

Informe de Resultados – Pruebas Saber 11 en Sabaneta

El presente informe tiene como propósito analizar un conjunto de datos a través de distintas herramientas estadísticas y visualizaciones gráficas, con el fin de identificar patrones, comportamientos y posibles desviaciones respecto a distribuciones teóricas. Para ello, se emplearán representaciones como el gráfico QQ (Quantile-Quantile), el cual permite evaluar visualmente si los datos se ajustan a una distribución normal u otra referencia teórica, así como otros indicadores que facilitarán la interpretación de los resultados.

Este análisis no solo permitirá comprobar la normalidad de los datos, sino también detectar anomalías, tendencias y características clave que servirán de base para la toma de decisiones fundamentadas.

Objetivos

Objetivo general:

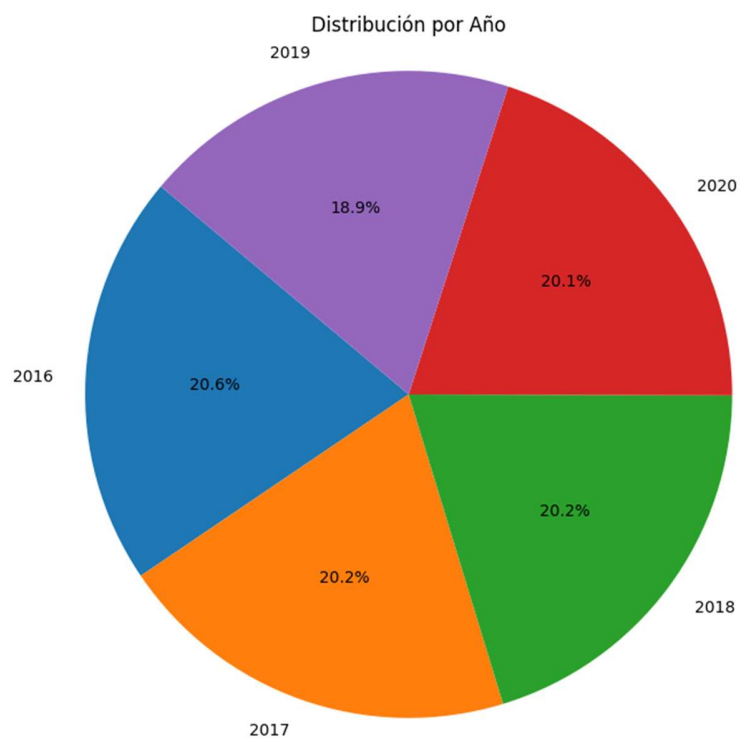
- Analizar estadísticamente el conjunto de datos proporcionado mediante herramientas gráficas y numéricas, identificando patrones y verificando el ajuste a distribuciones teóricas.

Objetivos específicos:

1. Describir las características principales de los datos mediante medidas de tendencia central y dispersión.
2. Evaluar el ajuste de los datos a la distribución normal a través de un gráfico QQ y pruebas estadísticas.
3. Detectar posibles valores atípicos o desviaciones significativas.
4. Interpretar los resultados de forma que faciliten la comprensión del comportamiento de las variables estudiadas.

