

## ¿PARA QUÉ QUIERO UNA BASE DE DATOS?

Profesor: Teófilo Chambilla

Usted cuenta con dos datasets en formato CSV en los cuales se ha recogido datos de servicios de taxi de la ciudad de Nueva York en diciembre del 2017. El primer dataset “driver.csv” contiene datos de los conductores como el nombre del chofer, numero de licencia de automovil, nombre de la empresa de transporte que pertenece, etc. El segundo dataset “service.csv” corresponde a un historial de servicios de taxi incluyendo datos de hora y lugar de partida, cantidad de pasajeros, distancia recorrida, monto de pago, etc. Los servicios de taxi estan relacionados con los choferes mediante el campo “Vehicle License Number”. Ademas, se adjunta dos PDFs (driver info, service info) que contiene la descripción de los campos(diccionario de datos)

Mediante el uso de cualquier herramienta que considere util (lenguaje de programación C++, Python, R) menos un DBMS y Excel, responda las siguientes preguntas:

- C1. ¿Cuales son los servicios de taxi que realizo Paulino Ramon?
- C2. ¿Cuáles son los choferes que hicieron la mayor cantidad de servicios de taxi? Cruce la información de ambos archivos para obtener los resultados requeridos. Orden de mayor a menor y reportar el Top 10.
- C3. ¿Algún chofer No Activo realizo servicios de taxi?. Reporte los choferes inactivos y la cantidad de servicios realizados.
- C4. ¿Cuántas empresas de taxi han registrado servicios de taxi? Verifique cuantas empresas (Base Name) tienen al menos un servicio de taxi asociado.
- C5. ¿Qué empresa(s) de taxi (Base Name) reportaron servicio(s) el día 28 de diciembre?

**Entregable:** Usted deberá entregar al final del laboratorio un archivo PDF escrito en **LATEX** con las respuestas a las consultas indicadas líneas más abajo. Para obtener dichas respuestas, Ud. puede utilizar cualquier herramienta que considere útil (C++, Python, otro lenguaje de programación o búsqueda manual) menos un Sistema de Base de datos(SGBD) y Excel teniendo en cuenta lo siguiente:

1. Integrantes del grupo (máximo 4).
  2. Método empleado: explique el procedimiento que usó para resolver cada consulta.
  3. Muestre el resultado de cada consulta de forma clara.
-