



# UNIVERSIDAD NACIONAL DE COLOMBIA

## ESPECIALIZACIÓN EN CIENCIAS ESTADÍSTICA

DEPARTAMENTO DE ESTADÍSTICA

FACULTAD DE CIENCIAS

### — MÉTODOS MULTIVARIADOS APLICADOS —

*Análisis Descriptivo*

*Para las pruebas Saber-11 del año 2015-2,  
en el Departamento del Cesar  
tarea 01*

---

**Estudiante:**

Hugo Hernan Rodriguez mesa C.C. 1022952094

Medellín, Colombia

Medellin, Febrero 15 de 2024

# Índice general

Índice de Figuras . . . . .	III
Índice de Tablas . . . . .	V
<b>1. Metodologia</b>	<b>1</b>
<b>2. Resultados</b>	<b>3</b>
2.1. Estructura de los datos . . . . .	3
2.1.1. Head . . . . .	3
2.1.2. Tail . . . . .	4
2.1.3. Puntajes promedio . . . . .	4
2.1.4. Desviaciones Estándar: . . . . .	4
2.1.5. Distribución de promedios por tipos de prueba: . . . . .	4
2.1.6. Correlación de datos . . . . .	4
2.1.7. Tablas de frecuencias: . . . . .	5
2.1.8. Grafico de frecuencia . . . . .	5
2.1.9. Tabla de frecuencias absolutas por 2 variables . . . . .	5
2.1.10. Tabla de frecuencias relativas por 2 variables . . . . .	6
2.1.11. Tabla de frecuencias relativas por fila . . . . .	6
2.1.12. Grafico frecuencias absolutas por fila . . . . .	7
<b>3. Conclusiones</b>	<b>9</b>
<b>4. Conclusiones</b>	<b>11</b>
4.1. Generales . . . . .	11
4.2. Recomendaciones Futuras . . . . .	11
<b>Bibliografía</b>	<b>13</b>
->	
->	



# Índice de figuras



# Índice de tablas

2.1.	Tabla head de Datos . . . . .	3
2.2.	Tabla cola de Datos . . . . .	4
2.3.	Tabla de promedio por tipo de prueba . . . . .	4
2.4.	Tabla de desviación estándar por tipo de prueba . . . . .	4
2.5.	Summary por tipo de prueba . . . . .	4
2.6.	Tabla de correlación . . . . .	5
2.7.	Tabla de frecuencia relativa/absoluta por clasificacion puntaje prueba ingles	5
2.8.	Tabla de frecuencia absoluta por clasificacion puntaje prueba ingles y naturaleza del colegio. . . . .	6
2.9.	Tabla de frecuencia relativa por clasificacion puntaje prueba ingles y natu- raleza del colegio. . . . .	6
2.10.	Tabla de frecuencia relativa por clasificacion puntaje prueba ingles y natu- raleza del colegio. . . . .	6



# Capítulo 1

## Metodologia

Los datos utilizados en este análisis descriptivo fueron extraídos de la página oficial del examen Saber 11, correspondientes al periodo 2015-2. Se limitaron los análisis únicamente al departamento de Cesar, Colombia.

Inicialmente, se accedió a la plataforma en línea proporcionada por el Instituto Colombiano para la Evaluación de la Educación (ICFES) para obtener los datos específicos de interés. Los datos fueron recopilados de manera sistemática, centrándose en las variables relevantes para el análisis descriptivo, que incluyen, entre otras, los puntajes promedio en las distintas pruebas (Lectura Crítica, Matemáticas, Sociales y Ciudadanas, Ciencias Naturales e Inglés). Con las siguientes características

Nombre del Campo	Significado
CODIGO	Código del establecimiento educativo.
INSTITUCION	Nombre del establecimiento educativo.
MUNICIPIO	Código correspondiente al municipio donde está ubicado el establecimiento educativo.
MUNICIPIO	Nombre del municipio donde está ubicada la sede del establecimiento educativo.
DEPARTAMENTO	Nombre del departamento donde está ubicado el establecimiento educativo.
CALENDARIO	Calendario Académico al que pertenece el establecimiento educativo (A=calendario A; B=calendario B; O=otro).
NATURALEZA	Sector al que pertenece el establecimiento educativo: Oficial o No Oficial.
JORNADA	Jornada de la sede del establecimiento educativo en la que están matriculados los estudiantes que presentaron las pruebas: completa, mañana, tarde, nocturna, fin de semana, única.
LECTURACRITICA	Puntaje promedio del establecimiento educativo en la prueba de Lectura Crítica.
MATEMATICAS	Puntaje promedio del establecimiento educativo en la prueba de Matemáticas.
SOCIALES	Puntaje promedio del establecimiento educativo en la prueba de Sociales y Ciudadanas.
CIENCIAS	Puntaje promedio del establecimiento educativo en la prueba de Ciencias Naturales.



