

Principios de Estadística Computacional

Docentes: Guadalupe González, Lucas Peitton

Institución: *FLACSO*

Cohorte 2025/Año: 2026

## **¿La brecha digital es también educativa?**

Evidencia empírica con learningtower (PISA) - Chile (2018-2022)

Autora: Gabriela Viola

Fecha: 2026

# Índice

<b>1. Introducción</b>	<b>1</b>
1.1. Problema y motivación . . . . .	1
1.2. Objetivos . . . . .	1
1.3. Datos, alcance y unidad de análisis . . . . .	1
<b>2. Datos y metodología</b>	<b>1</b>
2.1. Fuente de datos y preparación . . . . .	1
2.2. Variables y definiciones . . . . .	1
2.3. Recodificaciones y faltantes . . . . .	2
<b>3. Resultados</b>	<b>2</b>
3.1. Análisis exploratorio . . . . .	2
3.1.1. Resumen descriptivo por año . . . . .	2
3.1.2. Distribución de matemática por año . . . . .	3
3.1.3. Distribución de computadoras en el hogar por año . . . . .	3
3.1.4. Comparación (Matemática) entre 2018 y 2022 . . . . .	6
3.2. Análisis de correlación . . . . .	6
3.3. Análisis de regresión lineal . . . . .	7
3.3.1. Modelo 1: regresión lineal simple . . . . .	7
3.3.2. Modelo 2: sensibilidad al controlar por escs . . . . .	8
3.3.3. Comparación de ajuste (Modelo 1 vs Modelo 2) . . . . .	8
3.4. Evaluación de supuestos y sensibilidad . . . . .	9
3.4.1. Robustez de la inferencia (errores estándar robustos) y no linealidad . . . . .	14
3.5. Visualización de la relación (dispersión y recta de regresión) . . . . .	14
3.6. Pruebas de hipótesis (por disponibilidad/periodicidad) . . . . .	14
3.6.1. Prueba de medias: cambio en matemática (2018 vs 2022) . . . . .	14
3.6.2. Prueba de proporciones: cambio en acceso a internet (2018 vs 2022) . . . . .	16
<b>4. Discusión y limitaciones</b>	<b>17</b>
<b>5. Conclusiones</b>	<b>17</b>
<b>6. Referencias</b>	<b>17</b>
<b>7. Anexo (reproducibilidad)</b>	<b>17</b>

# 1. Introducción

## 1.1. Problema y motivación

Este trabajo evalúa si la brecha digital (recursos tecnológicos disponibles en el hogar) se asocia con una brecha educativa (rendimiento en matemática). La pregunta que organiza el análisis es: **¿La brecha digital es también educativa?**

## 1.2. Objetivos

**Objetivo general.** Analizar la asociación entre disponibilidad de recursos digitales en el hogar y rendimiento en matemática, utilizando datos PISA muestras (learningtower) para Chile en 2018 y 2022.

**Objetivos específicos.**

1. Describir la distribución de rendimiento (`math`) y recursos digitales (`computer_n_num`, `internet01`) por año.
2. Estimar y comparar (i) correlación y (ii) regresión lineal entre recursos digitales y matemática, evaluando sensibilidad al controlar por estatus socioeconómico y cultural (`escs`).

## 1.3. Datos, alcance y unidad de análisis

Se utiliza el paquete `learningtower`, que contiene subconjuntos de datos PISA muestras por año. En esta versión se encuentran disponibles los objetos `student_subset_2018` y `student_subset_2022`, por lo que el análisis se realiza para Chile (CHL) en 2018 y 2022.

**Nota conceptual (población vs muestra).** Los resultados se interpretan como evidencia sobre la muestra disponible (subconjunto muestras) y no como parámetros exactos de la población total.

# 2. Datos y metodología

## 2.1. Fuente de datos y preparación

El conjunto analizado incluye 100 observaciones: 50 en 2018 y 50 en 2022.

**Nota metodológica (tamaño muestral efectivo).** Aunque el archivo contiene 100 observaciones (50 por año), los análisis utilizan eliminación por lista (casos completos) en las variables involucradas. Por ello, el tamaño efectivo es  $n = 91$  en correlación/regresión con `computer_n_num`, y  $n = 95$  en la prueba asociada a `internet01`.

## 2.2. Variables y definiciones

Se trabaja con las siguientes variables:

- **Resultado educativo (Y):** `math` (puntaje en matemática).
- **Brecha digital (X):** `computer_n_num` (cantidad de computadoras en el hogar, recodificada como 0-3).
- **Complemento de brecha digital:** `internet01` (acceso a internet en el hogar, 0/1).
- **Control contextual (sensibilidad):** `escs` (estatus socioeconómico y cultural).

## 2.3. Recodificaciones y faltantes

```
faltantes_out <- faltantes %>% dplyr::rename("N total" = n_total,
  "NA matemática" = na_math, "NA computadoras" = na_computer_n,
  "NA internet" = na_internet, "NA ESCS" = na_escs)
knitr::kable(faltantes_out, caption = "Faltantes en variables claves
(matemática, computadoras en el hogar,
internet en el hogar, nivel socioeconómico). Fuente: learningtower (PISA).",
,booktabs = TRUE)
```

**Cuadro 1:** Faltantes en variables claves (matemática, computadoras en el hogar, internet en el hogar, nivel socio-económico). Fuente: learningtower (PISA).

N total	NA matemática	NA computadoras	NA internet	NA ESCS
100	0	9	5	4

## 3. Resultados

### 3.1. Análisis exploratorio

#### 3.1.1. Resumen descriptivo por año

```
resumen_anio <- pais_df %>%
  group_by(year) %>%
  summarise(n = n(),
    math_media = mean(math, na.rm = TRUE),
    math_sd = sd(math, na.rm = TRUE),
    math_mediana = median(math, na.rm = TRUE),
    comp_media = mean(computer_n_num, na.rm = TRUE),
    comp_mediana = median(computer_n_num, na.rm = TRUE),
    internet_prop = mean(internet01 == 1, na.rm = TRUE),
    escs_media = mean(escs, na.rm = TRUE), .groups = "drop") %>%
  mutate(internet_prop = scales::percent(internet_prop, accuracy = 0.1))
resumen_anio_out <- resumen_anio %>%
  dplyr::transmute(`Año` = year,
  `N` = n, `Matemática (media)` = math_media,
  `Matemática (DE)` = math_sd,
  `Matemática (mediana)` = math_mediana,
  `Computadoras (media)` = comp_media,
  `Computadoras (mediana)` = comp_mediana,
  `Internet en el hogar` = internet_prop,
  `ESCS (media)` = escs_media)
knitr::kable(resumen_anio_out,
  caption = "Resumen descriptivo por año
(matemática, computadoras, internet y nivel socioeconómico).
Fuente: learningtower (PISA).",
  booktabs = TRUE, digits = 2)
```

**Cuadro 2:** Resumen descriptivo por año (matemática, computadoras, internet y nivel socioeconómico). Fuente: learningtower (PISA).

