Los Angeles, California 😎. The City of Angels. Tinseltown. The Entertainment Capital of the World!

Known for its warm weather, palm trees, sprawling coastline, and Hollywood, along with producing some of the most iconic films and songs. However, as with any highly populated city, it isn't always glamorous and there can be a large volume of crime. That's where you can help!

You have been asked to support the Los Angeles Police Department (LAPD) by analyzing crime data to identify patterns in criminal behavior. They plan to use your insights to allocate resources effectively to tackle various crimes in different areas.

The Data

They have provided you with a single dataset to use. A summary and preview are provided below.

It is a modified version of the original data, which is publicly available from Los Angeles Open Data.

crimes.csv

Column	Description
`'DR_NO'`	Division of Records Number: Official file number made up of a 2-digit year, area ID, and 5 digits.
`'Date Rptd'`	Date reported - MM/DD/YYYY.
`'DATE OCC'`	Date of occurrence - MM/DD/YYYY.
`'TIME OCC'`	In 24-hour military time.
`'AREA NAME'`	The 21 Geographic Areas or Patrol Divisions are also given a name designation that references a landmark or the surrounding community that it is responsible for. For example, the 77th Street Division is located at the intersection of South Broadway and 77th Street, serving neighborhoods in South Los Angeles.
`'Crm Cd Desc'`	Indicates the crime committed.
`'Vict Age'`	Victim's age in years.
`'Vict Sex'`	Victim's sex: `F`: Female, `M`: Male, `X`: Unknown.
`'Vict Descent'`	Victim's descent: • `A` - Other Asian • `B` - Black • `C` - Chinese • `D` - Cambodian • `F` - Filipino • `G` - Guamanian • `H` - Hispanic/Latin/Mexican • `I` - American Indian/Alaskan Native • `J` - Japanese • `K` - Korean • `L` - Laotian • `O` - Other • `P` - Pacific Islander • `S` - Samoan • `U` - Hawaiian • `V` - Vietnamese • `W` - White • `X` - Unknown • `Z` - Asian Indian
`'Weapon Desc'`	Description of the weapon used (if applicable).
`'Status Desc'`	Crime status.

Column	Description
'LOCATION'	Street address of the crime.

```
# Import required libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
crimes = pd.read_csv("crimes.csv", dtype={"TIME OCC": str})
crimes.head()
```

	D	D…	D…		AR…	Crm Cd Desc			Vict …	We…
4	220213256	2022-07-14	2020-07-14	0900	Rampart	THEFT OF IDENTITY	79	M	B	null
0	220314085	2022-07-22	2020-05-12	1110	Southwest	THEFT OF IDENTITY	27	F	B	null
2	220614831	2022-08-18	2020-08-17	1200	Hollywood	THEFT OF IDENTITY	28	M	H	null
1	222013040	2022-08-06	2020-06-04	1620	Olympic	THEFT OF IDENTITY	60	M	H	null
3	231207725	2023-02-27	2020-01-27	0635	77th Street	THEFT OF IDENTITY	37	M	H	null

Rows: 5 Expand Table

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

#Question 1: Which hour has the highest frequency of crimes? Store as an integer variable called peak_crime_hour.
crimes['Hour OCC']  = crimes['TIME OCC'].str[:2].astype(int) #created new Column, making sure to only take first two digits as those represent the hour, changed type to integer
crimehour= crimes['Hour OCC'].value_counts() #counted values of each hour
print(crimehour)
peak_crime_hour = 12 #selected hour with highest value count from above print


#Question 2: Which area has the largest frequency of night crimes (crimes committed between 10pm and 3:59am)? Save as a string variable called peak_night_crime_location.
crimelate=crimes[crimes['Hour OCC'].isin([22,23,0,1,2,3])] #selected hours where crime was between 10pm(22) and 3:59 am
peak_night_crime_location= crimelate.groupby("AREA NAME",as_index=False)["Hour OCC"].count().sort_values("Hour OCC",ascending=False).iloc[0]["AREA NAME"] #grouped the data by area name and then counted the number of crimes that happened in each hour. After sorting values by the hour in descending order (highest to lowest), I made sure to select the first row using iloc and the column that represented the area of that coluimn to get which area had the most crime at night
print(peak_night_crime_location)


# Question 3: Identify the number of crimes committed against victims of different age groups. Save as a pandas Series called victim_ages, with age group labels "0-17", "18-25", "26-34", "35-44", "45-54", "55-64", and "65+" as the index and the frequency of crimes as the values.
age_bins = [0, 17, 25, 34, 44, 54, 64, np.inf] #created bins list
age_labels = ["0-17", "18-25", "26-34", "35-44", "45-54", "55-64", "65+"] #created labels list
crimes["Age Bracket"] = pd.cut(crimes["Vict Age"],
                               bins=age_bins,
                               labels=age_labels) #added new column using pd.cut() to bin values into discrete intervals
victim_ages = crimes["Age Bracket"].value_counts() #counted values of each occurrence in each age bracket to find frequency
print(victim_ages)
```

```
12    13663
18    10125
17     9964
20     9579
15     9393
19     9262
16     9224
14     8872
11     8787
0      8728
21     8701
22     8531
13     8474
10     8440
8      7523
23     7419
9      7092
1      5836
6      5621
7      5403
2      4726
3      3943
4      3238
5      3171
Name: Hour OCC, dtype: int64
Central
26-34    47470
35-44    42157
45-54    28353
```