

Assignment 1: Relational Database Normalization

김이경 (2022-28616)

In this assignment, I used medical data, which includes information of hospitals and pharmacies in South Korea. The data is downloaded from Open Government Data Portal, and data preprocessing is performed on the raw data before data normalization. There are 4 different tables used for this assignment, and each of them was normalized in order to reduce data redundancy and improve data integrity. Then, some interesting questions were answered by utilizing SQL on the table.

1. Raw Data

“건강보험심사평가원_전국 병의원 및 약국현황_20211231”

Resource: <https://www.data.go.kr/data/15051059/fileData.do>

- “1. 병원정보서비스_2021.12” (Renamed as “1_medical_service” afterwards)
- “2. 약국정보서비스_2021.12” (Renamed as “2_pharmacy_service” afterwards)
- “3. 의료기관별상세정보서비스 01시설정보 2021.12” (Renamed as “3_medical_facility_info” afterwards)
- “10. 의료기관별상세정보서비스 08특수진료정보 2021.12” (Renamed as “4_special_case_info” afterwards)

2. Data Preprocessing on Each Table

1_medical_service (Original name: 1. 병원정보서비스_2021.12)

A. Changed Types of Variables

- 암호화요약기호: VARCHAR(128) -> VARCHAR(1024)
- 요약기관명: VARCHAR(32) -> VARCHAR(128)
- 종별코드명: VARCHAR(4) -> VARCHAR(8)
- 개설일자: VARCHAR(16) -> date

B. Deleted Unused Columns

: 주소, 읍면동, 우편번호, 전화번호, 병원url, X좌표, Y좌표

	ncs 암호화요약기호	ncs 요약기관명	T	123 종별코드	T	ncs 종별코드명	T	122 시도코드	T	ncs 시도코드명	T	123 시군구코드	T	ncs 시군구코드명	T	125 개설일자	T	122 총의사수	T	122 의과일반의 인원수	T	
1	JDQ4MTYyMm4MSMI	경북대학교병원		1	상급종합		230,000	대구	230,006	대구중구	1910-09-07	481	5									
2	JDQ4MTYyMm4MSMI	계명대학교동산병원		1	상급종합		230,000	대구	230,007	대구달서구	1968-04-06	416	1									
3	JDQ4MTg4MSM1MSM	경희대학교병원		1	상급종합		110,000	서울	110,007	동대문구	1971-10-05	456	3									
4	JDQ4MTg4MSM1MSM	한경대학교병원		1	상급종합		110,000	서울	110,011	성동구	1972-04-24	402	3									
5	JDQ4MTg4MSM1MSM	고려중앙학원 고려대학교의과대학부속병원(안암병원)		1	상급종합		110,000	서울	110,012	성북구	1974-03-25	550	7									
6	JDQ4MTg4MSM1MSM	학교법인연세대학교의과대학서브란스병원		1	상급종합		110,000	서울	110,010	서대문구	1974-03-25	1,268	8									
7	JDQ4MTYyMm1MSMk	연세대학교 원주세브란스기독병원		1	상급종합		320,000	강원	320,400	원주시	1977-05-01	429	16									
8	JDQ4MTAxMm1MSMk	인제대학교 원주세브란스기독병원		1	상급종합		110,000	서울	110,016	종로구	1979-03-24	394	4									
9	JDQ4MTAxMm1MSMk	인제대학교부산백병원		1	상급종합		210,000	부산	210,004	부산진구	1979-06-01	369	2									
10	JDQ4MTg4MSM1MSM	학교법인가톨릭학원가톨릭대학교서울성모병원		1	상급종합		110,000	서울	110,021	서초구	1980-05-02	856										
	123 의과인턴 인원수	T	123 의사레지던트 인원수	T	123 의과전문의 인원수	T	123 치과일반의 인원수	T	123 치과인턴 인원수	T	123 치과레지던트 인원수	T	123 치과전문의 인원수	T	123 한방일반의 인원수	T						
	65		193		217		0		0		0		1		0							
	35		138		237		3		0		0		2		0							
	65		162		226		0		0		0		0		0							
	46		138		202		0		2		4		7		0							
	36		159		326		0		4		10		8		0							
	117		292		851		0		0		0		0		0							
	40		119		225		0		6		13		10		0							
	26		131		226		3		0		0		4		0							
	36		116		210		1		0		0		4		0							
	55		225		543		0		5		15		11		0							

123 한방인턴 인원수	123 한방레지던트 인원수	123 한방전문의 인원수
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0

2_pharmacy_service (Original name: 2. 약국정보서비스_2021.12)

A. Changes Types for Variables

- 암호화요양기호: VARCHAR(128) -> VARCHAR(1024)
- 요양기관명: VARCHAR(16) -> VARCHAR(64)

