Cleaning Cam's Data

August 29, 2020

```python
[6]: import numpy as np
     import pandas as pd
     import math
```

```python
[7]: data = pd.read_csv('2020ARRESTDATAusethis-csv.csv')
```

```python
[8]: data
```

```
[8]:        Report ID Report Type   Arrest Date      Time  Area ID    Area Name  \
       0     202111570         RFC   7/9/2020 0:00  1830.0       21      Topanga
       1       5906894     BOOKING  3/25/2020 0:00  2045.0        6    Hollywood
       2     201511771         RFC  6/23/2020 0:00   930.0       15  N Hollywood
       3     200911637         RFC  6/24/2020 0:00  2322.0        9     Van Nuys
       4     200612567         RFC  6/23/2020 0:00  1210.0        6    Hollywood
       ...         ...         ...            ...     ...      ...          ...
       40018   5955800     BOOKING  6/17/2020 0:00   122.0       15  N Hollywood
       40019   5957296     BOOKING  6/19/2020 0:00  1130.0       21      Topanga
       40020   5957111     BOOKING  6/19/2020 0:00   100.0        7     Wilshire
       40021   5955822     BOOKING  6/17/2020 0:00   200.0       16     Foothill
       40022   5955230     BOOKING  6/16/2020 0:00   300.0       12   77th Street

              Reporting District  Age Sex Code Descent Code  ...  Census Tracts  \
       0                    2125   29        M            H  ...          281.0
       1                     645   24        M            W  ...          414.0
       2                    1581   60        F            O  ...          346.0
       3                     901   46        M            H  ...          220.0
       4                     642   27        M            B  ...         2105.0
       ...                   ...  ...      ...          ...  ...            ...
       40018                1591   21        F            W  ...          352.0
       40019                2156   28        M            H  ...          297.0
       40020                 758   25        M            H  ...          648.0
       40021                1698   39        M            W  ...          167.0
       40022                1258   14        F            H  ...          792.0

                             Location Precinct Boundaries   Booking Date  \
       0      POINT (-118.6031 34.2129)               887.0           NaN
       1      POINT (-118.3434 34.1015)              1303.0  3/25/2020 0:00
       2      POINT (-118.4071 34.1478)              1505.0           NaN
```

1

```
3      POINT (-118.4662 34.2085)            1323.0          NaN
4      POINT (-118.3593 34.0944)             NaN            NaN
...                          ...             ...            ...
40018  POINT (-118.4097 34.1447)            1236.0  6/17/2020 0:00
40019  POINT (-118.6059 34.1905)             434.0  6/19/2020 0:00
40020  POINT (-118.3308 34.0482)            1345.0  6/19/2020 0:00
40021   POINT (-118.365 34.2203)            1374.0  6/17/2020 0:00
40022  POINT (-118.2739 33.9718)             998.0  6/16/2020 0:00

       LA Specific Plans Booking Time Council Districts      Booking Location  \
0                   NaN          NaN             4.0                       NaN
1                   NaN       2324.0             8.0                 HOLLYWOOD
2                   NaN          NaN             5.0                       NaN
3                   NaN          NaN             3.0                       NaN
4                   NaN          NaN             NaN                       NaN
...                 ...          ...             ...                       ...
40018               6.0        347.0             5.0  METRO - JAIL DIVISION
40019               2.0       1617.0             4.0        VALLEY - JAIL DIV
40020               NaN        406.0            12.0  METRO - JAIL DIVISION
40021               NaN        512.0             5.0  METRO - JAIL DIVISION
40022               7.0        434.0            13.0                   77TH ST

       Neighborhood Councils (Certified)  Booking Location Code
0                                   37.0                    NaN
1                                   34.0                 4206.0
2                                   84.0                    NaN
3                                   18.0                    NaN
4                                    NaN                    NaN
...                                  ...                    ...
40018                               84.0                 4273.0
40019                               49.0                 4279.0
40020                               52.0                 4273.0
40021                               21.0                 4273.0
40022                               46.0                 4212.0

[40023 rows x 31 columns]
```

[9]: `len(set(data["Charge Description"])) #764 --> 754`

[9]: 754

[10]: `data.dtypes`

[10]:
```
Report ID                        int64
Report Type                     object
Arrest Date                     object
Time                           float64
Area ID                          int64
Area Name                       object
```

```
Reporting District                      int64
Age                                      int64
Sex Code                                object
Descent Code                            object
Charge Group Code                      float64
Charge Group Description               object
Arrest Type Code                        object
Charge                                  object
Charge Description                      object
Disposition Description                 object
Address                                 object
Cross Street                            object
LAT                                    float64
Zip Codes                              float64
LON                                    float64
Census Tracts                          float64
Location                                object
Precinct Boundaries                    float64
Booking Date                            object
LA Specific Plans                      float64
Booking Time                           float64
Council Districts                      float64
Booking Location                        object
Neighborhood Councils (Certified)      float64
Booking Location Code                  float64
dtype: object
```

```python
[11]: data['Descent Code'] = data['Descent Code'].map({'H': 'Hispanic/Latin/Mexican',
                            'A': 'Other Asian',
                            'B': 'Black',
                            'C' : 'Chinese',
                            'D' : 'Cambodian',
                            'F' : 'Filipino',
                            'G' : 'Guamanian',
                            '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'})
```

```
[12]: data['Report Type'] = data['Report Type'].map({
                             'RFC': 'RFC',
                             'BOOKING': 'BOOKED',
                             })
```

```
[13]: #data['Arrest Date'] = pd.to_datetime(data['Arrest Date'])
```

```
[14]: data.head(2)
```

```
[14]:    Report ID Report Type Arrest Date    Time   Area ID   Area Name  \
      0  202111570          RFC  2020-07-09  1830.0        21     Topanga
      1    5906894       BOOKED  2020-03-25  2045.0         6   Hollywood

         Reporting District  Age Sex Code              Descent Code  ...  \
      0                2125   29        M  Hispanic/Latin/Mexican  ...
      1                 645   24        M                   White  ...

         Census Tracts                   Location Precinct Boundaries  \
      0          281.0  POINT (-118.6031 34.2129)               887.0
      1          414.0  POINT (-118.3434 34.1015)              1303.0

           Booking Date LA Specific Plans Booking Time Council Districts  \
      0            NaN               NaN         NaN               4.0
      1  3/25/2020 0:00               NaN      2324.0               8.0

        Booking Location  Neighborhood Councils (Certified)  Booking Location Code
      0             NaN                               37.0                    NaN
      1       HOLLYWOOD                               34.0                 4206.0

      [2 rows x 31 columns]
```

```
[15]: data.dtypes
```

```
[15]: Report ID                              int64
      Report Type                           object
      Arrest Date                   datetime64[ns]
      Time                                 float64
      Area ID                                int64
      Area Name                             object
      Reporting District                     int64
      Age                                    int64
      Sex Code                              object
      Descent Code                          object
      Charge Group Code                    float64
      Charge Group Description              object
      Arrest Type Code                      object
      Charge                                object
      Charge Description                    object
      Disposition Description               object
```

```
Address                              object
Cross Street                         object
LAT                                 float64
Zip Codes                           float64
LON                                 float64
Census Tracts                       float64
Location                             object
Precinct Boundaries                 float64
Booking Date                         object
LA Specific Plans                   float64
Booking Time                        float64
Council Districts                   float64
Booking Location                     object
Neighborhood Councils (Certified)   float64
Booking Location Code               float64
dtype: object
```

[16]:
```python
#set(data['Time'])
```

[17]:
```python
sum(data['Time'].isnull())
```

[17]: 2

[18]:
```python
data['Time'] = data['Time'].fillna(0.0)
```

[19]:
```python
sum(data['Time'].isnull())
```

[19]: 0

[20]:
```python
data['Time'] = data['Time'].apply(lambda x: int(x))
```

[21]:
```python
data['Time'] = data['Time'].apply(lambda x: 0 if x == 2400 else x)
```

