

Analysis Report of Insurance Market Data & Analytics

Group 2

Menghong Han, Xia Dai, Rui Cao, Mengchun Li, Xuhui Bai, Bo Chen

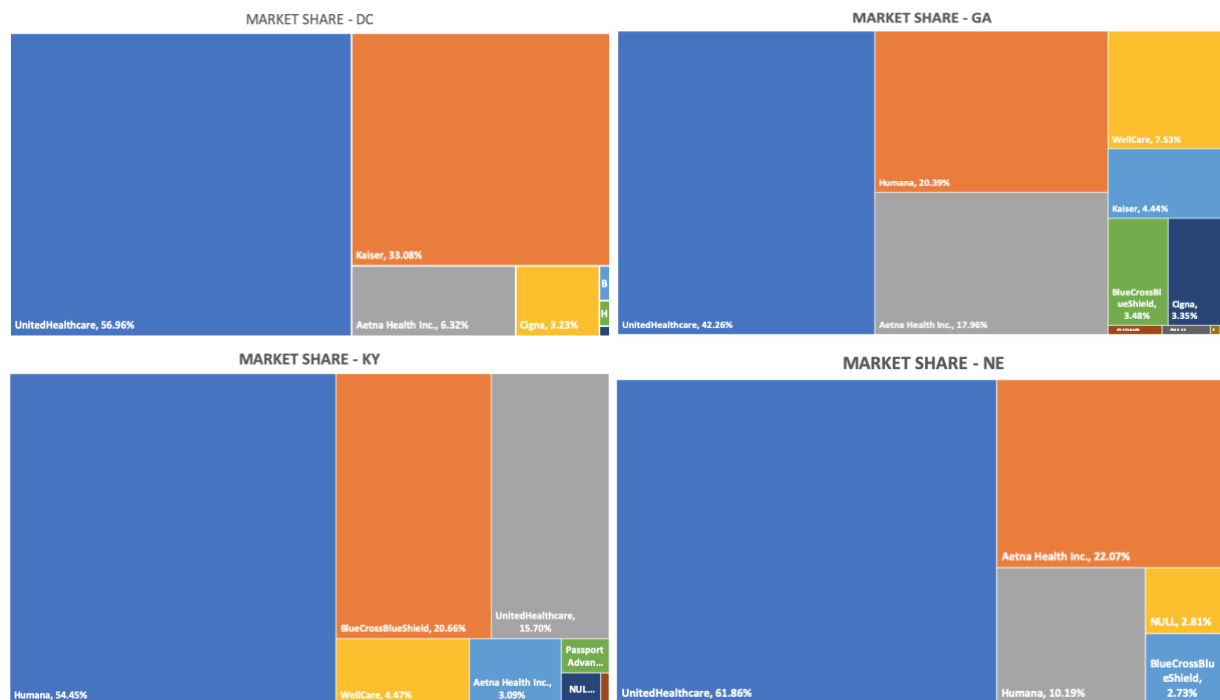
Declaration

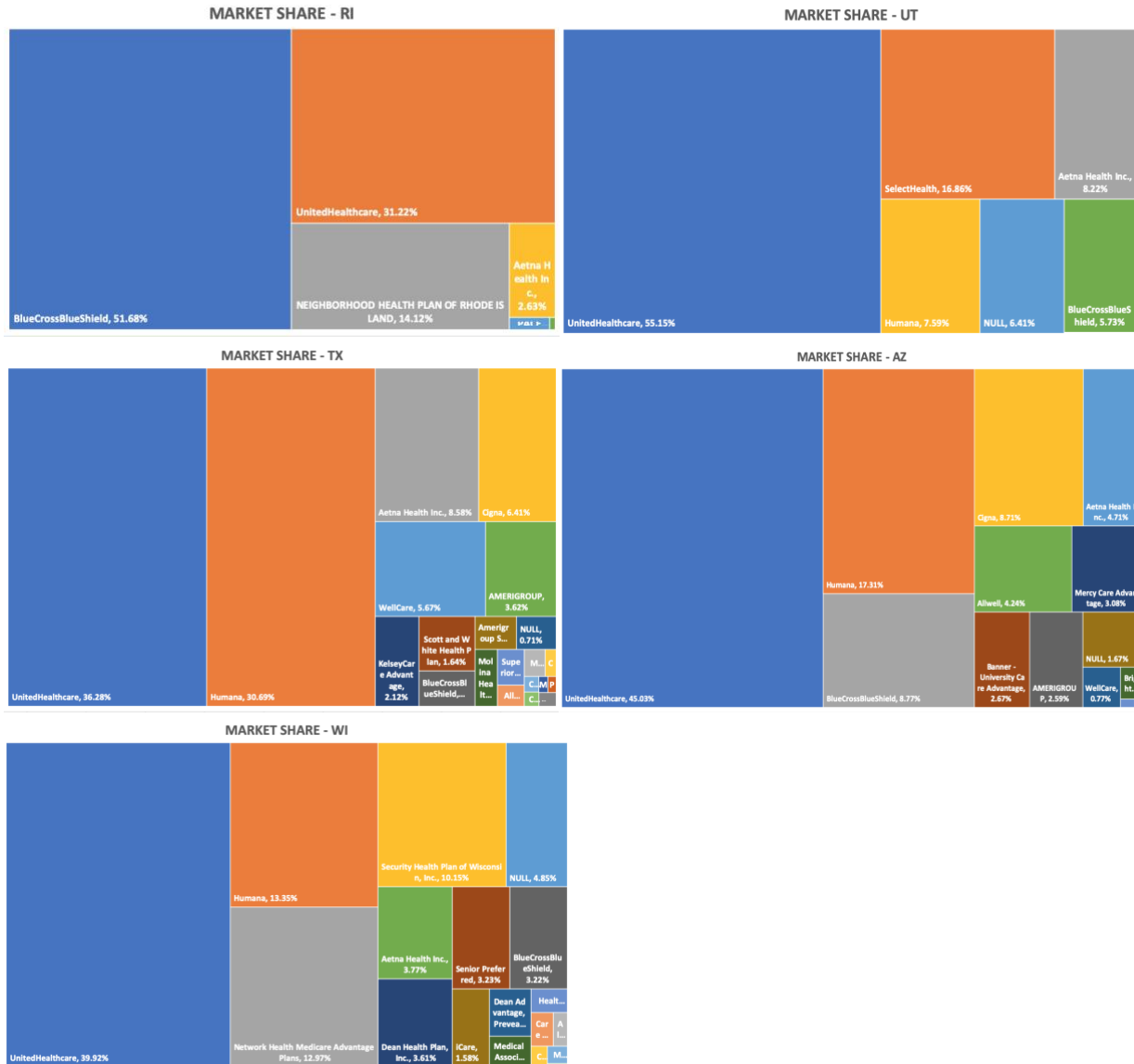
For this analysis report, we focus on nine states: AZ, GA, DC, KY, NE, TX, WI, UT and RI.

