Data Extract Feature

Summary

The Total Expert analytics data model (ADM) provides a view into your marketing, engagement, sales, and productivity data. It provides this data in a format that is easy to consume into a data warehouse or directly into business intelligence and reporting tools.

Data is provided at the most granular level possible to enable detailed reporting and integration with data sources not available in the Total Expert platform. The model loosely follows a familiar star schema approach and is composed of:

- *Facts* – The measurable events or activities that make up a business process
- *Dimensions* – The attributes or nouns that describe the process

Each table in the model is provided as a single daily file, delivered directly to your chosen destination, as described below. You can subscribe to one, many, or all tables in the model.

The tables are described in detail in *Data Lake Extracts Data Dictionary*.

Delivery Methods

You can have your data delivered via either:

- **S3** – Delivery to an S3 bucket you own, using AWS access key and secret for authentication. To use S3 delivery, you must provide the following:
  - The bucket's Amazon Resource Name (ARN)
  - The bucket's prefix (path) where files should be delivered
  - The access key and secret with write access to the bucket/prefix

> **Note**
> By default, files will be written with `bucket-owner-full-control` permission in S3.

- **SFTP** – Delivery via FTP over SSH using a username/password authentication. TE does not currently support asymmetric key–based authentication. To use SFTP delivery, you must provide the following:
  - SFTP host name
  - Port number
  - Destination path where the files should be delivered
  - User name and password for a user with write access to the destination path

File Formats

You can have your data delivered in any of the following formats:

- **Comma-separated variable (CSV)** – Text file with data points delimited by commas (,). Includes a header row.

- **Delimited** – Similar to a CSV file, but with a pipe symbol (|) or other configurable delimiter instead of a comma.
- **Parquet** – Binary, compressed, columnar file format optimized for analytics workloads.

Extract Types

Data extracts are provided as either

- **Initial full extract** – A full extract containing all records for a given table of your data.
  - This is typically delivered as the first extract when an extract is requested. Thereafter, delta extracts are provided.
- **Delta** – Daily extract containing only the changed records since the last successful delivery.
  - These contain an additional `change_type` attribute that indicates the type of operation that occurred, where:
    - `I` = insert
    - `U` = update
    - `D` = delete