**Insights using MySQL to explore Global Trade Item Number (GTIN) data structures**

Randy Lisbona, Marvin Scott, Vinh Le

**Abstract**

GS1 is a global organization that maintains standards for numbering systems used between trading partners. 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. Commonly referred to as UPC codes in the USA. This project will provide insight on the format and usage of GTIN product codes.

An open source GTIN subset > 100K records will be downloaded and installed on a local mySQL database instance. Open source documentation will be reviewed to better understand the GTIN data. 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.

Figure - GTIN formats



# GS1.org History

GS1.org is the global standards organization that manages a system of unique identification numbers used between trading partners in over 150 countries worldwide. GS1 traces its roots to the Uniform Product Code Council (UCC) which was established in the USA in 1973 to manage UPC barcodes used in North America, and to the European Article Numbering Association (EAN International) established in 1977 to develop a compatible barcode identification system outside of North America. GS1 was launched in 2005 to combine the two standards organizations into one international organization. [1]

# GTIN formats

GTIN (Global Trade Item Number) is just one of the numbering systems managed by GS1:

* **Global Trade Item Number (GTIN).**
* Global Location Number (GLN).
* Serial Shipping Container Code (SSCC).
* Global Returnable Asset Identifier (GRAI).
* Global Individual Asset Identifier (GIAI).
* Global Service Relation Number (GSRN).
* Global Document Type Identifier (GDTI).
* Global Shipment Identification Number (GSIN).
* Global Item Number for Consignment (GINC).
* Global Coupon Number (GCN).
* Component / Part Identifier (CPID).

A GTIN consists of four parts, an optional Application Identifier, the Company Prefix, the Item Reference, and a check digit. See Figure 1.

* GTIN-8 is a truncated 8 digit GSI identification key used on packages with limited label space, such as chewing gum.
* GTIN-12 is 12 digit GSI identification key consisting of a U.P.C. company prefix, item reference, and check digit.
* GTIN-13 is a 13 digit GSI identification key consisting of a GS1 Company prefix, item reference, and check digit
* GTIN-14 is a 14 digit GSI identification key consisting of an indicator digit (1-9), GS1 company prefix, item key, and check digit.

## Indicator Prefix



Table - Figure 3 Decoding

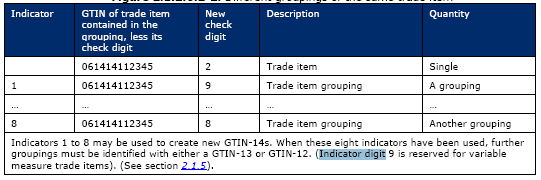


Figure - Indicators

The Indicator Prefix Digit is used to define a grouping, or packaging level. For instance, in Figure 2, we see an example GTIN of 061414112345, let’s say this was a 2oz bag of candy corn. Adding an indicator prefix of “1” could denote a bulk pack of 12 2oz packages, indicator prefix “2” could be used for a 24 pack, indicator prefix “3” could be a “Gross” pack (144) of 2oz candy corn packages, and so on. If more than 8 packaging levels are needed, then a new GTIN-12 or GTIN-13 is required.

## Variable Measure Trade Items

Indicator prefix “9” is reserved for variable measure trade items. Variable measure items are trade items that cannot guarantee consistent weight, size, or length due to the production process (e.g. meat, bulk cheese).

## Data Carrier

A data carrier is a means of representing GS1 identification codes in a machine readable form. 1D Linear barcodes such as UPC-A or EAN-13 are widely used. Matrix symbols such as QR code and Data Matrix are also supported, as well as RFID tags.

## Examples

Figure 3 shows an example barcode from a box of chocolates. The format is GTIN-13 and the data carrier is an EAN-13 Linear barcode.



Figure - EAN-13 barcode symbol

Table 1 shows decoded values from figure 3. The first three digits (400) indicate the country that issued the Global Company Prefix. In this case it was Germany. The country that issued the GCP is not necessarily the home country of the company. For instance, a US based company may get a GCP issued in a foreign country. The Prefix plus the company code 81555 results in a unique company code worldwide (40081555).

The item code is 02020, Companies assign their own item codes. The calculated check digit is 7. Check digit calculations are outlined in section 7.9 of the GS1\_General\_Specifications document. [1]

Figure 4 is a chocolate candy bar from the same company, note the item code of 03300 and check digit 9.

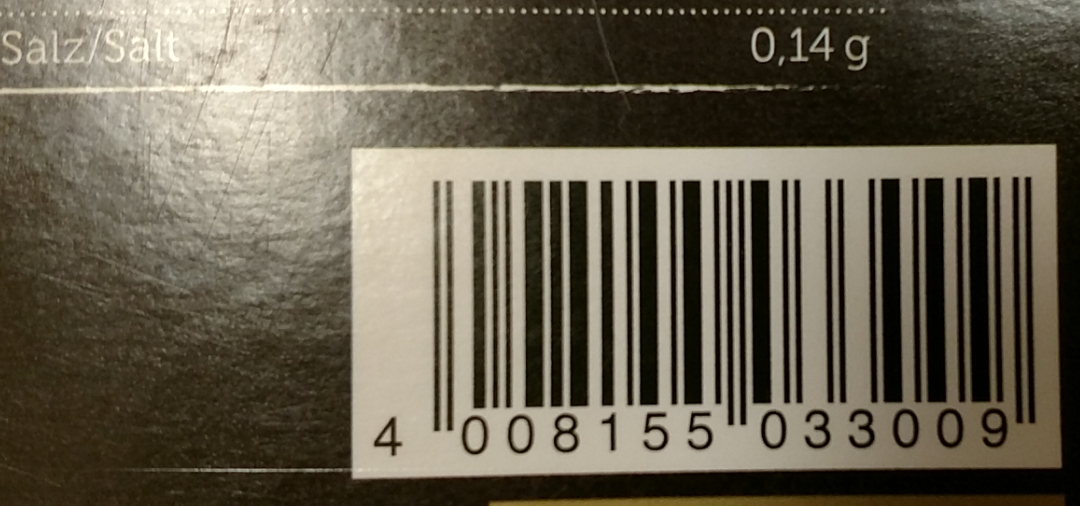


Figure - EAN-13 barcode symbol 2

Figure 5 shows a GTIN-12 encoded in a UPC-A linear barcode. 8 indicates the GPC was assigned in the USA, 84912 is the GPC for Post Consumer Brands, LLC. The Item number is 00471 and it was from a 581-gram box of Grape Nuts Cereal.



Figure - UPC-A barcode symbol

# Product Open Data



Figure - Product Open Data

GS1 does not maintain a worldwide list of Global Trade Item numbers (GTIN), GS1 simply assigns Global Company Code prefixes to ensure unique GTIN values. Each company using GTIN maintains their own list of item numbers and shares this with their trading partners. There are several initiatives to create a global open source GTIN list. For this project the Product Open Data database was used. This is far from a complete list in its current state, but does demonstrate the type of data linked to GTIN’s.

## Download and install POD

The POD database was downloaded and installed locally in MySQL. [2] The Specification file was also downloaded. [3] Queries were written to pull sample records from each table.

webhosted mySQL like <https://www.000webhost.com/>

Background on GTIN from [www.gs1.org](http://www.gs1.org)

, who uses it, can we find/estimate worldwide usage. Describe various formats for GTIN =(companycode+Itemcode) How many unique companies and itemID’s can it encode. We have GTIN-13, there are a few others. Appears that there is not an open source version of the complete database, trading partners publish their data to companies selling the products? Can we explain why? Explain why price is not in this database.

EER diagram– Randy working on this. Steps in MySQL to create EER, Reverse engineering the schema in MySQL didn’t bring in the tables. Data Export to SQL script, import mysql reverse engineer script does work, steps to explore and define relationships. Need explanation of nearly empty skeleton tables found with just a few records. Base our queries on the

EER Graphic

Queries: Basic stats on each table, records, keys, do we need to add indexes? File size of zipped, unzipped, and MySQL database. Create several

example summary queries, a few charts:

1. ItemCount by Company
2. Brand Count by company
3. Average Package size ?

Use Excel PowerQuery for charts, need to install SQL connector <https://support.office.com/en-gb/article/Connect-to-a-MySQL-database-Power-Query-8760c647-88b9-409d-b312-6ea8f84a269b?ui=en-US&rs=en-GB&ad=GB>

Insights:

1. Average package size
2. Min max avg number of items per company.
3. Brands per company
4. Ease of using MySQL for relationships, queries, data exploration

Summary, what we learned with each of us working on our own local copy, vs a web version if we can get that working.

Max length of paper = 4 pages

# References

|  |  |
| --- | --- |
| [1] | GS1.org, "The global Language of Business," 11 2016. [Online]. Available: http://www.gs1.org/gs1-source/latest GS1\_General\_Specifications.pdf. |
| [2] | "Product Open Data," 01 01 2014. [Online]. Available: http://www.product-open-data.com/download/ "POD Database - Dump". [Accessed 11 2016]. |
| [3] | "Product Open Data," 13 11 2013. [Online]. Available: http://www.product-open-data.com/download/ "POD Database - Specifications". [Accessed 11 2016]. |

**Appendix A Data sources and related works**

Github Repository <https://github.com/rlisbona/MSDS-7330-Term-Paper-1>

Figure Data Sources

|  |  |
| --- | --- |
| GS1.org – Standards Organization for GTIN | <http://www.gs1.org/> |
| Product Open Data – Subset of GTIN | <http://www.product-open-data.com/en/1-home.html> |
| POD database SQL Create and Load | <http://www.product-open-data.com/docs/pod_web_2014.01.01_01.sql.gz> |
| POD database Specification | <http://www.product-open-data.com/docs/POD-SPECS-2013.11.13_01.xlsx> |

Figure - Related Works

|  |  |
| --- | --- |
| Open EAN/GTIN Database | http://opengtindb.org/ |
| Outpan | https://www.outpan.com/ |
| Datakick | https://www.datakick.org/ |
| EAN-Search | http://www.ean-search.org/ |

**Appendix B POD database Schema**

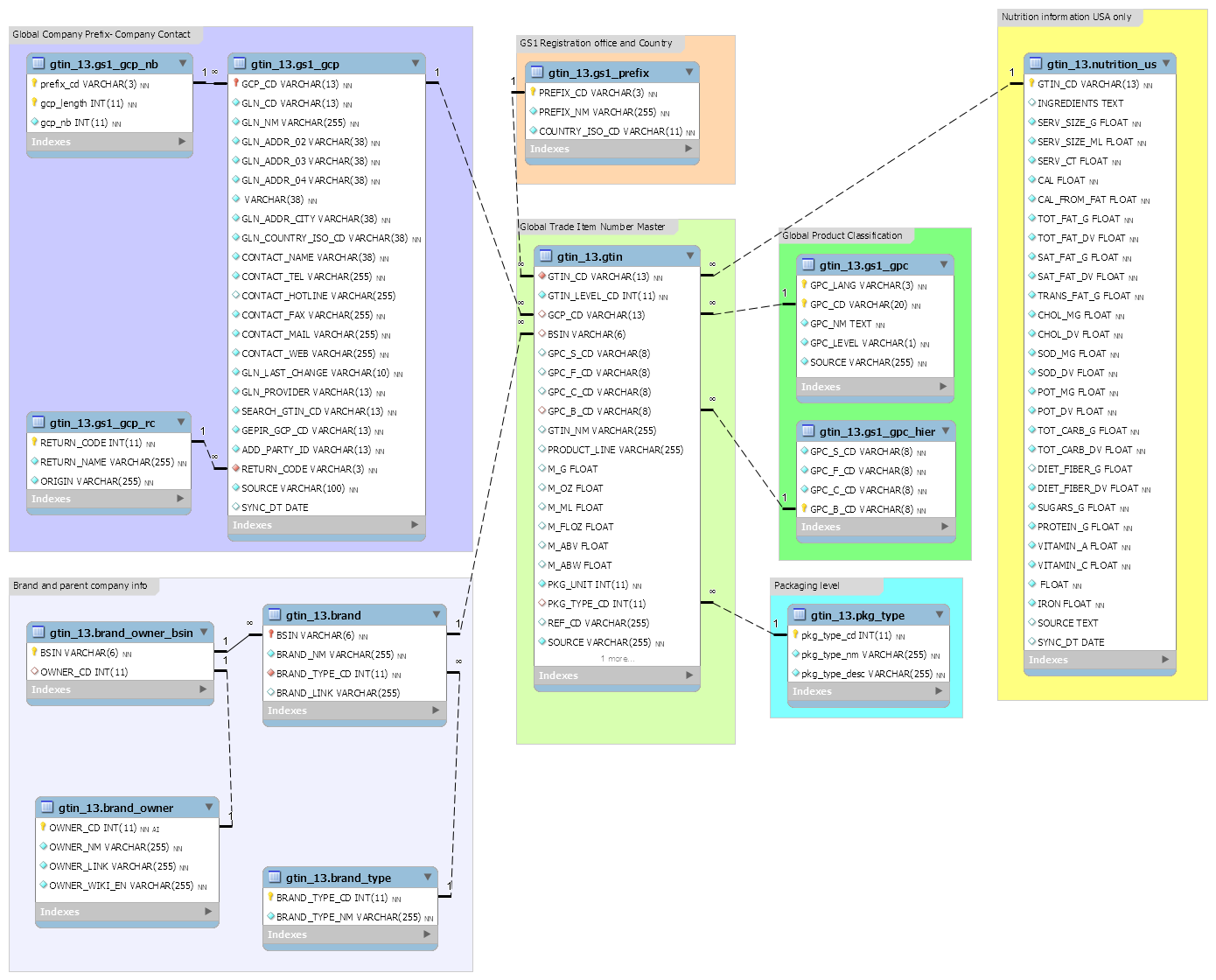


Figure - POD database Schema