Part 1. Finding a partner for a “private single-payer” proposal

We used treemap below to show the main insurance companies in the nine states. Some small-sized companies are not fully presented in the graph, please refer to the appendix for the full list of market share in each state. The size of each company indicates the company’s market share in that state: the larger size, the higher market share.

As shown below, every state has 1 - 3 “big players” that take most market share in that state. UnitedHealthcare is the most frequent name shown in our graph: it is the biggest insurance company in 7 out of 9 states. Humana and BlueCrossBlueShield are also strong: even though they are not the biggest one, they still play as the 2nd or 3rd biggest insurance company in many states. TX, WI, and AZ have a relatively diverse insurance market: these three states have more insurance companies than the rest six states.





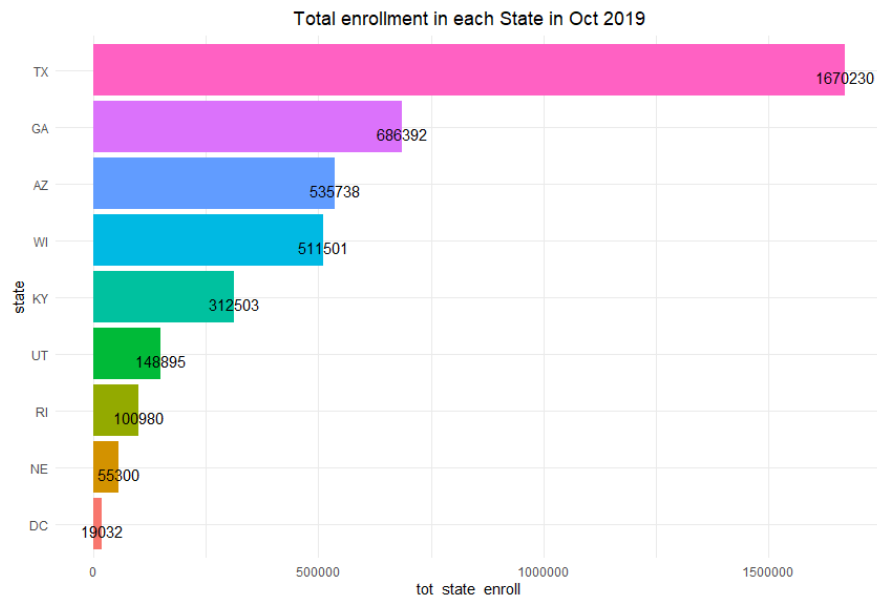
The HHI result shows that top 4 concentrated states are NE, DC, RI, and KY. The company having “Lion Share” in these states are listed below.

state	HHI	MajorInsuranceOrgName	market_share
NE	0.4433	UnitedHealthcare	61.86%
DC	0.4389	UnitedHealthcare	56.96%
RI	0.3852	BlueCrossBlueShield	51.68%
KY	0.3669	Humana	54.45%
UT	0.3525		
GA	0.2624		
AZ	0.2547		
TX	0.2428		
WI	0.2119		

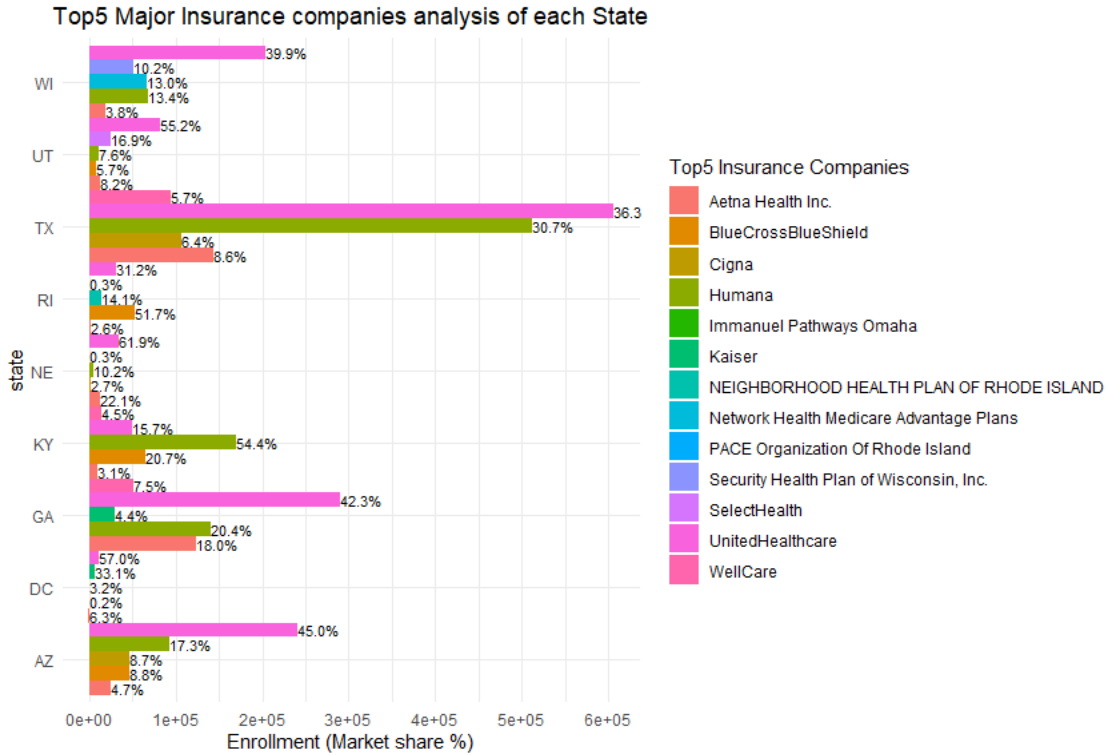
Part 2. Examine the Insurance Benefit Package

