```
import pandas as pd

meteorites = pd.read_csv('Meteorite_Landings.csv',nrows=5)
meteorites
```

Out[378...

	name	id	nametype	recclass	mass (g)	fall	year	reclat	reclong	
0	Aachen	1	Valid	L5	21	Fell	01/01/1880 12:00:00 AM	50.77500	6.08333	
1	Aarhus	2	Valid	H6	720	Fell	01/01/1951 12:00:00 AM	56.18333	10.23333	
2	Abee	6	Valid	EH4	107000	Fell	01/01/1952 12:00:00 AM	54.21667	-113.00000	
3	Acapulco	10	Valid	Acapulcoite	1914	Fell	01/01/1976 12:00:00 AM	16.88333	-99.90000	
4	Achiras	370	Valid	L6	780	Fell	01/01/1902 12:00:00 AM	-33.16667	-64.95000	

◄ ━━━━━━━━━━━━━━━━━━━━━ ►

```
meteorites.name
```

Out[379...
```
0      Aachen
1      Aarhus
2        Abee
3    Acapulco
4     Achiras
Name: name, dtype: object
```

```
meteorites['name']
```

Out[380...
```
0      Aachen
1      Aarhus
2        Abee
3    Acapulco
4     Achiras
Name: name, dtype: object
```

```
meteorites.columns
```

Out[381...
```
Index(['name', 'id', 'nametype', 'recclass', 'mass (g)', 'fall', 'year',
       'reclat', 'reclong', 'GeoLocation'],
      dtype='object')
```

```
meteorites.index
```

```
Out[382…    RangeIndex(start=0, stop=5, step=1)
```

```python
In [383…   import requests

           response = requests.get(
               'https://data.nasa.gov/resource/gh4g-9sfh.json',
               params={'$limit': 50_000}
           )

           if response.ok:
               payload = response.json()
           else:
               print(f'Request was not successful and returned code: {response.status_code}.')
               payload = None
```

```python
In [384…   payload[0]
```

```
Out[384…   {'name': 'Aachen',
            'id': '1',
            'nametype': 'Valid',
            'recclass': 'L5',
            'mass': '21',
            'fall': 'Fell',
            'year': '1880-01-01T00:00:00.000',
            'reclat': '50.775000',
            'reclong': '6.083330',
            'geolocation': {'latitude': '50.775', 'longitude': '6.08333'}}
```

```python
In [385…   df = pd.DataFrame(payload)
           df.head(3)
```

Out[385…

	name	id	nametype	recclass	mass	fall	year	reclat	reclong	g
0	Aachen	1	Valid	L5	21	Fell	1880-01-01T00:00:00.000	50.775000	6.083330	
1	Aarhus	2	Valid	H6	720	Fell	1951-01-01T00:00:00.000	56.183330	10.233330	
2	Abee	6	Valid	EH4	107000	Fell	1952-01-01T00:00:00.000	54.216670	-113.000000	

```python
In [386…   meteorites = pd.read_csv('Meteorite_Landings.csv')
```

```python
In [387…   meteorites.shape
```

```
Out[387…   (45716, 10)
```

```
In [388… meteorites.columns
```

```
Out[388… Index(['name', 'id', 'nametype', 'recclass', 'mass (g)', 'fall', 'year',
               'reclat', 'reclong', 'GeoLocation'],
              dtype='object')
```

```
In [389… meteorites.dtypes
```

```
Out[389… name             object
         id                int64
         nametype         object
         recclass         object
         mass (g)        float64
         fall             object
         year             object
         reclat          float64
         reclong         float64
         GeoLocation      object
         dtype: object
```

```
In [390… meteorites.head(10)
```

	name	id	nametype	recclass	mass (g)	fall	year	reclat	reclong
0	Aachen	1	Valid	L5	21.0	Fell	01/01/1880 12:00:00 AM	50.77500	6.08333
1	Aarhus	2	Valid	H6	720.0	Fell	01/01/1951 12:00:00 AM	56.18333	10.23333
2	Abee	6	Valid	EH4	107000.0	Fell	01/01/1952 12:00:00 AM	54.21667	-113.00000
3	Acapulco	10	Valid	Acapulcoite	1914.0	Fell	01/01/1976 12:00:00 AM	16.88333	-99.90000
4	Achiras	370	Valid	L6	780.0	Fell	01/01/1902 12:00:00 AM	-33.16667	-64.95000
5	Adhi Kot	379	Valid	EH4	4239.0	Fell	01/01/1919 12:00:00 AM	32.10000	71.80000
6	Adzhi-Bogdo (stone)	390	Valid	LL3-6	910.0	Fell	01/01/1949 12:00:00 AM	44.83333	95.16667
7	Agen	392	Valid	H5	30000.0	Fell	01/01/1814 12:00:00 AM	44.21667	0.61667
8	Aguada	398	Valid	L6	1620.0	Fell	01/01/1930 12:00:00 AM	-31.60000	-65.23333
9	Aguila Blanca	417	Valid	L	1440.0	Fell	01/01/1920 12:00:00 AM	-30.86667	-64.55000

```
meteorites.tail(5)
```

	name	id	nametype	recclass	mass (g)	fall	year	reclat	r
45711	Zillah 002	31356	Valid	Eucrite	172.0	Found	01/01/1990 12:00:00 AM	29.03700	17
45712	Zinder	30409	Valid	Pallasite, ungrouped	46.0	Found	01/01/1999 12:00:00 AM	13.78333	8
45713	Zlin	30410	Valid	H4	3.3	Found	01/01/1939 12:00:00 AM	49.25000	17
45714	Zubkovsky	31357	Valid	L6	2167.0	Found	01/01/2003 12:00:00 AM	49.78917	41
45715	Zulu Queen	30414	Valid	L3.7	200.0	Found	01/01/1976 12:00:00 AM	33.98333	-115

