TP Business Intelligence

L’ingestion de données désigne un processus d’absorption de donnée provenant de diverses sources et qui pourront être déposer (stocker) ou analyser. Le traitement des données permettra par la suite de regrouper les données. Il permet de les illustrer à l’aide de diagramme, de graphe (etc…) mais aussi dans le même temps d’obtenir un résumé numérique. Cependant ces données doivent être nettoyées dans un format qui est facile d’utilisation.

* Ingestion par lots :

Par lots, les données sont transférées à intervalles réguliers et par portions

* Ingestion en temps réel (streaming)

Etant sensible au facteur temps, les données sont importés tels quelles sont énoncées. En effet elles sont extraites, traitées, et stockées telles qu’elles sont émises.

* Architecture Lambda

Elle regroupe les deux premiers en en traitant les données par lot pour étudier l’ensemble des données lots. Dans le même temps il utilise ingestion en temps réel pour étudier les informations liées au facteur de temps.

SOURCES :

<https://whatis.techtarget.com/fr/definition/ingestion-de-donnees>

<https://docs.microsoft.com/fr-fr/azure/data-explorer/ingest-data-overview>

<https://blog.stack-labs.com/code/data_ingestion_state_the_art/>