**Resumen**

El flujo de trabajo del data science esta compuesto de

Untitled.png

.  
Puede existir profesiones que se enfoquen mas a cada una de fases, ***no existe un perfil de data science que se encargue a todo el flujo de trabajo***.  
.

**¿En que partes del flujo de trabajo se necesita de estadística?**

.  
Todos las partes del flujo requiere del conocimiento en ciertas ramas de la estadística. La estadística descriptiva se va a emplear más en los dos primeros bloques de trabajo.

.

* **Ingesta de datos y Validación :** Se encarga de todo el procesamiento de **ETL (Extract Transform Load)** obtener los datos, limpiarlos y estructurarlos, crear pipelines de análisis automatizado, es decir que transformaciones vamos a realizar a los datos para que estén listos para el caso especifico de estudio que vamos a realizar.
* **Preparación y entrenamiento del modelo:** En este bloque se va a realizar un análisis exploratorio de los datos con estadística descriptiva, entender correlaciones y realizar posibles reducciones de datos.
* **Evaluar el modelo, Producción e Interacción:** esta parte del flujo se basa mas en la estadística inferencial.

Responder