

# Rompiendo el molde: Identificando los diversos perfiles de éxito académico en Colombia con Machine Learning



Elaborado por: Juan Pablo Gómez Morales

## Introducción

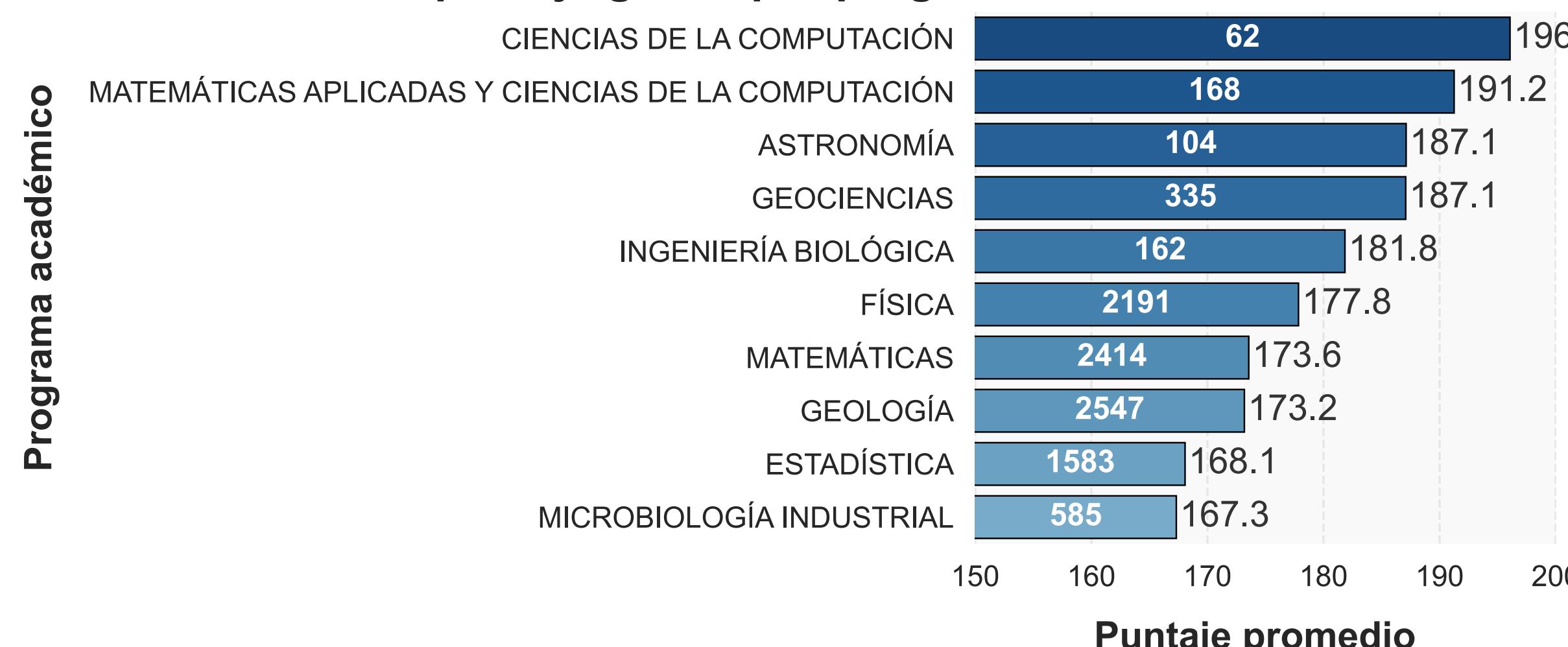
Este estudio integra el análisis estadístico descriptivo tradicional con técnicas avanzadas de Machine Learning (MCA, PCA, K-Means) para ofrecer una caracterización sólida y multidimensional del rendimiento académico en Colombia y en el área de Matemáticas y Ciencias Naturales, utilizando datos de las Pruebas Saber Pro desde 2016 hasta 2024. El análisis descriptivo presentado en las gráficas de distribución y rendimiento no solo contextualiza los datos, sino que valida y refuerza significativamente los patrones complejos identificados por los modelos multivariados.

La convergencia de ambos enfoques se visualiza en este póster, donde mapas de calor, distribuciones por género y desempeño confirman los perfiles estudiantiles obtenidos computacionalmente, proporcionando una evidencia robusta para la toma de decisiones en política educativa.

