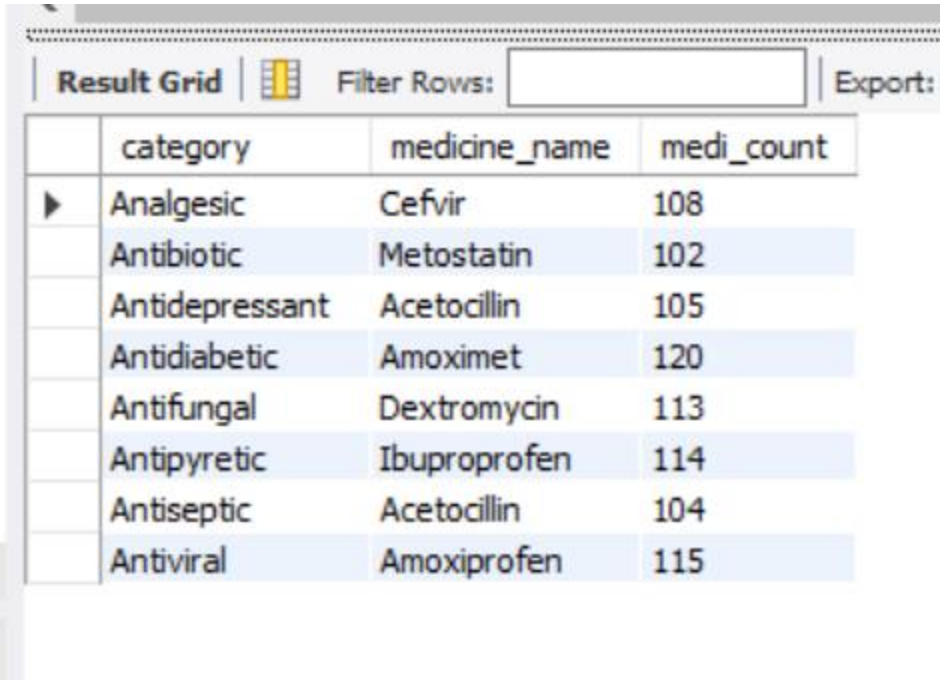
MEDICINE/DRUGS DATASET ANALYSIS

```
-- Clean data --

-- Remove 'mg' from Strength column

UPDATE medicine_dataset

SET Strength = cast(REMOVE(Strength, ' mg','')AS UNSIGNED);


-- set Strength column to Integer

ALTER TABLE medicine_dataset

MODIFY COLUMN Strength INT;


-- change 'name' column to medicine_name

ALTER TABLE medicine_dataset

CHANGE COLUMN name medicine_name CHAR (100);


-- change 'dosage form' column to 'dosage_form'

ALTER TABLE medicine_dataset

CHANGE COLUMN `dosage form` dosage_form CHAR(100);


-- 1. Most Common Medicines in Each Category:

-- Identify which specific medicines are most frequently listed within each therapeutic category.

WITH common_medicine as

(SELECT category, medicine_name, count(*) as medi_count

from medicine_dataset

GROUP BY category, medicine_name

),

rank_med AS

(SELECT category,medicine_name, medi_count,

        ROW_NUMBER () OVER (PARTITION BY category ORDER BY medi_count DESC) AS `rank`

FROM common_medicine)

SELECT category, medicine_name, medi_count

FROM rank_med

where `rank` = 1;
```

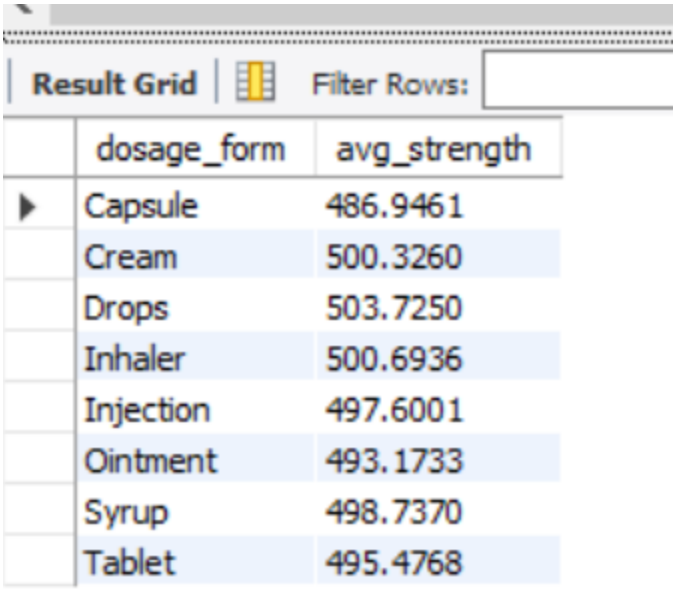


Result Grid | Filter Rows: | Export:

category	medicine_name	medi_count
Analgesic	Cefvir	108
Antibiotic	Metostatin	102
Antidepressant	Acetocillin	105
Antidiabetic	Amoximet	120
Antifungal	Dextromycin	113
Antipyretic	Ibuproprofen	114
Antiseptic	Acetocillin	104
Antiviral	Amoxiprofen	115

-- 2. Average Strength by Dosage Form: Determine the average strength of medicines for each dosage form.

```
WITH average_med AS
( SELECT dosage_form, avg(strength) AS avg_strength
FROM medicine_dataset
GROUP BY dosage_form
),
rank_avg AS
(SELECT dosage_form, avg_str,
        ROW_NUMBER() OVER (PARTITION BY dosage_form ORDER BY avg_str desc) as `rank`
FROM average_med
)
SELECT dosage_form, avg_str
FROM rank_avg
where `rank` = 1;
```



Result Grid | Filter Rows:

dosage_form	avg_strength
Capsule	486.9461
Cream	500.3260
Drops	503.7250
Inhaler	500.6936
Injection	497.6001
Ointment	493.1733
Syrup	498.7370
Tablet	495.4768

-- 3. Distribution of Medicines by Manufacturer: Count how many different medicines each manufacturer produces.

```
WITH manufacturer AS
( SELECT manufacturer, count(*) as medicine_count
FROM medicine_dataset
GROUP BY manufacturer
),

rank_manufacturer AS
( select manufacturer, medicine_count,
        ROW_NUMBER () OVER (PARTITION BY manufacturer ORDER BY medicine_count desc) AS `rank`
   FROM manufacturer
)
SELECT manufacturer, medicine_count
FROM rank_manufacturer
where `rank` = 1;
```

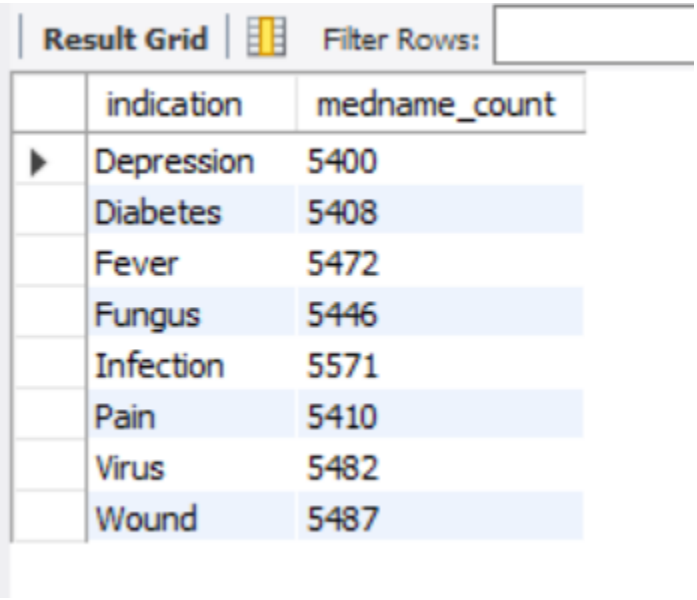
Result Grid | Filter Rows: | Export

manufacturer	medicine_count
AbbVie Inc.	2196
Amgen Inc.	2218
AstraZeneca plc	2217
Bayer AG	2182
Biogen Inc.	2227
Boehringer Ingelheim GmbH	2224
Bristol-Myers Squibb Company	2220
CSL Limited	2171
Eli Lilly and Company	2160
Gilead Sciences, Inc.	2181
GlaxoSmithKline plc	2195
Johnson & Johnson	2150
Merck & Co., Inc.	2146
Novartis AG	2154
Novo Nordisk A/S	2198
Pfizer Inc.	2218
Roche Holding AG	2143

