SQL queries are a widespread technique to analyse datasets. In BigDataStack, data skipping for SQL queries has been further researched and developed. This technology is relevant when the dataset resides in an Object Store. This research has already been contributed to the IBM SQL Query service for now as a closed beta.

**Input**

An SQL query against a given dataset which is stored in an Object Store.

**Output**

The result of the query

**Initial TRL**

IBM was already working at integrating with the IBM SQL Query Service, however this was for the basic data skipping technology. BigDataStack permitted to enhance in a much non-trivial way the basic technology. These enhancements have been integrated within the IBM SQL Query Service.

**Final TRL**

TRL 9
End Users
Since our technology is already in advanced stage of integration with the IBM SQL Query service, the customers of this service are our natural customers.

Key Features and Benefits
This technology reduces the size of the data scanned for answering the SQL query. Thus it a) it reduces the price to be paid by the customer b) accelerates the query

Essential Information for Users
Prior to using this technology, the dataset has to be indexed as explained in the documentation of the IBM SQL Query Service.

Standards involved in the development of the component
None

Implementation in BigDataStack Use Cases
For now, the Data Skipping technology has been tested on Real time Shipping, however the Connected Consumer seems to be an even better use case for it. It is also planned to apply it in the third use case (insurance).