

집단면역 형성 국가 추론

- WHO 백신 접종 데이터를 바탕으로 -

2조

데이터사이언스학과

201804217 유제우, 201804235 이현중, 202004191 김소현

분석 목적 |

코로나 시국에 있어서, 현재 대한민국은 가장 백신 접종률이 높으며, 동시에 가장 확진자 수가 많아 전 세계에서 첫 번째로 전문가들이 코로나 바이러스의 풍토병화를 예상하는 상태에 직면했다.

따라서 이번 분석에서는 WHO 에 업데이트 되어있는 가장 최근 보고된 국가 별 백신 접종률에 관한 데이터(Vaccination Data)를 Json으로 변환한 뒤 분석하여, 코로나 바이러스가 풍토병화 되어가는 한국의 백신 접종률과, 다른 국가의 백신접종률을 비교하여 다음으로 코로나 바이러스에 대항하는 집단면역 형성에 유리한 국가들을 살펴볼 것이다.

데이터 살펴보기 |

Vaccination DB에는 다음과 같은 도큐먼트들이 있다.

분야 명	유형	설명
COUNTRY	String	국가, 영토, 지역
ISO3	String	ISO Alpha-3 국가 코드
WHO_REGION	String	WHO 지역 사무소: WHO 회원국은 6개의 WHO 지역으로 그룹화됨 아프리카 지역 사무소(AFRO) 아메리카 지역 사무소(AMRO) 동남아시아 지역 사무소(SEARO) 유럽 지역 사무소(EURO) 동부 지중해 지역 사무소(EMRO) 서태평양 지역 사무소(WPRO)
DATA_SOURCE	String	데이터 출처를 나타냄 [출처 종류] - report:: 회원국이 보고한 데이터 혹은 공식

		보고서 - OWID: Our World in Data에서 < https://ourworldindata.org/covid-vaccinations >
DATE_UPDATED	data	마지막 업데이트 날짜
TOTAL_VACCINATIONS	Integer	투여된 누적 총 백신 투여량
PERSONS_VACCINATED_1PLUS_DOSE	Decimal	1회 이상 예방접종을 받은 사람의 누적 수
TOTAL_VACCINATIONS_PER100	Integer	인구 100명당 투여된 누적 총 백신 접종량
PERSONS_VACCINATED_1PLUS_DOSE_PER100	Decimal	인구 100명당 1회 이상 접종한 누적인원
PERSONS_FULLY_VACCINATED	Integer	예방접종을 완료한 누적인원
PERSONS_FULLY_VACCINATED_PER100	Decimal	인구 100명당 예방접종을 완료한 누적인원
VACCINES_USED	String	백신의 약칭: "회사 - 제품명"
FIRST_VACCINE_DATE	Data	첫 번째 예방 접종 날짜 (한 국가에서 접종된 첫 번째 백신의 시작/출시 날짜와 동일)
NUMBER_VACCINES_TYPES_USED	Integer	국가, 지역, 지역별 사용되는 백신 종류 수
PERSONS_BOOSTER_ADD_DOSE	Integer	추가 접종 또는 부스터 샷을 맞은 사람
PERSONS_BOOSTER_ADD_DOSE_PER100	Decimal	인구 100명당 추가 접종 또는 부스터 샷을 맞은 사람

```
db.newCol.find().limit(3).pretty()

{
  "_id" : ObjectId("6250104e3ad92d6a725ba8ad"),
  "COUNTRY" : "American Samoa",
  "ISO3" : "ASM",
  "WHO_REGION" : "WPRO",
  "DATA_SOURCE" : "REPORTING",
  "DATE_UPDATED" : "2022-03-17",
  "TOTAL_VACCINATIONS" : 97346,
  "PERSONS_VACCINATED_1PLUS_DOSE" : 43637,
  "TOTAL_VACCINATIONS_PER100" : 176.361,
  "PERSONS_VACCINATED_1PLUS_DOSE_PER100" : 79.057,
  "PERSONS_FULLY_VACCINATED" : 39137,
  "PERSONS_FULLY_VACCINATED_PER100" : 70.904,
  "VACCINES_USED" : "Janssen - Ad26.COV 2-S,Moderna - Spikevax,Pfizer BioNTech - Comirnaty",
  "FIRST_VACCINE_DATE" : "2020-12-21",
  "NUMBER_VACCINES_TYPES_USED" : 3,
  "PERSONS_BOOSTER_ADD_DOSE" : 15135,
  "PERSONS_BOOSTER_ADD_DOSE_PER100" : 27.42
}
```

위와 같이 국가명, 해당 국가가 소속되어있는 WHO 지부, 데이터 출처, 데이터 업데이트 일자, 사용 백신 종류, 백신 접종 인구 및 백분율 등으로 이루어져 있다.

데이터 분석 |

```
> db.newCol.find({"ISO3":"KOR"}).pretty()
{
  "_id" : ObjectId("6270aec31da680094dc45924"),
  "COUNTRY" : "Republic of Korea",
  "ISO3" : "KOR",
  "WHO_REGION" : "WPRO",
  "DATA_SOURCE" : "REPORTING",
  "DATE_UPDATED" : "2022-04-04",
  "TOTAL_VACCINATIONS" : 120600744,
  "PERSONS_VACCINATED_1PLUS_DOSE" : 44894841,
  "TOTAL_VACCINATIONS_PER100" : 235.23,
  "PERSONS_VACCINATED_1PLUS_DOSE_PER100" : 87.567,
  "PERSONS_FULLY_VACCINATED" : 44421396,
  "PERSONS_FULLY_VACCINATED_PER100" : 86.643,
  "VACCINES_USED" : "AstraZeneca - Vaxzevria, Janssen - Ad26.COV 2-S, Moderna - Spikevax, Novavax-NUVAXOVID, Pfizer BioNTech - Comirnaty",
  "FIRST_VACCINE_DATE" : "2021-02-26",
  "NUMBER_VACCINES_TYPES_USED" : 5,
  "PERSONS_BOOSTER_ADD_DOSE" : 32800214,
  "PERSONS_BOOSTER_ADD_DOSE_PER100" : 63.976
}
```

먼저 우리나라(KOR)의 Collection을 살펴보면 인구의 87퍼센트가 1차 백신을 접종을 완료했으며, 인구의 63퍼센트가 2차 백신까지 접종 완료 한 것을 볼 수 있다. 전체 백신 접종률은 235.23퍼센트로, 이것은 3차 접종 백분율까지 포함한 것으로 보인다.