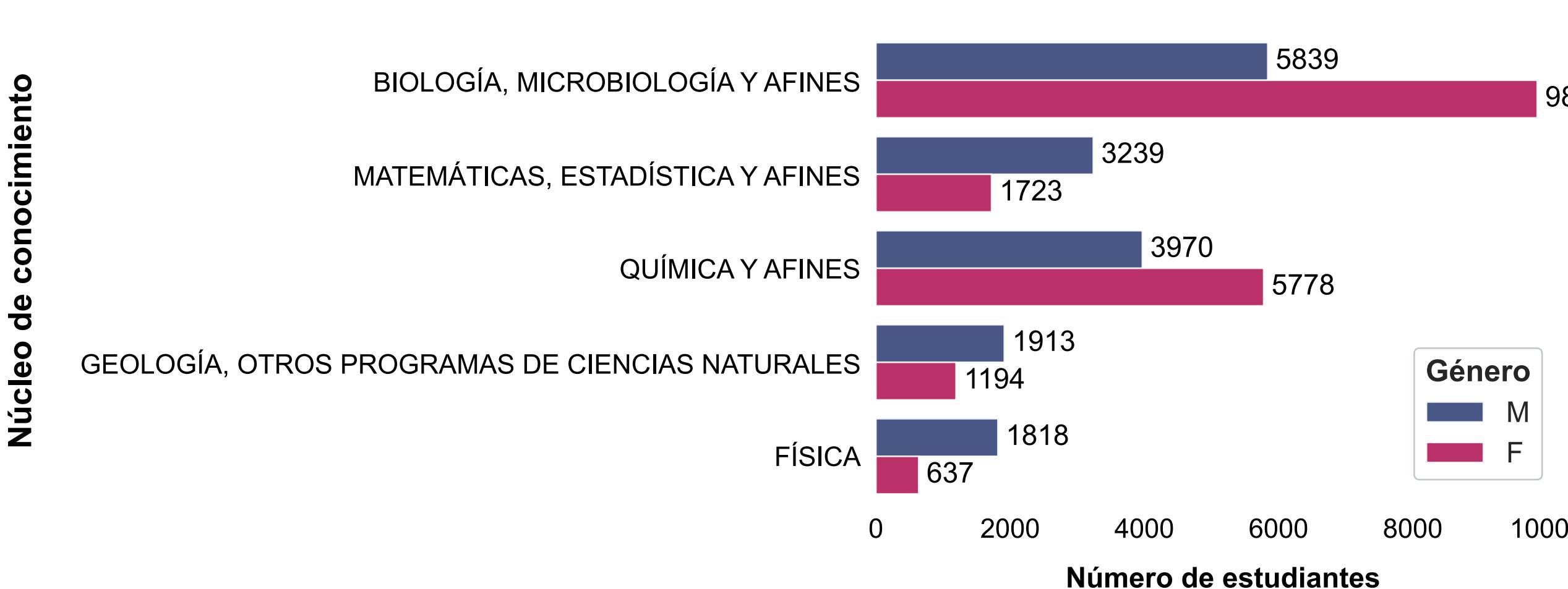
## Área de Matemáticas y Ciencias Naturales

Este análisis examina la participación y el rendimiento académico en las áreas de Matemáticas y Ciencias Naturales, considerando la distribución de estudiantes por género, el puntaje global por núcleo de conocimiento y el desempeño promedio por programa académico específico.

