Econ 470 - HW1

Ron Huang

Table of contents

1 Github: https://github.com/Rorn001/Econ470.git

2 Enrollment Data

```
C:\Users\huang\AppData\Local\Temp\ipykernel_9152\1604708242.py:1: DtypeWarning: Columns (11)
  data = pd.read_csv(r'data\output\enrollment data-Py\full_ma_data.csv',
```

2.1 How many observations exist in your current dataset?

```
19126783
```

3 How many different plan_types exist in the data?

```
'the number of distinct values in plan_type is 26'
```

```
array(['HCPP - 1833 Cost', 'Pilot',
       'Employer/Union Only Direct Contract PDP', 'Employer Direct PFFS',
       'Local PPO', 'HMO/HMOPOS', 'National PACE',
       'Continuing Care Retirement Community', '1876 Cost', 'PFFS',
       'PSO (State License)', 'MA Health Senior Care Options',
       'MN Disability Health Options', 'MN Senior Health Options', 'SHMO',
       'PSO (Federal Waiver of State License)', 'WI Partnership Program',
       'ESRD I', 'ESRD II', 'MSA', 'MSA Demo', 'Regional PPO',
       'Medicare Prescription Drug Plan',
       'Employer/Union Only Direct Contract PFFS', nan, 'RFB PFFS',
       'Medicare-Medicaid Plan HMO/HMOPOS'], dtype=object)
```

3.1 Provide a table of the count of plans under each plan type in each year.

year plan_type	2010	2011	2012	2013	2014	2015
1876 Cost	6035	6851	7633	7731	7069	7157
Continuing Care Retirement Community	142	0	0	0	0	0
ESRD I	117	0	0	0	0	0
ESRD II	8	0	0	0	0	0
Employer/Union Only Direct Contract PDP	28700	28697	28669	25526	25528	25630
Employer/Union Only Direct Contract PFFS	3332	3329	3323	0	0	0
HCPP - 1833 Cost	3604	11	11	10	9	9

year plan_type	2010	2011	2012	2013	2014	2015
HMO/HMOPOS	506802	528473	507272	530909	523304	479275
Local PPO	417551	515700	636701	633884	664716	704993
MSA	135	6421	6416	6431	6449	6518
Medicare Prescription Drug Plan	893609	771694	815223	826907	1122209	991457
Medicare-Medicaid Plan HMO/HMOPOS	0	0	0	265	1319	4130
National PACE	717	781	858	953	1118	1216
PFFS	385733	45781	36423	31919	24905	13658
PSO (State License)	123	176	171	0	0	0
Pilot	53	3	3	2	2	2
Regional PPO	24442	22773	21602	19970	19773	17578

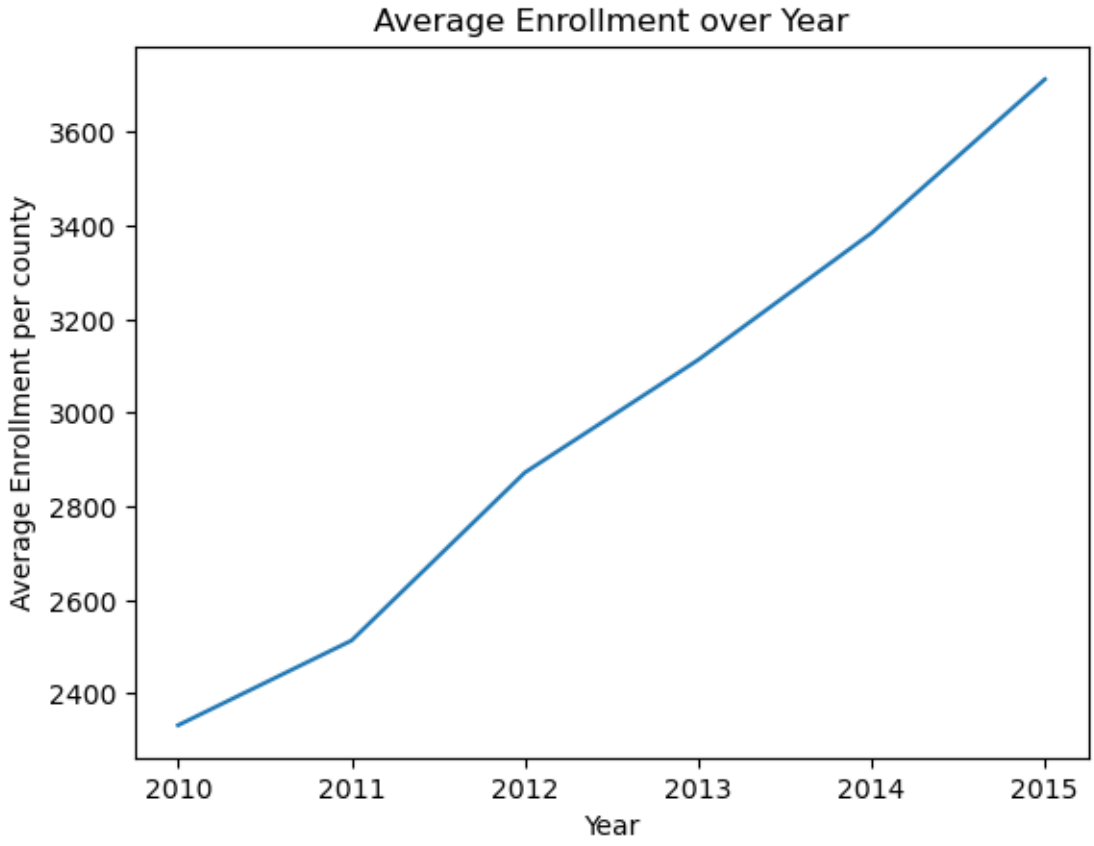
3.2 Remove all special needs plans (SNP), employer group plans (eghp), and all "800-series" plans.

year plan_type	2010	2011	2012	2013	2014	2015
1876 Cost	4923	5829	6647	6759	6207	6329
Continuing Care Retirement Community	64	0	0	0	0	0
ESRD I	117	0	0	0	0	0
HMO/HMOPOS	34460	33931	37551	37179	38893	36588
Local PPO	11652	13874	17030	17089	17169	16728
MSA	68	131	132	145	163	232
Medicare Prescription Drug Plan	391205	295458	289044	278091	301082	269153
Medicare-Medicaid Plan HMO/HMOPOS	0	0	0	265	1319	4130
National PACE	717	781	858	953	1118	1216
PFFS	54119	22038	17449	12945	6053	4232
PSO (State License)	97	141	143	0	0	0
Regional PPO	10659	10995	11279	9660	10420	8531

3.3 Merge the contract service area data to the enrollment data, and restrict the data only to contracts that are approved in their respective counties. Limiting your dataset only to plans with non-missing enrollment data, provide a graph showing the average number of Medicare Advantage enrollees per county from 2010 to 2015.

```
C:\Users\huang\AppData\Local\Temp\ipykernel_9152\2472351519.py:2: DtypeWarning: Columns (10)
  area_data = pd.read_csv('data\output\contract_service_area.csv', encoding='ISO-8859-1')
```

	contractid	org_name	org_type
889797	90091	UNITED MINE WORKERS OF AMERICA HEALTH & RETIRE...	HCPP - 1833 Co
889798	E5088	DESERET HEALTHCARE EMPLOYEE BENEFITS TRUST	Employer/Union
889799	E5088	DESERET HEALTHCARE EMPLOYEE BENEFITS TRUST	Employer/Union
889800	E5088	DESERET HEALTHCARE EMPLOYEE BENEFITS TRUST	Employer/Union
889801	E5088	DESERET HEALTHCARE EMPLOYEE BENEFITS TRUST	Employer/Union

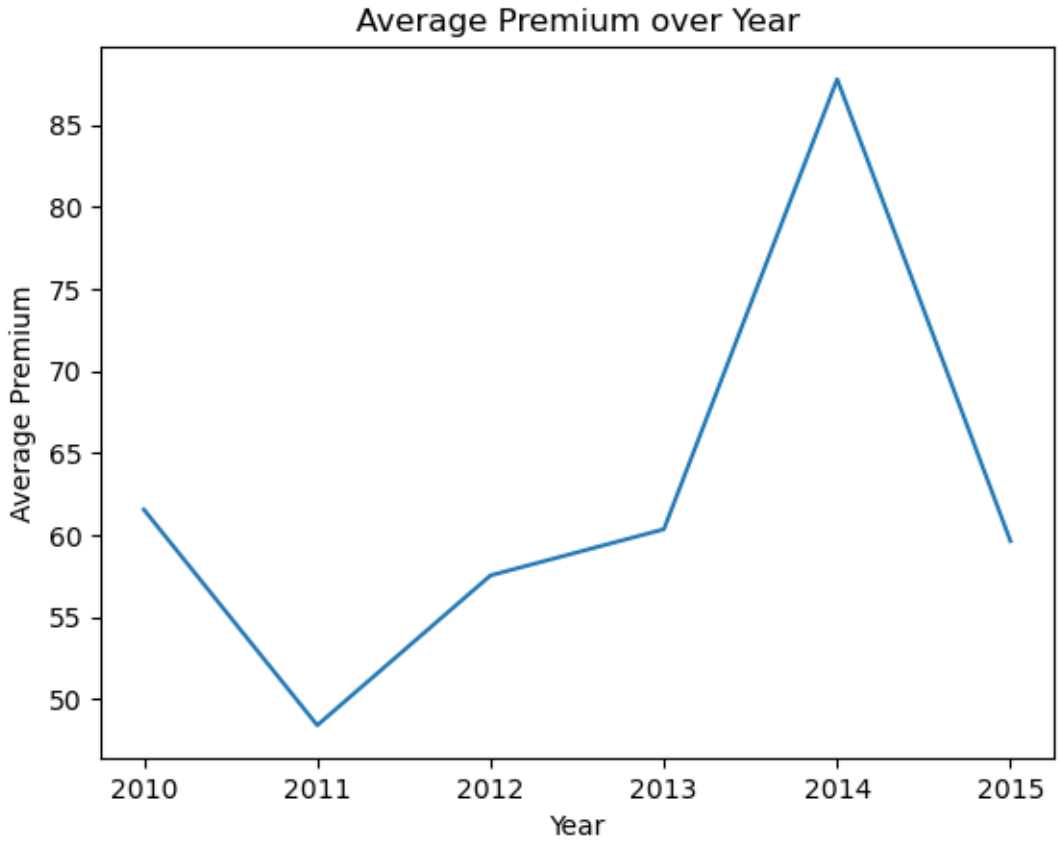


4 Premium Data

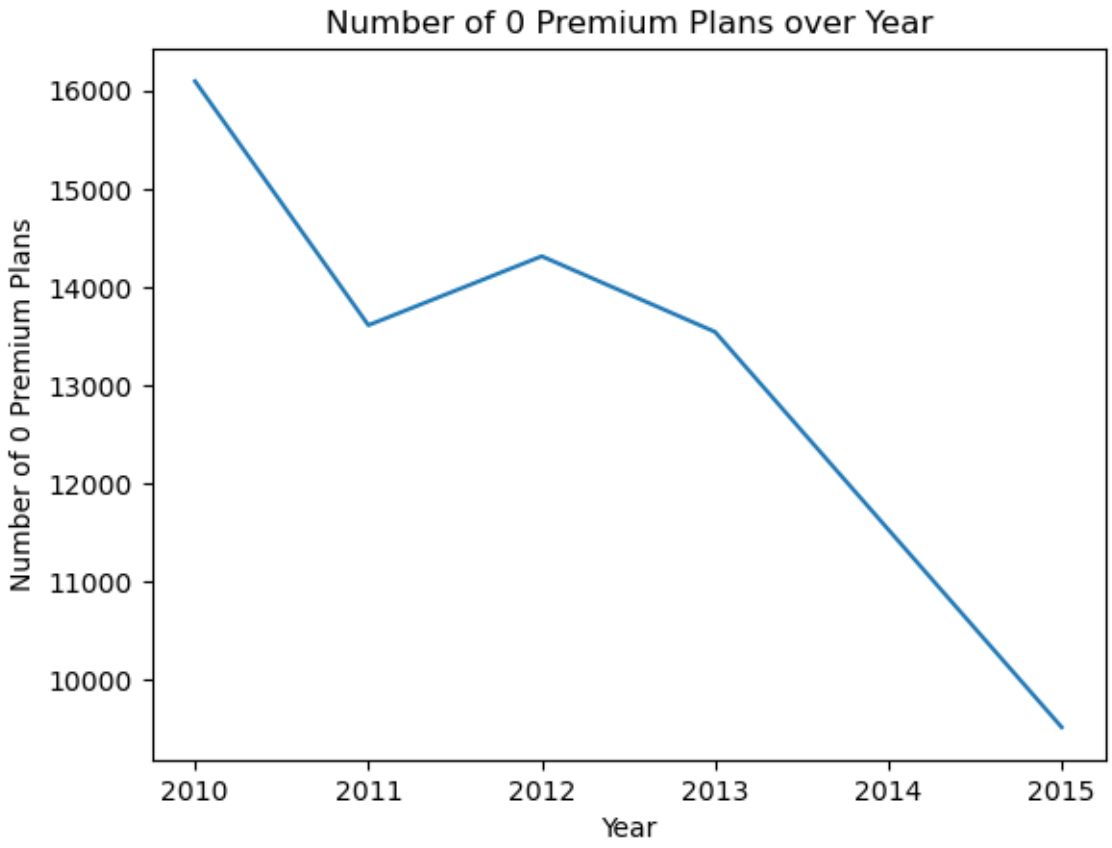
4.1 Merge the plan characteristics data to the dataset you created in Step 5 above. Note that you'll need to join the Market Penetration Data in order to get the information you need to merge the plan characteristics. Provide a graph showing the average premium over time.

	contractid	planid	state	county	premium	premium_partc	premium_partd_basic	premium
0	H0104	4	Alabama	Autauga	41.0	23.3	17.7	0.0
1	H0104	8	Alabama	Autauga	142.0	107.7	26.5	7.8
2	H0104	2	Alabama	Autauga	0.0	0.0	0.0	0.0
3	H2762	15	Alabama	Autauga	80.0	NaN	NaN	NaN
4	H2762	19	Alabama	Autauga	135.0	NaN	NaN	NaN

	Unnamed: 0	contractid	planid	org_type_x	plan_type_x	partd	snp	eghp_x	org_name_x
0	5884037	H0084	1.0	Local CCP	Local PPO	Yes	No	No	CARE IMPROV
1	5884038	H0084	1.0	Local CCP	Local PPO	Yes	No	No	CARE IMPROV
2	5884039	H0084	1.0	Local CCP	Local PPO	Yes	No	No	CARE IMPROV
3	5884041	H0084	1.0	Local CCP	Local PPO	Yes	No	No	CARE IMPROV
4	5884042	H0084	1.0	Local CCP	Local PPO	Yes	No	No	CARE IMPROV



4.2 Provide a graph showing the percentage of $0 premium plans over time.



5 Summary Questions

5.1 Why did we drop the "800-series" plans?

year	2011
plan_type	
PFFS	1774

800-series plans are coded as the employer group waiver plans, but some of them are not coded as true in the egwp column. As shown in the above tabel, there are 1774 plans that are 800-series plans but not egwp. Therefore, we drop the 800-series plans to exclude all egwp plans.

5.2 Why do so many plans charge a $0 premium? What does that really mean to a beneficiary?

0 premium plans may be designed to provide health insurance to those who cannot afford other expensive plans or to those who think they are healthy and do not need to pay for higher premium. It seems it is a good way to contain cost for the enrolles, but 0 premium plans may still result in high out-of-pocket cost due to high deductible, copayment, or coinsurance rate. Also those plans may have limited coverage and network, so out-of-network payment and surprsing bills could also occur.

5.3 Briefly describe your experience working with these data (just a few sentences). Tell me one thing you learned and one thing that really aggravated you.

The most troublesome issue when dealing with these data is the size the data. It takes very long to load data, and there are raw data such that we have to spend quite a lot of time to clean it before any analysis. Sometimes my computer may run out of memory and I have to restart it.