**Insights on mySQL using Global Trade Item Number (GTIN) data structures**

Randy Lisbona, Marvin Scott, Vinh Le

**Abstract**

MySQL is a popular open source database format. GTIN (Global Trade Item Number) is a family of product ID codes, used worldwide between trading partners, to identify products via linear or 2D barcodes on the product or packaging. This project will provide insight and first impressions on the ease of use of mySQL using GTIN example data.

An open source GTIN subset > 100K records (.csv, .sql, or JSON format) will be downloaded and installed on a local mySQL database

instance. Open source documentation will be reviewed to better understand the GTIN data. The schema will be checked for normalization and modified if necessary, Primary and foreign key indexes will be created/modified as needed. A graphical schema of the database will be created with example data from each table to visualize the relationship between tables. Exploratory queries will be created to summarize selected fields such as brand, and packaging level. A summary list of mySQL commands used in the analysis will be provided along with impressions on ease of use, intuitiveness, and effectiveness

**Introduction:**

As a group this project was selected due to the similarity of the tasks that were performed and learned during the live sessions. The exercises included in the GTIN data sets were discovered sufficient to illustrate the level of effort put together as a team and demonstrates our understanding of Database Management and Design. The following examples provided by our group of the GTIN datasets involves detailed information of the tables and Database that were created, the relationships between each entity, entity and schema metadata, description of each table and a bit of analysis. A description of the project can be found in entirety, please refer to the GITHUB URL listed in the Appendix.

A total of 15 tables were created and analyzed in the GTIN\_13 Database. The **BRAND** entity has four columns and a total 4151 tuple count. The entity provides website details of the several brand names such as Clorox and Glade products as well as many others. The Primary Key of **BRAND** entity is the **BSIN** attribute and is a unique identifier for each brand name or product. The **BRAND\_GROUP** entity has three columns and only 3 tuples and was not found to provide helpful details of the GTIN datasets, the Primary Key of **BRAND\_GROUP** entity was also the **BSIN** attribute. The **BRAND\_OWNER** entity has four columns and a total tuple count of 32. The entity provides website details of the several owners such as P&G and Coca-Cola and has a Primary Key on the **OWNER\_CD** attribute which is a unique identifier for each owner. The **BRAND\_OWNER\_BSIN** entity has two columns and a total 581 tuple count. The entity provides a unique relationship between the **BSIN** ids with their **OWNER\_CD** ids showing a one to many relationships between the BSIN primary key and the owners. The **BRAND\_TYPE** entity has two columns and only 2 tuples, the Primary Key **BRAND\_TYPE\_CD** gives descriptive details on the **BRAND\_TYPE\_NM.** The **GS1\_GCP\_RC** entity has three columns and only 16 tuples and a Primary Key **RETURN\_CODE** and provides descriptive categorical details of return options. The **NUTRITION\_US** entity has a total of 30 attributes and 231 tuples. The Primary key for this entity is the GTIN\_CD, this entity provides nutritional facts of the products catalogue in the datasets provided.