[22]:
```python
max(data['Time'])
```

[22]: 2359

[23]:
```python
def time_clean(miltime):
    num = miltime
    if num < 60:
        return pd.to_datetime(num, format='%M').strftime('%H:%M')
    if num > 60 and num < 1000:
        return pd.to_datetime('0' + str(num), format='%H%M').strftime('%H:%M')
    if num >= 1000:
        return pd.to_datetime(num, format='%H%M').strftime('%H:%M')
    else:
        return
```

[24]:
```python
newtime = data['Time'].apply(time_clean)
```

[25]:
```python
data['Time'] = newtime
```

[26]:
```python
data.dtypes
```

```
[26]: Report ID                                  int64
      Report Type                               object
      Arrest Date                       datetime64[ns]
      Time                                      object
      Area ID                                    int64
      Area Name                                 object
      Reporting District                         int64
      Age                                        int64
      Sex Code                                  object
      Descent Code                              object
      Charge Group Code                        float64
      Charge Group Description                  object
      Arrest Type Code                          object
      Charge                                    object
      Charge Description                        object
      Disposition Description                   object
      Address                                   object
      Cross Street                              object
      LAT                                      float64
      Zip Codes                                float64
      LON                                      float64
      Census Tracts                            float64
      Location                                  object
      Precinct Boundaries                      float64
      Booking Date                              object
      LA Specific Plans                        float64
      Booking Time                             float64
      Council Districts                        float64
      Booking Location                          object
      Neighborhood Councils (Certified)        float64
      Booking Location Code                    float64
      dtype: object
```

```
[27]: data
```

```
[27]:         Report ID Report Type Arrest Date    Time  Area ID     Area Name  \
      0       202111570         RFC  2020-07-09   18:30       21      Topanga
      1         5906894      BOOKED  2020-03-25   20:45        6    Hollywood
      2       201511771         RFC  2020-06-23   09:30       15  N Hollywood
      3       200911637         RFC  2020-06-24   23:22        9     Van Nuys
      4       200612567         RFC  2020-06-23   12:10        6    Hollywood
      ...           ...         ...         ...     ...      ...          ...
      40018     5955800      BOOKED  2020-06-17   01:22       15  N Hollywood
      40019     5957296      BOOKED  2020-06-19   11:30       21      Topanga
      40020     5957111      BOOKED  2020-06-19   01:00        7     Wilshire
      40021     5955822      BOOKED  2020-06-17   02:00       16     Foothill
      40022     5955230      BOOKED  2020-06-16   03:00       12   77th Street
```

```
       Reporting District  Age Sex Code           Descent Code  ...  \
0                    2125   29        M  Hispanic/Latin/Mexican  ...
1                     645   24        M                   White  ...
2                    1581   60        F                   Other  ...
3                     901   46        M  Hispanic/Latin/Mexican  ...
4                     642   27        M                   Black  ...
...                   ...  ...      ...                     ...  ...
40018                1591   21        F                   White  ...
40019                2156   28        M  Hispanic/Latin/Mexican  ...
40020                 758   25        M  Hispanic/Latin/Mexican  ...
40021                1698   39        M                   White  ...
40022                1258   14        F  Hispanic/Latin/Mexican  ...

       Census Tracts                  Location Precinct Boundaries  \
0              281.0  POINT (-118.6031 34.2129)                887.0
1              414.0  POINT (-118.3434 34.1015)               1303.0
2              346.0  POINT (-118.4071 34.1478)               1505.0
3              220.0  POINT (-118.4662 34.2085)               1323.0
4             2105.0  POINT (-118.3593 34.0944)                  NaN
...              ...                        ...                  ...
40018          352.0  POINT (-118.4097 34.1447)               1236.0
40019          297.0  POINT (-118.6059 34.1905)                434.0
40020          648.0  POINT (-118.3308 34.0482)               1345.0
40021          167.0   POINT (-118.365 34.2203)               1374.0
40022          792.0  POINT (-118.2739 33.9718)                998.0

         Booking Date LA Specific Plans Booking Time Council Districts  \
0                 NaN               NaN          NaN               4.0
1      3/25/2020 0:00               NaN       2324.0               8.0
2                 NaN               NaN          NaN               5.0
3                 NaN               NaN          NaN               3.0
4                 NaN               NaN          NaN               NaN
...               ...               ...          ...               ...
40018  6/17/2020 0:00               6.0        347.0               5.0
40019  6/19/2020 0:00               2.0       1617.0               4.0
40020  6/19/2020 0:00               NaN        406.0              12.0
40021  6/17/2020 0:00               NaN        512.0               5.0
40022  6/16/2020 0:00               7.0        434.0              13.0

         Booking Location  Neighborhood Councils (Certified)  \
0                     NaN                               37.0
1               HOLLYWOOD                               34.0
2                     NaN                               84.0
3                     NaN                               18.0
4                     NaN                                NaN
...                   ...                                ...
40018  METRO - JAIL DIVISION                             84.0
```

```
40019        VALLEY - JAIL DIV                                    49.0
40020  METRO - JAIL DIVISION                                      52.0
40021  METRO - JAIL DIVISION                                      21.0
40022               77TH ST                                       46.0

       Booking Location Code
0                       NaN
1                    4206.0
2                       NaN
3                       NaN
4                       NaN
...                     ...
40018                4273.0
40019                4279.0
40020                4273.0
40021                4273.0
40022                4212.0

[40023 rows x 31 columns]
```

[28]: `data['Arrest Hour'] = pd.to_datetime(data['Time'],format= '%H:%M' ).dt.hour`

[29]: `data['Arrest Date'].dt.month`

[29]:
```
0        7
1        3
2        6
3        6
4        6
        ..
40018    6
40019    6
40020    6
40021    6
40022    6
Name: Arrest Date, Length: 40023, dtype: int64
```

[30]: `data['Month'] = data['Arrest Date'].dt.month`

[31]: `data.dtypes`

[31]:
```
Report ID                        int64
Report Type                     object
Arrest Date             datetime64[ns]
Time                            object
Area ID                          int64
Area Name                       object
Reporting District               int64
Age                              int64
Sex Code                        object
```

```
Descent Code                                      object
Charge Group Code                                 float64
Charge Group Description                          object
Arrest Type Code                                  object
Charge                                            object
Charge Description                                object
Disposition Description                           object
Address                                           object
Cross Street                                      object
LAT                                               float64
Zip Codes                                         float64
LON                                               float64
Census Tracts                                     float64
Location                                          object
Precinct Boundaries                               float64
Booking Date                                      object
LA Specific Plans                                 float64
Booking Time                                      float64
Council Districts                                 float64
Booking Location                                  object
Neighborhood Councils (Certified)                 float64
Booking Location Code                             float64
Arrest Hour                                        int64
Month                                              int64
dtype: object
```

[32]: ```
data
```

[32]:
```
        Report ID Report Type Arrest Date   Time Area ID   Area Name \
0   202111570   RFC 2020-07-09 18:30   21   Topanga
1   5906894   BOOKED 2020-03-25 20:45   6   Hollywood
2   201511771   RFC 2020-06-23 09:30   15   N Hollywood
3   200911637   RFC 2020-06-24 23:22   9   Van Nuys
4   200612567   RFC 2020-06-23 12:10   6   Hollywood
...   .
40018   5955800   BOOKED 2020-06-17 01:22   15   N Hollywood
40019   5957296   BOOKED 2020-06-19 11:30   21   Topanga
40020   5957111   BOOKED 2020-06-19 01:00   7   Wilshire
40021   5955822   BOOKED 2020-06-17 02:00   16   Foothill
40022   5955230   BOOKED 2020-06-16 03:00   12   77th Street

       Reporting District Age Sex Code   Descent Code ... \
0   2125   29   M Hispanic/Latin/Mexican ...
1   645   24   M   White ...
2   1581   60   F   Other ...
3   901   46   M Hispanic/Latin/Mexican ...
4   642   27   M   Black ...
...   .
```

```
40018                     1591  21     F                   White  ...
40019                     2156  28     M  Hispanic/Latin/Mexican  ...
40020                      758  25     M  Hispanic/Latin/Mexican  ...
40021                     1698  39     M                   White  ...
40022                     1258  14     F  Hispanic/Latin/Mexican  ...