In [392…

```
meteorites.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 45716 entries, 0 to 45715
Data columns (total 10 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   name         45716 non-null  object
 1   id           45716 non-null  int64
 2   nametype     45716 non-null  object
 3   recclass     45716 non-null  object
 4   mass (g)     45585 non-null  float64
 5   fall         45716 non-null  object
 6   year         45425 non-null  object
 7   reclat       38401 non-null  float64
 8   reclong      38401 non-null  float64
 9   GeoLocation  38401 non-null  object
dtypes: float64(3), int64(1), object(6)
memory usage: 3.5+ MB
```

In [393…

```
meteorites[['name','mass (g)']]
```

	name	mass (g)
0	Aachen	21.0
1	Aarhus	720.0
2	Abee	107000.0
3	Acapulco	1914.0
4	Achiras	780.0
...	...	...
45711	Zillah 002	172.0
45712	Zinder	46.0
45713	Zlin	3.3
45714	Zubkovsky	2167.0
45715	Zulu Queen	200.0

45716 rows × 2 columns

```
meteorites[100:104]
```

	name	id	nametype	recclass	mass (g)	fall	year	reclat	reclon
100	Benton	5026	Valid	LL6	2840.0	Fell	01/01/1949 12:00:00 AM	45.95000	-67.5500
101	Berduc	48975	Valid	L6	270.0	Fell	01/01/2008 12:00:00 AM	-31.91000	-58.3283
102	Béréba	5028	Valid	Eucrite-mmict	18000.0	Fell	01/01/1924 12:00:00 AM	11.65000	-3.6500
103	Berlanguillas	5029	Valid	L6	1440.0	Fell	01/01/1811 12:00:00 AM	41.68333	-3.8000

```
meteorites.iloc[100:104, [0,3,4,6]]
```

	name	recclass	mass (g)	year
100	Benton	LL6	2840.0	01/01/1949 12:00:00 AM
101	Berduc	L6	270.0	01/01/2008 12:00:00 AM
102	Béréba	Eucrite-mmict	18000.0	01/01/1924 12:00:00 AM
103	Berlanguillas	L6	1440.0	01/01/1811 12:00:00 AM

```python
meteorites.loc[100:104, 'mass (g)':'year']
```

	mass (g)	fall	year
100	2840.0	Fell	01/01/1949 12:00:00 AM
101	270.0	Fell	01/01/2008 12:00:00 AM
102	18000.0	Fell	01/01/1924 12:00:00 AM
103	1440.0	Fell	01/01/1811 12:00:00 AM
104	960.0	Fell	01/01/2004 12:00:00 AM

```python
meteorites.iloc[-1, [-1]]
```

```
GeoLocation    (33.98333, -115.68333)
Name: 45715, dtype: object
```

```python
(meteorites['mass (g)'] > 50) & (meteorites.fall == 'Found')
```

```
0        False
1        False
2        False
3        False
4        False
         ...
45711     True
45712    False
45713    False
45714     True
45715     True
Length: 45716, dtype: bool
```

```python
meteorites[(meteorites['mass (g)'] > 1e6) & (meteorites.fall == 'Fell')]
```

Out[399...

	name	id	nametype	recclass	mass (g)	fall	year	reclat	reclong
29	Allende	2278	Valid	CV3	2000000.0	Fell	01/01/1969 12:00:00 AM	26.96667	-105.3166
419	Jilin	12171	Valid	H5	4000000.0	Fell	01/01/1976 12:00:00 AM	44.05000	126.1666
506	Kunya-Urgench	12379	Valid	H5	1100000.0	Fell	01/01/1998 12:00:00 AM	42.25000	59.2000
707	Norton County	17922	Valid	Aubrite	1100000.0	Fell	01/01/1948 12:00:00 AM	39.68333	-99.8666
920	Sikhote-Alin	23593	Valid	Iron, IIAB	23000000.0	Fell	01/01/1947 12:00:00 AM	46.16000	134.6533

◀ ━━━━━━━━━━━━━━━━━━━━━━━━━━ ▶

In [400...
```python
meteorites.query("`mass (g)` > 1e6 & fall == 'Fell'")
```

Out[400...

