Database Overview
June 2021 HA Database

The HA Database is the source of data for the HA Website and HA Desktop. The HA Database is located in a Microsoft Azure datacenter and stored as an Azure SQL Database. The HA Database can be accessed over an internet connection with a User ID and Password. The HA Database is created from publicly available government data without any HIPPA restrictions. The HA Database has 4 sources of data. The first source is the POS (Provider of Service) data used to identify all the Medicare Providers that should file Cost Reports. The second source is CMS Medicare Cost Reports data obtained directly from the www.CMS.gov website. The third source is the Census Bureau Population Estimate for States and Counties data obtained directly from the FTP2.Census.gov website. The fourth source is the CMS Final Rule data in a collection of tables used to provider additional information. These 4 sources of input data are processed/converted into tables in Microsoft SQL Server. At this point, the input data is significantly enhanced with additional description information and analytical information to create a complete solution for all CMS Medicare Cost Reports for the 8 providers required to file CMS Medicare Cost Reports. The Providers are: Hospital, SNF, HHA, ESRD, Hospice, CMHC, FQHC and Clinic. The HA Database information is available for 12 years.

HA Database Documentation

This Database Overview document is the first part of the HA Database documentation. This document is part of the working documentation for HA Website, HA Desktop and the HA Database. Over time, this documentation may be out of date as updates are completed. Comments may be included to identify problems that require attention. In addition, future enhancements may be defined for future features. This documentation is available on the HA Website to anyone needing information about the HA Database of Healthcare Almanac.

HA Database Input Tables

The input files for the HA Database are from two Federal Government agencies. All the data is available to the public in a CSV format. The two government agencies are www.CMS.gov (Center for Medicare Medicaid Services) and the www.Census.gov (Census Bureau) website. The input tables are downloaded directly from the government websites. The primary input tables are the Cost Report files from the CMS website. These tables are updated each quarter for each provider and the data is available for over 20 years. The next input table is the POS (Provider of Service) table that is the basic information from all Medicare Providers that can submit invoices/bills to CMS. Physicians are not providers and are paid in another part of CMS.

HA Website and HA Desktop Tables

The CMS Cost Reports data is created from the Input Tables from CMS. The Detail Tables are created for the HA Desktop to show every detailed line-item record from the CMS Input Tables in a CMS Cost Report. The Summary Tables are created for the HA Desktop to provide Profile

Database Overview
June 2021 HA Database

Reports, Financial Reports, Departmental Reports and Healthcare Analytics of the CMS Cost Reports. The HA Website uses the Summary tables of the HA Database to provide quick access to provider information through the Profile Screens. The HA Website uses the Geography Tables to provide National, State, County and CBSA healthcare analytics. The HA Website uses the Summary Worksheet Tables to provide summary line item Cost Report information. The HA Website uses the KPI Calculations and Line Values to provider healthcare analytics of the Cost Report information.

Special HA Databases

Two special HA Databases are created to support the HA Database. The first database is a special User Database stored on your local computer to provide selection lists of providers and worktable storage. The location of the database is in the /Halmanac directory in the users Documents directory. The User Database is a SQL Compact Server version 4.1. The second database is the HA Demo Database created to provide a demonstration of the HA Desktop software. The HA Demo Database is currently the State of Florida with 250+ Hospitals and the other providers. The Demo Database in stored in Microsoft SQL Azure.

Conversion Software (Tab2)

Conversion software is used to convert the downloaded government website CSV files into usable file formats. The Conversion Software is processed 4 times a year. For the June and December quarters, Hospital, SNF, HHA, ESRD, Hospice, CMHC, FQHC and Clinic are processed. For the March and September quarters, Hospital is processed. The last five years of data is processed in each conversion. A provider Cost Report must be filed with the Fiscal Intermediary 5 months after the Fiscal year end of a provider. Provider Cost Reports can be amended for three years and sometime longer. Usually, several providers have processing delays due to provider issues or events like fires or hurricanes. In Healthcare Almanac, normally 12 years of data are available, but the last 5 years are processed in each quarter.

The steps for processing the Cost Report data.

1. The CMS Cost Report Headers, Numeric and Text tables are converted from CSV files into database work tables. The Header record is extended with control information for future processing and control.
2. The converted Numeric tables Line Numbers and Columns for Hospitals and SNF are updated with Cost Center conversions.
3. The Cost Reports require Worksheet Description Names, Worksheet Part Description Names and Worksheet Type Description Names. Special Tables are created from the Numeric and Text tables and the missing descriptions are inserted into the name tables.

