

---

# Proyecto No. 1

## *Introduccion a la minería de datos*

Milton Rolando Vásquez Hernández

noviembre de 2024

# Informe de proyecto

<b>1. Evaluación de graduandos</b>	<b>2</b>
1.1. Exploración inicial . . . . .	2
1.2. Reglas de asociación . . . . .	7
1.3. Agrupación . . . . .	9
1.4. Propuestas . . . . .	13
<b>2. Documentación</b>	<b>14</b>
2.1. Entorno de trabajo . . . . .	14
<b>Bibliografía</b>	<b>17</b>

Esta documentación corresponde al primer proyecto del curso **Introducción a la Minería de Datos** de la *Maestría en Ingeniería para la Industria con Especialización en Ciencia de la Computación* impartida en la Facultad de Ingeniería de la Universidad de San Carlos de Guatemala, durante el cuarto trimestre del año 2024.

Se solicito al estudiante aplicar los algoritmos *Apriori* y *FP-growth* para encontrar cuatro reglas de asociación con cada e identificar grupos de variables asociadas mediante el algoritmo *K-means*, con el fin de interpretar la información contenida en el conjunto de datos asignado y hacer propuestas de valor en base a los conclusiones extraídas.

El conjunto de datos asignado para este trabajo corresponde a los [resultados de las evaluaciones](#) que realizó el Ministerio de Educación de Guatemala a los graduandos de diversificado al finalizar el ciclo escolar del año 2023.

# Capítulo 1

## Evaluación de graduandos

Se presenta a continuación la sección del informe que contiene la aplicación e interpretación de los algoritmos sobre el conjunto de datos así junto las propuestas elaboradas.

### 1.1 Exploración inicial

#### 1.1.1 Contexto

El Ministerio de Educación de Guatemala (MINEDUC) a través de la Dirección General de Evaluación e Investigación Educativa (DIGEDUCA) realiza anualmente una evaluación a los graduandos de nivel diversificado para medir su rendimiento sus en matemáticas así como en lenguaje y comprensión lectora.

Es una evaluación diagnostica que deben realizar todos los estudiantes que están por graduarse, tiene objetivo de establecer en que medida los estudiantes lograron alcanzar los objetivos de aprendizaje, recolectar información sobre el estado del sistema educativo de Guatemala e identificar líneas de acción para mejorarlo.

Durante la pandemia de COVID-19 no fue posible realizar esta evaluación, en su lugar se utilizaron otras herramientas por lo que no hay disponibles datos para los años 2020 y 2021.

#### 1.1.2 Conjunto de datos

Los datos pueden ser obtenidos en la sección de [resultados de las evaluaciones](#) de la página web del MINEDUC, cada conjunto de datos esta disponible por año, en formato `.xlsx`, `.xls` o `.sav`, junto su diccionario de variables; para esta exploración básica se tomaron los datos de los años 2023, 2022 y 2019.

Utilizando el lenguaje de programación R se cargaron, limpiaron y serializaron los datos, esto último para recuperarlos de forma más rápida en diferentes scripts. Abajo se muestra la cantidad de registros de cada uno de los *datasets* que representan la cantidad de estudiantes evaluados ese año.

Los datos del 2022 difieren en varios aspectos comparado con los otros dos años debido a ser este conjunto al que corresponde a la primera evaluación luego de la pandemia, mas adelante se hablara

de esto, pero se pudo ver que este año se cuenta con registros con una sola evaluación matemática o lectura, posiblemente por esto pueda atribuirse a que un grupo de estudiantes se evaluó únicamente en una sola área matemáticas o lectura..

```
# Datos 2023
[1] "Total estudiantes: 144899"

# Datos 2022
[1] "Total estudiantes: 136144"
[1] "Evaluaciones mate: 135285"
[1] "Evaluaciones lect: 131358"

# Datos 2019
[1] "Total estudiantes: 163825"
```

### 1.1.3 Variables de interés

Los datos para cada año cuentan con una gran cantidad de variables la gran mayoría categóricas, también difiere el número de variables disponibles el 2023 cuenta con 74, el 2022 con 276 y el 2019 con 185. En el 2022 al realizarse la primera evaluación después la pandemia se agregaron variables asociadas a las condiciones de la educación en ese momento, como por ejemplo la forma de acceder a internet o el uso de plataformas virtuales (Zoom, Google Meet, etc.) es debido a ello la mayor diferencia de variables.

Debido a la cantidad de datos disponible, que requiere una cantidad de tiempo y recursos considerable si se desean analizar en profundidad, se seleccionaron algunas variables para limitar la cantidad de información.

Al revisar las descripciones de los datos para cada año se podría clasificar las variables dentro de alguna de las siguientes categorías.