-- 4. Count of Medicines by Indication: Find out how many medicines are available for each medical indication.

```
WITH med_indication AS
(SELECT indication, count(*) as medname_count
FROM medicine_dataset
GROUP BY indication
),

rank_indication AS
( SELECT indication, medname_count,
        ROW_NUMBER () OVER (PARTITION BY indication ORDER BY medname_count desc) AS `rank`
   FROM med_indication
)
```

```
SELECT indication, medname_count
FROM rank_indication
WHERE `rank` = 1;
```



Result Grid | Filter Rows:

indication	medname_count
Depression	5400
Diabetes	5408
Fever	5472
Fungus	5446
Infection	5571
Pain	5410
Virus	5482
Wound	5487

-- 5. Prescription vs. Over-the-Counter Medicines by Category:

-- Compare the number of prescription and over-the-counter medicines within each therapeutic category.

```
WITH med_classification AS
(select classification, count(*) as medname_count
FROM medicine_dataset
GROUP BY classification
),
rank_classification AS
(SELECT classification, medname_count,
        ROW_NUMBER() OVER (PARTITION BY classification ORDER BY medname_count desc) AS `rank`
   FROM med_classification
)
SELECT classification, medname_count
FROM rank_classification
WHERE `rank` = 1;
```

Result Grid | Filter Rows:

classification	medname_count
Over-the-Counter	21829
Prescription	21847

-- 6. Top Manufacturers for Specific Indications:

-- Identify which manufacturers produce the most medicines for particular medical conditions.

```
WITH manu_indication AS
(SELECT indication, manufacturer, count(*) AS medname_count
        FROM medicine_dataset
   GROUP BY indication, manufacturer
),
rank_manu_indication AS
(SELECT indication, manufacturer, medname_count,
        ROW_NUMBER() OVER (PARTITION BY indication ORDER BY medname_count desc) as `rank`
   FROM manu_indication
)
SELECT indication, manufacturer, medname_count
FROM rank_manu_indication
WHERE `rank` = 1;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

indication	manufacturer	medname_count
Depression	Teva Pharmaceutical Industries Ltd.	286
Diabetes	Bristol-Myers Squibb Company	296
Fever	Novo Nordisk A/S	301
Fungus	Takeda Pharmaceutical Company Limited	289
Infection	Pfizer Inc.	315
Pain	Roche Holding AG	297
Virus	AbbVie Inc.	309
Wound	Bristol-Myers Squibb Company	314

-- 7. Strength Range for Each Dosage Form: Determine the range of strengths available for each dosage form.

```
SELECT dosage_form, min(Strength), max(strength)

FROM medicine_dataset

GROUP BY dosage_form;
```

Result Grid | Filter Rows:

dosage_form	min(Strength)	max(strength)
Cream	1	999
Injection	1	999
Ointment	1	999
Syrup	1	999
Tablet	1	999
Inhaler	1	999
Capsule	1	999
Drops	1	999

-- 8. Medicines with the Highest and Lowest Strengths: Identify which medicines have the highest and lowest strengths.

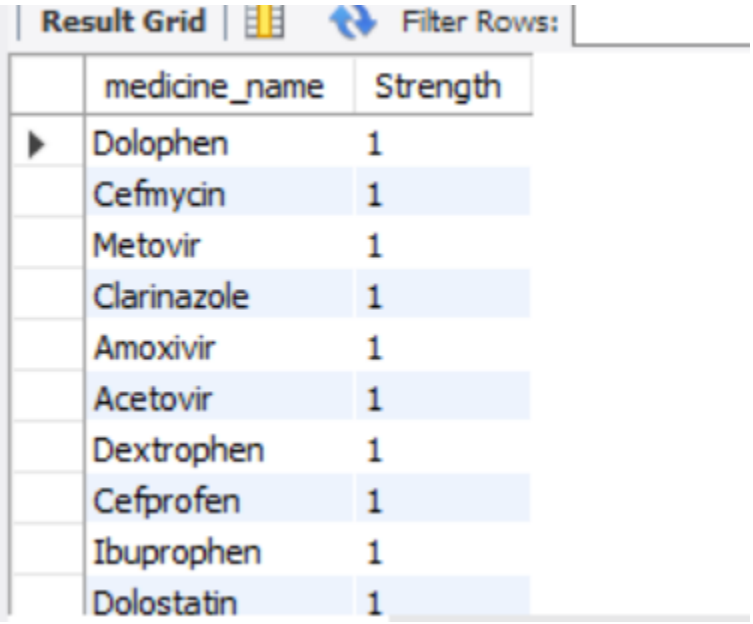
-- Lowest

```
SELECT medicine_name, Strength

FROM medicine_dataset

WHERE Strength = (SELECT min(Strength) FROM medicine_dataset);
```

medicine_name	Strength
Dolophen	1
Cefmycin	1
Metovir	1
Clarinazole	1
Amoxivir	1
Acetovir	1
Dextrophen	1
Cefprofen	1
Ibuprophen	1
Dolostatin	1

-- Highest

```
SELECT medicime_name, Strength
FROM medicine_dataset
WHERE Strength = (SELECT max(Strength) FROM medicine_dataset);
```



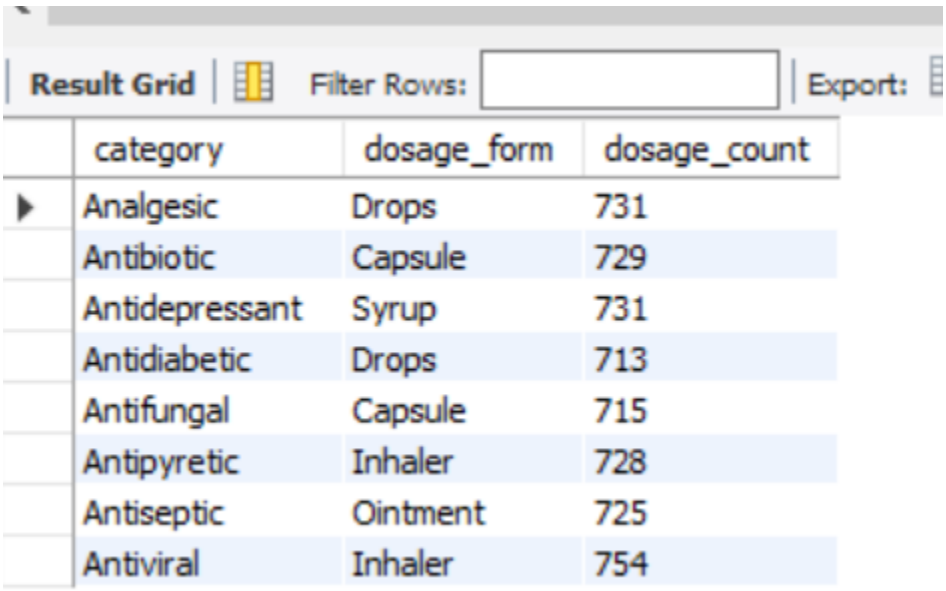
Result Grid | Filter Rows:

	medicine_name	Strength
▶	Dolophen	1
	Cefmycin	1
	Metovir	1
	Clarinazole	1
	Amoxivir	1
	Acetovir	1
	Dextrophen	1
	Cefprofen	1
	Ibuprophen	1
	Dolostatin	1

-- 9. Most Common Dosage Forms for Specific Categories:

-- Find out which dosage forms are most commonly associated with each therapeutic category.

```
WITH dosage_category AS
(SELECT dosage_form, category, count(*) AS dosage_count
FROM medicine_dataset
GROUP BY dosage_form, category
),

rank_dosage_category AS
(SELECT dosage_form, category, dosage_count,
        ROW_NUMBER () OVER (PARTITION BY category ORDER BY dosage_count desc) AS `rank`
   FROM dosage_category
)
SELECT category, dosage_form, dosage_count
FROM rank_dosage_category
WHERE `rank` = 1;
```



Result Grid | Filter Rows: | Export:

	category	dosage_form	dosage_count
▶	Analgesic	Drops	731
	Antibiotic	Capsule	729
	Antidepressant	Syrup	731
	Antidiabetic	Drops	713
	Antifungal	Capsule	715
	Antipyretic	Inhaler	728
	Antiseptic	Ointment	725
	Antiviral	Inhaler	754

Data Source: https://www.kaggle.com/datasets/ujjwalaggarwal402/medicine-dataset

GitHub Repository: https://github.com/Florence93/Medicine-Drugs-Data-Analysis