	name	id	nametype	recclass	mass (g)	fall	year	reclat	reclong
29	Allende	2278	Valid	CV3	2000000.0	Fell	01/01/1969 12:00:00 AM	26.96667	-105.3166
419	Jilin	12171	Valid	H5	4000000.0	Fell	01/01/1976 12:00:00 AM	44.05000	126.1666
506	Kunya-Urgench	12379	Valid	H5	1100000.0	Fell	01/01/1998 12:00:00 AM	42.25000	59.2000
707	Norton County	17922	Valid	Aubrite	1100000.0	Fell	01/01/1948 12:00:00 AM	39.68333	-99.8666
920	Sikhote-Alin	23593	Valid	Iron, IIAB	23000000.0	Fell	01/01/1947 12:00:00 AM	46.16000	134.6533

◀ ━━━━━━━━━━━━━━━━━━━━━━━━━━ ▶

In [401...
```python
meteorites.fall.value_counts()
```

Out[401...
```
fall
Found    44609
Fell      1107
Name: count, dtype: int64
```

In [402...
```python
meteorites.value_counts(subset=['nametype','fall'], normalize = True)
```

```
Out[402…   nametype   fall
           Valid      Found      0.974145
                      Fell       0.024215
           Relict     Found      0.001641
           Name: proportion, dtype: float64
```

In [403…
```python
meteorites['mass (g)'].mean()
```

Out[403…
```
13278.078548601512
```

In [404…
```python
meteorites['mass (g)'].quantile(0.75)
```

Out[404…
```
202.6
```

In [405…
```python
meteorites['mass (g)'].median()
```

Out[405…
```
32.6
```

In [406…
```python
meteorites['mass (g)'].max()
```

Out[406…
```
60000000.0
```

In [407…
```python
meteorites.loc[meteorites['mass (g)'].idxmax()]
```

Out[407…
```
name                              Hoba
id                               11890
nametype                         Valid
recclass                     Iron, IVB
mass (g)                    60000000.0
fall                             Found
year           01/01/1920 12:00:00 AM
reclat                       -19.58333
reclong                       17.91667
GeoLocation      (-19.58333, 17.91667)
Name: 16392, dtype: object
```

In [408…
```python
meteorites.recclass.nunique()
```

Out[408…
```
466
```

In [409…
```python
meteorites.recclass.unique()[:14]
```

Out[409…
```
array(['L5', 'H6', 'EH4', 'Acapulcoite', 'L6', 'LL3-6', 'H5', 'L',
       'Diogenite-pm', 'Unknown', 'H4', 'H', 'Iron, IVA', 'CR2-an'],
      dtype=object)
```

In [410…
```python
meteorites.name.nunique()
```

Out[410…
```
45716
```

In [411…
```python
meteorites.describe(include='all')
```

Out[411...

	name	id	nametype	recclass	mass (g)	fall	year	
count	45716	45716.000000	45716	45716	4.558500e+04	45716	45425	3840
unique	45716	NaN	2	466	NaN	2	266	
top	Aachen	NaN	Valid	L6	NaN	Found	01/01/2003 12:00:00 AM	
freq	1	NaN	45641	8285	NaN	44609	3323	
mean	NaN	26889.735104	NaN	NaN	1.327808e+04	NaN	NaN	-39
std	NaN	16860.683030	NaN	NaN	5.749889e+05	NaN	NaN	46
min	NaN	1.000000	NaN	NaN	0.000000e+00	NaN	NaN	-87
25%	NaN	12688.750000	NaN	NaN	7.200000e+00	NaN	NaN	-76
50%	NaN	24261.500000	NaN	NaN	3.260000e+01	NaN	NaN	-71
75%	NaN	40656.750000	NaN	NaN	2.026000e+02	NaN	NaN	0
max	NaN	57458.000000	NaN	NaN	6.000000e+07	NaN	NaN	81

In [412...

```python
#Excercise (Part 1)

#1.)
df = pd.read_csv('2019_Yellow_Taxi_Trip_Data.csv')
df
```

	vendorid	tpep_pickup_datetime	tpep_dropoff_datetime	passenger_count	trip_distan
0	2	2019-10-23T16:39:42.000	2019-10-23T17:14:10.000	1	7.
1	1	2019-10-23T16:32:08.000	2019-10-23T16:45:26.000	1	2.
2	2	2019-10-23T16:08:44.000	2019-10-23T16:21:11.000	1	1.
3	2	2019-10-23T16:22:44.000	2019-10-23T16:43:26.000	1	1.
4	2	2019-10-23T16:45:11.000	2019-10-23T16:58:49.000	1	1.
...	...	...	...	...	
9995	1	2019-10-23T17:39:59.000	2019-10-23T17:49:26.000	2	1.
9996	1	2019-10-23T17:53:02.000	2019-10-23T18:00:45.000	1	1.
9997	1	2019-10-23T17:07:16.000	2019-10-23T17:11:35.000	1	0.
9998	1	2019-10-23T17:38:26.000	2019-10-23T17:49:28.000	2	2.
9999	1	2019-10-23T17:22:14.000	2019-10-23T17:52:09.000	1	3.

