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.
`'LOCATION'`	Street address of the crime.

```python
# Re-run this cell
# 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...	Sta...	LOCATION
0	220314085	2022-07-22	2020-05-12	1110	Southwest	THEFT OF IDENTITY	27	F	B	null	Invest Cont	2500 S SYCAMORE AV
1	222013040	2022-08-06	2020-06-04	1620	Olympic	THEFT OF IDENTITY	60	M	H	null	Invest Cont	3300 SAN MARINO ST
2	220614831	2022-08-18	2020-08-17	1200	Hollywood	THEFT OF IDENTITY	28	M	H	null	Invest Cont	1900 TRANSIENT
3	231207725	2023-02-27	2020-01-27	0635	77th Street	THEFT OF IDENTITY	37	M	H	null	Invest Cont	6200 4TH AV
4	220213256	2022-07-14	2020-07-14	0900	Rampart	THEFT OF IDENTITY	79	M	B	null	Invest Cont	1200 W 7TH ST

Rows: 5   ⤢ Expand

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

# Load the data
crimes = pd.read_csv('crimes.csv')
crimes.head()

# Extract the hour from "TIME OCC" (e.g., 2230 → 22)
crimes['hour'] = crimes['TIME OCC'] // 100

# Find the hour with the highest frequency of crimes
peak_crime_hour = crimes['hour'].value_counts().idxmax()

# Store as integer
peak_crime_hour = int(peak_crime_hour)

# Output the result
print("The hour that has the highest frequency of crimes is " + str(peak_crime_hour))

# Use a list for the isin method
peak_night_crime_hours = ["2200", "2300", "2400", "0000", "0100", "0200", "0300"]
peak_night_crime_location = crimes[crimes["TIME OCC"].isin(peak_night_crime_hours)]

# Filter for night crimes: 10:00 PM (2200) to 3:59 AM (0359)
night_crimes = crimes[(crimes["TIME OCC"] >= 2200) | (crimes["TIME OCC"] <= 359)]

# Find the area with the highest number of night crimes
peak_night_crime_location = night_crimes["AREA NAME"].value_counts().idxmax()

# Output the result
print("The area that has the largest frequency of night crimes (crimes committed between 10pm and 3:59am) is " + str(peak_night_crime_location))

# Define age bins and labels
age_bins = [0, 17, 25, 34, 44, 54, 64, float('inf')]
age_labels = ["0-17", "18-25", "26-34", "35-44", "45-54", "55-64", "65+"]

# Bin the victim ages
crimes["age_group"] = pd.cut(crimes["Vict Age"], bins=age_bins, labels=age_labels, right=True)

# Count the number of crimes per age group
victim_ages = crimes["age_group"].value_counts().sort_index()

# Display the result
print(victim_ages)
```

```
The hour that has the highest frequency of crimes is 12
The area that has the largest frequency of night crimes (crimes committed between 10pm and 3:59am) is Central
0-17       4528
18-25     28291
26-34     47470
35-44     42157
45-54     28353
55-64     20169
65+       14747
Name: age_group, dtype: int64
```

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