

Análisis de datos Escuela

Resumen – El presente informe describe el análisis de datos académicos recolectados de una institución educativa, utilizando la herramienta KNIME como apoyo para el procesamiento y exploración de la información. El estudio se centra en la organización, limpieza y análisis de los datos con el fin de identificar patrones relevantes relacionados con el rendimiento y comportamiento estudiantil. Los resultados obtenidos permiten obtener una visión general del contexto educativo y sirven como base para la toma de decisiones orientadas a la mejora de los procesos académicos.

I. INTRODUCCIÓN

El análisis de datos se ha convertido en una herramienta fundamental para la toma de decisiones informadas en distintos ámbitos, incluido el sector educativo. Las instituciones educativas generan grandes cantidades de información relacionada con estudiantes, rendimiento académico, asistencia y características demográficas, la cual puede ser aprovechada para identificar patrones y mejorar los procesos educativos.

En este informe se presenta el análisis de un conjunto de datos recolectados de una escuela, utilizando la herramienta KNIME como plataforma principal para el procesamiento, análisis y visualización de la información. KNIME permite construir flujos de trabajo visuales que facilitan la limpieza, transformación y análisis de datos sin necesidad de programación compleja.

II. DESARROLLO DE CONTENIDOS

La información generada dentro de una institución educativa constituye un recurso fundamental para comprender el comportamiento académico y formativo de los estudiantes. Datos como calificaciones, asistencia y características generales del alumnado reflejan patrones que, si son analizados adecuadamente, pueden aportar una visión más clara sobre el rendimiento y las dinámicas escolares.

Sin embargo, en muchos casos estos datos se encuentran dispersos o son utilizados únicamente con fines administrativos, sin un análisis profundo que permita extraer conclusiones relevantes. Frente a esta problemática, el análisis de datos se presenta como una herramienta clave para transformar registros académicos en información útil para la toma de decisiones.

En este proyecto, el uso de la herramienta KNIME permitió procesar y analizar los datos recolectados de una escuela de manera estructurada y eficiente. A través de flujos de trabajo visuales, fue posible limpiar, organizar y explorar la información, identificando tendencias y relaciones que no son evidentes a simple vista. Este enfoque facilita una comprensión más integral del contexto educativo y contribuye al desarrollo de estrategias orientadas a la mejora del desempeño académico y la gestión escolar.

A. Objetivos y justificación

Objetivo General:

Analizar los datos académicos de una escuela mediante la herramienta KNIME, con el fin de obtener información relevante que apoye la toma de decisiones educativas.

Objetivos Específicos:

- Comprender la estructura y características del conjunto de datos recolectado.
- Aplicar procesos de limpieza y transformación de datos en KNIME.
- Identificar patrones y tendencias en la información académica.
- Generar conclusiones y recomendaciones basadas en los resultados obtenidos.

B. Justificación

El análisis de datos en el ámbito educativo permite transformar información bruta en conocimiento útil para la toma de decisiones. En una escuela, los datos relacionados con estudiantes, rendimiento académico y asistencia constituyen una fuente valiosa para identificar fortalezas, debilidades y oportunidades de mejora en los procesos de enseñanza–aprendizaje.

El uso de la herramienta KNIME se justifica por su capacidad para integrar, limpiar, analizar y visualizar datos mediante flujos de trabajo intuitivos y reproducibles. Esto facilita el análisis sistemático de la información sin requerir conocimientos avanzados de programación, permitiendo enfocarse en la interpretación de los resultados.

Este proyecto es relevante porque aplica conceptos de análisis de datos a un caso real, fortaleciendo las habilidades prácticas adquiridas en la asignatura. Además, los resultados obtenidos pueden servir como apoyo para mejorar la gestión académica y contribuir a una toma de decisiones basada en datos dentro de la institución educativa.

C. Explicación de los datos

Los datos analizados corresponden a información recolectada de una escuela e incluyen variables relacionadas con estudiantes y su desempeño académico. El conjunto de datos fue importado a KNIME desde un archivo estructurado (por ejemplo, CSV o Excel) y posteriormente procesado mediante un flujo de trabajo.