10000 rows × 18 columns

```
#2)

df.shape
```

```
(10000, 18)
```

```
#3)

df1 = df.iloc[:, [4, 10,13,14,16]]
df1.describe()
```

	trip_distance	fare_amount	tip_amount	tolls_amount	total_amount
count	10000.000000	10000.000000	10000.000000	10000.000000	10000.000000
mean	3.015250	15.106313	2.634494	0.623447	22.564659
std	4.148063	13.954762	3.409800	6.437507	19.209255
min	0.000000	-52.000000	0.000000	-6.120000	-65.920000
25%	0.920000	7.000000	0.000000	0.000000	12.375000
50%	1.500000	10.000000	2.000000	0.000000	16.300000
75%	2.760000	16.000000	3.250000	0.000000	22.880000
max	38.110000	176.000000	43.000000	612.000000	671.800000

In [415...

```
#4)

df1.loc[df1['trip_distance'].idxmax()]
```

Out[415...

```
trip_distance      38.11
fare_amount       176.00
tip_amount         18.29
tolls_amount        6.12
total_amount      201.21
Name: 8338, dtype: float64
```

Observation: I experienced difficulties in Practicing the codes discussed, but as I started to get the flow, I hadn't had to go back and see the codes done for reference, I remembered some of the codes although not all.

In [416...

```
import pandas as pd

newdf = pd.read_csv('2019_Yellow_Taxi_Trip_Data.csv')
```

In [417...

```
mask = newdf.columns.str.contains('id$|store_and_fwd_flag', regex = True)
columns_to_drop = newdf.columns[mask]
columns_to_drop
```

Out[417...

```
Index(['vendorid', 'ratecodeid', 'store_and_fwd_flag', 'pulocationid',
       'dolocationid'],
      dtype='object')
```

In [418...

```
taxis = newdf.drop(columns=columns_to_drop)
taxis.head()
```

	tpep_pickup_datetime	tpep_dropoff_datetime	passenger_count	trip_distance	payment_t
0	2019-10-23T16:39:42.000	2019-10-23T17:14:10.000	1	7.93	
1	2019-10-23T16:32:08.000	2019-10-23T16:45:26.000	1	2.00	
2	2019-10-23T16:08:44.000	2019-10-23T16:21:11.000	1	1.36	
3	2019-10-23T16:22:44.000	2019-10-23T16:43:26.000	1	1.00	
4	2019-10-23T16:45:11.000	2019-10-23T16:58:49.000	1	1.96	

In [419...
```python
taxis = taxis.rename(
    columns={
        'tpep_pickup_datetime': 'pickup_time',
        'tpep_dropoff_datetime': 'dropoff_time'
    }
)
taxis.columns
```

Out[419...
```
Index(['pickup_time', 'dropoff_time', 'passenger_count', 'trip_distance',
       'payment_type', 'fare_amount', 'extra', 'mta_tax', 'tip_amount',
       'tolls_amount', 'improvement_surcharge', 'total_amount',
       'congestion_surcharge'],
      dtype='object')
```

In [420...
```python
taxis[['pickup_time','dropoff_time']] = taxis[['pickup_time','dropoff_time']].apply
taxis.dtypes
```

Out[420...
```
pickup_time              datetime64[ns]
dropoff_time             datetime64[ns]
passenger_count                   int64
trip_distance                   float64
payment_type                      int64
fare_amount                     float64
extra                           float64
mta_tax                         float64
tip_amount                      float64
tolls_amount                    float64
improvement_surcharge           float64
total_amount                    float64
congestion_surcharge            float64
dtype: object
```

In [421...
```python
taxis = taxis.assign(
    elapsed_time=lambda x: x.dropoff_time - x.pickup_time,
    cost_before_tip=lambda x: x.total_amount - x.tip_amount,
    tip_pct=lambda x: x.tip_amount / x.cost_before_tip,
    fees=lambda x: x.cost_before_tip - x.fare_amount,
    avg_speed=lambda x: x.trip_distance.div(x.elapsed_time.dt.total_seconds()/60/60
```

```
                                        )
    )
```

```
taxis.dtypes
```

```
pickup_time              datetime64[ns]
dropoff_time             datetime64[ns]
passenger_count                   int64
trip_distance                   float64
payment_type                      int64
fare_amount                     float64
extra                           float64
mta_tax                         float64
tip_amount                      float64
tolls_amount                    float64
improvement_surcharge           float64
total_amount                    float64
congestion_surcharge            float64
elapsed_time             timedelta64[ns]
cost_before_tip                 float64
tip_pct                         float64
fees                            float64
avg_speed                       float64
dtype: object
```

```
taxis.sort_values(['passenger_count','pickup_time'],ascending=[False,True]).head()
```

	pickup_time	dropoff_time	passenger_count	trip_distance	payment_type	fare_amou
5997	2019-10-23 15:55:19	2019-10-23 16:08:25	6	1.58	2	1(
443	2019-10-23 15:56:59	2019-10-23 16:04:33	6	1.46	2	]
8722	2019-10-23 15:57:33	2019-10-23 16:03:34	6	0.62	1	!
4198	2019-10-23 15:57:38	2019-10-23 16:05:07	6	1.18	1	]
8238	2019-10-23 15:58:31	2019-10-23 16:29:29	6	3.23	2	1!

```
taxis.head()
```