---

Nombre del Campo	Significado
INGLÉS	Puntaje promedio del establecimiento educativo en la prueba de Inglés.

---

Posteriormente, se procedió a realizar un proceso de limpieza y organización de los datos para asegurar su integridad y coherencia. Donde se descargo un archivo Excel y se filtraron solo los datos correpondiente al Departamento del Cesar y luego se exporto en una base .csv.

Una vez completada la fase de preparación de datos, se procedió a realizar el análisis descriptivo. Este análisis incluyó la generación de estadísticas descriptivas, como medias, desviaciones estándar, percentiles y distribuciones de frecuencia, con el fin de caracterizar el desempeño académico de los estudiantes en el departamento de Córdoba en el periodo mencionado.

Es importante destacar que este análisis se centra exclusivamente en la descripción de los datos disponibles y no incluye análisis inferencial ni la realización de pruebas de hipótesis. El objetivo principal es proporcionar una visión general del desempeño académico en el departamento de Córdoba durante el periodo de estudio.

# Capítulo 2

## Resultados

En esta sección se presentan los resultados del análisis descriptivo de los puntajes obtenidos en el examen Saber 11 en el departamento de Cesar durante el periodo 2015-2. El objetivo de este análisis es proporcionar una visión general del desempeño académico de los estudiantes en las diferentes áreas evaluadas, así como identificar posibles tendencias o patrones destacados.

Se analizarán los puntajes promedio en las pruebas de Lectura Crítica, Matemáticas, Sociales y Ciudadanas, Ciencias Naturales e Inglés. Además, se proporcionará una distribución detallada de los puntajes, incluyendo el mínimo, máximo y percentiles.

Es importante tener en cuenta que estos resultados se basan en datos públicos proporcionados por el Instituto Colombiano para la Evaluación de la Educación (ICFES) y se limitan exclusivamente al departamento de Córdoba y al periodo mencionado.

A continuación se presentan los resultados detallados del análisis descriptivo:

### 2.1. Estructura de los datos

```
'data.frame': 264 obs. of 9 variables:
 $ CALENDARIO      : Factor w/ 1 level "A": 1 1 1 1 1 1 1 1 1 1 ...
 $ NATURALEZA      : Factor w/ 2 levels "OFICIAL","NO OFICIAL": 1 1 1 1 1 2 1 1 ...
 $ JORNADA         : Factor w/ 5 levels "COMPLETA U ORDINARIA",...: 2 2 3 3 2 3 ...
 $ EVALUADOS       : int 89 79 78 78 90 40 18 4 54 57 ...
 $ PROMLECTURACRITICA : num 52.9 49.4 46.6 51.5 48 ...
 $ PROMMATEMATICA   : num 55.4 49.9 48.5 53.2 49.7 ...
 $ PROMSOCIALESYCIUDADANAS: num 53.6 49.1 45.5 51.9 49.6 ...
 $ PROMCIENCIASNATURALES : num 53.7 49.9 47.7 51.3 48.6 ...
 $ PROMINGLES       : num 51.7 48.9 44.4 52 48.4 ...
```

#### 2.1.1. Head

Tabla 2.1: Tabla head de Datos

CALENDARIO	NATURALEZA	JORNADA	EVALUADOS	PROMLECTURACRITICA	PROMMATEMATICA	PROMSOCIALESYCIUDADANAS	PROMCIENCIASNATURALES	PROMINGLES
A	OFICIAL	MAÑANA	89	52.87	55.36	53.62	53.69	51.67
A	OFICIAL	MAÑANA	79	49.37	49.85	49.11	49.87	48.85
A	OFICIAL	TARDE	78	46.58	48.49	45.49	47.67	44.42
A	OFICIAL	TARDE	78	51.49	53.17	51.86	51.32	52.05
A	OFICIAL	MAÑANA	90	48.03	49.68	49.63	48.61	48.37

A	NO OFICIAL	TARDE	40	56.55	59.88	59.18	57.40	58.63
---	------------	-------	----	-------	-------	-------	-------	-------