Año	N	Matemática (media)	Matemática (DE)	Matemática (mediana)	Computadoras (media)	Computadoras (mediana)	Internet en el hogar	ESCS (media)
2018	50	430.23	82.49	436.32	1.86	2	95.9%	-0.24
2022	50	435.56	88.17	416.15	0.60	1	93.5%	-0.37

En esta muestra, el puntaje medio de matemática es similar entre 2018 y 2022, mientras que la disponibilidad promedio de computadoras (`computer_n_num`) es menor en 2022. El acceso a internet (`internet01`) se mantiene alto en ambos años, aunque la presencia de faltantes implica que el tamaño efectivo varía según la técnica aplicada.

### 3.1.2. Distribución de matemática por año

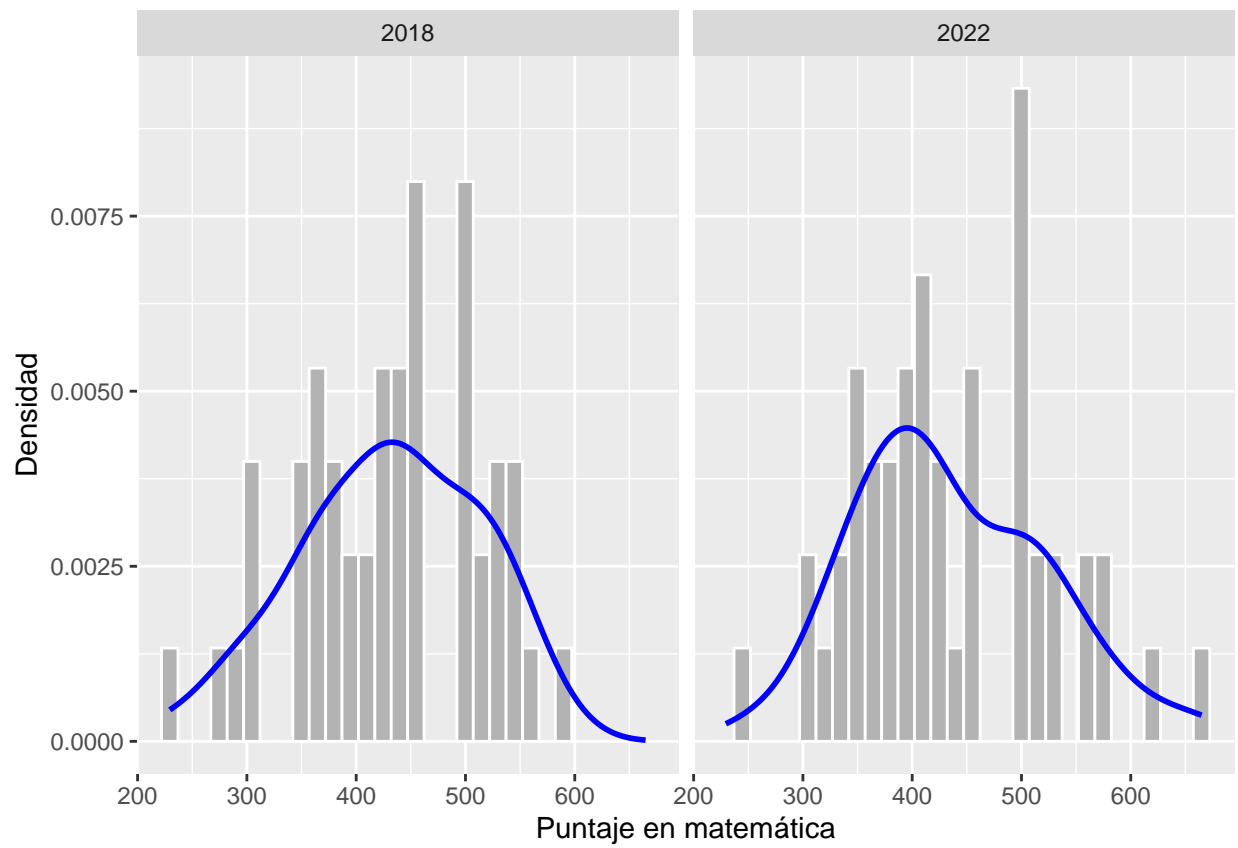
```
ggplot(pais_df, aes(x = math)) +
  geom_histogram(aes(y = after_stat(density)),
    bins = 30, fill = "grey70", color = "white") +
  geom_density(color = "blue", linewidth = 1) +
  facet_wrap(~year) +
  labs(x = "Puntaje en matemática",
    y = "Densidad")
```

En la Figura 1 se presenta la distribución de `math` por año. Se observa dispersión y valores extremos en ambos años; la comparación visual sugiere formas globales comparables, aunque con variaciones en colas.