Database Overview
June 2021 HA Database

4. The Cost Reports require Worksheet Line Description and Worksheet Column Descriptions. Special tables are created from the Numeric and Text tables to ensure all lines and columns are created. The Worksheet Lines are five digits and can have subscripted cost center codes in the lines. The lines are summarized into a 3 digit line number to get a complete total for each line. Then the line and column descriptions are inserted into the Line and Column Descriptions.
5. The final Headers, Numeric and Text tables are created by joining the Numeric and Text records to the selected header records for a calendar year.
6. The next step is to create clustered primary indexes to ensure the SQL tables can be uploaded to the Microsoft Azure SQL Server.
7. The last step is to create a Control Table of providers from the POS tables to identify the providers that have filed, not filed or exempt for filing Cost Reports. In addition, record counts of the numeric and text records are placed in each Control Table record.

This concludes the Cost Report conversion from downloaded CSV files into a local Microsoft SQL Server database for each year of Cost Reports. The last step is to upload the local Microsoft SQL Server databases to Microsoft Azure SQL Server. At this time, record counts of the numeric and text record counts in the SQL server are stored in the Cost Report Controls Table are balanced to the Microsoft Azure SQL Server tables.

HA Control Totals (Tab 3)

The HA Control Totals are a special step to balance key system control totals in the HA Database. Special fields in the HA Database store the HA Control Totals. This is a work in progress as control totals issues are identified. Most of the control total issues are data problems in the CMS data. New control total information is stored in the HA Database with each processing of the CMS data. There is control information in the Summary Tables that is key to balancing from the CSV header, numeric and text tables to Azure SQL Server tables. In the September 2021 processing, new fields will be added to the HA Database. The new fields will be column totals from Worksheet A, B and C. The Worksheet tables will have new fields for the Cost Report Ranking field and Worksheet Numbers for the 12 years of information. The goal is to balance the totals of every amount by worksheet in the HA Database. The controls will be retested and saved in the HA Database. This is a very important step to ensure the reliability of the HA Database. At present, the HA Control Totals focus on 11 key fields to balance control totals through all the processing steps. Another HA Control Total test is the record counts of the Numeric and Text tables by worksheet. Another HA Control Total test is the Column Totals for the Worksheets. As software upgrades occur and database information is expanded, the HA Control Totals are continually tested to ensure all processing steps are balanced and accurate.

Database Overview
June 2021 HA Database

HA Database Tables

System Tables (Tab 4)

The HA System Tables are a small group of tables in the HA Database and User Database that are necessary to supporting the HA Desktop and HA Website. The HA Database is a read only database with users limited to a data reader role. The HA Desktop uses the User Database stored on the local computer for three System Tables. The System tables are the System Log, Error Log and Reports Table. The HA Desktop has additional tables that are used for descriptions of selected fields. These tables are the States, Carriers and Counties tables used for descriptions of fields and improve the provider tagging features.

The Impact table is a CMS Excel Spreadsheet from the Final Rule each year for Hospitals that contains Area Wage Index (AWI), CBSA Codes related to Area Wage Index determination and Case Mix Index (CMI) used to determine a hospitals complexity of cases. The CMI is based on CMS Medicare Cases. The Final Rule is released each year and is the table used to calculate the PPS DRG rates for a provider.

The Census Bureau Population Estimate is used to get a population estimate by State and Counties. The tables are downloaded from the ftp2.census.gov website and are in the POPEST directory. These tables are used to insert information into the National, State, County and CBSA Geography Tables. The Population Estimate is a calculation by the Census Bureau based on July 31 population plus births minus deaths and migration adjustment to get an ending population at July 31 each year. The Population Estimate begins at the last Census and adds years each year.

Summary Tables (Tab 5)

The Summary Tables are the main source of information for the HA Website. Several Summary Tables are also used by the HA Desktop in the Summary Section. In the table names listed below, are table names with asterisks on key words. The Provider* name is the 5 Providers: Hospital, SNF, HHA, ESRD and Hospice. The Year* is the last 12 years (2020 to 2009) of Provider information. The Geographies* are the 4 Geographies; National, States, Counties and CBSAs. The last item is the BaseYear* which is the current year for the website. The Baseyear is updated when provider filings start receiving data for a new year. There are 7 Groups of Tables in the Summary tables.

- **REF_Summary_Provider*_Year*** - The REF_Summary tables are created for each provider and each year. There are five providers; Hospital, SNF, HHA, ESRD and Hospice. There is a table for each year and currently goes from 2020 back to 2009. Therefore, there are 60 tables for the five providers and 12 years. The REF_Summary table has 368 fields of information for a provider for each year. New fields are added every year and all Summary tables are updated each year. You can use SQL

Database Overview
June 2021 HA Database

Management Studio to see the provider fields. The file format is the same for each provider. Therefore, special fields for hospital, SNF, HHA are not normalized. The loss of space from unused fields is not worth the effort to have a file format for each provider. These tables are used for HA Desktop to create the Provider Profiles and Provider KPI reports.

- **REF_Geography*_Provider*_BaseYear*** - The REF_Geography tables are created each year for the base year of the website with 12 years of information. The REF_Geography tables are created for 4 geographies; National, States, Counties and CBSAs. The Geography tables are created for each of the five providers: Hospital, SNF, HHA, ESRD and Hospice. The tables contain 12 years of information. The National table contains one record for each year. The States table contains 58 state records for each year. The Counties tables contains over 3,100 county records and the CBSAs table contain over 950 records.

- **REF_KPI_Provider*_BaseYear*** and **REF_KPI_Values_Provider*_BaseYear*** - The REF_KPI Tables are created each year for the base year of the website for 12 years of information. The REF_KPI Tables are summaries of key information from the 60 REF_Summary_Provider Tables. The KPI_Values Table has the values used in calculations for the KPIs.

- **REF_CostReports_Provider*_BaseYear*** - The REF_CostReports table has all the Worksheet Numbers of all the Cost Reports filed by the Providers. The table has key information about the filing status of each cost report.

- **REF_Worksheet_Letter*_Provider*_BaseYear*** - The REF_Worksheet are three tables created each year for the base year of the website for 12 years of information. The Worksheet Letters are Worksheet A – Trial Balance of Expenses, Worksheet B – Cost Allocation and Worksheet C – Ratio of Cost to Charges. These tables provide detailed cost and changes information. These tables are used in the HA Desktop for the Provider Department reports.

- **REF_Financial_Provider*_BaseYear*** - The REF_Financial is created each year for the base year of the website for 12 years of information. The table is created from Worksheet G/F. The Financial information in the REF_Summary tables is updated from this table. The website Balance Sheet, Income Statement and Charges Statement are from this table. These tables are used in the HA Desktop for the Provider Financial reports.

- **REF_Compare_Provider*_BaseYear*** - The REF_Compare tables are used for selecting providers for the KPI_Compare reports on the HA Desktop. The table is created each year for the base year of the website and 6 years of selection information. The table is used in the HA Desktop and not used on the website.

Database Overview
June 2021 HA Database

Cost Report Tables (Tab 6)
The CMS Cost Report Tables are the primary source of detailed Cost Report information for the HA Website and HA Desktop. The CMS Cost Report Tables are created from CSV files downloaded from the www.CMS.gov website. The downloaded tables are by provider and year based on the CMS fiscal year of October 1 to September 30. In Healthcare Almanac, all Cost Report dates are converted to calendar years. The HA Database contains 8 Providers for up to 12 years. In addition, each provider and year has 12 to 14 tables. Therefore, the total Cost Report Tables are over 1,000 tables.

The original CMS input files are Headers, Numeric and Text information in a CSV file format. These tables are loaded with their original field names. The Header table is extended with additional information taken from the Numeric and Text input tables. The table names listed below have table names with asterisks in key words. The Provider* name is the 8 Providers: Hospital, SNF, HHA, ESRD, Hospice, CMHC, FQHC and Clinic. Over the years, CMS has updated the CMS reporting forms for most of the providers. When this occurs, a V2 is added to the provider name. The providers for updated Cost Reports are; Hospital_V2 in 2011, SNF_V2 in 2011, ESRD_V2 in 2011, Hospice_V2 in 2015, Clinic_V2 in 2017 and CMHC_V2 in 2017. The providers for CMHC and Clinic are in the process of being updated in the June 2021 processing. The Year* is the last 12 years (2020 to 2009) of Provider information. There are 4 Groups of Tables; CMS Input Tables, Worksheet Tables, Line/Column Tables and a Control Table for the Cost Report tables.

CMS Input Tables
- **CR_Provider*_Header_Year*** - The CR_Provider_Header tables are created for each provider and each year. There are eight providers; Hospital, SNF, HHA, ESRD, Hospice and CMHC, FQHC and Clinic. There is a table for each year and currently goes from 2020 back to 2009. This table is extended with extra information taken from the Numeric and Text tables. Also, control record counts are added to the table to count the records from the Numeric and Text tables.
- **CR_Provider*_Text_Year*** - The CR_Provider_Text tables are created for each provider and each year. There are eight providers; Hospital, SNF, HHA, ESRD, Hospice and Clinic, CMHC and FQHC. There is a table for each year and currently goes from 2020 back to 2009. This table is extended with three fields for the Worksheet Line number, Worksheet Column number and a Flag is set when the line number is changed from the Worksheet Line number based on the Cost Center Table.
- **CR_Provider*_Numeric_Year*** - The CR_Provider_Numeric tables are created for each provider and each year. There are eight providers; Hospital, SNF, HHA, ESRD, Hospice and Clinic, CMHC and FQHC. There is a table for each year and currently goes

Database Overview
June 2021 HA Database

from 2020 back to 2009. This table is extended with three fields for the Worksheet Line number, Worksheet Column number and a Flag is set when the line number is changed from the Worksheet Line number based on the Cost Center Table.

Worksheet Tables
- **CR_Provider*_Worksheets_Year*** - The CR_Provider_Worksheets tables are created for each provider and each year. These records are the titles of the Worksheets.
- **CR_Provider*_Workparts_Year*** - The CR_Provider_Workparts tables are created for each provider and each year. These records are the titles of the Workparts.
- **CR_Provider*_Worktypes_Year*** - The CR_Provider_Worktypes tables are created for each provider and each year. These records are the titles of the Worktypes.

Line/Column Description Tables
- **CR_Provider*_Lines_Year*** - The CR_Provider_Lines tables are created for each provider and each year. These records are the Line Descriptions for the lines on the Cost Report worksheets. These Lines are the actual 5 digit line number from the CMS Cost Reports. The CMS Cost Reports use line number subscripting to provide additional information for submitting a Cost Report. This table also contain information for formatting and describing the numeric values in the lines.
- **CR_Provider*_CC_Lines_Year*** - The CR_Provider_CC_Lines tables are created for Hospital and SNF each year. These records are the conversion of line numbers to worksheet line numbers or Cost Center line numbers. The actual worksheets lines are converted to cost center line numbers.
- **CR_Provider*_WS_Lines_Year*** - The CR_Provider_WS_Lines are created by using the first three digits of a line number or the CC Lines table is used to convert the line number to a Worksheet line or Cost Center line. These Line Numbers also appear on the CMS Cost Reports worksheet form. This table also contain information for formatting and describing the numeric values in the lines.
- **CR_Provider*_Columns_Year*** - The CR_Provider_Columns tables are created for each provider and each year. These records are the Column Descriptions for the Worksheet Columns on the worksheets. This table also contain information for formatting and describing the numeric values in the columns.
- **CR_Provider*_CC_Columns_Year*** - The CR_Provider_CC_Columns tables are created for Hospital and SNF each year. These records are the conversion of column numbers to worksheet column numbers for the Cost Center Column numbers. The CMS documentation implies the column numbers can be changed based on the Cost Center, but all the column numbers are never changed.

Database Overview
June 2021 HA Database

- **CR_Provider*_Overrides_Year*** - The CR_Provider_Overrides tables are created for each provider and each year. These records override the default Line Descriptions and default Column Descriptions for the lines and columns on the CMS Cost Report worksheets. This table can also override the line and column format of a numeric value. The Overrides customize the Cost Report Reference report to the actual CMS Cost Report Forms.
- **CR_Provider*_Titles_Year*** - The CR_Provider_Titles tables are created for each provider and each year. This table is not currently used but was intended to insert descriptions into the Cost Reports to show additional information contained on the CMS Cost Report worksheets.

Controls Table
- **CR_Provider*_Controls_Year*** - The CR_Provider_Controls table is created for each provider and each year. These controls are used with the POS Tables to identify all the providers that should file a Cost Report. The filing status of all Cost Reports are listed in this table. The table also contains fields for record counts of the numeric and text tables. These record counts are used to count all header, numeric and text records uploaded from CMS to a local Microsoft SQL Server to Azure SQL Database.