### 2.1.2. Tail

Tabla 2.2: Tabla cola de Datos

	CALENDARIO	NATURALEZA	JORNADA	EVALUADOS	PROMLECTURACRITICA	PROMMATEMATICA	PROMSOCIALESYCIUDADANAS	PROMCIENCIASNATURALES	PROMINGLES
259	A	NO OFICIAL	COMPLETA U	24	46.04	47.08	44.33	45.92	50.08
260	A	NO OFICIAL	ORDINARIA COMPLETA U	79	49.15	49.65	47.72	51.48	47.78
261	A	NO OFICIAL	ORDINARIA COMPLETA U	96	46.66	47.53	46.22	47.25	46.86
262	A	NO OFICIAL	ORDINARIA MAÑANA	43	57.81	59.49	58.63	58.58	70.16
263	A	NO OFICIAL	MAÑANA	2	51.50	47.00	54.00	51.50	51.00
264	A	NO OFICIAL	COMPLETA U ORDINARIA	35	55.11	53.40	56.11	52.51	59.66

### 2.1.3. Puntajes promedio

Tabla 2.3: Tabla de promedio por tipo de prueba

Pruebas	Promedios
Promedio Lectura Critica	46.95913
Promedio Matemática	46.85413
Promedio Sociales y Ciudadanas	46.53811
Promedio Ciencias Naturales	47.38924
Promedio Ingles	47.32826

### 2.1.4. Desviaciones Estándar:

Tabla 2.4: Tabla de desviación estándar por tipo de prueba

Pruebas	Dataframe.sd
Promedio Lectura Critica	4.298761
Promedio Matemática	5.678502
Promedio Sociales y Ciudadanas	5.356387
Promedio Ciencias Naturales	4.830339
Promedio Ingles	5.680974

### 2.1.5. Distribución de promedios por tipos de prueba:

Tabla 2.5: Summary por tipo de prueba

PROMLECTURACRITICA	PROMMATEMATICA	PROMSOCIALESYCIUDADANAS	PROMCIENCIASNATURALES	PROMINGLES
Min. :38.00	Min. :34.00	Min. :28.00	Min. :38.16	Min. :40.25
1st Qu.:44.17	1st Qu.:43.27	1st Qu.:43.23	1st Qu.:44.22	1st Qu.:44.41
Median :46.73	Median :46.40	Median :46.12	Median :46.94	Median :45.84
Mean :46.96	Mean :46.85	Mean :46.54	Mean :47.39	Mean :47.33
3rd Qu.:48.75	3rd Qu.:49.17	3rd Qu.:49.48	3rd Qu.:49.29	3rd Qu.:47.96
Max. :61.52	Max. :68.66	Max. :65.07	Max. :65.66	Max. :80.07

### 2.1.6. Correlación de datos

Correlación entre las variables de los promedios”Lectura critica”, “Ciencias sociales”, “Ciencias naturales”, “Matematicas” e “Ingles”.

Tabla 2.6: Tabla de correlación

	PROMLECTURACRITICA	PROMMATEMATICA	PROMSOCIALESYCIUDADANAS	PROMCIENCIASNATURALES	PROMINGLES
PROMLECTURACRITICA	1.0000000	0.9000400	0.9212412	0.9007762	0.8079816
PROMMATEMATICA	0.9000400	1.0000000	0.8824403	0.9345844	0.7880849
PROMSOCIALESYCIUDADANAS	0.9212412	0.8824403	1.0000000	0.9108803	0.7498446
PROMCIENCIASNATURALES	0.9007762	0.9345844	0.9108803	1.0000000	0.8160976
PROMINGLES	0.8079816	0.7880849	0.7498446	0.8160976	1.0000000

### 2.1.7. Tablas de frecuencias:

Observando las prueba de ingles que contiene la mayor varianza vamos a crear una clasificación para segun el puntaje de esta aprueba:

- Puntajes promedio menores al 1st Qu(44.41) son considerados como **bajo**,
- Puntajes promedio mayores/iguales al 1st Qu(44.41) y menores/iguales al 3rd Qu(47.96) son considerados como **promedio**,
- Puntajes promedio mayores al 3rd Qu 47.96 son considerados como **alto**.