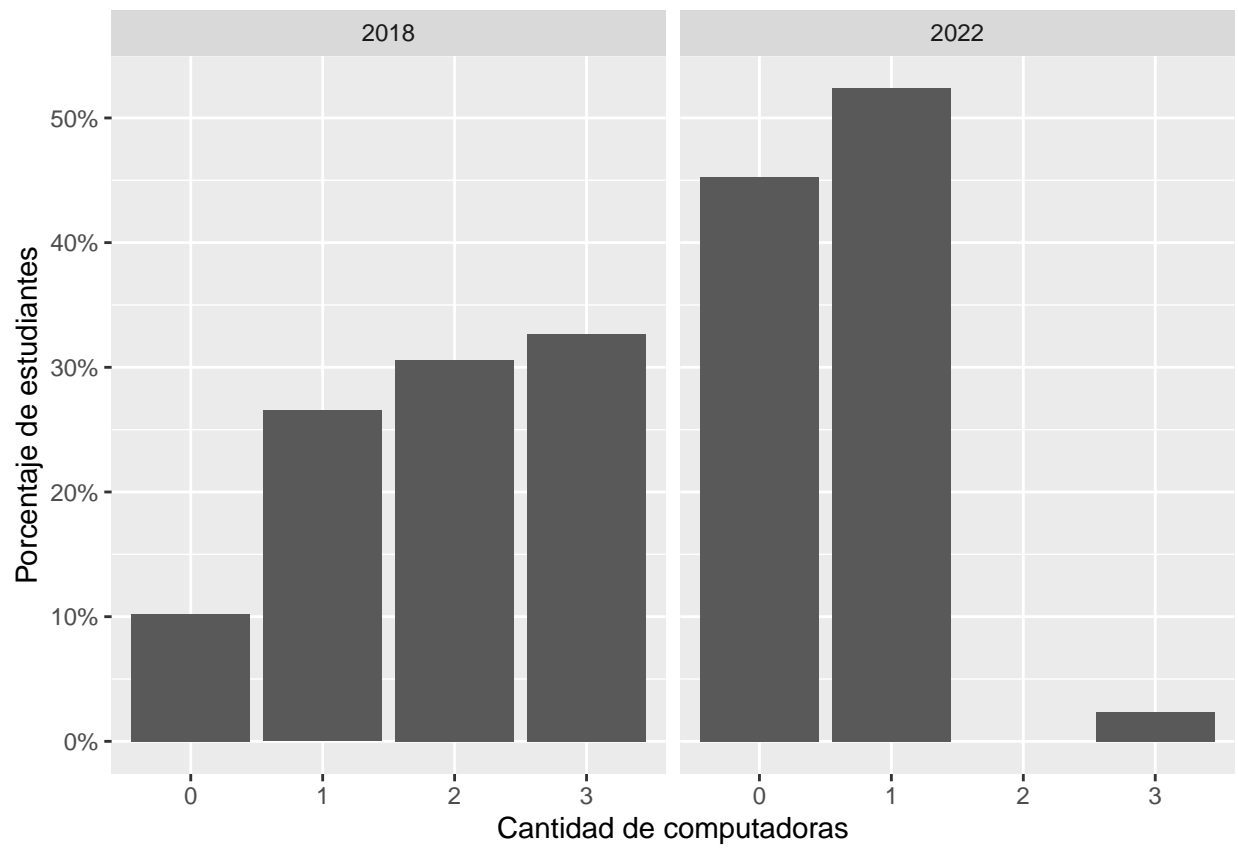
### 3.1.3. Distribución de computadoras en el hogar por año

```
dist_comp <- pais_df %>%
  filter(!is.na(computer_n_num)) %>%
  mutate(computer_n_num = factor(computer_n_num, levels = 0:3)) %>%
  count(year, computer_n_num) %>%
  group_by(year) %>%
  mutate(prop = n / sum(n)) %>%
  ungroup()
ggplot(dist_comp, aes(x = computer_n_num, y = prop)) +
  geom_col() +
  facet_wrap(~year) +
  scale_y_continuous(labels = scales::percent_format(accuracy = 1)) +
  labs(x = "Cantidad de computadoras",
    y = "Porcentaje de estudiantes")
```

En la Figura 2 se presenta la distribución de `computer_n_num` por año. Dado que `computer_n_num` es discreta (0-3), se observa concentración por categorías. Esta discreción explica que, en gráficos de dispersión, los puntos formen “bandas” verticales.



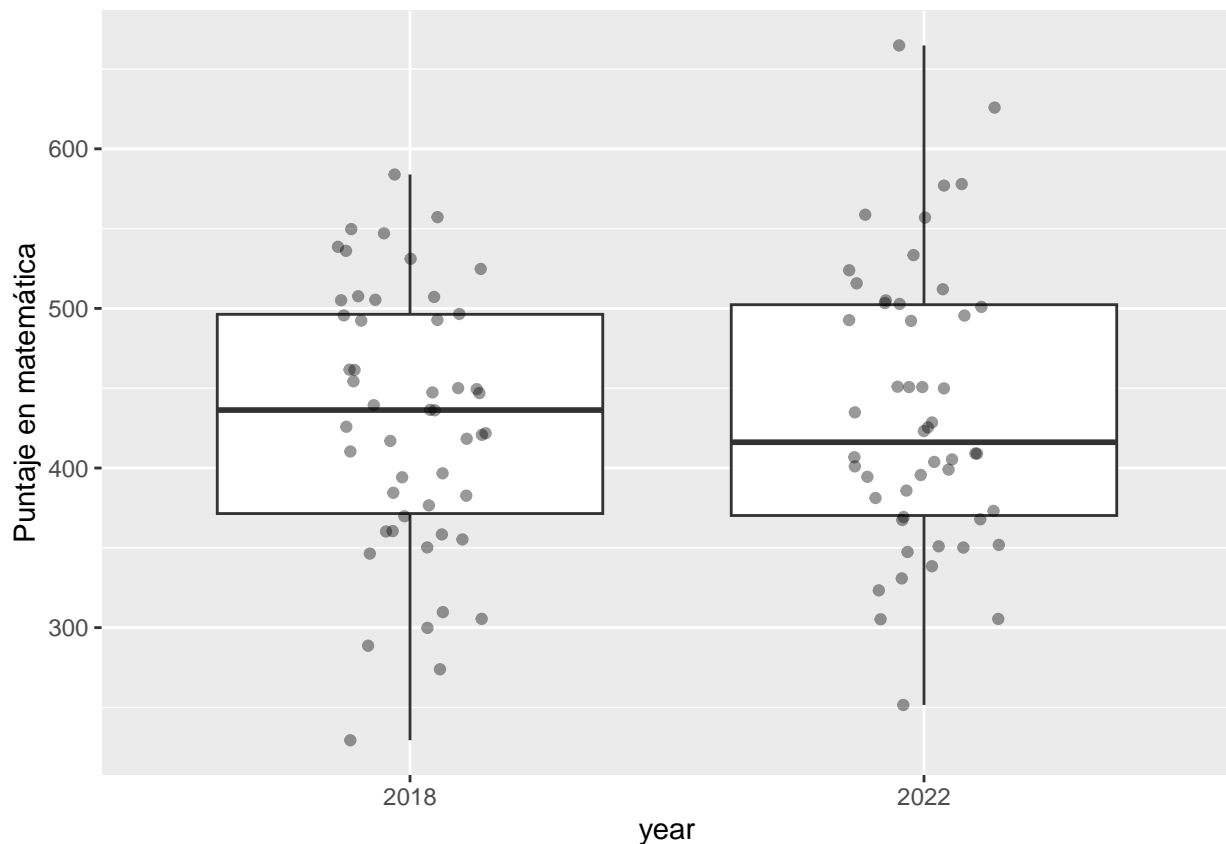
**Figura 1:** Distribución del puntaje en matemática por año. Fuente: learningtower (PISA).



**Figura 2:** Distribución de computadoras en el hogar por año, en porcentaje. Fuente: learningtower (PISA).

### 3.1.4. Comparación (Matemática) entre 2018 y 2022

```
ggplot(pais_df, aes(x = year, y = math)) +
  geom_boxplot() +
  geom_jitter(width = 0.15, alpha = 0.4) +
  labs(y = "Puntaje en matemática")
```



**Figura 3:** Comparación del puntaje en matemática entre 2018 y 2022. Fuente: learningtower (PISA).

En la Figura 3 se compara `math` entre años mediante boxplots. La comparación de medianas e IQR sugiere diferencias acotadas entre años, con presencia de valores atípicos en ambos periodos.

