

```
In [2]: import pandas as pd
```

```
In [3]: crime = pd.read_csv(r'C:\Users\jeeva\OneDrive\Desktop\udemy\crime_boston.csv')
```

```
In [4]: crime
```

Out[4]:

	INCIDENT_NUMBER	OFFENSE_CODE	OFFENSE_CODE_GROUP	OFFENSE_DESCRIPTION
0	I192064813	3006	Medical Assistance	SICK/INJURED/MEDICAL - PERSON
1	I192064810	3410	Towed	TOWED MOTOR VEHICLE
2	I192064808	2003	Offenses Against Child / Family	CHILD ENDANGERMENT (NO ASSAULT)
3	I192064803	802	Simple Assault	ASSAULT SIMPLE BATTERING
4	I192064800	3114	Investigate Property	INVESTIGATE PROPERTY
...	...	...	...	...
<b>413667</b>	I020691165-00	3125	Warrant Arrests	WARRANT ARRESTS
<b>413668</b>	I020674147-00	3125	Warrant Arrests	WARRANT ARRESTS
<b>413669</b>	I020669075-00	3125	Warrant Arrests	WARRANT ARRESTS
<b>413670</b>	I020662996-00	3125	Warrant Arrests	WARRANT ARRESTS
<b>413671</b>	I010370257-00	3125	Warrant Arrests	WARRANT ARRESTS

413672 rows × 17 columns



```
In [5]: #to check the missing values by using method is null
crime.isnull()
```

Out[5]:

	INCIDENT_NUMBER	OFFENSE_CODE	OFFENSE_CODE_GROUP	OFFENSE_DESCRIPTION
0		False	False	False
1		False	False	False
2		False	False	False
3		False	False	False
4		False	False	False
...	...	...	...	...
413667		False	False	False
413668		False	False	False
413669		False	False	False
413670		False	False	False
413671		False	False	False

413672 rows × 17 columns



In [6]: `crime.isnull().any(axis=1)`

Out[6]:

0	True
1	True
2	True
3	True
4	True
...	
413667	True
413668	True
413669	True
413670	True
413671	True

Length: 413672, dtype: bool

In [7]: `rows_with_missing_values = crime.isnull().any(axis=1)`  
`crime[rows_with_missing_values]`

Out[7]:

	INCIDENT_NUMBER	OFFENSE_CODE	OFFENSE_CODE_GROUP	OFFENSE_DESCRIPTION
0	I192064813	3006	Medical Assistance	SICK/INJURED/MEDICAL - PERSON
1	I192064810	3410	Towed	TOWED MOTOR VEHICLE
2	I192064808	2003	Offenses Against Child / Family	CHILD ENDANGERMENT (NO ASSAULT)
3	I192064803	802	Simple Assault	ASSAULT SIMPLE BATTERING
4	I192064800	3114	Investigate Property	INVESTIGATE PROPERTY
...	...	...	...	...
<b>413667</b>	I020691165-00	3125	Warrant Arrests	WARRANT ARRESTS
<b>413668</b>	I020674147-00	3125	Warrant Arrests	WARRANT ARRESTS
<b>413669</b>	I020669075-00	3125	Warrant Arrests	WARRANT ARRESTS
<b>413670</b>	I020662996-00	3125	Warrant Arrests	WARRANT ARRESTS
<b>413671</b>	I010370257-00	3125	Warrant Arrests	WARRANT ARRESTS

412042 rows × 17 columns



In [8]: `# to drop the column year, month, hour  
cleaned_crime = crime.drop(columns=['YEAR','MONTH','HOUR'])`

In [9]: `cleaned_crime`

Out[9]:

	INCIDENT_NUMBER	OFFENSE_CODE	OFFENSE_CODE_GROUP	OFFENSE_DESCRIPTION
0	I192064813	3006	Medical Assistance	SICK/INJURED/MEDICAL - PERSON
1	I192064810	3410	Towed	TOWED MOTOR VEHICLE
2	I192064808	2003	Offenses Against Child / Family	CHILD ENDANGERMENT (NO ASSAULT)
3	I192064803	802	Simple Assault	ASSAULT SIMPLE BATTERING
4	I192064800	3114	Investigate Property	INVESTIGATE PROPERTY
...	...	...	...	...
<b>413667</b>	I020691165-00	3125	Warrant Arrests	WARRANT ARRESTS
<b>413668</b>	I020674147-00	3125	Warrant Arrests	WARRANT ARRESTS
<b>413669</b>	I020669075-00	3125	Warrant Arrests	WARRANT ARRESTS
<b>413670</b>	I020662996-00	3125	Warrant Arrests	WARRANT ARRESTS
<b>413671</b>	I010370257-00	3125	Warrant Arrests	WARRANT ARRESTS

413672 rows × 14 columns



In [12]: `#to check missing spelling in the columns  
cleaned_crime['OFFENSE_CODE_GROUP'].unique()`

```
Out[12]: array(['Medical Assistance', 'Towed', 'Offenses Against Child / Family',
   'Simple Assault', 'Investigate Property', 'Verbal Disputes',
   'Larceny', 'Fire Related Reports', 'Missing Person Reported',
   'Robbery', 'Firearm Violations', 'Residential Burglary',
   'Investigate Person', 'Drug Violation', 'Aggravated Assault',
   'Violations', 'Motor Vehicle Accident Response', 'Property Found',
   'Larceny From Motor Vehicle', 'Liquor Violation', 'Other',
   'Harassment', 'Vandalism', 'Property Lost',
   'Assembly or Gathering Violations', 'Warrant Arrests',
   'Landlord/Tenant Disputes', 'Restraining Order Violations',
   'Recovered Stolen Property', 'Auto Theft Recovery',
   'Firearm Discovery', 'Other Burglary', 'Auto Theft',
   'Commercial Burglary', 'Service', 'Police Service Incidents',
   'Disorderly Conduct', 'Property Related Damage',
   'Harbor Related Incidents', 'Fraud', 'Ballistics',
   'Missing Person Located', 'Search Warrants', 'Confidence Games',
   'Counterfeiting', 'License Violation', 'HOME INVASION',
   'Operating Under the Influence', 'License Plate Related Incidents',
   'Evading Fare', 'Embezzlement', 'Prisoner Related Incidents',
   'Aircraft', 'Arson', 'Explosives', 'Homicide', 'Prostitution',
   'Criminal Harassment', 'Bomb Hoax', 'Phone Call Complaints',
   'HUMAN TRAFFICKING - INVOLUNTARY SERVITUDE', 'Manslaughter',
   'HUMAN TRAFFICKING', 'Biological Threat', 'Gambling',
   'INVESTIGATE PERSON', 'Burglary - No Property Taken'], dtype=object)
```

```
In [11]: cleaned_crime['OFFENSE_DESCRIPTION'].unique()
```

```
Out[11]: array(['SICK/INJURED/MEDICAL - PERSON', 'TOWED MOTOR VEHICLE',
   'CHILD ENDANGERMENT (NO ASSAULT)', 'ASSAULT SIMPLE - BATTERY',
   'INVESTIGATE PROPERTY', 'VERBAL DISPUTE', 'LARCENY SHOPLIFTING',
   'FIRE REPORT/ALARM - FALSE', 'SUDDEN DEATH', 'MISSING PERSON',
   'ROBBERY - STREET', 'WEAPON - OTHER - CARRYING / POSSESSING, ETC',
   'FIRE REPORT - HOUSE, BUILDING, ETC.',
   'BURGLARY - RESIDENTIAL - NO FORCE', 'INVESTIGATE PERSON',
   'LARCENY PURSE SNATCH - NO FORCE ',
   'DRUGS - POSS CLASS B - COCAINE, ETC.', 'DRUGS - POSS CLASS E',
   'ASSAULT - AGGRAVATED - BATTERY',
   'VAL - OPERATING W/O AUTHORIZATION LAWFUL',
   'M/V - LEAVING SCENE - PROPERTY DAMAGE', 'PROPERTY - FOUND',
   'LARCENY THEFT FROM MV - NON-ACCESSORY', 'M/V ACCIDENT - OTHER',
   'LIQUOR - DRINKING IN PUBLIC', 'BURGLARY - RESIDENTIAL - FORCE',
   'M/V ACCIDENT - INVOLVING \xa0BICYCLE - INJURY',
   'THREATS TO DO BODILY HARM', 'HARASSMENT',
   'LARCENY THEFT OF BICYCLE', 'VANDALISM', 'ROBBERY - OTHER',
   'GRAFFITI', 'PROPERTY - LOST', 'DEMONSTRATIONS/RIOT',
   'TRESPASSING', 'WARRANT ARREST', 'LANDLORD - TENANT SERVICE',
   'ASSAULT - AGGRAVATED', 'VIOL. OF RESTRAINING ORDER W NO ARREST',
   'LARCENY THEFT FROM BUILDING',
   'STOLEN PROPERTY - BUYING / RECEIVING / POSSESSING',
   'M/V ACCIDENT - PERSONAL INJURY', 'LARCENY ALL OTHERS',
   'BURGLARY - RESIDENTIAL - ATTEMPT',
   'DRUGS - POSS CLASS A - INTENT TO MFR DIST DISP',
   'LARCENY THEFT OF MV PARTS & ACCESSORIES',
   'VAL - OPERATING UNREG/UNINS \xa0CAR',
   'VAL - OPERATING AFTER REV/SUSP.',
   'RECOVERED - MV RECOVERED IN BOSTON (STOLEN OUTSIDE BOSTON)',
   'FIREARM/WEAPON - FOUND OR CONFISCATED',
   'M/V ACCIDENT - PROPERTY \xa0DAMAGE',
   'DRUGS - SICK ASSIST - HEROIN', 'BURGLARY - OTHER - FORCE',
   'DRUGS - POSS CLASS A - HEROIN, ETC. ', 'AUTO THEFT',
   'SICK/INJURED/MEDICAL - POLICE', 'ROBBERY - COMMERCIAL',
   'VAL - VIOLATION OF AUTO LAW - OTHER',
   'VAL - OPERATING WITHOUT LICENSE',
   'DRUGS - SICK ASSIST - OTHER HARMFUL DRUG',
   'BURGLARY - COMMERCIAL - FORCE',
   'M/V ACCIDENT INVOLVING PEDESTRIAN - INJURY', 'ANIMAL INCIDENTS',
   'OTHER OFFENSE', 'SERVICE TO OTHER PD INSIDE OF MA.',
   'DISTURBING THE PEACE', 'DISORDERLY CONDUCT',
   'M/V ACCIDENT - POLICE VEHICLE', 'SUICIDE / SUICIDE ATTEMPT',
   'M/V - LEAVING SCENE - PERSONAL INJURY',
   'AUTO THEFT - MOTORCYCLE / SCOOTER',
   'DRUGS - POSS CLASS E - INTENT TO MFR DIST DISP',
   'PROPERTY - ACCIDENTAL DAMAGE', 'HARBOR INCIDENT / VIOLATION',
   'FRAUD - FALSE PRETENSE / SCHEME', 'PROPERTY - MISSING',
   'POSSESSION OF BURGLARIOUS TOOLS', 'BALLISTICS EVIDENCE/FOUND',
   'MISSING PERSON - LOCATED', 'M/V ACCIDENT - OTHER CITY VEHICLE',
   'INVESTIGATION FOR ANOTHER AGENCY',
   'DRUGS - POSS CLASS B - INTENT TO MFR DIST DISP',
   'VIOLATION - CITY ORDINANCE',
   'M/V ACCIDENT - INVOLVING BICYCLE - NO INJURY',
   'WEAPON - FIREARM - OTHER VIOLATION',
   'DRUGS - SALE / MANUFACTURING', 'SEARCH WARRANT',
   'FRAUD - CREDIT CARD / ATM FRAUD', 'DRUGS - OTHER',
```

'DRUGS - POSS CLASS D - INTENT TO MFR DIST DISP',  
'DRUGS - POSS CLASS C', 'DEATH INVESTIGATION',  
'FORGERY / COUNTERFEITING', 'AUTO THEFT - LEASED/RENTED VEHICLE',  
'FRAUD - IMPERSONATION',  
'M/V ACCIDENT - INVOLVING PEDESTRIAN - NO INJURY',  
'ASSAULT - SIMPLE', 'SERVICE TO OTHER PD OUTSIDE OF MA.',  
'LICENSE PREMISE VIOLATION', 'HOME INVASION',  
'WEAPON - FIREARM - CARRYING / POSSESSING, ETC',  
'OPERATING UNDER THE INFLUENCE DRUGS',  
'MISSING PERSON - NOT REPORTED - LOCATED',  
'PROPERTY - STOLEN THEN RECOVERED', 'ANNOYING AND ACCOSTING',  
'VIOL. OF RESTRAINING ORDER W ARREST',  
'WEAPON - OTHER - OTHER VIOLATION', 'M/V PLATES - LOST',  
'DRUGS - CLASS A TRAFFICKING OVER 18 GRAMS',  
'ANIMAL CONTROL - DOG BITES - ETC.', 'DRUGS - POSS CLASS D',  
'FRAUD - WIRE', 'CHILD ABANDONMENT (NO ASSAULT)',  
'NOISY PARTY/RADIO-NO ARREST', 'BURGLARY - OTHER - NO FORCE',  
'ROBBERY - HOME INVASION', 'KIDNAPPING - ENTICING OR ATTEMPTED',  
'PROPERTY - LOST THEN LOCATED', 'EVADING FARE',  
'DRUGS - CLASS B TRAFFICKING OVER 18 GRAMS',  
'DRUGS - CONSP TO VIOL CONTROLLED SUBSTANCE',  
'OPERATING UNDER THE INFLUENCE ALCOHOL',  
'DRUGS - POSS CLASS C - INTENT TO MFR DIST DISP',  
'DRUGS - POSSESSION OF DRUG PARAPHRANALIA',  
'OBSCENE MATERIALS - PORNOGRAPHY', 'AFFRAY', 'EMBEZZLEMENT',  
'FUGITIVE FROM JUSTICE', 'DANGEROUS OR HAZARDOUS CONDITION',  
'BURGLARY - OTHER - ATTEMPT', 'LARCENY PICK-POCKET',  
'AIRCRAFT INCIDENTS', 'ARSON', 'DRUGS - POSSESSION',  
'EXPLOSIVES - TURNED IN OR FOUND', 'ROBBERY - BANK',  
'BURGLARY - COMMERCIAL - ATTEMPT',  
'FIRE REPORT - CAR, BRUSH, ETC.', 'CHINS',  
'PROTECTIVE CUSTODY / SAFEKEEPING', 'ROBBERY - CAR JACKING',  
'RECOVERED STOLEN PLATE', 'LIQUOR LAW VIOLATION',  
'EXTORTION OR BLACKMAIL', 'INTIMIDATING WITNESS',  
'CHILD ENDANGERMENT', 'KIDNAPPING/CUSTODIAL KIDNAPPING',  
'MURDER, NON-NEGIGENT MANSLAUGHTER',  
'BURGLARY - COMMERCIAL - NO FORCE',  
'VIOLATION - HAWKER AND PEDDLER',  
'DRUGS - SICK ASSIST - OTHER NARCOTIC', 'ANIMAL ABUSE',  
'PROSTITUTION - SOLICITING', 'FIREARM/WEAPON - LOST',  
'REPORT AFFECTING OTHER DEPTS.', 'CRIMINAL HARASSMENT',  
'BOMB THREAT', 'EXPLOSIVES - POSSESSION OR USE', 'STALKING',  
'NOISY PARTY/RADIO-ARREST',  
'WEAPON - FIREARM - SALE / TRAFFICKING',  
'CONTRIBUTING TO DELINQUENCY OF MINOR',  
'INJURY BICYCLE NO M/V INVOLVED', 'CONSPIRACY EXCEPT DRUG LAW',  
'OBSCENE PHONE CALLS', 'FRAUD - WELFARE',  
'PRISONER - SUICIDE / SUICIDE ATTEMPT',  
'VIOLATION - CITY ORDINANCE CONSTRUCTION PERMIT',  
'DRUGS - CLASS D TRAFFICKING OVER 50 GRAMS',  
'PROSTITUTION - COMMON NIGHTWALKER',  
'HUMAN TRAFFICKING - INVOLUNTARY SERVITUDE',  
'MANSLAUGHTER - VEHICLE - NEGLIGENCE', 'TRUANCY / RUNAWAY',  
'PROPERTY - CONCEALING LEASED', 'ABDUCTION - INTICING',  
'PROSTITUTION', 'FIREARM/WEAPON - ACCIDENTAL INJURY / DEATH',  
'LARCENY THEFT FROM COIN-OP MACHINE',

```
'PROSTITUTION - ASSISTING OR PROMOTING',
'PRISONER ESCAPE / ESCAPE & RECAPTURE',
'HUMAN TRAFFICKING - COMMERCIAL SEX ACTS',
'GATHERING CAUSING ANNOYANCE', 'BIOLOGICAL THREATS',
'PRISONER ATTEMPT TO RESCUE', 'GAMBLING - BETTING / WAGERING',
'CUSTODIAL KIDNAPPING', 'MANSLAUGHTER - NON-VEHICLE - NEGLIGENCE',
'DRUGS - GLUE INHALATION', 'KILLING OF FELON BY POLICE',
'COUNTERFEITING', 'LARCENY NON-ACCESSORY FROM VEH. $50 TO $199',
'ANNOYING AND ACCOSTIN', 'LARCENY BICYCLE $200 & OVER',
'DRUGS - POSS CLASS A - HEROIN, ETC.',
'LARCENY SHOPLIFTING UNDER $50',
'A&B HANDS, FEET, ETC. - MED. ATTENTION REQ.',
'ASSAULT & BATTERY D/W - OTHER',
'DRUGS - POSS CLASS D - MARIJUANA, ETC.',
'LARCENY OTHER $200 & OVER', 'LARCENY IN A BUILDING $200 & OVER',
'FIREARM/WEAPON - POSSESSION OF DANGEROUS',
'FIREARM/WEAPON - CARRY - SELL - RENT',
'ASSAULT & BATTERY D/W - OTHER ON POLICE OFFICER',
'A&B ON POLICE OFFICER', 'ASSAULT D/W - OTHER',
'ASSAULT & BATTERY', 'AUTO THEFT - RECOVERED IN BY POLICE',
'ROBBERY ATTEMPT - KNIFE - BANK',
'ASSAULT D/W - KNIFE ON POLICE OFFICER',
'ROBBERY - UNARMED - STREET', 'LARCENY SHOPLIFTING $200 & OVER',
'FORGERY OR UTTERING', 'FRAUD - FALSE PRETENSE',
'AUTO THEFT LEASE/RENT VEHICLE', 'VIOLATION - RESTRAINING ORDER',
'ROBBERY - UNARMED - CHAIN STORE',
'LARCENY IN A BUILDING $50 TO $199', 'LARCENY OTHER $50 TO $199',
'PROPERTY - RECEIVING STOLEN',
'AUTO THEFT - OUTSIDE - RECOVERED IN BOSTON',
'ROBBERY - KNIFE - STREET',
'LARCENY NON-ACCESSORY FROM VEH. $200 & OVER',
'LARCENY NON-ACCESSORY FROM VEH. UNDER $50',
'DRUGS - POSS CLASS E INTENT TO MF DIST DISP',
'ASSAULT & BATTERY D/W - KNIFE',
'B&E NON-RESIDENCE DAY - NO FORCE', 'LARCENY OTHER UNDER $50',
'LARCENY IN A BUILDING UNDER $50', 'DISORDERLY PERSON',
'B&E RESIDENCE NIGHT - ATTEMPT FORCE', 'CHILD ABUSE',
'ROBBERY - UNARMED - BUSINESS', 'AUTO THEFT OTHER',
'FRAUDS - ALL OTHER', 'B&E RESIDENCE DAY - NO PROP TAKEN',
'LARCENY SHOPLIFTING $50 TO $199', 'ROBBERY - UNARMED - RESIDENCE',
'B&E NON-RESIDENCE DAY - FORCIBLE',
'B&E NON-RESIDENCE DAY - NO PROP TAKEN',
'B&E RESIDENCE DAY - NO FORCE', 'B&E NON-RESIDENCE NIGHT - FORCE',
'PROSTITUTE - COMMON NIGHTWALKER',
'DRUGS - POSS CLASS D - INTENT MFR DIST DISP'], dtype=object)
```

In [13]: `cleaned_crime['DAY_OF_WEEK'].unique()`

Out[13]: `array(['Saturday', 'Friday', 'Wednesday', 'Thursday', 'Monday', 'Tuesday', 'Sunday'], dtype=object)`

In [15]: `cleaned_crime = cleaned_crime.drop(columns=['Location'])`

In [16]: `cleaned_crime`

INCIDENT_NUMBER	OFFENSE_CODE	OFFENSE_CODE_GROUP	OFFENSE_DESCRIPTION
0	I192064813	3006	Medical Assistance SICK/INJURED/MEDICAL - PERSO
1	I192064810	3410	Towed TOWED MOTOR VEHIC
2	I192064808	2003	Offenses Against Child / Family CHILD ENDANGERMENT (NO ASSAUL
3	I192064803	802	Simple Assault ASSAULT SIMPLE BATTER
4	I192064800	3114	Investigate Property INVESTIGATE PROPERTY
...	...	...	...
<b>413667</b>	I020691165-00	3125	Warrant Arrests WARRANT ARRESTS
<b>413668</b>	I020674147-00	3125	Warrant Arrests WARRANT ARRESTS
<b>413669</b>	I020669075-00	3125	Warrant Arrests WARRANT ARRESTS
<b>413670</b>	I020662996-00	3125	Warrant Arrests WARRANT ARRESTS
<b>413671</b>	I010370257-00	3125	Warrant Arrests WARRANT ARRESTS

413672 rows × 13 columns