우리나라에서 사용한 백신은 AstraZeneca, Janssen, Moderna, Pfizer이며, 첫번째 백신 접종일은 2021년 2월 26일인 것을 볼 수 있다. 이와 같이 우리나라는 1차 접종률이 70퍼센트가 넘어서 집단면역 달성 기준을 초과하였으며, 거기에 2차, 3차 접종까지 수월하게 이루어졌다.

2차 접종에 대한 데이터는 많은 국가들에서 2차 접종까지 활성화하지 못한 탓에 NULL값으로 두고 있다. 따라서 2차 접종에 대한 결측치를 0으로 통일하여 처리했다.

```
db.newCol.updateMany({"PERSONS_BOOSTER_ADD_DOSE":null},{ $set:{"PERSONS_BOOSTER_ADD_DOSE":0}})
```

```
> db.vac.updateMany({PERSONS_BOOSTER_ADD_DOSE:null}, {$set:{PERSONS_BOOSTER_ADD_DOSE:0}})
{ "acknowledged" : true, "matchedCount" : 49, "modifiedCount" : 49 }
> db.vac.find({PERSONS_BOOSTER_ADD_DOSE:0}).pretty()
{
  "_id" : ObjectId("6256a3a45ad856e2f5e90815"),
  "COUNTRY" : "Afghanistan",
  "ISO3" : "AFG",
  "WHO_REGION" : "EMRO",
  "DATA_SOURCE" : "REPORTING",
  "DATE_UPDATED" : "2022-04-04",
  "TOTAL_VACCINATIONS" : 5872684,
  "PERSONS_VACCINATED_1PLUS_DOSE" : 5188057,
  "TOTAL_VACCINATIONS_PER100" : 15.086,
  "PERSONS_VACCINATED_1PLUS_DOSE_PER100" : 13.327,
  "PERSONS_FULLY_VACCINATED" : 4532577,
  "PERSONS_FULLY_VACCINATED_PER100" : 11.643,
  "VACCINES_USED" : "Beijing CNBG - BBIBP-CoV, Janssen - Ad26.COV 2-S, Pfizer BioNTech - Comirnaty, SII - Covishield",
  "FIRST_VACCINE_DATE" : "2021-02-22",
  "NUMBER_VACCINES_TYPES_USED" : 4,
  "PERSONS_BOOSTER_ADD_DOSE" : 0,
  "PERSONS_BOOSTER_ADD_DOSE_PER100" : null
}
```

```
db.newCol.updateMany({"PERSONS_BOOSTER_ADD_DOSE_PER100":null},
{$set{"PERSONS_BOOSTER_ADD_DOSE_PER100":0}})
```

```
> db.vac.updateMany({PERSONS_BOOSTER_ADD_DOSE_PER100:null}, {$set:{PERSONS_BOOSTER_ADD_DOSE_PER100:0}})
{ "acknowledged" : true, "matchedCount" : 49, "modifiedCount" : 49 }
> db.vac.find({PERSONS_BOOSTER_ADD_DOSE:0, PERSONS_BOOSTER_ADD_DOSE_PER100:0}).pretty()
{
  "_id" : ObjectId("6256a3a45ad856e2f5e90815"),
  "COUNTRY" : "Afghanistan",
  "ISO3" : "AFG",
  "WHO_REGION" : "EMRO",
  "DATA_SOURCE" : "REPORTING",
  "DATE_UPDATED" : "2022-04-04",
  "TOTAL_VACCINATIONS" : 5872684,
  "PERSONS_VACCINATED_1PLUS_DOSE" : 5188057,
  "TOTAL_VACCINATIONS_PER100" : 15.086,
  "PERSONS_VACCINATED_1PLUS_DOSE_PER100" : 13.327,
  "PERSONS_FULLY_VACCINATED" : 4532577,
  "PERSONS_FULLY_VACCINATED_PER100" : 11.643,
  "VACCINES_USED" : "Beijing CNBG - BBIBP-CorV,Janssen - Ad26.COV 2-S,Pfizer BioNTech - Comirnaty,SII - Covishield",
  "FIRST_VACCINE_DATE" : "2021-02-22",
  "NUMBER_VACCINES_TYPES_USED" : 4,
  "PERSONS_BOOSTER_ADD_DOSE" : 0,
  "PERSONS_BOOSTER_ADD_DOSE_PER100" : 0
}
```

```
db.newCol.find({},{"_id":0,"COUNTRY":1,"PERSONS_FULLY_VACCINATED_PER100":1}).sort(
{"PERSONS_FULLY_VACCINATED_PER100":-1})
```

```
type "it" for more
> db.newCol.find({}, {"_id":0,"COUNTRY":1,"PERSONS_FULLY_VACCINATED_PER100":1}).sort({"PERSONS_FULLY_VACCINATED_PER100":-1})
{ "COUNTRY" : "Gibraltar", "PERSONS_FULLY_VACCINATED_PER100" : 122.614 }
{ "COUNTRY" : "Palau", "PERSONS_FULLY_VACCINATED_PER100" : 98.568 }
{ "COUNTRY" : "United Arab Emirates", "PERSONS_FULLY_VACCINATED_PER100" : 98.25 }
{ "COUNTRY" : "Brunei Darussalam", "PERSONS_FULLY_VACCINATED_PER100" : 92.641 }
{ "COUNTRY" : "Puerto Rico", "PERSONS_FULLY_VACCINATED_PER100" : 91.797 }
{ "COUNTRY" : "Chile", "PERSONS_FULLY_VACCINATED_PER100" : 90.93 }
{ "COUNTRY" : "Qatar", "PERSONS_FULLY_VACCINATED_PER100" : 90.325 }
{ "COUNTRY" : "Cayman Islands", "PERSONS_FULLY_VACCINATED_PER100" : 88.947 }
{ "COUNTRY" : "Niue", "PERSONS_FULLY_VACCINATED_PER100" : 87.577 }
{ "COUNTRY" : "Cuba", "PERSONS_FULLY_VACCINATED_PER100" : 87.53 }
{ "COUNTRY" : "Republic of Korea", "PERSONS_FULLY_VACCINATED_PER100" : 86.643 }
{ "COUNTRY" : "Malta", "PERSONS_FULLY_VACCINATED_PER100" : 86.146 }
{ "COUNTRY" : "Portugal", "PERSONS_FULLY_VACCINATED_PER100" : 85.59 }
{ "COUNTRY" : "Singapore", "PERSONS_FULLY_VACCINATED_PER100" : 84.867 }
{ "COUNTRY" : "China", "PERSONS_FULLY_VACCINATED_PER100" : 84.68 }
{ "COUNTRY" : "Cambodia", "PERSONS_FULLY_VACCINATED_PER100" : 83.991 }
{ "COUNTRY" : "New Zealand", "PERSONS_FULLY_VACCINATED_PER100" : 83.796 }
{ "COUNTRY" : "Faroe Islands", "PERSONS_FULLY_VACCINATED_PER100" : 83.69 }
{ "COUNTRY" : "Cook Islands", "PERSONS_FULLY_VACCINATED_PER100" : 83.557 }
{ "COUNTRY" : "Australia", "PERSONS_FULLY_VACCINATED_PER100" : 83.453 }
type "it" for more
```

다음으로, 인구 100명당 예방접종을 완료한

누적인원(PERSONS_FULLY_VACCINATED_PER100) 값을 확인한 후, 대한민국의 접종 수치에 근접하게 백신 접종을 완료한 국가의 경우 코로나 바이러스의 풍토병화를 예상해볼 수 있다,

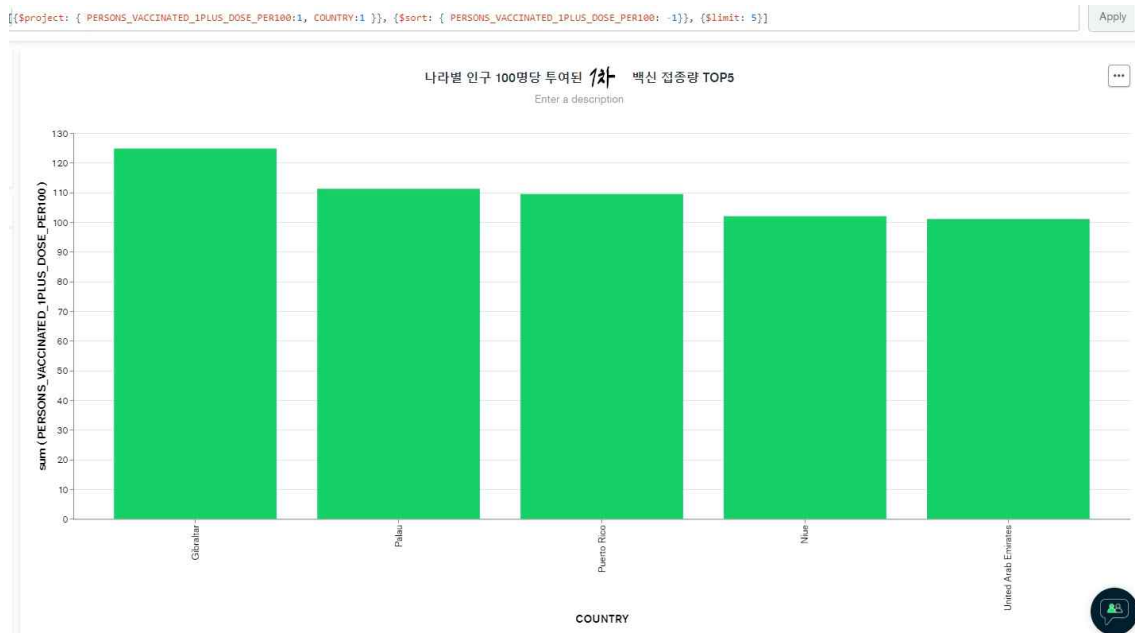
기준이 되는 우리나라에서 2차까지 백신 접종을 완료한 비율은 86.643 퍼센트이다. 그 위로는 쿠바, 푸에르토 리코, 아랍 등의 국가들이 있고, 아래로는 포르투갈, 싱가포르, 중국 등이 있다. 화면에 보이는 나라들의 2차까지의 접종률이 우리나라의 수치와 가장 근사하다.

```
db.newCol.find({}, {"_id":0, "COUNTRY":1, "PERSONS_BOOSTER_ADD_DOSE_PER100":1}).sort({"PERSONS_BOOSTER_ADD_DOSE_PER100":-1})
```

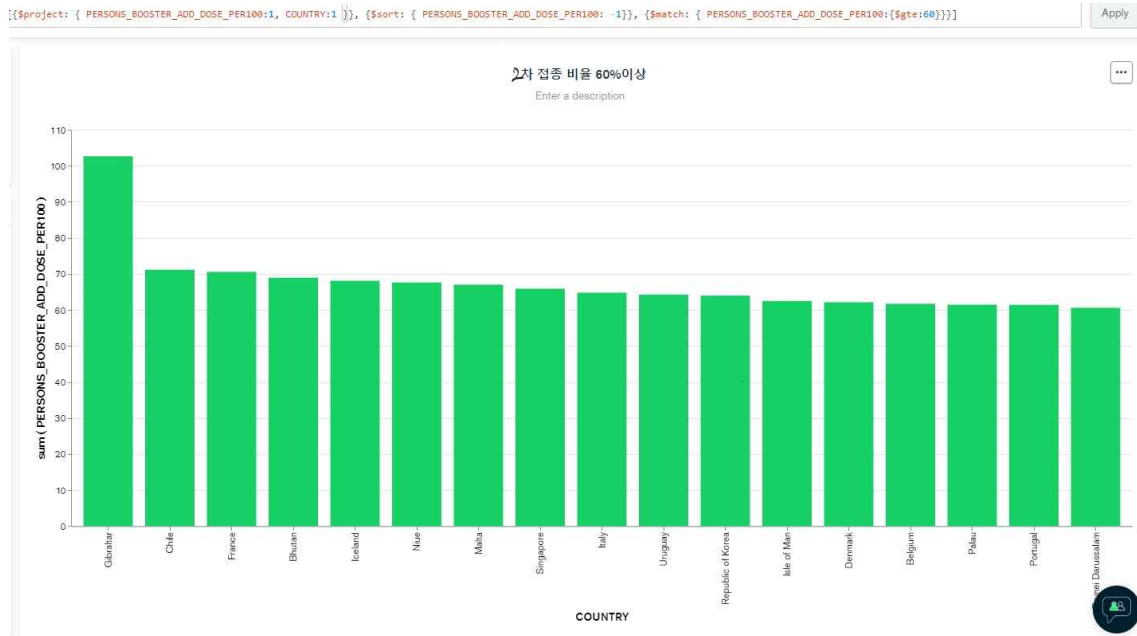
```
> db.newCol.find({}, {"_id":0, "COUNTRY":1, "PERSONS_BOOSTER_ADD_DOSE_PER100":1}).sort({"PERSONS_BOOSTER_ADD_DOSE_PER100":-1})
{"COUNTRY": "Gibraltar", "PERSONS_BOOSTER_ADD_DOSE_PER100": 102.654 }
{"COUNTRY": "Chile", "PERSONS_BOOSTER_ADD_DOSE_PER100": 71.152 }
{"COUNTRY": "France", "PERSONS_BOOSTER_ADD_DOSE_PER100": 70.55 }
{"COUNTRY": "Bhutan", "PERSONS_BOOSTER_ADD_DOSE_PER100": 68.945 }
{"COUNTRY": "Iceland", "PERSONS_BOOSTER_ADD_DOSE_PER100": 68.094 }
{"COUNTRY": "Niue", "PERSONS_BOOSTER_ADD_DOSE_PER100": 67.614 }
{"COUNTRY": "Malta", "PERSONS_BOOSTER_ADD_DOSE_PER100": 67.009 }
{"COUNTRY": "Singapore", "PERSONS_BOOSTER_ADD_DOSE_PER100": 65.892 }
{"COUNTRY": "Italy", "PERSONS_BOOSTER_ADD_DOSE_PER100": 64.793 }
{"COUNTRY": "Uruguay", "PERSONS_BOOSTER_ADD_DOSE_PER100": 64.256 }
{"COUNTRY": "Republic of Korea", "PERSONS_BOOSTER_ADD_DOSE_PER100": 63.976 }
{"COUNTRY": "Isle of Man", "PERSONS_BOOSTER_ADD_DOSE_PER100": 62.476 }
{"COUNTRY": "Denmark", "PERSONS_BOOSTER_ADD_DOSE_PER100": 62.132 }
{"COUNTRY": "Belgium", "PERSONS_BOOSTER_ADD_DOSE_PER100": 61.7 }
{"COUNTRY": "Palau", "PERSONS_BOOSTER_ADD_DOSE_PER100": 61.469 }
{"COUNTRY": "Portugal", "PERSONS_BOOSTER_ADD_DOSE_PER100": 61.428 }
{"COUNTRY": "Brunei Darussalam", "PERSONS_BOOSTER_ADD_DOSE_PER100": 60.584 }
{"COUNTRY": "Ireland", "PERSONS_BOOSTER_ADD_DOSE_PER100": 59.052 }
{"COUNTRY": "Germany", "PERSONS_BOOSTER_ADD_DOSE_PER100": 58.778 }
{"COUNTRY": "Austria", "PERSONS_BOOSTER_ADD_DOSE_PER100": 57.385 }
```

예상 외로 우리나라를 제외한 다른 국가들에서 2차 접종까지 원활하게 이루어지지 않은 경우들이 많았다. 기존 코로나 바이러스의 출현 이후, 오미크론 변종으로 인한 2차 접종이 시행되었다. 따라서 우리나라와 비슷하게 2차 접종이 원활히 이루어진 국가들을 찾아보았다.

위에서부터 칠레, 프랑스, 싱가포르, 이탈리아, 우리나라, 덴마크, 벨기에 등 초기에 백신접종으로 집단면역 형성이 기대되었던 유럽 지역의 국가들에서 2차 접종 비율이 높은 것을 볼 수 있었다.



1차 백신 접종률 TOP 5개국을 살펴보았다, 1위부터, 지브롤터(영국의 해외 영토) 팔라우(오세아니아의 국가), 푸에르토리코(카리브 해 미국령 섬), 니우에(오세아니아 섬), 아랍 에미레이트(중동 국가)로 나왔으나 TOP 5개국 이 섬 혹은 인구 수가 적은 국가이다. 백분율이 크게 의미가 없기 때문에, 실질적인 2차 접종률을 비교하기로 했다.



일반 백신 기준, 집단 면역 형성 기준은 국민의 70퍼센트로 정의되어 있지만, 2차 접종의 경우 추가적인 접종이기 때문에 70퍼센트를 넘는 국가들이 많지 않다. 따라서 60퍼센트로 기준을 낮추어 해당되는 모든 국가들을 살펴보았다. 228개국 중 2차 접종 비율이 60퍼센트가 넘는 국가는 총 17개국이며, 주로 유럽 국가들이 많이 속해있고 우리나라는 11번째에 속한다.

```
db.vac.find({"HERD_IMMUNITY":"TRUE"}, {_id:0, COUNTRY:1,
PERSONS_VACCINATED_1PLUS_DOSE_PER100:1, HERD_IMMUNITY:1,
VACCINES_USED:1}).pretty()
```

```
db.vac.updateMany({"PERSONS_VACCINATED_1PLUS_DOSE_PER100":{"gte":60}}, {"$set":{"HERD_IMMUNITY":"TRUE"}}, {upsert:true})
{"acknowledged":true,"matchedCount":93,"modifiedCount":93}
> db.vac.find({"HERD_IMMUNITY":"TRUE"}, {_id:0, COUNTRY:1, PERSONS_VACCINATED_1PLUS_DOSE_PER100:1, HERD_IMMUNITY:1, VACCINES_USED:1}).pretty()

{"COUNTRY": "Andorra",
"PERSONS_VACCINATED_1PLUS_DOSE_PER100": 75.922,
"VACCINES_USED": "AstraZeneca - Vaxzevria, Moderna - Spikevax, Pfizer - BioNTech - Comirnaty",
"HERD_IMMUNITY": "TRUE"}

{"COUNTRY": "Argentina",
"PERSONS_VACCINATED_1PLUS_DOSE_PER100": 90.541,
"VACCINES_USED": "AstraZeneca - Vaxzevria, Beijing OIBG - BBIBP-CorV, CanSino - Convidecia, Gamaleya - Gam-Covid-Vac, Moderna - Spikevax, Pfizer - BioNTech - Comirnaty, SII - Covishield",
"HERD_IMMUNITY": "TRUE"}

{"COUNTRY": "Aruba",
"PERSONS_VACCINATED_1PLUS_DOSE_PER100": 82.347,
"VACCINES_USED": "Janssen - Ad26.COV.2-S, Pfizer - BioNTech - Comirnaty",
"HERD_IMMUNITY": "TRUE"}

{"COUNTRY": "American Samoa",
"PERSONS_VACCINATED_1PLUS_DOSE_PER100": 79.057,
"VACCINES_USED": "Janssen - Ad26.COV.2-S, Moderna - Spikevax, Pfizer - BioNTech - Comirnaty",
"HERD_IMMUNITY": "TRUE"}

{"COUNTRY": "Australia",
"PERSONS_VACCINATED_1PLUS_DOSE_PER100": 87.125,
"VACCINES_USED": "AstraZeneca - Vaxzevria, Janssen - Ad26.COV.2-S, Moderna - Spikevax, Novavax-NUVAXOVID, Pfizer - BioNTech - Comirnaty",
"HERD_IMMUNITY": "TRUE"}

Windows 정품 인증
```

1차 접종 백분율에 대하여 집단면역 기준인 인구의 70% 가 넘는지를 기준으로 판별하는 변수 HERD_IMMUNITY 를 upsert 하여 1차 접종에서의 집단면역 형성 가능 여부를 TRUE로 나타내어 집단 면역 가능성에 해당하는 국가들을 추려보고자 한다.

먼저, 1차 접종률 (PERSONS_VACCINATED_1PLUS_DOSE_PER100)을 기준으로 70이 넘으면 TRUE 값을 갖는 HERD_IMMUNITY 변수를 upsert 했다. (insert는 document 단위로만 삽입이 가능하다는 제한이 있어서 upsert(update+insert)를 이용하여 조건에 해당되는

도큐먼트에만 HERD_IMMUNITY를 추가하였다.

1차 접종률이 기준(70%)에 부합하지 않아 HERD_IMMUNITY 값을 가지지 않는 국가들은 분석 목적상 활용하기 어려우므로 delete 하였다. 삭제된 국가는 처음의 총 228개국 중 136개국으로, 전세계에서 59퍼센트에 달하는 국가들이 집단면역 기준을 달성하지 못한 것으로 보인다.

```
db.newCol.deleteMany({"TOTAL_VACCINATIONS_PER100" : {$lt: 70}})
"acknowledged" : true, "deletedCount" : 136 }
```

```
db.vac.find({"HERD_IMMUNITY":"TRUE"}, {_id:0, COUNTRY:1,
PERSONS_VACCINATED_1PLUS_DOSE_PER100:1, HERD_IMMUNITY:1,
VACCINES_USED:1}).pretty()
```

```
db.vac.updateMany({'PERSONS_VACCINATED_1PLUS_DOSE_PER100' : {$lt: 70}}, {$set: {'HERD_IMMUNITY' : "TRUE"}}, {upsert:true})
{ "acknowledged" : true, "matchedCount" : 93, "modifiedCount" : 93 }
> db.vac.find({'HERD_IMMUNITY':'TRUE'}, {_id:0, COUNTRY:1, PERSONS_VACCINATED_1PLUS_DOSE_PER100:1, HERD_IMMUNITY:1, VACCINES_USED:1}).pretty()
{
  "COUNTRY" : "Andorra",
  "PERSONS_VACCINATED_1PLUS_DOSE_PER100" : 75.922,
  "VACCINES_USED" : "AstraZeneca - Vaxzevria, Moderna - Spikevax, Pfizer BioNTech - Comirnaty",
  "HERD_IMMUNITY" : "TRUE"
}
{
  "COUNTRY" : "Argentina",
  "PERSONS_VACCINATED_1PLUS_DOSE_PER100" : 90.541,
  "VACCINES_USED" : "AstraZeneca - Vaxzevria, Beijing CNBG - BBIBP-CorV, CanSino - Convidecia, Gamaleya - Gam-Covid-Vac, Moderna - Spikevax, Pfizer BioNTech - Comirnaty, SII - Covishield",
  "HERD_IMMUNITY" : "TRUE"
}
{
  "COUNTRY" : "Aruba",
  "PERSONS_VACCINATED_1PLUS_DOSE_PER100" : 82.347,
  "VACCINES_USED" : "Janssen - Ad26.COV.2-S, Pfizer BioNTech - Comirnaty",
  "HERD_IMMUNITY" : "TRUE"
}
{
  "COUNTRY" : "American Samoa",
  "PERSONS_VACCINATED_1PLUS_DOSE_PER100" : 79.057,
  "VACCINES_USED" : "Janssen - Ad26.COV.2-S, Moderna - Spikevax, Pfizer BioNTech - Comirnaty",
  "HERD_IMMUNITY" : "TRUE"
}
{
  "COUNTRY" : "Australia",
  "PERSONS_VACCINATED_1PLUS_DOSE_PER100" : 87.125,
  "VACCINES_USED" : "AstraZeneca - Vaxzevria, Janssen - Ad26.COV.2-S, Moderna - Spikevax, Novavax-NVX0VID, Pfizer BioNTech - Comirnaty",
  "HERD_IMMUNITY" : "TRUE"
}
Windows 정품 인증
```

다음으로, HERD_IMMUNITY 값이 TRUE인 집단면역 대상국의 1차 접종률 및 사용 백신에 대해 알아보았다. 사용 백신으로는 모더나와 화이자가 가장 보편적이다.

```
db.newCol.find({"$text":{"$search":"Moderna"}}, {"VACCINES_USED":1,
"COUNTRY":1}).pretty()
```

```
> db.newCol.find({"$text":{"$search":"Moderna"}}, {"VACCINES_USED":1, "COUNTRY":1}).pretty()

  "_id" : ObjectId("6251069d841dc4efbab30a21"),
  "COUNTRY" : "Wallis and Futuna",
  "VACCINES_USED" : "Moderna - Spikevax"
}

  "_id" : ObjectId("6251069d841dc4efbab30991"),
  "COUNTRY" : "Greenland",
  "VACCINES_USED" : "Moderna - mRNA-1273"
}

  "_id" : ObjectId("6251069d841dc4efbab3098a"),
  "COUNTRY" : "French Guiana",
  "VACCINES_USED" : "Moderna - Spikevax,Pfizer BioNTech - Comirnaty"
}

  "_id" : ObjectId("6251069d841dc4efbab3095c"),
  "COUNTRY" : "Bonaire, Sint Eustatius and Saba",
  "VACCINES_USED" : "Moderna - Spikevax,Pfizer BioNTech - Comirnaty"
}

  "_id" : ObjectId("6251069d841dc4efbab30987"),
  "COUNTRY" : "Fiji",
  "VACCINES_USED" : "AstraZeneca - Vaxzevria,Moderna - Spikevax,SII - Covishield"
}

  "_id" : ObjectId("6251069d841dc4efbab30986"),
  "COUNTRY" : "Faroe Islands",
  "VACCINES_USED" : "Moderna - mRNA-1273, Pfizer BioNTech - Comirnaty"
}

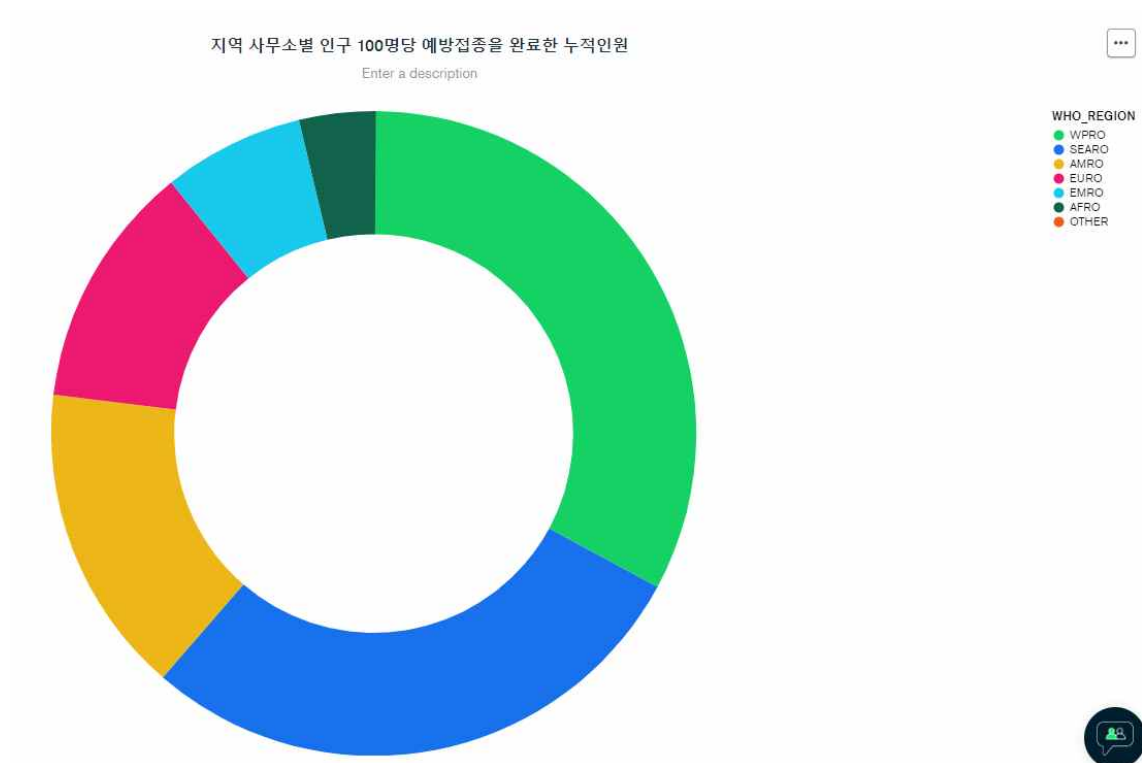
  "_id" : ObjectId("6251069d841dc4efbab30a25"),
  "COUNTRY" : "The United Kingdom",
  "VACCINES_USED" : "AstraZeneca - Vaxzevria,Moderna - Spikevax,Pfizer BioNTech - Comirnaty"
}
```

사용성이 가장 높은 모더나 백신의 사용국을 살펴보았다.

Aggregation 분석 |

WHO 회원국은 6개의 WHO 지역으로 그룹화되어 관리된다. 아프리카 지역 사무소(AFRO), 아메리카 지역 사무소(AMRO), 동남아시아 지역 사무소(SEARO), 유럽 지역 사무소(EURO), 동부 지중해 지역 사무소(EMRO) 및 서태평양 지역 사무소(WPRO)로 나뉜다. 전 세계 국가들을 지역 사무소 단위로 나누어 사무소 별로 접종 합계의 비율을 살펴보았다.

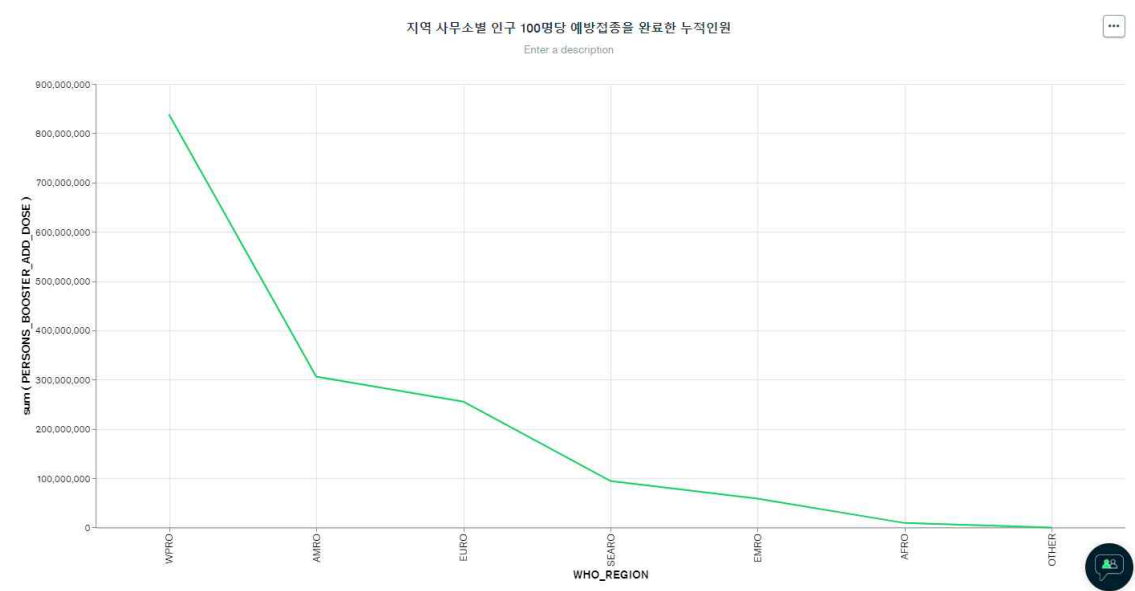
```
[{$project: {WHO_REGION: 1, PERSONS_VACCINATED_1PLUS_DOSE_PER100: 1}},  
{$group: {_id: '$WHO_REGION', avg: {$avg:  
'$PERSONS_VACCINATED_1PLUS_DOSE_PER100'}}}, {$sort: {avg: -1}}]
```



```

[{$project: {WHO_REGION: 1,PERSONS_VACCINATED_1PLUS_DOSE_PER100: 1}},
{$group: {_id: '$WHO_REGION', avg: {$avg:
'$PERSONS_VACCINATED_1PLUS_DOSE_PER100'}}}, {$sort: {avg: -1}}]

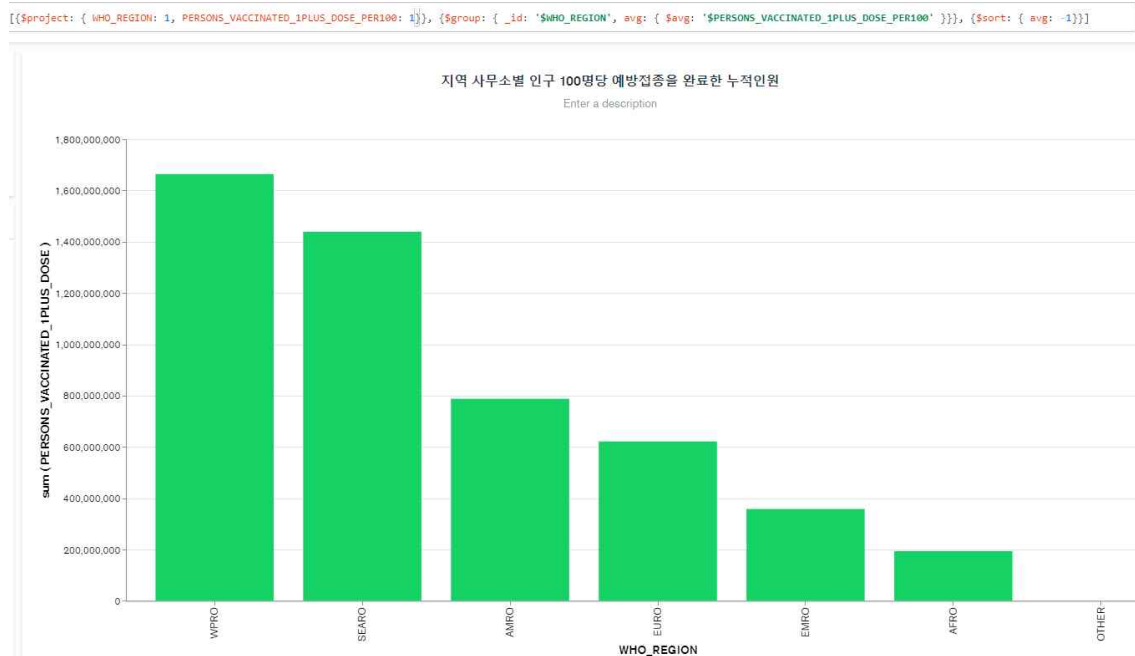
```



```

[{$project: {WHO_REGION: 1,PERSONS_VACCINATED_1PLUS_DOSE_PER100: 1}},
{$group: {_id: '$WHO_REGION', avg: {$avg:
'$PERSONS_VACCINATED_1PLUS_DOSE_PER100'}}}, {$sort: {avg: -1}}]

```



첫 번째로 백신 접종완료 합계가 가장 높은 사무소는 서태평양 지역 사무소(WPRO)이며, 우리나라, 일본, 중국과 베트남, 필리핀 등의 대표적 동남아 국가들이 포함되어 있다. 하지만

대부분의 동남아 국가에서는 의료 체계의 불안정으로 백신 물량 수급, 관리 및 접종이 원활하게 이루어지지 않았기 때문에, 우리나라, 일본, 중국의 접종자 합계가 압도적으로 높았을 것으로 생각된다.

두 번째로, 동남아시아 지역 사무소(SEARO)에는 북한, 인도, 스리랑카, 네팔, 부탄 등이 속한다. 부탄의 경우 위에서 접종완료 누적 인원 수치가 높은 편에 속하는 것을 확인하였다.

세 번째로는 유럽 지역 사무소(EURO)의 비율이 높고, 영국, 독일, 프랑스, 이탈리아, 스페인, 네덜란드 등이 이에 속해있다. 실제로 유럽 지역 국가들에서는 백신 접종이 원활히 진행되었으나 집단 면역 형성 정책에 중점을 두었고, 감염률이 높은 변종이 나타나면서 문제가 되었다.

분석 결과 |

1. 각국의 접종률을 집단면역 조건과 비교하여 형성 가능성을 예측하려고 하였다. 하지만 국가 단위의 접종률을 이용하여 비교하기에는 인구 수가 비정상적으로 적은 국가(섬 등)가 포함되기 때문에 접종률이 비정상적으로 높게 나온다. 따라서 의도하고자 하는 결론에 적절하지 않았다.

