

**INTELIGENCIA DE NEGOCIOS
Actividad Evaluativa 1: Perfilado de Datos – Diseño de ETL**

Profesor: Mauricio Olivares Faúndez

Objetivo:

Desarrollar, de forma correcta y completa, el perfilado de dos conjuntos de datos, mediante una inspección general (estadística) y otra visual, a partir del cual poder diseñar las condiciones a considerar en un futuro proceso ETL.

Dinámica de la Tarea:

Parte 0: Perfilado de Datos (trabajo de coordinación)

De común acuerdo al interior de cada grupo:

- Cada integrante debe escoger uno de los sets de datos dados
- Entre todos, seleccionar los mismos dos atributos a graficar: coordinarse para que cada atributo sea visualizado con el mismo tipo de gráfico por todos, para efectos de comparación; y que cada atributo sea visualizado de forma distinta; ej. atributo 1 sólo con gráficos de dispersión y el atributo 2 solamente con gráficos de torta

Parte 1: Perfilado de Datos (trabajo individual)

- Realizar una inspección de los datos, completando una tabla resumen con las propiedades (estadísticas) de cada atributo presente, para generar un perfil de los mismos.
- Desarrollar la visualización de los dos atributos escogidos, en base a los tipos de gráficos conversados entre todos.

Parte 2: Diseño preliminar de ETL (trabajo grupal)

A partir de los resultados individuales, realizar un trabajo grupal de integración para generar un diseño preliminar de lo que sería el trabajo ETL de los sets de datos. Para tal efecto, considerar las actividades propias de este tipo de trabajo, vistas en la clase anterior (limpieza, integración, transformación y reducción de datos).