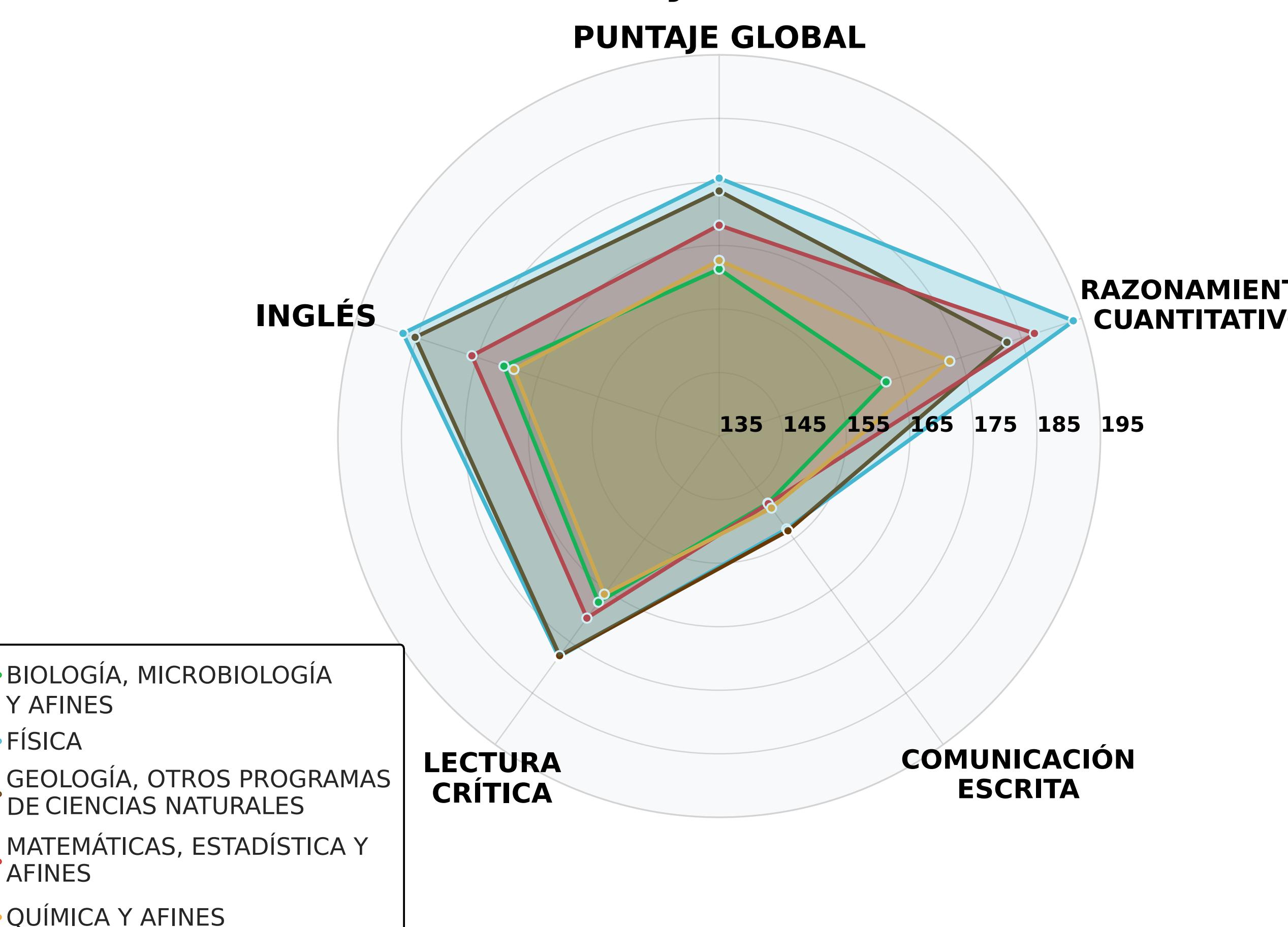
### Promedio del puntaje global por programa académico de MCN



### Cantidad de estudiantes por género en el área de Matemáticas y Ciencias Naturales



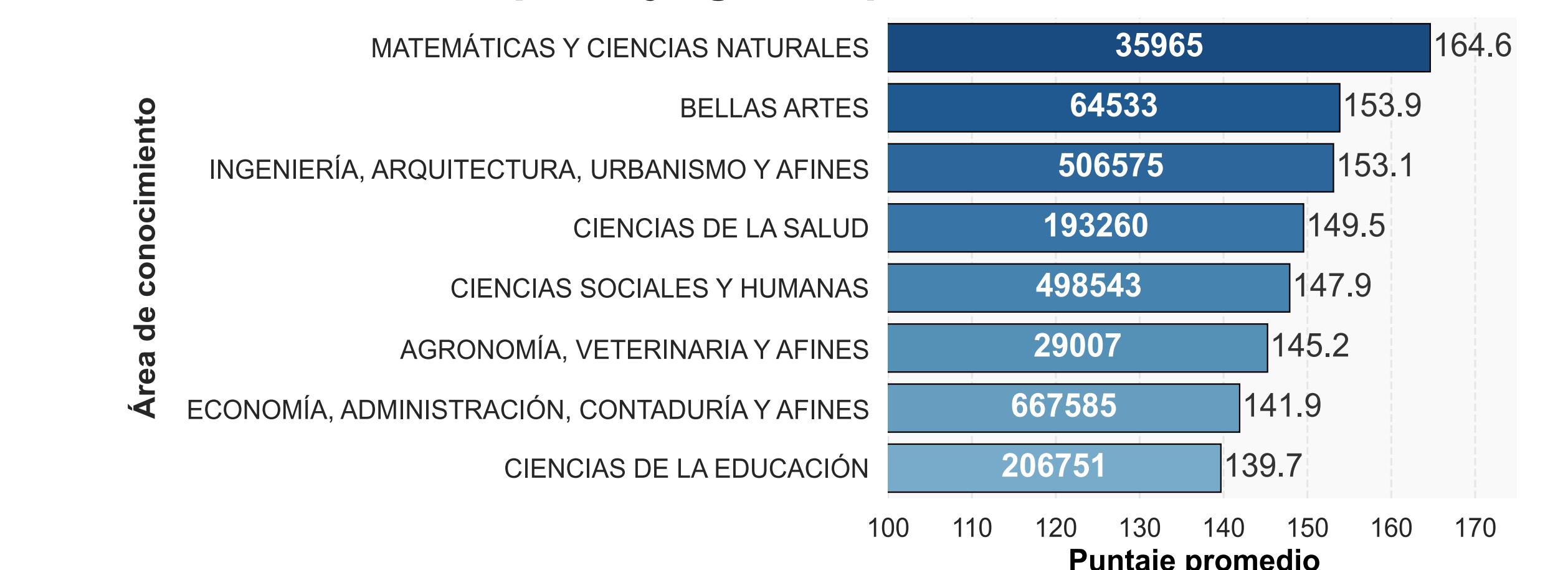
### Rendimiento de cada núcleo en el área de Matemáticas y Ciencias Naturales