- Identificación (por código de estudiante, establecimientos, etc)
- Ubicación
- Aspectos socioeconómicos
- Aspectos educativos
- Resultados

A continuación se muestran ejemplos de las variables que conforman los, sobre las que podría encontrarse relaciones para comprender mejor los resultados obtenidos en las evaluaciones. Las variables para aplicar los algoritmos solicitados en el trabajo serán escogidas por representar de alguna manera las categorías antes mencionadas.

Variable	Descripción
Logro_Lect	Logro en Lectura
Logro_Mate	Logro en Matemáticas
Desempeño_Mate	Desempeño en Matemáticas
Desempeño_Lect	Desempeño en Lectura
Cod_Region	Código de Región donde se encuentra ubicado el Establecimiento
Jornada	Jornada del Establecimiento
Cod_Sector	Sector del Establecimiento(privado, oficial, ...)
CC_Familia_Automovil	¿Su familia tiene vehículo propio?
Edad_RECO	Edad del Estudiante (Recodificada en Intervalos)
Sexo_RECO	Código Género del Estudiante
Cod_Area	Código del Área del Establecimiento (urbano, rural)
Fm_Grado_Alcanzo_Papa_RECC	Marque el grado más alto que completó su papá.
Fm_Grado_Alcanzo_Mama_REC	Marque el grado más alto que completó su mamá.
Ed_Repitio_Algun_Grado	¿Repitió algún grado?
Sc_Servicio_Internet	Servicios que tiene en su casa: Internet
Tipo_Eval	Tipo de Evaluación (en línea, presencial)

### 1.1.4 Descripción

Mediante la aplicación de la función `summary()` de R, para describir rápidamente los datos, en este caso se obtienen por consola las tablas de frecuencias para las variables, como se muestra a continuación de forma recordada.

Cod_Region	Edad_Reco	CC_Servicio_Internet
Región 1 o Metropolitana:59512	16 años o Menos :18486	No :92627
Región 6 o Suroccidental:35846	17 años :44933	Si :70069
Región 5 o Central :18572	18 años :44144	NA's: 1129
Región 4 o Suroriental :12812	19 años :21464	
Región 3 o Nororiental :11985	20 años :10735	
Región 7 o Noroccidental:10766	Mayor de 20 años:24040	
(Other) :14332	NA's : 23	

### Observaciones

Con las tablas de frecuencia se puede hacer algunas observaciones rápidas sobre la información en distintos años.

- Para los tres años la mayoría de los estudiantes evaluados corresponde a la región Metropolitana.
- La mayoría de los estudiantes tienen una edad entre 17 y 18 años al momento de graduarse.
- Para el año 2019 la mayoría de los estudiante indicaron no tener servicio de internet en su casa, esta variable se invierte para los años 2022 y 2023.

- Para los años 2023 y 2022 la mayoría de los estudiantes evaluados pertenecen al sector privado, también la mayoría para ambos años fuero evaluados en línea. Las variables que indican sector y tipo de evaluación no están presentes en los datos del año 2019.
- Durante los tres años la mayoría de los estudiante evaluados estudiaron en el sector privado.

### 1.1.5 Desempeño

Debido a que el objetivo de este trabajo es aplicar los los algoritmos mencionados previamente para buscar relaciones en los datos, se hace evidente que la mejor forma de abordarlo es buscando relaciones que expliquen los resultados de desempeño de los estudiantes en la prueba o si se logró aprobar la misma.

Por ello a continuación en las figuras 1.1, 1.2 y 1.3 se presentan gráficos de los resultados obtenidos en los tres años para ayudar a orientar la búsqueda de relaciones en la información.

