**Tables Manifest**

Since we have so many tables, we wanted to provide an overview of all of them to make navigation easier:

**Core F2 Tables**

|  |  |
| --- | --- |
| **Table** | **Description** |
| DOC.csv | Standard documents table |
| LIB.csv | Library table; Includes standard features, and also an added variable of “period” which reflects the general time-period of the cookbook |
| TOKEN.csv | Standard TOKEN table; includes part of speech tagging and term\_str |
| VOCAB.csv | Vocab table; includes word frequency, a number dummy variable, a stop-word dummy variable, stems, and three separate TFIDF values based on bags of time period, book, and recipe. |

**Embeddings**

|  |  |
| --- | --- |
| **Table** | **Description** |
| Embeddings\_mid1800s.csv | Contains word embeddings for the corpus of cookbooks written in the mid-1800s. The vector column represents the embeddings generated from Word2Vec, and the x and y are the coordinates generated by T-SNE; links to VOCAB table via “term\_str” |
| Emeddings\_late1800s.csv | Contains word embeddings for the corpus of cookbooks written in the late-1800s. The vector column represents the embeddings generated from Word2Vec, and the x and y are the coordinates generated by T-SNE; links to VOCAB table via “term\_str” |
| Embeddings\_1900s.csv | Contains word embeddings for the corpus of cookbooks written in the early-1900s. The vector column represents the embeddings generated from Word2Vec, and the x and y are the coordinates generated by T-SNE; links to VOCAB table via “term\_str” |

**Sentiment**

|  |  |
| --- | --- |
| **Table** | **Description** |
| Emolex\_sentiment.csv | The emolex lexicon. This was not generated by us, but is necessary for our code to run; links to the VOCAB table through “term\_str” |
| Sentiment\_book.csv | Contains overall sentiment scores for each book; the NRC values come from the emolex lexicon, while the VADER scores come from the VADER engine |
| Sentiment\_period.csv | Contains overall sentiment scores for each time period; the NRC values come from the emolex lexicon, while the VADER scores come from the VADER engine |