Crime LA Project

Introduction to SQL

Problem Statement

Business Context

People throughout Los Angeles are concerned about recent reports of crimes in numerous locations. The mayor of Los Angeles has established a new Criminal Investigation Division to study how and why crime is on the rise and the elements that contribute to it so that officials may take the necessary steps to keep the city's residents safe.

Objective

As a member of the Analytics Division, you are aware that numerous questions must be answered utilizing CID data. Import the dump file that contains various tables that are present in the database. Using the data, react to the questions, and write a detailed report for the authorities to utilize in taking action against crimes in Los Angeles.

Business Questions

Question 1: What is the total number of crimes for each crime status?

Solution Query:

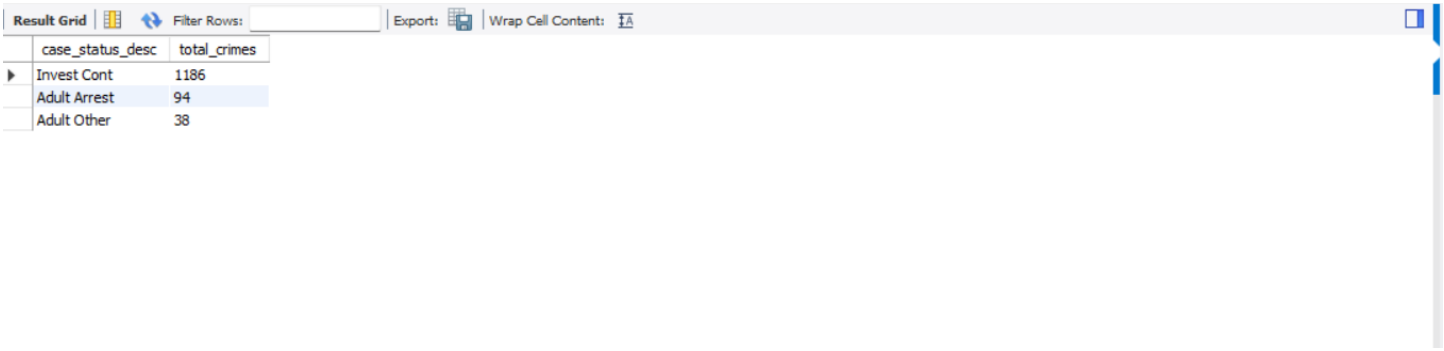
<Write your SQL query here>

Example:-

```
SELECT case_status_desc, COUNT(*) AS total_crimes
FROM report_t
GROUP BY case_status_desc;
```

Output:

<Attach the screenshot of your output table.>



case_status_desc	total_crimes
Invest Cont	1186
Adult Arrest	94
Adult Other	38

Observations and Insights for Question 1:

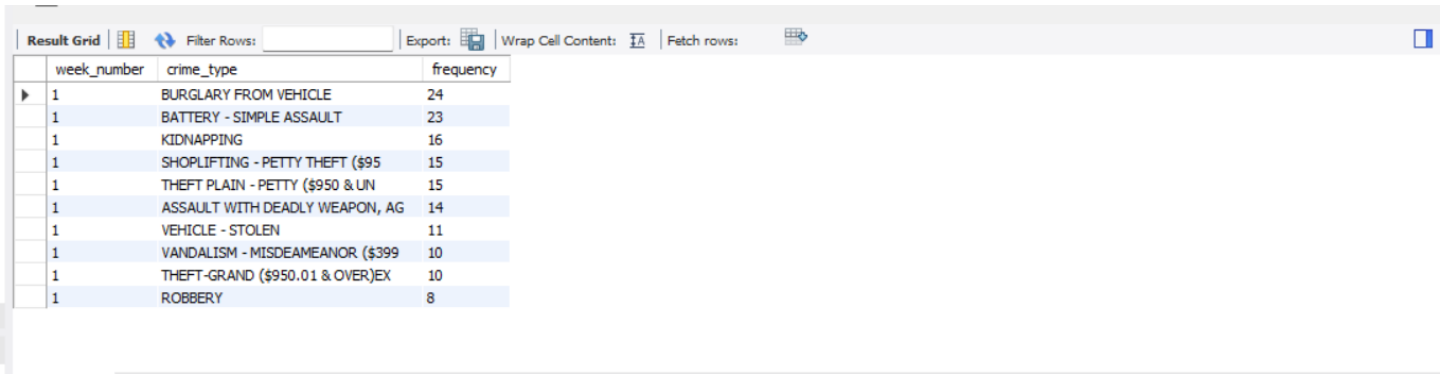
- "Invest Cont" (Investigation Continued) accounts for the majority of cases (1186), suggesting that a significant number of crimes remain unresolved or under active review.
- "Adult Arrest" cases (94) reflect successful identification and apprehension of suspects.
- "Adult Other" (38) could include administrative closures or unresolved cases outside standard classifications.
- The data indicates that law enforcement is still working on a large number of cases — possibly due to lack of evidence, offender escape, or complexity.