Entre los principales datos analizados se encuentran:

- Información general de los estudiantes (edad, género, curso o nivel).
- Resultados académicos (calificaciones por materia o promedio general).
- Datos de asistencia o registro académico, según disponibilidad del conjunto de datos.

Durante el análisis en KNIME se realizaron las siguientes acciones:

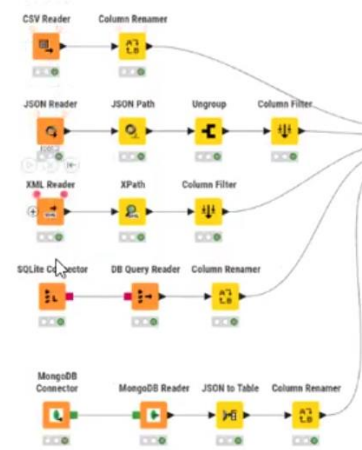
- Limpieza de datos para eliminar registros incompletos o inconsistentes.
- Conversión de tipos de datos para asegurar un análisis correcto.
- Agrupación y filtrado de información para facilitar la interpretación de resultados.
- Generación de estadísticas descriptivas y visualizaciones que permiten identificar tendencias y comportamientos relevantes.

El uso de nodos específicos de KNIME permitió automatizar el proceso de análisis, garantizando un flujo ordenado y reproducible, lo cual es esencial para trabajos de análisis de datos.

D. Capturas de funcionalidad

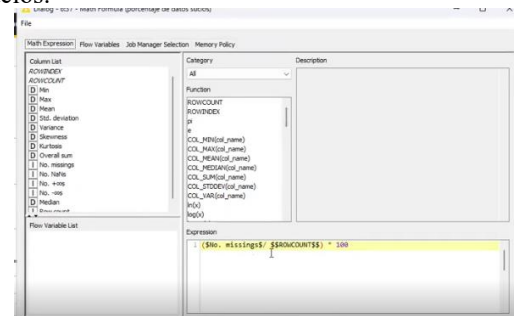
A continuación, explicaremos el funcionamiento, se mostrará media capturas.

1. Los primeros nodos que se observan en la imagen son nodos de lectura para los archivos xml, json, y csv. Los 2 nodos restantes, son para conexión de bases de datos, SQLite y MongoDB
2. Como nodos secundarios tenemos nodos que cambian y convierten los diferentes datos a datos para leerlos con tablas.
3. Un datatlle importante es que todas las columnas deben estar con el mismo nombre, además con el mismo tipo de dato (string, int, float).
4. Como todo lo anterior está listo entonces se parece a concatenar como lo vemos en la imagen



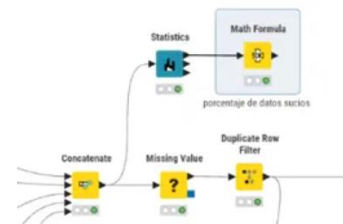
#	RowID	ID	Nombre	Apellido	correo	Sexo	PartTime	Faltas	Extra_acti...	Honestad...	Aspiración	Matemática
129	Row12	122	Jorge	Acosta	jorge.acosta@...	male	true	2	true	19	Accountant	81
125	Row12	123	Laura	Pera	laura.pera@gt...	female	true	2	true	28	Scientist	88
127	Row12	124	Alfredo	Naranjo	alfredo.naranj...	male	true	1	true	26	Software Engin...	76
128	Row12	125	Marina	Campos	marina.campos@...	female	false	3	false	18	Social Worker	72
129	Row12	126	Tomas	Roldan	tomas.rolsan@...	male	true	0	true	31	Data Scientist	94
130	Row12	127	Gabriela	Mendez	gabriela.mende...	female	true	2	true	24	Economist	86
131	Row13	128	Luis	Herrera	luis.herrera@gt...	male	false	4	true	16	Public Policy Ar...	70
132	Row13	148	Mariana	Torres	mariana.torres...	female	false	4	true	19	IoT Researcher	82
133	Row13	150	Fernando	Lozano	fernando.lozan@...	male	true	1	false	26	AI Engineer	94
134	Row13	151	Isabella	Flores	isabella.flores...	female	true	0	true	23	Neuroscientist	97
135	Row13	152	Raul	Herrera	raul.herrera.152...	male	false	8	false	11	Retail Manager	69

5. Luego de concatenar datos, procedemos a usar el nodo Statics, sirven para operaciones matematicas
6. Se usa el nodo Math Formula para contar los valores sucios.

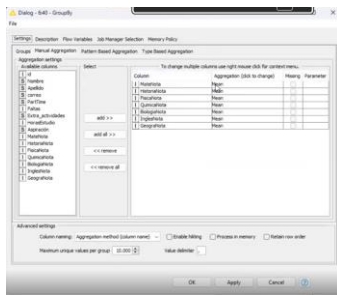


	Variance	Skewness	Kurtosis	Overall no.	No. min.	No. Nulls	No. == 0	No. == 1	No. == 2	Median	Row count	Histogram
125	1.258 033	-0.162	-0.209	6.059	0	0	0	0	0	125	125	
4.865	0.685	0.149	427	1	0	0	0	0	0	125	125	
70.679	-0.278	0.256	2.958	0	0	0	0	0	0	125	125	
103.193	-0.765	0.291	11.258	2	0	0	0	0	0	125	125	
81.92	-0.872	0.847	10.935	2	0	0	0	0	0	125	125	
103.096	-0.156	-1.265	10.947	1	0	0	0	0	0	125	125	
101.677	-0.915	-0.812	10.765	0	0	0	0	0	0	125	125	

7. Luego usamos otro nodo que es missing Value. Sirve para cambiar los valores perdidos o no existentes y convertirlos en 0, ademas usamos otro nodo para filtrar filas repetidas y eliminarlas



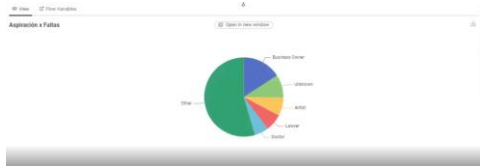
8. Se procede a usar el nodo group by, para agrupar por materia.



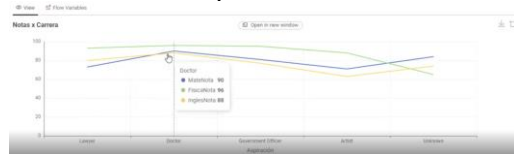
9. La grafica que se muestra es un bar chart



10. Con el nodo de math Formula nos ayudó a crear un "Pai Chart"

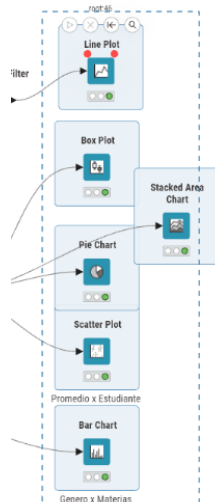


11. De igual manera creamos un line Plot donde se visualiza notas por carrera

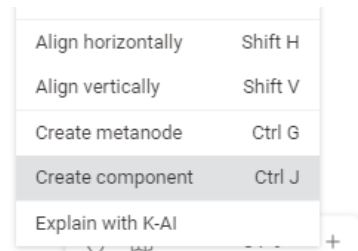


E. Despliegue

A continuación, se adjuntan las capturas que muestran **paso a paso** cómo se realizó el despliegue en KNIME.



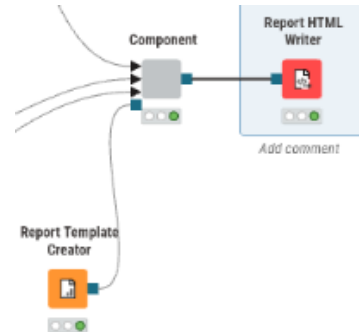
En primer lugar, se deben seleccionar todos los nodos de gráficos implementados. Posteriormente, se hace clic derecho sobre la selección y se elige la opción "Create Component".



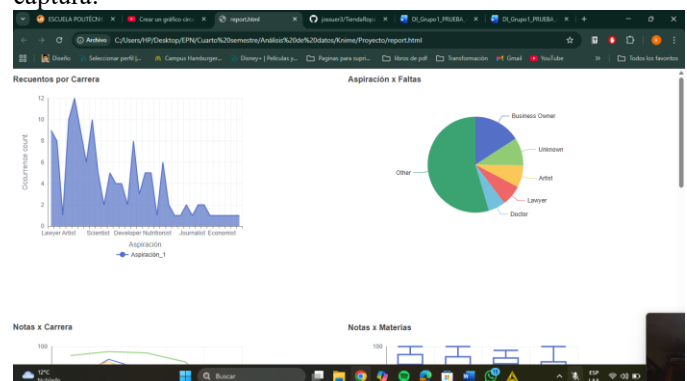
Luego, se realiza la configuración del componente creado, asegurando que los gráficos se ejecuten correctamente. Después, se busca el nodo "Report Template Creator", el cual permite generar un reporte basado en los gráficos proporcionados. Es importante conectar este nodo al componente creado previamente.



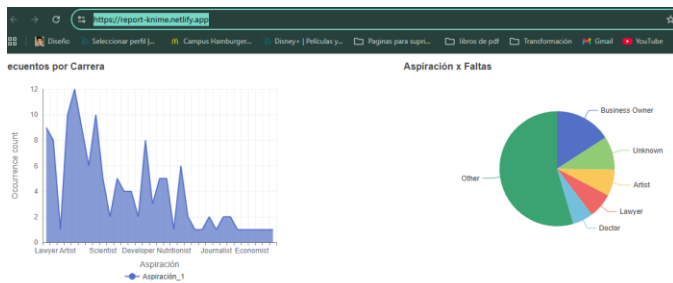
A continuación, el flujo se conecta al nodo "Report HTML Writer", que permite crear un archivo HTML con el reporte generado y almacenarlo en el directorio seleccionado:



Al abrir el reporte generado por el nodo "Report HTML Writer", la visualización es la que se muestra en la siguiente captura.



Finalmente, se realiza el despliegue en Render, lo que permite obtener una URL pública para acceder al reporte.



III. CONCLUSIONES

A partir del análisis realizado, se pudo observar que el uso de herramientas de análisis de datos como KNIME facilita la comprensión de la información académica y permite identificar patrones que no son evidentes a simple vista. El procesamiento visual de los datos ayuda a resumir grandes volúmenes de información de manera clara y estructurada.

El análisis permitió obtener una visión general del rendimiento académico de los estudiantes y de las características del conjunto de datos, lo cual puede servir como base para evaluaciones más profundas o estudios posteriores. Además, se evidenció la importancia de contar con datos bien estructurados y completos para obtener resultados confiables.

F. Recomendaciones

- Mejorar la calidad de los datos recolectados, asegurando consistencia y completitud en los registros.
- Utilizar de forma continua herramientas como KNIME para el análisis periódico de la información académica.
- Complementar el análisis con más variables relevantes que permitan un estudio más detallado del desempeño estudiantil.

Emplear los resultados obtenidos como apoyo para la toma de decisiones educativas y la mejora de procesos internos de la institución.

IV. REFERENCIAS

- Link del despliegue manual: <https://analisis-graficos.netlify.app/>
- Link del reporte generado por KNIME: <https://report-knime.netlify.app/>
- “Student Scores Sample Data (CSV, JSON, XLSX, XML) - Sling Academy,” Slingacademy.com, 2023. Available: <https://www.slingacademy.com/articulo/student-scores-sample-data-csv-json-xlsx-xml/>. [Accessed: Jan. 26, 2026]

- Youtube.com, 2026. Available: <https://www.youtube.com/watch?v=8SCjywiiwUs>. [Accessed: Jan. 26, 2026]
- “KNIME Community Hub,” KNIME Community Hub, 2026. Available: <https://hub.knime.com/>. [Accessed: Jan. 26, 2026]