Analysis of Medicare Plans
Project for Hadoop IBM Certification

Author: Robert Chen

Under the Guidance of: Prashanth Shantigrama

Document History

Revision History

Revision Number	Revision Date	Summary of Changes	Revised By
1.0	06/30/2015	Document Creation	Prashanth Shantigrama
1.1	11/19/2015	Filled in work done for project for IBM Certification in Hadoop.	Robert Chen

Table of Contents

1. Introduction

1.1 Problem Statement

Multiple Medicare plans are available for senior citizens and other qualified members to enroll into every year, offered by different health insurance companies. While a lot of information regarding individual plans exist, it is difficult to compare plans based on various criteria to make an informed choice to suit unique situations of individual members as well as for plan benefit designers to compare plans and design benefits that meet unique requirements and are competitive in different markets. The purpose of this document is to detail the analysis of Medicare plan data across the US and provide useful summary details.

1.2 Motivation

- The primary purpose of this project is to facilitate analysis of Medicare plans to provide meaningful insights that help in choosing appropriate Medicare plans by comparing all relevant details regarding plans that are available in each county throughout the country.
- While CMS provides rich details on all the plans that are offered county wise, it makes for better analysis when each plan that is offered is compared on the finer descriptions of its cost and coverage details.

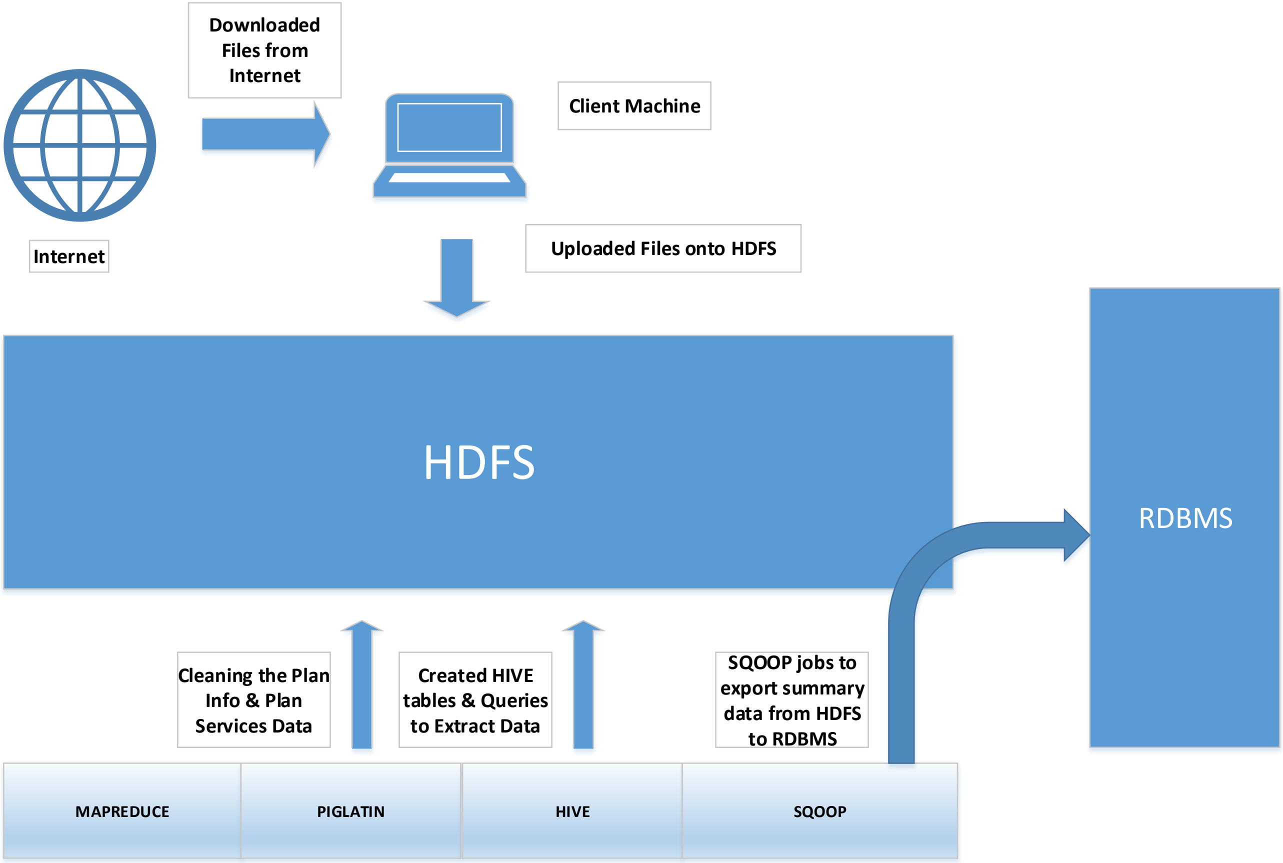
1.3 Objectives

- Implement an efficient system to extract, load and transform all data related to Medicare plans to perform analytics.
- Compare plan offerings by various criteria to select a suitable plan for the Members.
- Compare plan offerings to design a suitable benefit plan for different regions.

1.4 Analysis Required for this project

1. Identify the top 5 plans with the lowest premiums for a given county across the US.
2. Find plans that have the highest co-pays for doctor’s visits in a given county.
3. Compare plans based on features such as free ambulance services.
4. Compare plans based on features such as the benefits available for diabetes under a specific plan.
5. Compare plan benefits on diabetes and mental healthcare offered by all companies in a particular county.

2. Solution Architecture


Downloaded Files from Internet
Client Machine
Internet
Uploaded Files onto HDFS
HDFS
RDBMS
Cleaning the Plan Info & Plan Services Data
Created HIVE tables & Queries to Extract Data
SQOOP jobs to export summary data from HDFS to RDBMS
MAPREDUCE
PIGLATIN
HIVE
SQOOP

2.1 Architecture Components

The proposed solution uses the Hadoop framework and its ecosystem tools to implement a distributed storage system and to perform the analysis.

Client Machine: The Medicare files are downloaded from the Medicare website onto the client machine.

HDFS: The downloaded files are uploaded onto the Hadoop Distributed File System. The files are uploaded onto a Cloudera VMWare setup which runs in Pseudonode mode.

PIG: The raw Medicare files on the HDFS are cleansed and transformed on HDFS and spooled into CSV files on HDFS using Pig Latin Scripts.

HIVE: The Partioned HIVE tables are built over the cleansed Medicare plan info and plan services CSV files that are stored in HDFS which were generated by the Pig Latin scripts.

SQOOP: Squoop can be used to export out the analyzed summary data stored in HIVE tables to a traditional RDBMS.

2.2 System Requirements

Software: Pig Latin, HIVE, Sqoop and WINSCP

Hardware: Cloudera VMWare

Cluster: Pseudo mode Cluster

2.3 Data Source

The Medicare data can be downloaded from the Medicare government website from the following location:

https://www.medicare.gov/download/downloaddb.asp

Specifically, the following two CSV files consists of the plan information and plan services data from where the analysis of the data can be performed:

1. **PlanInfoCounty_FipsCodeMoreThan30000**

   This file contains the contractid, planid, segmented, planname and county ID where the county ID is more than 30000. Sample data follows:

   "H0022","001","0","2015","Buckeye Health Plan - MyCareOhio","Buckeye Health Plan - MyCareOhio (Medicare-Medicaid Plan)","","Cleveland Dayton Toledo and surrounding counties","1","For Profit","Con Fines de Lucro","48","Medicare-Medicaid Plan","http://mmp.bchpohio.com","","www.bchpohio.com","www.bchpohio.com","Approved by Medicare and Medicaid","Aprobadopor Medicare y Medicaid","FALSE","TRUE","FALSE","88","&nbsp;","","","<ul><li>This is a Medicare-Medicaid plan for people with both Medicare and Medicaid. Contact the plan for details.</li></ul>","","","","BUCKEYE COMMUNITY HEALTH PLAN INC.","","","","","4349 Easton Way Suite 200","Columbus","OH","43219","nslee@centene.com","1-866-549-8289","1-866-549-8289","711","711","nslee@centene.com","1-866-549-8289","1-866-549-8289","711","711","","4349 Easton Way Suite 200","Columbus","OH","43219","nslee@centene.com","1-866-549-8289","1-866-549-8289","711","711","nslee@centene.com","1-866-549-8289","1-866-549-8289","711","711","FALSE","FALSE","0","Not SNP","No hay planes para necesidades especiales","$0.00","$0.00","$0.00","$0.00","FALSE","39023"

   "H0022","001","0","2015","Buckeye Health Plan - MyCareOhio","Buckeye Health Plan - MyCareOhio (Medicare-Medicaid Plan)","","Cleveland Dayton Toledo and surrounding counties","1","For Profit","Con

Fines de Lucro","48","Medicare-Medicaid Plan","http://mmp.bchpohio.com","","www.bchpohio.com","www.bchpohio.com","Approved by Medicare and Medicaid","Aprobadopor Medicare y Medicaid","FALSE","TRUE","FALSE","88","&nbsp;","","","<ul><li>This is a Medicare-Medicaid plan for people with both Medicare and Medicaid. Contact the plan for details.</li></ul>","","","","BUCKEYE COMMUNITY HEALTH PLAN INC.","","","","","4349 Easton Way Suite 200","Columbus","OH","43219","nslee@centene.com","1-866-549-8289","1-866-549-8289","711","711","nslee@centene.com","1-866-549-8289","1-866-549-8289","711","711","","4349 Easton Way Suite 200","Columbus","OH","43219","nslee@centene.com","1-866-549-8289","1-866-549-8289","711","711","nslee@centene.com","1-866-549-8289","1-866-549-8289","711","711","FALSE","FALSE","0","Not SNP","No hay planes para necesidades especiales","$0.00","$0.00","$0.00","$0.00","FALSE","39035"

2. **PlanInfoCounty_FipsCodeLessThan30000**

This file contains the contractid, planid, segmented, planname and county ID where the county ID is less than 30000. Sample data follows:

"H0028","001","0","2015","CHA HMO Inc.","Humana Gold Plus H0028-001 (HMO)","","Cedar Rapids Metro Area","1","For Profit","Con Fines de Lucro","1","HMO","www.humana-medicare.com","www.humana.com","https://www.humana.com/pharmacy/medicare/tools/druglist/","https://www.humana.com/pharmacy/medicare/","Approved by Medicare","Aprobadopor Medicare","FALSE","TRUE","FALSE","88","&nbsp;","","","<ul><li>This plan does not charge an annual deductible for all drugs. The $320 annual deductible only applies to drugs on certain tiers.</li></ul>","","","","CHA HMO INC.","","1501-2000 physicians and providers.","1501-2000 médicos y proveedores.","","500 West Main Street","Louisville","KY","40202","HumanaOCS@humana.com","1-800-833-2364","1-800-833-2364","711","711","","1-800-457-4708","1-800-457-4708","711","711","","500 West Main Street","Louisville","KY","40202","HumanaOCS@humana.com","1-800-833-2364","1-800-833-2364","711","711","","1-800-457-4708","1-800-457-4708","711","711","FALSE","FALSE","0","Not SNP","No hay planes para necesidades especiales","$0.00","$0.00","$0.00","$0.00","FALSE","19011"

"H0028","001","0","2015","CHA HMO Inc.","Humana Gold Plus H0028-001 (HMO)","","Cedar Rapids Metro Area","1","For Profit","Con Fines de Lucro","1","HMO","www.humana-medicare.com","www.humana.com","https://www.humana.com/pharmacy/medicare/tools/druglist/","https://www.humana.com/pharmacy/medicare/","Approved by Medicare","Aprobadopor Medicare","FALSE","TRUE","FALSE","88","&nbsp;","","","<ul><li>This plan does not charge an annual deductible for all drugs. The $320 annual deductible only applies to drugs on certain tiers.</li></ul>","","","","CHA HMO INC.","","1501-2000 physicians and providers.","1501-2000 médicos y proveedores.","","500 West Main Street","Louisville","KY","40202","HumanaOCS@humana.com","1-800-833-2364","1-800-833-2364","711","711","","1-800-457-4708","1-800-457-4708","711","711","","500 West Main Street","Louisville","KY","40202","HumanaOCS@humana.com","1-800-833-2364","1-800-833-2364","711","711","","1-800-457-4708","1-800-457-4708","711","711","FALSE","FALSE","0","Not SNP","No hay planes para necesidades especiales","$0.00","$0.00","$0.00","$0.00","FALSE","19019"

3. **vwPlanServices**

This file contains the contractid, category description, category code and benefit information. Sample data follows:

English,2015,H0001,001,0,Monthly Premium Deductible and Limits on How Much You Pay for Covered Services,1,In 2015 the monthly Part B Standard Premium is <b>$104.90</b>,Base Plan,000,1

Spanish,2015,H0001,001,0,Prima Mensual Deducible y Límites a la Cantidad que UstedPagaporlosServicios Cubiertos,1,La prima estándar que pagapor la Parte B en el 2015 es<b>$104.90</b>,Plan Básico,000,1

3. Project Work Flow

Source Files

The source files were downloaded from the Medicare government website to the client machine and decompressed to CSV files.

HDFS

HDFS is run using a Cloudera VMWare image (v5.4.2) running in pseudo Mode. The Medicare files were uploaded onto HDFS using the Hadoop FS commands.

Clean up files using PIG Latin Scripts

- The data in the plan info and plan services files had double quotes enclosed around each field. The files were cleaned using Pig scripts (listed in the next section).
- The data in the plan services file had 2 rows for each record -- one in English and the other in Spanish. The file was cleaned using a Pig script.
- Plan info & Plan services files were filtered with records where contractId, SegmentId, PlanId were null using pig scripts. The details will be listed in the next section.

HIVE tables to analyze tables:

Hive tables were built with the cleaned output files from the Pig latin scripts.

Export Data using Sqoop:

The summary files data from HIVE tables can be exported out of HDFS into traditional databases using Sqoop. However, this step is not in the scope of the project but can be included in future use cases.

4. Analysis Performed

The following is a summary of the analysis performed:

- Grouping the plans based on companies managing the plans and the counties where offered.
- Finding plans that offer specific services like free ambulance service in its coverage descriptions.
- Comparing plans based on the benefits offered for specific conditions like Diabetes.
- Comparing plans based on premiums and co-pays for specific coverage criteria like doctors co-pays.
- Comparing plans based on premiums for specific coverage criteria like doctors co-pays.

5. Code

5.1 Pig Latin Scripts

```
/* These are Pig commands used to preprocess the Medicare database.
   Two preprocessing operations are done.

   First, the 2 county CSV files are combined into one file.  The
   10 fields needed for analysis are extracted.  Also, all entries
   with blank values for contract_id, planid, or segmentid are filtered
   out.

   Second, the services CSV file is filtered so that only English entries
   are stored.  The 5 fields needed for analysis are extracted and any
   entries with blank values for contract_id, planid, or segmentid are not
   used.

   These commands assume that the 2 county files and services file are
   copied into HDFS in /medicare and renamed to "County1.csv", "County2.csv",
   and "vwPlanServices.csv".

   The output of these commands are 2 CSV files in HDFS in /medicare:
   "County.csv" and "Services.csv".
*/


/* Needed for CSVLoader */
REGISTER /usr/lib/pig/piggybank.jar;

/* Access special CSVLoader that accounts for "," inside of quotes */
DEFINE CSVLoader org.apache.pig.piggybank.storage.CSVLoader();

/* Load County data part 1 */
A = LOAD '/medicare/County1.csv' USING CSVLoader AS (contractid, planid,
      segmentid, year, org, plan, sp_plan, geo, ts, tsd, sp_tsd, type,
      type_desc, web, partd_web, form_web, pharm_web, fas, sp_fas,
      pos, moa, cgo, cgi, cgd, contract_note, sp_contract_note,
      plan_note, sp_plan_note, seg_note, sp_seg_note, legal, trade,
      network, sp_network, contact, address, city, state, zip,
      email_pro, lphone_pro, tphone_pro, ltty_pro, tftty_pro,
      email_cur, lphone_cur, tphone_cur, ltty_cur, tftty_cur,
      contact_pd, address_pd, city_pd, state_pd, zip_pd,
      email_pro_pd, lphone_pro_pd, tphone_pro_pd, ltty_pro_pd, tftty_pro_pd,
      email_cur_pd, lphone_cur_pd, tphone_cur_pd, ltty_cur_pd, tftty_cur_pd,
      mapd, ppopd, snpid, snpdesc, sp_snpdesc, lis100, lis75, lis50, lis25,
      region, county);

/* Filter out header and records with null ID values */
B = FILTER A BY contractid != 'contract_id' AND contractid !='' AND
```

```
      planid != '' AND segmentid != '';

/* Only save the 10 fields needed for future queries */
C = FOREACH B GENERATE contractid, planid, segmentid, org, plan,
      address, city, state, zip, county;

/* Load County data part 2 */
A1 = LOAD '/medicare/County2.csv' USING CSVLoader AS (contractid, planid,
       segmentid, year, org, plan, sp_plan, geo, ts, tsd, sp_tsd, type,
      type_desc, web, partd_web, form_web, pharm_web, fas, sp_fas,
      pos, moa, cgo, cgi, cgd, contract_note, sp_contract_note,
      plan_note, sp_plan_note, seg_note, sp_seg_note, legal, trade,
      network, sp_network, contact, address, city, state, zip,
      email_pro, lphone_pro, tphone_pro, ltty_pro, tftty_pro,
      email_cur, lphone_cur, tphone_cur, ltty_cur, tftty_cur,
      contact_pd, address_pd, city_pd, state_pd, zip_pd,
      email_pro_pd, lphone_pro_pd, tphone_pro_pd, ltty_pro_pd, tftty_pro_pd,
      email_cur_pd, lphone_cur_pd, tphone_cur_pd, ltty_cur_pd, tftty_cur_pd,
      mapd, ppopd, snpid, snpdesc, sp_snpdesc, lis100, lis75, lis50, lis25,
      region, county);

/* Filter out header and records with null ID values */
B1 = FILTER A1 BY contractid != 'contract_id' AND contractid !='' AND
       planid != '' AND segmentid != '';

/* Only save the 10 fields needed for future queries */
C1 = FOREACH B1 GENERATE contractid, planid, segmentid, org, plan,
      address, city, state, zip, county;

/* Combine the 2 parts into 1 table */
U  = UNION C, C1;

/* Store the combined table */
store U into '/medicare/County.csv' USING PigStorage(',');

/* Load the Services table */
D  = LOAD '/medicare/vwPlanServices.csv' USING CSVLoader AS (language, year,
       contractid, planid, segmentid, category, code, benefit, package_name,
       package_id, sso);

/* Only take the English records and those with non-null entries for ID */
E  = FILTER D BY language == 'English' AND contractid !='' AND planid != ''
      AND segmentid != '';

/* Save 5 fields for future use */
F  = FOREACH E GENERATE contractid, planid, segmentid, category, benefit;

/* Perform the store */
store F into '/medicare/Services.csv' USING PigStorage(',');
```

5.2 HIVE Scripts

```
-- These are a set of Hive commands to do analysis of the Medicare
-- database.  It assumes 2 tables, "County" and "Services", are
-- created by Pig.   The following analysis is done:
--
--    (1) Show plans grouped by organization.
--    (2) Show plans grouped by county.
--    (3) Show plans with free ambulance service.
--    (4) Show plans with diabetes options.
--    (5) Show plans with mental health care options.
--    (6) Show plans with most expensive copays for a specific county.
--    (7) Show plans with least expensive premiums for a specific county.


-- /****************************************/
-- /* Load County table from Pig into Hive */
-- /****************************************/

CREATE TABLE countyTable (
  contractid  STRING,
  planid      STRING,
  segmentid   STRING,
  org         STRING,
  plan        STRING,
  address     STRING,
  city        STRING,
  state       STRING,
  zip         STRING,
  county      STRING)
  COMMENT 'This is medicare county data'
  ROW FORMAT DELIMITED
    FIELDS TERMINATED BY ','
    LINES TERMINATED BY '\n'
  STORED AS TEXTFILE
  LOCATION '/medicare/County.csv';


-- /*************************/
-- /* Query by organization */
-- /*************************/

INSERT OVERWRITE LOCAL DIRECTORY '/home/cloudera/ibm/hive/querybyorg'
  ROW FORMAT DELIMITED
  FIELDS TERMINATED BY ','
  SELECT org, plan FROM countyTable GROUP BY org,plan;


-- /*******************/
-- /* Query by county */
-- /*******************/
```

```
INSERT OVERWRITE LOCAL DIRECTORY '/home/cloudera/ibm/hive/querybycounty'
  ROW FORMAT DELIMITED
  FIELDS TERMINATED BY ','
  SELECT county, plan, org FROM countyTable
    GROUP BY county, plan, org;


-- /************************************************************************/
-- /* Load Services table from Pig into Hive -- Needed for next 2 queries. */
-- /************************************************************************/

CREATE TABLE servicesTable (
  contractid  STRING,
  planid      STRING,
  segmentid   STRING,
  category    STRING,
  benefit     STRING)
  COMMENT 'This is medicare services data'
  ROW FORMAT DELIMITED
    FIELDS TERMINATED BY ','
    LINES TERMINATED BY '\n'
  STORED AS TEXTFILE
  LOCATION '/medicare/Services.csv';


-- /*************************************************/
-- /* Query for plans with free ambulance services. */
-- /*************************************************/

-- Group by is used to get rid of duplicate entries.
INSERT OVERWRITE LOCAL DIRECTORY '/home/cloudera/ibm/hive/queryfreeambulance'
  ROW FORMAT DELIMITED
  FIELDS TERMINATED BY ','
  SELECT plan, org, benefit FROM countyTable JOIN servicesTable
    ON countyTable.contractID = servicesTable.contractID
    WHERE benefit = 'If you are admitted to the hospital  you do not have to pay for
the ambulance services.'
    GROUP BY plan, org, benefit;


-- /******************************************/
-- /* Query for plans with diabetes options. */
-- /******************************************/

-- Group by is used to get rid of duplicate entries.
INSERT OVERWRITE LOCAL DIRECTORY '/home/cloudera/ibm/hive/querydiabetes'
  ROW FORMAT DELIMITED
  FIELDS TERMINATED BY ','
  SELECT plan, org, benefit FROM countyTable JOIN servicesTable
    ON countyTable.contractID = servicesTable.contractID
    WHERE category = 'Diabetes Supplies and Services'
    GROUP BY plan, org, benefit;


-- /****************************************************/
-- /* Query for plans with mental health care options. */
-- /****************************************************/

-- Group by is used to get rid of duplicate entries.
INSERT OVERWRITE LOCAL DIRECTORY '/home/cloudera/ibm/hive/querymentalhealth'
  ROW FORMAT DELIMITED
  FIELDS TERMINATED BY ','
  SELECT plan, org, benefit FROM countyTable JOIN servicesTable
    ON countyTable.contractID = servicesTable.contractID
    WHERE category = 'Mental Health Care'
    GROUP BY plan, org, benefit;


-- /********************/
-- /* Query for copays */
-- /********************/

-- First, set up a table with copay information
CREATE TABLE copays (
  contractid  STRING,
  org         STRING,
  plan        STRING,
  county      STRING,
  benefit     STRING,
  copay       INT)
  COMMENT 'This is a table of copay data created to do sorting by copay'
  ROW FORMAT DELIMITED
    FIELDS TERMINATED BY ','
    LINES TERMINATED BY '\n'
  STORED AS TEXTFILE
  LOCATION '/medicare/Copays.csv';

-- Second, extract copay information for Doctor's visits and populate the copays
-- table
INSERT OVERWRITE TABLE copays
  SELECT countyTable.contractid, org, plan, county, benefit,
    regexp_extract(benefit, '([A-Za-z :<>&-]*.&nbsp.<b>.)([0-9]*)', 2) copay
  FROM countyTable JOIN servicesTable
  ON countyTable.contractid = servicesTable.contractid
  WHERE servicesTable.category = "Doctor's Office Visits";

-- Finally, do a query grouping by copay and sorting from cheapest to most
-- expensive copay for a specific county.
INSERT OVERWRITE LOCAL DIRECTORY '/home/cloudera/ibm/hive/querycopay'
  ROW FORMAT DELIMITED
  FIELDS TERMINATED BY ','
  SELECT county, plan, org, benefit, copay
```

```
    FROM copays
    WHERE county = "39113"
    GROUP BY county, plan, org, benefit, copay
    ORDER BY copay DESC;


-- /**********************/
-- /* Query for premiums */
-- /**********************/

-- First, set up a table with premium information
CREATE TABLE premiums (
  contractid  STRING,
  org         STRING,
  plan        STRING,
  county      STRING,
  benefit     STRING,
  premium     INT)
  COMMENT 'This is a table of premium data created to do sorting by premium'
  ROW FORMAT DELIMITED
    FIELDS TERMINATED BY ','
    LINES TERMINATED BY '\n'
  STORED AS TEXTFILE
  LOCATION '/medicare/Premiums.csv';

-- Second, extract premium information for Doctor's visits and populate the
-- premiums table.  The WHERE clauses are used to narrow down which entries
-- in the servicesTable have monthly premium information.
INSERT OVERWRITE TABLE premiums
  SELECT countyTable.contractid, org, plan, county, benefit,
    regexp_extract(benefit,'....([0-9]*.[0-9]*).... per month', 1) premium
  FROM countyTable JOIN servicesTable
  ON countyTable.contractid = servicesTable.contractid
  WHERE category = "Monthly Premium  Deductible  and Limits on How Much You Pay for
Covered Services" AND instr(benefit, "per month") <> 0 AND substr(benefit, 1, 4) =
"<b>$";

-- Finally, do a query grouping by premium and sorting from cheapest to most
-- expensive copay for a specific county.
INSERT OVERWRITE LOCAL DIRECTORY '/home/cloudera/ibm/hive/querypremium'
  ROW FORMAT DELIMITED
  FIELDS TERMINATED BY ','
  SELECT county, plan, org, benefit, premium
    FROM premiums
    WHERE county = "39113"
    GROUP BY county, plan, org, benefit, premium
    ORDER BY premium;
```

6. Sample Results

(1) Identify the plans with the lowest premiums for a given county across the US.

```
39113,PrimeTime Health Plan Employer Group HMO MA-Only (HMO-POS),PrimeTime Health
Plan,<b>$0.00</b> per month.  In addition  you must keep paying your Medicare Part B
premium.,0

39113,Today's Options PPO - MAPD Group Plan (PPO),Universal American
Corp.,<b>$0.00</b> per month.  In addition  you must keep paying your Medicare Part
B premium.,0

39113,PrimeTime Health Plan Employer Group - Non CY (HMO-POS),PrimeTime Health
Plan,<b>$0.00</b> per month.  In addition  you must keep paying your Medicare Part B
premium.,0

39113,Touchstone Health Medicare Power (HMO),Touchstone Health,<b>$0.00</b> per
month.  In addition  you must keep paying your Medicare Part B premium.,0

39113,Cigna-HealthSpring Premier Rx FY (HMO-POS),Cigna-HealthSpring,<b>$0</b> per
month.  In addition  you must keep paying your Medicare Part B premium.,0
```

(2) Find plans that have the highest co-pays for doctor’s visits in a given county.

```
39113,Empire MediBlue Freedom (PPO),Empire BlueCross BlueShield,<ul><li>Out-of-
network:&nbsp;&nbsp;<b>$75</b> copay</li></ul>,75

39113,Blue Cross Medicare Advantage Employer Expanded (PPO),Blue Cross and Blue
Shield of Montana,<ul><li>Out-of-network:&nbsp;&nbsp;<b>$65</b> copay</li></ul>,65

39113,Anthem Medicare Preferred (PPO),Anthem Blue Cross and Blue Shield,<ul><li>Out-
of-network:&nbsp;&nbsp;<b>$65</b> copay</li></ul>,65

39113,Blue Cross Medicare Advantage Employer Ideal (PPO),Blue Cross and Blue Shield
of Montana,<ul><li>Out-of-network:&nbsp;&nbsp;<b>$65</b> copay</li></ul>,65

39113,Anthem MediBlue Preferred (PPO),Anthem Blue Cross and Blue Shield,<ul><li>Out-
of-network:&nbsp;&nbsp;<b>$65</b> copay</li></ul>,65
```

(3) Find plans that include free ambulance services.

```
ADVANTRA SILVER (PPO),HealthAmerica,If you are admitted to the hospital  you do not
have to pay for the ambulance services.

ATRIO Bronze (PPO),ATRIO Health Plans,If you are admitted to the hospital  you do
not have to pay for the ambulance services.
```

```
ATRIO Bronze (Rogue) (PPO),ATRIO Health Plans,If you are admitted to the hospital you do not have to pay for the ambulance services.

ATRIO Bronze (Willamette) (PPO),ATRIO Health Plans,If you are admitted to the hospital  you do not have to pay for the ambulance services.

ATRIO Bronze Rx (Basin) (PPO),ATRIO Health Plans,If you are admitted to the hospital you do not have to pay for the ambulance services.
```

(4) Find plans that have benefits available for diabetes.

```
'Ohana Choice (HMO-POS),'Ohana Health Plan,<ul><li>Out-of-network:&nbsp;&nbsp;<b>20%</b> of the cost</li></ul>

'Ohana Choice (HMO-POS),'Ohana Health Plan,Diabetes monitoring supplies:<ul><li>In-network:&nbsp;&nbsp;<b>20%</b> of the cost</li></ul>

'Ohana Choice (HMO-POS),'Ohana Health Plan,Diabetes monitoring supplies:<ul><li>In-network:&nbsp;&nbsp;<b>You pay nothing</b></li></ul>

'Ohana Choice (HMO-POS),'Ohana Health Plan,Diabetes self-management training:<ul><li>In-network:&nbsp;&nbsp;<b>You pay nothing</b></li></ul>

'Ohana Choice (HMO-POS),'Ohana Health Plan,Therapeutic shoes or inserts:<ul><li>In-network:&nbsp;&nbsp;<b>20%</b> of the cost</li></ul>

'Ohana Choice (HMO-POS),'Ohana Health Plan,Therapeutic shoes or inserts:<ul><li>In-network:&nbsp;&nbsp;<b>You pay nothing</b></li></ul>

'Ohana Liberty (HMO-POS SNP),'Ohana Health Plan,<ul><li>Out-of-network:&nbsp;&nbsp;<b>20%</b> of the cost</li></ul>

'Ohana Liberty (HMO-POS SNP),'Ohana Health Plan,Diabetes monitoring supplies:<ul><li>In-network:&nbsp;&nbsp;<b>20%</b> of the cost</li></ul>

'Ohana Liberty (HMO-POS SNP),'Ohana Health Plan,Diabetes monitoring supplies:<ul><li>In-network:&nbsp;&nbsp;<b>You pay nothing</b></li></ul>

'Ohana Liberty (HMO-POS SNP),'Ohana Health Plan,Diabetes self-management training:<ul><li>In-network:&nbsp;&nbsp;<b>You pay nothing</b></li></ul>

'Ohana Liberty (HMO-POS SNP),'Ohana Health Plan,Therapeutic shoes or inserts:<ul><li>In-network:&nbsp;&nbsp;<b>20%</b> of the cost</li></ul>

'Ohana Liberty (HMO-POS SNP),'Ohana Health Plan,Therapeutic shoes or inserts:<ul><li>In-network:&nbsp;&nbsp;<b>You pay nothing</b></li></ul>
```

(5) Find plans that have benefits available for mental health care:

```
'Ohana Choice (HMO-POS),'Ohana Health Plan,</ul>

'Ohana Choice (HMO-POS),'Ohana Health Plan,<li><b>$250</b> copay per day for days 1
through 6</li>

'Ohana Choice (HMO-POS),'Ohana Health Plan,<li><b>$350</b> copay per day for days 1
through 4</li>

'Ohana Choice (HMO-POS),'Ohana Health Plan,<li><b>You pay nothing</b> per day for
days 5 through 90</li>

'Ohana Choice (HMO-POS),'Ohana Health Plan,<li><b>You pay nothing</b> per day for
days 7 through 90</li>

'Ohana Choice (HMO-POS),'Ohana Health Plan,<li>In-network:&nbsp;&nbsp;</li><ul
style="list-style-type:disc">

'Ohana Choice (HMO-POS),'Ohana Health Plan,<ul>

'Ohana Choice (HMO-POS),'Ohana Health Plan,<ul><li>In-network:&nbsp;&nbsp;<b>You pay
nothing</b></li></ul>

'Ohana Choice (HMO-POS),'Ohana Health Plan,Inpatient visit:

'Ohana Choice (HMO-POS),'Ohana Health Plan,Our plan also covers 60 "lifetime reserve
days."  These are "extra" days that we cover.  If your hospital stay is longer than
90 days  you can use these extra days.  But once you have used up these extra 60
days  your inpatient hospital coverage will be limited to 90 days.

'Ohana Choice (HMO-POS),'Ohana Health Plan,Our plan covers 90 days for an inpatient
hospital stay.

'Ohana Choice (HMO-POS),'Ohana Health Plan,Our plan covers up to 190 days in a
lifetime for inpatient mental health care in a psychiatric hospital.  The inpatient
hospital care limit does not apply to inpatient mental services provided in a
general hospital.

'Ohana Choice (HMO-POS),'Ohana Health Plan,Outpatient group therapy
visit:<ul><li>In-network:&nbsp;&nbsp;<b>$20</b> copay</li></ul>

'Ohana Choice (HMO-POS),'Ohana Health Plan,Outpatient group therapy
visit:<ul><li>In-network:&nbsp;&nbsp;<b>$35</b> copay</li></ul>

'Ohana Choice (HMO-POS),'Ohana Health Plan,Outpatient group therapy
visit:<ul><li>In-network:&nbsp;&nbsp;<b>You pay nothing</b></li></ul>

'Ohana Choice (HMO-POS),'Ohana Health Plan,Outpatient individual therapy
visit:<ul><li>In-network:&nbsp;&nbsp;<b>$30</b> copay</li></ul>

'Ohana Choice (HMO-POS),'Ohana Health Plan,Outpatient individual therapy
visit:<ul><li>In-network:&nbsp;&nbsp;<b>$40</b> copay</li></ul>

'Ohana Choice (HMO-POS),'Ohana Health Plan,Outpatient individual therapy
visit:<ul><li>In-network:&nbsp;&nbsp;<b>You pay nothing</b></li></ul>
```

(5) List plans by group offering the plan:

```
'Ohana Health Plan,'Ohana Choice (HMO-POS)
'Ohana Health Plan,'Ohana Liberty (HMO-POS SNP)
'Ohana Health Plan,'Ohana Value (HMO-POS)
ADVANTAGE Health Solutions  Inc.,ADVANTAGE Choice (PPO) (PPO)
ADVANTAGE Health Solutions  Inc.,ADVANTAGE Enhanced (PPO) (PPO)
ADVANTAGE Health Solutions  Inc.,ADVANTAGE Network (HMO) (HMO)
ADVANTAGE Health Solutions  Inc.,ADVANTAGE Preferred (PPO) (PPO)
ADVANTAGE Health Solutions  Inc.,ADVANTAGE Preferred Group (PPO) (PPO)
ADVANTAGE Health Solutions  Inc.,ADVANTAGE Preferred Plus Group (PPO) (PPO)
ADVANTAGE Health Solutions  Inc.,ADVANTAGE Select (PPO) (PPO)
ADVANTAGE Health Solutions  Inc.,ADVANTAGE Select Group (PPO) (PPO)
ADVANTAGE Health Solutions  Inc.,ADVANTAGE Select Plus Group (PPO) (PPO)
ADVANTAGE Health Solutions  Inc.,ADVANTAGE Special Needs Plan (HMO SNP) (HMO SNP)
AHF,PHP (HMO SNP)
ALBRIGHT CARE SERVICES,Albright LIFE (PACE)
ALBRIGHT CARE SERVICES,Albright LIFE Medicare Only (PACE)
AMERIGROUP,Amerivantage Classic + Rx (HMO)
AMERIGROUP,Amerivantage Specialty + Rx (HMO SNP)
ARKANSAS BLUE CROSS AND BLUE SHIELD,AR Blue Cross - Medi-Pak Advantage  MA (PFFS)
ARKANSAS BLUE CROSS AND BLUE SHIELD,AR Blue Cross - Medi-Pak Advantage MA (PFFS)
ARKANSAS BLUE CROSS AND BLUE SHIELD,AR Blue Cross - Medi-Pak Advantage MA-PD (PFFS)
ARKANSAS BLUE CROSS AND BLUE SHIELD,Medi-Pak Advantage Elite (PPO)
ARKANSAS BLUE CROSS AND BLUE SHIELD,Medi-Pak Advantage Essential (PPO)
```

(5) List plans by county the plan is offered in:

```
10001,ActiveSaver MSA (MSA),ActiveSaver MSA or BlueSaver MSA
10001,Advantra (HMO),Coventry Health Care
10001,Advantra (HMO-POS),Coventry Health Care
10001,Advantra (PPO),Coventry Health Care
10001,Advantra EmpGrp/NoRx/Cyear (PPO),Coventry Health Care
10001,Advantra EmpGrp/Rx/Cyear (PPO),Coventry Health Care
10001,Advantra Employer Group (HMO),Coventry Health Care
10001,Advantra Freedom Employer Group (PPO),Coventry Health Care
10001,Advantra Gold (HMO),Coventry Health Care
10001,Advantra Gold (HMO),HealthAmerica
10001,Advantra Gold (PPO),Coventry Health Care
10001,Advantra Gold (PPO),HealthAmerica
10001,Advantra MA Integrated Calendar (PPO),Coventry Health Care of Illinois  Inc.
10001,Advantra MA Only (HMO),Coventry Health Care
10001,Advantra MA Only (HMO),HealthAmerica
10001,Advantra MA Only (PPO),Coventry Health Care
10001,Advantra MA Only (PPO),Coventry Health Care of Illinois  Inc.
10001,Advantra MA Only (PPO),HealthAmerica
10001,Advantra Preferred PPO (PPO),Coventry Health Care
10001,Aetna Medicare PPO Plan with NCY Rx (PPO),Aetna Medicare
10001,Aetna Medicare Part B Only Plan (PPO),Aetna Medicare
```

7. Issues & Workarounds

- When using REGEXP_EXTRACT in Hive to find the entries in the Services table with copay and premium information, it was found that some regular expression commands that work in Python do not work in Hive. This is worked around by using alternate regular expression commands. For example, “[A-Za-z]” was used instead of “\w”, “[0-9]” was used instead of “\d”, and “<>” was used instead of “!=”.
- For the highest copays report, plans where the copay is $0 are ignored.
- Some plans have an entry where there is a space between the $ and the dollar amount (“$ 100” instead of “$100”). These entries were relatively few and for the purposes of this analysis are ignored.