B. Deleted Unused Columns

: 주소, 읍면동, 우편번호, 전화번호, X좌표, Y좌표

	asc 암호화요양기호	asc 요양기관명	123 종별코드	asc 종별코드명	123 시도코드	asc 시도코드명	123 시군구코드	asc 시군구코드명	개설일자
1	JDQ4MTAxMiM1MS	우리들약국	81	약국	210,000	부산	210,010	부산사하구	2016-06-01
2	JDQ4MTAxMiM1MS	현대온누리약국	81	약국	210,000	부산	210,005	부산북구	2007-04-17
3	JDQ4MTAxMiM1MS	길약국	81	약국	210,000	부산	210,100	부산기장군	1991-08-28
4	JDQ4MTAxMiM1MS	정한약국	81	약국	210,000	부산	210,100	부산기장군	2001-09-05
5	JDQ4MTAxMiM1MS	건강약국	81	약국	210,000	부산	210,100	부산기장군	2001-07-02
6	JDQ4MTAxMiM1MS	동양약국	81	약국	210,000	부산	210,009	부산해운대구	2012-05-04
7	JDQ4MTAxMiM1MS	녹십자약국	81	약국	210,000	부산	210,015	부산사상구	2007-08-01
8	JDQ4MTAxMiM1MS	미소약국	81	약국	210,000	부산	210,010	부산사하구	1994-10-11
9	JDQ4MTAxMiM1MS	신의약국	81	약국	210,000	부산	210,010	부산사하구	1995-01-21
10	JDQ4MTAxMiM1MS	중앙약국	81	약국	210,000	부산	210,100	부산기장군	1990-03-07

3_medical_facility_info (Original name: 3. 의료기관별상세정보서비스_01시설정보_2021.12)

A. Changes Types for Variables

- 암호화요양기호: VARCHAR(128) -> VARCHAR(1024)
- 요양기관명: VARCHAR(32) -> VARCHAR(128)
- 종별코드명: VARCHAR(4) -> VARCHAR(8)
- 설립구분코드명: VARCHAR(8) -> VARCHAR(16)
- 개설일자: VARCHAR(16) -> date

B. Deleted Unused Columns

: 주소, 읍면동, 우편번호, 전화번호, 병원url

	asc 암호화요양기호	asc 요양기관명	123 종별코드	asc 종별코드명	123 설립구분코드	asc 설립구분코드명	123 시도코드	asc 시도코드명	123 시군구코드	asc 시군구코드명	개설일자
1	JDQ4MTg4MSM1MSMkM	서울대학교병원	1	상급종합	3	공립	110,000	서울	110,016	종로구	1981-07-09
2	JDQ4MTAxMiM1MSMKMS	부산대학교병원	1	상급종합	3	공립	210,000	부산	210,006	부산서구	1981-08-27
3	JDQ4MTTyMiM4MSMKMS	칠곡경북대학교병원	1	상급종합	3	공립	230,000	대구	230,003	미구북구	2010-10-29
4	JDQ4MTTyMiM4MSMKMS	경북대학교병원	1	상급종합	3	공립	230,000	대구	230,006	미구중구	1910-09-07
5	JDQ4MTTyMiM3MSMKMS	전남대학교병원	1	상급종합	3	공립	240,000	광주	240,001	광주동구	1983-04-15
6	JDQ4MTTyMiM2MSMKMS	충남대학교병원	1	상급종합	3	공립	250,000	대전	250,005	대전중구	1984-11-05
7	JDQ4MTTyMiM1MSMKMS	분당서울대학교병원	1	상급종합	3	공립	310,000	경기	310,403	성남분당구	2003-04-15
8	JDQ4MTTyMiM1MSMKMS	충북대학교병원	1	상급종합	3	공립	330,000	충북	330,104	청주서원구	1991-05-31
9	JDQ4MTTyMiM2MSMKMS	전북대학교병원	1	상급종합	3	공립	350,000	전북	350,402	전주덕진구	1983-01-28
10	JDQ4MTTyMiM3MSMKMS	화순전남대학교병원	1	상급종합	3	공립	360,000	전남	360,022	화순군	2004-04-06
	123 일반입원실상급병상수	123 일반입원실일반병상수	123 성인증환자병상수	123 소아증환자병상수	123 신생아증환자병상수	123 정신과폐쇄상급병상수	123 정신과폐쇄일반병상수	123 격리병상병상수	123 무균치료실병상수	123 분만실병상수	
	117	1,391	122	24	43	2	39	11	22	3	
	65	940	105	0	20	1	18	28	12	12	
	32	588	18	8	31	0	0	288	0	4	
	44	733	88	0	13	0	24	8	2	2	
	30	850	108	8	45	0	30	7	0	8	
	65	1,060	85	5	34	0	26	14	19	5	
	57	1,076	85	0	40	1	14	18	18	4	
	36	608	62	0	25	1	18	36	4	2	
	18	993	73	10	25	0	0	8	12	3	
	33	601	32	0	1	0	0	5	12	1	

수술실병상수	응급실병상수	물리치료실병상수
44	48	16
26	37	21
34	31	15
24	34	15
23	37	7
30	40	63
39	42	13
17	49	22
21	48	32
13	24	10

4_special_case_info (Original name: 10. 의료기관별상세정보서비스 08특수진료정보 2021.12)

A. Changes Types for Variables

- 암호화요양기호: VARCHAR(128) -> VARCHAR(1024)
- 요양기관명: VARCHAR(32) -> VARCHAR(128)
- 검색코드: Integer -> VARCHAR(32)

