

Impacto de la pandemia en el rendimiento académico: análisis comparativo de los resultados ICFES saber 11 pre y post-COVID-19 (2019-2021) en el departamento del Caquetá

Hector Fabian Cerquera Castañeda

Director

Mireya Garcia Garcia

Universidad Nacional Abierta y a Distancia UNAD

Escuela de Ciencias Básicas Tecnologías e Ingenierías ECBTI

Especialización en Ciencia de Datos y Analítica

2025

Dedicatoria

Quiero dedicar este trabajo a mis padres, cuya sabiduría, paciencia y amor incondicional han sido

mi fuente constante de fortaleza. Gracias por ser el ejemplo de esfuerzo y dedicación, por

enseñarme el valor de la educación y por acompañarme en cada paso de este camino.

A mis tutores y asesores, por su guía invaluable y por inspirarme a alcanzar nuevas metas. Su

compromiso con mi formación ha sido un pilar fundamental en mi crecimiento académico,

profesional y personal.

A todos los estudiantes y docentes que, en medio de la pandemia, enfrentaron obstáculos sin

precedentes. Este trabajo está dedicado a ustedes, quienes, con resiliencia y esfuerzo, han

demostrado que el espíritu humano puede superar cualquier adversidad.

Agradecimientos

A Dios, por ser mi guía y darme la fortaleza y sabiduría necesarias para afrontar cada reto con fe y perseverancia.

A mis padres, por su apoyo incondicional, por creer en mis capacidades y por motivarme a dar siempre lo mejor de mí. Sus enseñanzas y ejemplo han sido fundamentales para llegar hasta aquí.

A mis profesores y tutores, quienes, con su paciencia y dedicación, me brindaron el conocimiento y las herramientas necesarias para superar los desafíos académicos. Gracias por su invaluable orientación y por inspirarme a seguir creciendo como profesional.

A la Universidad, por ofrecerme un espacio de aprendizaje y crecimiento. Agradezco especialmente a los miembros del equipo docente que hicieron posible este proyecto, por su tiempo y disposición a lo largo del proceso.

Finalmente, a todos mis amigos y compañeros, por su compañía, palabras de aliento y apoyo en los momentos difíciles. Este logro es también fruto de su motivación y amistad.

Resumen

Este proyecto analiza cómo la pandemia de COVID-19 ha impactado en el rendimiento académico de los estudiantes del departamento del Caquetá, Colombia, utilizando los resultados de las pruebas ICFES Saber 11 calendario A entre los años 2019 y 2021. Debido a la crisis sanitaria, las instituciones educativas tuvieron que cambiar rápidamente de la educación tradicional a la modalidad educativa (virtual-híbrida), presentando numerosos desafíos.

El estudio compara los resultados de las pruebas antes, durante y después de la pandemia para determinar si los estudiantes han mejorado, empeorado o mantenido su desempeño académico. Además, evalúa la eficacia de la enseñanza con modalidad (virtual, híbrida) y su impacto en el resultado de las pruebas.

Para ello, se utilizarán análisis estadísticos que investigarán factores como el acceso a la tecnología, la correlación entre las competencias de matemáticas, inglés y lectura crítica y las condiciones socioeconómicas de los estudiantes. El objetivo es proporcionar información valiosa para el desarrollo de políticas educativas que mejoren la calidad y la equidad de la educación, y preparar al sistema educativo para futuras crisis.

Esperamos que esta investigación no solo contribuya al conocimiento académico, sino que también ofrezca recomendaciones prácticas para mejorar el sistema educativo, asegurando que todos los estudiantes tengan las mismas oportunidades de aprendizaje, sin importar las circunstancias.

Palabras clave: Pandemia de COVID-19, Rendimiento académico, ICFES Saber 11, Educación virtual, Análisis de datos, Calidad educativa, Equidad en la educación.

Abstract

This project analyzes how the COVID-19 pandemic has impacted the academic performance of students in the department of Caquetá, Colombia, using the results of the ICFES Saber 11 tests for calendar A between the years 2019 and 2021. Due to the health crisis, educational institutions had to rapidly shift from traditional education to remote or hybrid modalities, presenting numerous challenges. The study compares test results before, during, and after the pandemic to determine whether students have improved, worsened, or maintained their academic performance. Additionally, it evaluates the effectiveness of virtual and hybrid teaching methods and their impact on test outcomes.

For this purpose, statistical analyses will be conducted to investigate factors such as access to technology, the correlation between math, English, and critical reading skills, and the socioeconomic conditions of the students. The goal is to provide valuable insights for the development of educational policies that improve the quality and equity of education, while preparing the education system for future crises. We hope this research not only contributes to academic knowledge but also offers practical recommendations for improving the educational system, ensuring that all students have equal learning opportunities, regardless of the circumstances.

Keywords: COVID-19 Pandemic, Academic Performance, ICFES Saber 11, Virtual Education, Data Analysis, Educational Quality, Equity in Education.

Tabla de Contenido

Introducción	14
Justificación	15
Objetivos.....	16
Objetivo General	16
Objetivos Específicos.....	16
Marco Conceptual y Teórico	17
Marco Legal y Políticas Educativas en Colombia	19
Metodología	22
Diseño de la Investigación	22
Población de Interés	22
Muestra.....	22
Recolección de Datos	23
Fuentes Primarias	23
Técnicas y Modelos de Análisis.....	23
Análisis Estadístico	23
Pruebas de Hipótesis	23
Modelos de Regresión.....	23
Análisis Diferencial.....	24
Modelos Predictivos.....	24
Regresión Múltiple	24
Descripción de las Variables.....	25
Resultados por Años	26

Análisis del Periodo 2019	26
Pruebas ANOVA.....	29
Prueba de T-test.....	31
Análisis del Periodo 2020	34
Pruebas ANOVA.....	37
Pruebas T-test.....	40
Análisis del Periodo 2021	42
Pruebas ANOVA.....	46
Pruebas T-test.....	48
Comparación de los Resultados por Años 2019-2021	51
¿Afectó la Pandemia en 2020 los Resultados de las Pruebas ICFES Saber 11?.....	53
Comparación de los Resultados en Matemáticas, Inglés y Lectura Crítica	53
Análisis por Estrato Social	55
Estrato 1.....	57
Estrato 2.....	57
Estrato 3.....	57
Estrato 4.....	58
Estrato 5.....	58
Estrato 6.....	58
Análisis Segmentado por Estrato Socioeconómico.....	59
Comparación por Tpo de Colegio (Oficial y no Oficial)	61
Colegios no Oficiales	62
Colegios Oficiales	62

Análisis Segmentado por Área Geográfica (Urbana/Rural).....	63
Análisis del Uso de Herramientas Tecnológicas, Internet- Computador.....	64
Modelos de Regresión.....	67
Regresión Lineal.....	67
Interpretación de los Coeficientes.....	68
Evaluación del Modelo.....	69
Evaluación de Residuos.....	70
Prueba de Homocedasticidad.....	71
Multicolinealidad.....	72
Regresión Logística.....	73
Interpretación Coeficientes.....	74
Evaluación del Modelo.....	75
Evaluación de Residuos.....	75
Prueba de Homocedasticidad.....	77
Multicolinealidad.....	77
Regresión Múltiple.....	78
Interpretación Coeficientes.....	79
Conclusiones.....	82
Cumplimiento de los Objetivos del Estudio.....	84
Objetivo General.....	84
Objetivo Específico 1.....	84
Objetivo Específico 2.....	84
Objetivo Específico 3.....	84

Recomendaciones	87
Referencias Bibliográficas	90

Lista de Tablas

Tabla 1 <i>Descripción Detallada de las Variables</i>	25
--	----

Lista de Figuras

Figura 1 <i>Summary 2019</i>	26
Figura 2 <i>Diagrama de Barras Resultados 2019</i>	29
Figura 3 <i>Prueba ANOVA Puntaje Global- Internet- Computador 2019</i>	29
Figura 4 <i>Gráfico de Cajas Prueba ANOVA Puntaje Global- Internet 2019</i>	30
Figura 5 <i>Gráfico de Cajas Prueba ANOVA Puntaje Global- Computador 2019</i>	30
Figura 6 <i>T-test Resultado Global Acceso a Internet y Computador 2019</i>	32
Figura 7 <i>Comparación T-test 2019</i>	32
Figura 8 <i>Summary 2020</i>	34
Figura 9 <i>Diagrama de Barras Resultados 2020</i>	37
Figura 10 <i>Prueba ANOVA Puntaje Global- Internet- Computador 2020</i>	38
Figura 11 <i>Gráfico de Cajas Prueba ANOVA Puntaje Global- Internet 2020</i>	38
Figura 12 <i>Gráfico de Cajas ANOVA Puntaje Global- Computador 2020</i>	39
Figura 13 <i>T-test Resultado Global Acceso a Internet Y Computador 2020</i>	40
Figura 14 <i>Comparación T-test 2020</i>	41
Figura 15 <i>Summary 2021</i>	43
Figura 16 <i>Diagrama de Barras Resultados 2021</i>	46
Figura 17 <i>Prueba ANOVA Puntaje Global- Internet- Computador 2021</i>	46
Figura 18 <i>Gráfico de Cajas ANOVA Puntaje Global- Internet 2021</i>	47
Figura 19 <i>Gráfico de Cajas ANOVA Puntaje Global- Computador 2021</i>	47
Figura 20 <i>T-test Resultado Global Acceso a Internet y Computador 2021</i>	49
Figura 21 <i>Comparación T-test 2021</i>	49
Figura 22 <i>Resultados Globales por Año</i>	51

Figura 23 <i>Gráfica de Resultados Globales por Año</i>	51
Figura 24 <i>Comparación por Areas</i>	53
Figura 25 <i>Gráfica de Comparación por Areas</i>	54
Figura 26 <i>Comparación por Estrato Social</i>	55
Figura 27 <i>Gráfica de Comparación por Estrato Social</i>	56
Figura 28 <i>Análisis Segmentado por Estrato Socioeconómico</i>	59
Figura 29 <i>Comparación por Tipo de Colegio</i>	61
Figura 30 <i>Gráfica de Comparación por Tipo de Colegio</i>	61
Figura 31 <i>Análisis Área Geográfica (Urbana/Rural)</i>	63
Figura 32 <i>Gráfica de Comparación por Acceso a Internet</i>	64
Figura 33 <i>Gráfica de Comparación por Acceso a Computador</i>	65
Figura 34 <i>Regresión Lineal</i>	67
Figura 35 <i>Gráfico de Regresión Lineal</i>	67
Figura 36 <i>Evaluación Residuos Regresión Lineal</i>	70
Figura 37 <i>Histograma de Evaluación Residuos Regresión Lineal</i>	71
Figura 38 <i>Prueba de Homocedasticidad Regresión Lineal</i>	71
Figura 39 <i>Multicolinealidad Regresión Lineal</i>	72
Figura 40 <i>Regresión Logística</i>	73
Figura 41 <i>Gráfico Regresión Logística</i>	73
Figura 42 <i>Evaluación de Residuos Regresión Logística</i>	76
Figura 43 <i>Histograma de Residuos Regresión Logística</i>	76
Figura 44 <i>Prueba de Homocedasticidad Regresión Logística</i>	77
Figura 45 <i>Multicolinealidad Regresión Logística</i>	77

Figura 46 <i>Regresión Múltiple</i>	78
Figura 47 <i>Gráfico Regresión Múltiple</i>	80

Introducción

La pandemia de COVID-19 ha transformado radicalmente la vida diaria y los sistemas sociales en todo el mundo, siendo la educación uno de los sectores más impactados. En Colombia, y específicamente en el departamento del Caquetá, las instituciones educativas enfrentaron el reto de adaptar sus métodos tradicionales de enseñanza a modalidades virtuales e híbridas en un período de tiempo extremadamente corto. Este cambio abrupto, y motivado por las medidas de aislamiento y distanciamiento social, generó que se plantearan diversos desafíos que afectaron el rendimiento académico de los estudiantes.

El presente estudio tiene como objetivo analizar el impacto de la transición de la educación presencial a la virtual durante la pandemia, utilizando como referencia los resultados de las pruebas ICFES Saber 11 entre los años 2019 y 2021. Estas pruebas han sido a lo largo de la historia un indicador clave del rendimiento académico y son cruciales para el acceso a la educación superior. A través de un análisis comparativo de los resultados pre y post pandemia, se busca determinar si los estudiantes han mejorado, empeorado o mantenido su desempeño académico.

Además, se examinarán factores como el acceso a la tecnología y su influencia en el aprendizaje, así como la efectividad de las diferentes modalidades de enseñanza implementadas durante la crisis sanitaria. Este estudio no solo pretende contribuir al conocimiento académico sobre el impacto de la pandemia en la educación, sino también ofrecer recomendaciones prácticas para mejorar la calidad y equidad educativa en futuras crisis.

Justificación

La pandemia de COVID-19 tocó profundamente cada rincón de nuestra sociedad, afectando especialmente a la educación. Frente a los desafíos impuestos por el aislamiento y el distanciamiento social, el sistema educativo en Colombia, al igual que en muchos otros sectores, tuvieron que transformar rápidamente sus métodos tradicionales de enseñanza presencial en modalidades virtuales. Esta transición, abrupta y sin precedentes, nos lleva a preguntarnos sobre su impacto en el aprendizaje de nuestros estudiantes, especialmente a través de indicadores clave como lo son los resultados de las pruebas ICFES Saber 11, las cuales son fundamentales para acceder a la educación superior.

Académicamente, este proyecto contribuye al conocimiento sobre el impacto de crisis en la educación, con un enfoque específico en una región vulnerable. Al analizar las variaciones en el desempeño antes, durante y después de la pandemia, proporcionará datos valiosos para futuras investigaciones.

Socialmente, la investigación aborda la equidad educativa en un contexto donde las desigualdades se han profundizado debido a la pandemia. Al identificar los factores que han afectado el rendimiento permitirá desarrollar políticas públicas que mejoren la resiliencia del sistema educativo y garantice una educación de calidad, incluso en situaciones adversas.

Objetivos

Objetivo General

Evaluar el impacto de la transición de la educación tradicional a la modalidad educativa (virtual - híbrida), a raíz de la pandemia de COVID-19, en el rendimiento de los estudiantes en las pruebas ICFES Saber 11 Calendario A en los años 2019 a 2021 en el departamento del Caquetá.

Objetivos Específicos

Comparar los resultados de las pruebas ICFES Saber 11 calendario A entre los años 2019 y 2021 para identificar variaciones significativas en el rendimiento académico de los estudiantes del departamento del Caquetá, con un enfoque en las competencias de matemáticas, inglés y lectura crítica, antes, durante y después de la pandemia de COVID-19.

Evaluar la eficacia de las modalidades de enseñanza virtual implementadas durante la pandemia mediante la correlación entre el tipo de modalidad educativa (virtual, híbrida, presencial) y los resultados obtenidos en las pruebas ICFES Saber 11 en el período 2019-2021.

Analizar cómo el uso de herramientas tecnológicas influye en los resultados de las pruebas ICFES Saber 11, midiendo la cantidad de estudiantes que disponen de dichos recursos durante el período 2019-2021.

Marco Conceptual y Teórico

La aparición de la pandemia de COVID-19 no solo representó un desafío sanitario global sin precedentes, sino que también actuó como un agente de cambio para una transformación radical en el ámbito educativo en Colombia. Esta emergencia obligó a escuelas, colegios y universidades a reemplazar las aulas físicas por espacios virtuales, una transición que, aunque necesaria, presentó complicaciones, desafíos y errores. En este sentido, Ávila Palacio, Ospino Gutiérrez y Páez Reales (2021) destacan la urgencia de evaluar cómo estos cambios impactaron en los resultados de las pruebas ICFES Saber 11, resaltando la variabilidad en el acceso y la calidad de la educación virtual implementada durante la emergencia.

El cambio forzado y repentino de la educación tradicional a la educación virtual reveló e intensificó las desigualdades en el acceso a la tecnología y la capacitación en habilidades digitales, aspectos cruciales para una educación de calidad. Según Inciarte González y colaboradores (2020), la pandemia no solo transformó los métodos de enseñanza, sino que también puso en evidencia significativas brechas en el acceso a recursos tecnológicos, lo que implicó una afectación profunda en la motivación y el bienestar emocional de los estudiantes.

Este escenario puso en evidencia la fragilidad de nuestro sistema educativo frente a emergencias lo que nos lleva a la necesidad de mejorar las políticas que garanticen una educación equitativa y de calidad para todos, independientemente de su situación socioeconómica o geográfica.

Los resultados de las pruebas ICFES Saber 11, las cuales entre muchas otras áreas como lo son matemáticas y ciencias, se han convertido en indicadores clave del rendimiento y la calidad educativa durante la pandemia. Los estudios de Demarchi Sánchez (2023) ilustran cómo las condiciones del hogar y la disponibilidad de recursos de aprendizaje han jugado un papel

importante en el rendimiento académico de los estudiantes durante la pandemia. Este enfoque destaca cómo el ambiente doméstico puede influir significativamente en la capacidad de los estudiantes para adaptarse a los nuevos métodos de aprendizaje, revelando un panorama complejo de los desafíos y necesidades emergentes.

La literatura que existe sobre la educación durante crisis similares nos proporciona una perspectiva a través de la cual podemos examinar y estudiar los desafíos de la actualidad. Ávila Palacio, Ospino Gutiérrez y Páez Reales (2021) investigaron los efectos de los cambios en la educación durante la pandemia en los resultados de las pruebas ICFES en Colombia, encontrando que las diferencias en el rendimiento académico reflejaban no sólo diferencias en la calidad de la enseñanza virtual, sino también en el acceso a los recursos esenciales como internet y dispositivos tecnológicos adecuados. Estos hallazgos son fundamentales para entender y abordar las variaciones de los resultados académicos en nuestro estudio en el departamento del Caquetá.

En una exploración profunda de cómo las decisiones políticas han influido en las realidades educativas durante la pandemia, investigadores como Ávila Palacio, Ospino Gutiérrez y Páez Reales (2021) se han adentrado en el análisis de las políticas educativas implementadas para enfrentar los retos de la emergencia. Este estudio enfatiza en entender las estrategias que el gobierno y las instituciones educativas implementaron para superar los problemas que acarrea una enseñanza virtual. A través de sus hallazgos, se muestra cómo estas medidas han buscado no solo mantener sino también mejorar el rendimiento académico y garantizar una mayor equidad en el acceso a la educación. Las experiencias de los estudiantes y docentes, marcadas por las políticas durante la crisis, revelan un relato de la resiliencia y adaptación. Estas historias

personales, llenas de desafíos y esperanzas, dan vida y forma a la educación en estos tiempos difíciles.