        Precinct Boundaries    Booking Date LA Specific Plans Booking Time  \
0                     887.0             NaN              NaN          NaN
1                    1303.0  3/25/2020 0:00              NaN       2324.0
2                    1505.0             NaN              NaN          NaN
3                    1323.0             NaN              NaN          NaN
4                       NaN             NaN              NaN          NaN
...                     ...             ...              ...          ...
40018                1236.0  6/17/2020 0:00              6.0        347.0
40019                 434.0  6/19/2020 0:00              2.0       1617.0
40020                1345.0  6/19/2020 0:00              NaN        406.0
40021                1374.0  6/17/2020 0:00              NaN        512.0
40022                 998.0  6/16/2020 0:00              7.0        434.0

        Council Districts      Booking Location  \
0                     4.0                   NaN
1                     8.0             HOLLYWOOD
2                     5.0                   NaN
3                     3.0                   NaN
4                     NaN                   NaN
...                   ...                   ...
40018                 5.0  METRO - JAIL DIVISION
40019                 4.0      VALLEY - JAIL DIV
40020                12.0  METRO - JAIL DIVISION
40021                 5.0  METRO - JAIL DIVISION
40022                13.0               77TH ST

        Neighborhood Councils (Certified) Booking Location Code  Arrest Hour  \
0                                    37.0                   NaN           18
1                                    34.0                4206.0           20
2                                    84.0                   NaN            9
3                                    18.0                   NaN           23
4                                     NaN                   NaN           12
...                                   ...                   ...          ...
40018                                84.0                4273.0            1
40019                                49.0                4279.0           11
40020                                52.0                4273.0            1
40021                                21.0                4273.0            2
40022                                46.0                4212.0            3

        Month
0           7
```

```
1           3
2           6
3           6
4           6
...        ...
40018       6
40019       6
40020       6
40021       6
40022       6
```

```
[40023 rows x 33 columns]
```

[33]: `set(data['Sex Code'])`

[33]: `{'F', 'M'}`

[34]:
```python
bureau_dict = {
    "Valley": [21, 17, 19, 16, 10, 9, 15],
    "West": [8, 6, 7, 20, 14],
    "Central": [11, 2, 1, 13, 4],
    "South" : [3, 12, 18, 5]
}
```

[35]:
```python
bureau_names = []
for area_id in data['Area ID']:
    for key in bureau_dict:
        ids = bureau_dict[key]

        if area_id in ids:
            bureau_names.append(key)
```

[36]: `#bureau_names`

[37]: `data['Bureau']= bureau_names`

[38]: `data.head(10)`

[38]:
```
   Report ID Report Type Arrest Date   Time  Area ID     Area Name  \
0  202111570         RFC  2020-07-09  18:30       21       Topanga
1    5906894      BOOKED  2020-03-25  20:45        6     Hollywood
2  201511771         RFC  2020-06-23  09:30       15   N Hollywood
3  200911637         RFC  2020-06-24  23:22        9      Van Nuys
4  200612567         RFC  2020-06-23  12:10        6     Hollywood
5  200612602         RFC  2020-07-10  14:30        6     Hollywood
6  200611612         RFC  2020-06-17  18:45        6     Hollywood
7  201512657         RFC  2020-07-13  09:50       15   N Hollywood
8    5956479      BOOKED  2020-06-18  09:50       14       Pacific
9  202012596         RFC  2020-07-24  12:50       20       Olympic

   Reporting District  Age Sex Code        Descent Code  ...  \
```

```
0               2125   29       M  Hispanic/Latin/Mexican  ...
1                645   24       M                   White  ...
2               1581   60       F                   Other  ...
3                901   46       M  Hispanic/Latin/Mexican  ...
4                642   27       M                   Black  ...
5                666   73       M                   Black  ...
6                646   31       M                   Black  ...
7               1505   46       F                   White  ...
8               1431   44       F                   White  ...
9               2004   24       M  Hispanic/Latin/Mexican  ...

      Booking Date LA Specific Plans Booking Time Council Districts  \
0              NaN              NaN          NaN               4.0
1   3/25/2020 0:00              NaN       2324.0               8.0
2              NaN              NaN          NaN               5.0
3              NaN              NaN          NaN               3.0
4              NaN              NaN          NaN               NaN
5              NaN              NaN          NaN               8.0
6              NaN              NaN          NaN               8.0
7              NaN              NaN          NaN               5.0
8   6/18/2020 0:00             10.0       1108.0              10.0
9              NaN              NaN          NaN               7.0

   Booking Location Neighborhood Councils (Certified) Booking Location Code  \
0              NaN                              37.0                   NaN
1         HOLLYWOOD                              34.0                4206.0
2              NaN                              84.0                   NaN
3              NaN                              18.0                   NaN
4              NaN                               NaN                   NaN
5              NaN                              34.0                   NaN
6              NaN                              82.0                   NaN
7              NaN                              21.0                   NaN
8          PACIFIC                              30.0                4214.0
9              NaN                              89.0                   NaN

   Arrest Hour  Month  Bureau
0           18      7  Valley
1           20      3    West
2            9      6  Valley
3           23      6  Valley
4           12      6    West
5           14      7    West
6           18      6    West
7            9      7  Valley
8            9      6    West
9           12      7    West
```

```
[10 rows x 34 columns]
```

[39]: `data.dtypes`

[39]:
```
Report ID                              int64
Report Type                            object
Arrest Date                            datetime64[ns]
Time                                   object
Area ID                                int64
Area Name                              object
Reporting District                     int64
Age                                    int64
Sex Code                               object
Descent Code                           object
Charge Group Code                      float64
Charge Group Description               object
Arrest Type Code                       object
Charge                                 object
Charge Description                     object
Disposition Description                object
Address                                object
Cross Street                           object
LAT                                    float64
Zip Codes                              float64
LON                                    float64
Census Tracts                          float64
Location                               object
Precinct Boundaries                    float64
Booking Date                           object
LA Specific Plans                      float64
Booking Time                           float64
Council Districts                      float64
Booking Location                       object
Neighborhood Councils (Certified)      float64
Booking Location Code                  float64
Arrest Hour                            int64
Month                                  int64
Bureau                                 object
dtype: object
```

[40]: `data = data.rename(columns = {'Descent Code': 'Ethnicity'})`

[41]: `data`

[41]:
```
      Report ID Report Type Arrest Date   Time  Area ID    Area Name  \
0     202111570         RFC  2020-07-09  18:30       21      Topanga
1       5906894      BOOKED  2020-03-25  20:45        6    Hollywood
2     201511771         RFC  2020-06-23  09:30       15  N Hollywood
3     200911637         RFC  2020-06-24  23:22        9     Van Nuys
```

```
4        200612567        RFC   2020-06-23  12:10        6    Hollywood
...            ...        ...         ...    ...      ...         ...
40018    5955800       BOOKED   2020-06-17  01:22       15  N Hollywood
40019    5957296       BOOKED   2020-06-19  11:30       21      Topanga
40020    5957111       BOOKED   2020-06-19  01:00        7     Wilshire
40021    5955822       BOOKED   2020-06-17  02:00       16     Foothill
40022    5955230       BOOKED   2020-06-16  03:00       12  77th Street

       Reporting District  Age Sex Code              Ethnicity  ...  \
0                    2125   29        M  Hispanic/Latin/Mexican  ...
1                     645   24        M                   White  ...
2                    1581   60        F                   Other  ...
3                     901   46        M  Hispanic/Latin/Mexican  ...
4                     642   27        M                   Black  ...
...                   ...  ...      ...                     ...  ...
40018                1591   21        F                   White  ...
40019                2156   28        M  Hispanic/Latin/Mexican  ...
40020                 758   25        M  Hispanic/Latin/Mexican  ...
40021                1698   39        M                   White  ...
40022                1258   14        F  Hispanic/Latin/Mexican  ...