	암호화요양기호	요양기관명	검색코드	검색코드명
1	JDQ4MTAxMiM1MSMkMSMkMCMkMDMkMzgxNzAyIzIxlyQxlyQ	제이드약국	99	소아야간진료(20시 이후)
2	JDQ4MTAxMiM1MSMkMSMkMCMkMTMkMzgxMTkxIzExlyQxlyQ	재단법인 자생의료재단 해운대자생한방병원	53	한방척추질환
3	JDQ4MTAxMiM1MSMkMSMkMCMkMTMkMzgxMTkxIzExlyQxlyQ	대연더존한방병원	99	소아야간진료(20시 이후)
4	JDQ4MTAxMiM1MSMkMSMkMCMkMTMkMzgxNzAyIzMxlyQxlyQ	광혜한의원	99	소아야간진료(20시 이후)
5	JDQ4MTAxMiM1MSMkMSMkMCMkMTMkNDgxMzUxIzExlyQxlyQ	물편안한의원	99	소아야간진료(20시 이후)
6	JDQ4MTAxMiM1MSMkMSMkMCMkMTMkNDgxMzUxIzIxlyQxlyQ	하늘마음한의원	99	소아야간진료(20시 이후)
7	JDQ4MTAxMiM1MSMkMSMkMCMkMTMkNDgxMzUxIzMxlyQxlyQ	비른몸청담한의원	99	소아야간진료(20시 이후)
8	JDQ4MTAxMiM1MSMkMSMkMCMkMTMkNDgxOTyIzExlyQxlyQxlyQ	빙빙한의원	99	소아야간진료(20시 이후)
9	JDQ4MTAxMiM1MSMkMSMkMCMkNjIkMzgxMzUxIzExlyQxlyQz	우리들40플란트과병원	17	축두하악관절 자극요법
10	JDQ4MTAxMiM1MSMkMSMkMCMkNjIkMzgxMzUxIzExlyQxlyQz	굿윌치과병원	17	축두하악관절 자극요법

3. Data Normalization

3-1. 1_medical_service

- Primary Key (Candidate Key): {"암호화요양기호"}
- 1NF: Since each column of a table has a single value, this table satisfies the first normal form
- 2NF: In this table, all of the attributes that are not part of the candidate key {"암호화요양기호"} depends on the candidate key. Since there is not any partial-dependency and non-candidate keys depend on the whole candidate key, this table is in the second normal form.
- 3NF: There are functional dependencies among the non-key attributes, so the table is normalized to meet the third normal form.
 - “종별코드” is dependent on “종별코드명”
 - “시도코드” is dependent on “시도코드명”
 - “시군구코드” is dependent on “시군구코드명”
- After Normalization:

(Table 1) Medical Center Info

SQL:

```
select 암호화요양기호, 요양기관명, 종별코드명, 시도코드명, 시군구코드명, 개설일자, 총의사수, ms."의과일반의 인원수", ms."의과인턴 인원수", ms."의과레지던트 인원수", ms."의과전문의 인원수", ms."치과일반의 인원수", ms."치과인턴 인원수", ms."치과레지던트 인원수", ms."치과전문의 인원수", ms."한방일반의 인원수", ms."한방인턴 인원수", ms."한방레지던트 인원수", ms."한방전문의 인원수"
```

```
from "1_medical_service" ms
```

	암호화요양기호	요양기관명	종별코드명	시도코드명	시군구코드명	개설일자	총의사수	의과일반의 인원수	의과인턴 인원수	의과레지던트 인원수	의과전문의 인원수
1	JDQ4MTTyMmM4MSMK	경북대학교병원	상급종합	대구	대구중구	1910-09-07	481	5	65	193	217
2	JDQ4MTTyMmM4MSMK	경경대학교동산병원	상급종합	대구	대구달서구	1968-04-06	416	1	35	138	237
3	JDQ4MTg4MSMTMSMI	경희대학교병원	상급종합	서울	동대문구	1971-10-05	456	3	65	162	226
4	JDQ4MTg4MSMTMSMI	한양대학교교병원	상급종합	서울	성동구	1972-04-24	402	3	46	138	202
5	JDQ4MTg4MSMTMSMI	학교법인 고려중앙학원 고려대학교	상급종합	서울	성북구	1974-03-25	550	7	36	159	326
6	JDQ4MTg4MSMTMSMI	학교법인 민족대학교의과대학서브	상급종합	서울	서대문구	1974-03-25	1,268	8	117	292	851
7	JDQ4MTTyMmM4MSMK	연세대학교 원주세브란스기독병원	상급종합	강원	원주시	1977-05-01	429	16	40	119	225
8	JDQ4MTg4MSMTMSMI	강북삼성병원	상급종합	서울	종로구	1979-03-24	394	4	26	131	226
9	JDQ4MTAxMmM4MSMK	인제대학교부산백병원	상급종합	부산	부산진구	1979-06-01	369	2	36	116	210
10	JDQ4MTg4MSMTMSMI	학교법인 가톨릭학원가톨릭대학교	상급종합	서울	서초구	1980-05-02	856	2	55	225	543
	치과일반의 인원수	치과인턴 인원수	치과레지던트 인원수			치과전문의 인원수		한방일반의 인원수	한방인턴 인원수	한방레지던트 인원수	한방전문의 인원수
0	0	0	0	1	0	0	0	0	0	0	0
3	0	0	0	2	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	2	4	7	0	0	0	0	0	0	0	0
0	4	10	8	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0
0	6	13	10	0	0	0	0	0	0	0	0
3	0	0	4	0	0	0	0	0	0	0	0
1	0	0	4	0	0	0	0	0	0	0	0
0	5	15	11	0	0	0	0	0	0	0	0