Demarchi Sánchez (2023) analizó el impacto del entorno doméstico en el aprendizaje de los estudiantes durante la pandemia. Este estudio reveló que algunos factores como el espacio físico adecuado para estudiar y el apoyo familiar eran muy importantes para mantener el rendimiento académico, destacando las desigualdades en las experiencias de aprendizaje entre estudiantes de diferentes estratos socioeconómicos.

La historia ha demostrado que la capacidad del sistema educativo para adaptarse y responder eficazmente a crisis inesperadas es crucial para mantener la continuidad y la calidad del aprendizaje. Por ejemplo, estudios que hablan sobre las respuestas educativas a desastres naturales y crisis económicas resaltan que la resiliencia y flexibilidad del sistema educativo son fundamentales para reducir y mitigar los impactos negativos en el rendimiento académico de los estudiantes. Según los hallazgos de investigaciones realizadas por autores como Inciarte González, Paredes-Chacín y Zambrano Villada (2020), intervenciones educativas que sean bien planificadas y coordinadas han demostrado ser cruciales y de gran importancia para reducir las interrupciones en el aprendizaje y facilitar una recuperación académica rápida y eficaz. Este antecedente de conocimiento nos ayuda e invita a considerar estrategias proactivas y bien estructuradas que nos ayuden a enfrentar las actuales y futuras crisis educativas, asegurando que los sistemas educativos no solo sobrevivan, sino que también se fortalezcan ante adversidades.

Marco Legal y Políticas Educativas en Colombia

Durante la pandemia de COVID-19, Colombia implementó y adoptó una serie de políticas educativas para hacer que su sistema educativo resistiera las restricciones sanitarias producidas por la crisis. Estas políticas estuvieron enfocadas en hacer más fácil la transición

hacia la educación virtual y apoyar tanto a estudiantes como a docentes en este desafiante proceso. Se desarrollaron estrategias específicas para mejorar el acceso a la educación a distancia, proporcionando recursos digitales y ofreciendo capacitación en herramientas virtuales a los docentes (Banco Mundial, 2020).

Además, se implementaron medidas para asegurar que las poblaciones más vulnerables no quedarán desamparadas, con el fin de evitar que la brecha educativa fuera más grande en estos grupos durante la emergencia sanitaria. Este enfoque incluyó políticas para reforzar el soporte socioemocional a los estudiantes, reconociendo la importancia del bienestar psicológico en el rendimiento académico (Banco Mundial, 2020).

En la transición de la educación tradicional a la educación virtual durante la pandemia, se ha observado una disminución amplia y general en el rendimiento académico, especialmente en áreas clave como las matemáticas y las ciencias. Este comportamiento puede atribuirse a varios factores incluyendo el estrés y la adaptación a nuevas formas de aprendizaje. Estudios recientes sugieren que los desafíos asociados con la educación virtual, como la sobrecarga de tareas y la falta de interacción directa con la sociedad, han contribuido a esta disminución en el rendimiento académico (Mendieta, Cairo & Cairo, 2020). Por lo tanto, es fundamental investigar más a fondo estas causas y diseñar nuevas estrategias que sean efectivas para mejorar el aprendizaje virtual, especialmente en asignaturas que requieren un alto grado de interacción y apoyo conceptual.

En un esfuerzo por entender cómo la pandemia por COVID-19 ha cambiado la educación a nivel mundial, hace que la UNESCO brinde un informe detallando los diferentes enfoques que países alrededor del mundo han adoptado frente a los desafíos educativos presentados. Este análisis comparativo es de vital importancia para apreciar las distintas estrategias y sus variados éxitos y limitaciones en el mantenimiento de la continuidad educativa (UNESCO, 2021).

Chan et al. (2021) exploran en gran profundidad las experiencias de los estudiantes internacionales durante la pandemia, resaltando cómo los cambios forzosos en la educación y la movilidad estudiantil han afectado no solo su aprendizaje sino también su bienestar emocional y social. Este estudio nos proporciona un valioso contexto para entender las consecuencias humanas de la transición abrupta a la educación en línea (Chan, Bista & Allen, 2021).

Metodología

Este proyecto tiene como objetivo comprender cómo la pandemia de COVID-19 en el año 2020 afectó el rendimiento académico de los estudiantes en el departamento del Caquetá, utilizando los resultados de las pruebas ICFES Saber 11 del calendario A.

Diseño de la Investigación

Tipo de estudio: Cuantitativo, comparativo y correlacional. Se llevará a cabo un análisis de los resultados de las pruebas ICFES Saber 11 calendario A entre los años 2019 y 2021 en el departamento del Caquetá.

Enfoque: La investigación seguirá un enfoque cuantitativo, que se centrará en el análisis de los resultados de las pruebas ICFES Saber 11 y su relación con variables como el acceso a herramientas tecnológicas. Se utilizarán modelos estadísticos para identificar patrones y correlaciones significativas en los resultados académicos antes, durante y después de la pandemia de COVID-19.

Población de Interés

La población de estudio estará compuesta por los estudiantes que presentaron las pruebas ICFES Saber 11 entre los años 2019 y 2021 en el calendario A. El análisis se realizará exclusivamente con los datos cuantitativos proporcionados por ICFES, lo que permitirá evaluar cómo los cambios educativos y el acceso a la tecnología influyeron en el rendimiento académico durante este período de transición.

Muestra

Se seleccionará una muestra representativa de los resultados de las pruebas ICFES Saber 11 entre los años 2019 y 2021, correspondiente a estudiantes del calendario A en el departamento del Caquetá. La muestra incluirá datos de estudiantes de diferentes municipios y tipos de

instituciones educativas, tanto públicas como privadas, para así garantizar una perspectiva amplia sobre el impacto educativo en la región.

Recolección de Datos

Fuentes Primarias

Resultados de las pruebas ICFES: Se recopilaron y analizaron los resultados de las pruebas ICFES Saber 11 calendario A de los años 2019 a 2021.

Técnicas y Modelos de Análisis

Análisis Estadístico

Análisis Descriptivo: Se emplearán medidas de tendencia central (media, mediana) y de dispersión (desviación estándar) para describir los resultados de las pruebas ICFES antes, durante y después de la pandemia. Este análisis permitirá obtener una visión general de las variaciones en los puntajes durante estos períodos.

Pruebas de Hipótesis

Se aplicarán pruebas ANOVA para determinar si el acceso a herramientas tecnológicas tuvo un impacto significativo en los resultados de las pruebas. Estas pruebas permitirán evaluar si existen diferencias significativas en el rendimiento académico según el acceso a tecnología durante los años 2019-2021.

Modelos de Regresión

Se utilizarán modelos de regresión para explorar la relación entre variables independientes (modalidad educativa, acceso a tecnología) y el rendimiento académico. Los modelos de regresión lineal se enfocarán en resultados continuos, mientras que la regresión logística analizará probabilidades de determinados rendimientos académicos.

Análisis Diferencial

Se realizará un análisis diferencial para evaluar la correlación entre el desempeño en competencias como matemáticas, inglés y lectura crítica en las pruebas ICFES. Este análisis ayudará a identificar cómo el rendimiento en una competencia puede influir en las demás, brindando una visión más integral del rendimiento académico global durante el período de estudio.

Modelos Predictivos***Regresión Múltiple***

Se implementará para evaluar cómo factores individuales (como el uso de la tecnología) y contextuales (como el tipo de institución) interactúan para influir en los resultados académicos de los estudiantes.

Descripción de las Variables

Tabla 1

Descripción Detallada de las Variables

Variables	Descripción	Tipo de Dato
ESTU_GENERO	1 = Masculino, 0 = Femenino	Categorico (binario)
PERIODO	Año en que se presentó	Numérico
ESTU_TIENEETNIA	1 = Sí, 0 = No	Categorico (binario)
ESTU_MCPIO_RESIDE	Nombre del municipio de residencia	Categorico (texto)
FAMI ESTRATOVIVIENDA	Estrato socioeconómico	Numérico
FAMI_TIENEINTERNET	1 = Sí, 0 = No	Categorico (binario)
FAMI_TIENECOMPUTADOR	1 = Sí, 0 = No	Categorico (binario)
COLE_NATURALEZA	1 = Oficial, 0 = No oficial	Categorico (binario)
COLE_BILI0GUE	1 = Sí, 0 = No	Categorico (binario)
COLE_AREA_UBICACION	1 = Urbano, 0 = Rural	Categorico (binario)
PUNT_LLECTURA_CRITICA	Puntaje obtenido en lectura crítica	Numérico
PUNT_MATEMATICAS	Puntaje obtenido en matemáticas	Numérico
PUNT_C_NATURALES	Puntaje obtenido en ciencias naturales	Numérico
PUNT_SOCIALES_CIUADANAS	Puntaje obtenido en sociales y ciudadanía	Numérico
PUNT_INGLES	Puntaje obtenido en inglés	Numérico
DESEMP_INGLES	Nivel de inglés alcanzado	Categorico (texto)
PUNT_GLOBAL	Puntaje global	Numérico

Nota. La tabla presenta una descripción de las variables utilizadas en el análisis de los datos de las pruebas ICFES Saber 11. Se incluyen tanto las variables categóricas como las numéricas, indicando su significado y las posibles categorías o rangos de valores.

Resultados por Años

Análisis del Periodo 2019

Se realizó el análisis de los datos correspondientes al año 2019, suministrados por el ICFES, los cuales contienen un total de 4633 registros. De estos, 675 registros presentan datos faltantes en las variables FAMI_TIENEINTERNET y FAMI_TIENECOMPUTADOR, lo que representa el 14.5% del total. Debido a la proporción significativa de datos faltantes, se decidió eliminar estos registros para así asegurar la integridad del análisis.

Figura 1

Summary 2019

```

> summary(DATA2019_CLEAN)
 ESTU_GENERO      PERIODO      ESTU_TIENEETNIA      ESTU_MCPIO_RESIDE      FAMI_ESTRATOVIVIENDA      FAMI_TIENEINTERNET
 Min.   :0.0000   Min.   :20194   Min.   :0.000000   Length:3958   Min.   :1.000   0:2127
 1st Qu.:0.0000   1st Qu.:20194   1st Qu.:0.000000   Class :character   1st Qu.:1.000   1:1831
 Median :0.0000   Median :20194   Median :0.000000   Mode  :character   Median :1.000
 Mean   :0.4411   Mean   :20194   Mean   :0.01541
 3rd Qu.:1.0000   3rd Qu.:20194   3rd Qu.:0.000000
 Max.   :1.0000   Max.   :20194   Max.   :1.000000
 FAMI_TIENECOMPUTADOR  COLE_NATURALEZA  COLE_BILIOGUE  COLE_AREA_UBICACION  PUNT_LECTURA_CRITICA  PUNT_MATEMATICAS
 0:2260                Min.   :0.0000   Min.   :0.000000   Min.   :0.0000   Min.   :23.0   Min.   :17.00
 1:1698                1st Qu.:1.0000   1st Qu.:0.000000   1st Qu.:1.0000   1st Qu.:44.0   1st Qu.:42.00
                Median :1.0000   Median :0.000000   Median :1.0000   Median :51.0   Median :49.00
                Mean   :0.8274   Mean   :0.009853   Mean   :0.8555   Mean   :50.4   Mean   :49.34
                3rd Qu.:1.0000   3rd Qu.:0.000000   3rd Qu.:1.0000   3rd Qu.:57.0   3rd Qu.:57.00
                Max.   :1.0000   Max.   :1.000000   Max.   :1.0000   Max.   :80.0   Max.   :83.00
 PUNT_C_NATURALES      PUNT_SOCIALES_CIUDADANAS  PUNT_INGLES      DESEMP_INGLES      PUNT_GLOBAL      puntaje_alto
 Min.   :20.00   Min.   :19.00   Min.   : 0.00   Length:3958   Min.   :134.0   Min.   :0.0000
 1st Qu.:39.00   1st Qu.:36.00   1st Qu.: 37.00   Class :character   1st Qu.:203.0   1st Qu.:0.0000
 Median :45.00   Median :44.00   Median : 44.00   Mode  :character   Median :234.0   Median :0.0000
 Mean   :46.16   Mean   :44.58   Mean   : 45.26
 3rd Qu.:52.00   3rd Qu.:52.00   3rd Qu.: 53.00
 Max.   :77.00   Max.   :80.00   Max.   :100.00
  
```

Nota. Esta figura muestra el resultado por consola del Summary

El análisis descriptivo de los resultados para el año 2019 se realizó utilizando medidas de tendencia central (media y mediana) y medidas de dispersión (desviación estándar) para así describir los puntajes obtenidos por los estudiantes en las pruebas ICFES para este año.

Puntaje Global:

- Media: 237.2
- Mediana: 234
- Desviación estándar: 42.82

Se encontró que el puntaje global promedio es de 237.2, y la mediana es de 234, lo que indica que existe una distribución simétrica en torno a estos valores. La desviación estándar de 42.82 indica una variabilidad moderada en los puntajes, lo que representa que la mayoría de los estudiantes se agrupan cerca de la media, con algunos casos dispersos.

Matemáticas:

- Media: 49.34
- Mediana: 49
- Desviación estándar: 10.64

Los puntajes en matemáticas presentan una distribución centrada alrededor de 49, con una desviación estándar moderada, lo que nos indica que la mayoría de los estudiantes puntúan dentro de un rango similar.

Lectura Crítica:

- Media: 50.4
- Mediana: 51
- Desviación estándar: 9.40

La media y la mediana en lectura crítica son cercanas, lo que nos sugiere una distribución simétrica. La desviación estándar es de 9.40, lo cual refleja una variabilidad moderada en los puntajes.

Ciencias Naturales:

- Media: 46.16
- Mediana: 45
- Desviación estándar: 9.21

Los puntajes en ciencias naturales son ligeramente más bajos que en matemáticas y lectura crítica. La desviación estándar de 9.21 indica una distribución moderada de los puntajes alrededor de la media.

Sociales y Ciudadanas:

- Media: 44.58
- Mediana: 44
- Desviación estándar: 11.03

Los puntajes en sociales y ciudadanas son más bajos en promedio, con una desviación estándar de 11.03, lo que indica una mayor variabilidad en los puntajes de los estudiantes.

Inglés:

- Media: 45.26
- Mediana: 44
- Desviación estándar: 9.28

Los puntajes en inglés son los más bajos de todas las áreas evaluadas. La desviación estándar de 9.28 sugiere que los puntajes están moderadamente dispersos alrededor de la media, con una distribución simétrica.

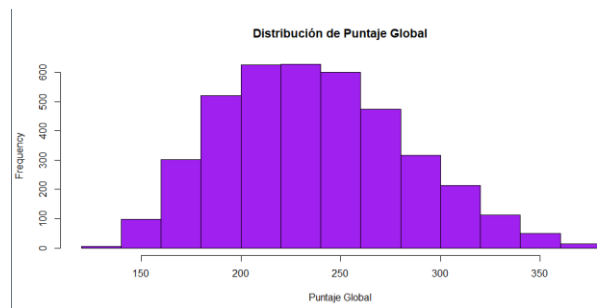
Conclusiones: Se concluye que, en el análisis del año 2019, se observan patrones de rendimiento académico que podrían reflejar rezagos asociados a procesos previos de adaptación educativa, posiblemente influenciados por la transición hacia modalidades no presenciales.

Los puntajes globales y por área muestran una distribución razonablemente simétrica, con diferencias notables en la dispersión entre las áreas. Lectura crítica y ciencias naturales presentan menor variabilidad, lo que nos indica que los estudiantes tienen tendencia a rendir de manera más homogénea en estas áreas. Por otro lado, sociales y ciudadanas presentan una mayor

variabilidad, lo que indica una dispersión más amplia entre los puntajes de los estudiantes. El rendimiento más bajo es en el área de inglés, lo que nos indica que es necesario tomar acciones para mitigar este déficit.

Figura 2

Diagrama de Barras Resultados 2019



Nota. Este diagrama de barras representa las métricas de resultado de las pruebas ICFES Saber 11 del año 2019.

Pruebas ANOVA

Creamos esta prueba para validar la relación entre el puntaje global y el acceso a internet y computador 2019.

Figura 3

Prueba ANOVA Puntaje Global- Internet- Computador 2019

```

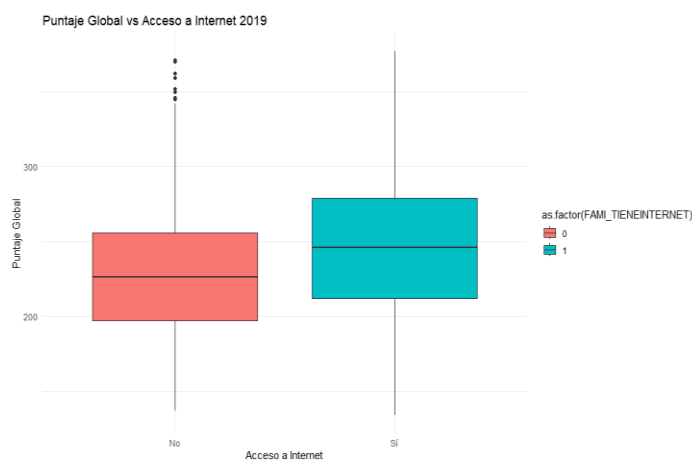
> summary(anova_tiencomputador)
          Df Sum Sq Mean Sq F value Pr(>F)
FAMI_TIENINTERNET  1  333249  333249  191.9 <2e-16 ***
Residuals        3399 5901442   1736
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> # Realiza el ANOVA por el acceso a computador
> anova_computador <- aov(PUNT_GLOBAL ~ FAMI_TIENCOMPUTADOR, data = data2020_clean)
>
> # Resumen del ANOVA
> summary(anova_computador)
          Df Sum Sq Mean Sq F value Pr(>F)
FAMI_TIENCOMPUTADOR  1  372347  372347  215.9 <2e-16 ***
Residuals        3399 5862343   1725
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> # ANOVA con acceso a Internet y computadores como factores
> anova_internet_computador <- aov(PUNT_GLOBAL ~ FAMI_TIENINTERNET * FAMI_TIENCOMPUTADOR, data = data2020_clean)
>
> # Resumen del ANOVA
> summary(anova_internet_computador)
          Df Sum Sq Mean Sq F value Pr(>F)
FAMI_TIENINTERNET  1  333249  333249  196.724 <2e-16 ***
FAMI_TIENCOMPUTADOR  1  144027  144027  85.022 <2e-16 ***
FAMI_TIENINTERNET:FAMI_TIENCOMPUTADOR  1  2918  2918  1.722 0.189
Residuals        3397 5754437   1694
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>

```

Nota. Esta figura muestra el resultado por consola de la prueba de ANOVA para el año 2019.