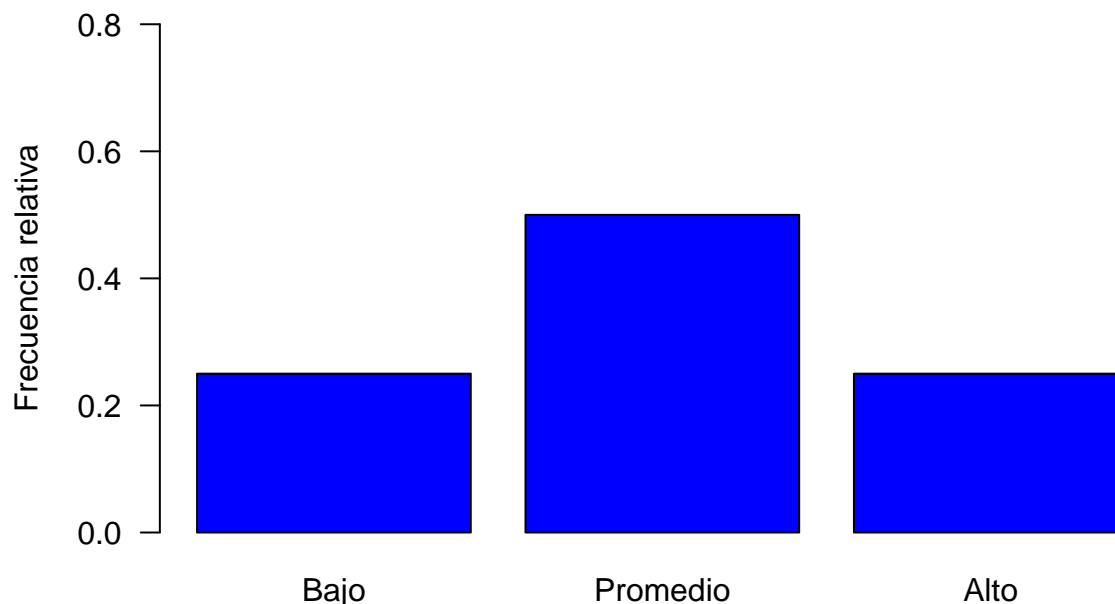
####Tabla de frecuencias absolutas y relativas

Tabla 2.7: Tabla de frecuencia relativa/absoluta por clasificacion puntaje prueba ingles

	Frec.clasi.ingles	Frec.Rel.clasi.ingles
Bajo	66	0.25
Promedio	132	0.50
Alto	66	0.25

### 2.1.8. Grafico de frecuencia

#### Clasificación según el promedio de la prueba de ingles



### 2.1.9. Tabla de frecuencias absolutas por 2 variables

Analizaremos ahora dos variables la clasificación segun el promedio de ingles y la naturaleza del colegio.

Tabla 2.8: Tabla de frecuencia absoluta por clasificación puntaje prueba inglés y naturaleza del colegio.

	Bajo	Promedio	Alto
OFICIAL	55	115	35
NO OFICIAL	11	17	31

### 2.1.10. Tabla de frecuencias relativas por 2 variables

Analizaremos ahora dos variables la clasificación según el promedio de inglés y la naturaleza del colegio.

Tabla 2.9: Tabla de frecuencia relativa por clasificación puntaje prueba inglés y naturaleza del colegio.

	Bajo	Promedio	Alto
OFICIAL	0.2083	0.4356	0.1326
NO OFICIAL	0.0417	0.0644	0.1174

### 2.1.11. Tabla de frecuencias relativas por fila

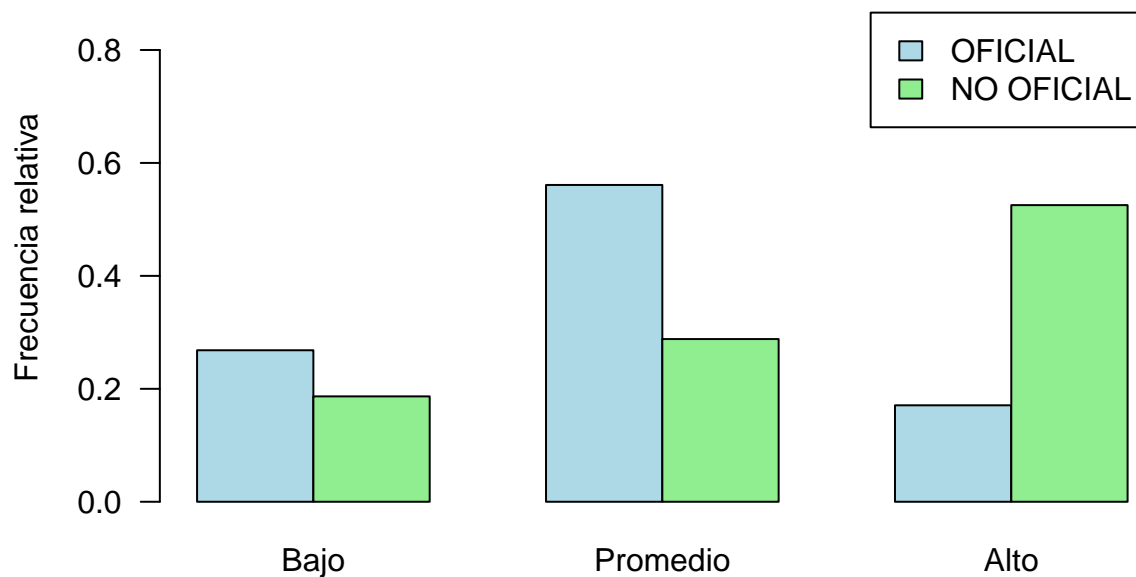
Analizaremos ahora dos variables la clasificación según el promedio de inglés y la naturaleza del colegio.

