

# Trabajo de Investigación Prueba SIMCE

Romina Mercado

9 de noviembre de 2020

## Introducción

**Disclaimer:** Cambié mi tema pues no pude manejar las base de datos de PISA ya que eran muy numerosas y mi pc colapsaba.

La Prueba del Sistema de Medición de la Calidad de la Educación (SIMCE), es una herramienta que permite al Estado de Chile a través de su Agencia de Calidad de la educación, medir el nivel de dominio de conocimientos de los escolares Chilenos del plan de estudio vigente. Esta prueba fue diseñada en la época del regimen militar con otro nombre y vigente hasta el día de hoy como “SIMCE”

Esta prueba tiene como finalidad conocer las necesidades educativas del país y ayudar a mejorar la calidad y equidad de la educación, para esto, mide los conocimientos de los estudiantes en las áreas de lectura, matemática y/o ciencia, en alumnos de los niveles: cuarto básico, sexto básico, segundo medio y tercero medio. Además de la rendición de la prueba, se aplican encuestas para alumnos y padres relacionada con el ámbito social, psicologico y socioeconómico de los estudiantes.

Los datos obtenidos a partir de la prueba SIMCE son muy relevantes, pues nos permite visualizar o analizar cuales son los factores más importantes en cuanto a la capacidad de desarrollar aprendizajes y aplicar estas competencias en la evaluación, y por ende podemos analizar el sistema educativo del país de Chile. Al evaluar esto, permite tomar decisiones importantes en cuanto a educación y políticas públicas asociadas a este tema. Mejorar la educación aporta beneficios económicos y sociales dentro del país y del mundo, además del desarrollo personal de cada ciudadano.

A partir de lo anterior, y con los datos de los cuartos básicos, de los años 2016, 2017 y 2018, en las pruebas de Lectura y Matemática, vamos a realizar un análisis descriptivo para analizar los rangos de puntajes obtenidos y su asociación y correlacion con distintas variables categóricas relacionadas al ambito social obtenidas de las encuestas, y además de un regresión lineal, para determinar las variables mas importantes que explican los resultados obtenidos en la Prueba SIMCE, estos analisis estadísticos nos permitirán obtener información importante para los objetivos del informe, por ejemplo: ¿Qué factores sobre el ambito social del estudiante afecta en el aprendizaje y aplicación del aprendizaje?

El siguiente informe comenzará explicando la conformación de las variables y una explicación de su significado, para luego explicar los distintos análisis de medias, y correlaciones que se realizarán, para concluir con la realización de una regresión lineal buscando el modelo que mejor ajuste para estos datos.

## Resultados

Los datos están conformados por 11 variables, de las cuales hay 27410 observaciones, en la siguiente tabla resumen es posible ver el contenido de cada variable:

Variable	Tipo de variable	Descripción
----------	------------------	-------------

-	-	-
---	---	---

Agno | numérica | Año de rendición de la prueba simce |  
 Grado | caracter | Nivel educación evaluado |  
 rbd | numérica | Código que identifica cada establecimiento educacional |  
 Grupo\_socioeconomico | Caracter | Grupo socioeconomico |  
 Ptje Lectura | numérica | Puntaje promedio obtenido por el establecimiento en la prueba de lectura |  
 Ptje Matematica | numérica | Puntaje promedio obtenido por el establecimiento en la prueba de matemática |  
 Autoestima | numérica/factor | Medición de la motivación y autoestima de los estudiantes de 0 a 100 |  
 Convivencia | numerica/factor | Medición del Clima de convivencia escolar del establecimiento de 0 a 100 |  
 Habitos Saludables | numerica/factor | Medición de hábitos saludables en el establecimiento de 0 a 100 |  
 Formacion Ciudadana | numérica/factor | Medición de la participacion y formacion ciudadana en el establecimiento |  
 Desempeno | Categórica | Criterio de Desempeño del establecimiento |

```

#Valores para tabla
min_Lec <- min(datos$PtjeLectura) #149
min_Mat <- min(datos[datos$PtjeMatematica>0,"PtjeMatematica"], na.rm=T) #129

#maximos puntajes
max_Lec <- max(datos$PtjeLectura) #396
max_Mat <- max(datos$PtjeMatematica) #376

#promedio lectura

media_Lec <- mean(datos$PtjeLectura) #261.8559

#promedio matematica

media_Mat <- mean(datos$PtjeMatematica) # 261.8559

#Promedio Desempeño MATESMATICAS

D_ins_Mat <- mean(datos$PtjeMatematica[datos$Desempeno=="INSUFICIENTE"]) #[1] 228.4617
D_Medbaj_Mat<- mean(datos$PtjeMatematica[datos$Desempeno=="MEDIO-BAJO" | datos$Desempeno=="MEDIO-BAJO (NU
D_medio_Mat <- mean(datos$PtjeMatematica[datos$Desempeno=="MEDIO"]) #[1] 255.6693
D_alto_Mat <- mean(datos$PtjeMatematica[datos$Desempeno=="ALTO"]) # [1] 280.4555

#Promedio Desempeño Lectura

D_ins_Lec <- mean(datos$PtjeLectura[datos$Desempeno=="INSUFICIENTE"]) #[1] 240.3822

D_medbaj_Lec <- mean(datos$PtjeLectura[datos$Desempeno=="MEDIO-BAJO" | datos$Desempeno=="MEDIO-BAJO (NU
D_medio_Lec <- mean(datos$PtjeLectura[datos$Desempeno=="MEDIO"]) #[1] 265.808
  
```

```
D_alto_Lec <- mean(datos$PtjeLectura[datos$Desempeno=="ALTO"]) # [1] 286.296
```

```
##Promedio puntaje segun nivel socioeconomico MATESMATICAS
```

```
mean(datos$PtjeMatematica[datos$Grupo_socioeconomico == "Bajo"]) # [1] 235.5431
```

```
## [1] 222.9445
```

```
mean(datos$PtjeMatematica[datos$Grupo_socioeconomico=="Medio bajo" | datos$Desempeno=="MEDIO-BAJO (NUEVO)"])
```

```
## [1] 245.4342
```

```
mean(datos$PtjeMatematica[datos$Grupo_socioeconomico=="Medio"]) # [1] 256.6302
```

```
## [1] 254.385
```

```
mean(datos$PtjeMatematica[datos$Grupo_socioeconomico=="Medio alto"]) # [1] 272.7193
```

```
## [1] 270.7688
```

```
mean(datos$PtjeMatematica[datos$Grupo_socioeconomico=="Alto"]) # [1] 290.5961
```

```
## [1] 288.7167
```

```
##Promedio puntaje segun nivel socioeconomico Lectura
```

```
mean(datos$PtjeLectura[datos$Grupo_socioeconomico == "Bajo"]) # [1] 253.0105
```

```
## [1] 239.4039
```

```
mean(datos$PtjeLectura[datos$Grupo_socioeconomico=="Medio bajo" | datos$Desempeno=="MEDIO-BAJO (NUEVO)"])
```

```
## [1] 257.6539
```

```
mean(datos$PtjeLectura[datos$Grupo_socioeconomico=="Medio"]) # [1] 265.0224
```

```
## [1] 262.5592
```

```
mean(datos$PtjeLectura[datos$Grupo_socioeconomico=="Medio alto"]) # [1] 280.7158
```

```
## [1] 278.6803
```

```
mean(datos$PtjeLectura[datos$Grupo_socioeconomico=="Alto"]) # [1] 296.0503
```

```
## [1] 294.7019
```