Figura 4

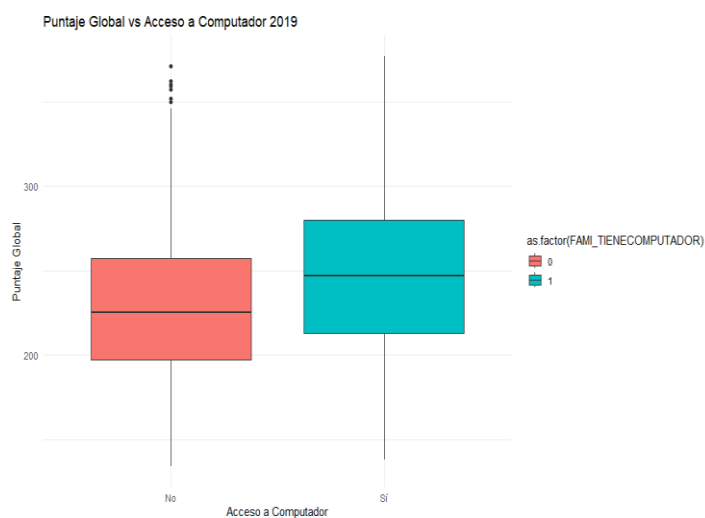
Gráfico de Cajas Prueba ANOVA Puntaje Global- Internet 2019



Nota. Este gráfico muestra la representación gráfica de la prueba de ANOVA puntaje global-internet para el año 2019.

Figura 5

Gráfico de Cajas Prueba ANOVA Puntaje Global- Computador 2019



Nota. Este gráfico muestra la representación gráfica de la prueba de ANOVA puntaje global-computador para el año 2019.

El análisis ANOVA reveló los siguientes resultados: El análisis ANOVA reveló resultados significativos en cuanto al impacto del acceso a Internet y computadores sobre los puntajes globales obtenidos por los estudiantes en las pruebas ICFES del año 2019 esto como se evidencia en las gráficas proporcionadas:

Acceso a Internet: El *Valor f* = 175.2 , junto con un *Valor p* < $2e - 16$, confirma que el acceso a Internet tiene un efecto significativo en los puntajes globales de los estudiantes. Este resultado indica que los estudiantes con acceso a Internet tienden a obtener puntajes más altos que aquellos que no tienen acceso.

Acceso a Computador: El *Valor f* = 166.1, acompañado de un *Valor p* < $2e - 16$, nos indica que el acceso a un computador también tiene un impacto muy significativo sobre los puntajes globales. Los estudiantes que cuentan con acceso a un computador en sus hogares obtienen puntajes más altos en comparación con los estudiantes que no tienen acceso a estas herramientas.

Interacción entre Acceso a Internet y Computadores: El *Valor f* = 12.53 con un *Valor p* = 0.000405, nos indica que existe una interacción significativa entre el acceso a Internet y el acceso a un computador. Esto nos indica que el efecto de tener acceso a ambas tecnologías es mayor que la suma de los efectos individuales de cada una. Los estudiantes que tienen acceso tanto a Internet como a un computador tienden a obtener mejores resultados que aquellos que tienen acceso solo a una de estas herramientas.

Prueba de T-test

Realizamos un T-test para comparar los puntajes globales basados en el acceso a internet y a computador.

Figura 6

T-test Resultado Global Acceso a Internet y Computador 2019

```

> # Mostrar los resultados del t-test
> print(t_test_internet)

welch Two Sample t-test

data: PUNT_GLOBAL by FAMI_TIENEINTERNET
t = -13.133, df = 3720.4, p-value < 2.2e-16
alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
95 percent confidence interval:
 -21.44077 -15.87073
sample estimates:
mean in group 0 mean in group 1
 228.5496      247.2054

>
> # Realizar un t-test para comparar los puntajes globales basados en el acceso a computador
> t_test_computador <- t.test(PUNT_GLOBAL ~ FAMI_TIENECOMPUTADOR, data = Data2019_clean)
>
> # Mostrar los resultados del t-test
> print(t_test_computador)

welch Two Sample t-test

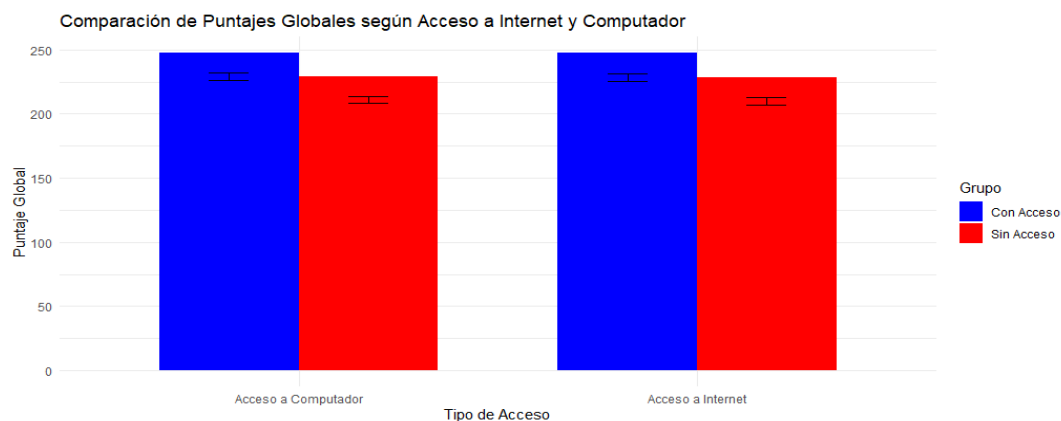
data: PUNT_GLOBAL by FAMI_TIENECOMPUTADOR
t = -12.692, df = 3428.2, p-value < 2.2e-16
alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
95 percent confidence interval:
 -21.15291 -15.49202
sample estimates:
mean in group 0 mean in group 1
 229.3195      247.6419

```

Nota. Esta figura muestra el resultado por consola del T-test para el año 2019.

Figura 7

Comparación T-test 2019



Nota. Este gráfico compara los puntajes según el acceso a internet y computador.

Acceso a Internet: El análisis de la prueba t muestra que el *Valor t* = -13.133 , lo que representa una diferencia bastante considerable en los puntajes globales entre los estudiantes con

y sin acceso a Internet. Con un *Valor p* $< 2.2e - 16$, podemos concluir que esta diferencia es altamente significativa. Los grados de libertad para esta prueba son 3720.4. El intervalo de confianza del 95% para la diferencia de medias oscila entre -21.44 y -15.87, lo que indica que, en promedio, los estudiantes sin acceso a Internet obtienen entre 15.87 y 21.44 puntos menos que aquellos con acceso. Las medias de los puntajes globales de las pruebas muestran que los estudiantes sin acceso tienen un promedio de 228.55 puntos, mientras que aquellos con acceso a Internet tienen una media de 247.21, lo que representa una diferencia de 18.66 puntos.

Acceso a Computador: La prueba t para esta variable arrojó un *Valor de t* = -12.693 , lo que indica una diferencia muy significativa en los puntajes globales entre los estudiantes con y sin acceso a un computador. El *valor de p* $< 2.2e16$, refuerza la significancia de esta diferencia. Los grados de libertad son 3428.2, y el intervalo de confianza del 95% para la diferencia de medias se sitúa entre -21.15 y -15.49. Esto indica que los estudiantes sin acceso a un computador tienen entre 15.49 y 21.15 puntos menos en sus puntajes globales que aquellos estudiantes que sí cuentan con esta herramienta. En términos de medias, los estudiantes sin acceso a un computador tienen un promedio de 229.32 puntos, mientras que los que disponen de uno alcanzan una media de 247.64, lo que refleja una diferencia de 18.32 puntos.

Conclusiones: Se concluye que, en el análisis correspondiente al año 2019, se identifican diferencias significativas en el rendimiento académico de los estudiantes, las cuales parecen estar asociadas al modelo de educación remota adoptado durante la pandemia. Las pruebas estadísticas respaldan esta hipótesis, al mostrar patrones consistentes que reflejan el impacto de las condiciones educativas excepcionales de ese periodo.

Los dos análisis confirman que el acceso a Internet y a un computador tiene un efecto significativo y positivo sobre los puntajes globales de los estudiantes. Las diferencias en los

puntajes son consistentes, con aquellos que tienen acceso a estas tecnologías obteniendo alrededor de 18 a 20 puntos más que aquellos sin acceso. Estos resultados confirman la importancia del acceso a herramientas tecnológicas en el rendimiento académico y podrían respaldar la implementación de políticas que promuevan una mayor disponibilidad de recursos tecnológicos para mejorar los resultados educativos.

Análisis del Periodo 2020

Se llevó a cabo el análisis de los datos del año 2020 proporcionados por el ICFES, con un total de 3868 registros. De estos, 467 presentaban valores ausentes en las variables FAMI_TIENEINTERNET y FAMI_TIENECOMPUTADOR, lo que equivale al 12% del conjunto de datos. Debido a la cantidad considerable de datos faltantes, se optó por excluir estos registros para garantizar la precisión del análisis.

Figura 8

Summary 2020

```
> summary(Data2020_clean)
 ESTU_GENERO      PERIODO      ESTU_TIENEETNIA      ESTU_MCPPIO_RESIDE      FAMI ESTRATOVIVIENDA
 Min.   :0.0000   Min.   :20204   Min.   :0.000000   Length:3401   Min.   :1.000
 1st Qu.:0.0000   1st Qu.:20204   1st Qu.:0.000000   Class :character   1st Qu.:1.000
 Median :0.0000   Median :20204   Median :0.000000   Mode  :character   Median :1.000
 Mean   :0.4408   Mean   :20204   Mean   :0.005587             Mean   :1.675
 3rd Qu.:1.0000   3rd Qu.:20204   3rd Qu.:0.000000             3rd Qu.:2.000
 Max.   :1.0000   Max.   :20204   Max.   :1.000000             Max.   :6.000

 FAMI_TIENEINTERNET FAMI_TIENECOMPUTADOR COLE_NATURALEZA COLE_BILIÓGUE COLE_AREA_UBICACION
 Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.000000   Min.   :0.0000
 1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:1.0000   1st Qu.:0.000000   1st Qu.:1.0000
 Median :1.0000   Median :1.0000   Median :1.0000   Median :0.000000   Median :1.0000
 Mean   :0.6095   Mean   :0.5072   Mean   :0.8571   Mean   :0.007645   Mean   :0.8565
 3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:0.000000   3rd Qu.:1.0000
 Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :1.000000   Max.   :1.0000

 PUNT_LECTURA_CRITICA PUNT_MATEMATICAS PUNT_C_NATURALES PUNT_SOCIALES_CIUDADANAS PUNT_INGLES
 Min.   : 24.00   Min.   : 19.00   Min.   :25.00   Min.   : 0.00   Min.   : 0.00
 1st Qu.: 43.00   1st Qu.: 42.00   1st Qu.:39.00   1st Qu.: 38.00   1st Qu.: 38.00
 Median : 50.00   Median : 49.00   Median :46.00   Median : 45.00   Median : 44.00
 Mean   : 50.27   Mean   : 49.38   Mean   :46.42   Mean   : 46.32   Mean   : 44.79
 3rd Qu.: 57.00   3rd Qu.: 57.00   3rd Qu.:53.00   3rd Qu.: 54.00   3rd Qu.: 49.00
 Max.   :100.00   Max.   :100.00   Max.   :76.00   Max.   :100.00   Max.   :100.00

 DESEMP_INGLES      PUNT_GLOBAL
 Length:3401       Min.   :114.0
 Class :character   1st Qu.:207.0
 Mode  :character   Median :235.0
                   Mean   :239.2
                   3rd Qu.:268.0
                   Max.   :393.0
```

Nota. Esta figura muestra el resultado por consola del Summary.

Puntaje Global:

- Media: 239.22
- Mediana: 235
- Desviación estándar: 42.82

El puntaje global promedio es de 239.22, y la mediana es de 235, lo que indica una distribución simétrica en torno a estos valores. La desviación estándar de 42.82 indica una variabilidad moderada, lo que sugiere que la mayoría de los estudiantes se agrupan en torno a la media, con algunos casos más dispersos.

Matemáticas:

- Media: 49.38
- Mediana: 49
- Desviación estándar: 10.64

Los puntajes en matemáticas tienen una distribución centrada alrededor de 49, con una desviación estándar moderada, lo que indica que la mayoría de los estudiantes puntúan dentro de un rango relativamente estrecho.

Lectura Crítica:

- Media: 50.27
- Mediana: 50
- Desviación estándar: 9.40

La media y la mediana en lectura crítica son casi idénticas, lo que sugiere una distribución simétrica. La variabilidad en los puntajes es moderada, reflejada en una desviación estándar de 9.40.

Ciencias Naturales:

- Media: 46.42
- Mediana: 46
- Desviación estándar: 9.21

Los puntajes en ciencias naturales son ligeramente más bajos que en otras áreas, con una desviación estándar de 9.21, lo que indica que los puntajes están bien distribuidos en torno a la media.

Sociales y Ciudadanas:

- Media: 46.32
- Mediana: 45
- Desviación estándar: 11.03

Los puntajes en esta área son más bajos en comparación con otras, con una mayor variabilidad en los puntajes, lo que podría indicar diferencias más marcadas entre los estudiantes.

Inglés:

- Media: 44.79
- Mediana: 44
- Desviación estándar: 9.28

Los puntajes en inglés son los más bajos en promedio, con una desviación estándar de 9.28, lo que sugiere una variabilidad moderada y una distribución simétrica.

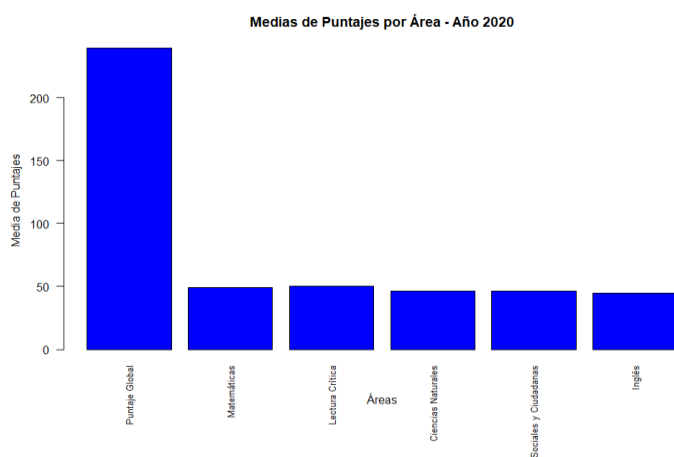
Conclusiones: Se concluye que, en el análisis del año 2020, se identifican diferencias internas en el rendimiento académico entre áreas, posiblemente influenciadas por los efectos acumulados del modelo de educación remota implementado durante la pandemia. Los puntajes globales en este año por área muestran una distribución razonablemente simétrica, con diferencias algo notables en la dispersión entre las áreas. Áreas como lectura crítica y ciencias

naturales presentan menor dispersión, lo que denota que los estudiantes tienden a rendir de manera más homogénea. Sin embargo, sociales y ciudadanas presentan la mayor variación, lo que podría advertir diferencias más marcadas entre los estudiantes en esta área.

El rendimiento más bajo en inglés podría ser un punto de interés para investigaciones futuras, especialmente si esta tendencia se mantiene en los datos de otros años.

Figura 9

Diagrama de Barras Resultados 2020



Nota. Este diagrama de barras representa las métricas de resultado de las pruebas ICFES Saber 11 del año 2020.

Pruebas ANOVA

Creamos esta prueba para validar la relación entre el puntaje global y el acceso a internet y computador en el año 2020.

Figura 10

Prueba ANOVA Puntaje Global- Internet- Computador 2020

```

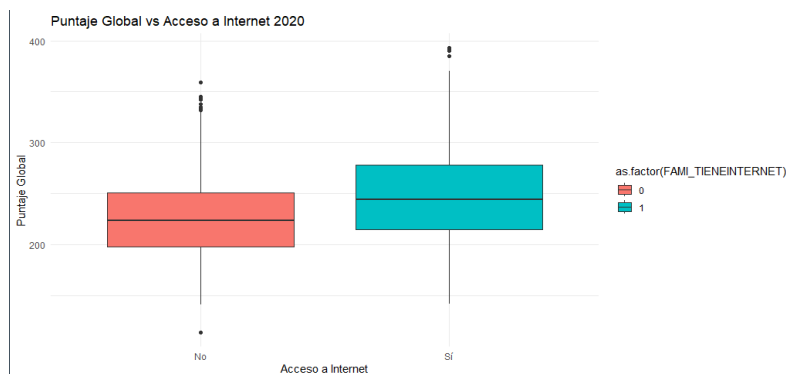
> summary(anova_internet)
              Df Sum Sq Mean Sq F value Pr(>F)
FAMI_TIENEINTERNET  1  333249  333249  191.9 <2e-16 ***
Residuals        3399 5901442   1736
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> # Realizar el ANOVA para el acceso a computador
> anova_computador <- aov(PUNT_GLOBAL ~ FAMI_TIENECOMPUTADOR, data = data2020_clean)
>
> # Resumen del ANOVA
> summary(anova_computador)
              Df Sum Sq Mean Sq F value Pr(>F)
FAMI_TIENECOMPUTADOR  1  372347  372347  215.9 <2e-16 ***
Residuals          3399 5862343   1725
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> # ANOVA con acceso a Internet y computadores como Factores
> anova_internet_computador <- aov(PUNT_GLOBAL ~ FAMI_TIENEINTERNET * FAMI_TIENECOMPUTADOR, data = data2020_clean)
>
> # Resumen del ANOVA
> summary(anova_internet_computador)
              Df Sum Sq Mean Sq F value Pr(>F)
FAMI_TIENEINTERNET  1  333249  333249  196.724 <2e-16 ***
FAMI_TIENECOMPUTADOR  1  144027  144027  85.022 <2e-16 ***
FAMI_TIENEINTERNET:FAMI_TIENECOMPUTADOR  1  2918  2918  1.722 0.189
Residuals          3397 5754497   1694
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>

```

Nota. Esta figura muestra el resultado por consola de la prueba de ANOVA para el año 2020.

