

Trabajo en grupo 2
Estadística descriptiva y análisis exploratorio de
datos

¿Qué variables influyen en el tamaño de un
archivo en disco?



Kenneth Roy Cabrera Torres

3 de septiembre de 2013

1. Objetivo

Unificar y depurar las bases de los distintos grupos sobre el espacio en disco que ocupa un archivo.

Realizar una análisis exploratorio de la base de datos para mostrar las variables que influyen en el tamaño de un archivo en disco.

2. Materiales

1. Bases de datos de los demás grupos.
2. Lenguaje estadístico R.

3. Actividades

1. Se organizan en grupos de mínimo tres (3) estudiantes y máximo (5) estudiantes. No necesariamente los mismos grupos anteriores.
2. Unificar las bases de datos en una sola base de datos.
3. Depurar y unificar criterios de las diferentes variables de la base de datos unificada.
4. Realizar un análisis exploratorio para mostrar la manera en la cual las variables influyen el tamaño de los archivos en disco.

4. Entregables

1. Un trabajo escrito en formato “.pdf” que tenga como mínimo los siguientes elementos:
 - a) Portada: (Nombres de los integrantes, indentificación, universidad, asignatura, etc)
 - b) Resumen del trabajo (un párrafo corto que resuma el trabajo que se presenta).
 - c) Criterios relevantes de unificación de variables o consideraciones a la hora de realizar agrupaciones de niveles en las variables categóricas.
 - d) Desarrollo del trabajo en donde se muestra un análisis exploratorio que indique la manera como las variables influyen el tamaño de los archivos en disco. En análisis exploratorio incluye gráficas y tablas donde se considere necesario.
 - e) Conclusiones.
 - f) Recomendaciones.
 - g) Bibliografía y cibergrafía utilizada en el desarrollo del trabajo.
2. Una base de datos unificada y depurada en formato “.csv”.
3. Un *script* o guión en R con la unificación de las bases de datos y la depuración.
4. Un *script* o guión con el análisis exploratorio de la base de datos unificada.

5. Se deberá enviar al siguiente correo: krcabrer@unal.edu.co ambos archivos. Se colocará la fecha de entrega tan pronto se normalice la situación de la Universidad. Por favor estar atentos al los correos y al desarrollo de los acontecimientos. Recuerde colocar en asunto la sigla EDAED.