         Booking Date LA Specific Plans Booking Time Council Districts  \
0                 NaN               NaN          NaN              4.0
1       3/25/2020 0:00              NaN       2324.0              8.0
2                 NaN               NaN          NaN              5.0
3                 NaN               NaN          NaN              3.0
4                 NaN               NaN          NaN              NaN
...               ...               ...          ...              ...
40018   6/17/2020 0:00              6.0        347.0              5.0
40019   6/19/2020 0:00              2.0       1617.0              4.0
40020   6/19/2020 0:00              NaN        406.0             12.0
40021   6/17/2020 0:00              NaN        512.0              5.0
40022   6/16/2020 0:00              7.0        434.0             13.0

          Booking Location Neighborhood Councils (Certified)  \
0                     NaN                                37.0
1               HOLLYWOOD                                34.0
2                     NaN                                84.0
3                     NaN                                18.0
4                     NaN                                 NaN
...                   ...                                 ...
40018  METRO - JAIL DIVISION                              84.0
40019      VALLEY - JAIL DIV                              49.0
40020  METRO - JAIL DIVISION                              52.0
40021  METRO - JAIL DIVISION                              21.0
40022             77TH ST                                46.0
```

```
      Booking Location Code  Arrest Hour  Month  Bureau
0                      NaN           18      7  Valley
1                   4206.0           20      3    West
2                      NaN            9      6  Valley
3                      NaN           23      6  Valley
4                      NaN           12      6    West
...                    ...          ...    ...     ...
40018               4273.0            1      6  Valley
40019               4279.0           11      6  Valley
40020               4273.0            1      6    West
40021               4273.0            2      6  Valley
40022               4212.0            3      6   South

[40023 rows x 34 columns]
```

[42]:
```python
#area names --> zip code of police station
list(set(data['Area Name']))
```

[42]:
```
['Pacific',
 '77th Street',
 'Olympic',
 'Mission',
 'Hollenbeck',
 'Northeast',
 'West Valley',
 'Southwest',
 'Van Nuys',
 'Southeast',
 'Rampart',
 'Wilshire',
 'Topanga',
 'Devonshire',
 'Foothill',
 'West LA',
 'Central',
 'Harbor',
 'Hollywood',
 'Newton',
 'N Hollywood']
```

[43]:
```python
areazip_dict = {'N Hollywood': 91601,
 'Rampart': 90017,
 'Southwest': 90062,
 'West Valley': 91335,
 'Wilshire':  90019,
 'Olympic': 90006,
 'Central': 90014,
 'Devonshire': 91325,
```

```
    'Southeast': 90061,
    'Hollywood': 90028,
    'Hollenbeck':  90033,
    'Mission': 91345,
    'Northeast': 90065,
    'Pacific': 90066,
    'West LA': 90025,
    'Topanga': 91304,
    'Van Nuys': 91401,
    'Harbor':  90731,
    'Newton': 90011,
    'Foothill': 91331,
    '77th Street': 90003}
```

[44]:
```python
zipincome_dict = {91601 : 52607 ,
90017 : 31423,
90062 : 44835,
91335 : 60413,
90019 : 52625,
90006 : 35411,
90014 : 35857,
91325 : 68642,
90061 : 39739,
90028 : 43998,
90033 : 35548,
91345 : 75969,
90065 : 63147,
90066 : 79088,
90025 : 83655,
91304 : 68282,
91401 : 54058,
90731 : 51161,
90011 : 36954,
91331 : 60360,
90003 : 37226
}
```

[45]:
```python
data['Area Zip'] = data['Area Name'].map(areazip_dict)
```

[46]:
```python
data['Zip Median Income'] = data['Area Zip'].map(zipincome_dict)
```

[47]:
```python
data.head(10)
```

[47]:
```
   Report ID Report Type Arrest Date   Time  Area ID    Area Name  \
0  202111570         RFC  2020-07-09  18:30       21      Topanga
1    5906894      BOOKED  2020-03-25  20:45        6    Hollywood
2  201511771         RFC  2020-06-23  09:30       15  N Hollywood
3  200911637         RFC  2020-06-24  23:22        9     Van Nuys
4  200612567         RFC  2020-06-23  12:10        6    Hollywood
```

```
5  200612602      RFC   2020-07-10  14:30      6    Hollywood
6  200611612      RFC   2020-06-17  18:45      6    Hollywood
7  201512657      RFC   2020-07-13  09:50     15  N Hollywood
8    5956479   BOOKED   2020-06-18  09:50     14      Pacific
9  202012596      RFC   2020-07-24  12:50     20      Olympic

   Reporting District  Age Sex Code               Ethnicity  ...  \
0                2125   29       M  Hispanic/Latin/Mexican  ...
1                 645   24       M                   White  ...
2                1581   60       F                   Other  ...
3                 901   46       M  Hispanic/Latin/Mexican  ...
4                 642   27       M                   Black  ...
5                 666   73       M                   Black  ...
6                 646   31       M                   Black  ...
7                1505   46       F                   White  ...
8                1431   44       F                   White  ...
9                2004   24       M  Hispanic/Latin/Mexican  ...

   Booking Time  Council Districts Booking Location  \
0          NaN                4.0              NaN
1       2324.0                8.0         HOLLYWOOD
2          NaN                5.0              NaN
3          NaN                3.0              NaN
4          NaN                NaN              NaN
5          NaN                8.0              NaN
6          NaN                8.0              NaN
7          NaN                5.0              NaN
8       1108.0               10.0           PACIFIC
9          NaN                7.0              NaN

   Neighborhood Councils (Certified) Booking Location Code Arrest Hour Month  \
0                               37.0                   NaN          18     7
1                               34.0                4206.0          20     3
2                               84.0                   NaN           9     6
3                               18.0                   NaN          23     6
4                                NaN                   NaN          12     6
5                               34.0                   NaN          14     7
6                               82.0                   NaN          18     6
7                               21.0                   NaN           9     7
8                               30.0                4214.0           9     6
9                               89.0                   NaN          12     7

   Bureau  Area Zip  Zip Median Income
0  Valley     91304              68282
1    West     90028              43998
2  Valley     91601              52607
3  Valley     91401              54058
```

```
4      West       90028              43998
5      West       90028              43998
6      West       90028              43998
7    Valley       91601              52607
8      West       90066              79088
9      West       90006              35411

[10 rows x 36 columns]
```

[48]: `data.dtypes`

[48]:
```
Report ID                              int64
Report Type                           object
Arrest Date                   datetime64[ns]
Time                                  object
Area ID                                int64
Area Name                             object
Reporting District                     int64
Age                                    int64
Sex Code                              object
Ethnicity                             object
Charge Group Code                    float64
Charge Group Description              object
Arrest Type Code                      object
Charge                                object
Charge Description                    object
Disposition Description               object
Address                               object
Cross Street                          object
LAT                                  float64
Zip Codes                            float64
LON                                  float64
Census Tracts                        float64
Location                              object
Precinct Boundaries                  float64
Booking Date                          object
LA Specific Plans                    float64
Booking Time                         float64
Council Districts                    float64
Booking Location                      object
Neighborhood Councils (Certified)    float64
Booking Location Code                float64
Arrest Hour                            int64
Month                                  int64
Bureau                                object
Area Zip                               int64
Zip Median Income                      int64
dtype: object
```

```python
[49]: #booked  =  data[data["Booking Time"].notnull()]
      #booked
```

```python
[50]: sum(data['Booking Time'].isnull())
```

```
[50]: 8476
```

```python
[51]: data['Booking Time'] = data['Booking Time'].fillna(87654)
```

```python
[52]: sum(data['Booking Time'].isnull())
```

```
[52]: 0
```

```python
[53]: data['Booking Time'] = data['Booking Time'].apply(lambda x: int(x))
```

```python
[54]: max(data['Booking Time'])
```

```
[54]: 87654
```

```python
[55]: def booking_time_clean(btime):
          num = btime
          if num == 87654:
              return 87654
          if num < 60:
              return pd.to_datetime(num, format='%M').strftime('%H:%M')
          if num > 60 and num < 1000:
              return pd.to_datetime('0' + str(num), format='%H%M').strftime('%H:%M')
          if num >= 1000:
              return pd.to_datetime(num, format='%H%M').strftime('%H:%M')
          else:
              return
```

