Question - 1

WRITE A QUERY TO CHECK NULL VALUES.

```sql
SELECT
    COUNT(*) FILTER (WHERE province IS NULL),
    COUNT(*) FILTER (WHERE country_region IS NULL),
    COUNT(*) FILTER (WHERE latitude IS NULL) ,
    COUNT(*) FILTER (WHERE longitude IS NULL),
    COUNT(*) FILTER (WHERE date IS NULL),
    COUNT(*) FILTER (WHERE confirmed IS NULL) ,
    COUNT(*) FILTER (WHERE deaths IS NULL) ,
    COUNT(*) FILTER (WHERE recovered IS NULL)
FROM corona_virus;
```

COVID-19 CORONAVIRUS

	count bigint	count bigint	count bigint	count bigint	count bigint	count bigint	count bigint	count bigint
1	0	0	0	0	0	0	0	0

Question - 2

FINDS THE EARLIEST AND LATEST DATES IN THE DATASET.

```
select
min(date) as start_date,
max(date) as End_date
from corona_virus;
```

COVID-19
CORONAVIRUS

	start_date date	end_date date
1	2020-01-22	2021-06-13

Question - 3

FIND MONTHLY AVERAGE FOR CONFIRMED, DEATHS, RECOVERED.

```
SELECT
    TO_CHAR(date, 'YYYY-MM') AS year_month,
    AVG(confirmed)::NUMERIC(10,4)   AS avg_confirmed,
    AVG(deaths)::NUMERIC(10,4)      AS avg_deaths,
    AVG(recovered)::NUMERIC(10,4)   AS avg_recovered
FROM corona_virus
GROUP BY year_month
ORDER BY year_month;
```

	year_month text	avg_confirmed numeric (10,4)	avg_deaths numeric (10,4)	avg_recovered numeric (10,4)
1	2020-01	4.1455	0.1234	0.0929
2	2020-02	15.2960	0.5936	7.0320
3	2020-03	161.1303	8.6607	27.8739
4	2020-04	505.8004	41.5223	171.6422
5	2020-05	574.8498	30.2809	318.2964
6	2020-06	859.2281	29.8175	548.7916
7	2020-07	1432.3611	35.1096	983.0582
8	2020-08	1611.8429	37.5367	1299.2947
9	2020-09	1784.5874	34.7773	1438.9067
10	2020-10	2412.1996	36.7583	1420.6431
11	2020-11	3592.1944	56.7634	1985.3446
12	2020-12	4050.4397	71.2183	2497.8850
13	2021-01	3911.2285	84.1837	1919.6370
14	2021-02	2433.3636	69.1649	1558.3917
15	2021-03	2916.7972	59.1998	1652.2859
16	2021-04	4699.3552	78.4387	3074.7851
17	2021-05	4005.2541	76.7803	4007.5078
18	2021-06	2508.6324	66.2622	2769.4496

Question - 4

FIND MAXIMUM VALUES FOR CONFIRMED, DEATHS, RECOVERED PER MONTH.