## 3.2. Análisis de correlación

```
cor_tabla_out <- cor_tabla %>%
  dplyr::rename(`Año` = year, `Correlación(computadoras, matemática)`
    = cor_comp_math)
knitr::kable(cor_tabla_out, caption =
  "Correlación entre computadoras en el hogar y matemática
  (total y por año). Fuente: learningtower (PISA).", booktabs = TRUE, digits = 3)
```

**Cuadro 3:** Correlación entre computadoras en el hogar y matemática (total y por año). Fuente: learningtower (PISA).

Año	Correlación(computadoras, matemática)
Total (2018-2022)	0.230
2018	0.428
2022	0.204

En el total, la correlación es 0.23, lo que indica una asociación lineal **positiva y débil**. La interpretación es descriptiva: correlación no implica causalidad.

### 3.3. Análisis de regresión lineal

#### 3.3.1. Modelo 1: regresión lineal simple

```
m1_tabla_out <- m1_tabla %>% dplyr::rename(`Término` = term,
  `Estimación` = estimate, `Error estándar` = std.error,
  `Estadístico t` = statistic, `p-valor` = p.value, `IC 95% (inf.)` =
  conf.low, `IC 95% (sup.)` = conf.high)
knitr::kable(m1_tabla_out, caption = "Modelo 1 (simple):
  coeficientes con IC del 95%. Fuente: learningtower (PISA).",
  booktabs = TRUE, digits = 3)
```

**Cuadro 4:** Modelo 1 (simple): coeficientes con IC del 95%. Fuente: learningtower (PISA).

Término	Estimación	Error estándar	Estadístico t	p-valor	IC 95 % (inf.)	IC 95 % (sup.)
Intercepto	412.778	13.717	30.092	0.000	385.522	440.035
computer_n_num	18.555	8.309	2.233	0.028	2.045	35.066

```
ajuste_m1 <- tibble::tibble(
  r_squared = m1_glance$r_squared,
  adj_r_squared = m1_glance$adj.r_squared,
  n = m1_glance$nobs)
ajuste_m1_out <- ajuste_m1 %>%
  dplyr::rename(`R^2` = r_squared,
  `R^2 ajustado` = adj_r_squared,
  `N` = n )
knitr::kable(ajuste_m1_out,
  caption = "Modelo 1 (simple): medidas de ajuste. Fuente: learningtower (PISA).",
  booktabs = TRUE,
  digits = 3)
```

**Cuadro 5:** Modelo 1 (simple): medidas de ajuste. Fuente: learningtower (PISA).

$R^2$	$R^2$ ajustado	N
0.053	0.042	91

Interpretación: en el modelo simple, una computadora adicional se asocia con un aumento promedio de 18.6 puntos en `math` ( $p = 0.0281$ ). Sin embargo, el poder explicativo es bajo ( $R^2 = 0.053$ ), por lo que la mayor parte de la variabilidad de `math` no es explicada por `computer_n_num`.

### 3.3.2. Modelo 2: sensibilidad al controlar por `escs`

```
m2_tabla_out <- m2_tabla %>% dplyr::rename(`Término` = term, `Estimación` = estimate,
  `Error estándar` = std.error, `Estadístico t` =
  statistic, `p-valor` = p.value, `IC 95% (inf.)` =
  conf.low, `IC 95% (sup.)` = conf.high)
knitr::kable(m2_tabla_out, caption = "Modelo 2 (con control): coeficientes
  con IC del 95%. Fuente: learningtower (PISA).", booktabs = TRUE,digits = 3)
```

**Cuadro 6:** Modelo 2 (con control): coeficientes con IC del 95%. Fuente: learningtower (PISA).

Término	Estimación	Error estándar	Estadístico t	p-valor	IC 95 % (inf.)	IC 95 % (sup.)
Intercepto	453.618	13.114	34.591	0.000	427.557	479.679
<code>computer_n_num</code>	-2.567	7.680	-0.334	0.739	-17.830	12.695
<code>escs</code>	51.684	8.144	6.346	0.000	35.499	67.870

```
ajuste_m2 <- tibble(r_squared = m2_glance$r_squared,
  adj_r_squared = m2_glance$adj.r_squared,
  n = m2_glance$nobs)
ajuste_m2_out <- ajuste_m2 %>% dplyr::rename(`R^2` = r_squared, `R^2 ajustado`
  = adj_r_squared, `N` = n)
knitr::kable(ajuste_m2_out,caption = "Modelo 2 (con control):
  medidas de ajuste. Fuente: learningtower (PISA).",
  booktabs = TRUE,digits = 3)
```

**Cuadro 7:** Modelo 2 (con control): medidas de ajuste. Fuente: learningtower (PISA).

$R^2$	$R^2$ ajustado	N
0.35	0.336	91

Interpretación: al incorporar `escs`, el coeficiente de `computer_n_num` pasa a -2.6 y deja de ser estadísticamente significativo ( $p = 0.739$ ). En cambio, `escs` muestra una asociación positiva sustantiva con `math` ( $\beta = 51.7$ ,  $p < 0.001$ ). El ajuste mejora ( $R^2 = 0.35$ ), sugiriendo que el componente socioeconómico explica una porción relevante de la variabilidad del rendimiento.

### 3.3.3. Comparación de ajuste (Modelo 1 vs Modelo 2)

```
comparacion_ajuste <- tibble::tibble(
  modelo = c(
    "Modelo 1: math ~ computer_n_num",
    "Modelo 2: math ~ computer_n_num + escs"),
  r_squared = c(m1_glance$r_squared, m2_glance$r_squared),
```

```

adj_r_squared = c(m1_glance$adj.r.squared, m2_glance$adj.r.squared),
AIC = c(m1_glance$AIC, m2_glance$AIC),
BIC = c(m1_glance$BIC, m2_glance$BIC))

comparacion_ajuste_out <- comparacion_ajuste %>%
  dplyr::rename(
    `Modelo` = modelo,
    `$R^2$` = r_squared,
    `R^2 ajustado` = adj_r_squared,
    `AIC` = AIC,
    `BIC` = BIC)

knitr::kable( comparacion_ajuste_out,
  caption = "Comparación de ajuste entre modelos. Fuente:
  learningtower (PISA).",
  booktabs = TRUE,
  digits = 3)

