**StreamSets Transformer for Snowflake**

# Healthcare and Life Sciences

## Overview

The pharmaceutical and life sciences industry deals with vast amounts of data daily. This data can be in different formats, including structured, unstructured, and semi-structured data. With the proliferation of electronic health records, clinical trials data, and other data sources, it has become more complex and harder to analyze, leading to a greater need for data transformations.

Data transformation is the process of converting data from one format to another. The purpose of data transformation is to ensure that data is in a consistent format, making it easier to analyze and derive insights. Inconsistent data models/formats can lead to errors, making it difficult to draw insights and make informed decisions. For pharmaceutical and life sciences organizations, data transformations play a critical role in influencing improved patient outcomes, accelerating drug research and discovery, as well as optimizing supply chain management.

## Example of how StreamSets Transformer for Snowflake can be leveraged in healthcare:

StreamSets Transformer for Snowflake is a highly versatile tool that aids in processing and visualization of complex datasets in the healthcare sector, making it possible to comprehend patterns and trends in medication and prescription use. In this context, the tool is used to ingest and integrate data from multiple tables within Snowflake, where patient, provider, and prescription information is stored but not optimized for analytics. Given that these tables consist of both transactional data - like patient and provider details, medications prescribed, as well as specific patient-

provider encounters at varying points in time, the process involves bringing together disparate types of information in a meaningful way.

The power of StreamSets Transformer for Snowflake lies in its ability to perform operations such as joins and unions effortlessly, without requiring the user to write any code or deal with the complex nature of SQL. Normalized tables containing patient and historical prescription data can be quickly prepared, allowing for the creation of analytical models with dimensions and facts that are ready to be consumed. Furthermore, deduplication eliminates redundant data and retains only the unique, relevant information. This tool is instrumental in generating a unified, clean dataset that can be used to identify which doctor is prescribing a certain medication more often than others, or which patients, broken down by age and gender, are using a certain medication more frequently. The key takeaway is that StreamSets is not just about moving data from one place to another, but rather transforming raw, disparate data into actionable insights, enabling data-driven decision-making across healthcare.

## Other data transformation examples:

**Slowly changing dimensions (SCDs)** is a critical data transformation technique in the pharmaceutical and life sciences industry. SCDs can be used to track changes in patient data over time, such as changes in patient demographics, disease progressions, and treatment efficacy. By capturing changes from separate tables, users can derive more accurate analysis of patient outcomes. For example, SCDs can help identify patterns in patient data that can indicate the effectiveness of certain treatments or medication.

This can ultimately improve patient outcomes by enabling more personalized and targeted treatment plans.

**Denormalization** is a data transformation technique that involves combining multiple tables into a single table to reduce the number of joins required to access data. For healthcare organizations, denormalization can be used to optimize supply chain management by providing faster access to critical supply chain data. For example, they may have multiple databases containing information about suppliers, inventory, and distribution. By combining datasets into a single table, organizations can improve query performance and enable faster access to critical supply chain insights. This process helps them reduce costs and improve overall operational efficiency.

**Window Functions** - Transformer for Snowflake seamlessly supports the application of window functions to help derive insights from complex datasets. For example, to assess medication sales across different regions, window functions can calculate which regions record higher sales for a particular medication in comparison to other parts of the state or country. This enables users to identify areas where a specific medication might have lower prescription rates, which can present opportunities for expansion or reinforcing the benefits of the medication in those areas. The ratio-to-report function can be utilized to compute and analyze sales.

**Dynamic Pivots** – Another application of this tool in healthcare is through the support of dynamic pivots. Transformer for Snowflake allows users to apply parameters to compare medical exams requested by doctors for specific symptoms. For example, if a patient had a fever, the tool would allow the hospital to administer the correct examination in a timely manner for the specified symptom. The data for all the exams and the corresponding symptoms would be in varying formats, but the dynamic pivot capability would allow users to visualize the data in a more structured and accessible format for easier decision making.

## About StreamSets

StreamSets, a Software AG company, eliminates data integration friction in complex hybrid and multi-cloud environments to keep pace with need-it-now business data demands. Our platform lets data teams unlock data—without ceding control—to enable a data-driven enterprise. Resilient and repeatable pipelines deliver analytics-ready data that improve real-time decision-making and reduce the costs and risks associated with data flow across an organization. That's why the largest companies in the world trust StreamSets to power millions of data pipelines for modern analytics, smart applications, and hybrid integration.

**To learn more,** visit www.streamsets.com and follow us on LinkedIn.