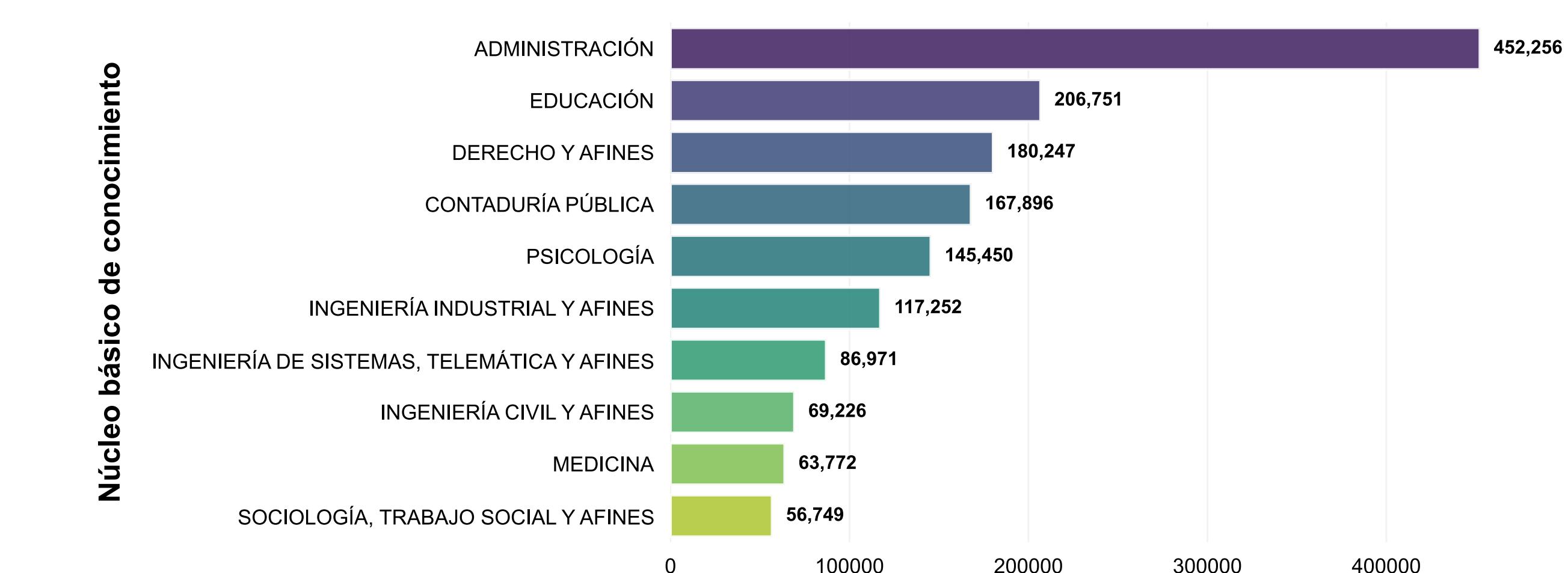
## Análisis estadístico descriptivo - Pruebas genéricas a nivel nacional

Se ofrece una visión general del panorama de la educación superior en Colombia, analizando la distribución de estudiantes, el rendimiento académico y las diferencias por áreas de conocimiento, programas específicos e instituciones. A través de gráficas descriptivas, se presentan datos clave sobre preferencias de estudio y desempeño en pruebas estandarizadas. Estos hallazgos proporcionan el contexto inicial para la posterior aplicación de técnicas multivariadas y métodos de agrupamiento.