Figura 11

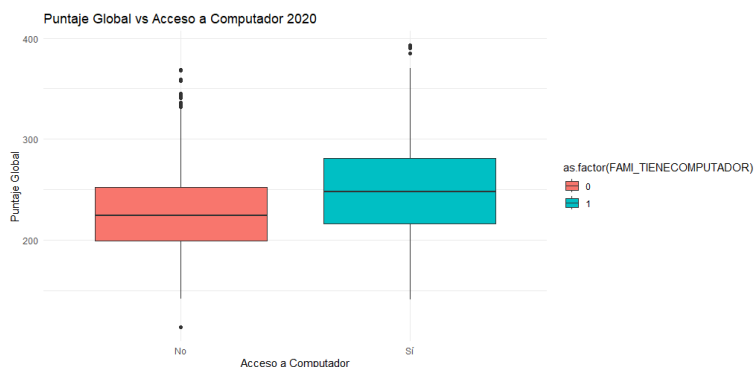
Gráfico de Cajas Prueba ANOVA Puntaje Global- Internet 2020



Nota. Este gráfico muestra la representación gráfica de la prueba de ANOVA puntaje global- internet para el año 2020.

Figura 12

Gráfico de Cajas ANOVA Puntaje Global- Computador 2020



Nota. Este gráfico muestra la representación gráfica de la prueba de ANOVA puntaje global-computador para el año 2020.

El análisis ANOVA reveló los siguientes resultados: Acceso a Internet: El *Valor f* = 196.724, acompañado de un *Valor p* < $2e - 16$, confirma que el acceso a Internet tiene un efecto significativo en los puntajes globales de los estudiantes. Este resultado confirma que los estudiantes con acceso a Internet tienden a obtener puntajes más altos en comparación con aquellos que no tienen acceso.

Acceso a Computador: De manera similar, el *Valor f* = 85.022, con un *Valor p* < $2e - 16$, también indica un impacto significativo del acceso a un computador sobre los puntajes globales. Los estudiantes que cuentan con un computador en sus hogares obtienen puntajes notablemente mejores que aquellos que no lo tienen.

Interacción entre Acceso a Internet y Computadores: El *Valor f* = 1.722, con un *Valor p* = 0.189, sugiere que no existe una interacción significativa entre el acceso a Internet y el acceso a un computador en términos de impacto en los puntajes globales. Esto significa que el

efecto de tener acceso a ambas tecnologías no es mayor que la suma de los efectos individuales de cada una.

La desviación residual, lo que indica que se ajusta mejor que el modelo nulo. Sin embargo, el AIC de 4518.6 indica que el modelo podría mejorar con la inclusión de más variables predictoras.

Pruebas T-test

Realizamos un T-test para comparar los puntajes globales basados en el acceso a internet y a computador en el año 2020.

Figura 13

T-test Resultado Global Acceso a Internet Y Computador 2020

```
> print(t_test_internet)

welch Two Sample t-test

data: PUNT_GLOBAL by FAMI_TIENEINTERNET
t = -14.295, df = 3109.3, p-value < 2.2e-16
alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
95 percent confidence interval:
 -23.07338 -17.50726
sample estimates:
mean in group 0 mean in group 1
 226.8539      247.1442

> # Realizar un t-test para comparar los puntajes globales basados en el acceso a computador
> t_test_computador <- t.test(PUNT_GLOBAL ~ FAMI_TIENECOMPUTADOR, data = data2020_clean)
>
> # Mostrar los resultados del t-test
> print(t_test_computador)

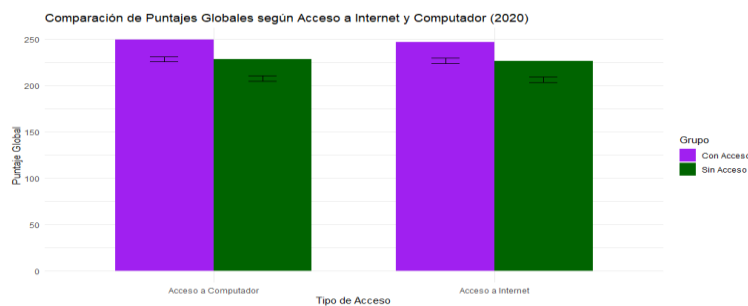
welch Two Sample t-test

data: PUNT_GLOBAL by FAMI_TIENECOMPUTADOR
t = -14.725, df = 3349.6, p-value < 2.2e-16
alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
95 percent confidence interval:
 -23.71562 -18.14212
sample estimates:
mean in group 0 mean in group 1
 228.6062      249.5351
```

Nota. Esta figura muestra el resultado por consola del T- test para el año 2020.

Figura 14

Comparación T-test 2020



Nota. Esta gráfica compara los resultados globales según el acceso a internet y computador en el año 2020.

Acceso a Internet:

Estadística t: El *Valor t* = -14.295 , lo que denota una diferencia significativa en los puntajes globales entre los estudiantes con y sin acceso a Internet.

Grados de Libertad: 3109.3.

Valor p < $2.2e - 16$, lo que sugiere que la diferencia en los puntajes es altamente significativa a cualquier nivel convencional.

Intervalo de Confianza (95%): La diferencia de medias está entre -23.07 y -17.51 , lo que indica que, en promedio, los estudiantes sin acceso a Internet puntúan entre 17.51 y 23.07 puntos menos que aquellos con acceso.

Medias: Los estudiantes sin acceso a Internet tienen una media de 226.85 puntos, mientras que aquellos con acceso tienen una media de 247.14, reflejando una diferencia de 20.29 puntos.

Acceso a Computador:

Estadística t: El *Valor t* = -14.725 , indicando una diferencia significativa en los puntajes globales entre los estudiantes con y sin acceso a un computador.

Grados de Libertad: 3349.6.

Valor p < $2.2e - 16$, confirmando una diferencia estadísticamente significativa.

Intervalo de Confianza (95%): La diferencia de medias está entre -23.72 y -18.14 , lo que implica que los estudiantes sin acceso a un computador puntúan entre 18.14 y 23.72 puntos menos que aquellos con acceso.

Media: La media de puntajes para los estudiantes sin acceso a computador es 228.61 , mientras que para los que tienen acceso es 249.54 , con una diferencia de 20.93 puntos.

Conclusiones: Se concluye que, en el análisis, se evidencian diferencias significativas en el rendimiento académico de los estudiantes, probablemente asociadas a las condiciones particulares del modelo de educación remota implementado durante la pandemia. Estos análisis confirman que el acceso a Internet y a computadores tiene un impacto significativo y positivo en los puntajes globales. Las diferencias observadas son consistentes y considerables, con los estudiantes que tienen acceso a estas tecnologías obteniendo alrededor de 20 puntos más que aquellos sin acceso. Estos resultados refuerzan la importancia del acceso a tecnología como un factor clave en el rendimiento académico, lo que podría apoyar políticas que mejoren la disponibilidad de recursos tecnológicos como una estrategia para elevar los resultados educativos.

Análisis del Periodo 2021

Se realizó el análisis de los datos correspondientes al año 2021, proporcionados por el ICFES, con un total de 3718 registros. De estos, 482 presentaban valores ausentes en las variables FAMI_TIENEINTERNET y FAMI_TIENECOMPUTADOR, lo que representa el

12.9% del total. Dado que esta proporción de datos faltantes no supera el umbral del 15%, se decidió eliminar estos registros para asegurar la integridad y precisión del análisis.

Figura 15

Summary 2021

```

> summary(Data2021_clean)
ESTU_GENERO      PERIODO      ESTU_TIENEETNIA      ESTU_MCPIO_RESIDE      FAMI_ESTRATOVIENDA      FAMI_TIENEINTERNET
Min.   :0.0000   Min.   :20214   Min.   :0.000000   Length:3236   Min.   :1.0   Min.   :0.0000
1st Qu.:0.0000   1st Qu.:20214   1st Qu.:0.000000   Class :character   1st Qu.:1.0   1st Qu.:0.0000
Median :0.0000   Median :20214   Median :0.000000   Mode  :character   Median :1.0   Median :1.0000
Mean   :0.4481   Mean   :20214   Mean   :0.008653             Mean :1.7   Mean   :0.6706
3rd Qu.:1.0000   3rd Qu.:20214   3rd Qu.:0.000000             3rd Qu.:2.0   3rd Qu.:1.0000
Max.   :1.0000   Max.   :20214   Max.   :1.000000             Max.   :6.0   Max.   :1.0000

FAMI_TIENECOMPUTADOR  COLE_NATURALEZA  COLE_BILILOGUE  COLE_AREA_UBICACION  PUNT_LECTURA_CRITICA  PUNT_MATEMATICAS
Min.   :0.0000   Min.   :0.0000   Min.   :0.00000   Min.   :0.0000   Min.   :22.00   Min.   :17.00
1st Qu.:0.0000   1st Qu.:1.0000   1st Qu.:0.00000   1st Qu.:1.0000   1st Qu.:44.00   1st Qu.:42.00
Median :0.0000   Median :1.0000   Median :0.00000   Median :1.0000   Median :50.00   Median :49.00
Mean   :0.4666   Mean   :0.8733   Mean   :0.01082   Mean   :0.8708   Mean   :50.81   Mean   :48.94
3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:0.00000   3rd Qu.:1.0000   3rd Qu.:58.00   3rd Qu.:56.00
Max.   :1.0000   Max.   :1.0000   Max.   :1.00000   Max.   :1.0000   Max.   :81.00   Max.   :100.00

PUNT_C_NATURALES  PUNT_SOCIALES_CIUDADANAS  PUNT_INGLES  DESEMP_INGLES  PUNT_GLOBAL
Min.   :24.00   Min.   :20.00   Min.   :27.00   Length:3236   Min.   :131.0
1st Qu.:41.00   1st Qu.:37.00   1st Qu.:38.00   Class :character   1st Qu.:207.0
Median :46.00   Median :45.00   Median :45.00   Mode  :character   Median :235.0
Mean   :46.81   Mean   :45.34   Mean   :46.27             Mean   :239.2
3rd Qu.:52.00   3rd Qu.:53.00   3rd Qu.:53.00             3rd Qu.:268.0
Max.   :79.00   Max.   :100.00   Max.   :100.00             Max.   :407.0

```

Nota. Esta figura muestra el resultado por consola de Summary.

El análisis de los resultados de las pruebas ICFES para el año 2021 se realizó empleando medidas de tendencia central (media y mediana) y de dispersión (desviación estándar), con el fin de analizar el rendimiento de los estudiantes en cada área evaluada.

Puntaje Global:

- Media: 239.2
- Mediana: 235
- Desviación estándar: 42.82

El puntaje global promedio es de 239.2, con una mediana de 235, lo que señala una distribución simétrica en torno a estos valores. La desviación estándar de 42.82 denota una dispersión moderada, sugiriendo que la mayoría de los estudiantes se encuentran en un rango cercano a la media, con algunos puntajes más extremos.

Matemáticas:

- Media: 48.94

- Mediana: 49
- Desviación estándar: 10.64

Los puntajes en matemáticas muestran una media de 48.94 y una mediana de 49, lo que señala una distribución centrada alrededor de ese valor. La variabilidad es moderada, con la mayoría de los estudiantes puntuando dentro de un rango cercano.

Lectura Crítica:

- Media: 50.81
- Mediana: 50
- Desviación estándar: 9.40

La media y la mediana en lectura crítica son prácticamente idénticas, sugiriendo una distribución balanceada. La desviación estándar de 9.40 indica una dispersión moderada en los puntajes.

Ciencias Naturales:

- Media: 46.81
- Mediana: 46
- Desviación estándar: 9.21

Los puntajes en ciencias naturales son ligeramente más bajos que en matemáticas y lectura crítica. La desviación estándar de 9.21 sugiere una distribución moderada en torno a la media.

Sociales y Ciudadanas:

- Media: 45.34
- Mediana: 45
- Desviación estándar: 11.03

Los puntajes en sociales y ciudadanas son un poco más bajos en comparación con otras áreas. La desviación estándar de 11.03 indica una mayor variabilidad en los resultados.

Inglés:

- Media: 46.27
- Mediana: 45
- Desviación estándar: 9.28

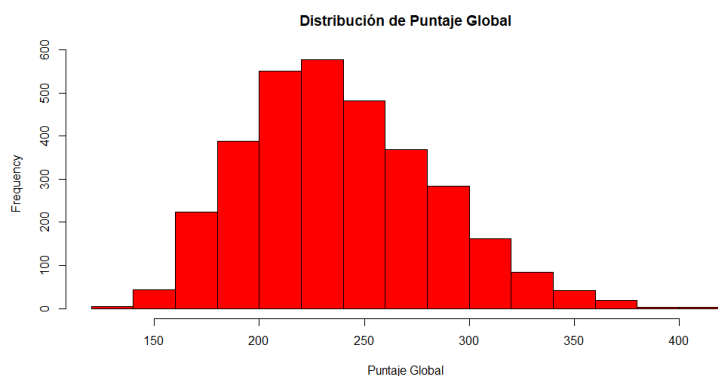
Los puntajes en inglés muestran una media de 46.27 y una mediana de 45, con una desviación estándar de 9.28, lo que refleja una dispersión moderada alrededor de la media.

Conclusiones: El análisis del año 2021 permite identificar patrones diferenciados en el rendimiento académico de los estudiantes, posiblemente vinculados a los efectos prolongados de la educación remota. La evidencia estadística respalda la hipótesis inicial, al mostrar tendencias consistentes en la forma en que distintas áreas del conocimiento se ven afectadas bajo estas condiciones.

Los puntajes globales y por área muestran una distribución generalmente simétrica. Áreas como lectura crítica y ciencias naturales presentan una menor dispersión, lo que denota que los estudiantes tienden a rendir de manera más homogénea en estas áreas. En contraste, sociales y ciudadanas tienen la mayor variabilidad, lo que refleja diferencias más pronunciadas entre los estudiantes. El rendimiento en inglés, aunque más bajo en promedio, tiene una variabilidad moderada.

Figura 16

Diagrama de Barras Resultados 2021



Nota. Este diagrama de barras representa las métricas de resultado de las pruebas ICFES Saber 11 del año 2021.

Pruebas ANOVA

Creamos esta prueba para validar la relación entre el puntaje global y el acceso a internet y computadora para el año 2021.

Figura 17

Prueba ANOVA Puntaje Global- Internet- Computador 2021

```

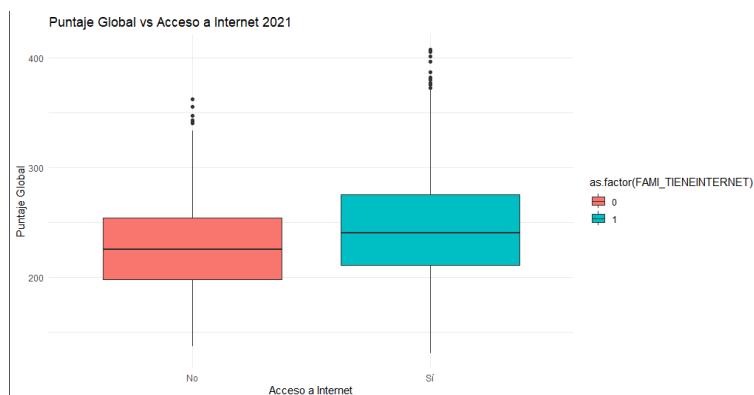
> # Resumen del ANOVA
> summary(anova_internet)
          Df Sum Sq Mean Sq F value Pr(>F)
FAMI_TIENEINTERNET  1 180423  180423  93.23 <2e-16 ***
Residuals        3234  6258379  1935
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
> # Realizar el ANOVA para el acceso a computador
> anova_computador <- aov(PUNT_GLOBAL ~ FAMI_TIENECOMPUTADOR, data = Data2021_clean)
>
> # Resumen del ANOVA
> summary(anova_computador)
          Df Sum Sq Mean Sq F value Pr(>F)
FAMI_TIENECOMPUTADOR  1 408630  408630  219.2 <2e-16 ***
Residuals          3234  6030172  1865
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
> # ANOVA con acceso a internet y computadores como factores
> anova_internet_computador <- aov(PUNT_GLOBAL ~ FAMI_TIENEINTERNET * FAMI_TIENECOMPUTADOR, data = Data2021_clean)
>
> # Resumen del ANOVA
> summary(anova_internet_computador)
          Df Sum Sq Mean Sq F value Pr(>F)
FAMI_TIENEINTERNET  1 180423  180423  97.324 <2e-16 ***
FAMI_TIENECOMPUTADOR  1 257377  257377 138.835 <2e-16 ***
FAMI_TIENEINTERNET:FAMI_TIENECOMPUTADOR  1  9410  9410  5.076 0.0243 *
Residuals          3232  5991592  1854
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>

```

Nota. Esta figura muestra el resultado por consola de la prueba de ANOVA para el año 2021.

Figura 18

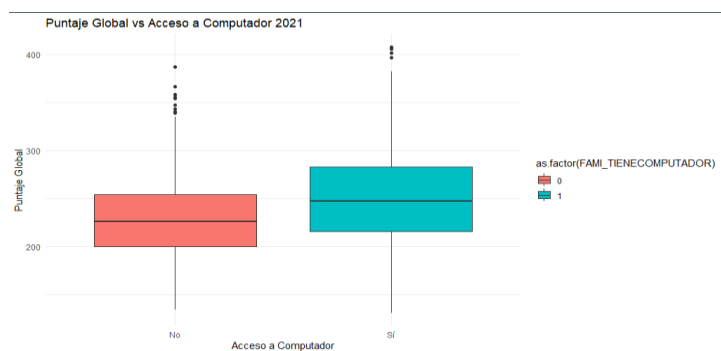
Gráfico de Cajas ANOVA Puntaje Global- Internet 2021



Nota. Este gráfico muestra la representación gráfica de la prueba de ANOVA puntaje global-internet para el año 2021.