2. 누적인원 집계 가장 정확했다. 모든 국가를 비교하기에는 인구 수로 인한 편차가 너무 클 것으로 예상되어 WHO 지역 사무소 단위로 그룹화하여 접종 누적 인원을 분석하였다. 해당 지역 사무소에 속하는 국가의 특성을 감안하여 분석한 결과, 접종 누적 인원이 월등히 많음에도 코로나 사태가 안정화되지 못했던 이유까지 살펴볼 수 있었다.

3. 집단면역 형성 국가에 포함되는 국가임에도 국가 정책 및 접종 환경, 접종 백신 등이 요인에 따라 코로나 바이러스의 안정화 정도가 크게 달라질 수 있다. 중국의 경우에는 효력이 없는 중국산 백신을 사용하였으며, 유럽 몇 개국의 경우에는 마스크 착용 정책 폐지 및 무리한 집단 면역 정책을 시행하여 변종 바이러스에 전혀 대비하지 못했다. 네팔의 경우, 접종률이 상당히 높았지만 방역 정책으로 저위험군인 젊은 층 사이 감염 확산을 허용하면서 앞선 사례와 같이 변종 바이러스에 전혀 대비하지 못했던 것이 컸다.

이렇게 집단 면역 형성이 가장 유력한 국가 추론과 동시에 접종 수치와는 다르게 폭증하는 감염률의 원인에 대하여 파악할 수 있었다.

분석을 거듭할수록 우리나라에서 코로나 바이러스가 신종 인플루엔자와 같이 안정화될 수 있었던 것은 타국과 비교하여 정부의 시기마다 적절한 방역 정책 및 견고한 의료시스템, 백신 수급 및 관리가 잘되었던 그 요인이 결정적이었던 것을 알게 되었다.

MongoDB 장단점 |

[장점]

사용성

RDMS와 비교하여, MongoDB는 database - collections - documents 구조로 이루어져 있어 문서지향적이므로 뛰어난 확장성으로 scale-out에 용이하다.

MongoDB와 같은 도큐먼트 지향적인 데이터 모델에서는 객체를 자연스럽게 모아 놓는 형태로 표현함으로써 객체를 전체적으로 작업할 수 있다는 점이 있다. 구조가 다양하다는 점 이외에도 도큐먼트에는 미리 정해진 스키마 정의가 없기에 프로젝트 초기 단계에서 개발 시간을 단축할 수 있어서 델을 실험해 보기 쉬워 수십 개의 데이터 모델을 시도해 본 후 최선의 모델을 선택할 수 있다.

가변적인 속성을 갖는 데이터를 표현 할 수 있기 때문에, 도큐먼트로 데이터 모델링을 하게 되면 조인이 필요 없고 새로운 속성은 한 도큐먼트에 동적으로 추가 할 수 있다.

확장 가능

데이터베이스의 확장은 결국 더 큰 장비로 성능 확장할지 혹은 여러 장비에 데이터를 나눠 분산 확장할지 결정해야 한다. MongoDB는 분산 확장을 염두에 두고 설계되었기에 document 지향데이터 모델로써 데이터를 여러 서버에 쉽게 분산해 준다. 이는 클러스터 내 데이터 양과 부하를 조절할 수 있게 한다.

인덱싱

MongoDB는 일반적인 보조 인덱스를 지원하면 고유, 복합, 공간 정박, 전문 인덱싱 기능도 제공한다. 중첩된 도큐먼트 및 배열과 같은 계층 구조의 보조 인덱스도 지원하며, 개발자는 모델링 기능을 자신의 애플리케이션에 가장 적합한 방식으로 최대한 활용할 수 있다.

집계

MongoDB는 데이터 처리 파이프라인 개념을 기반으로 한 집계 프레임워크를 제공한다. 집계 파이프라인은 데이터베이스 최적화를 최대한 활용해, 서버 측에서 비교적 간단한 일련의 단계로 데이터를 처리함으로써 복잡한 분석 엔진을 구축할 수 있게 한다.

특수한 유형

MongoDB는 로그와 같은 최신 데이터를 유지하고자 세션이나 고정 크기 컬렉션과 같이 특정 시간에 만료해야 하는 데이터에 대한 유효시간(TTL) 컬렉션을 지원한다. 또한 기준 필터와 일치하는 도큐먼트에 한정된 부분 인덱스를 지원하여 효율성을 높이고 필요한 저장 공간을 줄일 수 있다.

[단점]

mongoDB를 사용하여 프로젝트를 진행함에 있어 다양한 단점 또한 느낄 수 있었다.

데이터 자체가 즉각적으로 바뀌기는 경우가 생기는데, 이때 되돌리기 기능을 사용하고 싶지만 되돌리기 기능이 없어 insert한 경우에는 번거롭게 delete를 이용해 삭제할 수 있었지만, 많은 doc가 삽입된 경우엔 새롭게 json파일을 불러와 처음부터 진행하는 경우도 발생하였다.

deleteMany를 이용해 잘못 삭제했을때 해당하는 모든 doc가 삭제되어 복구하기 힘들때가 생긴다. insert의 경우엔 삽입 조건 설정(필드 상의 비교)이 다양하지 못했다.

모든 doc가 _id값을 가지고 있다.

mongoDB는 CASCADE가 불가능했다. 자동으로 이루어지지 않으니 수동으로 일일이 찾아서 바꿔줘야 한다. ObjectId를 활용하기에는 join이 필수적이다. 하지만 mongoDB에서 join은 관계형 데이터베이스에서처럼 데이터베이스 상에서 이루어지는 게 아니라 자바스크립트로 이루어지기 때문에 느린편이다.

데이터 업데이트 중 장애 발생 시, 분산 처리 시스템으로 인해 데이터가 손실될 수 있다. 기본적으로, 데이터 양이 많을 경우에 분산 저장이나 데이터 복제 기능의 오류로 일부 데이터가 손실될 가능성이 있다는 것이 가장 위험부담이 크다. 비효율적인 key의 중복 입력으로 인해 데이터의 공간 소모가 RDBMS에 비하여 컸다.