Out[424... | | pickup_time | dropoff_time | passenger_count | trip_distance | payment_type | fare_amount |
|---|---|---|---|---|---|---|
| **0** | 2019-10-23 16:39:42 | 2019-10-23 17:14:10 | 1 | 7.93 | 1 | 29.5 |
| **1** | 2019-10-23 16:32:08 | 2019-10-23 16:45:26 | 1 | 2.00 | 1 | 10.5 |
| **2** | 2019-10-23 16:08:44 | 2019-10-23 16:21:11 | 1 | 1.36 | 1 | 9.5 |
| **3** | 2019-10-23 16:22:44 | 2019-10-23 16:43:26 | 1 | 1.00 | 1 | 13.0 |
| **4** | 2019-10-23 16:45:11 | 2019-10-23 16:58:49 | 1 | 1.96 | 1 | 10.5 |

In [425...
```python
taxis.nlargest(3,'elapsed_time')
```

Out[425... | | pickup_time | dropoff_time | passenger_count | trip_distance | payment_type | fare_amou |
|---|---|---|---|---|---|---|
| **7576** | 2019-10-23 16:52:51 | 2019-10-24 16:51:44 | 1 | 3.75 | 1 | 1? |
| **6902** | 2019-10-23 16:51:42 | 2019-10-24 16:50:22 | 1 | 11.19 | 2 | 39 |
| **4975** | 2019-10-23 16:18:51 | 2019-10-24 16:17:30 | 1 | 0.70 | 2 | ? |

In [426...
```python
taxis.nlargest(3, 'trip_distance')
```

Out[426... | | pickup_time | dropoff_time | passenger_count | trip_distance | payment_type | fare_amou |
|---|---|---|---|---|---|---|
| **8338** | 2019-10-23 16:50:53 | 2019-10-24 15:32:55 | 1 | 38.11 | 1 | 176 |
| **9965** | 2019-10-23 17:34:29 | 2019-10-23 18:48:00 | 1 | 37.86 | 2 | 52 |
| **1656** | 2019-10-23 16:04:45 | 2019-10-23 19:11:40 | 3 | 37.57 | 1 | 52 |

In [427...
```python
#Seatwork 2
```

In [428...
```python
met = pd.read_csv('Meteorite_Landings.csv')
met.head()
```

Out[428...

	name	id	nametype	recclass	mass (g)	fall	year	reclat	reclong
0	Aachen	1	Valid	L5	21.0	Fell	01/01/1880 12:00:00 AM	50.77500	6.08333
1	Aarhus	2	Valid	H6	720.0	Fell	01/01/1951 12:00:00 AM	56.18333	10.23333
2	Abee	6	Valid	EH4	107000.0	Fell	01/01/1952 12:00:00 AM	54.21667	-113.00000
3	Acapulco	10	Valid	Acapulcoite	1914.0	Fell	01/01/1976 12:00:00 AM	16.88333	-99.90000
4	Achiras	370	Valid	L6	780.0	Fell	01/01/1902 12:00:00 AM	-33.16667	-64.95000

In [429...

```python
met = met.rename(
    columns ={
    'mass (g)': 'mass'
}
)
met.head()
```

Out[429...

	name	id	nametype	recclass	mass	fall	year	reclat	reclong
0	Aachen	1	Valid	L5	21.0	Fell	01/01/1880 12:00:00 AM	50.77500	6.08333
1	Aarhus	2	Valid	H6	720.0	Fell	01/01/1951 12:00:00 AM	56.18333	10.23333
2	Abee	6	Valid	EH4	107000.0	Fell	01/01/1952 12:00:00 AM	54.21667	-113.00000
3	Acapulco	10	Valid	Acapulcoite	1914.0	Fell	01/01/1976 12:00:00 AM	16.88333	-99.90000
4	Achiras	370	Valid	L6	780.0	Fell	01/01/1902 12:00:00 AM	-33.16667	-64.95000

In [430...

```python
met = met.drop(columns=['reclat','reclong'])
met.head()
```

	name	id	nametype	recclass	mass	fall	year	GeoLocation
0	Aachen	1	Valid	L5	21.0	Fell	01/01/1880 12:00:00 AM	(50.775, 6.08333)
1	Aarhus	2	Valid	H6	720.0	Fell	01/01/1951 12:00:00 AM	(56.18333, 10.23333)
2	Abee	6	Valid	EH4	107000.0	Fell	01/01/1952 12:00:00 AM	(54.21667, -113.0)
3	Acapulco	10	Valid	Acapulcoite	1914.0	Fell	01/01/1976 12:00:00 AM	(16.88333, -99.9)
4	Achiras	370	Valid	L6	780.0	Fell	01/01/1902 12:00:00 AM	(-33.16667, -64.95)

In [431...

```python
met = met.sort_values('mass',ascending=False)
met.head()
```

Out[431...

