

## Análisis exploratorio de datos

Realice un análisis exploratorio completo de la información disponible en alguna base de datos referida a un problema que esté relacionado a su carrera. Dicha base se puede obtener de algún sitio de internet, de libros, pedir a algún profesional que trabaje con datos o armar a partir de información disponible. La misma debe contener al menos una variable de tipo categórica, una cuantitativa discreta y una cuantitativa continua y como máximo 2 variables de cada tipo.

### Presentar el trabajo con la siguiente estructura:

- **Carátula** (título, nombre de los integrantes del grupo, materia y profesores)

- **Introducción**

Objetivo del estudio  
Variables (definición, clasificación)  
Población o muestra en estudio  
Unidad en estudio

- **Análisis**

Por variable:

Tabla (título, frecuencias, porcentajes)  
Gráfico/s (título)  
Medidas descriptivas (medidas de posición y de dispersión)  
Interpretación (se debe hacer para cada tabla, gráfico y medida en base al problema planteado)

Cruzar dos variables:

Tabla (título, frecuencia)  
Gráfico/s (título)  
Medidas descriptivas (medidas de posición y de dispersión)  
Interpretación (se debe hacer para la tabla, el gráfico y las medidas en base al problema planteado)

- **Anexo**

Datos (fuente).  
Si se quiere presentar las sentencias de R (opcional) sólo se pueden encontrar en el anexo no en el cuerpo del trabajo.

**Nota:** las sentencias utilizadas en R para obtener los resultados no tienen que estar en el trabajo, si desean las pueden agregar en el Anexo.

Fecha límite para mostrar la base de datos: 6 de abril de 2017 (mandar por mail a [catalano@fceia.unr.edu.ar](mailto:catalano@fceia.unr.edu.ar) para su aceptación)

Fecha límite de entrega: 19 de abril de 2017 (en clases).

Se debe realizar en equipo de 3/4 personas.

A continuación se adjunta una parte de una base de datos para que sirva de guía.

CPU	Velocidad Bus	Velocidad de reloj	Cache L1	Cache L2	Trans/Millones	Forma - Pines	Voltaje	\$Precio
AMD Athlon 64	800Mhz	2000MHz	128 K	512K/1 Mb	105.9	754	1.5V	128000
AMD Mobile 64	400Mhz	2400MHz	128 K	1Mb	105.9	754	1.45 V	300.000
AMD Athlon 64 FX	800Mhz	2600MHz	128 K	1Mb	105.9	939	1.5V	752.000

Para trabajar los datos en el software R se puede volcar la información en una planilla Excel (como se muestra arriba) y luego guardar el archivo con extensión ".csv". Para importarlo en R se usa la siguiente sentencia:

```
read.csv("C:/data/tp.csv",header=T)
```

#se indica en donde se encuentra el archivo, header es para indicar que las primeras celdas tienen el nombre de las variables.