Figura 19

Gráfico de Cajas ANOVA Puntaje Global- Computador 2021



Nota. Este gráfico muestra la representación gráfica de la prueba de ANOVA puntaje global-computador para el año 2021.

El análisis ANOVA reveló los siguientes resultados:

Acceso a Internet: El *Valor f* = 93.23 con un *Valor p* < $2e - 16$, lo que señala que el acceso a Internet tiene un efecto significativo en los puntajes globales. Esto significa que los estudiantes con acceso a Internet obtienen puntajes significativamente más altos en las pruebas ICFES en comparación con aquellos que no tienen acceso.

Acceso a Computador: El análisis del acceso a un computador arroja un *Valor f* = 219.2 , también con un *Valor p* < $2e - 16$, lo que confirma un impacto significativo en los puntajes globales. Los estudiantes con acceso a un computador muestran un rendimiento significativamente mejor en comparación con aquellos que no cuentan con esta herramienta.

Interacción entre Acceso a Internet y Computador: El análisis de la interacción entre el acceso a Internet y a un computador muestra un *Valor f* = 5.076 y un *Valor p* = 0.0243, lo que marca que la interacción entre ambas variables es estadísticamente significativa. Aunque este efecto es menor en comparación con los efectos individuales del acceso a Internet o a un computador, la interacción muestra que tener acceso a ambas tecnologías tiene un impacto adicional en los puntajes globales de los estudiantes.

Pruebas T-test

Realizamos un T-test para comparar los puntajes globales basados en el acceso a internet y a computador para el año 2021.

Figura 20

T-test Resultado Global Acceso a Internet y Computador 2021

```

> # Mostrar los resultados del t-test
> print(t_test_internet)

      Welch Two Sample t-test

data:  PUNT_GLOBAL by FAMI_TIENEINTERNET
t = -10.156, df = 2420.5, p-value < 2.2e-16
alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
95 percent confidence interval:
 -18.95448 -12.81952
sample estimates:
mean in group 0 mean in group 1
 228.5582      244.4452

>
> # Realizar un t-test para comparar los puntajes globales basados en el acceso a computador
> t_test_computador <- t.test(PUNT_GLOBAL ~ FAMI_TIENECOMPUTADOR, data = data2021_clean)
>
> # Mostrar los resultados del t-test
> print(t_test_computador)

      Welch Two Sample t-test

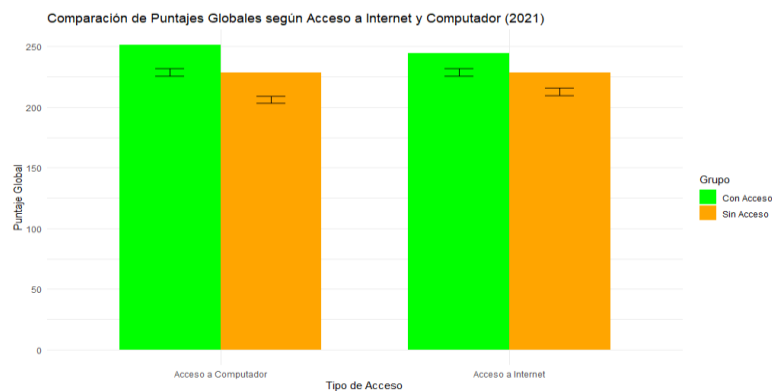
data:  PUNT_GLOBAL by FAMI_TIENECOMPUTADOR
t = -14.611, df = 2922, p-value < 2.2e-16
alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
95 percent confidence interval:
 -25.54751 -19.50206
sample estimates:
mean in group 0 mean in group 1
 228.7010      251.2258

```

Nota. Esta figura muestra el resultado por consola del T-test para el año 2021.

Figura 21

Comparación T-test 2021



Nota. Esta gráfica compara los resultados globales según el acceso a internet y computador en el año 2021.

Los resultados de las pruebas t muestran que tanto el acceso a Internet como a un computador tienen un efecto significativo en los puntajes globales de los estudiantes.

Acceso a Internet: El *Valor t* = -10.156 , con un *Valor p* $< 2.2e - 16$, lo cual denota que hay una diferencia significativa en los puntajes globales entre los estudiantes con y sin acceso a Internet. Los estudiantes sin acceso a Internet tienen una media de 228.56 puntos, mientras que aquellos que cuentan con acceso tienen una media de 244.45 puntos. El intervalo de confianza es del 95% para la diferencia de medias está entre -18.95 y -12.82, lo que marca que, en promedio, los estudiantes sin acceso a Internet puntúan entre 12.82 y 18.95 puntos menos.

Acceso a Computador: El *Valor t* = -14.611 , con un *Valor p* $< 2.2e - 16$, lo que muestra una diferencia significativa en los puntajes globales entre los estudiantes con y sin acceso a un computador. Los estudiantes sin acceso a un computador tienen una media de 228.70 puntos, mientras que aquellos con acceso obtienen una media de 251.23 puntos. El intervalo de confianza del 95% para la diferencia de medias está entre -25.55 y -19.50, lo que indica que los estudiantes sin acceso a un computador puntúan entre 19.50 y 25.55 puntos menos.

Conclusiones: A partir del análisis realizado para el año 2021, se identifican asociaciones claras entre el rendimiento académico y las condiciones de acceso a recursos tecnológicos. La evidencia estadística respalda la hipótesis de que el entorno digital de aprendizaje desempeñó un papel clave en los resultados obtenidos por los estudiantes.

Los dos análisis confirman que los estudiantes con acceso a Internet y a un computador obtienen puntajes significativamente más altos que aquellos sin acceso, con diferencias notables en los puntajes promedio.

Comparación de los Resultados por Años 2019-2021

Figura 22

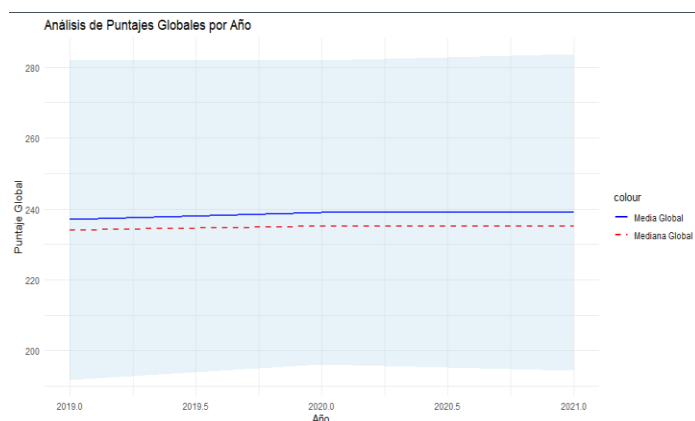
Resultados Globales por Año

```
# A tibble: 3 × 4
  PERIODO Media_Global Mediana_Global Desvstd_Global
  <dbl>    <dbl>         <dbl>         <dbl>
1  2019      237.           234            45.2
2  2020      239.           235            42.8
3  2021      239.           235            44.6
```

Nota. Esta figura muestra el resultado por consola de los puntajes globales.

Figura 23

Gráfica de Resultados Globales por Año



Nota. Esta gráfica de líneas muestra los resultados globales por años.

Nos muestra la evolución de los puntajes globales en los años 2019, 2020 y 2021, con la media y la mediana de los puntajes representados. La línea azul indica la media global y la línea roja discontinua muestra la mediana global. Se identifica que tanto la media como la mediana de los puntajes globales se mantienen bastante estables a lo largo del tiempo, sin grandes

variaciones entre los tres años, esto indica que la pandemia en 2020 no tuvo un impacto significativo en los resultados globales de los estudiantes. Sin embargo, la leve disminución en la desviación estándar en 2020 podría indicar una menor dispersión de los puntajes en ese año.

Los resultados de los puntajes globales entre los años 2019, 2020 y 2021 muestran lo siguiente:

Año 2019:

- Media: 237
- Mediana: 234
- Desviación estándar: 45.2

Año 2020:

- Media: 239
- Mediana: 235
- Desviación estándar: 42.8

Año 2021:

- Media: 239
- Mediana: 235
- Desviación estándar: 44.6

Conclusiones: El análisis comparativo de los tres años no muestra variaciones significativas en los puntajes globales, lo cual sugiere cierta estabilidad en el rendimiento académico promedio a pesar de los cambios en el contexto educativo. Sin embargo, las pruebas estadísticas permiten identificar ligeras fluctuaciones en la dispersión de los resultados, que podrían estar relacionadas con el impacto progresivo del modelo de educación remota durante la pandemia.

En términos generales, no se observan variaciones significativas en los puntajes globales entre los tres años. La media y la mediana se mantuvieron prácticamente iguales durante este período, lo que sugiere que el rendimiento académico promedio de los estudiantes no cambió sustancialmente a lo largo del tiempo. La desviación estándar, la cual mide la variabilidad en los puntajes, también mostró una leve disminución en 2020, lo que indica una menor dispersión en los resultados en ese año, seguido de un leve aumento en 2021.

¿Afectó la Pandemia en 2020 los Resultados de las Pruebas ICFES Saber 11?

A pesar de las interrupciones causadas por la pandemia en 2020, los resultados muestran que no hubo un impacto negativo notable en los puntajes globales. La media y la mediana en 2020 fueron ligeramente superiores a las de 2019, y la desviación estándar disminuyó, lo que sugiere que, en promedio, los estudiantes mantuvieron un rendimiento estable durante este período. Esto indica que, al menos en términos generales, la pandemia no afectó de manera negativa los resultados académicos de los estudiantes, al menos en los puntajes globales del examen ICFES.

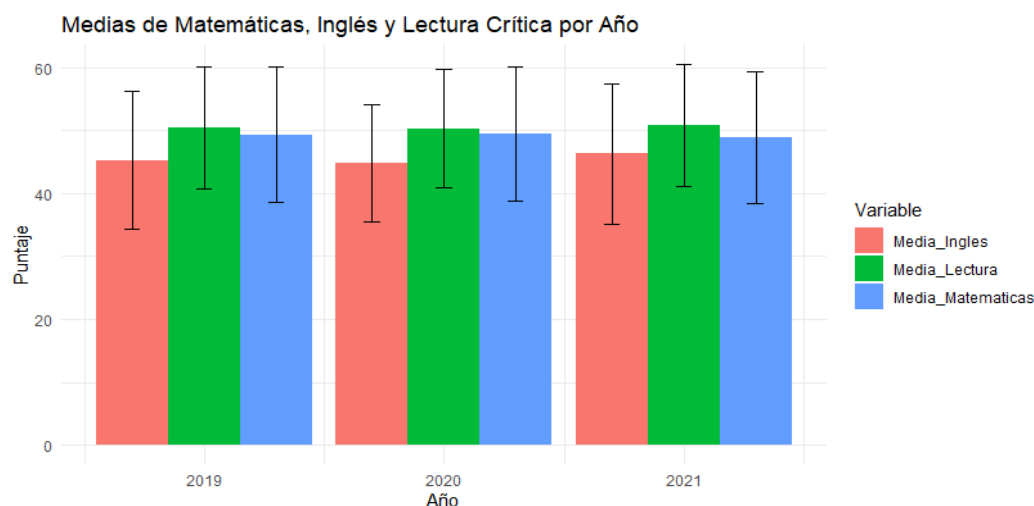
Comparación de los Resultados en Matemáticas, Inglés y Lectura Crítica

Figura 24

Comparación por Areas

```
# A tibble: 3 x 10
  PERIODO Media_Matematicas Mediana_Matematicas Desvstd_Matematicas Media_Ingles Mediana_Ingles Desvstd_Ingles Media_Lectura Mediana_Lectura Desvstd_Lectura
  <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1 2019 49.3 49 10.8 45.3 44 11.0 50.4 51 9.62
2 2020 49.4 49 10.6 44.8 44 9.28 50.3 50 9.40
3 2021 48.9 49 10.5 46.3 45 11.1 50.8 50 9.73
```

Nota. Esta figura muestra el resultado por consola, las comparaciones de los puntajes por áreas.

Figura 25*Gráfica de Comparación por Areas*

Nota. Esta gráfica ilustra la comparación de los resultados por áreas específicas.

Matemáticas: La media se mantiene relativamente constante en los tres años. En 2019, la media fue de 49.3, y aunque en 2020 subió ligeramente a 49.4, en 2021 bajó a 48.9. Las desviaciones estándar son similares en los tres años, lo que denota que la variabilidad en los puntajes no cambió significativamente a lo largo del tiempo.

Inglés: En 2019, la media fue de 45.3, pero en 2020 bajó ligeramente a 44.8. Sin embargo, en 2021, la media subió notablemente a 46.3. Esto sugiere que, en promedio, los estudiantes obtuvieron mejores resultados en el área de inglés en 2021 en comparación con los dos años anteriores. Las desviaciones estándar también son similares, lo que sugiere consistencia en la variabilidad de los puntajes.

Lectura Crítica: Al igual que en Matemáticas, los puntajes de Lectura Crítica han sido bastante estables en los tres años. La media en 2019 fue de 50.4, ligeramente baja en 2020 a 50.3, y volvió a subir a 50.8 en 2021. Esto sugiere que no hubo cambios significativos en los resultados

de Lectura Crítica durante el período, y las desviaciones estándar son también bastante consistentes.

Conclusiones: El análisis de las áreas evaluadas entre 2019 y 2021 muestra una tendencia general de estabilidad en el rendimiento académico. No obstante, las pruebas estadísticas permiten identificar ligeras variaciones, como el incremento en los puntajes de inglés y la leve disminución en matemáticas en 2021, que podrían reflejar ajustes en los procesos de enseñanza-aprendizaje durante y después del periodo de educación remota.

Los resultados en las tres áreas se han mantenido relativamente estables a lo largo de los tres años. Sin embargo, cabe destacar el aumento en la media de los puntajes de inglés en 2021, lo cual podría indicar una mejora en el rendimiento en esa área específica. Por otro lado, Matemáticas ha tenido una ligera disminución en 2021, pero sigue siendo bastante consistente en general.

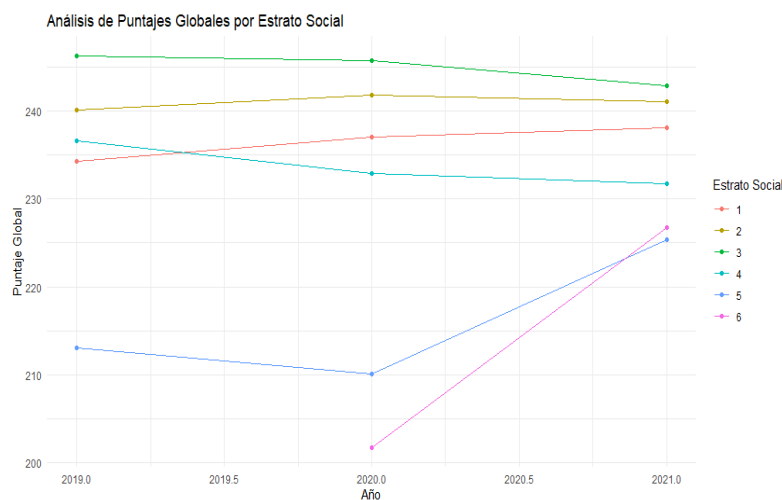
Análisis por Estrato Social

Figura 26

Comparación por Estrato Social

```
# Groups: FAMI ESTRATOVIVIENDA [6]
  FAMI ESTRATOVIVIENDA PERIODO Media_Global Mediana_Global DesvStd_Global
    <int>    <dbl>      <dbl>          <dbl>          <dbl>
1         1      2019       234.           232            42.6
2         1      2020       237.           234            40.4
3         1      2021       238.           235            41.9
4         2      2019       240.           238            44.9
5         2      2020       242.           237            43.2
6         2      2021       241.           237            45.8
7         3      2019       246.           242            49.7
8         3      2020       246.           241            47.8
9         3      2021       243.           236            49.9
10        4      2019       237.           224            61.0
11        4      2020       233.           224            53.0
12        4      2021       232.           221            53.8
13        5      2019       213.           199            51.4
14        5      2020       210.           198            47.3
15        5      2021       225.           215            38.9
16        6      2020       202.           201            36.9
17        6      2021       227.           218            50.1
```

Nota. Esta figura muestra el resultado por consola de la comparación por estrato social.

Figura 27*Gráfica de Comparación por Estrato Social*

Nota. Esta gráfica de líneas muestra la evolución de los puntajes globales por estrato social a lo largo de los años.

El gráfico indica el análisis de los puntajes globales de las pruebas ICFES por estrato social a lo largo de los años 2019, 2020 y 2021.

Podemos observar que los estudiantes de estratos 1 y 2 mantienen puntajes relativamente estables, con ligeras variaciones de un año a otro. En cambio, los estratos 3 y 4 presentan una tendencia levemente decreciente, especialmente en 2020, lo que podría estar relacionado con el impacto de la pandemia. Sin embargo, los estratos 5 y 6 muestran un comportamiento particular: el estrato 5 cae drásticamente en 2020, pero luego recupera de manera significativa en 2021, mientras que el estrato 6, a pesar de ser el menos representado, también muestra una notable recuperación en 2021 tras un descenso en 2020.

Esto indica que los efectos de la pandemia variaron según el estrato social, afectando más a los estratos intermedios- altos (4 y 5) en términos de estabilidad en los puntajes globales, aunque algunos lograron recuperarse en 2021.

Estrato 1

Tendencia: En 2019, la media del puntaje global fue de 234, aumentando ligeramente en 2020 (237) y 2021 (238).

Variabilidad: La desviación estándar se mantuvo relativamente constante alrededor de 42-43, lo que indica una variación moderada en los puntajes dentro de este estrato.

Mediana: La mediana también ha sido constante, en torno a los 232-235, lo que denota que la mayoría de los estudiantes se mantienen en una tendencia estable en su rendimiento.

Estrato 2

Tendencia: Este estrato también muestra un ligero aumento en el puntaje global entre 2019 (240) y 2020 (242), aunque en 2021 disminuye levemente (241).

Variabilidad: La desviación estándar ha oscilado entre 43 y 46, lo que indica una ligera variación en los resultados a lo largo de los años.