```

**Cuadro 8:** Comparación de ajuste entre modelos. Fuente: learningtower (PISA).

Modelo	$R^2$	$R^2$ ajustado	AIC	BIC
Modelo 1: $\text{math} \sim \text{computer\_n\_num}$	0.053	0.042	1066.778	1074.311
Modelo 2: $\text{math} \sim \text{computer\_n\_num} + \text{escs}$	0.350	0.336	1034.488	1044.532

### 3.4. Evaluación de supuestos y sensibilidad

A continuación se presentan diagnósticos estándar para el **modelo simple**. Estos gráficos ayudan a evaluar posibles violaciones de linealidad, normalidad de residuos, homocedasticidad e influencia.

```

ggplot(aug, aes(x = .fitted, y = .resid)) +
  geom_point(alpha = 0.6) +
  geom_smooth(se = FALSE) +
  geom_hline(yintercept = 0, linetype = 2) +
  labs(x = "Valores ajustados", y = "Residuos")

```

```

ggplot(aug, aes(sample = resid_std)) +
  stat_qq() +
  stat_qq_line() +
  labs(x = "Cuantiles teóricos", y = "Cuantiles muestrales")

```

```

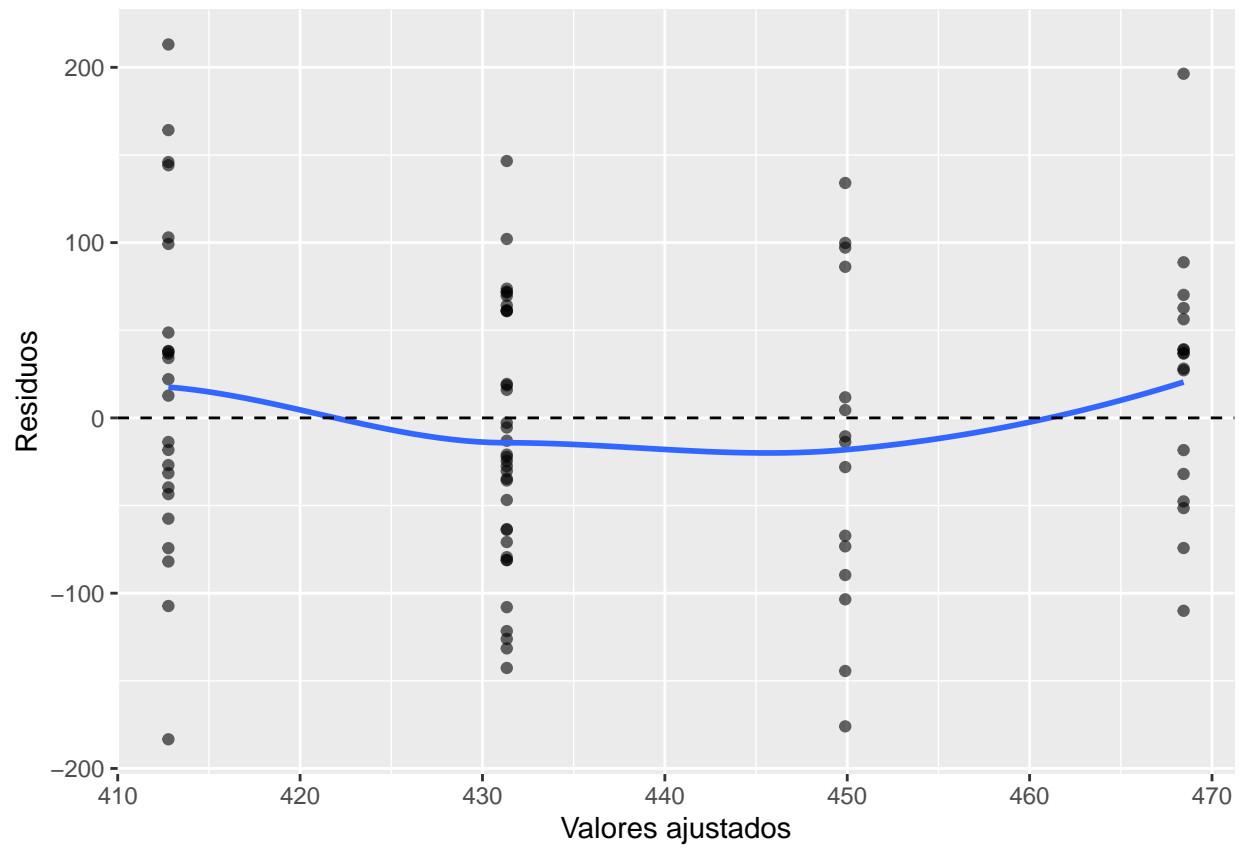
ggplot(aug, aes(x = .fitted, y = escala_loc)) +
  geom_point(alpha = 0.6) +
  geom_smooth(se = FALSE) +
  labs(x = "Valores ajustados", y = "Raíz(|residuo estandarizado|)")

```

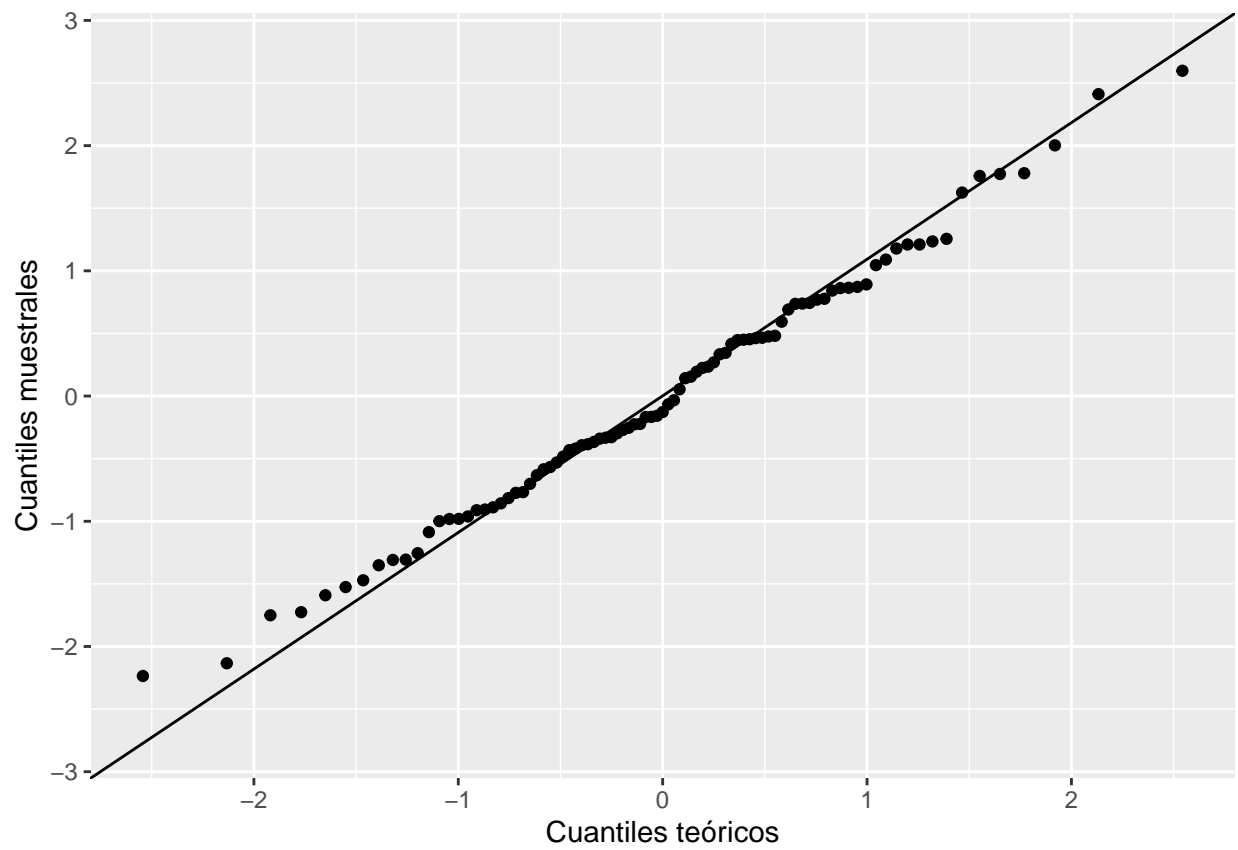
```

ggplot(aug, aes(x = leverage, y = resid_std)) +
  geom_point(aes(size = cook), alpha = 0.6) +
  geom_hline(yintercept = 0, linetype = 2) +
  labs(x = "Apalancamiento", y = "Residuos estandarizados", size = "Distancia de Cook")

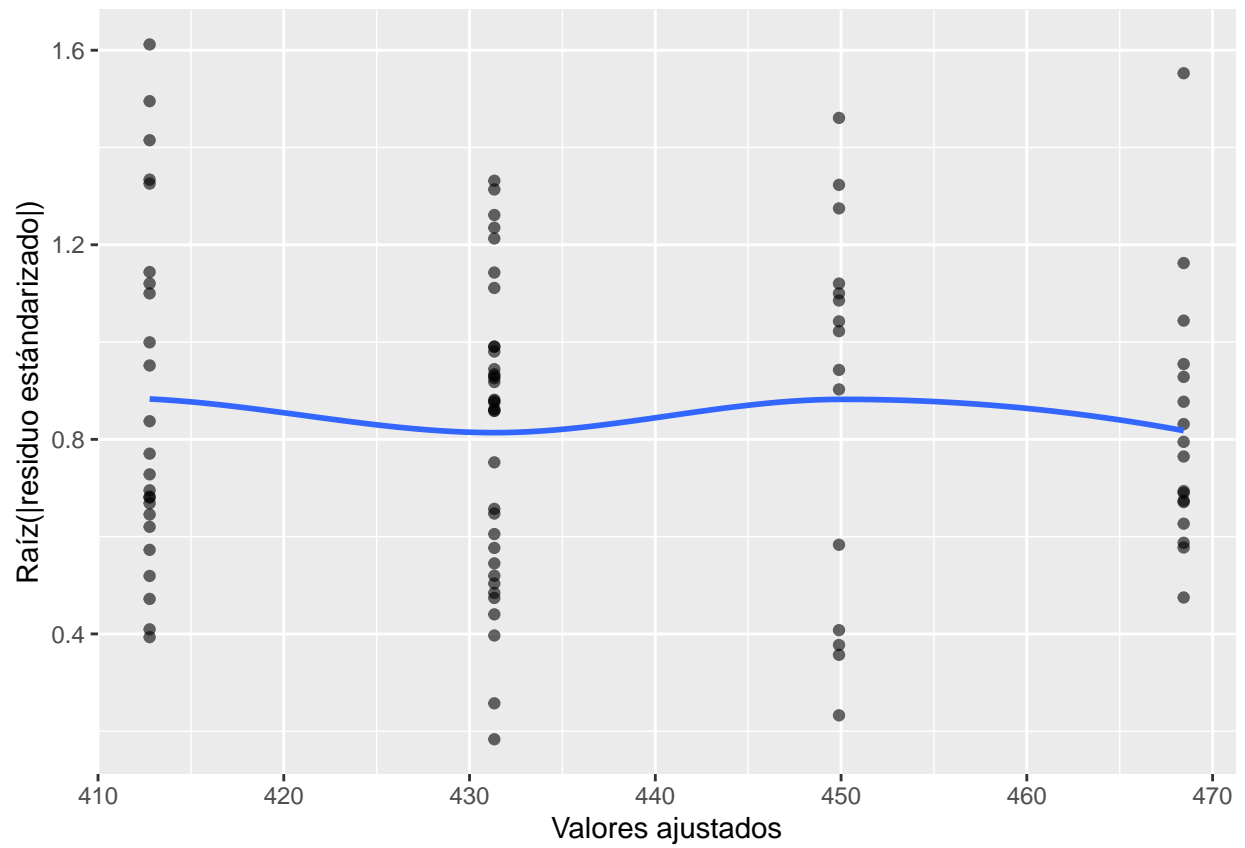
```