Data Table for Top 5 Major Insurance Companies in Each State with Market share and Preventive Dental and Comprehensive Dental data details								Total Enrollment in each State	
major_org_revised	state	major_state	prev_major_stat	comp_major_stat	mkt_share	prev_major_state_pc	comp_major_state_pct	state	tot_state_enroll
UnitedHealthcare	AZ	241,240	167,350	51,146	45.03%	63.37%	21.20%	AZ	535,738
Humana	AZ	92,743	85,304	24,183	17.31%	91.98%	26.08%	DC	19,032
BlueCrossBlueShield	AZ	46,965	17,242	13	8.77%	36.71%	0.03%	GA	686,392
Cigna	AZ	46,638	45,980	45,980	8.71%	98.53%	98.53%	KY	312,503
Aetna Health Inc.	AZ	25,236	16,749	16,749	4.71%	66.37%	66.37%	NE	55,300
UnitedHealthcare	DC	10,840	10,123	10,123	56.96%	93.39%	93.39%	RI	100,980
Kaiser	DC	6,296	3,530	3,530	33.08%	56.07%	56.07%	TX	1,670,230
Aetna Health Inc.	DC	1,202	393	393	6.32%	32.70%	32.70%	UT	148,895
Cigna	DC	615	512	315	3.23%	83.25%	51.22%	WI	148,895
BlueCrossBlueShield	DC	39	-	-	0.20%	0.00%	0.00%		511,501
UnitedHealthcare	GA	290,077	17,544	17,544	42.26%	6.05%	6.05%		
Humana	GA	139,959	120,347	120,347	20.39%	85.99%	85.99%		
Aetna Health Inc.	GA	123,242	84,003	84,003	17.96%	68.16%	68.16%		
WellCare	GA	51,693	23,705	23,705	7.53%	45.86%	45.86%		
Kaiser	GA	30,470	24,090	24,090	4.44%	79.06%	79.06%		
Humana	KY	170,173	106,938	106,938	54.45%	62.84%	62.84%		
BlueCrossBlueShield	KY	64,562	63,968	9,932	20.66%	99.08%	15.38%		
UnitedHealthcare	KY	49,077	4,636	4,636	15.70%	9.57%	9.57%		
WellCare	KY	13,971	13,971	13,971	4.47%	100.00%	100.00%		
Aetna Health Inc.	KY	9,652	2,194	2,194	3.09%	22.73%	22.73%		
UnitedHealthcare	NE	34,209	22,575	22,575	61.86%	65.99%	65.99%		
Aetna Health Inc.	NE	12,203	11,496	11,496	22.07%	94.21%	94.21%		
Humana	NE	5,637	5,351	5,351	10.19%	94.93%	94.93%		
BlueCrossBlueShield	NE	1,510	1,447	1,447	2.73%	95.83%	95.83%		
Immanuel Pathways Omaha	NE	173	-	-	0.31%	0.00%	0.00%		
BlueCrossBlueShield	RI	52,190	24,257	24,257	51.68%	46.48%	46.48%		
UnitedHealthcare	RI	31,530	26,098	24,807	31.22%	82.77%	78.68%		
NEIGHBORHOOD HEALTH PLAN OF RHODE ISLAND	RI	14,263	-	14,263	14.12%	0.00%	100.00%		
Aetna Health Inc.	RI	2,659	2,219	2,219	2.63%	83.45%	83.45%		
PACE Organization Of Rhode Island	RI	299	-	-	0.30%	0.00%	0.00%		
UnitedHealthcare	TX	605,915	466,880	377,994	36.28%	77.05%	62.38%		
Humana	TX	512,546	273,517	273,517	30.69%	53.36%	53.36%		
Aetna Health Inc.	TX	143,264	38,589	38,589	8.58%	26.94%	26.94%		
Cigna	TX	107,110	106,509	106,509	6.41%	99.44%	99.44%		
WellCare	TX	94,652	94,083	94,083	5.67%	99.40%	99.40%		
UnitedHealthcare	UT	82,110	79,464	54,440	55.15%	96.78%	66.30%		
SelectHealth	UT	25,101	24,865	24,865	16.86%	99.06%	99.06%		
Aetna Health Inc.	UT	12,240	10,488	10,488	8.22%	85.69%	85.69%		
Humana	UT	11,306	11,168	4,447	7.59%	98.78%	33.33%		
BlueCrossBlueShield	UT	8,526	7,796	7,396	5.73%	91.44%	86.75%		
UnitedHealthcare	WI	204,195	160,433	130,035	39.92%	78.57%	63.68%		
Humana	WI	68,310	63,125	41,858	13.35%	92.41%	61.28%		
Network Health Medicare Advantage Plans	WI	66,340	36,579	2,561	12.97%	55.14%	3.86%		
Security Health Plan of Wisconsin, Inc.	WI	51,935	48,158	48,158	10.15%	92.73%	92.73%		
Aetna Health Inc.	WI	19,284	5,989	5,989	3.77%	31.06%	31.06%		

For your reference, above is the data snapshot we extracted for analysis.