Mediana: La mediana se mantiene cercana a los 237, lo que sugiere estabilidad en el rendimiento.

Estrato 3

Tendencia: Los puntajes en este estrato han sido consistentes, con una media de 246 en 2019 y 2020, y una ligera disminución en 2021 (243).

Variabilidad: La desviación estándar es mayor que en los estratos más bajos, alrededor de 47-50, lo que sugiere una mayor dispersión de los resultados.

Mediana: La mediana se ha mantenido estable, aunque en 2021 bajó a 236.

Estrato 4

Tendencia: Los estudiantes de este estrato muestran un descenso en los puntajes medios desde 2019 (237) hasta 2021 (232).

Variabilidad: La desviación estándar es relativamente alta, lo que indica una mayor variabilidad en los puntajes (61 en 2019, 53 en 2020 y 2021).

Mediana: La mediana disminuye ligeramente de 224 a 221 entre 2019 y 2021.

Estrato 5

Tendencia: Este estrato muestra una ligera disminución en los puntajes en 2020 (210), pero una recuperación en 2021, con una media de 225.

Variabilidad: La desviación estándar disminuyó en 2021, lo que sugiere una mayor consistencia en los resultados en ese año.

Mediana: En 2021, la mediana subió a 215, lo que refleja una mejora en los puntajes.

Estrato 6

Tendencia: Aunque en 2020 los puntajes son bajos (202), en 2021 se observa un aumento significativo a 227.

Variabilidad: La desviación estándar también aumentó considerablemente en 2021, lo que sugiere una mayor variabilidad en los puntajes.

Mediana: La mediana subió de 201 en 2020 a 218 en 2021, lo que indica una mejora en el rendimiento promedio.

Conclusiones: Se concluye que existe una tendencia de ligera mejora o estabilidad en los puntajes globales de los estudiantes de los estratos más bajos (1-3) a lo largo de los años, mientras que los estratos más altos muestran variaciones más notables, con una disminución en

los puntajes en 2020 seguida de una recuperación en 2021. La variabilidad en los puntajes también es mayor en los estratos superiores.

Análisis Segmentado por Estrato Socioeconómico

Figura 28

Análisis Segmentado por Estrato Socioeconómico



Nota. Esta gráfica de cajas muestra la distribución de los puntajes globales por estrato socioeconómico.

Este gráfico de boxplot y los valores estadísticos nos proporcionan una visión clara de cómo los puntajes globales en las pruebas ICFES varían según el estrato socioeconómico. Los puntajes globales tienden a ser más altos en estratos más elevados (3 y 4), con un ligero descenso en los estratos más altos (5 y 6). Esto podría reflejar una mayor calidad educativa o acceso a recursos en los niveles intermedios.

La dispersión es mayor en los estratos intermedios y altos, lo que nos indica una mayor variabilidad en el rendimiento académico. Esto nos sugiere que no todos los estudiantes de

estratos altos tienen un desempeño uniforme, posiblemente debido a diferencias en el acceso a tecnología, apoyo educativo o calidad de las instituciones educativas.

Estrato Bajo (1 y 2): Los estudiantes de estratos bajos presentan puntajes globales más bajos en promedio, con una distribución más estrecha y menos valores atípicos. Esto refleja limitaciones en el acceso a recursos educativos y tecnológicos, lo que impacta negativamente en el rendimiento académico.

Estratos intermedios (3 y 4): Estos grupos muestran un mejor desempeño académico, con medias superiores a los estratos bajos. Esto podría ser indicativo de una mayor inversión familiar en educación y acceso a tecnología.

Estratos Altos (5 y 6): A pesar de tener recursos y ventajas educativas, los puntajes no presentan un incremento significativo con respecto a los estratos intermedios, lo que podría deberse a factores como expectativas más altas o mayor estrés académico.

La relación positiva entre el nivel socioeconómico y el rendimiento académico es evidente, aunque con ciertas limitaciones en los estratos más altos. Esto nos sugiere la necesidad de diseñar intervenciones específicas para los estratos más bajos, como subsidios educativos, programas de acceso a tecnología y fortalecimiento de las instituciones públicas.

En resumen, estos resultados destacan cómo el contexto socioeconómico influye significativamente en el rendimiento académico, subrayando la importancia de las políticas públicas que reducen las brechas educativas entre los diferentes estratos.

Comparación por Tpo de Colegio (Oficial y no Oficial)

Figura 29

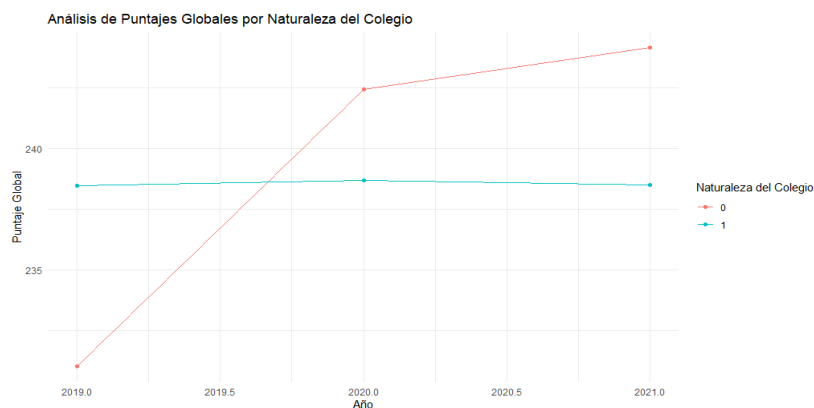
Comparación por Tipo de Colegio

```
# Groups: COLE_NATURALEZA [2]
COLE_NATURALEZA PERIODO Media_Global Mediana_Global DesvStd_Global
<int> <dbl> <dbl> <dbl> <dbl>
1 0 2019 231. 222 53.1
2 0 2020 242. 232. 53.3
3 0 2021 244. 234. 59.5
4 1 2019 238. 237 43.2
5 1 2020 239. 235 40.8
6 1 2021 238. 235 42.0
```

Nota. Esta figura muestra el resultado por consola de la comparación de los resultados por tipo de colegio 0= No oficial, 1= Oficial.

Figura 30

Gráfica de Comparación por Tipo de Colegio



Nota. Esta gráfica de líneas muestra la comparación de los puntajes globales según la naturaleza del colegio.

Se observa en el gráfico que los colegios oficiales (línea azul) mantienen una tendencia muy estable en los puntajes globales, sin variaciones importantes durante los tres años, con una ligera disminución en 2021. Por otro lado, los colegios no oficiales (línea roja) muestran un

incremento significativo en los puntajes globales, especialmente entre 2019 y 2020, y este aumento se mantiene en 2021.

Este comportamiento denota que los colegios no oficiales mejoraron sus resultados de manera notable durante el periodo analizado, mientras que los colegios oficiales mantuvieron un rendimiento más constante, sin cambios significativos en los puntajes globales.

Colegios no Oficiales

Tendencia: En 2019, la media de los puntajes globales fue de 231, aumentando en 2020 a 242 y manteniendo una tendencia positiva en 2021 con una media de 244.

Variabilidad: La desviación estándar es alta en todos los años, con valores que oscilan entre 53.1 en 2019 y 59.5 en 2021, lo que denota una mayor dispersión en los puntajes de los estudiantes de colegios no oficiales. Este aumento en la desviación estándar refleja una creciente heterogeneidad en los resultados de los estudiantes.

Mediana: La mediana ha aumentado de 222 en 2019 a 234 en 2021, lo que indica que los puntajes de los estudiantes en colegios no oficiales han mejorado en general.

Colegios Oficiales

Tendencia: Los colegios oficiales muestran una estabilidad en los puntajes globales entre 2019 y 2021. La media fue de 238 en 2019, aumentó ligeramente a 239 en 2020, y volvió a 238 en 2021. Esto marca una tendencia más estable en comparación con los colegios no oficiales.

Variabilidad: La desviación estándar es menor que en los colegios no oficiales, con valores entre 40.8 y 43.2, lo que indica que los puntajes en colegios oficiales son más homogéneos.

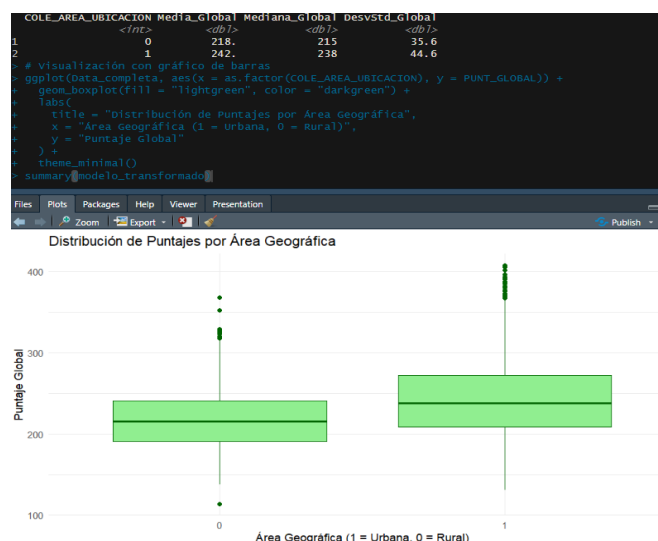
Mediana: La mediana muestra estabilidad, manteniéndose en torno a los 235-237 a lo largo de los tres años. Esto indica que la mayoría de los estudiantes en colegios oficiales se mantienen en un rango de puntajes similar año tras año.

Conclusiones: Los resultados concluyen que los estudiantes de colegios no oficiales experimentaron un aumento en los puntajes globales a lo largo de los años, con una mayor variabilidad en sus resultados, lo que sugiere diferencias más pronunciadas en el rendimiento de los estudiantes de estos colegios. Por otro lado, los estudiantes de colegios oficiales presentaron una mayor estabilidad en sus puntajes, con una variabilidad más baja y un rendimiento más uniforme. Esto podría sugerir que las condiciones de los colegios oficiales proporcionan un entorno más homogéneo para el aprendizaje, mientras que en los colegios no oficiales existe una mayor diversidad en los resultados académicos.

Análisis Segmentado por Área Geográfica (Urbana/Rural)

Figura 31

Análisis Área Geográfica (Urbana/Rural)



Nota. Esta gráfica de cajas muestra la distribución de los puntajes globales según el área geográfica.

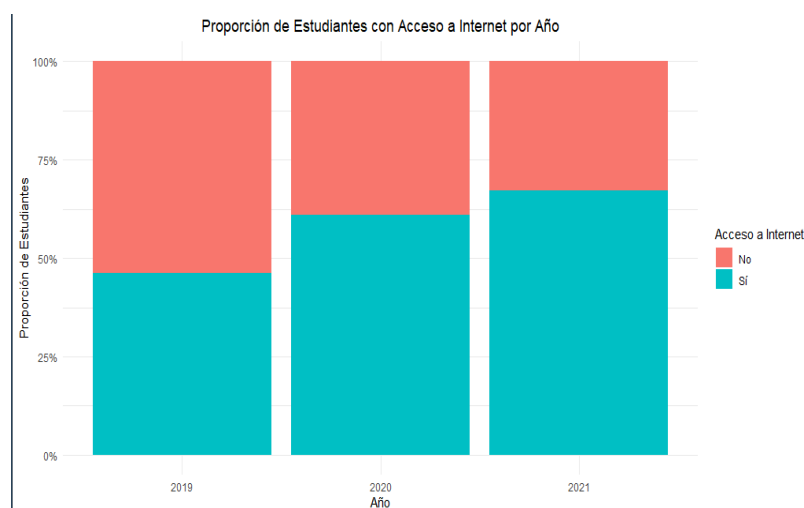
El análisis de los puntajes globales por área geográfica evidencia una clara diferencia entre estudiantes de zonas urbanas y rurales. Los estudiantes de áreas urbanas presentan un puntaje promedio más alto (242 puntos) en comparación con los de áreas rurales (215 puntos). Esta brecha podría deberse a un mejor acceso a recursos educativos, tecnología e infraestructura en las zonas urbanas.

En las áreas rurales, los puntajes son más homogéneos, reflejando limitaciones estructurales comunes, mientras que en las urbanas existe mayor variabilidad, posiblemente asociada a la heterogeneidad en la calidad de las instituciones.

Análisis del Uso de Herramientas Tecnológicas, Internet- Computador

Figura 32

Gráfica de Comparación por Acceso a Internet



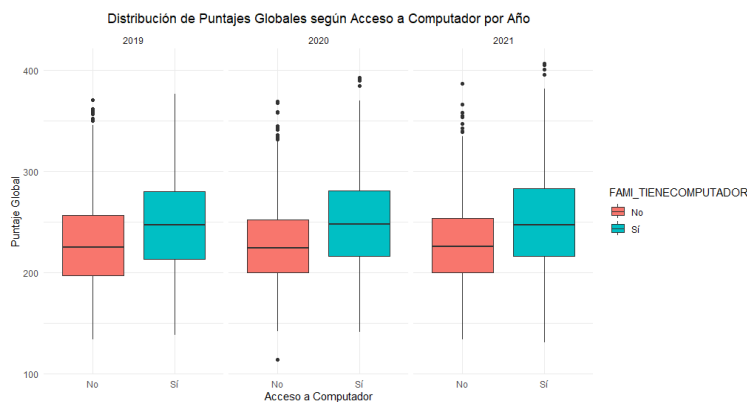
Nota. Esta gráfica apilada muestra la proporción de estudiantes con y sin acceso a Internet a lo largo de los años.

El gráfico refleja la proporción de estudiantes con acceso a Internet en cada año. Se observa que el porcentaje de estudiantes con acceso a Internet aumenta progresivamente,

especialmente en 2020 y 2021, posiblemente debido al impulso de la educación virtual durante la pandemia. Sin embargo, todavía hay una proporción significativa de estudiantes sin acceso, lo que podría ser un obstáculo para mejorar los resultados generales de las pruebas.

Figura 33

Gráfica de Comparación por Acceso a Computador



Nota. Esta gráfica muestra la distribución de los puntajes globales según el acceso a computadora por año.

El gráfico muestra que los estudiantes con acceso a un computador notoriamente obtienen mejores puntajes globales en comparación con aquellos que no tienen acceso, independientemente del año. Esta tendencia se mantiene durante los tres años, lo que indica una relación positiva entre el acceso a esta tecnología y el rendimiento académico. Incluso durante el año 2020, cuando la pandemia afectó la educación, la diferencia en puntajes entre quienes tenían acceso a un computador y quienes no lo tenían sigue siendo clara y notable.

Conclusiones: El análisis de los datos entre 2019 y 2021 evidencia una relación consistente entre el acceso a recursos tecnológicos y el rendimiento académico. Las pruebas estadísticas respaldan la hipótesis de que disponer de Internet y computador influye positivamente en los puntajes obtenidos, incluso durante los años más críticos de la pandemia.

Los estudiantes con acceso tanto a Internet como a un computador obtuvieron consistentemente puntajes globales más altos en comparación con aquellos sin acceso a estas herramientas. En particular, los estudiantes que contaban con ambas tecnologías (Internet y computador) en 2019, 2020 y 2021 presentaron una media de puntaje global superior a los 250 puntos, mientras que aquellos sin ninguna de estas herramientas oscilaron entre 224 y 227 puntos. Esto reafirma la importancia del acceso a tecnologías tanto para la continuidad educativa como para mejorar el rendimiento académico.

El acceso a computadores e Internet tiene un impacto significativo y positivo en los resultados de las pruebas ICFES. Los estudiantes que cuentan con estas herramientas obtienen puntajes globales más altos, lo cual se mantiene incluso durante los años de pandemia. Por lo tanto, se podría recomendar políticas que busquen aumentar el acceso a estos recursos tecnológicos, especialmente en contextos de bajos recursos, para mitigar las brechas educativas existentes.

Modelos de Regresión

Regresión Lineal

Figura 34

Regresión Lineal

```

Call:
lm(formula = log(PUNT_GLOBAL) ~ FAMI_TIENEINTERNET + FAMI_TIENECOMPUTADOR +
  COLE_NATURALEZA + FAMI ESTRATOVIVIENDA, data = Data_completa)

Residuals:
    Min       1q   Median       3Q      Max
-0.6616 -0.1256  0.0024  0.1276  0.5433

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.394906   0.006258  862.025 < 2e-16 ***
FAMI_TIENEINTERNETsí  0.050102   0.004018  12.470 < 2e-16 ***
FAMI_TIENECOMPUTADORsí  0.066031   0.003971  16.628 < 2e-16 ***
COLE_NATURALEZA  0.024948   0.004960   5.030 4.97e-07 ***
FAMI ESTRATOVIVIENDA -0.011016   0.002040  -5.399 6.84e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

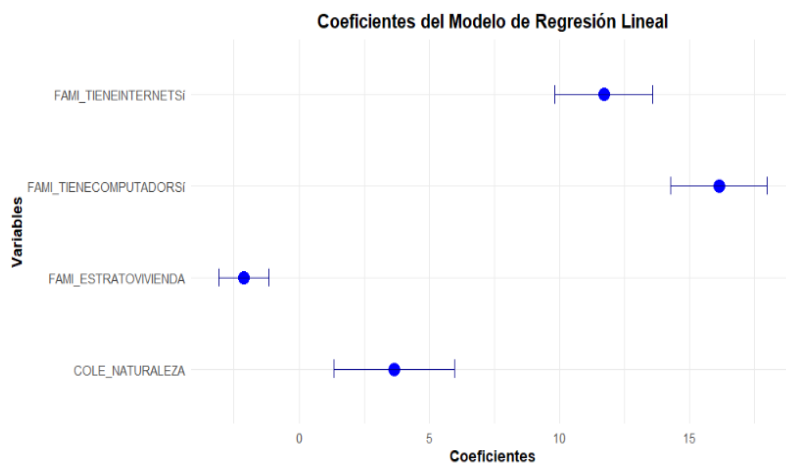
Residual standard error: 0.1793 on 10590 degrees of freedom
Multiple R-squared:  0.06729, Adjusted R-squared:  0.06694
F-statistic: 191 on 4 and 10590 DF, p-value: < 2.2e-16

```

Nota. Esta figura muestra el resultado por consola de la regresión lineal.

Figura 35

Gráfico de Regresión Lineal



Nota. Esta gráfica muestra los coeficientes del modelo de regresión lineal.