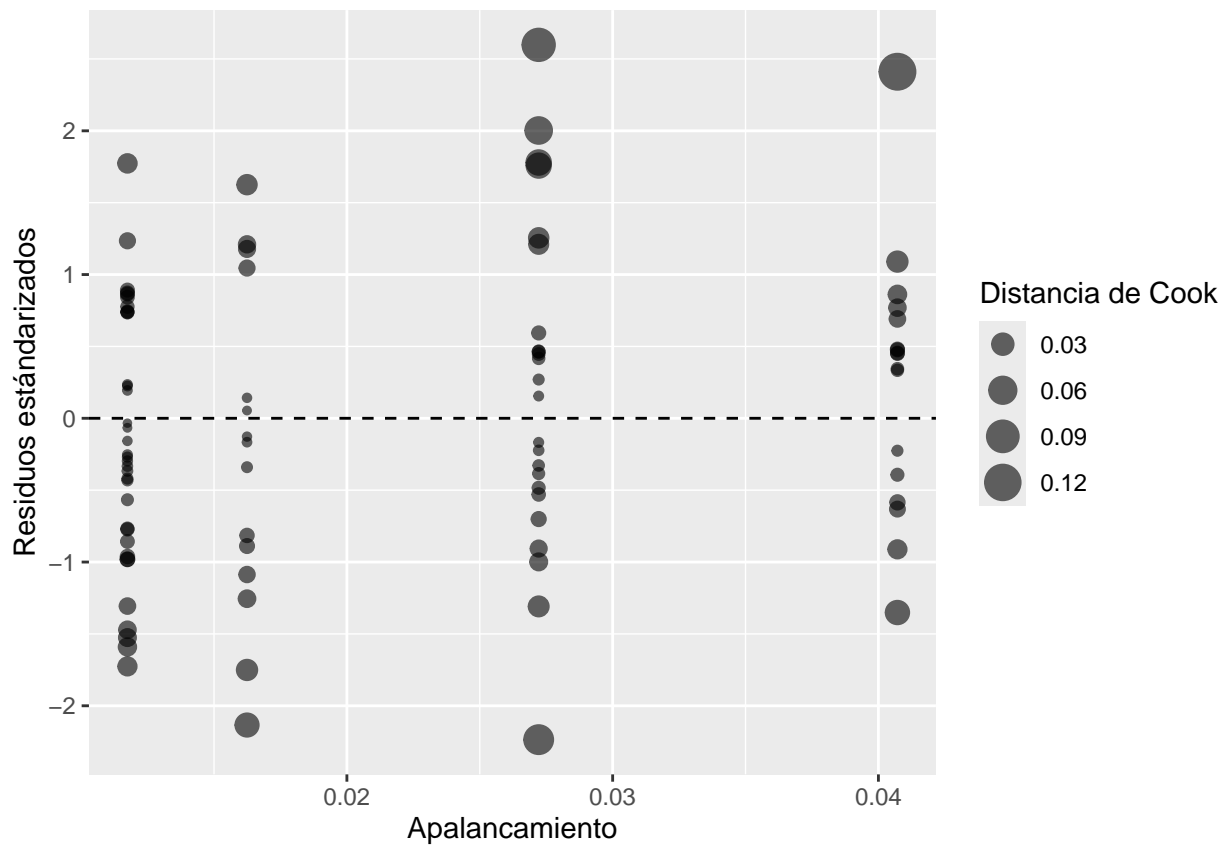
**Figura 4:** Residuos vs valores ajustados (chequeo de linealidad). Fuente: learningtower (PISA).



**Figura 5:** Gráfico Q-Q de residuos estandarizados (normalidad). Fuente: learningtower (PISA).



**Figura 6:** Escala-localización (homocedasticidad). Fuente: learningtower (PISA).



**Figura 7:** Residuos estandarizados vs apalancamiento (influencia; el tamaño del punto representa distancia de Cook).  
 Fuente: learningtower (PISA).

### 3.4.1. Robustez de la inferencia (errores estándar robustos) y no linealidad

```
robusto_tabla_out <- robusto_tabla %>% dplyr::rename(`Término` = term, `Estimación`
  = estimate, `Error estándar` = std.error,
  `Estadístico t` = statistic, `p-valor` = p.value)
knitr::kable(robusto_tabla_out,
  caption = "Modelo 1 con errores estándar robustos
  (HC3). Fuente: learningtower (PISA).",
  booktabs = TRUE, digits = 3)
```

**Cuadro 9:** Modelo 1 con errores estándar robustos (HC3). Fuente: learningtower (PISA).

Término	Estimación	Error estándar	Estadístico t	p-valor
Intercepto	412.778	14.731	28.021	0.000
computer_n_num	18.555	8.553	2.170	0.033

El uso de errores robustos (HC3) permite verificar estabilidad de la inferencia frente a heterocedasticidad. En este caso, la significancia del coeficiente de `computer_n_num` se mantiene en el mismo orden de magnitud.

Como sensibilidad adicional, se ajusta un término cuadrático. La comparación ANOVA sugiere una mejora marginal ( $p = 0.0506$ ), coherente con posible curvatura leve.

## 3.5. Visualización de la relación (dispersión y recta de regresión)

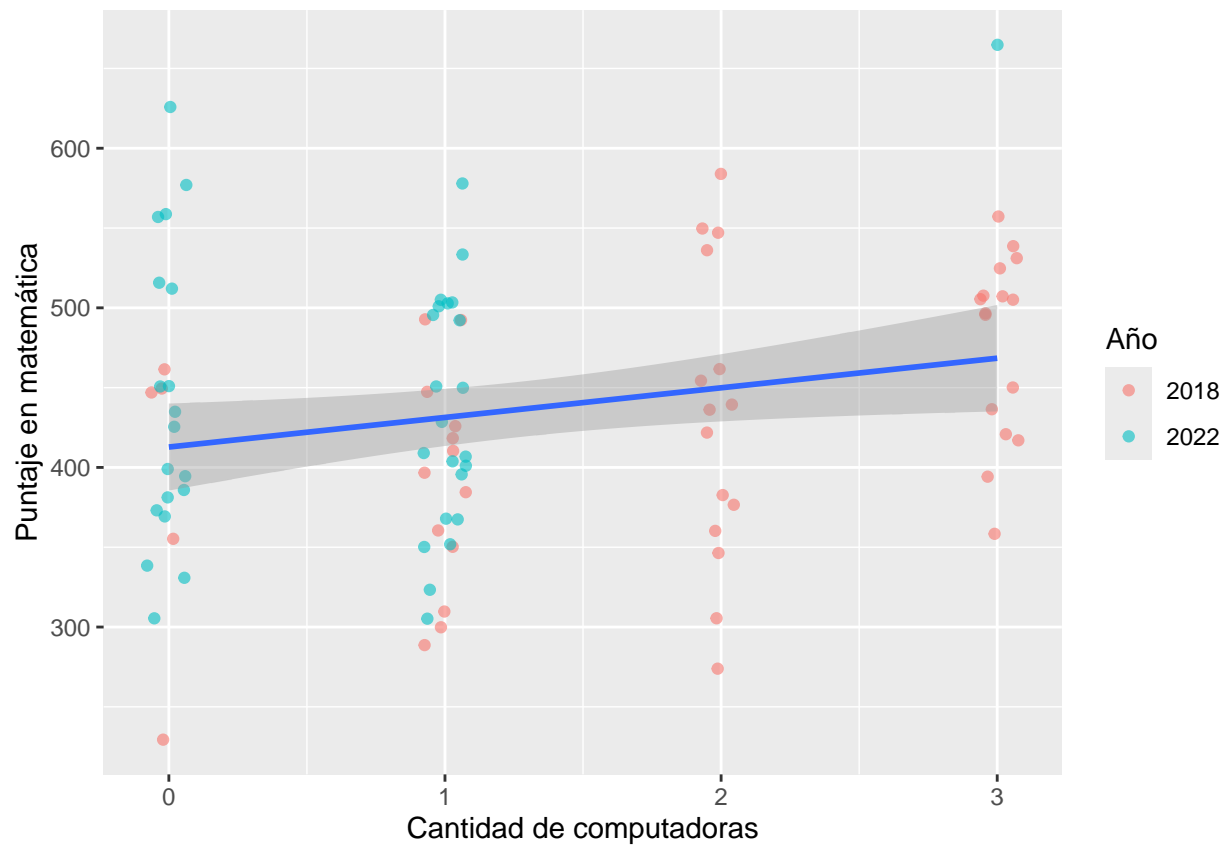
```
ggplot(df_modelo, aes(x = computer_n_num, y = math)) +
  geom_point(aes(color = year),
  position = position_jitter(width = 0.08, height = 0), alpha = 0.6) +
  geom_smooth(method = "lm", se = TRUE) +
  labs(x = "Cantidad de computadoras",
  y = "Puntaje en matemática",
  color = "Año")
```