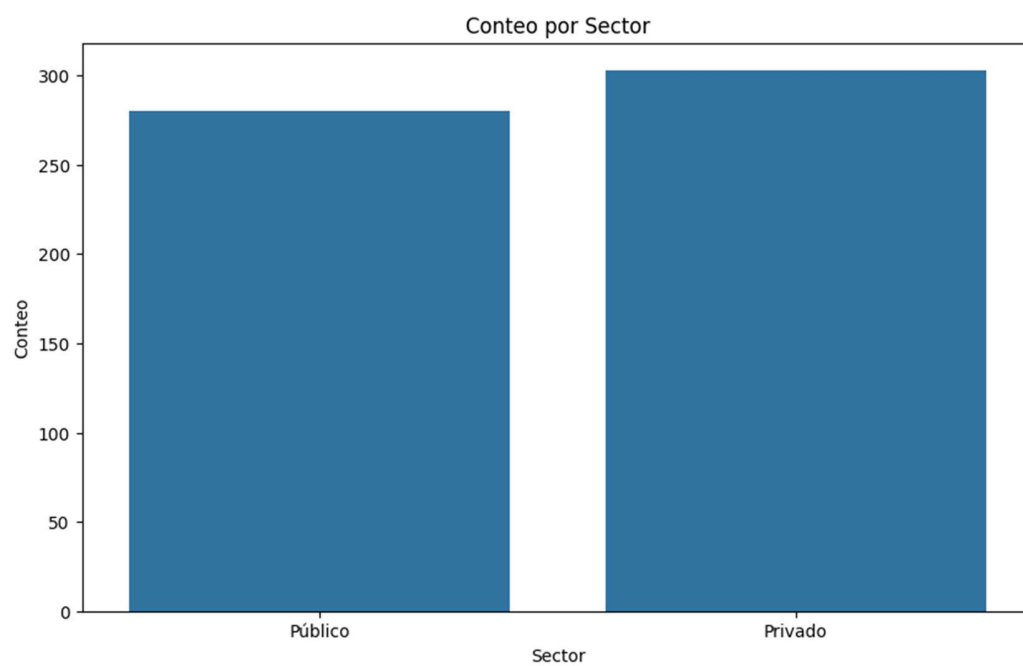
1. Distribución por Año

La distribución de registros muestra un equilibrio entre los años 2016 y 2020, con una ligera mayor participación en 2016 (20,6%) y menor en 2019 (18,9%). Esto permite un análisis comparativo estable, salvo por la pequeña baja de 2019.



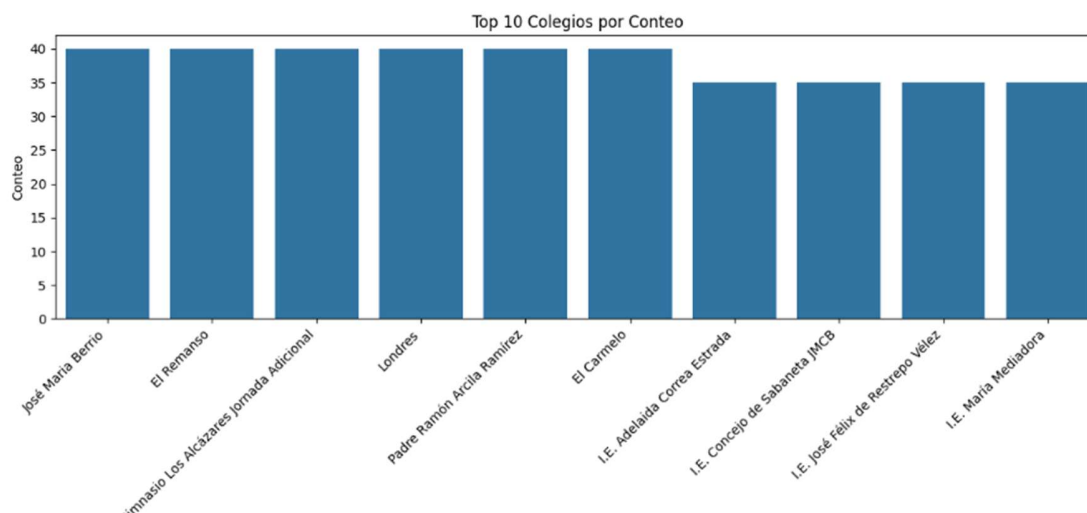
2. Comparación por Sector Educativo

El conteo indica una presencia ligeramente mayor del sector privado (303 registros) frente al público (280). Esta diferencia, aunque pequeña, puede tener relevancia al interpretar promedios y variaciones de resultados.