(only 10 rows are shown as above for reference)

(Table 2) Medical Center TypeSQL:`select distinct 종별코드명, 종별코드 from "1_medical_service" ms`

	asc 종별코드명 ↴	123 종별코드 ↵
1	정신병원	29
2	치과의원	51
3	보건소	71
4	종합병원	11
5	보건진료소	73
6	의원	31
7	한의원	93
8	치과병원	41
9	한방병원	92
10	조산원	61

(only 10 rows are shown as above for reference)

(Table 3) CitySQL:`select distinct 시도코드명, 시도코드 from "1_medical_service" ms`

	asc 시도코드명 ↴	123 시도코드 ↵
1	대구	230,000
2	전남	360,000
3	광주	240,000
4	부산	210,000
5	경기	310,000
6	인천	220,000
7	충남	340,000
8	전북	350,000
9	세종시	410,000
10	강원	320,000

(only 10 rows are shown as above for reference)

(Table 4) County (only 10 rows are captured in below attachment)SQL:`select distinct 시군구코드명, 시군구코드 from "1_medical_service" ms`

	asc 시군구코드명 ↴	123 시군구코드 ↵
1	정선군	320,009
2	홍성군	340,015
3	수원팔달구	310,603
4	군포시	311,400
5	강진군	360,001
6	강서구	110,003
7	고양일산서구	311,902
8	수원권선구	310,601
9	양양군	320,005
10	증평군	330,011

(only 10 rows are shown as above for reference)

3-2. 2_pharmacy_service

- Primary Key (Candidate Key): {"암호화요양기호"}
- 1NF: Since each column of a table has a single value, this table satisfies the first normal form
- 2NF: In this table, all of the attributes that are not part of the candidate key {"암호화요양기호"} depends on the candidate key. Since there is not any partial-dependency and non-candidate keys depend on the whole candidate key, this table is in the second normal form.
- 3NF: There are functional dependencies among the non-key attributes, so the table is normalized to meet the third normal form.
 - "종별코드" is dependent on "종별코드명"
 - "시도코드" is dependent on "시도코드명" (*Same as 3-1 normalized table*)
 - "시군구코드" is dependent on "시군구코드명" (*Same as 3-1 normalized table*)
- After Normalization:

(Table 5) Pharmacy Info

SQL:

`select ps.암호화요양기호, ps.요양기관명, ps.종별코드명, ps.시도코드명, ps.시군구코드명`

`from "2_pharmacy_service" ps`

	암호화요양기호	요양기관명	종별코드명	시도코드명	시군구코드명
1	JDQ4MTYyMiM2MSM	도안참사랑약국	약국	대전	대전유성구
2	JDQ4MTYyMiM2MSM	동아약국	약국	충남	보령시
3	JDQ4MTYyMiM2MSM	신성태평양약국	약국	대전	대전유성구
4	JDQ4MTYyMiM2MSM	신세계약국	약국	충남	청양군
5	JDQ4MTYyMiM2MSM	누리약국	약국	충남	논산시
6	JDQ4MTYyMiM2MSM	내포우리약국	약국	충남	홍성군
7	JDQ4MTYyMiM2MSM	서문약국	약국	충남	서산시
8	JDQ4MTYyMiM2MSM	엘약국	약국	세종시	세종시
9	JDQ4MTYyMiM2MSM	신중앙약국	약국	충남	홍성군
10	JDQ4MTYyMiM2MSM	미르온누리약국	약국	충남	천안동남구

(only 10 rows are shown as above for reference)

(Table 6) Pharmacy Type

SQL:

`select distinct ps.종별코드, ps.종별코드명 from "2_pharmacy_service" ps`

	종별코드	종별코드명
1	81	약국

(only 10 rows are shown as above for reference)

3-3. 3_medical_facility_info

- Primary Key (Candidate Key): {"암호화요양기호"}
 - 1NF: Since each column of a table has a single value, this table satisfies the first normal form
 - 2NF: In this table, all of the attributes that are not part of the candidate key {"암호화요양기호"} depends on the candidate key. Since there is not any partial-dependency and non-candidate keys depend on the whole candidate key, this table is in the second normal form.
 - 3NF: There are functional dependencies among the non-key attributes, so the table is normalized to meet the third normal form.
 - “종별코드” is dependent on “종별코드명” (*Same as 3-1 normalized table*)
 - “설립구분코드” is dependent on “설립구분코드명”
 - “시도코드” is dependent on “시도코드명” (*Same as 3-1 normalized table*)
 - “시군구코드” is dependent on “시군구코드명” (*Same as 3-1 normalized table*)

(Table 7) Facility Info

sol:

select mfi.암호화요양기호, mfi.요양기관명, mfi.종별코드명, mfi.설립구분코드명, mfi.시도코드명, mfi.시군구코드명, mfi.개설일자, mfi.일반입원실상급병상수, mfi.일반입원실일반병상수, mfi.성인중환자병상수, mfi.소아중환자병상수, mfi.신생아중환자병상수, mfi.정신과폐쇄상급병상수, mfi.정신과폐쇄일반병상수, mfi.격리병실병상수, mfi.무균치료실병상수, mfi.분만실병상수, mfi.수술실병상수, mfi.응급실병상수, mfi.물리치료실병상수

from "3 medical facility info" mfi

idx	임화호요약기호	T1	ssk 요약기판명	T1	ssk 총별도코드명	T1	ssk 설립구분코드명	T1	ssk 시도코드명	T1	ssk 시군구코드명	T1	개설일자	T1	12_일반인원실상급병상수	T1	12_일반입원실일반병상수	T1	12_성인중환자병상수
1	JJD4MTG4MSM1MSM	서울대학교병원	상급종합	衿편	서울	종로구	1981-07-09	117	1,391									122	
2	JJD4MTAxM1M1MSM	부산대학교병원	상급종합	공립	부산	부산서구	1981-08-27	65	940									105	
3	JJD4MTY7M1M4MSM	불곡경북대학교병원	상급종합	광립	대구	대구북구	2010-10-29	32	588									18	
4	JJD4MTY7M1M4MSM	경북대학교병원	상급종합	광립	대구	대구중구	1910-09-07	44	733									88	
5	JJD4MTY7M1M3MSM	전주대학교병원	상급종합	衿편	광주	광주동구	1983-04-15	30	850									108	
6	JJD4MTY7M1M2MSM	충북대학교병원	상급종합	衿편	대전	대전중구	1984-11-05	65	1,060									85	
7	JJD4MTY7M1M1MSM	부산대학병원	상급종합	衿립	경기	성남분당구	2003-04-15	57	1,076									95	
8	JJD4MTY7M1M1MSM	충북대학교병원	상급종합	衿편	충북	청주시서원구	1991-05-31	36	608									62	
9	JJD4MTY7M1M2MSM	전북대학교병원	상급종합	衿편	전북	진주지역진구	1983-01-28	18	993									73	
10	JJD4MTY7M1M3MSM	최순진남대학교병원	상급종합	衿립	전남	회현동	2004-04-06	33	601									32	
12	소아중환자병상수	T1	12_신생아중환자병상수	T1	12_정신과폐쇄상급병상수	T1	12_정신과폐쇄일반병상수	T1	12_격리병상병상수	T1	12_무균치료실병상수	T1	12_분만실병상수	T1	12_수술실병상수	T1	12_응급실병상수	T1	12_물리치료실병상수
	24		43		2		39		11		22		3		44		48		16
	0		20		1		18		28		12		12		26		37		21
	8		31		0		0		288		0		4		34		31		15
	0		13		0		24		8		2		2		24		34		15
	8		45		0		30		7		0		8		23		37		15
	5		34		0		26		14		19		5		30		40		63
	0		40		1		14		18		18		4		39		42		13
	0		25		1		18		36		4		2		17		49		22
	10		25		0		8		12		3		21		48		32		32

(only 10 rows are shown as above for reference)

(Table 8) Established Type

SQL:

select distinct mfi 설립구분코드, **mfi** 설립구분코드명

from "3 medical facility info" mfi order by mfi.설립구분코드

	123 설립구분코드	설립구분코드명
1	1	국립
2	3	공립
3	4	학교법인
4	5	특수법인
5	6	종교법인
6	7	사회복지법인
7	8	사단법인
8	9	재단법인
9	10	회사법인
-	11	국민 법인

(only 10 rows are shown as above for reference)

3-4. 4_special_case_info

- Primary Key (Candidate Key): {"암호화요양기호"}
- 1NF: Since each column of a table has a single value, this table satisfies the first normal form
- 2NF: In this table, all of the attributes that are not part of the candidate key {"암호화요양기호"} depends on the candidate key. Since there is not any partial-dependency and non-candidate keys depend on the whole candidate key, this table is in the second normal form.
- 3NF: There are functional dependencies among the non-key attributes, so the table is normalized to meet the third normal form.
 - "검색코드" is dependent on "검색코드명"
- After Normalization:

(Table 9) Special Info

SQL:

```
select sci.암호화요양기호, sci.요양기관명, sci.검색코드명 from "4_special_case_info" sci
```

	asc 암호화요양기호	asc 요양기관명	asc 검색코드명
1	JDQ4MTAxMi1MSMKMSMKMCMkMDMkMzgxNzAylzllyQxlyQxlyQ4OSQzNjEyMjjNDEjJDEjJDgjJDg	제이드약국	소아야간진료(20시 이후)
2	JDQ4MTAxMi1MSMKMSMKMCMkMTMkMzgxMTkxlzExlyQxlyQzlyQ2MiQzNjEyMjjNDEjJDEjJDgjJDg	재단법인 자생의료재단 해운대자생한방병원	한방척추질환
3	JDQ4MTAxMi1MSMKMSMKMCMkMTMkMzgxMTkxlzExlyQxlyQzlyQ3MiQ0NjEwMDljNDEjJDEjJDgjJDg	다면더존한방병원	소아야간진료(20시 이후)
4	JDQ4MTAxMi1MSMKMSMKMCMkMTMkMzgxNzAylzllyQxlyQzlyQ4OSQyNjEyMjjNTEjJDEjJDljjDg	광혜한의원	소아야간진료(20시 이후)
5	JDQ4MTAxMi1MSMKMSMKMCMkMTMkNDgxMzUxlzExlyQxlyQ3lyQ4MiQyNjEyMjjODEjJDEjJDYjJDg	몸편안한의원	소아야간진료(20시 이후)
6	JDQ4MTAxMi1MSMKMSMKMCMkMTMkNDgxMzUxlzExlyQxlyQ5lyQ5MiQ0NjEwMDljNTEjJDEjJDljjDg	하늘마음한의원	소아야간진료(20시 이후)
7	JDQ4MTAxMi1MSMKMSMKMCMkMTMkNDgxMzUxlzExlyQxlyQzlyQ3MiQyNjEwMDljNDEjJDEjJDYjJDg	비른몸청담한의원	소아야간진료(20시 이후)
8	JDQ4MTAxMi1MSMKMSMKMCMkMTMkNDgxOTYxlzExlyQxlyQzlyQ3MiQyNjE0ODEjNjEjJDEjJDAjJDg	빙빙한의원	소아야간진료(20시 이후)
9	JDQ4MTAxMi1MSMKMSMKMCMkNjikMzgxMzUxlzExlyQxlyQzlyQ2MiQyNjE4MzljNDEjJDEjJDgjJDg	우리들40플란트치과병원	축두하악관절 자극요법
10	JDQ4MTAxMi1MSMKMSMKMCMkNjikMzgxMzUxlzExlyQxlyQzlyQ3OSQzNjE4MzljNjEjJDEjJDgjJDg	굿월치과병원	축두하악관절 자극요법

(only 10 rows are shown as above for reference)

(Table 10) Search Type

SQL:

```
select distinct mfi.설립구분코드, mfi.설립구분코드명
from "3_medical_facility_info" mfi order by mfi.설립구분코드
```

	asc 검색코드	asc 검색코드명
1	02	제3차 의료급여기관
2	03	응급의료병원
3	04	성인·소아 중환자실
4	05	신생아 중환자실
5	06	개방병원
6	08	가정간호실시병원
7	09	보훈위탁병원
8	10	의료서비스지원사업기관
9	12	알코올질환입원치료병원
10	13	혈액투석

(only 10 rows are shown as above for reference)

4. Interesting Questions

**Q1. In which hospital, the workload for intern and resident would be the highest?
(Out of the hospitals where the number of medical specialists, resident, and intern is greater than 0 for each, which hospital has the lowest ratio of (intern + resident) to specialist?)**

- The interns and residents working at ‘의료법인 인화재단 한국병원’ has the highest level of workload, by having only two intern and residents working for 62 specialists.
- SQL:

```
select mci.요양기관명, mci."의과인턴 인원수", mci."의과레지던트 인원수", mci."의과전문의 인원수", ( 1.0 * (mci."의과인턴 인원수"+mci."의과레지던트 인원수"))/ mci."의과전문의 인원수 ) as ratio_workload
from medical_center_info mci
where mci."의과인턴 인원수" > 0 and mci."의과레지던트 인원수" > 0 and mci."의과전문의 인원수" > 0
order by ratio_workload asc
```

	abc 요양기관명	123 의과인턴 인원수	123 의과레지던트 인원수	123 의과전문의 인원수	123 ratio_workload
1	의료법인 인화재단 한국병원	1	1	62	0.0322580645
2	의료법인영훈의료재단 유성선병원	2	1	81	0.037037037
3	의료법인 동강의료재단 동강병원	1	4	91	0.0549450549
4	의료법인 온그룹의료재단 온종합병원	2	2	62	0.064516129
5	대동병원	3	2	65	0.0769230769
6	오산한국병원	2	3	58	0.0862068966
7	삼육부산병원	2	2	39	0.1025641026
8	의료법인 인당의료재단 부민병원	2	4	57	0.1052631579
9	의료법인 은성의료재단 좋은강안병원	5	4	84	0.1071428571
10	포항성모병원	4	6	89	0.1123595506

Q2. Nowadays, a lot of nursing hospitals are opening as senior population is growing. How many new nursing hospitals opened in last three years?

- There are 1,464 nursing hospitals in South Korea according to the following SQL code:

```
select count(distinct(mci.암호화요양기호)) from medical_center_info mci
where mci.종별코드명 = '요양병원'
```

- 143 nursing hospitals opened in last three years, and it is almost 10% of total number of nursing hospitals.

- SQL:

```
select count(distinct(t4."요양기관명",t4."개설일자"))
from(
select mci.요양기관명, mci.시도코드명, mci.종별코드명, mci.개설일자, (date_part('year',
'2021-12-31)::date)-date_part('year', mci.개설일자)) as no_years
from medical_center_info mci
where mci.종별코드명 = '요양병원'
) as t4
where t4.no_years < 3
```

Q3. In which county the average revenue of the pharmacy will be the greatest? (Let's assume the ratio of number of medical centers to pharmacy will be considered to expect the revenue)

- ‘인천옹진군’ is greatest in ratio of number of medical center to number of pharmacy.

So we expect high revenue for pharmacy in 인천옹진군

- SQL:

```
select t66."시도코드명",t66."시군구코드명", (no_medical / no_pharm) as ratio_richpharm
from (
    select mci.시도코드명 as 시도코드명, mci.시군구코드명 as 시군구코드명,
    count(distinct(mci.암호화요양기호)) as no_medical, t6.no_pharm as no_pharm
    from (
        select pi2."시군구코드명", count(distinct(pi2.암호화요양기호)) as no_pharm
        from pharmacy_info pi2
        group by pi2."시군구코드명"
    ) as t6
    join medical_center_info mci on t6."시군구코드명" = mci.시군구코드명
    group by mci.시도코드명, mci.시군구코드명, t6.no_pharm
    order by mci.시도코드명, mci.시군구코드명
) as t66
order by ratio_richpharm desc
```

	시도코드명	시군구코드명	ratio_richpharm
1	인천	인천옹진군	7
2	전남	신안군	6
3	서울	강남구	5
4	서울	서초구	5
5	경북	영양군	5
6	전북	순창군	4
7	부산	부산진구	4
8	경남	산청군	4
9	경남	합천군	4
10	전북	임실군	4

Q4. In emergency medical situation, which hospital provide the best service?

- ‘재단법인아산사회복지재단 서울아산병원’ provides the best service in emergency situation.
- I took the number of emergency room from facility info data and the number of medical specialists from the medical center info data. Then it is considered that the summation of the two number represents the emergency service.

- SQL:

```
select t88.name_hospital , (t88.no_specialist + t88.no_emergency) as no_emergency
from(
    select mci.요양기관명 as name_hospital, mci."의과전문의 인원수" as no_specialist
    ,t8.응급실병상수 as no_emergency
    from facility_info fi ) as t8
    join medical_center_info mci on mci.암호화요양기호 = t8.암호화요양기호
) as t88
```

`order by no_emergency desc`

▲	ABC name_hospital	T↑	123 no_emergency T↑
1	재단법인아산사회복지재단 서울아산병원		1,151
2	학교법인연세대학교의과대학세브란스병원		908
3	삼성서울병원		894
4	서울대학교병원		847
5	학교법인가톨릭학원가톨릭대학교서울성모병원		589
6	분당서울대학교병원		570
7	아주대학교병원		417
8	고려대학교의과대학부속구로병원		361
9	학교법인 고려중앙학원 고려대학교의과대학부속병원(인)		361
10	의료법인 길의료재단 길병원		361

Q5. Which county will be considered as a good county to have a child? The search type for '소아야간진료(20시이후)' is 99 according to the search type data.

➤ I took the search type from the special case info data, and then joined the result with medical info data to take the result by each county.

➤ '서울 강동구' has 12 hospitals which perform '소아야간진료', and it is the highest result.

➤ SQL:

```
select mci.시도코드명, mci.시군구코드명, count(distinct(mci.암호화요양기호)) as
no_kidsfriendly
from(
select sci.암호화요양기호 as 암호화요양기호, sci.요양기관명, sci.검색코드명
from "4_special_case_info" sci
where sci.검색코드 = '99'
) as t7
join medical_center_info mci on mci.암호화요양기호 = t7.암호화요양기호
group by mci.시도코드명, mci.시군구코드명
order by no_kidsfriendly desc
```

	ABC 시도코드명 T↑	ABC 시군구코드명 T↑	123 no_kidsfriendly T↑
1	서울	강동구	12
2	서울	노원구	12
3	경기	부천시	11
4	인천	인천남동구	10
5	대전	대전서구	10
6	경기	수원영통구	10
7	제주	제주시	10
8	경기	평택시	9
9	충북	청주흥덕구	9
10	서울	동대문구	9

Q6. How many big-sized hospital including university hospital and general hospital with more than 50 doctors located in each city?

➤ As it is shown in the below table, most of the big-sized hospitals are located in Seoul and Gyeonggi by comprising 75 out of 166 of all big-sized hospitals, which is nearly 45%.

➤ SQL:

```
select mci.시도코드명, count(distinct(mci.암호화요양기호)) as no_of_hos
from medical_center_info mci
```

`where mci.종별코드명 ='상급종합' or mci.종별코드명 = '종합병원' and mci.총의사수 >50
group by mci.시도코드명
order by count(mci.요양기관명) desc`

	enc 시도코드명	123 no_of_hos
1	서울	38
2	경기	37
3	부산	19
4	인천	10
5	경남	8
6	대구	7
7	경북	7
8	대전	7
9	충북	6
10	전북	6
11	충남	5
12	강원	4
13	광주	4
14	전남	3
15	제주	2
16	울산	2
17	세종시	1

Q7. Which hospital is the oldest hospital out of educational foundations and how many doctors working for the hospital currently?

- ‘가톨릭대학교 은평성모병원’ is the oldest educational foundational hospital in Korea. It has been 60 years since it is built, and there are 254 doctors currently working for the hospital.