En la Figura 8 se visualiza la relación entre `computer_n_num` y `math`. Se observa alta dispersión: aún con tendencia promedio positiva en el modelo simple, el solapamiento entre categorías y años es amplio.

## 3.6. Pruebas de hipótesis (por disponibilidad/periodicidad)

### 3.6.1. Prueba de medias: cambio en matemática (2018 vs 2022)

```
ttest_res_out <- ttest_res %>%
  dplyr::rename(`t` = estadistico_t,
  `gl` = gl,
  `p-valor` = p_value,
  `IC 95% (inf.)` = ic_95_inf,
  `IC 95% (sup.)` = ic_95_sup,
  `Media 2018` = media_2018,
  `Media 2022` = media_2022 )
```



**Figura 8:** Relación entre computadoras en el hogar y puntaje en matemática, con recta de regresión. Los puntos se muestran con jitter por X discreta. Fuente: learningtower (PISA).

```
knitr::kable( ttest_res_out,
  caption = "Prueba t de Welch:
  comparación de medias de matemática
  (2018 vs 2022). Fuente: learningtower (PISA).",
  booktabs = TRUE,digits = 3)
```

**Cuadro 10:** Prueba t de Welch: comparación de medias de matemática (2018 vs 2022). Fuente: learningtower (PISA).

t	gl	p-valor	IC 95 % (inf.)	IC 95 % (sup.)	Media 2018	Media 2022
-0.312	97.569	0.756	-39.219	28.557	430.227	435.558

En esta muestra, la prueba no aporta evidencia de diferencia en medias ( $p = 0.756$ ).

### 3.6.2. Prueba de proporciones: cambio en acceso a internet (2018 vs 2022)

```
tab_internet_out <- tab_internet %>%
  dplyr::rename(
    `Año` = year,
    `Sin internet` = `0`,
    `Con internet` = `1`)

knitr::kable( tab_internet_out,
  caption = "Tabla de contingencia para acceso a internet en
  el hogar por año. Fuente: learningtower (PISA).",
  booktabs = TRUE)
```

**Cuadro 11:** Tabla de contingencia para acceso a internet en el hogar por año. Fuente: learningtower (PISA).

Año	Sin internet	Con internet
2018	2	47
2022	3	43

```
fisher_tabla_out <- fisher_tabla %>%
  dplyr::rename(`p-valor` = p_value,
    `Odds ratio` = odds_ratio )

knitr::kable( fisher_tabla_out,
  caption = "Prueba exacta de Fisher:
  cambio en acceso a internet (2018 vs 2022). Fuente: learningtower (PISA).",
  booktabs = TRUE,digits = 3)
```

**Cuadro 12:** Prueba exacta de Fisher: cambio en acceso a internet (2018 vs 2022). Fuente: learningtower (PISA).

p-valor	Odds ratio
0.671	0.613

La prueba de Fisher no aporta evidencia de cambio en acceso a internet entre años ( $p = 0.671$ ) en esta muestra.

## 4. Discusión y limitaciones

1. **Asociación digital-educativa (magnitud y estabilidad).** En el total, la correlación (0.23) es positiva y débil. El modelo simple sugiere una pendiente positiva (beta 18.6), pero el ajuste es bajo ( $R^2$  0.053).
2. **Rol del contexto socioeconómico.** Al controlar por `escs`, la asociación de `computer_n_num` con `math` deja de ser significativa, mientras que `escs` mantiene un efecto positivo fuerte. Esto sugiere que parte de la asociación del modelo simple puede estar capturando diferencias socioeconómicas.
3. **Comparación temporal.** Para esta muestra, no se observa evidencia estadística de cambio en la media de `math` ( $p = 0.756$ ) ni en la proporción de acceso a internet ( $p = 0.671$ ).
4. **Limitaciones.**
  - Tamaño muestral reducido ( $n = 50$  por año) y subconjunto muestreo del paquete.
  - Faltantes en variables clave (en especial `computer_n_num` e `internet01`).
  - Variable `computer_n_num` discreta (0-3), lo que puede limitar sensibilidad a patrones no lineales finos.

**Próximos pasos.** Extender a más países o años (si estuviesen disponibles), evaluar medidas alternativas de brecha digital (por ejemplo, combinando dispositivos e internet en un índice) y considerar modelos no lineales o interacciones (p. ej., `computer_n_num * escs`).

## 5. Conclusiones

- En esta muestra, existe una asociación lineal **positiva y débil** entre disponibilidad de computadoras en el hogar y rendimiento en matemática ( $r = 0.23$ ).
- El modelo simple sugiere un aumento promedio de aproximadamente 18.6 puntos en `math` por computadora adicional, pero con **bajo poder explicativo** ( $R^2$  0.053).
- Al incorporar `escs`, el efecto de `computer_n_num` deja de ser significativo, mientras que `escs` mantiene un efecto positivo fuerte; por lo tanto, la evidencia sugiere que la brecha digital no se sostiene como explicación independiente una vez considerado el contexto socioeconómico.
- No se observa evidencia estadística de cambios entre 2018 y 2022 en la media de `math` ( $p = 0.756$ ) ni en el acceso a internet (Fisher  $p = 0.671$ ) para la muestra analizada.

## 6. Referencias

- **learningtower**: paquete de R que provee submuestras de PISA por año (objetos `student_subset_2018` y `student_subset_2022`).
- **OCDE / PISA**: evaluación internacional de estudiantes y marco conceptual utilizado para interpretar los resultados.
- **sandwich / lmtest**: cálculo de errores estándar robustos y pruebas asociadas para modelos lineales.

## 7. Anexo (reproducibilidad)