**Apéndice**

**Punto 2**

*Eliminación de campos*

* estu\_etnia
* estu\_fecha\_nacimiento
* estu\_mes\_termino\_bachill
* estu\_recibe\_salario
* estu\_veces\_estado\_estudiantil

Estos campos se eliminaron porque en su gran mayoría son valores nulos.

* percentil\_c\_naturales
* percentil\_global
* percentil\_ingles
* percentil\_lectura\_critica
* percentil\_matematicas
* percentil\_sociales\_ciudadanas

Estos campos se eliminaron porque lo que nos interesa es analizar los resultados (que son campos aparte en la tabla) y los percentiles están completamente determinados por los resultados. Por esto es que los percentiles no son relevantes.

* cole\_cod\_dane\_institucion
* cole\_cod\_icfes
* cole\_nombre\_sede

Estos datos se eliminaron porque son id’s de las instituciones. Como para este conjunto de datos nuestro interés es analizar cada estudiante individualmente, no es interesante saber cuál fue su institución.

* periodo

Este campo se eliminó porque es el mismo para todos los registros (2016-2).

* estu\_consecutivo
* estu\_tipo\_documento

Este campo se eliminó porque solo identifica un estudiante, como es un número asignado al estudiante bajo algún criterio, es evidente que no influye de ninguna manera en los resultados que haya obtenido.

* estu\_cod\_reside\_depto

Este campo se eliminó porque la gran mayoría de filas en este campo tenían un valor nulo.

* punt\_ingles
* punt\_sociales\_ciudadanas
* punt\_c\_naturales
* punt\_matematicas
* punt\_lectura\_critica

Estos campos se eliminaron porque analizaremos los resultados en forma general, es decir, usaremos las columnas de desempeño (1-4 que representan insatisfactorio, mínimo, satisfactorio y avanzado respectivamente). Estos campos de desempeño corresponden a una agrupación en intervalos de los campos de puntaje.

Se eliminaron todas las columnas que representaban una discapacidad en particular (estu\_limita\_\*) por una columna llamada “discapacitada” con valores booleanos que indican si el estudiante tiene alguna discapacidad (cualquiera).