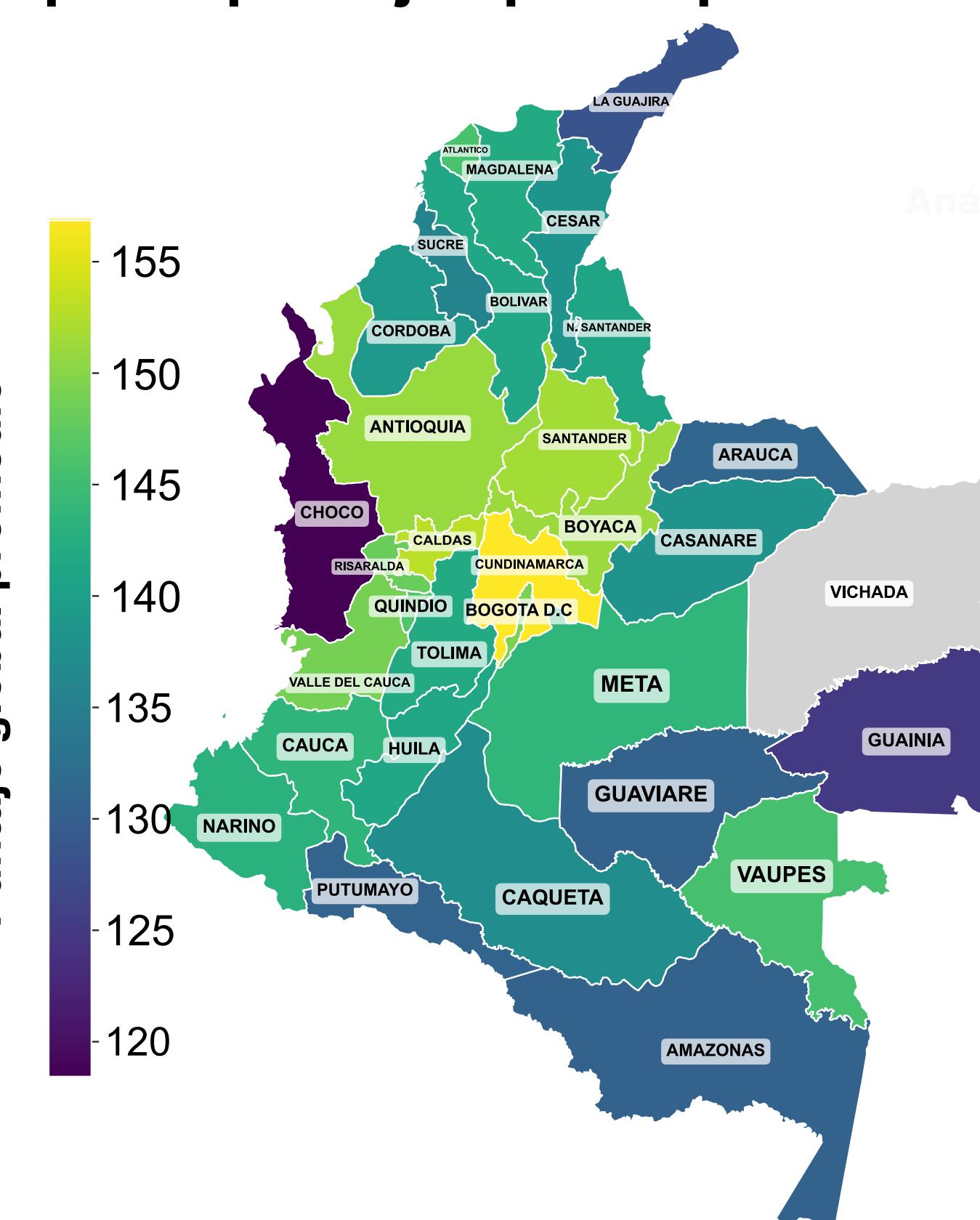
### Promedio del puntaje global por área de conocimiento



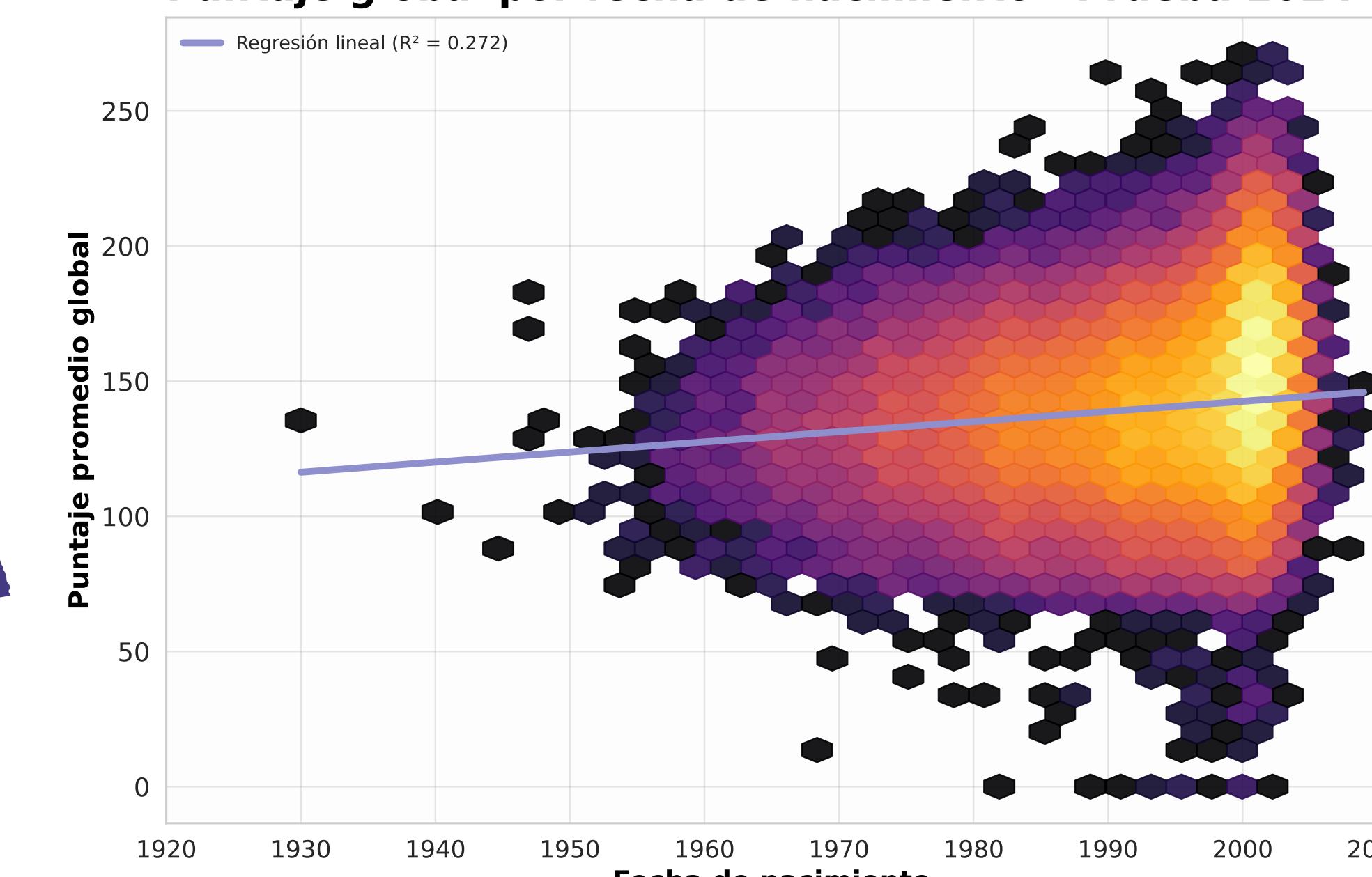
### Núcleos básicos de conocimiento con mayor cantidad de estudiantes