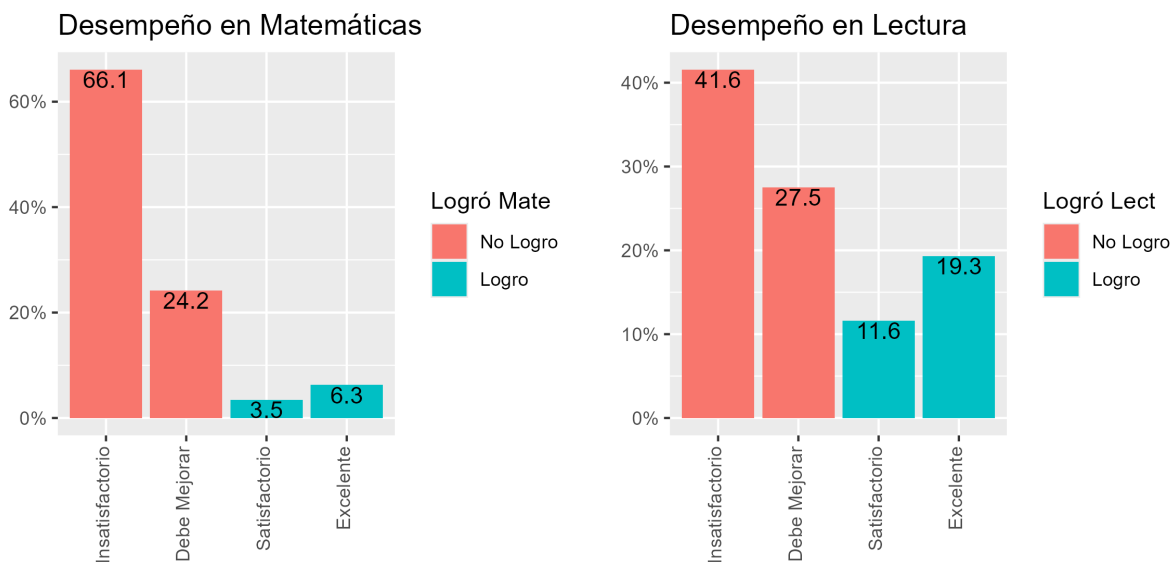


Fig. 1.1.: Resultados de evaluaciones 2023

En los tres años se puede observar como la mayoría de estudiantes obtienen resultados insatisfactorios en ambas evaluaciones siendo el porcentaje de aprobación más alto los para lectura que para matemáticas. Se puede observar como el porcentaje de aprobación en lectura decayó del año 2019 al 2022 con la pandemia de por medio y de 2022 al 2023 bajo el porcentaje de aprobación para matemática.

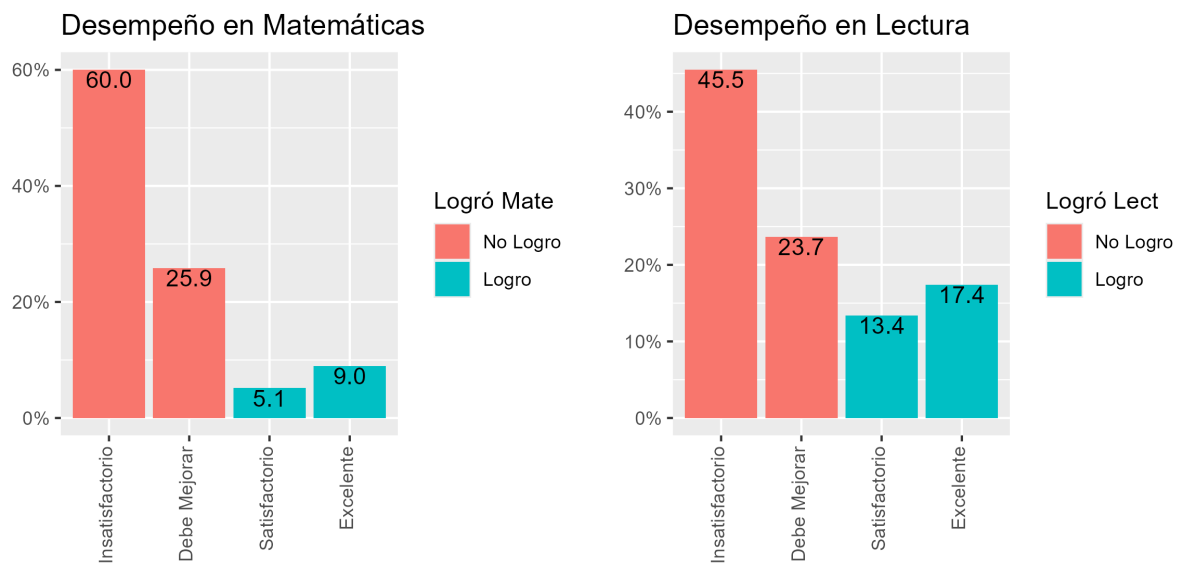


Fig. 1.2.: Resultados de evaluaciones 2022

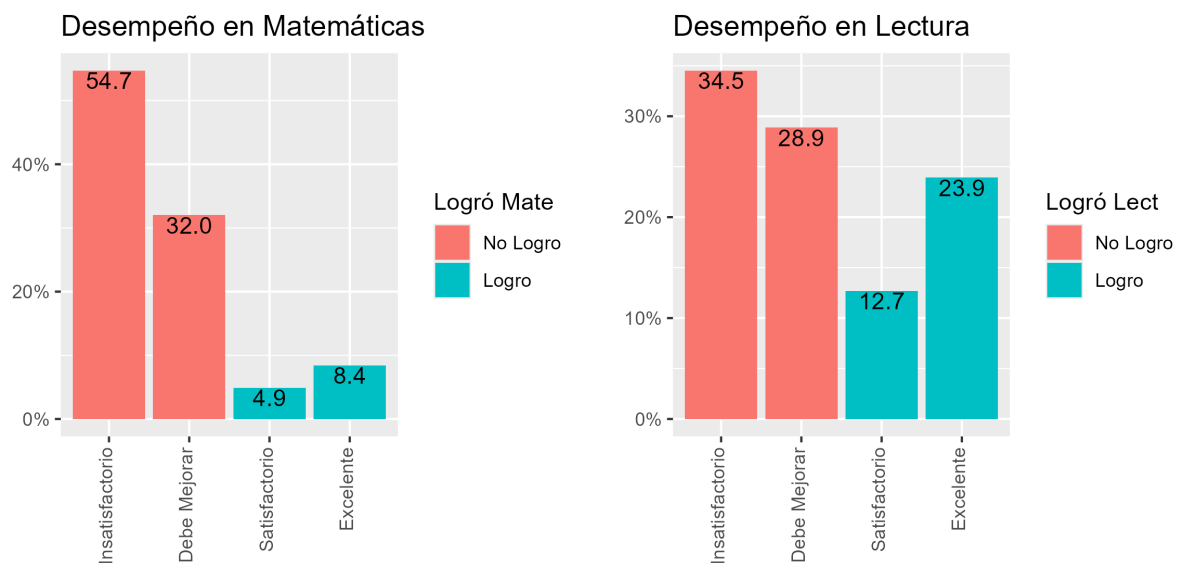


Fig. 1.3.: Resultados de evaluaciones 2019



## 1.2 Reglas de asociación

Dado que la mayoría de los estudiantes fallaron las evaluaciones, se buscaron reglas de asociación primero con los datos sin segmentar y luego enfocándose en los estudios que obtuviera buenos resultados para buscar asociaciones que explique su existencia en la prueba.

### 1.2.1 Apriori

Al buscar reglas de asociación con los datos sin segmentarlos, debido a la diversidad y cantidad de información, las reglas con mayor soporte que encuentra el algoritmo son simples y que se podrían buscar por otros medios. Aquellas reglas más complejas que asocian varias variables tienen poco soporte pero aún así aportan algo de información sobre los datos.

Debido a la mayoría de los estudiantes no lograron buenos resultados en ambas pruebas, el algoritmo sin filtrar, asocia distintas variables con un resultado negativo en matemáticas más que con resultado negativo en lectura dado que los estudiantes fallan más frecuentemente la prueba de matemática que la de lectura.

#### ■ Regla 1

```
# Año 2023
{Sc_Electricidad=Si} => {Cod_Area=Urbana}
0.8214097 0.8744000 0.9393981 1.0045790 113679

{Ed_Asistio_Preprimaria=Si, Sc_Electricidad=Si} => {Logro_Mate=No Logro}
0.7404964 0.8933765 0.8288739 0.9922223 102481
```

Reglas como las mostradas arriba son las que indican mayor soporte el conjunto de datos para el año 2023, sabiendo que la mayoría de los estudiantes evaluados están ubicados en la región metropolitana o en una área urbana es de esperarse que la mayoría cuente con servicios básicos como la electricidad. La siguiente es una asociación un poco inesperada sobre haber cursado preprimaria y la falla en la prueba en matemáticas. Debido al alto porcentaje de fallo en la prueba (90 %) no se puede afirmar con certeza y habría que realizar más análisis en esa dirección pero se podría hacer la observación de que haber iniciado vida estudiantil a temprana edad no refleja un mejor resultado en la evaluación.

#### ■ Regla 2

```
# Año 2023
{Ed_Asistio_Preprimaria=Si, Sc_Electricidad=Si, Computadora=Si,
Cod_Area=Urbana, Sc_Servicio_Internet=Si, Grad_Seguir_Estudiando=SI}
=> {Logro_Mate=No Logro}
0.3256259 0.8412515 0.3870732 0.9343300 4506
```

Entre las reglas que asocian una mayor cantidad de variables pero con menos soporte, se muestra esta regla con un conjunto de variables que se podría pensar estaría asociadas a un mejor resultado pero más bien están asociadas a un resultado negativo. Nos muestra que una parte de los estudiantes a pesar de contar con algunos recursos tecnológicos (computadora, acceso a internet) y con la intención de continuar sus estudios obtuvieron un resultado negativo.

■ Regla 3

```
{Ed_Asistio_Preprimaria=Si, Computadora=Si, Sc_Servicio_Internet=Si,  
Grad_Seguir_Estudiando=SI, Lectura_Gusto=Si} => {Cod_Area=Urbana}  
0.3474337 0.8832192 0.3933720 1.0648264 2579
```

Al filtrar los datos para el sector *Oficial* (Estatual) y para buscar asociaciones dentro del grupo de alumnos que aprobó el examen de lectura resalta la asociaciones sobre el area del establecimiento, como la mostrada que no indica que sabiendo que el estudiante logró aprobar la evaluación de lectura, con acceso a algunos recursos tecnológicos y con la intención de seguir estudiante pertenece al área urbana.

■ Regla 4

```
{Logro_Lect=Logro} => {Computadora=Si}  
0.6850520 0.8561439 0.8001599 1.0418638 857  
  
{Computadora=Si,Cod_Area=Urbana,Sc_Servicio_Internet=Si,  
Grad_Seguir_Estudiando=SI} => {Logro_Lect=Logro}  
0.4700240 0.8724036 0.5387690 1.0902866 588
```

Buscando relaciones con los estudiantes que ganaron la prueba de matemática y pertenecen al sector *Oficial* se puede ver como el acceso a la tecnología pueda favorecer obtener un buen resultado en la prueba, también que en general los estudiantes que obtuvieron un buen resultado en matemática también lo hicieron en lectura pero esta asociación no existe en la otra dirección.

### 1.2.2 FP-Growth

Al igual que sucede con el algoritmo apriori las reglas obtenidas con FP-Growth que tienen un mayor confianza y soporte son evidente o aportan poca información adicional.

■ Regla 1

```
{Logro_Lect=No Logro} => {Logro_Mate=No Logro}  
0.6672640 0.9757402 1.0836990 92346 1026
```

Esta regla muestra como hay asociaciones entre fallar las pruebas, parece implicar que si un alumno no es capaz de aprobar la prueba de lectura seguramente falle en la de matemáticas. Los datos muestran como hay un porcentaje de probación mas alto en la prueba de lectura, podría indicar que aprobar esta prueba es más fácil que la de matemáticas.

■ Regla 2

```
{Ed_Asistio_Preprimaria=Si, Cod_Sector=Privado,  
CC_Material_Predomina_En_Piso_Casa=Piso cerámico} => {Cod_Area=Urbana}  
0.3630333 0.9225317 1.0598765 50242
```

Con datos sin segmentar se encuentra esta asociación. Se sabe que la mayoría de los estudiantes que se someten a la evaluación pertenecen al sector privado y con uns cuantas condiciones

adicionales en general su establecimiento se encuentra ubicando en un área urbana. Dentro de la gran cantidad de reglas que muestra FP-Growth muchas relación las condiciones de los estudiantes con el sector privado y el área urbana indicando que existe poca representación de los otro sectores y del area rural en el conjunto de dato.

#### ■ Regla 3

```
Ed_Trabaja_Actualmente=No, Sc_Servicio_Internet=Si, Identificacion_Etnica_
↪ RECO=Ladino}
=> {Ed_Repitio_Algun_Grado=No}
0.3600970 0.8129562 1.0822407 2673
```

Filtrando los datos, tomando solo estudiantes que aprobaron lectura y pertenecen al sector **oficial**, surge la asociación que nos indica con un soporte bajo que si un estudian aprueba el examen de lectura y además cuenta con unos recursos mínimos (económicos y tecnológicos) seguramente nunca repitió un año escolar.

#### ■ Regla 4

```
{Logro_Lect=Logro, Ed_Asistio_Preprimaria=Si, Ed_Trabaja_Actualmente=No,
Computadora=Si} => {Ed_Repitio_Algun_Grado=No}
0.3501199 0.8358779 1.0995617 438
```

Nuevamente esta regl, como una condición de aprobación de matemática y pertenencia al sector **Oficial**, nos indica que si un estudiante aprobó ambas evaluaciones y además cuenta con unos recursos mínimos (económicos y tecnológicos) no repitió un año escolar. Estas dos últimas reglas muestra asociación entre entre la variable de renitencia y la aprobación tanto en la prueba de lectura como en la matemática.

## 1.3 Agrupación

Como en su mayoría las variables con las que cuenta el conjunto de datos son de tipo categórico, preguntas SI /NO o selección de la categoría más cercana, el algoritmo K-Means no es el más adecuado para generar clústeres, este algoritmo espera tipos de datos numéricos como parámetros de ingreso y se base en calcular la distancia euclidiana entre puntos lo cual no tiene mucho sentido para datos categóricos.

### 1.3.1 K - Means

En figura 4 se puede observar una gráfica con las únicas variables numéricas en los datos al aplicar K-Means junto con otras variables (convertidas a enteros) se puede ver que estas formaron clústeres muy cercanos al mismo punto. No es posible obtener mayores conclusiones de esta gráfica mas allá habilidad estimada para la mayoría de estudiante esta cerca del valor cero en ambas materias.

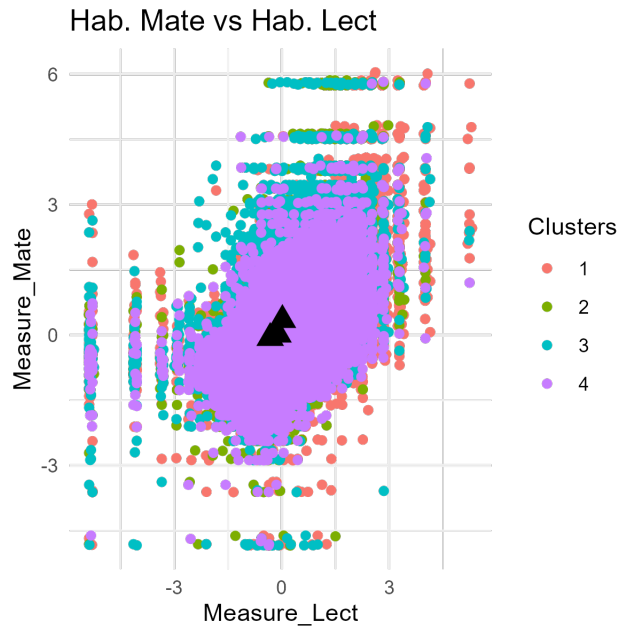


Fig. 1.4.: Agrupación mediante K-Means

### 1.3.2 K - Modes

Se utilizó el algoritmo K-Modes que esta diseñado para trabajar con datos categóricos ya que utiliza métricas de distancia adecuadas para datos categóricos, este calcula el centroide como el valor más frecuente para cada atributo en el cluster.

En las figuras 5 y 6 se puede ver los resultados de K-Modes en la primera se observa como los datos forma un grupo con que contiene a los estudiantes con mejor desempeño con los datos sin segmentar.

En la segunda se aplico un filtro para utilizar solamente los datos de los estudiante que ganaron alguna de las evaluaciones (logro en mate o en lectura) se puede ver como los datos nuevamente logran separar a los estudiantes con los mejores resultados.

### Observaciones

Al observar la tabla 1 que muestra los cinco modos que fue capas de encontrar el algoritmo sobre la los datos, se puede hacer algunas observaciones como:

- Los estudiantes con obtiene resultados tienden a tener padres con educación (diversificada o universitaria).
- Los estudiantes que logran aprobar ambos evaluaciones tienden a indicar que cuenta con cinco periodos de clase en ambas materias a la semana.
- Por lo general los estudiantes que aprueban alguna evaluación están ubicando en la Ciudad Capital. \

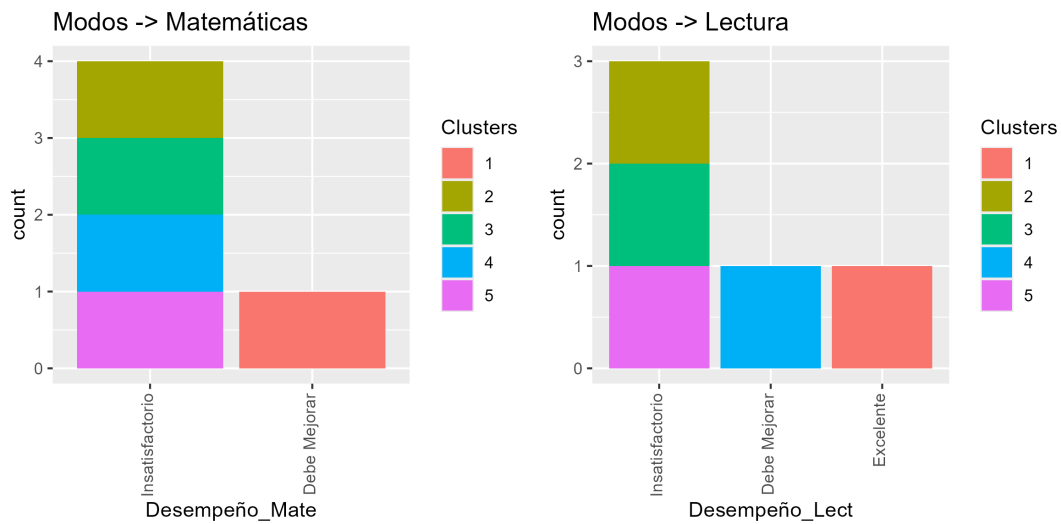


Fig. 1.5.: K-Modes con datos sin segmentar

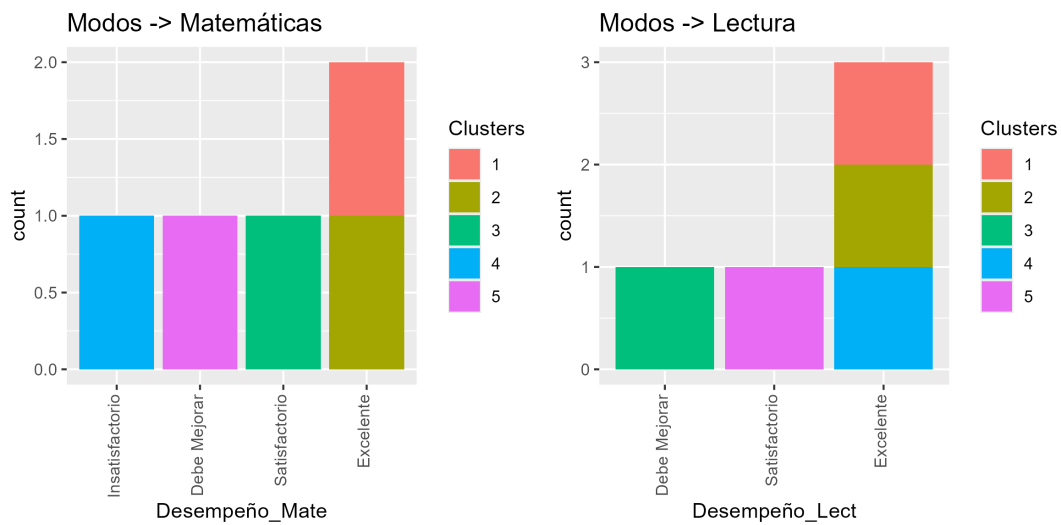


Fig. 1.6.: K-Modes con estudiantes que obtuvieron al menos un logro

Tabla 1.1.: Agrupaciones de estudiantes que obtuvieron al menos un logro

Variable	g1	g2	g3	g4	g5
Logro_Mate	Logro	Logro	Logro	No Logro	No Logro
Logro_Lect	Logro	Logro	No Logro	Logro	Logro
Ed_Astistio_Preprimaria	Si	Si	Si	Si	Si
Sc_Electricidad	Si	Si	Si	Si	Si
Desempeño_Mate	Excelente	Excelente	Satisfactorio	Insatisfactorio	Debe Mejorar
Desempeño_Lect	Excelente	Excelente	Debe Mejorar	Excelente	Satisfactorio
Cod_Depa	Ciudad Capital	Ciudad Capital	Ciudad Capital	Guatemala	Ciudad Capital
Lect_Periodos_Lectura_Ser	5 Períodos	Ninguno	1 Período	Ninguno	Ninguno
Ma-te_Periodos_Matematicas_1	Cinco periodos	Cuatro periodos	Cinco periodos	Ninguno	Cinco periodos
Jornada	Matutina	Matutina	Matutina	Vespertina	Matutina
Ed_Trabaja_Actualmente	No	No	No	No	No
Cod_Sector	Privado	Privado	Privado	Privado	Privado
CC_Material_Predomina_E	Piso cerámico	Torta de cemento	Piso cerámico	Torta de cemento	Piso cerámico
Computadora	Si	Si	Si	Si	Si
CC_Familia_Automovil	Si	No	Si	Si	Si
Edad_RECO	17 años	18 años	17 años	18 años	17 años
Sexo_RECO	Masculino	Femenino	Masculino	Femenino	Masculino
Cod_Area	Urbana	Urbana	Urbana	Urbana	Urbana
Sc_Servicio_Internet	Si	Si	Si	Si	Si
Fm_Grado_Alcanzo_Papa_	Universidad o más	Diversificado	Primaria	Primaria	Universidad o más
Fm_Grado_Alcanzo_Mama	Diversificado	Primaria	Primaria	Primaria	Universidad o más
Grad_Trabajar	SI	SI	SI	SI	SI
Grad_Seguir_Estudiando	SI	SI	SI	SI	SI
Lectura_Gusto	Si	Si	Si	Si	Si
Identificacion_Etnica_RECO	Ladino	Ladino	Ladino	Ladino	Ladino
Lect_Libros_Completos_Ha	Dos libros	Dos libros	Un libro	Un libro	Dos libros
Tipo_Eval	Linea	Linea	Linea	Linea	Linea
Ed_Repitio_Algun_Grado	No	No	No	No	No

## 1.4 Propuestas

---

**Nota:** Debido a que la mayoría en los estudiantes en Guatemala logran conseguir los resultados mínimos esperados en las pruebas de lectura y matemática cuando finaliza su formación estudiantil obligaría siendo este un problema grave que afecta a la población Guatemalteca, sería necesario un análisis más profundo de la información para realizar propuesta con un mejor respaldo.

---

1. La evaluación diagnostica debería contar con más variables que se centren en aspectos educativo y que puedan mediar la calidad de la educación de los establecimiento. La evaluación del año 2022 es la más sustanciosa en ese sentido dado que contenía variables como las estrategias de aprendizaje o si el estudiante recibía retroalimentación. Estas no se encuentran en los datos para los años 2023 o 2019.
2. La mayoría de la educación a nivel diversificado en las áreas urbanas es impartida por el sector privado, es necesaria una forma de evaluación que incluya a los colegios y no solamente al estudiante, un método adicional que permita medir la calidad educativa impartida por los colegios (no solo a través de los resultados obtenidos por los estudiantes) para se garantice que cumplen con los mínimos requeridos.
3. Los estudiantes que obtiene mejores calificaciones por lo general cuentan con el soporte necesario para su educación, cuentan con recursos, se puede ver por lo general los estudiantes que logran un desempeño excelente en ambas evaluaciones tienen padres con educación de diversificado o superior. Pero hay otra como la cantidad de periodos de clase que reciben semanalmente en la materia también asociada a un buen desempeño, asegurando entonces un mínimo de periodos de clase semanales a nivel general para cada una de las materias se podría ver una mejora.

## Capítulo 2

# Documentación

Esta sección contiene información técnica sobre las librerías utilizadas y el entorno de trabajo, para poder ejecutarlo.

## 2.1 Entorno de trabajo

### 2.1.1 Librerías R

#### 1. dplyr

Es una librería de R para la manipulación de datos. Ofrece funciones para filtrar, seleccionar, modificar y resumir datos de manera eficiente y legible.

```
install.packages("dplyr")
library(dplyr)

# Filtrar y seleccionar columnas
data(mtcars)
mtcars %>%
  filter(mpg > 20) %>%
  select(mpg, hp)
```

#### 2. here

La librería ayuda a gestionar rutas de archivos en proyectos R, facilitando la creación de rutas relativas desde el directorio principal del proyecto.

```
install.packages("here")
library(here)
```

(continúe en la próxima página)



(proviene de la página anterior)

```
# Obtener la ruta del archivo 'data.csv' en el directorio 'data'
file_path <- here("data", "data.csv")
```

### 3. foreign

Es una librería que permite leer y escribir archivos de datos en formatos de otros programas estadísticos como SPSS, SAS y Stata.

```
install.packages("foreign")
library(foreign)

# Leer un archivo .sav
data <- read.dta("path/to/file.sav")
```

### 4. patchwork

Es una librería que facilita la combinación de varios gráficos de `ggplot2` en una sola visualización.

```
install.packages("patchwork")
library(ggplot2)
library(patchwork)

# Crear dos gráficos y combinarlos
p1 <- ggplot(mtcars, aes(x = mpg, y = hp)) + geom_point()
p2 <- ggplot(mtcars, aes(x = mpg, y = wt)) + geom_point()

p1 + p2
```

### 5. klaR

Es una librería que proporciona herramientas para el análisis de datos, incluyendo clasificación, análisis discriminante y visualización.

```
install.packages("klaR")
library(klaR)

# Ejemplo:
modes <- kmodes(categorias, 4)
View(modes)
```

### 2.1.2 Docs

#### Sphinx

Es una herramienta de documentación para proyectos en Python, que permite generar documentación estática. Utiliza reStructuredText como su formato principal, permitiendo generar documentación en varios formatos (HTML, Latex, etc.) y soporta características avanzadas agregando extensiones.

#### MyST

Es una extensión de Sphinx que permite escribir documentación utilizando Markdown. MyST facilita la creación de documentación sencilla y accesible, mientras mantiene las características de Sphinx para generar documentación estructurada y dinámica

# Bibliografía

- [1] Ana Lucía Ola. *Esto dicen las pruebas del Mineduc acerca de las habilidades de los graduandos en Lectura y Matemáticas*. Prensa Libre, 2023. <https://www.prensalibre.com/guatemala/comunitario/esto-dicen-las-pruebas-del-mineduc-acerca-de-las-habilidades-de-los-graduandos-en-lectura-y-matematicas/>. [Accedido: 9 de noviembre de 2024].
- [2] Dirección General de Evvaluación e Investigación Educativa. *Resultados generales de la evaluación educativa*. 2023. [https://edu.mineduc.gob.gt/digeduca/documents/resultados/Resultados\\_generales.pdf](https://edu.mineduc.gob.gt/digeduca/documents/resultados/Resultados_generales.pdf). [Accedido: 9 de noviembre de 2024].
- [3] Ministerio de Educación de Guatemala. *Resultados de evaluaciones*. 2023. <https://edu.mineduc.gob.gt/digeduca/?p=resultadosevaluacionesMain.asp>. [Accedido: 9 de noviembre de 2024].