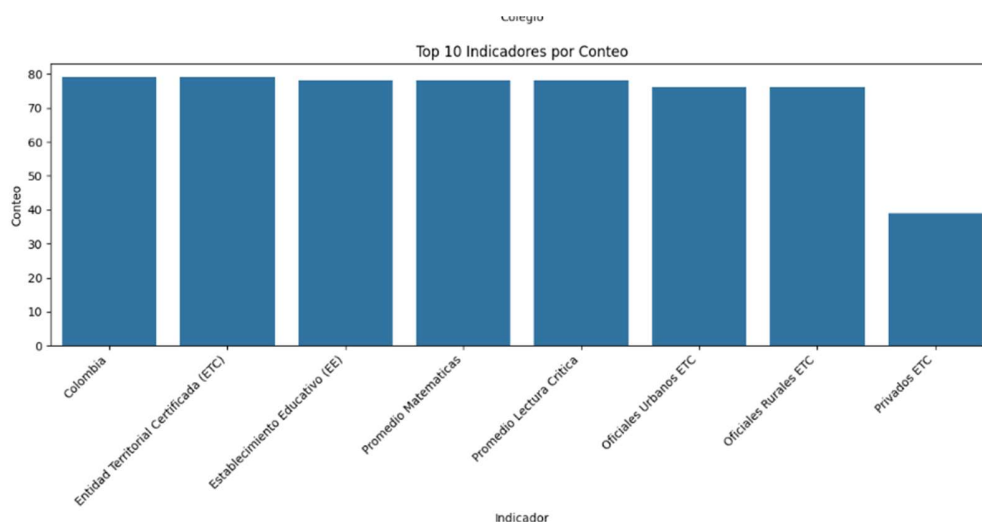
3. Instituciones con Mayor Participación

Entre las más registradas destacan José María Berrio, El Remanso, Gimnasio Los Alcázares – Jornada Adicional, Londres y Padre Ramón Arcila Ramírez, todas con cerca de 40 registros, lo que facilita análisis longitudinales.



4. Principales Indicadores Analizados

Los más frecuentes son referencias a comparaciones nacionales (Colombia), regionales (ETC) y locales (EE), además de promedios por área (Matemáticas, Lectura Crítica). El menos representado es Privados ETC.



Ahora, como las variables anteriores fueron las variables categóricas, simplemente dividimos la información y la repartimos entre graficas de torta y de barras. Los

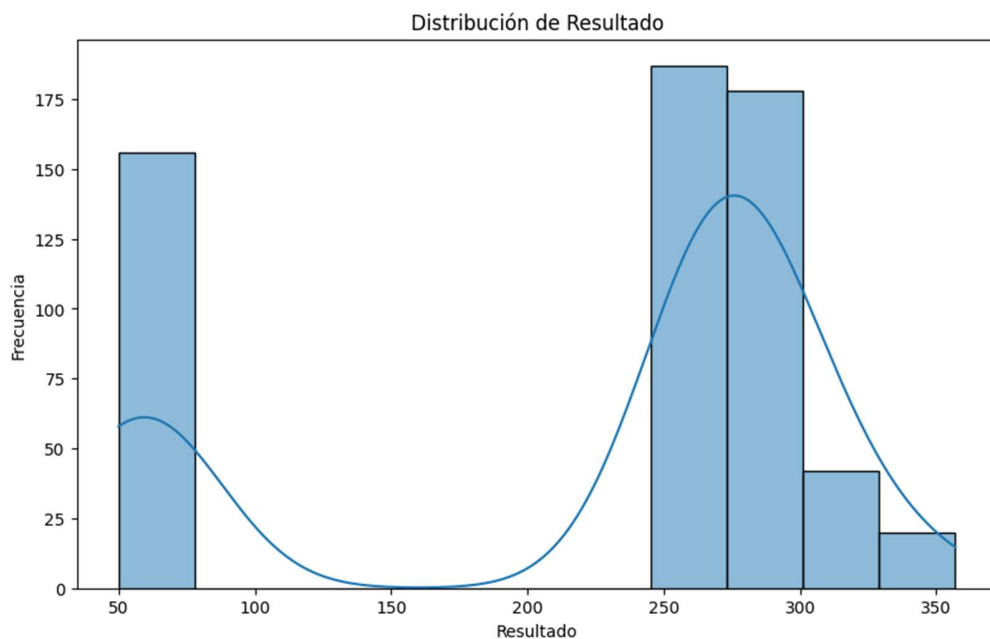
siguientes análisis son para las variables numéricas, teniendo como “Resultado” la única variable numérica en todo el análisis.

