General Overview:

1. The market sensing project needs 3 flat csv files loaded into Caspian Snowflake via STFP site. These 3 files each serve a different purpose
2. The file **SI\_ECON\_MNEMONIC\_MAPPING** provides mapping all of the mnemonics from FRED, IHS and Moody’s into 7 distinct tables in zone 4.
3. The file **SI\_CATEGORY\_MARKET\_SIZE** provides a calculated Total Market Size by Category as well as the SBD Market Share by Category.  **This is required to derive a scaling factor used for reporting and market share conversion**.
4. The file **SI\_BRAND\_MANUFACTURER\_MAPPING** provides mappings to identify brands that belong to SBD (and in the future, for competitors**).  This mapping is required to determine the percentage share by brand.**

**Current Existing Process** -

The current process is the market sensing work flow is pointed at these 3 files on a desktop.

The Market Sensing Team needs to be able to periodically realign and supplement these 3 files.

The proposed solution is **to extract a .CSV file monthly from an SFTP Folder** (to be created by development team).

**High Level Data Requirements –** At a high level, we need to acquire Reference data for 3 tables.Create a schema named **“SI\_MARKETPLACE\_REF\_DATA**” and create the following tables in Snowflake

1. SI\_ECON\_MNEMONIC\_MAPPING
2. SI\_CATEGORY\_MARKET\_SIZE
3. SI\_BRAND\_MANUFACTURER\_MAPPING

The data dictionary of the tables is included in Appendix A. **The data acquisition is to be done via sftp pull.**

**Ingestion and Acquisition**

**History Requirements & Frequency:**

SI\_ECON\_MNEMONIC\_MAPPING **One initial load** of the historical file is to be ingested.  Then **a delta load.**

SI\_CATEGORY\_MARKET\_SIZE One **initial load** of the file is needed then **the file needs to be appended every quarter to the existing dataset**

SI\_BRAND\_MANUFACTURER\_MAPPING

One initial load of the file is needed then **the file needs to be appended every quarter to the existing dataset**

**Acquisition :** The data pull mechanism is to be developed **within SBD sftp location for receiving the single CSV file.**

**Ingestion** The files will be comma delimited

**SI\_ECON\_MNEMONIC\_MAPPING**

|  |
| --- |
| Mnemonic Key |
| Mnemonic Category |
| Source System |
| Mnemonic Description |

****

**SI\_CATEGORY\_MARKET\_SIZE**

|  |
| --- |
| Date |
| Category |
| Metric |
| Value |

**SI\_BRAND\_MANUFACTURER\_MAPPING**

|  |
| --- |
| Brand |
| Manufacturer |