Provider of Service Tables (Tab 7)

The Provider of Service Tables (POS) are a source of information for the HA Website and HA Desktop. The POS tables provide limited information about Medicare Providers. The POS tables provide the POS numbers, names, address, etc. for all Medicare Providers. The new NPI tables were created by CMS over 10 years ago to contain the POS information. The NPI tables also contain Physician information. The NPI numbers are not in the HA Product.

The POS tables are created each year from two files (Providers and Labs) from the CMS website. The files are in CSV format and contain about 20+ types of providers. The tables have varying amounts of information for each provider. The table design of the POS tables were created in a time when disk space was expensive. The table design has fields that apply to one or more providers. The Hospital provider has about 100 fields and SNF has about 180 fields. About 50 of the fields are shared to save disk space. In processing, separate tables are created for each type of provider.

The POS Tables have two additional tables for Line Descriptions of each field in the POS Tables and a Decode table to provide a description for a code in a field. For example, Provider Type has a 2 digit code but the decode table provides the field description (Code 01 is Short Term Hospital).

Database Overview
June 2021 HA Database

The POS Tables also have special tables for Control Type Counts, Category Counts, State Counts and National Counts. A special table named POS_ALL_Year contains the first 36 fields that are common to all POS records. These tables are used to verify accurate processing.

- **POS_Provider*_Year*** - The POS_Provider tables are created for each provider and each year. There are 10 providers; Hospital, SNF, HHA, ESRD, Hospice, Mental, Rural, Therapy, Rehab, Nursing, Labs and Other. There is a table for each year and currently goes from 2020 back to 2008. The POS_Provider table has over 500 fields that are shared between the providers.
- **POS_LineDescription_Year*** - The POS_LineDescription tables are created each year for the annual POS files. The POS Line Description contains a field to identify which providers have which fields. The POS Record Layout from CMS provides a listing of the information. Each year, fields are added or dropped depending on legal requirements for the information.
- **POS_Decodes_Year*** - The POS_Decodes tables are created each year for the annual POS files. The POS Decodes contains the information to define the codes in fields. In the POS reporting, the decoded information is printed instead of the code.
- **POS_All_Year*** - The POS_All table is created each year for the annual POS files. The POS_All contains the first 36 fields of every provider record. The first 36 fields contain name, address, county, state, etc. This table is used for special processing.
- **POS_Control_Count_Year*** - The POS_Control_Count tables are created each year for the annual POS files. The Control Count contains information about the business organization (Company, Proprietorship, Partnership, County, State, Federal) by Type of Provider and business organization. This table is used to document processing results and compare year to year counts.
- **POS_Category_Count_Year*** - The POS_Category_Count tables are created each year for the annual POS files. The POS Category Count contains the information about the purpose of a provider like Hospitals are Short Term, Long Term, etc. This table is used to document processing results and compare year to year counts.
- **POS_National_Count_Year*** - The POS_National Count tables are created each year for the annual POS files. The POS National Count contains counts by type of provider. This table is used to document processing results and compare year to year counts.
- **POS_State_Count_Year*** - The POS_State Count tables are created each year for the annual POS files. The POS State Count contains counts by state and by type of provider. This table is used to document processing results and compare year to year counts.

Database Overview
June 2021 HA Database

HA Special Database Tables

User Database (Tab 8)

In the HA Desktop product, the User Database is created on the local computer in a SQL Server Compact Database product from Microsoft. The User Database is stored in the local computer Documents directory and is named Halmanac.sdf. There is also another user database named Halmanac_Demo.sdf. This User Database provides selection lists for the HA Desktop reports. When reports are created, the work tables are saved in the User Database. The User Database is created automatically by the HA Desktop software. If you save a report as a PDF or and XLSX file, the files are stored in this User Database.

Demo Database (Tab 9)

The HA Demo Database is created each year to provide a free access to information for the HA Desktop. The HA Desktop is the software part of the total solution for CMS Medicare Cost Reports. The HA Demo Database is created for the State of Florida with 260+ hospitals along with the other providers – SNF, HHA, ESRD and Hospice. This HA Demo Database provides enough information for a person to see the features of the HA Desktop. By using the HA Demo Database, you can see the reporting and selection features available in the HA Desktop.

Conclusion

The HA Database is the information part of the Healthcare Almanac total solution for CMS Medicare Cost Reports. This documentation is an index to the complete description of the features of the HA Database. The tabs listed in this overview point to additional documentation for parts of the HA Database. The document is prepared for the HA Website to provide documentation of the HA Database.

Robert Neal, CPA

Prepared 09-17-2021