- SQL:

```
select t9.설립구분코드명, mci.요양기관명, t9.no_years, mci.총의사수
from(
select fi.설립구분코드명 , fi.암호화요양기호, fi.요양기관명,fi.개설일자, (date_part('year',
'2021-12-31')::date)-date_part('year', fi.개설일자)) as no_years
from facility_info fi
where fi.설립구분코드명 = '학교법인'
) as t9
join medical_center_info mci on mci.암호화요양기호 = t9.암호화요양기호
order by no_years desc
```

	enc 설립구분코드명	enc 요양기관명	123 no_years	123 총의사수
1	학교법인	가톨릭대학교 은평성모병원	60	254
2	학교법인	가톨릭대학교 성빈센트병원	54	376
3	학교법인	계명대학교동산병원	53	416
4	학교법인	경희대학교한방병원	50	125
5	학교법인	가톨릭대학교의정부성모병원	50	271
6	학교법인	경희대학교병원	50	456
7	학교법인	경희대학교치과병원	50	146
8	학교법인	한양대학교병원	49	402
9	학교법인	학교법인 춘해병원	47	23
10	학교법인	학교법인연세대학교의과대학세브란스병원	47	1,268

Q8. Which county has the greatest number of hospital where the established type is personal?

- As we can imagine, '강남구' has the most number of personal hospital in Korea. The total sum of personal hospital in Korea is 68,769. And '강남구' takes 4% of the entire personal hospitals.

➤ SQL:

```
select mci.시도코드명, mci.시군구코드명, count(distinct(mci.암호화요양기호)) as
no_personal
from(
select fi.설립구분코드명, fi.암호화요양기호
from facility_info fi
where fi.설립구분코드명 ='개인') as t10
join medical_center_info mci on mci.암호화요양기호 =t10.암호화요양기호
group by mci.시도코드명, mci.시군구코드명
order by no_personal desc
```

	시도코드명	시군구코드명	no_personal
1	서울	강남구	2,756
2	서울	서초구	1,278
3	서울	송파구	1,198
4	경기	부천시	1,083
5	경기	성남분당구	934
6	대전	대전서구	872
7	서울	강서구	871
8	서울	강동구	838
9	대구	대구달서구	812
10	부산	부산진구	807

Q9. There are 15 different medical center types according to the normalized table. Which county has less than half of the total number of medical center types?

- There are 249 counties in South Korea according to the following SQL code:
- ```
select count(distinct(mci.시군구코드명, mci.시도코드명)) from medical_center_info
mci
```
- Interestingly, only 10 counties had less than 7 medical center type, and we could see that each county has at least 5 types of medical centers.

➤ SQL:

```
select * from
(select mci.시도코드명, mci.시군구코드명, count(distinct(mci.종별코드명)) as no_types
from medical_center_info mci
group by mci.시도코드명, mci.시군구코드명
order by mci.시도코드명, mci.시군구코드명) as t2
where t2.no_types < 7
```

|    | abc 시도코드명 | abc 시군구코드명 | 123 no_types |
|----|-----------|------------|--------------|
| 1  | 강원        | 고성군        | 6            |
| 2  | 강원        | 양양군        | 6            |
| 3  | 강원        | 화천군        | 6            |
| 4  | 경기        | 과천시        | 5            |
| 5  | 경북        | 군위군        | 6            |
| 6  | 경북        | 울릉군        | 6            |
| 7  | 인천        | 인천동구       | 5            |
| 8  | 인천        | 인천옹진군      | 6            |
| 9  | 전북        | 무주군        | 6            |
| 10 | 전북        | 장수군        | 6            |

**Q10. The medical center type code of 41 and 51 is for dental clinic. Which dental clinic is the oldest clinic? Use the nested query and normalized table for the answer.**

- According the following SQL query, the oldest dental clinic is '에스에프수정부부치과의원', which opened in 1900. However, it is quite not credible because the next oldest clinic opened in 1961. So it can be considered as a data error.

- SQL:

```
select distinct mci.요양기관명, mci.시도코드명, mci.종별코드명, t4."종별코드", mci.개설일자
from (select distinct ms.종별코드명, ms.종별코드
from "1_medical_service" ms
) as t4
join medical_center_info mci on t4."종별코드명" = mci.종별코드명
where t4."종별코드" = 41 or t4."종별코드" = 51
order by mci.개설일자 asc
```

|    | abc 요양기관명    | abc 시도코드명 | abc 종별코드명 | 123 종별코드 | 개설일자       |
|----|--------------|-----------|-----------|----------|------------|
| 1  | 에스에프수정부부치과의원 | 서울        | 치과의원      | 51       | 1900-01-01 |
| 2  | 이굿치과의원       | 대구        | 치과의원      | 51       | 1961-06-12 |
| 3  | 김봉호치과의원      | 부산        | 치과의원      | 51       | 1963-01-31 |
| 4  | 상산치과의원       | 경북        | 치과의원      | 51       | 1963-02-20 |
| 5  | 박치과의원        | 충남        | 치과의원      | 51       | 1963-09-29 |
| 6  | 선일치과의원       | 부산        | 치과의원      | 51       | 1965-03-31 |
| 7  | 마산치과의원       | 경남        | 치과의원      | 51       | 1965-08-17 |
| 8  | 참좋은치과의원      | 서울        | 치과의원      | 51       | 1965-08-30 |
| 9  | 남욱치과의원       | 대전        | 치과의원      | 51       | 1965-09-23 |
| 10 | 한국은행부속치과의원   | 서울        | 치과의원      | 51       | 1966-01-19 |