

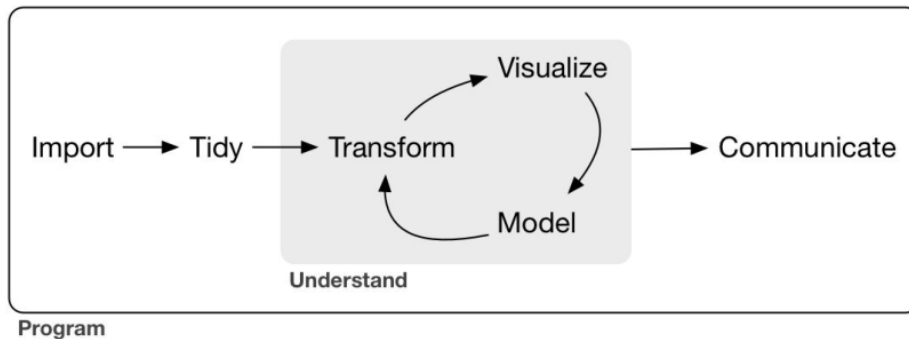
ICO 187 ANÁLISIS DE DATOS

CLASE 06: ¿QUÉ ES ANÁLISIS DE DATOS?

Año: 2021
Profesor: Sebastián Egaña

1. ¿Qué es análisis de datos?

Se puede definir **análisis de datos** como la búsqueda y generación de datos significativos que puedan ayudar a la toma de decisiones. Dicha definición sería principalmente orientada a los negocios. Por otra parte, en términos procedimentales, **análisis de datos** se define como una serie de pasos para la generación de información. El siguiente cuadro, responde a dicho procedimiento en base a Wickham & Grolemond (2017):



Estos pasos, a pesar de estar pensados en torno a Rstudio aplica para cualquier herramienta orientada al análisis de datos. Hablemos un poco de cada uno de estos pasos:

- Import: Importar refiere a la toma de la información ya sea por parte de alguna base de datos, conexión SQL, webscrapping o web API para ser cargados dentro del programa en donde trabajaremos.
- Tidy: Ordenar, refiere al desarrollo de procedimientos para que la data que se tiene sea consistente con la semántica de los datos que se obtuvieron. Recordar que cada columna es una variable y cada fila una observación. En dicho punto ya hemos avanzado en el curso.

Después de esto, existen dos formas de poder obtener información de la base de datos: visualizar o modelar.

- Visualization: Visualizar, refiere a la generación de una descripción visual de los datos obtenidos. Una buena visualización puede resumir de buena manera un set de datos más complejo, pero también puede esconder las preguntas relevantes sobre los datos. Su escalabilidad no es alta, debido a que necesitan de alguien para interpretar.