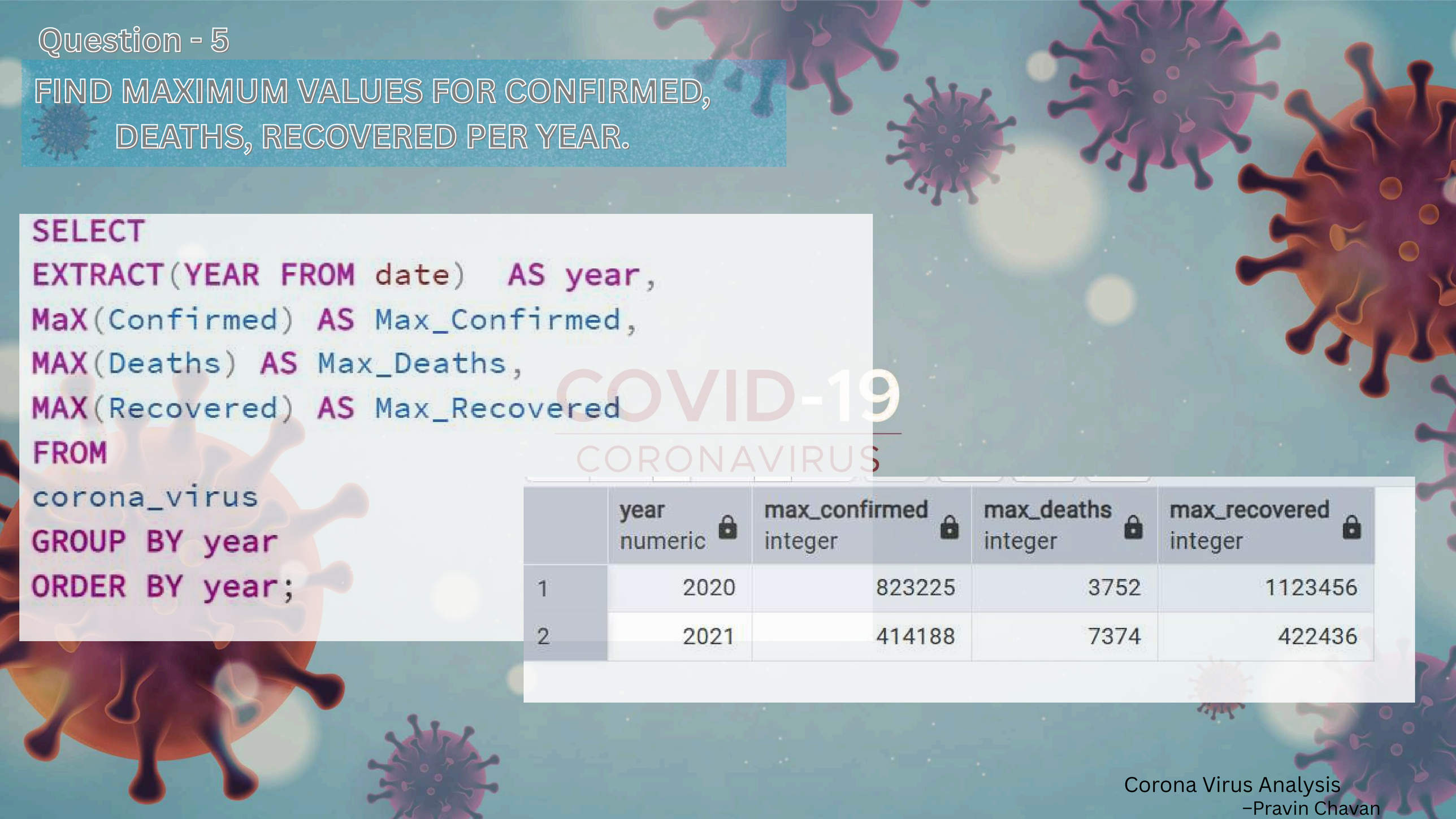
```
SELECT
 TO_CHAR(date, 'YYYY-MM') AS year_month,
MaX(Confirmed) AS Max_Confirmed,
MAX(Deaths) AS Max_Deaths,
MAX(Recovered) AS Max_Recovered
FROM
corona_virus
GROUP BY year_month
ORDER BY year_month;
```

	year_month text	max_confirmed integer	max_deaths integer	max_recovered integer
1	2020-01	2131	49	51
2	2020-02	14840	242	3418
3	2020-03	26314	1085	4289
4	2020-04	50740	2607	33227
5	2020-05	34907	2309	51717
6	2020-06	54771	2003	94305
7	2020-07	75866	1595	140050
8	2020-08	85687	1505	95881
9	2020-09	97894	1703	101468
10	2020-10	99264	3351	388340
11	2020-11	207933	2259	139292
12	2020-12	823225	3752	1123456
13	2021-01	300462	4475	87090
14	2021-02	134975	3907	98389
15	2021-03	100158	3869	102138
16	2021-04	401993	4249	299988
17	2021-05	414188	4529	422436
18	2021-06	134154	7374	231456

Question - 5

FIND MAXIMUM VALUES FOR CONFIRMED, DEATHS, RECOVERED PER YEAR.

```
SELECT
EXTRACT(YEAR FROM date)  AS year,
MaX(Confirmed) AS Max_Confirmed,
MAX(Deaths) AS Max_Deaths,
MAX(Recovered) AS Max_Recovered
FROM
corona_virus
GROUP BY year
ORDER BY year;
```



	year numeric	max_confirmed integer	max_deaths integer	max_recovered integer
1	2020	823225	3752	1123456
2	2021	414188	7374	422436