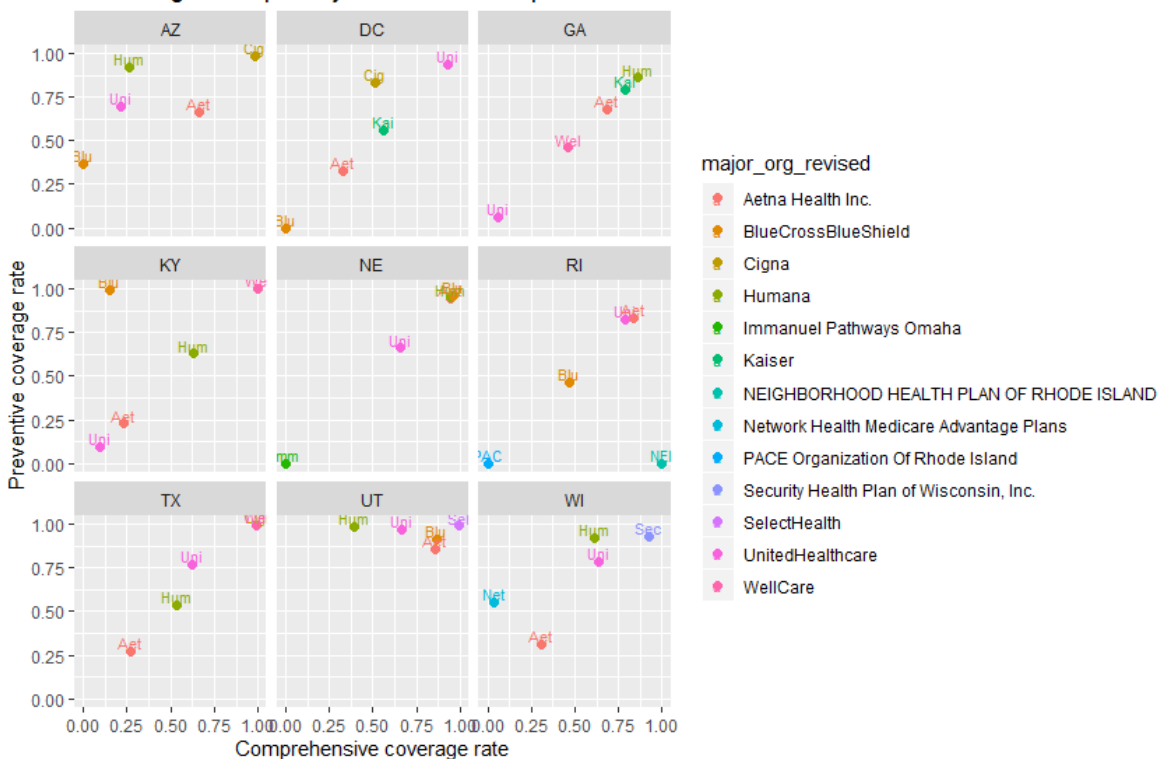


As the above graph shown, we explored the enrollment and plan data, and obtained the market size for each State our team work on. We can see that Texas has the largest market size in our 9 States and DC is comparably the smallest one because it is only a district.



In addition, we plot the top5 major insurance companies in each State with enrollment and market share. We can see that some of the companies have large enrollment and large market share in the State, like United Health Care in Arizona(45%), Georgia(42.3%), Texas(36.3%). However, some of the major companies have very small market share and enrollment. The reason might be the State is small or population is young or the people there are more distributed in choosing insurance companies.

Dental Coverage of Top5 Major Insurance companies each State



Finally, we plot the coverage rate correlation between Preventive dental item and Comprehensive dental item as above. Firstly, we would love to explain the meaning of two dental items. Preventive care includes dental x-ray, oral exam, cleaning, etc, while Comprehensive care includes treatments that maintain or restore dental health (e.g., diagnosis, treatment for missing teeth or prosthodontics, treatment for gum diseases, periodontics, etc.)

We analyzed the two dental items coverage in each state and offered recommendation for each State based on the company has high and balanced rate of dental service coverage between Preventive & Comprehensive Dental Items, to help make decisions on the choice of insurance companies regarding the dental items.

In state AZ, all top 5 insurance companies have Preventive & Comprehensive Dental Items as a supplemental benefit under Part C. Cigna (P/C: 98.59%) and Aetna Health Inc. (P/C: 66.37%) have the same proportion in Preventive & Comprehensive Dental Items, while Humana (P: 91.98%, C: 26.08%), UnitedHealthcare (P: 69.37%, C: 21.20%), BlueCrossBlueShield (P: 36.71%, C: 0.03%) have higher Preventive Dental Items percentage than that of Comprehensive one.

Recommendation for AZ: Cigna (98.59%).

In state DC, 4 of the top 5 insurance companies have Preventive & Comprehensive Dental Items as a supplemental benefit under Part C. UnitedHealthcare (P/C: 93.39%), Kaiser (P/C: 56.07%), Aetna Health Inc. (P/C: 32.70%) have the same proportion in Preventive &

Comprehensive Dental Items, while Cigna (P: 83.25%, C: 51.22%) has higher Preventive Dental Items percentage than that of Comprehensive one.

Recommendation for DC: UnitedHealthcare (P/C: 93.39%).

In state GA, all top 5 insurance companies have Preventive & Comprehensive Dental Items as a supplemental benefit under Part C. and they all have the same proportion in Preventive & Comprehensive Dental Items. Humana (P/C: 85.99%), Kaiser (P/C: 79.06%), Aetna Health Inc. (P/C: 68.16%), WellCare (P/C: 45.86%), UnitedHealthcare (P/C: 6.05%).

Recommendation for GA: Humana (P/C: 85.99%).

In state KY, all top 5 insurance companies have Preventive & Comprehensive Dental Items as a supplemental benefit under Part C. WellCare (P/C: 100.00%), Humana (P/C: 62.84%), Aetna Health Inc. (P/C: 22.73%), UnitedHealthcare (P/C: 9.57%) have the same proportion in Preventive & Comprehensive Dental Items, while BlueCrossBlueShield (P: 99.08%, C: 15.38%) has much higher Preventive Dental Items percentage than that of Comprehensive one. BlueCrossBlueShield (P/C: 99.08%)

Recommendation for KY: WellCare (P/C: 100.00%).

In state NE, 4 of the top 5 insurance companies have Preventive & Comprehensive Dental Items as a supplemental benefit under Part C. And they all have the same proportion in Preventive & Comprehensive Dental Items. BlueCrossBlueShield (P/C: 95.83%), Humana (P/C: 94.93%), Aetna Health Inc. (P/C: 94.21%), UnitedHealthcare (P/C: 65.99%).

Recommendation for NE: BlueCrossBlueShield (P/C: 95.83%)

In State RI, 3 of the top 5 insurance companies have Preventive & Comprehensive Dental Items as a supplemental benefit under Part C. BlueCrossBlueShield (P/C: 46.48%) and Aetna Health Inc. (P/C: 83.45%) have the same proportion in Preventive & Comprehensive Dental Items. While UnitedHealthcare (P: 82.77%, C: 78.68%) has a little higher rate in Preventive. Immanuel Pathways Omaha has no coverage for both items and NEIGHBORHOOD HEALTH PLAN OF RHODE ISLAND only 100% covers Comprehensive Dental Items.

Recommendation for RI: Aetna Health Inc. (P/C: 83.45%).

In State TX, all top 5 insurance companies have Preventive & Comprehensive Dental Items as a supplemental benefit under Part C. Only UnitedHealthcare (P: 77.05%, C: 62.38%) has a little higher rate in Preventive, others have the same proportion. Humana (P/C: 53.36%), Aetna Health Inc. (26.94%), Cigna (99.44%), WellCare (99.40%).

Recommendation for TX: Cigna (99.44%) and WellCare (99.40%)

In State UT, all top 5 insurance companies have Preventive & Comprehensive Dental Items as a supplemental benefit under Part C. UnitedHealthcare (P: 96.78%, C: 66.30%), Humana (P:

98.78%, C: 39.33%), BlueCrossBlueShield (91.44%, 86.75%) have a higher rate in Preventive. Other two have the same proportion. SelectHealth (P/C: 99.06%), Aetna Health Inc. (85.69%).

Recommendation for UT: SelectHealth (P/C: 99.06%).

In State WI, all top 5 insurance companies have Preventive & Comprehensive Dental Items as a supplemental benefit under Part C. Security Health Plan of Wisconsin, Inc.(P/C: 92.73%) and Aetna Health Inc. (31.06%) have the same proportion. While others UnitedHealthcare (P:78.57%, C:63.68%), Humana (P: 92.41%, C: 61.28%) and Network Health Medicare Advantage Plans (P: 55.14%, C: 3.86%) have a higher rate in Preventive.

Recommendation for WI: Security Health Plan of Wisconsin, Inc.(P/C: 92.73%).

In summary, from the chart and data analysis of Preventive & Comprehensive Dental Items proportion as a supplemental benefit under Part C, we can find that for most top 5 insurance companies (69%) in 9 states, the proportion of Preventive & Comprehensive Dental Items is nearly the same but the divisions between different companies are very fragmented, while the majority of the rest companies tend to have higher proportion of Preventive Dental Items (29%) than that of Comprehensive (2%).

Part 3. Quality of care and performance of the plans

On part 3, we get insights into the weighted average based on question1. Firstly, we just imported all useful data from python to SQL and tried to merge all useful tables together with corresponding variables including UOD rate, company name, contract, state and the number of enrollment in each contract. On top of that, we cleaned data by filtering out insignificant UOD rates. Afterward, we calculated the weighted average by summing up the product of UOD rate and the number of enrollment in each contract and dividing by the total number of enrollment in each company in each state. As a result, we obtained a data table like this:

State	Organization Marketing Name	Weighted	Total_Enrollment	Weighted_Average
AZ	UnitedHealthcare	1923.04	17	113.12
AZ	Allwell	27862.38	351	79.38
AZ	Humana	82810.4	1360	60.89
AZ	Cigna	658.92	12	54.91
AZ	Amerigroup	141388.8	4607	30.69

DC	Cigna	13327.05	197	67.65
DC	UnitedHealthcare	7619	190	40.1
GA	Humana	1519.35	21	72.35
GA	UnitedHealthcare	811.58	14	57.97
GA	Cigna	2839.8	60	47.33
GA	WellCare	1807.72	43	42.04
KY	Humana	33773.48	607	55.64
KY	UnitedHealthcare	1483.7	37	40.1
KY	Anthem HealthKeepers	144.1	11	13.1
NE	Rocky Mountain Health Plans	764.4	12	63.7
NE	Humana	12847.79	211	60.89
NE	UnitedHealthcare	2726.8	68	40.1
RI	UnitedHealthcare	695.64	12	57.97
TX	Humana	71484.86	1174	60.89
TX	UnitedHealthcare	985.49	17	57.97
TX	Amerigroup	126964.5	4137	30.69
TX	KelseyCare Advantage	4111.38	273	15.06
TX	WellCare	512.4	42	12.2
UT	Humana	42707.14	406	105.19
UT	SelectHealth	158526.8	1781	89.01
UT	UnitedHealthcare	218785.6	5456	40.1
WI	iCare	4494.1	26	172.85
WI	UnitedHealthcare	6798.88	88	77.26

According to the derived table, it is not hard to find that UnitedHealthCare, which has a weighted average of top3 among all companies in each state, is the most popular in

each state. Additionally, it is obvious that Humana is more likely to have the biggest weighted average among companies in each state.

Appendix

Part1: Clean and analyze the dataset in Mysql

```
CREATE DATABASE db_hc_assignment_2;
```

```
USE db_hc_assignment_2;
```

```
DELETE FROM monthly_report_by_plan_2019_10 WHERE Enrollment = '*';
```

```
DELETE FROM Enrollment WHERE Enrollment = '*';
```

```
drop table if exists Enrollment2;
```

```
CREATE TABLE Enrollment2
```

```
SELECT * FROM Enrollment WHERE state IN ('TX', 'GA', 'AZ', 'WI', 'KY', 'UT', 'NE', 'RI', 'DC');
```

```
DELETE FROM Enrollment2 WHERE `Contract Number` LIKE 'S%';
```

```
DELETE FROM monthly_report_by_plan_2019_10 WHERE `Contract_Number` LIKE 'S%';
```

```
drop table if exists bigguy_enrollment_by_state;
```

```
CREATE temporary TABLE bigguy_enrollment_by_state
```

```
SELECT d.MajorInsuranceOrgName, c.state, SUM(c.total_enrollment) FROM
```

```
    (SELECT Organization_Marketing_Name, state, SUM(Enrollment) AS total_enrollment FROM
```

```
        (SELECT a.Contract_Number, a.state, a.Enrollment, b.Organization_Marketing_Name
```

```
FROM Enrollment2 a
```

```
    LEFT JOIN monthly_report_by_plan_2019_10 b
```

```
        ON a.Contract_Number = b.Contract_Number and a.`Plan ID` = b.`plan id` ) k
```

```
    GROUP BY Organization_Marketing_Name, state) c
```

```
LEFT JOIN majorinsuranceorgs d
```

```
ON c.Organization_Marketing_Name = d.Organization_Marketing_Name
```

```
GROUP BY state, MajorInsuranceOrgName;
```

```
SELECT state, MajorInsuranceOrgName, ROUND(`SUM(c.total_enrollment)`/total_enrollment_by_state,  
4) AS market_share FROM
```

```
    (SELECT a.*, b.total_enrollment_by_state from bigguy_enrollment_by_state a
```

```
    left join (SELECT state, SUM(Enrollment) AS total_enrollment_by_state FROM(
```

```
        SELECT a.Contract_Number, a.state, a.Enrollment,
```

```
b.Organization_Marketing_Name FROM Enrollment2 a
```

```
    LEFT JOIN
```

```
        monthly_report_by_plan_2019_10 b
```

```
            ON a.Contract_Number = b.Contract_Number and a.`Plan ID` = b.`plan id`) k
```

```
    GROUP BY state) b
```

```
    ON a.state = b.state)l
```

```
order by state, market_share DESC;
```

```
SELECT state, ROUND(SUM(sqr_share), 4) as HHI FROM
```

```
(SELECT *, power(market_share, 2) AS sqr_share FROM
```

```
    (SELECT state, MajorInsuranceOrgName,
```

```
ROUND(`SUM(c.total_enrollment)`/total_enrollment_by_state, 4) AS market_share FROM
```

```
    (SELECT a.*, b.total_enrollment_by_state from bigguy_enrollment_by_state a
```

```

left join (SELECT state, SUM(Enrollment) AS total_enrollment_by_state FROM(
SELECT a.Contract_Number, a.state, a.Enrollment,
b.Organization_Marketing_Name FROM Enrollment2 a
LEFT JOIN
monthly_report_by_plan_2019_10 b
ON a.Contract_Number = b.Contract_Number and a.`Plan ID` = b.`plan id`) k
GROUP BY state) b
ON a.state = b.state) L
ORDER BY state, market_share DESC) l) f
group by state
ORDER BY HHI DESC;

```

Market share in each state

state	MajorInsuranceOrgName	market_share
AZ	UnitedHealthcare	45.03%
AZ	Humana	17.31%
AZ	BlueCrossBlueShield	8.77%
AZ	Cigna	8.71%
AZ	Aetna Health Inc.	4.71%
AZ	Allwell	4.24%
AZ	Mercy Care Advantage	3.08%
AZ	Banner - University Care Advantage	2.67%
AZ	AMERIGROUP	2.59%
AZ	NULL	1.67%
AZ	WellCare	0.77%
AZ	Bright Health	0.35%
AZ	Kaiser	0.06%
AZ	Lasso Healthcare	0.02%
AZ	Clover Health	0.01%

AZ	Excellus Health Plan, Inc	0.00%
DC	UnitedHealthcare	56.96%
DC	Kaiser	33.08%
DC	Aetna Health Inc.	6.32%
DC	Cigna	3.23%
DC	BlueCrossBlueShield	0.20%
DC	Humana	0.15%
DC	NULL	0.06%
GA	UnitedHealthcare	42.26%
GA	Humana	20.39%
GA	Aetna Health Inc.	17.96%
GA	WellCare	7.53%
GA	Kaiser	4.44%
GA	BlueCrossBlueShield	3.48%
GA	Cigna	3.35%
GA	Allwell	0.27%
GA	NULL	0.24%
GA	Clover Health	0.05%
GA	Clear Spring Health	0.03%
GA	AHF	0.00%
KY	Humana	54.45%
KY	BlueCrossBlueShield	20.66%
KY	UnitedHealthcare	15.70%
KY	WellCare	4.47%

KY	Aetna Health Inc.	3.09%
KY	Passport Advantage	0.83%
KY	NULL	0.64%
KY	Mutual of Omaha Medicare Advantage	0.07%
KY	Signature Advantage (HMO SNP)	0.06%
KY	Molina Healthcare of Ohio	0.01%
KY	Anthem HealthKeepers	0.00%
NE	UnitedHealthcare	61.86%
NE	Aetna Health Inc.	22.07%
NE	Humana	10.19%
NE	NULL	2.81%
NE	BlueCrossBlueShield	2.73%
NE	Immanuel Pathways Omaha	0.31%
NE	Rocky Mountain Health Plans	0.02%
RI	BlueCrossBlueShield	51.68%
RI	UnitedHealthcare	31.22%
RI	NEIGHBORHOOD HEALTH PLAN OF RHODE ISLAND	14.12%
RI	Aetna Health Inc.	2.63%
RI	PACE Organization Of Rhode Island	0.30%
RI	NULL	0.04%
TX	UnitedHealthcare	36.28%
TX	Humana	30.69%
TX	Aetna Health Inc.	8.58%
TX	Cigna	6.41%

TX	WellCare	5.67%
TX	AMERIGROUP	3.62%
TX	KelseyCare Advantage	2.12%
TX	Scott and White Health Plan	1.64%
TX	BlueCrossBlueShield	1.07%
TX	Amerigroup STAR+PLUS MMP	0.74%
TX	NULL	0.71%
TX	Molina Healthcare of Texas	0.68%
TX	Superior HealthPlan	0.52%
TX	Allwell	0.31%
TX	Memorial Hermann Health Plan	0.30%
TX	Care N' Care Insurance Company	0.16%
TX	Cigna-HealthSpring CarePlan	0.12%
TX	CHRISTUS Health Plan Generations	0.11%
TX	Molina Healthcare of Texas, Inc.	0.08%
TX	Prominence Health Plan	0.07%
TX	Bienvivir Senior Health Services	0.05%
TX	Imperial Insurance Company of Texas, Inc	0.03%
TX	Clover Health	0.02%
TX	Mutual of Omaha Medicare Advantage	0.01%
TX	Silver Star	0.01%
TX	The Basics at Jan Werner	0.01%
TX	Lasso Healthcare	0.00%
TX	Kaiser	0.00%