Question 2: Which was the most frequent crime committed each week?

Solution Query:

```
SELECT week_number, crime_type, COUNT(*) AS frequency

FROM report_t

GROUP BY week_number, crime_type

ORDER BY week_number, frequency DESC

LIMIT 10;
```

Output:



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

week_number	crime_type	frequency
1	BURGLARY FROM VEHICLE	24
1	BATTERY - SIMPLE ASSAULT	23
1	KIDNAPPING	16
1	SHOPLIFTING - PETTY THEFT ($95	15
1	THEFT PLAIN - PETTY ($950 & UN	15
1	ASSAULT WITH DEADLY WEAPON, AG	14
1	VEHICLE - STOLEN	11
1	VANDALISM - MISDEAMEANOR ($399	10
1	THEFT-GRAND ($950.01 & OVER)EX	10
1	ROBBERY	8

Observations and Insights:

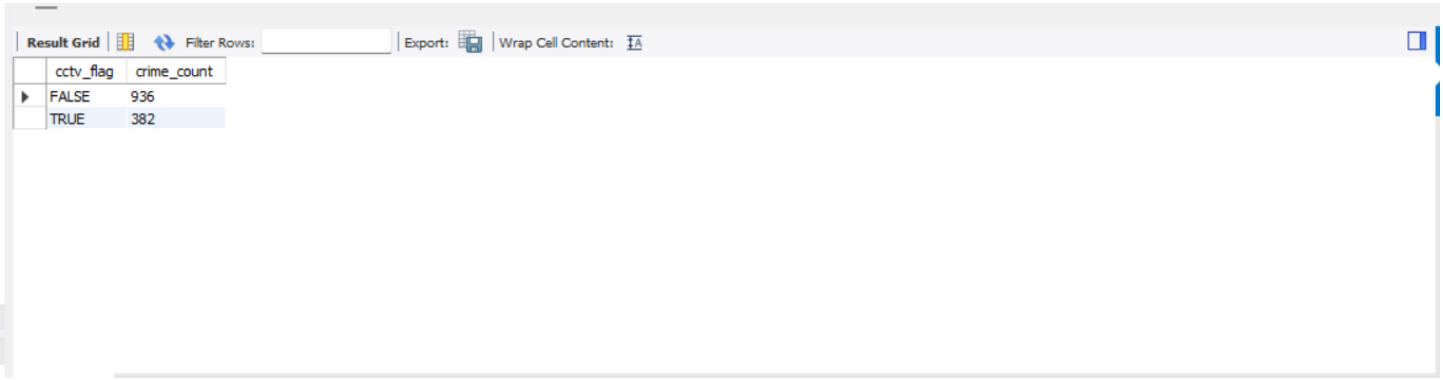
- In Week 1, the most frequent crime was Burglary from Vehicle, with 24 incidents, indicating theft from unattended cars is a common issue.
- Battery - Simple Assault (23 cases) and Kidnapping (16 cases) also occurred at high frequencies, raising concern for public safety and personal security.
- Petty thefts and assaults dominate the top list, which suggests a strong need for community awareness and prevention campaigns in public spaces and residential areas.
- The close counts between theft and assault categories highlight the dual problem of property and personal crimes.

Question 3: Does the existence of CCTV cameras deter crimes from happening?

Solution Query:

```
SELECT cctv_flag, COUNT(*) AS crime_count

FROM report_t

GROUP BY cctv_flag;
```

Output:



Observations and Insights:

- A significantly higher number of crimes occurred where CCTV was not present (936 cases without CCTV vs 382 cases with CCTV).
- This suggests that the presence of CCTV cameras may act as a deterrent to criminal activity.
- Areas without surveillance seem to experience more than double the crime, which supports the idea of installing more CCTV units in high-risk zones to discourage offenses.

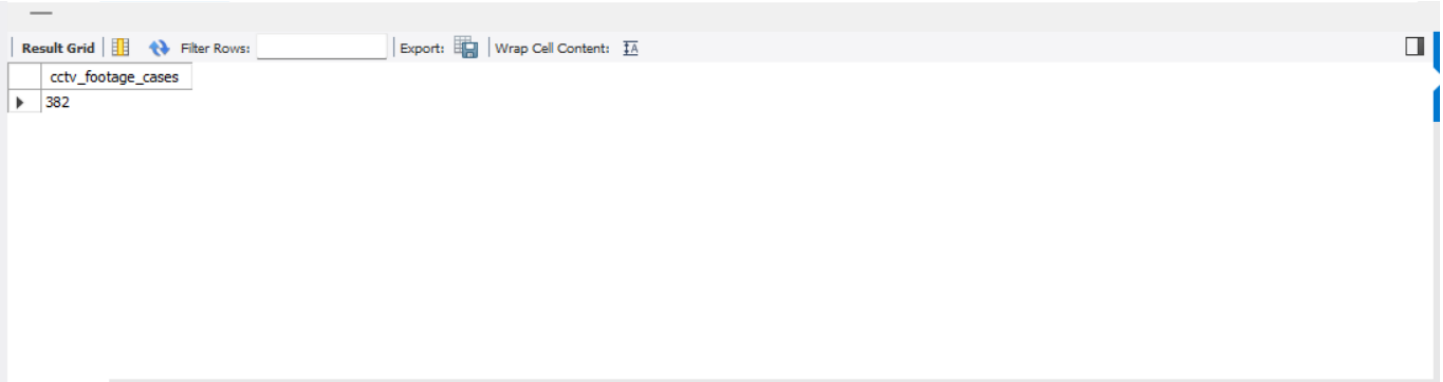
Question 4: How much footage has been recovered from the CCTV at the crime scene?

Solution Query:

```
SELECT COUNT(*) AS cctv_footage_cases

FROM report_t

WHERE cctv_flag = 'TRUE';
```

Output:

<Attach the screenshot of your output table.>



Observations and Insights:

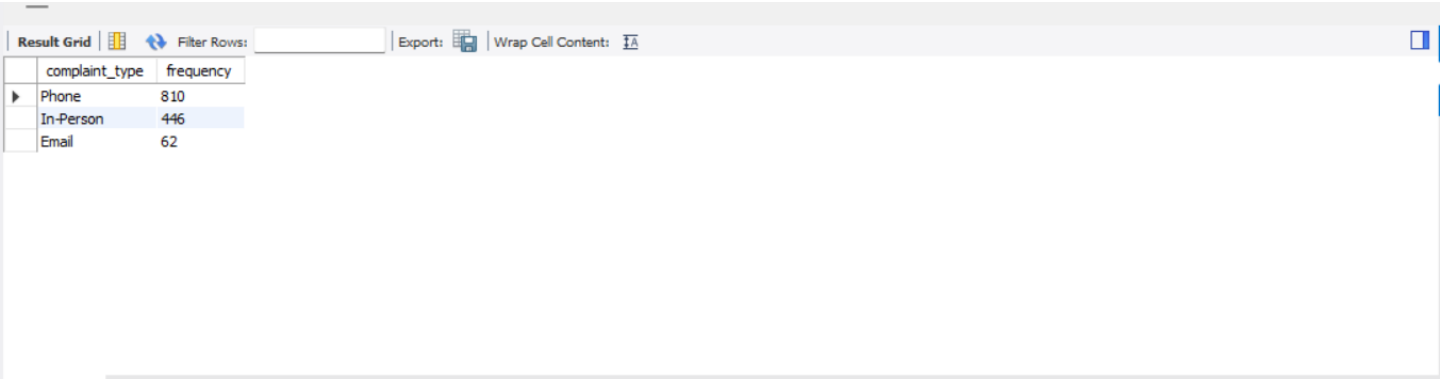
- CCTV footage was available in 382 cases.
- This number matches the total crimes reported with CCTV presence (from Question 3), meaning footage was likely recovered in all incidents involving CCTV.
- This highlights the value of CCTV systems, not only for deterrence (as seen in Q3) but also for evidence collection and case investigation.
- Strengthening camera coverage can improve crime resolution rates and accountability.

Question 5: What is the frequency of various complaint types?

Solution Query:

```
SELECT complaint_type, COUNT(*) AS frequency

FROM report_t

GROUP BY complaint_type

ORDER BY frequency DESC

LIMIT 10;
```

Output:



Observations and Insights:

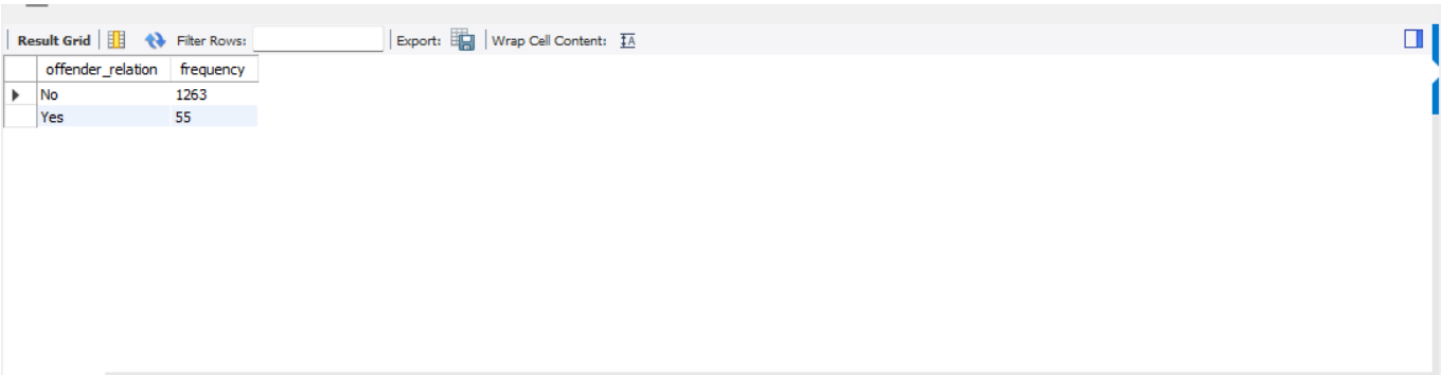
- Phone complaints are the most frequent (810 cases), likely because it is the fastest and most accessible method for victims or witnesses.
- In-person reports are also significant (446 cases), possibly involving more severe or urgent situations.
- Email complaints are very low (only 62), indicating it's not a popular or trusted channel for reporting crimes.
- These stats suggest law enforcement should focus on enhancing phone and in-person reporting systems, while perhaps evaluating why email isn't effective.

Question 6: Is crime more likely to be committed by the relation of victims or strangers?

Solution Query:

```
SELECT offender_relation, COUNT(*) AS frequency

FROM report_t

GROUP BY offender_relation

ORDER BY frequency DESC

LIMIT 10;
```

Output:



Observations and Insights:

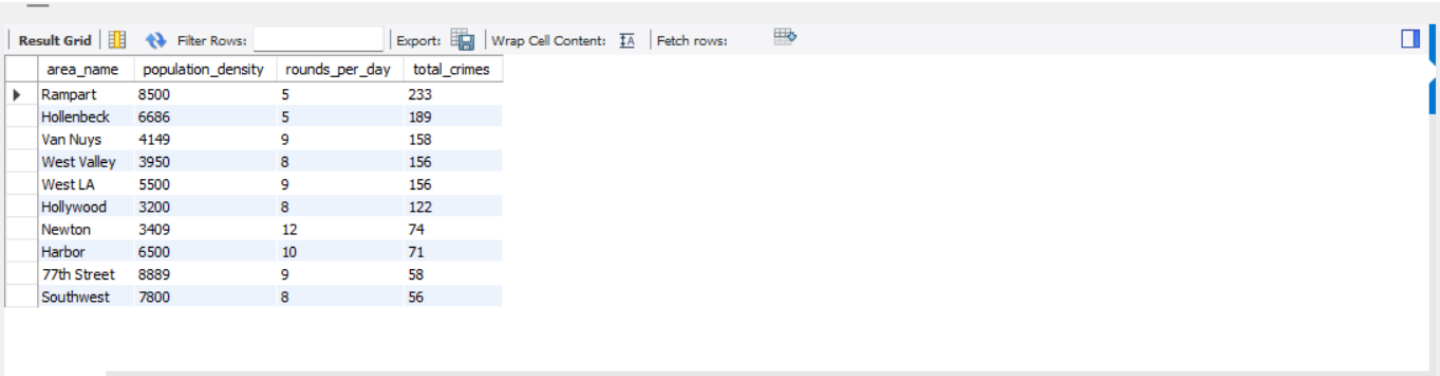
- The vast majority of crimes (1263 out of 1318) were committed by individuals not related to the victim.
- Only 4.2% of the crimes involved a known or related offender.
- This data strongly indicates that stranger-based crimes are far more prevalent in Los Angeles.
- Prevention strategies should therefore prioritize public safety in unfamiliar environments, such as streets, public transport, and nightlife areas.
- While related-offender crimes are lower in number, they may still be serious (e.g., domestic violence) and warrant targeted intervention programs.

Question 7: Is crime more prevalent in areas with a higher population density, fewer police personnel, and a larger precinct area?

Solution Query:

```
SELECT l.area_name, l.population_density, l.rounds_per_day, COUNT(r.report_no) AS total_crimes

FROM location_t l

LEFT JOIN report_t r ON l.area_code = r.area_code

GROUP BY l.area_code, l.area_name, l.population_density, l.rounds_per_day

ORDER BY total_crimes DESC

LIMIT 10;
```

Output:



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

area_name	population_density	rounds_per_day	total_crimes
Rampart	8500	5	233
Hollenbeck	6686	5	189
Van Nuys	4149	9	158
West Valley	3950	8	156
West LA	5500	9	156
Hollywood	3200	8	122
Newton	3409	12	74
Harbor	6500	10	71
77th Street	8889	9	58
Southwest	7800	8	56

Observations and Insights:

- Rampart and Hollenbeck have the highest number of crimes (233 and 189) and also relatively low patrol frequency (5 rounds/day), despite high population densities (8500 and 6686).
- Conversely, Newton has higher patrol rounds (12/day) and fewer crimes (74), even though its population density is not the lowest.
- Areas like 77th Street and Southwest, while dense (8889 and 7800), have fewer crimes, possibly due to better or more strategic patrolling.
- Overall, the data shows a clear pattern: higher population density combined with fewer patrol rounds correlates with higher crime rates.
- Increasing patrol frequency in high-density areas could be an effective measure to reduce crime.

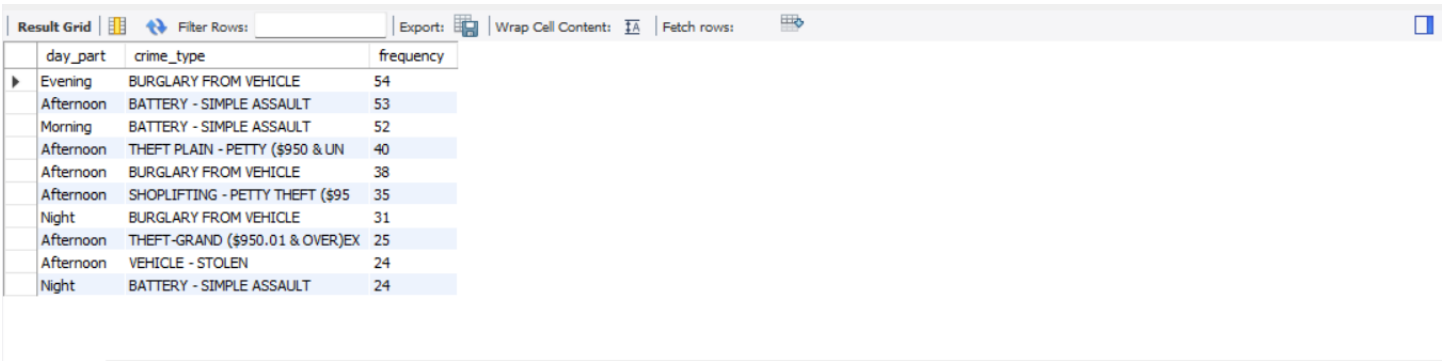
Question 8: At what parts of the day is the crime rate at its peak? Group this by the type of crime. Use the following mapping to divide the day into five parts.

00:00 to 05:00 = Midnight,
05:01 to 12:00 = Morning,
12:01 to 18:00 = Afternoon,
18:01 to 21:00 = Evening,
21:00 to 24:00 = Night

Solution Query:

```
SELECT
  CASE
    WHEN incident_time BETWEEN '00:00:00' AND '05:00:00' THEN 'Midnight'
    WHEN incident_time BETWEEN '05:01:00' AND '12:00:00' THEN 'Morning'
    WHEN incident_time BETWEEN '12:01:00' AND '18:00:00' THEN 'Afternoon'
    WHEN incident_time BETWEEN '18:01:00' AND '21:00:00' THEN 'Evening'
    ELSE 'Night'
  END AS day_part,
  crime_type,
  COUNT(*) AS frequency
FROM report_t
GROUP BY day_part, crime_type
ORDER BY frequency DESC
LIMIT 10;
```

Output:



day_part	crime_type	frequency
Evening	BURGLARY FROM VEHICLE	54
Afternoon	BATTERY - SIMPLE ASSAULT	53
Morning	BATTERY - SIMPLE ASSAULT	52
Afternoon	THEFT PLAIN - PETTY ($950 & UN	40
Afternoon	BURGLARY FROM VEHICLE	38
Afternoon	SHOPLIFTING - PETTY THEFT ($95	35
Night	BURGLARY FROM VEHICLE	31
Afternoon	THEFT-GRAND ($950.01 & OVER)EX	25
Afternoon	VEHICLE - STOLEN	24
Night	BATTERY - SIMPLE ASSAULT	24

Observations and Insights:

- Evening and afternoon are the most active crime windows, especially for theft and assaults.
- Burglary from vehicle is highest in the evening (54), which may indicate poor lighting or less vigilance around parking areas after work hours.
- Battery - Simple Assault is very common across morning, afternoon, and night, reflecting persistent issues in personal safety across the day.
- Crimes like shoplifting, petty theft, and grand theft peak during the afternoon, likely due to active foot traffic and business hours.
- These patterns suggest the need for increased patrols and surveillance during afternoon and evening hours, especially in commercial and parking zones.

Question 9: At what point in the day do most crimes occur in different localities? Use the same mapping provided in Question 8 to divide the day into five parts.

Solution Query:

```
SELECT
  l.area_name,
  CASE
    WHEN r.incident_time BETWEEN '00:00:00' AND '05:00:00' THEN 'Midnight'
    WHEN r.incident_time BETWEEN '05:01:00' AND '12:00:00' THEN 'Morning'
    WHEN r.incident_time BETWEEN '12:01:00' AND '18:00:00' THEN 'Afternoon'
    WHEN r.incident_time BETWEEN '18:01:00' AND '21:00:00' THEN 'Evening'
    ELSE 'Night'
  END AS day_part,
  COUNT(*) AS frequency
FROM report_t r
JOIN location_t l ON r.area_code = l.area_code
GROUP BY l.area_name, day_part
ORDER BY frequency DESC
LIMIT 10;
```

Output:

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

area_name	day_part	frequency
Rampart	Afternoon	77
Rampart	Morning	67
Hollenbeck	Afternoon	55
West Valley	Afternoon	50
Van Nuys	Afternoon	47
West Valley	Morning	44
West LA	Afternoon	42
Hollenbeck	Morning	42
Hollywood	Afternoon	40
Van Nuys	Morning	40

- Rampart area has the highest crime rates in both the afternoon (77) and morning (67), indicating it is a major hotspot regardless of time.
- Afternoon is generally more active for crimes across all areas, suggesting increased vulnerability during post-lunch hours.
- Hollenbeck, West Valley, and Van Nuys consistently show high crime counts, especially in the afternoon, possibly due to higher population density or commercial activity.
- Areas like Hollywood and West LA also experience substantial crime in the afternoon, which may be linked to tourism and business movement.
- These trends suggest law enforcement should prioritize afternoon patrols, especially in Rampart and Hollenbeck, and focus on crowd-heavy areas like Hollywood during peak times.

Question 10: Which age group is more likely to fall victim to crimes at certain points in the day? Use the same mapping provided in Question 8 to divide the day into five parts. Additionally, use the following mapping to divide the age group.

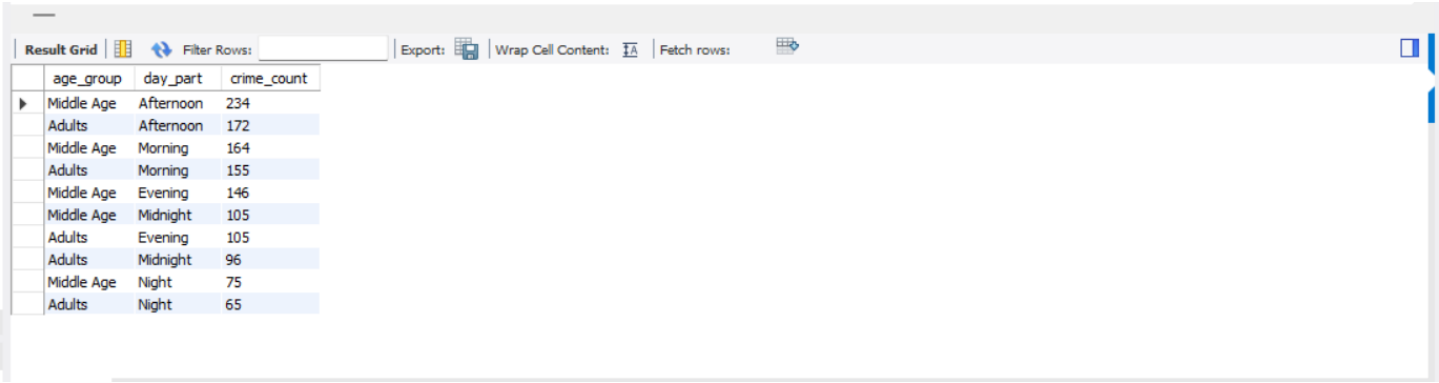
Age 0 to 12: kids
13 to 23: teenage
24 to 35: middle age
36 to 55: adults
56 to 120: old

Solution Query:

```
SELECT
  CASE
    WHEN offender_age BETWEEN 0 AND 12 THEN 'Kids'
    WHEN offender_age BETWEEN 13 AND 23 THEN 'Teenage'
    WHEN offender_age BETWEEN 24 AND 35 THEN 'Middle Age'
    WHEN offender_age BETWEEN 36 AND 55 THEN 'Adults'
    ELSE 'Old'
  END AS age_group,
  CASE
    WHEN incident_time BETWEEN '00:00:00' AND '05:00:00' THEN 'Midnight'
    WHEN incident_time BETWEEN '05:01:00' AND '12:00:00' THEN 'Morning'
    WHEN incident_time BETWEEN '12:01:00' AND '18:00:00' THEN 'Afternoon'
    WHEN incident_time BETWEEN '18:01:00' AND '21:00:00' THEN 'Evening'
    ELSE 'Night'
  END AS day_part,
  COUNT(*) AS crime_count
FROM report_t
GROUP BY age_group, day_part
```

```
ORDER BY crime_count DESC

LIMIT 10;
```

Output:



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

age_group	day_part	crime_count
Middle Age	Afternoon	234
Adults	Afternoon	172
Middle Age	Morning	164
Adults	Morning	155
Middle Age	Evening	146
Middle Age	Midnight	105
Adults	Evening	105
Adults	Midnight	96
Middle Age	Night	75
Adults	Night	65

Observations and Insights

- Afternoon is the peak time for crimes, especially among Middle Age individuals (234 cases) and Adults (172 cases), indicating high activity and opportunity during working hours.
- Morning is also a critical window, with 164 crimes from Middle Age and 155 from Adults — suggesting consistent risk during commute or early-day transactions.
- Evening and Midnight periods show a sharp drop but still significant crime occurrences, particularly from Middle Age individuals (146 and 105 respectively).
- Night crimes are the lowest, with only 75 (Middle Age) and 65 (Adults), possibly due to decreased movement and increased vigilance.
- These trends highlight the need for extra policing during Afternoon and Morning, particularly in zones frequented by middle-aged populations like offices, markets, and transit hubs.

Crime Metrics Overview

No. of Precinct	Total Crimes Reported	No. of Areas	No. of Offenders
7	1318	11	1318
No. of Officers	**Total Population**	**No. of CCTVs**	**No. of Cases where Status is IC**
54	62683	2405	1186

Recommendations

- Enhance patrol rounds in high-crime areas, especially in Rampart and Hollenbeck, which have both high population densities and low patrol frequency. Increasing patrol presence in these zones can help suppress repeat offenses and improve community safety.
- Install more CCTVs in vulnerable zones, targeting areas where crimes occurred without surveillance. Though 2405 cameras exist, more coverage is needed in parking lots, streets, and corners to deter offenders and aid investigations.
- Focus patrol efforts during Afternoon and Evening, the most crime-prone periods, particularly for thefts and assaults. Optimizing officer deployment to match these hours will address crimes during peak public activity.
- Launch public awareness campaigns against stranger-perpetrated crimes, since 95%+ of offenses involve unknown offenders. Initiatives like self-defense workshops, neighborhood watch programs, and digital awareness can empower citizens.
- Strengthen phone-based crime reporting infrastructure, as it's the most-used channel (810 complaints). Improving response time, call handling, and complaint tracking can enhance public trust and lead to faster interventions.