	name	id	nametype	recclass	mass	fall	year	GeoLocation
16392	Hoba	11890	Valid	Iron, IVB	60000000.0	Found	01/01/1920 12:00:00 AM	(-19.58333, 17.91667)
5373	Cape York	5262	Valid	Iron, IIIAB	58200000.0	Found	01/01/1818 12:00:00 AM	(76.13333, -64.93333)
5365	Campo del Cielo	5247	Valid	Iron, IAB-MG	50000000.0	Found	12/22/1575 12:00:00 AM	(-27.46667, -60.58333)
5370	Canyon Diablo	5257	Valid	Iron, IAB-MG	30000000.0	Found	01/01/1891 12:00:00 AM	(35.05, -111.03333)
3455	Armanty	2335	Valid	Iron, IIIE	28000000.0	Found	01/01/1898 12:00:00 AM	(47.0, 88.0)

In [432...

```python
#================================================#
```

In [433...

```python
taxis = taxis.set_index('pickup_time')
taxis.head(3)
```

pickup_time	dropoff_time	passenger_count	trip_distance	payment_type	fare_amount	ext
2019-10-23 16:39:42	2019-10-23 17:14:10	1	7.93	1	29.5	
2019-10-23 16:32:08	2019-10-23 16:45:26	1	2.00	1	10.5	
2019-10-23 16:08:44	2019-10-23 16:21:11	1	1.36	1	9.5	

◀ ▬▬▬▬▬▬▬▬▬ ▶

```
taxis = taxis.sort_index()
taxis.head()
```

pickup_time	dropoff_time	passenger_count	trip_distance	payment_type	fare_amount	ext
2019-10-23 07:05:34	2019-10-23 08:03:16	3	14.68	1	50.0	
2019-10-23 07:48:58	2019-10-23 07:52:09	1	0.67	2	4.5	
2019-10-23 08:02:09	2019-10-24 07:42:32	1	8.38	1	32.0	
2019-10-23 08:18:47	2019-10-23 08:36:05	1	2.39	2	12.5	
2019-10-23 09:27:16	2019-10-23 09:33:13	2	1.11	2	6.0	

◀ ▬▬▬▬▬▬▬▬▬ ▶

```
taxis['2019-10-23 07:45':'2019-10-23 08']
```

pickup_time	dropoff_time	passenger_count	trip_distance	payment_type	fare_amount	ext
2019-10-23 07:48:58	2019-10-23 07:52:09	1	0.67	2	4.5	
2019-10-23 08:02:09	2019-10-24 07:42:32	1	8.38	1	32.0	
2019-10-23 08:18:47	2019-10-23 08:36:05	1	2.39	2	12.5	

◀ ▬▬▬▬▬▬▬▬▬ ▶

```
taxis['2019-10-23':'2019-10-23']
```

Out[436...

pickup_time	dropoff_time	passenger_count	trip_distance	payment_type	fare_amount	ex
2019-10-23 07:05:34	2019-10-23 08:03:16	3	14.68	1	50.0	
2019-10-23 07:48:58	2019-10-23 07:52:09	1	0.67	2	4.5	
2019-10-23 08:02:09	2019-10-24 07:42:32	1	8.38	1	32.0	
2019-10-23 08:18:47	2019-10-23 08:36:05	1	2.39	2	12.5	
2019-10-23 09:27:16	2019-10-23 09:33:13	2	1.11	2	6.0	
...	...	...	...	...	...	
2019-10-23 17:59:53	2019-10-23 18:12:56	2	1.60	1	10.0	
2019-10-23 17:59:53	2019-10-23 18:19:12	1	2.39	2	14.0	
2019-10-23 18:00:03	2019-10-23 18:04:56	1	0.94	2	5.5	
2019-10-23 18:01:21	2019-10-23 18:08:00	5	1.25	1	6.5	
2019-10-23 18:03:03	2019-10-23 18:10:45	1	0.76	1	6.5	

9993 rows × 17 columns

In [437...
```python
taxis.loc['2019-10-23 08']
```

Out[437...

pickup_time	dropoff_time	passenger_count	trip_distance	payment_type	fare_amount	ex
2019-10-23 08:02:09	2019-10-24 07:42:32	1	8.38	1	32.0	
2019-10-23 08:18:47	2019-10-23 08:36:05	1	2.39	2	12.5	

In [438...
```python
taxis = taxis.reset_index()
taxis.head()
```

Out[438...

	pickup_time	dropoff_time	passenger_count	trip_distance	payment_type	fare_amount
0	2019-10-23 07:05:34	2019-10-23 08:03:16	3	14.68	1	50.0
1	2019-10-23 07:48:58	2019-10-23 07:52:09	1	0.67	2	4.5
2	2019-10-23 08:02:09	2019-10-24 07:42:32	1	8.38	1	32.0
3	2019-10-23 08:18:47	2019-10-23 08:36:05	1	2.39	2	12.5
4	2019-10-23 09:27:16	2019-10-23 09:33:13	2	1.11	2	6.0

In [439...
```python
#Excercise 3
```

In [440...
```python
met = met.sort_index()
met.head()
```

Out[440...

	name	id	nametype	recclass	mass	fall	year	GeoLocation
0	Aachen	1	Valid	L5	21.0	Fell	01/01/1880 12:00:00 AM	(50.775, 6.08333)
1	Aarhus	2	Valid	H6	720.0	Fell	01/01/1951 12:00:00 AM	(56.18333, 10.23333)
2	Abee	6	Valid	EH4	107000.0	Fell	01/01/1952 12:00:00 AM	(54.21667, -113.0)
3	Acapulco	10	Valid	Acapulcoite	1914.0	Fell	01/01/1976 12:00:00 AM	(16.88333, -99.9)
4	Achiras	370	Valid	L6	780.0	Fell	01/01/1902 12:00:00 AM	(-33.16667, -64.95)

In [441...
```python
met.dtypes
```

Out[441...
```
name           object
id              int64
nametype       object
recclass       object
mass          float64
fall           object
year           object
GeoLocation    object
dtype: object
```

In [442...
```python
metcopy = met
metcopy
```

	name	id	nametype	recclass	mass	fall	year	GeoLocation
0	Aachen	1	Valid	L5	21.0	Fell	01/01/1880 12:00:00 AM	(50.775 6.08333
1	Aarhus	2	Valid	H6	720.0	Fell	01/01/1951 12:00:00 AM	(56.18333 10.23333
2	Abee	6	Valid	EH4	107000.0	Fell	01/01/1952 12:00:00 AM	(54.21667 -113.0
3	Acapulco	10	Valid	Acapulcoite	1914.0	Fell	01/01/1976 12:00:00 AM	(16.88333 -99.9
4	Achiras	370	Valid	L6	780.0	Fell	01/01/1902 12:00:00 AM	(-33.16667 -64.95
...	...	...	...	...	...	...	...	.
45711	Zillah 002	31356	Valid	Eucrite	172.0	Found	01/01/1990 12:00:00 AM	(29.037 17.0185
45712	Zinder	30409	Valid	Pallasite, ungrouped	46.0	Found	01/01/1999 12:00:00 AM	(13.78333 8.96667
45713	Zlin	30410	Valid	H4	3.3	Found	01/01/1939 12:00:00 AM	(49.25 17.66667
45714	Zubkovsky	31357	Valid	L6	2167.0	Found	01/01/2003 12:00:00 AM	(49.78917 41.5046
45715	Zulu Queen	30414	Valid	L3.7	200.0	Found	01/01/1976 12:00:00 AM	(33.98333 -115.68333

45716 rows × 8 columns

```
metcopy1 = lambda x: x.str[6:10]
metcopy['year'] = metcopy1(metcopy['year'])
metcopy
```

	name	id	nametype	recclass	mass	fall	year	GeoLocation
0	Aachen	1	Valid	L5	21.0	Fell	1880	(50.775, 6.08333)
1	Aarhus	2	Valid	H6	720.0	Fell	1951	(56.18333, 10.23333)
2	Abee	6	Valid	EH4	107000.0	Fell	1952	(54.21667, -113.0)
3	Acapulco	10	Valid	Acapulcoite	1914.0	Fell	1976	(16.88333, -99.9)
4	Achiras	370	Valid	L6	780.0	Fell	1902	(-33.16667, -64.95)
...	...	...	...	...	...	...	...	...
45711	Zillah 002	31356	Valid	Eucrite	172.0	Found	1990	(29.037, 17.0185)
45712	Zinder	30409	Valid	Pallasite, ungrouped	46.0	Found	1999	(13.78333, 8.96667)
45713	Zlin	30410	Valid	H4	3.3	Found	1939	(49.25, 17.66667)
45714	Zubkovsky	31357	Valid	L6	2167.0	Found	2003	(49.78917, 41.5046)
45715	Zulu Queen	30414	Valid	L3.7	200.0	Found	1976	(33.98333, -115.68333)

45716 rows × 8 columns

```python
metcopy['year'] = metcopy['year'].apply(pd.to_numeric)
metcopy.dtypes
```

```
name           object
id              int64
nametype       object
recclass       object
mass          float64
fall           object
year          float64
GeoLocation    object
dtype: object
```

```python
metcopy["Observed"] = (metcopy.year < 1970) & (metcopy.fall == "Fell")
metcopy
```

	name	id	nametype	recclass	mass	fall	year	GeoLocation	O
0	Aachen	1	Valid	L5	21.0	Fell	1880.0	(50.775, 6.08333)	
1	Aarhus	2	Valid	H6	720.0	Fell	1951.0	(56.18333, 10.23333)	
2	Abee	6	Valid	EH4	107000.0	Fell	1952.0	(54.21667, -113.0)	
3	Acapulco	10	Valid	Acapulcoite	1914.0	Fell	1976.0	(16.88333, -99.9)	
4	Achiras	370	Valid	L6	780.0	Fell	1902.0	(-33.16667, -64.95)	
...	...	...	...	...	...	...	...	...	
45711	Zillah 002	31356	Valid	Eucrite	172.0	Found	1990.0	(29.037, 17.0185)	
45712	Zinder	30409	Valid	Pallasite, ungrouped	46.0	Found	1999.0	(13.78333, 8.96667)	
45713	Zlin	30410	Valid	H4	3.3	Found	1939.0	(49.25, 17.66667)	
45714	Zubkovsky	31357	Valid	L6	2167.0	Found	2003.0	(49.78917, 41.5046)	
45715	Zulu Queen	30414	Valid	L3.7	200.0	Found	1976.0	(33.98333, -115.68333)	

45716 rows × 9 columns

```
metcopy.reset_index()
```

	index	name	id	nametype	recclass	mass	fall	year	GeoLoca
0	0	Aachen	1	Valid	L5	21.0	Fell	1880.0	(50. 6.08
1	1	Aarhus	2	Valid	H6	720.0	Fell	1951.0	(56.18 10.23
2	2	Abee	6	Valid	EH4	107000.0	Fell	1952.0	(54.21 -1
3	3	Acapulco	10	Valid	Acapulcoite	1914.0	Fell	1976.0	(16.88 -9
4	4	Achiras	370	Valid	L6	780.0	Fell	1902.0	(-33.16 -64
...	...	...	...	...	...	...	...	...	
45711	45711	Zillah 002	31356	Valid	Eucrite	172.0	Found	1990.0	(29. 17.0
45712	45712	Zinder	30409	Valid	Pallasite, ungrouped	46.0	Found	1999.0	(13.78 8.96
45713	45713	Zlin	30410	Valid	H4	3.3	Found	1939.0	(49 17.66
45714	45714	Zubkovsky	31357	Valid	L6	2167.0	Found	2003.0	(49.78 41.5
45715	45715	Zulu Queen	30414	Valid	L3.7	200.0	Found	1976.0	(33.98 -115.68

45716 rows × 10 columns

◀ ▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬ ▶

```
metcopy.set_index('id')
```

id	name	nametype	recclass	mass	fall	year	GeoLocation	Observed
1	Aachen	Valid	L5	21.0	Fell	1880.0	(50.775, 6.08333)	True
2	Aarhus	Valid	H6	720.0	Fell	1951.0	(56.18333, 10.23333)	True
6	Abee	Valid	EH4	107000.0	Fell	1952.0	(54.21667, -113.0)	True
10	Acapulco	Valid	Acapulcoite	1914.0	Fell	1976.0	(16.88333, -99.9)	False
370	Achiras	Valid	L6	780.0	Fell	1902.0	(-33.16667, -64.95)	True
...	...	...	...	...	...	...	...	...
31356	Zillah 002	Valid	Eucrite	172.0	Found	1990.0	(29.037, 17.0185)	False
30409	Zinder	Valid	Pallasite, ungrouped	46.0	Found	1999.0	(13.78333, 8.96667)	False
30410	Zlin	Valid	H4	3.3	Found	1939.0	(49.25, 17.66667)	False
31357	Zubkovsky	Valid	L6	2167.0	Found	2003.0	(49.78917, 41.5046)	False
30414	Zulu Queen	Valid	L3.7	200.0	Found	1976.0	(33.98333, -115.68333)	False

45716 rows × 8 columns

```
metcopy = metcopy.sort_index()
metcopy
```

Out[448...

	name	id	nametype	recclass	mass	fall	year	GeoLocation	O
0	Aachen	1	Valid	L5	21.0	Fell	1880.0	(50.775, 6.08333)	
1	Aarhus	2	Valid	H6	720.0	Fell	1951.0	(56.18333, 10.23333)	
2	Abee	6	Valid	EH4	107000.0	Fell	1952.0	(54.21667, -113.0)	
3	Acapulco	10	Valid	Acapulcoite	1914.0	Fell	1976.0	(16.88333, -99.9)	
4	Achiras	370	Valid	L6	780.0	Fell	1902.0	(-33.16667, -64.95)	
...	...	...	...	...	...	...	...	...	
45711	Zillah 002	31356	Valid	Eucrite	172.0	Found	1990.0	(29.037, 17.0185)	
45712	Zinder	30409	Valid	Pallasite, ungrouped	46.0	Found	1999.0	(13.78333, 8.96667)	
45713	Zlin	30410	Valid	H4	3.3	Found	1939.0	(49.25, 17.66667)	
45714	Zubkovsky	31357	Valid	L6	2167.0	Found	2003.0	(49.78917, 41.5046)	
45715	Zulu Queen	30414	Valid	L3.7	200.0	Found	1976.0	(33.98333, -115.68333)	

45716 rows × 9 columns

◀ ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ ▶

In [455...
```python
metcopy.iloc[10036:10040]
```

Out[455...

	name	id	nametype	recclass	mass	fall	year	GeoLocation	Observed
10036	Elephant Moraine 90022	8432	Valid	CK5	15.5	Found	1990.0	(-76.28573, 156.45721)	False
10037	Elephant Moraine 90023	8433	Valid	CK5	31.5	Found	1990.0	(-76.27507, 156.41038)	False
10038	Elephant Moraine 90024	8434	Valid	Eucrite-br	22.8	Found	1990.0	(-76.28843, 156.47872)	False
10039	Elephant Moraine 90025	8435	Valid	CK5	45.8	Found	1990.0	(-76.282, 156.39926)	False

In [ ]: