



# Business Intelligence & Analytics

## TP Ingesta de datos



Ing. Javier Souza

[n.javier.souza@gmail.com](mailto:n.javier.souza@gmail.com)



# Actividad

---



1. Ingestar el archivo csv adjunto según el patrón “full load”:
  - a. Construir un artefacto que sirva de loader para ingestar el archivo en una base de datos a elección tomando el tiempo de comienzo del proceso y fin del mismo para determinar el tiempo total de procesamiento.
  - b. Realizar esta prueba 10 veces tomando los tiempos en cada iteración.
  - c. Documentar y compartir el tiempo promedio de ejecución bajo este patrón.
2. Usando la tabla creada y cargada en el punto 1, copiar los datos a una segunda tabla usando el patrón “incremental load”. Copiar intervalos de a una hora para que el proceso pueda iterar varias veces. Repetir esta operación 10 veces tomando los tiempos de ejecución. Usar el campo “updated\_at” como bookmark.
3. Repetir el paso 2 pero multiplicando el contenido de la tabla original de manera que en lugar de tener 100K registros, tenga 1 millón.

# Entregables

---



1. Código fuente del proceso.
2. Reporte de los tiempos de procesamiento de cada uno de los 3 escenarios distintos.
3. ¿Puede sacar algún tipo de conclusión respecto a los tiempos tomados? ¿Podría explicar el por qué de estas diferencias de tiempos?

# THANKS!

## Any questions?

You can find me at

[n.javier.souza@gmail.com](mailto:n.javier.souza@gmail.com)

