TMSIS Data Documentation

This website provides documentation for our TMSIS datasets for researchers within / collaborating with the Yale Medicaid Lab. On the site, you can find resources explaining the structure of the datasets, recommendations for working with large datasets, and example code from other projects that have used the same data.

Hive-Partitioned Datasets:

CMS gives us the data in a bespoke fixed-width file format called `.fts`. Working with these files is generally slow (as the file format is not optimized for database queries), and cumbersome (as one must implement their own parsing logic to read the files). We parse and format the TMSIS data into a series of hive partitioned parquet files, a standardized file format that is fast to read, supports modern database operations, and enforces strong types, so you never have to worry that your files have been read in incorrectly.

To get started, with the files, you can log into Milgram and open the directory where they are stored on the server. The files are currently located at `/gpfs/milgram/pi/medicaid_lab/data/cms/ingested/TMSIS_TAF`.

```
cd /gpfs/milgram/pi/medicaid_lab/data/cms/ingested/TMSIS_TAF
```

Opening this folder, you will see the following files:

```
TMSIS_taf/
    taf_demog_elig_base
    taf_demog_elig_dates
    taf_demog_elig_disability
    taf_demog_elig_hh_spo
    taf_demog_elig_mngd_care
    taf_demog_elig_mny_flw_prsn
    taf_demog_elig_waiver
    taf_inpatient_header
    taf_inpatient_line
    taf_inpatient_occurrence
    taf_long_term_header
    taf_long_term_line
    taf_long_term_occurrence
    taf_other_services_header
    taf_other_services_line
    taf_other_services_occurrence
    taf_rx_header
    taf_rx_line
```

The files are grouped into 4 groups, the pharmacy files (taf_rx_.*), the long-term care files (taf_long_term_.*), the inpatient files (taf_inpatient_.*), and

the demographic and eligibility files (taf_demog_elig_.*). You can find quick links to the resdac documentation for each file type at the bottom of this page.

Useful Data Links:

- Documentation for other services files
- Documentation for demographic and eligibility files
- Documentation for inpatient files
- Documentation for long term care files
- Documentation for pharmacy files