Tabla 2.10: Tabla de frecuencia relativa por clasificación puntaje prueba inglés y naturaleza del colegio.

	Bajo	Promedio	Alto
OFICIAL	0.2683	0.5610	0.1708
NO OFICIAL	0.1866	0.2881	0.5253

**2.1.12. Grafico frecuencias absolutas por fila**

**Frecuencias relativas de clasificacion prueba de ingles según la naturaleza del colegio.**





## Capítulo 3

### Conclusiones

Después de analizar los datos, se observa una tendencia interesante en cuanto al desempeño en las pruebas de inglés entre colegios oficiales y no oficiales. Los colegios no oficiales muestran una mayor propensión a obtener puntajes altos en la prueba de inglés en comparación con los colegios oficiales. Específicamente, el 52.53 % de los estudiantes en colegios no oficiales obtuvieron puntajes altos, mientras que solo el 17.08 % de los estudiantes en colegios oficiales lograron lo mismo.

Por otro lado, los colegios oficiales tienen una proporción más alta de estudiantes con puntajes bajos en la prueba de inglés en comparación con los colegios no oficiales. Un 26.83 % de los estudiantes en colegios oficiales obtuvieron puntajes bajos, mientras que solo un 18.66 % de los estudiantes en colegios no oficiales tuvieron resultados similares.

Estas conclusiones sugieren que el tipo de institución educativa puede influir en el desempeño de los estudiantes en la prueba de inglés. Aunque no se puede determinar la causa exacta de esta disparidad basada únicamente en estos datos, es posible que los colegios no oficiales tengan diferentes enfoques pedagógicos, recursos o condiciones que contribuyan a un mejor rendimiento en inglés. Estos hallazgos podrían ser útiles para informar políticas educativas y prácticas pedagógicas destinadas a mejorar el aprendizaje del idioma inglés en colegios oficiales. Sin embargo, se requiere una investigación más profunda para comprender completamente los factores subyacentes detrás de esta disparidad y desarrollar estrategias efectivas para abordarla.





# Capítulo 4

## Conclusiones

### 4.1. Generales

- xxxxxx
- xxxxxxxxxx
- dsxxxxxxx

### 4.2. Recomendaciones Futuras

- xxxxxx
- xxxxxxxxxx
- dsxxxxxxx



# Bibliografía

- Allaire, J., Xie, Y., Dervieux, C., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., Wickham, H., Cheng, J., Chang, W., and Iannone, R. (2023). *rmarkdown: Dynamic Documents for R*. R package version 2.25, <https://pkgs.rstudio.com/rmarkdown/>.
- RStudio Team (2015). *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA.
- Xie, Y. (2014). knitr: A comprehensive tool for reproducible research in R. In Stodden, V., Leisch, F., and Peng, R. D., editors, *Implementing Reproducible Computational Research*. Chapman and Hall/CRC. ISBN 978-1466561595.
- Xie, Y. (2015). *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition. ISBN 978-1498716963.
- Xie, Y. (2023). *knitr: A General-Purpose Package for Dynamic Report Generation in R*. R package version 1.45.
- Xie, Y., Allaire, J., and Golemund, G. (2018). *R Markdown: The Definitive Guide*. Chapman and Hall/CRC, Boca Raton, Florida.
- Xie, Y., Dervieux, C., and Riederer, E. (2020). *R Markdown Cookbook*. Chapman and Hall/CRC, Boca Raton, Florida.