución digital Mendenhall, W., Beaver, B. M., Beaver, R. J. (2015). Introducción a la probabilidad y estadística. Cengage Learning.
17. Ministerio de Educación Nacional, (2006). Estándares Básicos de Competencias en Lenguaje, Matemáticas, Ciencias y Ciudadanas: Guía sobre lo que los estudiantes deben saber y saber hacer con lo que aprenden
18. Navidi, W. (2022). Estadística para ingenieros y científicos. McGraw-Hill Interamericana.
19. Serrano Quezada, K. J., & Pugo Guamán, M. E. (2021). Diseño de un libro interactivo y material didáctico para generar conciencia en los niños y niñas sobre el cambio climático.
20. Valverde, M. M. M. (2013) Aportes de la investigación en educación estadística a la enseñanza de la estadística en la universidadPo

FINANCIACIÓN

Ninguna.

CONFLICTO DE INTERÉS

Ninguno.