UT	UnitedHealthcare	55.15%
UT	SelectHealth	16.86%
UT	Aetna Health Inc.	8.22%
UT	Humana	7.59%
UT	NULL	6.41%
UT	BlueCrossBlueShield	5.73%
UT	Lasso Healthcare	0.04%
UT	Kaiser	0.01%
WI	UnitedHealthcare	39.92%
WI	Humana	13.35%
WI	Network Health Medicare Advantage Plans	12.97%
WI	Security Health Plan of Wisconsin, Inc.	10.15%
WI	NULL	4.85%
WI	Aetna Health Inc.	3.77%
WI	Dean Health Plan, Inc.	3.61%
WI	Senior Preferred	3.23%
WI	BlueCrossBlueShield	3.22%
WI	iCare	1.58%
WI	Dean Advantage, Prevea360 Medicare Advantage	1.08%
WI	Medical Associates Clinic Health Plan of Wisconsin	0.71%
WI	HealthPartners	0.47%
WI	Care Wisconsin Health Plan, Inc.	0.41%
WI	Allwell	0.26%
WI	Community Care	0.19%

WI	Molina Healthcare of Wisconsin	0.19%
WI	UCare	0.04%

Part 2

SQL

1. Data cleaning and matching

```

delete from Major_Org_ana;
insert into Major_Org_ana
SELECT m.ContractNumber,
       m.PlanID,
       m.org_name,
       m.Enrollment_all,
       m.major_org,
       m.major_org_revised,
       s.state,
       s.state_enrollment
FROM (-- -enrollment segment
      SELECT t.ContractNumber,
             t.PlanID,
             t.State,
             sum(cast(t.Enrollment as real)) state_enrollment
      FROM cpsc_enrollment_info_2019_10 t
      WHERE t.Enrollment <> '*' AND
            t.State IN (
              SELECT DISTINCT state_code
              FROM state_submit
              WHERE team_number = 2 ) AND
            t.ContractNumber NOT LIKE 'S%'
      GROUP BY t.ContractNumber,
               t.PlanID,
               t.State ) s
LEFT JOIN
( --Org plans
  SELECT p.*,
         t2.MajorInsuranceOrgName Major_Org,
         CASE WHEN t2.MajorInsuranceOrgName IS NOT NULL THEN
t2.MajorInsuranceOrgName ELSE t2.OrganizationMarketingName END Major_org_revised
  FROM (
    SELECT t1.ContractNumber,
           t1.PlanID,
           upper([replace](t1.OrganizationMarketingName, ',', '')) ) Org_name,
           sum(CAST ([replace](t1.Enrollment, ',', '')) AS REAL) ) Enrollment_all
  FROM monthly_report_by_plan_2019_10 t1
  WHERE t1.Enrollment <> '*' AND
        t1.ContractNumber NOT LIKE 'S%'
  )

```



```

        GROUP BY t1.ContractNumber,
               t1.PlanID,
               upper(t1.OrganizationMarketingName) ) p
    LEFT JOIN
        major_insurance_orgs t2 ON p.org_name = upper(t2.OrganizationMarketingName)
    ) m ON m.ContractNumber = s.ContractNumber AND
        m.PlanID = s.PlanID
WHERE s.State IS NOT NULL

```

2. Compute the market share of insurance companies

```
delete from org_mkt_share;
```

```
insert into org_mkt_share
```

```

SELECT t1.*,
       s.tot_state_enroll,
       round(cast(t1.state_enrollment / s.tot_state_enroll as real),4) mkt_share
FROM (
    SELECT t.Major_org_revised,
           t.State,
           sum(cast(t.state_enrollment as real)) state_enrollment
    FROM Major_Org_ana t
    GROUP BY t.Major_org_revised,
             t.State
    ) t1
LEFT JOIN
(
    SELECT t.State,
           sum(cast(t.state_enrollment as real)) tot_state_enroll
    FROM Major_Org_ana t
    GROUP BY t.State
    ) s ON t1.state = s.state

```

3. Extract the top 5 Major Companies for each state

```
delete from top_5_major_org_list;
```

```
insert into top_5_major_org_list
```

```

select * from (SELECT *
FROM org_mkt_share o
WHERE o.Major_org_revised IN (
    SELECT major_org_revised
    FROM org_mkt_share i
    WHERE i.state = o.state
    and i.state not in('UT', 'NE', 'WI')
    ORDER BY mkt_share DESC
    LIMIT 5)

```

```
union all
```

```

SELECT *
FROM org_mkt_share o
WHERE o.Major_org_revised IN (
    SELECT major_org_revised
    FROM org_mkt_share i
    WHERE i.state = o.state
    and i.state in('UT', 'NE', 'WI')

```

```

ORDER BY mkt_share DESC
LIMIT 6
))ORDER BY State, mkt_share DESC, Major_org_revised
4. Match the dental plan data and export csv to R
delete from Top5_Major_State_Dental_Summary;
insert into Top5_Major_State_Dental_Summary
SELECT dent.*,
       b.tot_state_enroll,
       b.mkt_share,
       round(dent.prev_major_state / dent.major_state, 4) prev_major_state_pct,
       round(dent.comp_major_state / dent.major_state, 4) comp_major_state_pct
FROM (
  SELECT a.major_org_revised,
         a.state,
         sum(CAST (a.state_enrollment AS REAL) ) major_state,
         sum(CASE WHEN c.prev_dental == '1' THEN a.state_enrollment ELSE 0 END)
prev_major_state,
         sum(CASE WHEN c.Comp_dental == '1' THEN a.state_enrollment ELSE 0 END)
comp_major_state
  FROM Major_Org_ana a
  LEFT JOIN-- -min segment reps
  ( SELECT t.pbp_a_hnumber,
        t.pbp_a_plan_identfier,
        t.segment_id,
        t.pbp_b16a_bendesc_yn Prev_dental,
        t.pbp_b16b_bendesc_yn Comp_dental
  FROM pbp_b16_dental t
  WHERE (pbp_a_hnumber, pbp_a_plan_identfier, segment_id) IN (
        SELECT t.pbp_a_hnumber,
        t.pbp_a_plan_identfier,
        min(t.segment_id) segment_id
        FROM pbp_b16_dental t
        GROUP BY t.pbp_a_hnumber,
        t.pbp_a_plan_identfier)
  )
  c ON a.contractnumber = c.pbp_a_hnumber AND
  a.planid = c.pbp_a_plan_identfier
GROUP BY a.major_org_revised,a.state
) dent
JOIN (
  SELECT *
  FROM top_5_major_org_list
) b ON dent.state = b.state AND
dent.Major_org_revised = b.Major_org_revised;

R
setwd("E:/MBA@Brandeis/Syllabus/193HS-256F Healthcare Data Analytics and Data Mining/HW2")
getwd()

library(dplyr)