```python
[56]: newbookingtime = data['Booking Time'].apply(booking_time_clean)
```

```python
[57]: newbookingtime
```

```
[57]: 0        87654
      1        23:24
      2        87654
      3        87654
      4        87654
               ...
      40018    03:47
      40019    16:17
      40020    04:06
      40021    05:12
      40022    04:34
      Name: Booking Time, Length: 40023, dtype: object
```

```python
[60]: booking = []
      for time in newbookingtime:
          if time == 87654:
              booking.append(None)
```

```
        else:
            booking.append(time)
```

```
[61]: booking
```

```
[61]: [None,
 '23:24',
 None,
 None,
 None,
 None,
 None,
 None,
 '11:08',
 None,
 '00:16',
 None,
 None,
 '19:25',
 None,
 None,
 '22:12',
 '03:17',
 None,
 '22:13',
 None,
 None,
 '23:40',
 '05:29',
 '21:44',
 '18:55',
 '05:32',
 '20:47',
 '18:57',
 '10:27',
 '15:32',
 '05:35',
 '14:16',
 '03:50',
 '19:05',
 '01:09',
 '23:34',
 '01:55',
 '01:44',
 '13:38',
 '18:44',
 '03:35',
 '23:39',
```

```
'15:12',
'09:49',
'11:55',
'03:45',
'01:58',
'12:09',
'03:05',
'04:23',
'01:46',
'13:24',
'21:31',
'13:07',
'00:15',
'02:43',
'23:38',
'20:18',
'15:45',
'03:55',
'03:10',
'21:08',
'21:33',
'17:49',
None,
'15:11',
'02:38',
'22:56',
'14:43',
'03:16',
'21:24',
'00:33',
'11:08',
'21:33',
'00:16',
'04:35',
'23:24',
None,
'21:16',
'23:02',
'04:12',
'05:14',
'10:38',
'10:44',
'20:53',
'19:40',
'05:28',
None,
'21:36',
```

```
'04:56',
'08:24',
'04:37',
'12:28',
'16:57',
'23:45',
'21:31',
'00:22',
'20:05',
None,
'19:24',
'07:37',
'02:09',
'02:04',
'21:12',
'10:33',
'12:23',
'09:58',
'21:19',
'22:21',
'11:20',
'20:30',
'15:15',
'01:04',
'02:48',
'02:21',
'02:04',
'17:24',
'20:47',
'12:02',
'22:15',
'02:25',
'01:36',
'10:46',
'05:50',
'02:56',
'17:34',
'18:48',
'19:51',
'03:51',
'17:58',
'03:15',
'02:12',
'16:30',
'06:25',
'13:49',
'15:47',
```

```
'14:50',
'01:53',
'04:49',
None,
'22:06',
'21:19',
'23:01',
'04:59',
'22:50',
'10:56',
'18:12',
'09:32',
'01:38',
'07:01',
'22:51',
'05:06',
'17:09',
'20:52',
'15:30',
'00:38',
'03:45',
'19:40',
'01:57',
'17:13',
'14:27',
'15:59',
'22:09',
'02:00',
'15:24',
None,
'23:09',
'02:42',
'23:01',
'21:27',
'13:56',
'00:13',
'19:44',
'17:45',
'16:06',
'12:23',
None,
'13:06',
'03:09',
'16:34',
'15:31',
None,
None,
```

```
None,
'01:16',
'00:37',
'04:35',
'07:12',
None,
None,
'05:19',
'22:24',
None,
'23:33',
'06:32',
'04:41',
'02:10',
'21:26',
'18:12',
'10:36',
'10:03',
'23:38',
'21:36',
'02:31',
'20:12',
None,
'17:35',
'15:21',
'17:47',
None,
'04:12',
'01:24',
'07:35',
'12:19',
'04:03',
None,
'11:41',
'15:42',
'14:56',
'03:48',
'14:00',
'19:26',
'10:00',
'03:57',
None,
'15:08',
'14:49',
'15:27',
'07:19',
None,
```

'07:25',
'13:17',
'03:05',
'11:50',
None,
None,
'17:39',
'16:56',
'22:31',
'20:51',
None,
'03:30',
'04:12',
'05:33',
'08:37',
'08:29',
'18:24',
'20:46',
'00:20',
'04:23',
'15:04',
'05:33',
'20:52',
'20:59',
None,
'13:27',
'01:39',
None,
'02:53',
'16:06',
'22:52',
'15:15',
'13:31',
'22:43',
'19:51',
None,
'14:48',
'00:13',
'13:31',
'16:38',
None,
'14:33',
'23:44',
'06:18',
'16:00',
'21:40',
'17:22',

```
'06:55',
None,
'20:36',
None,
None,
None,
'23:28',
'20:45',
'17:38',
'23:40',
'16:36',
'19:44',
None,
'13:28',
'16:12',
'10:38',
None,
'23:16',
'06:54',
'22:17',
'19:19',
'19:50',
'23:40',
'03:46',
'11:06',
'20:43',
'18:51',
None,
'22:06',
'03:49',
'14:13',
'04:26',
None,
'19:21',
'15:06',
'11:40',
None,
'17:55',
'22:52',
'13:37',
None,
'13:49',
'03:17',
'15:30',
'23:30',
None,
'15:50',
```

```
'15:02',
'18:30',
'00:06',
None,
'09:35',
'01:17',
'01:07',
'21:37',
'11:19',
'16:06',
'14:14',
'05:05',
'20:54',
'18:22',
'02:24',
'05:56',
None,
'16:43',
'04:43',
'05:08',
'01:12',
'02:35',
'01:27',
'17:30',
'00:08',
'12:22',
'12:13',
'20:08',
'02:05',
None,
'22:08',
None,
'18:41',
'04:53',
'18:47',
'00:17',
'12:18',
None,
'11:33',
'08:43',
'07:36',
'18:51',
'17:53',
'23:00',
None,
'23:52',
'22:03',
```

'13:33',
'00:01',
'14:27',
'09:09',
'10:51',
'01:42',
'19:31',
'02:16',
'02:49',
'22:54',
'20:48',
'03:05',
None,
None,
'17:02',
None,
'00:37',
None,
None,
'16:56',
'04:38',
'20:03',
'09:45',
'01:46',
'19:23',
'01:55',
None,
'10:17',
'00:16',
'10:08',
'14:39',
'08:04',
'02:12',
'04:48',
'20:09',
'17:48',
'22:26',
'14:42',
'21:18',
'10:12',
'03:21',
'22:38',
'00:57',
'13:01',
'01:05',
'20:55',
None,

```
'00:29',
'19:33',
'16:13',
'22:44',
'07:06',
'03:16',
'05:55',
'15:29',
'07:14',
'05:53',
'12:23',
'19:41',
'21:45',
'07:56',
None,
None,
'11:20',
'20:19',
'05:57',
'22:49',
None,
'02:56',
'10:05',
None,
'16:09',
'12:55',
'09:37',
'09:11',
'16:00',
'14:57',
'23:18',
'23:39',
None,
'00:54',
'02:29',
'16:12',
'12:40',
'09:40',
'01:09',
None,
'04:04',
'02:25',
'04:13',
None,
'02:47',
'14:45',
'00:42',
```

```
'22:01',
None,
'04:07',
'02:16',
'04:46',
'22:54',
'23:46',
'04:59',
'22:57',
'01:32',
'10:20',
'18:58',
'12:27',
None,
'23:42',
None,
'00:27',
None,
None,
'11:01',
'21:07',
'15:18',
'15:20',
'14:53',
'18:05',
None,
'03:41',
'01:42',
'13:23',
None,
'18:19',
'02:42',
None,
'22:47',
'15:07',
'11:09',
'19:57',
'00:19',
'00:06',
'09:36',
'07:51',
'23:08',
'13:27',
'10:36',
'03:45',
'21:57',
'06:06',
```

