

```

USE crashes;

SELECT * FROM crashes;

DESCRIBE crashes;

SET sql_safe_updates = 0;

UPDATE crashes
SET crash_date = CASE
    WHEN crash_date REGEXP '^[0-9]{1,2}/[0-9]{1,2}/[0-9]{4} [0-9]{1,2}:[0-9]{1,2}$'
    THEN DATE_FORMAT(STR_TO_DATE(crash_date, '%m/%d/%Y %H:%i'), '%Y-%m-%d')
    WHEN crash_date REGEXP '^[0-9]{1,2}-[0-9]{1,2}-[0-9]{4} [0-9]{1,2}:[0-9]{1,2}$'
    THEN DATE_FORMAT(STR_TO_DATE(crash_date, '%m-%d-%Y %H:%i'), '%Y-%m-%d')
    ELSE DATE_FORMAT(STR_TO_DATE(crash_date, '%Y-%m-%d'), '%Y-%m-%d')
END;

ALTER TABLE crashes
MODIFY COLUMN crash_date DATE;

UPDATE crashes
SET date_police_notified = CASE
    WHEN date_police_notified REGEXP '^[0-9]{1,2}/[0-9]{1,2}/[0-9]{4}
[0-9]{1,2}:[0-9]{1,2}$' THEN DATE_FORMAT(STR_TO_DATE(date_police_notified,
'%m/%d/%Y %H:%i'), '%Y-%m-%d')
    WHEN date_police_notified REGEXP '^[0-9]{1,2}-[0-9]{1,2}-[0-9]{4}
[0-9]{1,2}:[0-9]{1,2}$' THEN DATE_FORMAT(STR_TO_DATE(date_police_notified,
'%m-%d-%Y %H:%i'), '%Y-%m-%d')
    ELSE DATE_FORMAT(STR_TO_DATE(date_police_notified, '%Y-%m-%d'), '%Y-%m-%d')
END;

ALTER TABLE crashes
MODIFY COLUMN date_police_notified DATE;

-- what are the total counts of recorded crashes in the dataset?

SELECT COUNT(*) As total_count
FROM crashes;

-- what is the earliest and latest date of recorded crashes?

SELECT MIN(crash_date) AS min, MAX(crash_date) AS max

```

```
FROM crashes;
```

```
-- What is the number of crashes reported per year?
```

```
SELECT YEAR(crash_date), COUNT(*)
FROM crashes
GROUP BY YEAR(crash_date)
ORDER BY YEAR(crash_date)
```

```
--
```

```
SELECT
    MONTHNAME(crash_date) AS crash_month,
    SUM(YEAR(crash_date) = 2017) AS count_2017,
    SUM(YEAR(crash_date) = 2018) AS count_2018,
    (SUM(YEAR(crash_date) = 2018) - SUM(YEAR(crash_date) = 2017)) AS count_diff,
    CASE
        WHEN SUM(YEAR(crash_date) = 2017) > 0 THEN
            ((SUM(YEAR(crash_date) = 2018) - SUM(YEAR(crash_date) = 2017)) /
SUM(YEAR(crash_date) = 2017)) * 100
        ELSE 0
    END AS percent_change
FROM
    crashes
WHERE
    YEAR(crash_date) IN (2017, 2018)
GROUP BY
    MONTHNAME(crash_date)
ORDER BY
    MONTHNAME(crash_date);
```

```
--
```

```
DROP TABLE IF EXISTS crash_new;
CREATE TABLE crash_new AS
(
    SELECT *
    FROM crashes
    WHERE YEAR(crash_date) BETWEEN '2018' AND '2022'
);
```

```
SELECT * FROM crash_new
```

```
SELECT COUNT(*) As total_count
FROM crash_new;
```

```
SELECT MIN(crash_date) AS min, MAX(crash_date) AS max
FROM crash_new;
```

```
TABLE crash_new  
INTO OUTFILE "C:/ProgramData/MySQL1/MySQL Server 8.0/Uploads/crash_new.csv"  
FIELDS TERMINATED BY ','  
OPTIONALLY ENCLOSED BY ''''  
ESCAPED BY ''  
LINES terminated by '\n'
```

```
SELECT 'crash_id', 'crash_date',  
'posted_speed_limit', 'traffic_control_device', 'device_condition', 'weather_condition',  
'lighting_condition', 'first_crash_type', 'traffic_way_type', 'lane_count', 'alignmen  
t', 'roadway_surface_condition',  
  
'road_defect', 'report_type', 'crash_type', 'hit_and_run', 'damage', 'date_police_notifi  
ed',  
'primary_cause', 'secondary_cause', 'street_direction', 'street_name', 'statement_take  
n', 'work_zone', 'work_zone_type', 'workers_present', 'number_unit',  
  
'most_severe_injury', 'injuries_total', 'injuries_fatal', 'injuries_incapacitated', 'in  
juries_non_incapacitated', 'injuries_reported_not_evident', 'crash_hour', 'crash_day_o  
f_week', 'crash_month'  
UNION ALL  
SELECT *  
INTO OUTFILE "C:/ProgramData/MySQL1/MySQL Server 8.0/Uploads/crash_new2.csv"  
FIELDS TERMINATED BY ','  
OPTIONALLY ENCLOSED BY ''''  
ESCAPED BY ''  
LINES TERMINATED BY '\n'  
FROM crash_new;
```