Question - 6

FIND HOW MONTHLY TOTAL CASES FOUND PER MONTH .

```
SELECT
EXTRACT(YEAR FROM date)  AS year,
EXTRACT(MONTH FROM date) AS month,
SUM(confirmed)   AS total_confirmed,
SUM(deaths)      AS total_deaths,
SUM(recovered)   AS total_recovered
FROM corona_virus
GROUP BY year ,month
ORDER BY year, month ;
```

	year numeric	month numeric	total_confirmed bigint	total_deaths bigint	total_recovered bigint
1	2020	1	6384	190	143
2	2020	2	68312	2651	31405
3	2020	3	769236	41346	133070
4	2020	4	2336798	191833	792987
5	2020	5	2744333	144561	1519547
6	2020	6	3969634	137757	2535417
7	2020	7	6838092	167613	4693120
8	2020	8	7694938	179200	6202833
9	2020	9	8244794	160671	6647749
10	2020	10	11515841	175484	6782150
11	2020	11	16595938	262247	9172292
12	2020	12	19336799	339996	11924903
13	2021	1	18672205	401893	9164347
14	2021	2	10492664	298239	6719785
15	2021	3	13924790	282620	7888013
16	2021	4	21711021	362387	14205507
17	2021	5	19121083	366549	19131842
18	2021	6	5022282	132657	5544438

Question - 7

CHECK HOW CORONA VIRUS SPREAD OUT WITH RESPECT TO CONFIRMED CASE.

```sql
SELECT
SUM(confirmed) AS total_confirmed_cases,
AVG(confirmed) AS avg_confirmed_cases,
VARIANCE(confirmed) AS variance_confirmed_cases,
STDDEV(confirmed) AS stdev_confirmed_cases
FROM corona_virus;
```

	total_confirmed_cases bigint	avg_confirmed_cases numeric	variance_confirmed_cases numeric	stdev_confirmed_cases numeric
1	169065144	2156.8283111780164825	157290931.69817455	12541.56815148

Question - 8

CHECK HOW CORONA VIRUS SPREAD OUT WITH RESPECT TO RECOVERD CASE .

```sql
SELECT
SUM(recovered) AS total_recovered_cases,
AVG(recovered) AS avg_recovered_cases,
VARIANCE(recovered) AS variance_recovered_cases,
STDDEV(recovered) AS stdev_recovered_cases
FROM corona_virus;
```

	total_recovered_cases bigint	avg_recovered_cases numeric	variance_recovered_cases numeric	stdev_recovered_cases numeric
1	113089548	1442.7263541959023295	107030888.69602982	10345.57338653

Question - 9

CHECK HOW CORONA VIRUS SPREAD OUT WITH RESPECT TO DEATH CASE.

```
SELECT
SUM(deaths) AS total_death_cases,
AVG(deaths) AS avg_death_cases,
VARIANCE(deaths) AS variance_death_cases,
STDDEV(deaths) AS stdev_death_cases
FROM corona_virus;
```

	total_death_cases bigint	avg_death_cases numeric	variance_death_cases numeric	stdev_death_cases numeric
1	3647894	46.5375704845252979	45892.604322956217	214.225592128850

Question - 10

CHECK HOW CORONA VIRUS SPREAD OUT WITH RESPECT TO DEATH CASE PER MONTH.

```
SELECT
EXTRACT(YEAR FROM date)  AS year,
EXTRACT(MONTH FROM date) AS month,
SUM(deaths) AS total_death_cases,
AVG(deaths)::NUMERIC(10,4) AS avg_death_cases,
VARIANCE(deaths)::NUMERIC(10,4) AS variance_death_cases,
STDDEV(deaths)::NUMERIC(10,4) AS stdev_death_cases
FROM corona_virus
GROUP BY year , month
ORDER BY year , month ;
```

	year numeric	month numeric	total_death_cases bigint	avg_death_cases numeric (10,4)	variance_death_cases numeric (10,4)	stdev_death_cases numeric (10,4)
1	2020	1	190	0.1234	4.2486	2.0612
2	2020	2	2651	0.5936	68.3372	8.2666
3	2020	3	41346	8.6607	3901.6095	62.4629
4	2020	4	191833	41.5223	40513.0372	201.2785
5	2020	5	144561	30.2809	20689.2454	143.8376
6	2020	6	137757	29.8175	16933.1109	130.1273
7	2020	7	167613	35.1096	21144.5841	145.4118
8	2020	8	179200	37.5367	23277.8724	152.5709
9	2020	9	160671	34.7773	20107.1214	141.7996
10	2020	10	175484	36.7583	17583.7543	132.6037
11	2020	11	262247	56.7634	27779.8065	166.6728
12	2020	12	339996	71.2183	65359.0598	255.6542
13	2021	1	401893	84.1837	102779.9614	320.5931
14	2021	2	298239	69.1649	68494.7562	261.7150
15	2021	3	282620	59.1998	54397.3642	233.2324
16	2021	4	362387	78.4387	94631.9540	307.6231
17	2021	5	366549	76.7803	131797.0766	363.0387
18	2021	6	132657	66.2622	113020.1266	336.1847

Corona Virus Analysis
–Pravin Chavan

Question - 11

CHECK HOW CORONA VIRUS SPREAD OUT WITH RESPECT TO RECOVERED CASE PER MONTH .

```
SELECT
EXTRACT(YEAR FROM date)  AS year,
EXTRACT(MONTH FROM date) AS month,
SUM(recovered) AS total_recovered_cases,
AVG(recovered)::NUMERIC(15,2) AS avg_recovered_cases,
VARIANCE(recovered)::NUMERIC(15,2) AS variance_recovered_cases,
STDDEV(recovered)::NUMERIC(15,2) AS stdev_recovered_cases
FROM corona_virus
GROUP BY year , month
ORDER BY year , month ;
```

	year numeric	month numeric	total_recovered_cases bigint	avg_recovered_cases numeric (15,2)	variance_recovered_cases numeric (15,2)	stdev_recovered_cases numeric (15,2)
1	2020	1	143	0.09	2.64	1.62
2	2020	2	31405	7.03	12449.45	111.58
3	2020	3	133070	27.87	40121.59	200.30
4	2020	4	792987	171.64	770059.71	877.53
5	2020	5	1519547	318.30	1978620.88	1406.63
6	2020	6	2535417	548.79	6531586.26	2555.70
7	2020	7	4693120	983.06	24849082.94	4984.89
8	2020	8	6202833	1299.29	40178838.38	6338.68
9	2020	9	6647749	1438.91	57035911.88	7552.21
10	2020	10	6782150	1420.64	73747150.17	8587.62
11	2020	11	9172292	1985.34	50738601.25	7123.10
12	2020	12	11924903	2497.89	326763170.52	18076.59
13	2021	1	9164347	1919.64	31500298.42	5612.51
14	2021	2	6719785	1558.39	24433077.90	4942.98
15	2021	3	7888013	1652.29	34904703.06	5908.02
16	2021	4	14205507	3074.79	224468171.33	14982.26
17	2021	5	19131842	4007.51	[illegible]	27483.34
18	2021	6	5544438	2769.45	[illegible]	[illegible]9.28

Question - 12

FIND COUNTRY HAVING HIGHEST NUMBER OF THE CONFIRMED ,DEATHS AND RECOVERED CASES.

```sql
SELECT
    country_region,
    SUM(confirmed)   AS total_confirmed,
    SUM(deaths)      AS total_deaths,
    SUM(recovered)   AS total_recovered
FROM corona_virus
GROUP BY country_region
ORDER BY
    total_confirmed DESC,
    total_deaths DESC,
    total_recovered DESC
LIMIT 1;
```

	country_region character varying (100)	total_confirmed bigint	total_deaths bigint	total_recovered bigint
1	US	33461982	599769	6303715

Question - 13

FIND COUNTRY HAVING LOWEST NUMBER OF THE DEATHS CASES.

```sql
WITH rankingCountry AS (
    SELECT
        country_region,
        SUM(deaths) AS total_death_cases,
        RANK() OVER (ORDER BY SUM(deaths) ASC) AS rank_no
    FROM corona_virus
    GROUP BY country_region
)
SELECT
    country_region,
    total_death_cases
FROM rankingCountry
WHERE rank_no = 1;
```

	country_region character varying (100)	total_death_cases bigint
1	Samoa	0
2	Kiribati	0
3	Dominica	0
4	Marshall Islands	0

Question - 14

FIND TOP 5 COUNTRIES HAVING MOST DEATH CASES.

```
select
country_region,
sum(deaths) as total_death_cases
from corona_virus
group by country_region
order by total_death_cases desc
limit 5;
```

COVID-19 CORONAVIRUS

	country_region character varying (100)	total_death_cases bigint
1	US	599769
2	Brazil	487401
3	India	370730
4	Mexico	230150
5	Peru	188708

Question - 15

FIND TOP 5 COUNTRIES HAVING MOST RECOVERED CASES.

```sql
select
country_region,
sum(recovered) as total_recovered_cases
from corona_virus
group by country_region
order by total_recovered_cases desc
limit 5;
```

	country_region character varying (100)	total_recovered_cases bigint
1	India	28089649
2	Brazil	15400169
3	US	6303715
4	Turkey	5202251
5	Russia	4745756