```
'00:44',
None,
'23:14',
'13:53',
'12:07',
'01:18',
'00:15',
'09:37',
'11:46',
'08:19',
'19:42',
'10:11',
'11:19',
'09:15',
'18:03',
'02:40',
'05:55',
'19:44',
None,
'13:25',
'15:28',
'23:11',
'23:49',
'16:00',
'07:10',
'12:19',
'22:02',
'22:46',
'19:33',
None,
'11:35',
'23:14',
'21:50',
'01:22',
None,
'01:11',
'23:55',
None,
'01:28',
'04:45',
'16:24',
'11:46',
'00:21',
'23:23',
'03:04',
'23:41',
'19:32',
```

```
'20:39',
None,
'23:24',
'02:10',
'00:22',
None,
'19:15',
'02:17',
'00:36',
'22:17',
'22:01',
'14:15',
'15:01',
'15:03',
'20:08',
'12:11',
'20:19',
None,
None,
'17:12',
'02:33',
'09:39',
'00:06',
'05:48',
'12:42',
'04:12',
'19:45',
'07:12',
'14:56',
'19:43',
'02:19',
'10:52',
'02:11',
'00:01',
None,
'23:28',
None,
'21:53',
None,
'15:06',
'23:15',
'16:09',
'01:02',
'19:06',
'22:06',
'12:29',
'22:53',
```

'01:57',
'19:28',
'23:46',
'04:19',
'19:47',
'22:15',
'00:25',
'20:49',
'03:40',
'00:58',
'17:18',
'18:48',
'21:55',
'20:46',
'19:30',
'02:32',
'20:12',
'04:33',
None,
'01:42',
'00:23',
'03:39',
'05:14',
'04:44',
'23:38',
'20:37',
'04:10',
'23:47',
'12:19',
'20:34',
'10:06',
'21:12',
'19:58',
'12:27',
'20:40',
None,
'21:59',
'01:02',
'04:25',
'16:09',
None,
'02:15',
'20:58',
None,
'23:27',
'20:40',
'11:40',

```
'05:10',
'17:38',
'16:19',
'11:40',
'08:31',
'22:19',
'01:01',
'06:03',
'16:35',
'18:34',
'01:36',
None,
None,
'00:41',
None,
'18:31',
'22:40',
'13:31',
'11:55',
'15:17',
'20:03',
'12:03',
'17:20',
'22:22',
None,
None,
'23:31',
'16:32',
'02:43',
'00:19',
'18:50',
'09:15',
'07:05',
'04:30',
'02:42',
'20:47',
'12:24',
'19:41',
'12:42',
'21:01',
'06:03',
None,
'20:55',
None,
'20:18',
'01:50',
'01:13',
```

```
'13:54',
'23:09',
'02:51',
'16:16',
None,
None,
'13:08',
'13:25',
'20:18',
'01:40',
'07:02',
'22:29',
None,
'03:19',
'23:32',
None,
'09:13',
'07:40',
'22:29',
'02:37',
None,
'14:41',
'21:09',
'11:28',
'02:44',
'11:59',
'01:50',
'18:56',
'11:28',
'23:11',
'15:11',
'05:26',
'10:23',
'23:13',
'05:47',
'15:31',
'05:43',
'19:40',
None,
'11:21',
'16:38',
'17:38',
'21:23',
None,
'23:19',
'08:45',
'00:01',
```

```
'14:37',
'03:32',
'00:45',
'03:55',
'01:49',
None,
'01:40',
'00:26',
'02:09',
'14:33',
None,
'14:44',
'13:44',
None,
'00:05',
'22:12',
'02:14',
'09:11',
'11:04',
'23:45',
'04:26',
'13:57',
'06:57',
'02:24',
'05:25',
'21:31',
None,
None,
'15:14',
'17:40',
None,
None,
'14:39',
'00:43',
None,
'12:28',
'07:39',
'13:58',
'17:00',
'23:10',
None,
'05:40',
'02:12',
None,
'07:24',
'00:40',
'20:09',
```

```
'10:24',
'07:06',
'02:15',
'15:16',
'02:14',
None,
'21:54',
'07:44',
'00:14',
'09:19',
None,
'15:16',
'21:06',
'15:49',
'20:14',
'12:21',
'06:40',
'17:17',
'09:04',
'01:03',
'18:53',
None,
'16:03',
None,
'10:39',
'01:18',
'01:07',
'14:42',
'11:58',
None,
'18:01',
'21:34',
'19:06',
'04:14',
'12:10',
'16:25',
'04:42',
'02:40',
'01:06',
'04:22',
'21:51',
'20:36',
'03:39',
'18:59',
'12:26',
'00:43',
'00:44',
```

```
'13:17',
'15:33',
'08:01',
None,
'17:06',
'20:16',
'02:20',
'15:41',
None,
'05:28',
'09:13',
'23:23',
'17:41',
'00:01',
'14:16',
'01:00',
'21:22',
'02:14',
'13:20',
None,
None,
'16:11',
'10:28',
'02:58',
None,
'19:47',
'14:02',
'16:05',
'14:31',
'08:33',
'15:52',
'14:53',
None,
'09:59',
None,
'00:17',
'22:26',
None,
'07:34',
'23:22',
'00:12',
'18:52',
'20:41',
'07:40',
'02:57',
'20:14',
'11:15',
```

```
'23:49',
'16:00',
'14:35',
'18:36',
'11:20',
'23:50',
'23:51',
'15:29',
None,
None,
'10:13',
'23:12',
None,
'03:47',
'21:35',
'16:05',
'17:25',
'00:09',
'04:07',
'05:06',
'12:02',
None,
'14:49',
'13:17',
'21:01',
'04:00',
'21:32',
'23:05',
'20:13',
None,
'04:51',
'19:02',
'11:36',
'12:14',
'09:45',
'21:17',
'08:03',
'11:45',
'14:06',
None,
None,
None,
'22:18',
'23:39',
None,
'17:49',
'18:12',
```

```
'07:06',
'01:27',
'06:57',
'23:29',
'10:20',
'17:13',
None,
'03:43',
'02:34',
'21:49',
'04:05',
'10:04',
'12:56',
'09:30',
'11:38',
'05:48',
None,
'13:59',
'07:55',
'22:44',
'22:00',
'06:59',
'16:07',
'17:43',
'12:59',
'23:45',
None,
'13:41',
'08:05',
'21:35',
None,
'14:43',
'14:17',
'20:05',
'04:14',
'16:11',
'22:23',
'06:05',
'13:26',
'20:14',
'01:31',
None,
'15:18',
'15:14',
'00:55',
'08:12',
'23:08',
```

```
    '23:12',
    '11:52',
    '20:35',
    '13:28',
    '11:32',
    '00:08',
    '08:24',
    None,
    '23:33',
    '01:02',
    '17:00',
    '05:08',
    '21:12',
    '20:41',
    '22:33',
    '00:47',
    '01:11',
    ...]
```

[62]: `data['Booking Time'] = booking`

[63]: `data.head(20)`

[63]:
```
    Report ID Report Type Arrest Date   Time  Area ID    Area Name   \
0   202111570         RFC  2020-07-09  18:30       21      Topanga
1     5906894      BOOKED  2020-03-25  20:45        6    Hollywood
2   201511771         RFC  2020-06-23  09:30       15  N Hollywood
3   200911637         RFC  2020-06-24  23:22        9     Van Nuys
4   200612567         RFC  2020-06-23  12:10        6    Hollywood
5   200612602         RFC  2020-07-10  14:30        6    Hollywood
6   200611612         RFC  2020-06-17  18:45        6    Hollywood
7   201512657         RFC  2020-07-13  09:50       15  N Hollywood
8     5956479      BOOKED  2020-06-18  09:50       14      Pacific
9   202012596         RFC  2020-07-24  12:50       20      Olympic
10    5973266      BOOKED  2020-07-16  20:00        7     Wilshire
11  200911454         RFC  2020-06-23  15:15        9     Van Nuys
12  201111323         RFC  2020-07-12  03:15       11    Northeast
13    5973109      BOOKED  2020-07-16  17:10       14      Pacific
14  200310427         RFC  2020-04-25  00:40        3    Southwest
15  201512885         RFC  2020-07-11  21:00       15  N Hollywood
16    5977505      BOOKED  2020-07-23  17:20       19      Mission
17    5960953      BOOKED  2020-06-26  01:00       13       Newton
18  201111894         RFC  2020-07-16  22:15       11    Northeast
19    5969081      BOOKED  2020-07-09  19:50       20      Olympic

    Reporting District  Age Sex Code              Ethnicity  ...  \
0                 2125   29        M  Hispanic/Latin/Mexican  ...
1                  645   24        M                   White  ...
```

41

```
2                  1581   60      F                 Other  ...
3                   901   46      M  Hispanic/Latin/Mexican  ...
4                   642   27      M                 Black  ...
5                   666   73      M                 Black  ...
6                   646   31      M                 Black  ...
7                  1505   46      F                 White  ...
8                  1431   44      F                 White  ...
9                  2004   24      M  Hispanic/Latin/Mexican  ...
10                  767   31      M  Hispanic/Latin/Mexican  ...
11                  941   36      M                 Black  ...
12                 1151   33      M  Hispanic/Latin/Mexican  ...
13                 1494   21      M                 Black  ...
14                  373   24      M                 Black  ...
15                 1544   32      M  Hispanic/Latin/Mexican  ...
16                 1999   42      M  Hispanic/Latin/Mexican  ...
17                 1333   28      M  Hispanic/Latin/Mexican  ...
18                 1151   30      M                 Other  ...
19                 2015   27      F  Hispanic/Latin/Mexican  ...

    Booking Time Council Districts        Booking Location  \
0           None               4.0                     NaN
1          23:24               8.0               HOLLYWOOD
2           None               5.0                     NaN
3           None               3.0                     NaN
4           None               NaN                     NaN
5           None               8.0                     NaN
6           None               8.0                     NaN
7           None               5.0                     NaN
8          11:08              10.0                 PACIFIC
9           None               7.0                     NaN
10         00:16              12.0                 77TH ST
11          None               7.0                     NaN
12          None               8.0                     NaN
13         19:25              10.0                 PACIFIC
14          None              12.0                     NaN
15          None               7.0                     NaN
16         22:12               3.0        VALLEY - JAIL DIV
17         03:17              13.0   METRO - JAIL DIVISION
18          None               8.0                     NaN
19         22:13               8.0               HOLLYWOOD

    Neighborhood Councils (Certified) Booking Location Code Arrest Hour Month  \
0                                37.0                   NaN           18     7
1                                34.0                4206.0           20     3
2                                84.0                   NaN            9     6
3                                18.0                   NaN           23     6
4                                 NaN                   NaN           12     6
```

5	34.0	NaN	14	7
6	82.0	NaN	18	6
7	21.0	NaN	9	7
8	30.0	4214.0	9	6
9	89.0	NaN	12	7
10	19.0	4212.0	20	7
11	18.0	NaN	15	6
12	58.0	NaN	3	7
13	16.0	4214.0	17	7
14	54.0	NaN	0	4
15	NaN	NaN	21	7
16	59.0	4279.0	17	7
17	6.0	4273.0	1	6
18	58.0	NaN	22	7
19	89.0	4206.0	19	7

	Bureau	Area Zip	Zip Median Income
0	Valley	91304	68282
1	West	90028	43998
2	Valley	91601	52607
3	Valley	91401	54058
4	West	90028	43998
5	West	90028	43998
6	West	90028	43998
7	Valley	91601	52607
8	West	90066	79088
9	West	90006	35411
10	West	90019	52625
11	Valley	91401	54058
12	Central	90065	63147
13	West	90066	79088
14	South	90062	44835
15	Valley	91601	52607
16	Valley	91345	75969
17	Central	90011	36954
18	Central	90065	63147
19	West	90006	35411

[20 rows x 36 columns]

[66]: `data.dtypes`

[66]:
```
Report ID                   int64
Report Type                object
Arrest Date                object
Time                       object
Area ID                     int64
Area Name                  object
```

```
Reporting District                      int64
Age                                      int64
Sex Code                                object
Ethnicity                               object
Charge Group Code                      float64
Charge Group Description                object
Arrest Type Code                        object
Charge                                  object
Charge Description                      object
Disposition Description                 object
Address                                 object
Cross Street                            object
LAT                                    float64
Zip Codes                              float64
LON                                    float64
Census Tracts                          float64
Location                                object
Precinct Boundaries                    float64
Booking Date                            object
LA Specific Plans                      float64
Booking Time                            object
Council Districts                      float64
Booking Location                        object
Neighborhood Councils (Certified)      float64
Booking Location Code                  float64
Arrest Hour                              int64
Month                                    int64
Bureau                                  object
Area Zip                                 int64
Zip Median Income                        int64
dtype: object
```

[65]:
```python
data['Arrest Date'] = data['Arrest Date'].astype(str)
```

[68]:
```python
data['Arrest Date and Time'] = pd.to_datetime(data['Arrest Date'] + ' ' +
 →data['Time'])
```

[74]:
```python
data['Booking Date and Time'] =pd.to_datetime(data['Booking Date'] + ' ' +
 →data['Booking Time'])
```

[78]:
```python
data['Time Difference'] = data['Booking Date and Time'] - data['Arrest Date and
 →Time']
```

[81]:
```python
data.head()
```

[81]:
```
   Report ID Report Type Arrest Date   Time  Area ID     Area Name  \
0  202111570         RFC  2020-07-09  18:30       21       Topanga
1    5906894      BOOKED  2020-03-25  20:45        6     Hollywood
2  201511771         RFC  2020-06-23  09:30       15   N Hollywood
3  200911637         RFC  2020-06-24  23:22        9      Van Nuys
```

```
4   200612567         RFC   2020-06-23  12:10          6      Hollywood


    Reporting District   Age Sex Code             Ethnicity  ...  \
0                 2125    29       M  Hispanic/Latin/Mexican  ...
1                  645    24       M                   White  ...
2                 1581    60       F                   Other  ...
3                  901    46       M  Hispanic/Latin/Mexican  ...
4                  642    27       M                   Black  ...


    Neighborhood Councils (Certified)  Booking Location Code  Arrest Hour  Month  \
0                               37.0                    NaN           18      7
1                               34.0                 4206.0           20      3
2                               84.0                    NaN            9      6
3                               18.0                    NaN           23      6
4                                NaN                    NaN           12      6


    Bureau Area Zip  Zip Median Income    Arrest Date and Time  \
0   Valley   91304              68282     2020-07-09 18:30:00
1     West   90028              43998     2020-03-25 20:45:00
2   Valley   91601              52607     2020-06-23 09:30:00
3   Valley   91401              54058     2020-06-24 23:22:00
4     West   90028              43998     2020-06-23 12:10:00


    Booking Date and Time  Time Difference
0                    NaT              NaT
1    2020-03-25 23:24:00         02:39:00
2                    NaT              NaT
3                    NaT              NaT
4                    NaT              NaT

[5 rows x 39 columns]
```

[84]: `data`

```
[84]:          Report ID Report Type Arrest Date    Time   Area ID     Area Name  \
       0      202111570         RFC  2020-07-09  18:30        21       Topanga
       1        5906894      BOOKED  2020-03-25  20:45         6     Hollywood
       2      201511771         RFC  2020-06-23  09:30        15   N Hollywood
       3      200911637         RFC  2020-06-24  23:22         9      Van Nuys
       4      200612567         RFC  2020-06-23  12:10         6     Hollywood
       ...          ...         ...         ...    ...       ...           ...
       40018    5955800      BOOKED  2020-06-17  01:22        15   N Hollywood
       40019    5957296      BOOKED  2020-06-19  11:30        21       Topanga
       40020    5957111      BOOKED  2020-06-19  01:00         7      Wilshire
       40021    5955822      BOOKED  2020-06-17  02:00        16       Foothill
       40022    5955230      BOOKED  2020-06-16  03:00        12   77th Street


