# Analysis of Medicare Advantage/Part D Contract Enrollment data

This repository is a simple demo of analyzing Enrollment data of Medicare Advantage/Part D contract to identify growth areas and popular plans, features using Hive and Impala.

Author: Than Tirouvingadame

**Overview**

A Medicare Advantage Plan is a type of Medicare health plan offered by a private company that contracts with Medicare to provide you with all your Part A and Part B benefits. Medicare Advantage Plans include Health Maintenance Organizations, Preferred Provider Organizations, Private Fee-for-Service Plans, Special Needs Plans, and Medicare Medical Savings Account Plans. We use Apache Hive and Impala (open-source tool) to analyze the data and come up with growth areas and business insights.

**Installation with Oracle VM Virtual Box**

To try this demo on the Oracle VM Virtual Box, follow these steps:

* Download the Oracle VM Virtual Box and install it on your machine, following the instructions here: <http://www.cloudera.com/content/cloudera/en/downloads/quickstart_vms/cdh-5-3-x.html>. We recommend at least 4GB to be allocated to the Virtual Machine. Whether you use VMWare, VirtualBox or KVM is your choice, the demo should work equally well on each of these products.
* The code is available in the code folder.
* The dataset is available in the data folder.
* The research document is available in the document file.