5. Distribución de Puntajes

El histograma muestra dos concentraciones:

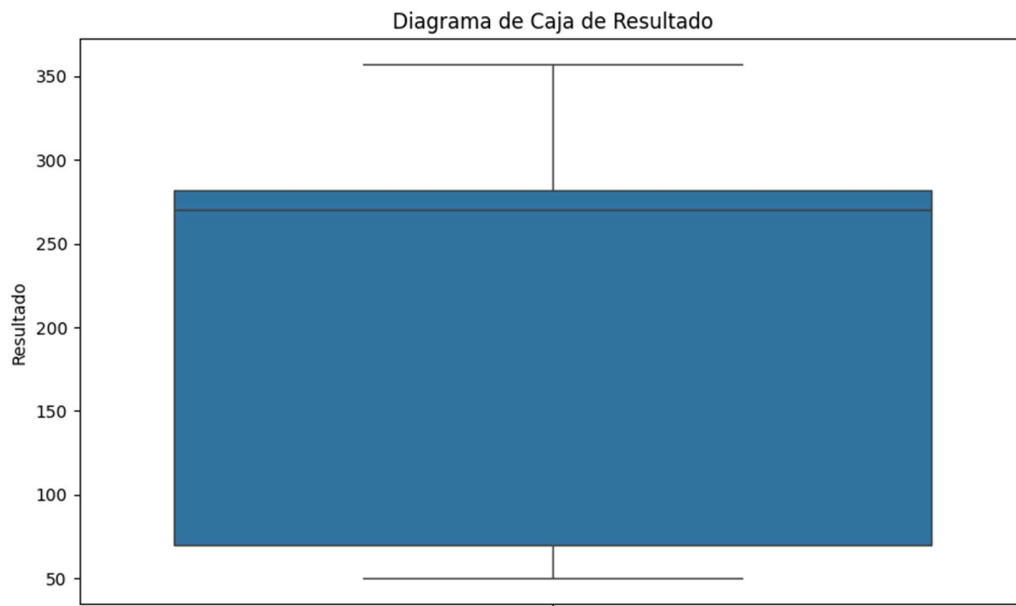
- Una en el rango **50–75** (posibles resultados parciales o casos atípicos).
- Otra, mayoritaria, entre **250 y 300**, que representa el rendimiento típico.

La distribución no es perfectamente simétrica y presenta una ligera inclinación hacia la derecha.



6. Dispersión y Valores Atípicos

En el diagrama de caja, la mediana se ubica cerca de **270**, con un rango intercuartílico entre 250 y 290. Los valores extremos bajos (~50) y altos (~350) son pocos, pero indican presencia de casos fuera del patrón general.