```

```

library(ggplot2)
library(reshape2)
library(scales)

dent_data <- read.csv(file.choose(), header = T)
summary(dent_data)
## Visualize enrollment of each state
enroll <- dent_data %>% select(state, tot_state_enroll) %>%
  distinct()
enroll$state <- factor(enroll$state, levels = enroll$state[order(enroll$tot_state_enroll)])
enroll %>%
  ggplot(aes(x=state, y=tot_state_enroll, fill = state)) +
  geom_col() +
  theme_minimal()+
  theme(legend.position="none")+
  geom_text(aes(label = tot_state_enroll),
            size = 4,
            hjust = 0.5,
            vjust = 1,
            position = "stack") +
  coord_flip()+
  ggtitle("Total enrollment in each State in Oct 2019") +
  theme(plot.title = element_text(hjust = 0.5))
## Top 5 Major Insurance Company each State
dent_data %>%
  select(state, major_org_revised, major_state, mkt_share) %>%
  distinct() %>% ggplot(aes(fill= major_org_revised, y=major_state, x= state))+
  geom_bar(position="dodge", width=1, stat="identity") +
  geom_text(aes(y=major_state, x= state,
                label= percent(mkt_share)),
            vjust= 0.5, hjust = 0, size = 3,
            position = position_dodge(width=1))+
  theme_minimal()+
  coord_flip()+
  scale_y_continuous(breaks = seq(0, 620000, 100000), name = 'Enrollment (Market share %)')+
  guides(fill=guide_legend(title='Top5 Insurance Companies')) +
  ggtitle("Top5 Major Insurance companies analysis of each State") +
  theme(plot.title = element_text(hjust = 0.5))

## Dental Coverage in Top 5 Major Insurance Company each State
pairs(dent_data[, 7:9])
dent_data %>%
  select(state, major_org_revised, prev_major_state_pct, comp_major_state_pct) %>%
  distinct() %>%
  ggplot(aes(color= major_org_revised,
            y=prev_major_state_pct,
            x= comp_major_state_pct))+
  geom_point(size = 2 ) +
  geom_text(label = substr(dent_data$major_org_revised, 1, 3),

```

```

size = 3,vjust= -.2, hjust = 0.5)+
scale_y_continuous(name = 'Preventive coverage rate') +
scale_x_continuous(name = 'Comprehensive coverage rate') +
ggtitle("Dental Coverage of Top5 Major Insurance companies each State") +
theme(plot.title = element_text(hjust = 0.5))+
guides(fill=guide_legend(title='Top5 Insurance Companies')) +
facet_wrap(~state)

```

Part3:

Import data to Mysql:

```

import pymysql
import pandas as pd
from sqlalchemy import create_engine

db = pymysql.connect(host='localhost',
                    user='root',
                    password='0707',
                    db='healthcare2',
                    charset='utf8mb4',
                    cursorclass=pymysql.cursors.DictCursor)
engine = create_engine('mysql+pymysql://root:0707@localhost/healthcare2')
enrollment = pd.read_csv("CPSC_Enrollment_Info_2019_10.csv")
pd.io.sql.to_sql(enrollment, 'enrollment', con=engine, index=False, if_exists='append')
uod= pd.read_csv("HEDIS2018.csv")
pd.io.sql.to_sql(uod, 'uod', con=engine, index=False, if_exists='append')
market=pd.read_csv("top 10 by state .csv")
pd.io.sql.to_sql(market, 'market', con=engine, index=False, if_exists='append')
contract = pd.read_csv("Monthly_Report_By_Contract_2019_10.csv")
pd.io.sql.to_sql(contract, 'contract', con=engine, index=False, if_exists='append')

```

SQL analysis:

```
use healthcare2;
```

```

drop table if exists table1;
create table table1
select e.`state`, c.`Organization Marketing Name`, u.`Contract Number`, e.`Enrollment`, u.`EOC170-0010`
from ((enrollment e inner join contract c on e.`Contract Number` = c.`Contract Number`)
     inner join uod u on e.`Contract Number` = u.`Contract Number`)
     inner join market m on (m.state=e.`state` AND m.MajorInsuranceOrgName= c.`Organization
Marketing Name`);
select * from table1;

```

```

drop table if exists tbl2;
create table tbl2
select * from table1 where Enrollment != '*' and `EOC170-0010` != 0.00 and `EOC170-0010` != 'NA' and
`EOC170-0010` != 'BR' and `EOC170-0010` != 'NB' and `EOC170-0010` != 'NQ';

```

```
select * from tbl2;
```

```
drop table if exists tbl3;
```

```
create table tbl3
```

```
select state, `Organization Marketing Name`, sum(Enrollment*`EOC170-0010`) over(partition by  
state, `Organization Marketing Name`) weighted, sum(Enrollment) over(partition by state, `Organization  
Marketing Name`) total_enrollment
```

```
from tbl2
```

```
group by state, `Organization Marketing Name`;
```

```
select *, ROW_NUMBER() OVER(PARTITION BY state
```

```
ORDER BY weighted_average desc)
```

```
from (select *, weighted/total_enrollment weighted_average from tbl3) as A;
```