En el modelo de regresión lineal se examinan los efectos del acceso a Internet, el acceso a computador, la naturaleza del colegio (oficial o no oficial) y el estrato socioeconómico sobre el puntaje global obtenido en las pruebas ICFES. Los resultados de este análisis nos proporcionan una visión detallada de cómo estos factores contribuyen al rendimiento académico en las pruebas Saber 11 entre 2019 y 2021

Interpretación de los Coeficientes

Intercepto (5.394906): Este valor representa el logaritmo del puntaje global promedio esperado en las pruebas ICFES para un estudiante que no tiene acceso a Internet ni a computadora, que estudia en un colegio no oficial y pertenece al estrato más bajo. Este valor sirve como base sobre la cual se añaden (o restan) los efectos de las otras variables.

FAMI_TIENEINTERNET (0.050102): Los estudiantes que cuentan con acceso a Internet tienen, en promedio, un aumento del 5.1% en su puntaje global en comparación con aquellos que no tienen acceso a Internet, manteniendo constantes las demás variables. Esto indica que el acceso a la conectividad a internet es un factor importante en el rendimiento académico, reflejando la importancia del acceso a información y recursos en línea para mejorar el aprendizaje.

FAMI_TIENECOMPUTADOR (0.066031): Tener un computador en casa es el factor más determinante en este análisis, con un incremento promedio del 6.6% en el puntaje global. Este hallazgo ratifica la relevancia del acceso a tecnología física para el aprendizaje, ya que un computador permite un acceso más eficiente y prolongado a herramientas educativas, investigación y plataformas de estudio.

COLE_NATURALEZA (0.024948): Estudiar en un colegio oficial también tiene un efecto positivo en el puntaje global, con un aumento promedio del 2.5% en comparación con los

estudiantes de colegios no oficiales. Si bien este efecto es más moderado en comparación con el acceso a tecnología, sigue siendo significativo y refleja el impacto positivo de los colegios oficiales en el desempeño de los estudiantes.

FAMI_ESTRATOVIVIENDA (-0.011016): Este coeficiente negativo indica que a medida que el estrato socioeconómico aumenta, se observa una ligera disminución del 1.1% en el puntaje global. Esto podría parecer contraintuitivo, pero refleja la heterogeneidad en la distribución de recursos educativos, la calidad educativa y otros factores dentro de diferentes niveles socioeconómicos.

Evaluación del Modelo

R^2 Ajustado de 0.067: Este valor sugiere que aproximadamente el 6.7% de la variabilidad en los puntajes globales se puede explicar mediante estas cuatro variables. Aunque el R^2 no es extremadamente alto, esto es común en estudios educativos donde múltiples factores externos (como el ambiente escolar, las características del estudiante, y otros factores socioeconómicos) influyen en el rendimiento académico.

Significancia: Todos los coeficientes son altamente significativos ($p < 0.001$), lo que refuerza la confiabilidad de los resultados obtenidos.

Conclusiones: El análisis del modelo lineal aplicado a los datos de los tres años permite identificar factores clave que inciden en el rendimiento académico. Las pruebas estadísticas respaldan la hipótesis de que el acceso a la tecnología y las condiciones educativas específicas tienen un impacto significativo en los puntajes globales de las pruebas ICFES. Además, el hecho de estudiar en un colegio oficial también muestra un efecto positivo, aunque más modesto, en los resultados. Estos hallazgos son coherentes con la premisa de que la brecha tecnológica y las diferencias en los recursos educativos pueden influir significativamente en el rendimiento

académico. La relación negativa observada con el estrato social también indica que, más allá del nivel socioeconómico, los recursos educativos como el acceso a la tecnología juegan un papel crucial en los resultados. Este análisis resalta la importancia de abordar la desigualdad en el acceso a la tecnología para mejorar el rendimiento educativo en poblaciones vulnerables.

Evaluación de Residuos

Analizamos los residuos del modelo lineal para verificar que cumpliera con los supuestos estadísticos necesarios. Esto incluyó:

Visualización de residuos: Generamos un Q-Q Plot y un histograma para evaluar si los residuos siguen una distribución normal.

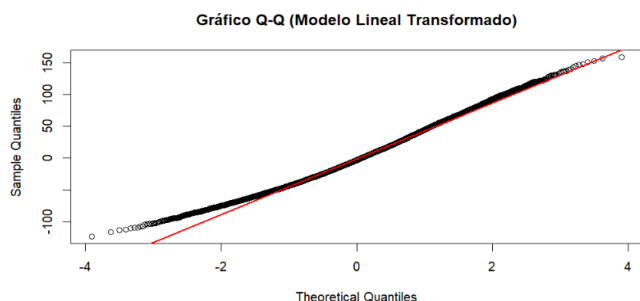
Prueba de Kolmogorov-Smirnov: Aplicamos esta prueba para confirmar la normalidad de los residuos de manera estadística.

Figura 36

Evaluación Residuos Regresión Lineal

```
Shapiro-wilk normality test
data: sample(residuos_log, 5000, replace = FALSE)
W = 0.99104, p-value < 2.2e-16

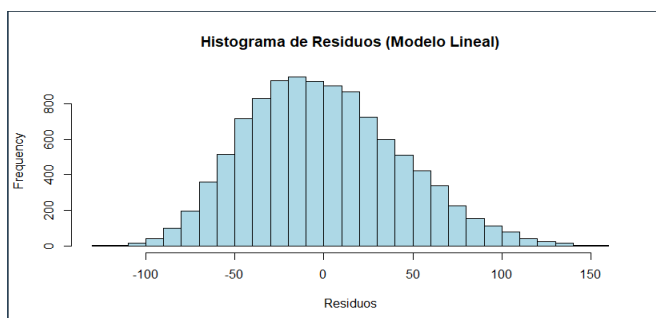
> residuos_log <- residuals(modelo_lineal, type = "pearson")
> qqnorm(residuos_log, main = "Gráfico Q-Q (Modelo Lineal Transformado)")
> qqline(residuos_log, col = "red", lwd = 2)
> summary(modelo_transformado)
```



Nota. Esta figura muestra el resultado por consola de la evaluación de residuos de la regresión lineal.

Figura 37

Histograma de Evaluación Residuos Regresión Lineal



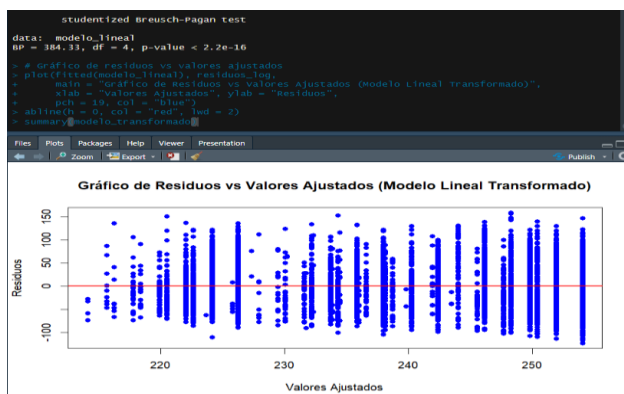
Nota. Este histograma muestra la distribución de los residuos del modelo lineal.

Vemos que el modelo nos ofrece una buena aproximación para analizar cómo variables como el acceso a tecnología, el tipo de institución y el estrato socioeconómico impactan en el rendimiento global. A pesar de las desviaciones observadas en los residuos, los resultados son sólidos y permiten identificar tendencias significativas.

Prueba de Homocedasticidad

Figura 38

Prueba de Homocedasticidad Regresión Lineal



Nota. Este gráfico de dispersión muestra la relación entre los residuos y los valores ajustados del modelo lineal transformado.

Encontramos que el modelo lineal es una herramienta robusta y efectiva para evaluar el impacto de variables clave como el acceso a Internet, computadores, la naturaleza del colegio y el estrato socioeconómico sobre el rendimiento académico. Si bien persisten pequeñas limitaciones técnicas, el modelo logra identificar patrones significativos y útiles que pueden orientar la toma de decisiones en políticas educativas.

Este modelo es un paso adelante en la comprensión de las dinámicas del rendimiento académico, ofreciendo una base sólida para futuras investigaciones y aplicaciones prácticas en el ámbito educativo.

Multicolinealidad

Figura 39

Multicolinealidad Regresión Lineal

```
> # Evaluar la multicolinealidad en el modelo lineal
> vif_lineal <- vif(modelo_lineal)
> print("VIF para el modelo lineal:")
[1] "VIF para el modelo lineal:"
> print(vif_lineal)
  FAMI_TIENEINTERNET  FAMI_TIENECOMPUTADOR  COLE_NATURALEZA  FAMI ESTRATO VIVIENDA
1.301304             1.293004             1.027989             1.078886
> # Interpretación
> if (any(vif_lineal > 5)) {
+   print("Algunas variables en el modelo lineal tienen problemas de multicolinealidad (VIF > 5).
Es necesario realizar ajustes.")
+ } else {
+   print("No se detecta multicolinealidad significativa (VIF <= 5) en el modelo lineal.")
+ }
[1] "No se detecta multicolinealidad significativa (VIF <= 5) en el modelo lineal."
```

Nota. Esta figura muestra el resultado por consola de la multicolinealidad de la regresión lineal.

Los factores de inflación de la varianza (VIF) obtenidos para el modelo logístico son los siguientes:

- FAMILIA_TIENEINTERNET: 1.27
- COMPUTADORA FAMILIAR: 1.27
- COLE_NATURALEZA: 1.03
- FAMILIA ESTRATO VIVIENDA: 1.09

Dado que todos los valores de VIF son menores a 5, no se detecta multicolinealidad significativa en el modelo logístico. Esto implica que las variables incluidas son suficientemente independientes entre sí, y no existe una relación excesiva entre las variables.

Regresión Logística

Figura 40

Regresión Logística

```

# Resumen del modelo
> summary(modelo_logistico)

call:
glm(formula = puntaje_alto ~ FAMI_TIENEINTERNET + FAMI_TIENECOMPUTADOR +
    COLE_NATURALEZA + FAMI ESTRATOVIVIENDA, family = "binomial",
    data = data_completa)

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -0.74330    0.07261  -10.237 < 2e-16 ***
FAMI_TIENEINTERNET  0.40645    0.04576   8.882 < 2e-16 ***
FAMI_TIENECOMPUTADOR  0.65505    0.04502  14.551 < 2e-16 ***
COLE_NATURALEZA    0.29025    0.05719   5.075 3.87e-07 ***
FAMI ESTRATOVIVIENDA -0.11601    0.02363  -4.910 9.13e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 14631  on 10594  degrees of freedom
Residual deviance: 14127  on 10590  degrees of freedom
AIC: 14137

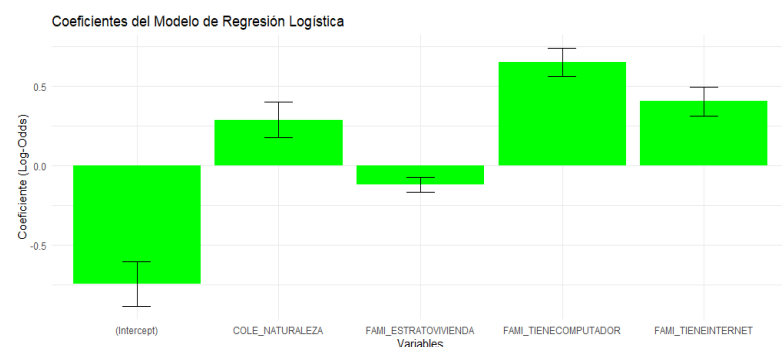
Number of Fisher scoring iterations: 4

```

Nota. Esta figura muestra el resultado por consola de la regresión logística.

Figura 41

Gráfico Regresión Logística



Nota. Esta gráfica muestra los coeficientes (Log-Odds) del modelo de regresión logística.

La regresión logística presentada en el gráfico y el resumen del modelo muestra cómo las variables de acceso a Internet, acceso a computadora, tipo de colegio y estrato socioeconómico influyen en la probabilidad de obtener un puntaje alto en las pruebas ICFES.

Interpretación Coeficientes

Intercepto: El coeficiente del intercepto es negativo (-0.74330), lo que advierte que, en ausencia de acceso a Internet, a computadores, en un colegio no oficial, y en un estrato bajo, la probabilidad de obtener un puntaje alto en las pruebas ICFES es baja.

FAMI_TIENEINTERNET: El coeficiente positivo (0.40645) sugiere que los estudiantes con acceso a Internet tienen una mayor probabilidad de obtener un puntaje alto en comparación con aquellos que no tienen acceso, manteniendo constantes las demás variables. Este efecto es altamente significativo ($p < 2e - 16$).

FAMI_TIENECOMPUTADOR: Con un coeficiente de 0.65505, este es el factor que más influye positivamente en la probabilidad de obtener un puntaje alto. Los estudiantes con acceso a una computadora tienen una probabilidad significativamente mayor de lograr un puntaje elevado ($p < 2e - 16$).

COLE_NATURALEZA: El coeficiente de 0.29025 indica que asistir a un colegio oficial aumenta la probabilidad de obtener un puntaje alto en comparación con asistir a un colegio no oficial, aunque este efecto es más moderado en comparación con el acceso a la tecnología. Este efecto también es significativo ($p < 3.87e - 07$).

FAMI ESTRATOVIVIENDA: Sorprendentemente, el coeficiente es negativo (-0.11601), lo que indica que a medida que aumenta el estrato socioeconómico, disminuye ligeramente la probabilidad de obtener un puntaje alto. Este resultado podría reflejar una

complejidad adicional en los factores que afectan el rendimiento académico en los diferentes estratos sociales.

Evaluación del Modelo

El modelo tiene un AIC de 14137, lo que indica un buen ajuste dado el tamaño de los datos.

La significancia de todas las variables ($p < 0,001$), refuerza la solidez del modelo.

Conclusiones: Se concluye que el acceso a la tecnología (Internet y computador) es un fuerte predictor del rendimiento académico, con el acceso a computadores teniendo el mayor impacto en la probabilidad de obtener un puntaje alto. También se observa que el tipo de colegio y el estrato socioeconómico tienen un papel importante, pero menos determinante. Este análisis apoya la conclusión de que la brecha en el acceso a recursos tecnológicos afecta significativamente el rendimiento académico en las pruebas ICFES.

Evaluación de Residuos

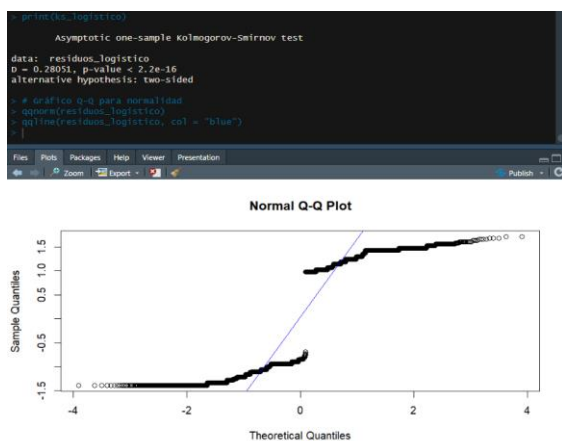
Evaluamos los residuos del modelo logístico para comprender mejor su ajuste:

Visualización de residuos deviance: Utilizamos un Q-Q Plot y un histograma para observar patrones de distribución.

Prueba de Kolmogorov-Smirnov: Verificamos si los residuos siguen una distribución normal, aunque sabemos que esto no es un requisito estricto en modelos logísticos.

Figura 42

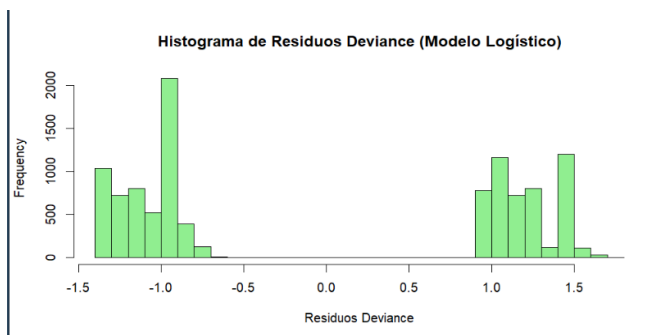
Evaluación de Residuos Regresión Logística



Nota. Esta figura muestra el resultado por consola de la evaluación de residuos regresión logística.

Figura 43

Histograma de Residuos Regresión Logística



Nota. Este histograma muestra la distribución de los residuos de desviación del modelo logístico.

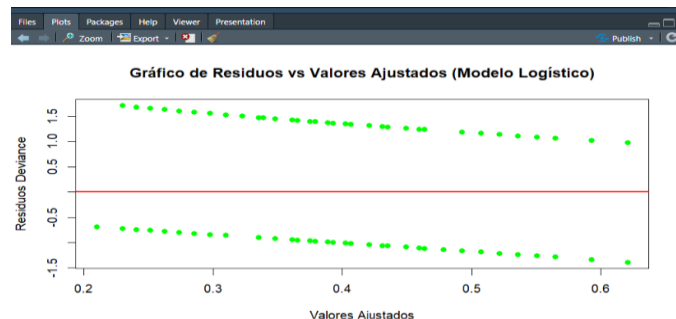
El modelo logístico confirma la influencia significativa de factores como el acceso a Internet, el uso de computadores y la naturaleza de la institución en las probabilidades de

alcanzar un alto rendimiento académico. Aunque los residuos no son normales, esto no afecta la validez del modelo, ya que su propósito es explorar relaciones y no predecir valores absolutos.

Prueba de Homocedasticidad

Figura 44

Prueba de Homocedasticidad Regresión Logística



Nota. Este gráfico de dispersión muestra la prueba de homocedasticidad para el modelo logístico.

En el modelo logístico, el análisis visual del gráfico de residuos vs. valores ajustados nos indica una distribución uniforme de los residuos en torno a la línea cero, lo cual es un indicativo positivo de homocedasticidad. Esto sugiere que las relaciones entre las variables predictoras y la variable dependiente están siendo modeladas de manera consistente, sin variaciones significativas en la dispersión de los residuos.