### Mapa de puntajes por departamento



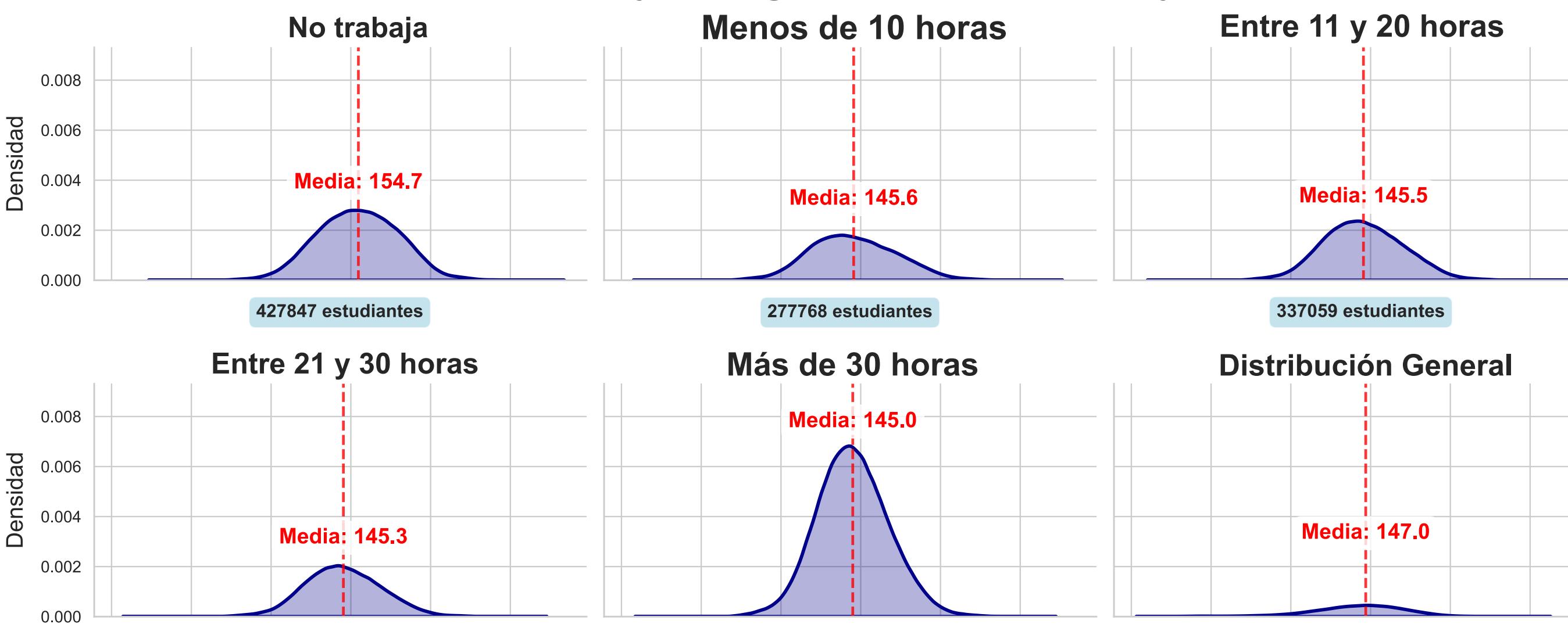
### Análisis de densidad y regresión lineal: Puntaje global por fecha de nacimiento - Prueba 2024



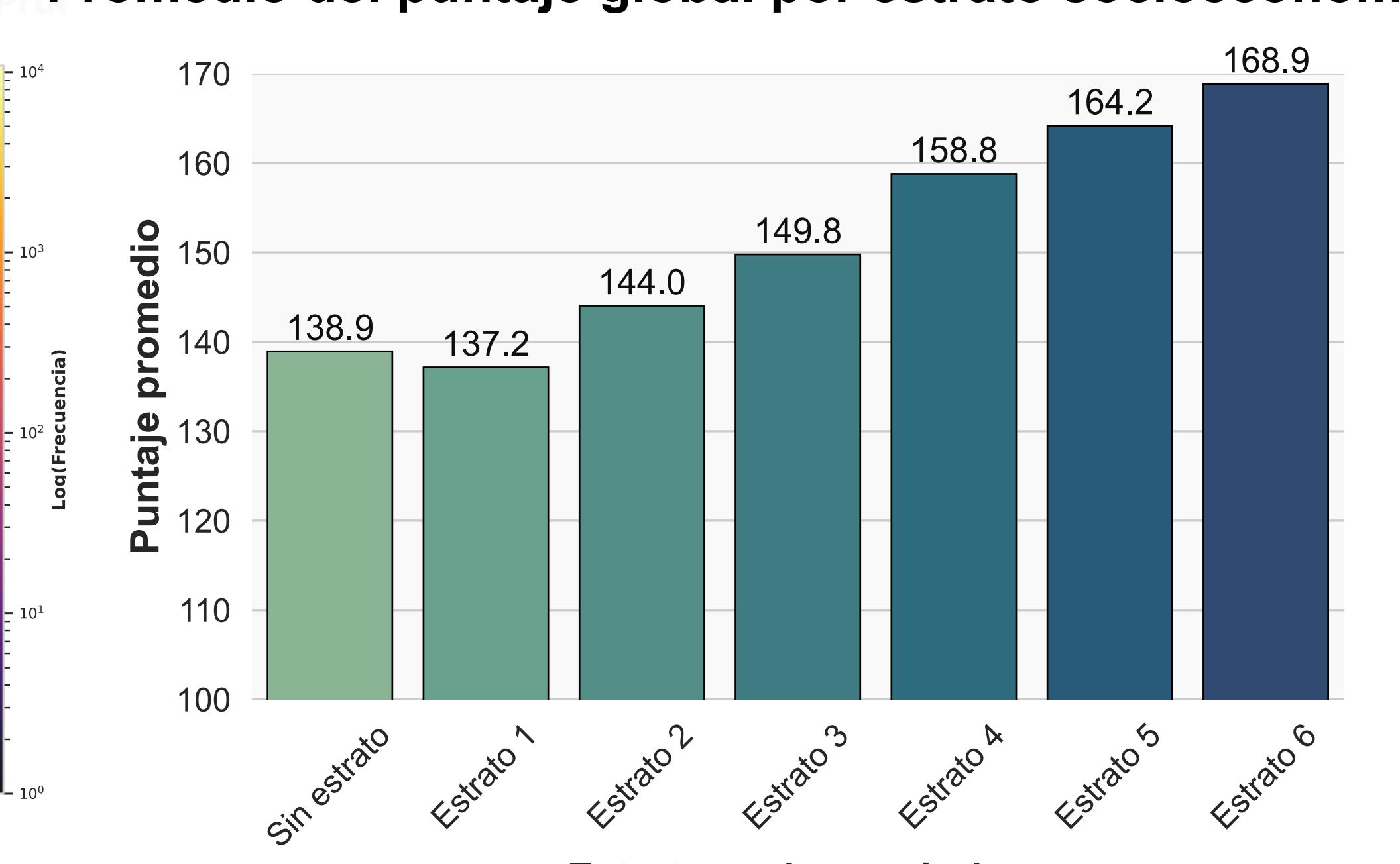
### Mayores puntajes promedio por universidad

UNIVERSIDAD DE LOS ANDES-BOGOTÁ D.C.	23604	188.9
UNIVERSIDAD NACIONAL DE COLOMBIA-BOGOTÁ D.C.	29871	183.4
UNIVERSIDAD EIA-MEDELLÍN	2329	182.7
COLEGIO DE ESTUDIOS SUPERIORES DE ADMINISTRACIÓN-CESA-BOGOTÁ D.C.	1973	182.6
UNIVERSIDAD DE LA SABANA-CHIA	14042	178.6
UNIVERSIDAD ICESI-CALI	7719	177.8
COLEGIO MAYOR DE NUESTRA SEÑORA DEL ROSARIO-BOGOTÁ D.C.	14458	177.7
UNIVERSIDAD NACIONAL DE COLOMBIA-MEDELLÍN	9942	176.7
PONTIFICIA UNIVERSIDAD JAVERIANA-BOGOTÁ D.C.	27389	173.9
UNIVERSIDAD DEL NORTE-BARRANQUILLA	18971	173.1
UNIVERSIDAD EXTERNADO DE COLOMBIA-BOGOTÁ D.C.	8701	171.8
UNIVERSIDAD EAFIT-MEDELLÍN	14259	171.2
ESCUELA COLOMBIANA DE INGENIERIA JULIO GARAVITO-BOGOTÁ D.C.	6279	169.9
PONTIFICIA UNIVERSIDAD JAVERIANA-CALI	10207	169.5
UNIVERSIDAD INDUSTRIAL DE SANTANDER-BUCARAMANGA	20629	167.2
UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS-BOGOTÁ D.C.	20645	167.2

### Distribución de puntajes según horas de trabajo semanal



### Promedio del puntaje global por estrato socioeconómico



**Cluster 1** - "Estudiantes promedio" (37.6%)  
Rendimiento académico medio: Puntaje Global = 153.35

Representan el grupo más numeroso. Son estudiantes típicos colombianos de estratos bajos con rendimiento sólido pero no excepcional. Probablemente asisten a universidades privadas de acceso masivo.  
**Cluster 2** - "Estudiantes en desventaja académica" (28.7%)  
Rendimiento académico bajo: Puntaje Global = 121.42 - el más bajo.  
El segundo grupo más grande. Estudiantes con significativas dificultades académicas.

**Cluster 3** - "Estudiantes adultos de bajo rendimiento" (12.9%)  
Rendimiento medio-bajo: Puntaje global = 133.96

Estudiantes que retoman educación superior después de tiempo. Posiblemente trabajan y estudian simultáneamente.  
**Cluster 4** - "Estudiantes de alto rendimiento" (20.8%)  
Rendimiento académico excelente: Puntaje global = 179.53 - el más alto

Interpretación: Estudiantes de élite académica. Probablemente de universidades de alto prestigio, carreras profesionales demandantes, y con mejor preparación previa y recursos económicos.