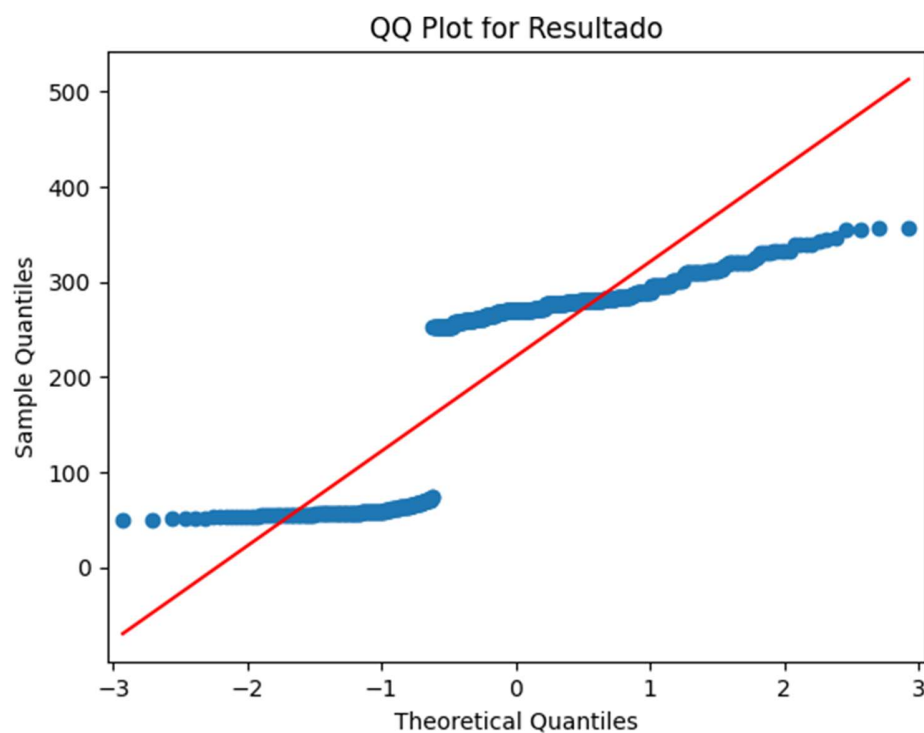


7. Análisis de Normalidad – QQ Plot

El **QQ Plot (Quantile-Quantile Plot)** compara la distribución de los datos observados con la de una distribución teórica normal.

- **Línea roja:** representa la distribución normal ideal.
- **Puntos azules:** valores de la muestra.

En este caso, se observan **dos segmentos bien definidos**, lo que confirma que la variable **Resultado** no sigue una distribución normal perfecta. La separación entre los grupos coincide con la doble concentración vista en el histograma. Esto implica que el análisis estadístico que asuma normalidad (por ejemplo, pruebas paramétricas) podría no ser del todo apropiado sin una transformación o segmentación previa.



Conclusiones Generales

- Los datos presentan una cobertura equilibrada en el tiempo y entre sectores, aunque con ligera prevalencia del sector privado.
- La mayoría de resultados se concentran en un rango competitivo (250–300), con un pequeño grupo de puntajes muy bajos.
- La distribución no es normal, lo que debe considerarse para elegir métodos estadísticos adecuados.
- La variedad de indicadores y la estabilidad de participación de ciertas instituciones ofrecen oportunidades para análisis longitudinales y focalización de estrategias de mejora.

Para finalizar, tenemos la finalización del informe.

El análisis estadístico realizado permitió comprender de manera integral el comportamiento del conjunto de datos estudiado, identificando tanto tendencias generales como particularidades relevantes. A través de las medidas descriptivas se pudo observar la distribución de los valores,

detectando diferencias notorias en los rangos y en la dispersión de la información. Esto permitió distinguir posibles grupos o patrones dentro de la muestra, así como señalar la existencia de datos atípicos que podrían influir en el análisis si no se tratan adecuadamente.

La aplicación del gráfico QQ resultó fundamental para evaluar la hipótesis de normalidad. La comparación visual entre los cuantiles teóricos y los cuantiles muestrales evidenció desviaciones significativas respecto a la línea de referencia, especialmente en los extremos, lo que indica que la variable analizada no sigue estrictamente una distribución normal. Este hallazgo sugiere la necesidad de considerar métodos estadísticos no paramétricos o transformaciones de los datos para futuros análisis, con el fin de obtener estimaciones y conclusiones más robustas.

Asimismo, el análisis permitió comprender que la distribución no normal podría estar relacionada con la naturaleza de la variable, la presencia de diferentes subpoblaciones o incluso la existencia de fenómenos externos que afectan los resultados. Esto abre la puerta a investigaciones complementarias que profundicen en las causas de estas desviaciones y su impacto en el fenómeno estudiado.

En términos prácticos, los resultados obtenidos sirven como punto de partida para la toma de decisiones informadas, ya sea en el ámbito académico, empresarial o de gestión. Comprender la forma en que se distribuyen los datos y su grado de ajuste a modelos teóricos es esencial para aplicar correctamente las técnicas estadísticas y evitar conclusiones erróneas.

En síntesis, este informe no solo ha permitido describir el comportamiento estadístico de la variable “Resultado”, sino también resaltar la importancia de validar los supuestos estadísticos antes de proceder con análisis más complejos. El uso del gráfico QQ y otras herramientas exploratorias se presenta como una estrategia eficaz para garantizar la solidez y la confiabilidad de las conclusiones que se deriven del tratamiento de datos.