Multicolinealidad

Figura 45

Multicolinealidad Regresión Logística

```
[1] "VIF para el modelo logístico:"
> print(vif_logistico)
  FAMI_TIENEINTERNET  FAMI_TIENECOMPUTADOR    COLE_NATURALEZA  FAMI ESTRATOVIVIENDA
1.274821             1.269421             1.033210             1.090067
> # Interpretación
> if (any(vif_logistico > 5)) {
+   print("Algunas variables en el modelo logístico tienen problemas de multicolinealidad (VIF >
+ 5). Es necesario realizar ajustes.")
+ } else {
+   print("No se detecta multicolinealidad significativa (VIF <= 5) en el modelo logístico.")
+ }
[1] "No se detecta multicolinealidad significativa (VIF <= 5) en el modelo logístico."
```

Nota. Esta figura muestra el resultado por consola de la multicolinealidad en la regresión logística.

Los VIF obtenidos para el modelo lineal son:

- FAMILIA_TIENEINTERNET: 1.30
- COMPUTADORA FAMILIAR: 1.29
- COLE_NATURALEZA: 1.03
- FAMILIA ESTRATO VIVIENDA: 1.08

Al igual que en el modelo logístico, los valores de VIF para el modelo lineal también están por debajo del umbral de 5. Esto indica que no se detecta multicolinealidad significativa en este modelo. Las variables predictoras son independientes entre sí, lo que asegura la validez de las estimaciones de los coeficientes en el modelo lineal.

Regresión Múltiple

Se considerará para evaluar cómo factores individuales (como el uso de la tecnología) y contextuales (como el tipo de institución) interactúan para influir en los resultados académicos.

Figura 46

Regresión Múltiple

```
> summary(modelo_regresion)
Call:
lm(formula = PUNT_GLOBAL ~ FAMI_TIENEINTERNET + FAMI_TIENECOMPUTADOR +
    COLE_NATURALEZA + FAMI ESTRATOVIVIENDA, data = Data_completa)

Residuals:
    Min       1Q   Median       3Q      Max
-123.070  -31.117   -3.165   28.054  158.711

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    224.7090     1.4921 150.600 < 2e-16 ***
FAMI_TIENEINTERNET  11.7061     0.9579  12.221 < 2e-16 ***
FAMI_TIENECOMPUTADOR  16.1228     0.9468  17.030 < 2e-16 ***
COLE_NATURALEZA     3.6567     1.1824   3.093  0.00199 **
FAMI ESTRATOVIVIENDA -2.1244     0.4864  -4.367  1.27e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 42.75 on 10590 degrees of freedom
Multiple R-squared:  0.06781, Adjusted R-squared:  0.06746
F-statistic: 192.6 on 4 and 10590 DF, p-value: < 2.2e-16
```

Nota. Esta figura muestra el resultado por consola de la regresión múltiple.

Interpretación Coeficientes

Intercepto (224.7090): Este valor representa el puntaje global esperado para un estudiante que no tiene acceso a Internet ni a un computador, que está en un colegio oficial y tiene un estrato social mínimo.

FAMI_TIENEINTERNET (11.7061): Los estudiantes con acceso a Internet tienden a obtener un puntaje global 11.71 puntos mayor, en promedio, en comparación con aquellos sin acceso, manteniendo los demás factores constantes. Este efecto es altamente significativo ($p < 2e - 16$).

FAMI_TIENECOMPUTADOR (16.1228): Los estudiantes con acceso a un computador obtienen, en promedio, 16.12 puntos más en el puntaje global que aquellos sin acceso a un computador, controlando por los demás factores. Este efecto también es altamente significativo ($p < 2e - 16$).

COLE_NATURALEZA (3.6567): Los estudiantes de colegios no oficiales (privados) tienden a obtener 3.66 puntos más en el puntaje global, en promedio, en comparación con los estudiantes de colegios oficiales ($p = 0.00199$). Aunque este efecto es significativo, es más pequeño en comparación con los factores tecnológicos.

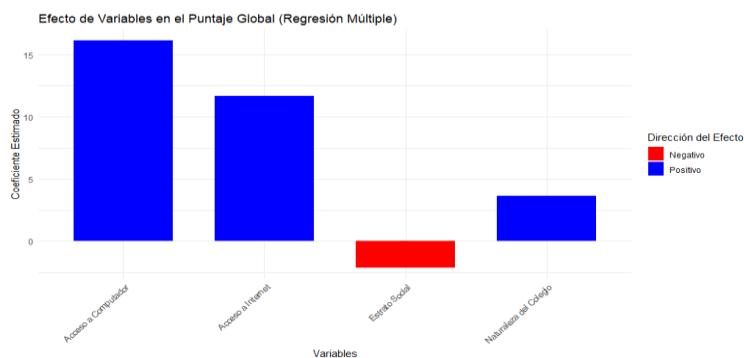
FAMI ESTRATOVIVIENDA (-2.1244): El coeficiente negativo indica que, a medida que el estrato social aumenta, el puntaje global disminuye en promedio en 2.12 puntos por cada unidad adicional de estrato, manteniendo constantes los demás factores. Este resultado es también significativo ($p = 1.27e - 05$), pero va en contra de lo esperado, por lo que podría requerir más análisis para comprenderlo mejor.

Conclusiones: Se concluye que el acceso a la tecnología (Internet y computador) tiene un impacto positivo considerable en los resultados de las pruebas ICFES. Por el contrario, la

naturaleza del colegio y el estrato social muestran un impacto menor, aunque también significativo. La tecnología es un factor clave para mejorar el rendimiento académico, y esto podría sugerir que políticas para aumentar el acceso a estos recursos pueden ser efectivas en elevar los puntajes.

Figura 47

Gráfico Regresión Múltiple



Nota. Esta gráfica de barras muestra el efecto de diferentes variables en el puntaje global según modelo de regresión múltiple.

En este gráfico de regresión múltiple, se visualiza cómo diferentes variables afectan el puntaje global de los estudiantes en las pruebas ICFES. Las barras azules representan efectos positivos, mientras que la barra roja indica un efecto negativo:

Acceso a Computador: Es la variable con el mayor efecto positivo, lo que denota que tener acceso a un computador incrementa significativamente el puntaje global.

Acceso a Internet: También tiene un impacto positivo considerable, aunque algo menor que el acceso a computador. Esto indica que el acceso a internet mejora los resultados en las pruebas.

Estrato Social: Muestra un impacto negativo moderado, lo que indica que a medida que el estrato social disminuye, los puntajes tienden a ser más bajos.

Naturaleza del Colegio: La naturaleza del colegio (si es oficial o no oficial) también tiene un impacto positivo, pero menos pronunciado en comparación con el acceso a tecnología.

Este gráfico confirma que el acceso a herramientas tecnológicas es un factor clave en el desempeño académico.

Conclusiones

A partir del análisis de los datos correspondientes a los años 2019, 2020 y 2021, se identificaron diversos factores que inciden en el rendimiento académico de los estudiantes en las pruebas ICFES Saber 11. Este estudio permitió observar patrones tanto de estabilidad como de variación en los puntajes, asociados a aspectos como el acceso a tecnología, el tipo de institución educativa, el estrato socioeconómico y la modalidad de enseñanza implementada durante el periodo de pandemia. A continuación, se presentan las conclusiones organizadas por cada una de las variables analizadas, con el fin de ofrecer una visión integral y detallada de los hallazgos más relevantes.

Impacto de la Modalidad Educativa en el Rendimiento Académico: La transición de la educación tradicional a las modalidades virtuales e híbridas durante la pandemia de COVID-19 mostró un efecto moderado en los resultados académicos de las pruebas ICFES Saber 11.

Aunque no se observaron cambios extremos en las competencias de matemáticas, inglés y lectura crítica, los resultados nos sugieren que el sistema educativo logró adaptarse a las nuevas modalidades. Sin embargo, esta adaptación también evidencia áreas de oportunidad para fortalecer la infraestructura y los recursos necesarios para una educación a distancia eficaz.

Acceso a Tecnología como Determinante en el Desempeño Académico: El acceso a Internet y a computadores estuvo regularmente asociado con mejores puntajes en las pruebas, lo cual destaca la importancia de las herramientas tecnológicas en el aprendizaje. Por esta razón enfatizamos la necesidad de que los estudiantes tengan acceso adecuado a la tecnología para reducir la brecha de rendimiento, particularmente en épocas atípicas de transición a modalidades educativas no presenciales.

Diferencias en los resultados de las pruebas según la Naturaleza de la Institución: Los estudiantes de instituciones no oficiales obtuvieron resultados superiores en comparación con aquellos de instituciones oficiales. Esto nos indica una disparidad en el acceso a recursos y en el apoyo pedagógico, lo que plantea la necesidad de políticas educativas enfocadas en fortalecer las capacidades y recursos de los colegios oficiales para proporcionar una educación de alta calidad y en igualdad de condiciones.

Influencia del Estrato Socioeconómico: Los hallazgos reflejan que los estudiantes provenientes de estratos socioeconómicos más altos tienden a obtener mejores puntajes en las pruebas ICFES saber 11. Esta relación entre el estrato social y rendimiento académico indica que, aunque la modalidad de enseñanza es un factor importante, las condiciones económicas y sociales también juegan un papel crítico en el desempeño académico. La implementación de programas de apoyo, acceso a educación de calidad y becas podría mitigar estas disparidades y contribuir a una educación más equitativa.

Cumplimiento de los Objetivos del Estudio

Objetivo General

El estudio evaluó satisfactoriamente el impacto de la transición educativa causado por la pandemia en el rendimiento académico, revelando tanto las fortalezas como las áreas de mejora del sistema educativo en el departamento del Caquetá.

Objetivo Específico 1

Se identificaron variaciones en los resultados de las pruebas ICFES entre 2019 y 2021, mostrando una ligera disminución en competencias claves como matemáticas e inglés durante la pandemia. Este análisis evidencia los desafíos de la educación remota, pero también el esfuerzo del sistema educativo por mantener los estándares de calidad.

Objetivo Específico 2

Se evaluó la eficacia de las modalidades virtuales e híbridas, mostrando que, aunque estas modalidades lograron adaptarse, enfrentaron limitaciones debido a la desigualdad en el acceso a recursos tecnológicos. Esto subraya la importancia de fortalecer las estrategias de aprendizaje remoto.

Objetivo Específico 3

El análisis confirma que el acceso a tecnología es un determinante crucial en el rendimiento académico. Los estudiantes con acceso a Internet y computadoras obtuvieron mejores resultados, destacando la necesidad de cerrar la brecha digital en comunidades rurales y de bajos ingresos.

En resumen, este estudio destaca varios factores claves que influyen en el rendimiento académico de los estudiantes en el contexto de la transición educativa causada por la pandemia de COVID-19. La investigación pone en evidencia que el acceso a los recursos tecnológicos

como el Internet y el computador tienen un impacto significativo en los resultados académicos, lo cual nos sugiere que el fortalecimiento de la infraestructura tecnológica en las instituciones educativas podría ser una estrategia fundamental para reducir las brechas de rendimiento.

Asimismo, el análisis comparativo entre instituciones oficiales y no oficiales nos muestra diferencias significativas en los resultados de los estudiantes, lo cual reafirma la necesidad de políticas educativas que busquen equilibrar los recursos y apoyos disponibles en estos contextos. Estas políticas podrían incluir la inversión en capacitación docente, mejoras en infraestructura, y programas de apoyo estudiantil específicos para instituciones oficiales, de manera que los estudiantes tengan oportunidades equitativas sin importar la naturaleza de su institución.

Además, los hallazgos sobre el efecto que tiene el estrato socioeconómico en los resultados de las pruebas refuerzan la importancia de considerar factores de contexto social en el diseño de intervenciones educativas. Las diferencias observadas en el rendimiento según el estrato socioeconómico de los estudiantes sugieren que hay necesidad de implementar programas que mitiguen las barreras económicas al acceso a una educación de calidad.

Finalmente, podemos decir que la resiliencia del sistema educativo durante la pandemia es destacable teniendo en cuenta que no se contaba con un plan de mitigación del riesgo en caso de emergencia, pero también se evidencian aspectos importantes para mejorar en la preparación para modalidades de enseñanza híbridas o remotas. La implementación de políticas y prácticas pedagógicas que optimicen estas modalidades será esencial para enfrentar futuros desafíos y garantizar la continuidad y efectividad del aprendizaje en circunstancias adversas.

Los resultados de este estudio proporcionan un marco de referencia para el diseño de políticas educativas inclusivas y equitativas, que atiendan tanto las necesidades tecnológicas

como las socioeconómicas y de infraestructura, contribuyendo así a la construcción de un sistema educativo más justo y accesible para todos los estudiantes.

Recomendaciones

Promover el Acceso Equitativo a Recursos Tecnológicos: Implementar programas que garanticen el acceso a Internet y dispositivos tecnológicos en colegios de todos los estratos socioeconómicos. Esto incluye alianzas público-privadas para subsidiar el costo de computadoras y conectividad en zonas rurales y urbanas vulnerables. Además, se propone la implementación de programas de subsidios tecnológicos directos para estudiantes de bajos recursos, garantizando así que estos recursos lleguen a quienes más los necesitan.

Fortalecer la Infraestructura de Colegios Oficiales: Se recomienda invertir en el mejoramiento de los recursos y la infraestructura de los colegios oficiales para cerrar la brecha en el rendimiento académico en comparación con los colegios no oficiales. Esto debe incluir planes específicos para la capacitación continua de docentes en herramientas tecnológicas y la dotación de equipos modernos. Además, se recomienda crear fondos gubernamentales dedicados al mantenimiento de infraestructura tecnológica y la adquisición de materiales pedagógicos avanzados.

Desarrollar Estrategias Educativas Inclusivas: Dado el impacto del estrato socioeconómico en el rendimiento académico, es importante diseñar políticas específicas para estudiantes de entornos desfavorecidos. Esto incluye becas de tecnología educativa, programas de apoyo académico focalizados y materiales educativos gratuitos. Se sugiere implementar estrategias como tutorías personalizadas en línea y recursos de aprendizaje digital que puedan ser utilizados incluso sin conexión a Internet.

Fomentar la Adaptación de Modalidades de Aprendizaje Híbridas: Establecer plataformas de aprendizaje digital que sean accesibles para estudiantes y docentes en todo momento, sin importar la modalidad educativa. Se recomienda diseñar programas de preparación para

emergencias educativas, que incluyan simulaciones de aprendizaje remoto, capacitación en plataformas digitales y acceso a herramientas tecnológicas para toda la comunidad escolar.

Implementar Programas de Evaluación y Monitoreo Continuo: Desarrollar sistemas integrales de monitoreo que evalúen regularmente el impacto de las intervenciones tecnológicas y educativas. Estos sistemas deben enfocarse en el análisis de resultados académicos por tipo de institución, área geográfica y estrato socioeconómico, permitiendo así una adaptación constante de las políticas públicas basadas en datos actualizados.

Fomentar la Investigación Educativa para Políticas Basadas en Evidencia: Apoyar investigaciones continuas sobre factores que afectan el rendimiento académico, con un enfoque especial en estudios longitudinales que evalúen el impacto a largo plazo de las intervenciones tecnológicas y educativas. Estudios comparativos por áreas geográficas y sectores socioeconómicos pueden ser cruciales para ajustar estrategias y programas.

Estrategias para Mitigar Brechas Tecnológicas en Colegios Rurales: Implementar programas que promuevan la donación tecnológica que proporcionen computadoras, dispositivos móviles y kits de aprendizaje a colegios rurales. Además, establecer centros comunitarios de acceso a Internet supervisados pedagógicamente, donde los estudiantes puedan utilizar recursos fuera del horario escolar. Los proyectos de conectividad, como el despliegue de redes satelitales y la construcción de torres de comunicación, deben ser una prioridad en comunidades remotas.

Orientación de Resultados para el Diseño de Programas en Zonas Vulnerables: Es importante priorizar la inversión en colegios rurales y zonas urbanas vulnerables, donde los puntajes académicos son más bajos. Diseñar programas integrales de acceso a tecnología que incluyan capacitaciones prácticas para docentes y estudiantes, así como iniciativas sostenibles para el mantenimiento continuo de equipos tecnológicos y redes de conectividad. Se recomienda

la creación de programas piloto regionales que puedan ser escalados a nivel nacional tras la evaluación de su éxito.

Estas recomendaciones no solo abordan las desigualdades tecnológicas y socioeconómicas detectadas, sino que también proponen estrategias concretas y sostenibles para garantizar que todos los estudiantes, sin importar su contexto, tengan acceso a una educación de calidad. Al implementar estas acciones, se busca construir un sistema educativo más inclusivo, equitativo y resiliente ante futuros desafíos.

Referencias Bibliográficas

- Ávila Palacio, L. K., Ospino Gutiérrez, E., & Páez Reales, A. J. (2021). *Análisis de resultados de las pruebas Saber 11 implementando técnicas de minería de datos*.
- Banco Mundial. (2020). *Estrategias para la educación a distancia en Colombia durante la pandemia de COVID19*.
<https://pubdocs.worldbank.org/en/641601599665038137/Colombia-COVID-education-final.pdf>.
- Banco Mundial. (2020). *Resumen de las políticas educativas implementadas en Colombia durante la pandemia de COVID-19*.
<https://pubdocs.worldbank.org/en/143771590756983343/Covid-19-Education-Summary-esp.pdf>.
- Chan, R. Y., Bista, K., & Allen, R. M. (2021). *Impacts of COVID-19 on international students and the future of student mobility: International perspectives and experiences*. Routledge.
<https://scholar.harvard.edu/roychan/publications/impacts-covid-19-international-students-and-future-student-mobility>.
- Demarchi Sánchez, G. D. (2023). *Factores que intervienen en el desempeño de los estudiantes en las Pruebas de Estado Saber Pro*. *El Ágora USB*, 23(2), 490–505.
<https://doi.org/10.21500/16578031.6224>.
- Inciarte González, A., Paredes-Chacín, A. J., & Zambrano Villada, L. M. (2020). *Docencia y tecnologías en tiempos de pandemia COVID-19*. *Utopía y Praxis Latinoamericana*, 25, 195–215. <https://doi.org/10.5281/zenodo.4087411>.

Mendieta, M. C., Cairo, A., & Cairo, O. (2020). *Estrés y afrontamiento ante las clases virtuales en estudiantes universitarios durante la contingencia sanitaria por COVID-19.*

<https://www.redalyc.org>.

UNESCO. (2021). *Educational disruption and response.*

<https://unesdoc.unesco.org/ark:/48223/pf0000380398>.