              Reporting District  Age Sex Code             Ethnicity  \
```

```
0                          2125   29         M  Hispanic/Latin/Mexican
1                           645   24         M                   White
2                          1581   60         F                   Other
3                           901   46         M  Hispanic/Latin/Mexican
4                           642   27         M                   Black
...                         ...   ...      ...                     ...
40018                      1591   21         F                   White
40019                      2156   28         M  Hispanic/Latin/Mexican
40020                       758   25         M  Hispanic/Latin/Mexican
40021                      1698   39         M                   White
40022                      1258   14         F  Hispanic/Latin/Mexican

       Charge Group Code         Charge Group Description Arrest Type Code  \
0                   17.0                      Liquor Laws                I
1                    4.0               Aggravated Assault                F
2                    NaN                              NaN                I
3                    NaN                              NaN                M
4                    NaN                              NaN                I
...                  ...                              ...              ...
40018               22.0           Driving Under Influence               M
40019               24.0  Miscellaneous Other Violations               M
40020               16.0               Narcotic Drug Laws               M
40021               23.0           Moving Traffic Violations             F
40022                5.0                         Burglary                F

               Charge                   Charge Description  \
0          25620(A)BP                      DRINKING ALCOHOL
1         273.5(A)PC  CORPORAL INJURY ON SPOUSE/COHABITANT/ETC
2             6344D4                                      NaN
3           1203.2PC                                      NaN
4         41.24(A)LAM                                     NaN
...              ...                                      ...
40018     23152(A)VC           DRUNK DRIVING ALCOHOL/DRUGS
40019     594(B)2APC           DEFACE/DAMAGE PROPERTY <$400
40020     11350(A)HS         POSSESSION OF CONTROLLED SUBSTANCE
40021      2800.2VC         EVADING PO/DRIVE IN RECKLESS MANNER
40022         459PC                                 BURGLARY

        Disposition Description                              Address  \
0       MISDEMEANOR COMPLAINT FILED  21900     ELKWOOD              ST
1               CITY ATTORNEY REJECT  1700 N  SYCAMORE              AV
2       MISDEMEANOR COMPLAINT FILED          LAUREL CANYON
3       MISDEMEANOR COMPLAINT FILED          SEPULVEDA             BL
4       MISDEMEANOR COMPLAINT FILED  1300 N  OGDEN                DR
...                             ...                              ...
40018   MISDEMEANOR COMPLAINT FILED          VENTURA
40019   MISDEMEANOR COMPLAINT FILED  6600    TOPANGA CANYON        BL
```

```
40020    MISDEMEANOR COMPLAINT FILED          PICO                          BL
40021          FELONY COMPLAINT FILED                           CROCKETT
40022  OTHER LAW ENFORCEMENT AGENCY    100 W   75TH                         ST
```

	Cross Street	LAT	Zip Codes	LON	Census Tracts
0	NaN	34.2129	4279.0	-118.6031	281.0
1	NaN	34.1015	23446.0	-118.3434	414.0
2	VENTURA	34.1478	8492.0	-118.4071	346.0
3	SATICOY	34.2085	19733.0	-118.4662	220.0
4	NaN	34.0944	23677.0	-118.3593	2105.0
...	...	...	...	...	...
40018	FAIRWAY	34.1447	8492.0	-118.4097	352.0
40019	NaN	34.1905	4278.0	-118.6059	297.0
40020	PLYMOUTH	34.0482	23080.0	-118.3308	648.0
40021	CLYBOURN	34.2203	19335.0	-118.3650	167.0
40022	NaN	33.9718	22352.0	-118.2739	792.0

	Location	Precinct Boundaries	Booking Date
0	POINT (-118.6031 34.2129)	887.0	NaN
1	POINT (-118.3434 34.1015)	1303.0	3/25/2020 0:00
2	POINT (-118.4071 34.1478)	1505.0	NaN
3	POINT (-118.4662 34.2085)	1323.0	NaN
4	POINT (-118.3593 34.0944)	NaN	NaN
...	...	...	...
40018	POINT (-118.4097 34.1447)	1236.0	6/17/2020 0:00
40019	POINT (-118.6059 34.1905)	434.0	6/19/2020 0:00
40020	POINT (-118.3308 34.0482)	1345.0	6/19/2020 0:00
40021	POINT (-118.365 34.2203)	1374.0	6/17/2020 0:00
40022	POINT (-118.2739 33.9718)	998.0	6/16/2020 0:00

	LA Specific Plans	Booking Time	Council Districts
0	NaN	None	4.0
1	NaN	23:24	8.0
2	NaN	None	5.0
3	NaN	None	3.0
4	NaN	None	NaN
...	...	...	...
40018	6.0	03:47	5.0
40019	2.0	16:17	4.0
40020	NaN	04:06	12.0
40021	NaN	05:12	5.0
40022	7.0	04:34	13.0

	Booking Location	Neighborhood Councils (Certified)
0	NaN	37.0
1	HOLLYWOOD	34.0
2	NaN	84.0

```
3                       NaN                                    18.0
4                       NaN                                    NaN
...                     ...                                    ...
40018  METRO - JAIL DIVISION                                  84.0
40019      VALLEY - JAIL DIV                                  49.0
40020  METRO - JAIL DIVISION                                  52.0
40021  METRO - JAIL DIVISION                                  21.0
40022            77TH ST                                      46.0

       Booking Location Code  Arrest Hour  Month  Bureau  Area Zip  \
0                        NaN           18      7  Valley     91304
1                     4206.0           20      3    West     90028
2                        NaN            9      6  Valley     91601
3                        NaN           23      6  Valley     91401
4                        NaN           12      6    West     90028
...                      ...          ...    ...     ...       ...
40018                 4273.0            1      6  Valley     91601
40019                 4279.0           11      6  Valley     91304
40020                 4273.0            1      6    West     90019
40021                 4273.0            2      6  Valley     91331
40022                 4212.0            3      6   South     90003

       Zip Median Income Arrest Date and Time Booking Date and Time  \
0                  68282  2020-07-09 18:30:00                   NaT
1                  43998  2020-03-25 20:45:00   2020-03-25 23:24:00
2                  52607  2020-06-23 09:30:00                   NaT
3                  54058  2020-06-24 23:22:00                   NaT
4                  43998  2020-06-23 12:10:00                   NaT
...                  ...                  ...                   ...
40018              52607  2020-06-17 01:22:00   2020-06-17 03:47:00
40019              68282  2020-06-19 11:30:00   2020-06-19 16:17:00
40020              52625  2020-06-19 01:00:00   2020-06-19 04:06:00
40021              60360  2020-06-17 02:00:00   2020-06-17 05:12:00
40022              37226  2020-06-16 03:00:00   2020-06-16 04:34:00

       Time Difference
0                  NaT
1             02:39:00
2                  NaT
3                  NaT
4                  NaT
...                ...
40018         02:25:00
40019         04:47:00
40020         03:06:00
40021         03:12:00
40022         01:34:00
```

```
[40023 rows x 39 columns]
```

[83]: `#pd.set_option('display.max_columns', None)`

[85]: `data.dtypes`

[85]:
```
Report ID                               int64
Report Type                            object
Arrest Date                            object
Time                                   object
Area ID                                 int64
Area Name                              object
Reporting District                      int64
Age                                     int64
Sex Code                               object
Ethnicity                              object
Charge Group Code                     float64
Charge Group Description               object
Arrest Type Code                       object
Charge                                 object
Charge Description                     object
Disposition Description                object
Address                                object
Cross Street                           object
LAT                                   float64
Zip Codes                             float64
LON                                   float64
Census Tracts                         float64
Location                               object
Precinct Boundaries                   float64
Booking Date                           object
LA Specific Plans                     float64
Booking Time                           object
Council Districts                     float64
Booking Location                       object
Neighborhood Councils (Certified)     float64
Booking Location Code                 float64
Arrest Hour                             int64
Month                                   int64
Bureau                                 object
Area Zip                                int64
Zip Median Income                       int64
Arrest Date and Time           datetime64[ns]
Booking Date and Time          datetime64[ns]
Time Difference               timedelta64[ns]
dtype: object
```

[50]: `data.to_csv('CLEAN_Arrest_USE_THIS.csv')`

[ ]: