Lecture 2: Introduction to Jupyter Notebook

First, get Jupyter notebook up and running.

If you've made it here, you've done this step!

Second, familiarize yourself with python.

As a first pass, you may want to play around with the examples in the cells below. Hit Shift+Enter to evaluate a cell.

```python
print("Hello World CMPE 255 - DATA Mining @ Shreek!")  # Hello world is really easy in python!
```

```
Hello World CMPE 255 - DATA Mining @ Shreek!
```

```python
A = [6,4,3,8,5] # A is a list.
print(A) # you can print it out!
```

```
[6, 4, 3, 8, 5]
```

```python
A[0] # lists are zero-indexed.
```

```
6
```

```python
# slicing up lists:
A = [6,4,3,8,5]
print(A[2:4]) # this is the list [A[2],A[3]] (it doesn't include A[4])
print(A[2:])  # this notation starts with A[2] and goes to the end
print(A[:4])  # this starts at the beginning and goes up until A[3]
print(A[:])   # this just returns a copy of the whole list
```

```
[3, 8]
[3, 8, 5]
[6, 4, 3, 8]
[6, 4, 3, 8, 5]
```

```python
len(A) # get the length of a list
```

```
5
```

```python
A.append(7) # this appends "7" to A
print(A)
# what happens if you evaluate this cell multiple times?
```

```
[6, 4, 3, 8, 5, 7]
```

```python
A = A[:5] # let's set A back to how it was.
print(A)
```

```
[6, 4, 3, 8, 5]
```

```
A = A + ["cat"]  # Python is totally cool with this
print(A)
```

```
[6, 4, 3, 8, 5, 'cat']
```

```
A = [6,4,3,8,5]
for x in A:  # we can iterate over items in a list to get a for loop
    print(2*x)
```

```
12
8
6
16
10
```

```
# Notice that there's no {} or ; or anything like that.
#Python uses the whitespace to tell what's in the loop and what's not.

for x in A:
    print(3*x)
print("This is outside the loop")

print("---")

for x in A:
    print(3*x)
    print("This is inside the loop")
```

```
18
12
9
24
15
This is outside the loop
---
18
This is inside the loop
12
This is inside the loop
9
This is inside the loop
24
This is inside the loop
15
This is inside the loop
```

```
T = range(5)  # the range function gives you a way to iterate over a
range of integers
```

```
for x in T:
    print(x)
```

```
0
1
2
3
4
```

```
for i in range(5):  # we can also use the range function to iterate
over A
    print(2*A[i])
```

```
12
8
6
16
10
```

```
for i in range(len(A)):  # and if we don't know how long A is to begin
with, we can just use len(A)
    print(2*A[i])
```

```
12
8
6
16
10
```

```
B = [] # make an empty list
for x in A:
    B.append(2*x)
print(B)
```

```
[12, 8, 6, 16, 10]
```

```
C = [ 2*x for x in A ]
# This makes exactly the same list B that we had before, but in just
one line.
print(C)
```

```
[12, 8, 6, 16, 10]
```

```
def f(x,y):  # this is how we define a function.  Notice that x and y
don't have types.
    return x + y

print(f(2,3))  # python has one version of + for integers
print(f([1,2,3],[4,5,6]))  # and another version for lists
print(f("hello ", "world"))  # and another version for strings
# what happens if you do f(2, "cat")?
```

```
5
[1, 2, 3, 4, 5, 6]
hello world
```

As a more serious pass, here is a nice tutorial: https://www.programiz.com/python-programming

Third, let's explore some data

```
!pip install matplotlib

Collecting matplotlib
  Downloading matplotlib-3.9.2-cp312-cp312-win_amd64.whl.metadata (11
kB)
Collecting contourpy>=1.0.1 (from matplotlib)
  Downloading contourpy-1.2.1-cp312-cp312-win_amd64.whl.metadata (5.8
kB)
Collecting cycler>=0.10 (from matplotlib)
  Downloading cycler-0.12.1-py3-none-any.whl.metadata (3.8 kB)
Collecting fonttools>=4.22.0 (from matplotlib)
  Downloading fonttools-4.53.1-cp312-cp312-win_amd64.whl.metadata (165
kB)
     ---------------------------------------- 0.0/165.9 kB ? eta
-:--:--
     ------- -------------------------------- 30.7/165.9 kB 1.4 MB/s
eta 0:00:01
     ---------------------------------------- 165.9/165.9 kB 2.5 MB/s
eta 0:00:00
Collecting kiwisolver>=1.3.1 (from matplotlib)
  Downloading kiwisolver-1.4.5-cp312-cp312-win_amd64.whl.metadata (6.5
kB)
Requirement already satisfied: numpy>=1.23 in c:\users\shree\appdata\
local\programs\python\python312\lib\site-packages (from matplotlib)
(2.1.0)
Requirement already satisfied: packaging>=20.0 in c:\users\shree\
appdata\local\programs\python\python312\lib\site-packages (from
matplotlib) (24.1)
Collecting pillow>=8 (from matplotlib)
  Downloading pillow-10.4.0-cp312-cp312-win_amd64.whl.metadata (9.3
kB)
Collecting pyparsing>=2.3.1 (from matplotlib)
  Downloading pyparsing-3.1.4-py3-none-any.whl.metadata (5.1 kB)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\shree\
appdata\local\programs\python\python312\lib\site-packages (from
matplotlib) (2.9.0.post0)
Requirement already satisfied: six>=1.5 in c:\users\shree\appdata\
local\programs\python\python312\lib\site-packages (from python-
dateutil>=2.7->matplotlib) (1.16.0)
Downloading matplotlib-3.9.2-cp312-cp312-win_amd64.whl (7.8 MB)
```

```
   ---------------------------------------- 0.0/7.8 MB ? eta -:--:--
   - -------------------------------------- 0.3/7.8 MB 6.7 MB/s eta
0:00:02
   ----- ---------------------------------- 1.1/7.8 MB 12.0 MB/s eta
0:00:01
   ---------- ----------------------------- 2.1/7.8 MB 19.1 MB/s eta
0:00:01
   ----------- ---------------------------- 2.2/7.8 MB 12.9 MB/s eta
0:00:01
   ------------- -------------------------- 2.6/7.8 MB 12.9 MB/s eta
0:00:01
   ------------- -------------------------- 2.6/7.8 MB 12.9 MB/s eta
0:00:01
   -------------- ------------------------- 2.8/7.8 MB 8.9 MB/s eta
0:00:01
   -------------- ------------------------- 2.8/7.8 MB 8.6 MB/s eta
0:00:01
   -------------- ------------------------- 2.8/7.8 MB 8.6 MB/s eta
0:00:01
   -------------- ------------------------- 2.8/7.8 MB 8.6 MB/s eta
0:00:01
   -------------------------- ------------- 5.2/7.8 MB 10.4 MB/s eta
0:00:01
   --------------------------------- ------ 6.6/7.8 MB 12.0 MB/s eta
0:00:01
   ---------------------------------------  7.8/7.8 MB 13.2 MB/s eta
0:00:01
   ---------------------------------------- 7.8/7.8 MB 12.5 MB/s eta
0:00:00
Downloading contourpy-1.2.1-cp312-cp312-win_amd64.whl (189 kB)
   ---------------------------------------- 0.0/189.9 kB ? eta -:--:--
   ---------------------------------------- 189.9/189.9 kB ? eta
0:00:00
Downloading cycler-0.12.1-py3-none-any.whl (8.3 kB)
Downloading fonttools-4.53.1-cp312-cp312-win_amd64.whl (2.2 MB)
   ---------------------------------------- 0.0/2.2 MB ? eta -:--:--
   ------------------------------- -------- 1.7/2.2 MB 54.5 MB/s eta
0:00:01
   ---------------------------------------- 2.2/2.2 MB 28.0 MB/s eta
0:00:00
Downloading kiwisolver-1.4.5-cp312-cp312-win_amd64.whl (56 kB)
   ---------------------------------------- 0.0/56.0 kB ? eta -:--:--
   ---------------------------------------- 56.0/56.0 kB 2.9 MB/s eta
0:00:00
Downloading pillow-10.4.0-cp312-cp312-win_amd64.whl (2.6 MB)
   ---------------------------------------- 0.0/2.6 MB ? eta -:--:--
   ----------------------------- ---------- 1.9/2.6 MB 60.4 MB/s eta
0:00:01
   ---------------------------------------- 2.6/2.6 MB 32.6 MB/s eta
```

```
0:00:00
Downloading pyparsing-3.1.4-py3-none-any.whl (104 kB)
   ---------------------------------------- 0.0/104.1 kB ? eta -:--:--
   ---------------------------------------- 104.1/104.1 kB 6.3 MB/s
eta 0:00:00
Installing collected packages: pyparsing, pillow, kiwisolver,
fonttools, cycler, contourpy, matplotlib
Successfully installed contourpy-1.2.1 cycler-0.12.1 fonttools-4.53.1
kiwisolver-1.4.5 matplotlib-3.9.2 pillow-10.4.0 pyparsing-3.1.4


[notice] A new release of pip is available: 24.0 -> 24.2
[notice] To update, run: python.exe -m pip install --upgrade pip
```

```python
# The usual preamble
import pandas as pd
# Open graphs in new cells in the page rather than in a separate
window
%matplotlib inline
# Make the graphs a bit prettier, and bigger
pd.set_option('display.width', 5000)
pd.set_option('display.max_columns', 60)
```

We're going to use a new dataset here, to demonstrate how to deal with larger datasets. This is a subset of the of 311 service requests from NYC Open Data.

```python
import pandas as pd

complaints = pd.read_csv('311-requests.csv', low_memory=False)
```

1.1 What's even in it? (the summary)

When you look at a large dataframe, instead of showing you the contents of the dataframe, it'll show you a *summary*. This includes all the columns, and how many non-null values there are in each column.

```
complaints
```

```
       Unique Key            Created Date              Closed Date
Agency                                             Agency Name
Complaint Type                        Descriptor         Location Type
Incident Zip          Incident Address        Street Name    Cross
Street 1                        Cross Street 2 Intersection Street 1
Intersection Street 2 Address Type       City Landmark Facility Type
Status                      Due Date Resolution Action Updated Date
Community Board    Borough   X Coordinate (State Plane)  Y Coordinate
(State Plane) Park Facility Name Park Borough  School Name School
Number School Region  School Code School Phone Number School Address
```

```
School City School State   School Zip School Not Found  School or
Citywide Complaint Vehicle Type Taxi Company Borough Taxi Pick Up
Location Bridge Highway Name Bridge Highway Direction Road Ramp Bridge
Highway Segment Garage Lot Name Ferry Direction Ferry Terminal Name
Latitude  Longitude                                  Location
0        26589651  10/31/2013 02:08:41 AM                     NaN
NYPD                    New York City Police Department  Noise -
Street/Sidewalk                  Loud Talking       Street/Sidewalk
11432         90-03 169 STREET         169 STREET        90 AVENUE
91 AVENUE                  NaN                   NaN      ADDRESS
JAMAICA      NaN      Precinct  Assigned  10/31/2013 10:08:41 AM
10/31/2013 02:35:17 AM       12 QUEENS     QUEENS
1042027.0                    197389.0        Unspecified       QUEENS
Unspecified   Unspecified   Unspecified  Unspecified
Unspecified    Unspecified  Unspecified  Unspecified  Unspecified
N                           NaN          NaN                  NaN
NaN                 NaN                      NaN       NaN
NaN             NaN             NaN                   NaN  40.708275 -
73.791604   (40.70827532593202, -73.79160395779721)
1        26593698  10/31/2013 02:01:04 AM                     NaN
NYPD                    New York City Police Department
Illegal Parking  Commercial Overnight Parking       Street/Sidewalk
11378               58 AVENUE          58 AVENUE         58 PLACE
59 STREET                  NaN                   NaN    BLOCKFACE
MASPETH      NaN      Precinct      Open  10/31/2013 10:01:04 AM
NaN       05 QUEENS     QUEENS                   1009349.0
201984.0        Unspecified       QUEENS  Unspecified   Unspecified
Unspecified  Unspecified         Unspecified    Unspecified
Unspecified  Unspecified  Unspecified               N
NaN          NaN                 NaN                    NaN
NaN                      NaN       NaN                      NaN
NaN             NaN                 NaN  40.721041 -73.909453
(40.721040535628305, -73.90945306791765)
2        26594139  10/31/2013 02:00:24 AM  10/31/2013 02:40:32 AM
NYPD                    New York City Police Department       Noise -
Commercial              Loud Music/Party   Club/Bar/Restaurant
10032          4060 BROADWAY           BROADWAY  WEST 171 STREET
WEST 172 STREET                 NaN                     NaN
ADDRESS  NEW YORK      NaN      Precinct    Closed  10/31/2013
10:00:24 AM         10/31/2013 02:39:42 AM    12 MANHATTAN  MANHATTAN
1001088.0                    246531.0        Unspecified    MANHATTAN
Unspecified   Unspecified   Unspecified  Unspecified
Unspecified    Unspecified  Unspecified  Unspecified  Unspecified
N                           NaN          NaN                  NaN
NaN                 NaN                      NaN       NaN
NaN             NaN             NaN                   NaN  40.843330 -
73.939144   (40.84332975466513, -73.93914371913482)
3        26595721  10/31/2013 01:56:23 AM  10/31/2013 02:21:48 AM
NYPD                    New York City Police Department          Noise
```

```
- Vehicle                Car/Truck Horn       Street/Sidewalk
10023           WEST 72 STREET     WEST 72 STREET  COLUMBUS AVENUE
AMSTERDAM AVENUE                   NaN                   NaN
BLOCKFACE  NEW YORK      NaN      Precinct    Closed  10/31/2013
09:56:23 AM         10/31/2013 02:21:10 AM    07 MANHATTAN  MANHATTAN
989730.0                    222727.0        Unspecified    MANHATTAN
Unspecified   Unspecified   Unspecified  Unspecified
Unspecified    Unspecified  Unspecified  Unspecified  Unspecified
N                           NaN          NaN                   NaN
NaN                 NaN                      NaN       NaN
NaN             NaN             NaN                NaN  40.778009 -
73.980213    (40.7780087446372, -73.98021349023975)
4        26590930  10/31/2013 01:53:44 AM                      NaN
DOHMH            Department of Health and Mental Hygiene
Rodent  Condition Attracting Rodents           Vacant Lot
10027          WEST 124 STREET    WEST 124 STREET     LENOX AVENUE
ADAM CLAYTON POWELL JR BOULEVARD                   NaN
NaN    BLOCKFACE  NEW YORK      NaN           NaN   Pending
11/30/2013 01:53:44 AM        10/31/2013 01:59:54 AM    10 MANHATTAN
MANHATTAN                    998815.0                    233545.0
Unspecified    MANHATTAN  Unspecified   Unspecified   Unspecified
Unspecified         Unspecified    Unspecified  Unspecified
Unspecified  Unspecified             N
NaN          NaN                 NaN                   NaN
NaN                      NaN       NaN                     NaN
NaN             NaN                NaN  40.807691 -73.947387
(40.80769092704951, -73.94738703491433)
...           ...                    ...                    ...     .
..                                                ...
...                           ...                    ...          ...
...                ...              ...
...                   ...                    ...          ...       ...
...           ...       ...                     ...
...             ...        ...                         ...
...                ...          ...          ...           ...
...          ...                 ...            ...          ...
...          ...              ...                            ...
...                  ...                    ...                 ...
...        ...                    ...             ...              ...
...         ...        ...                                     ...
49994    26524469  10/21/2013 12:00:00 AM                      NaN
HPD  Department of Housing Preservation and Develop...          PAINT
- PLASTER                       CEILING  RESIDENTIAL BUILDING
11235        2940 OCEAN AVENUE       OCEAN AVENUE         AVENUE Y
AVENUE Z                   NaN                  NaN      ADDRESS
BROOKLYN      NaN           NaN      Open                      NaN
10/21/2013 12:00:00 AM     15 BROOKLYN   BROOKLYN
998260.0                    154071.0        Unspecified     BROOKLYN
Unspecified   Unspecified   Unspecified  Unspecified
Unspecified    Unspecified  Unspecified  Unspecified  Unspecified
NaN                           NaN          NaN                  NaN
NaN                 NaN                      NaN       NaN
NaN             NaN             NaN                 NaN  40.589554 -
73.949557  (40.589554394535476, -73.94955717050078)
49995    26524470  10/21/2013 12:00:00 AM                     NaN
HPD  Department of Housing Preservation and Develop...
ELECTRIC               ELECTRIC-SUPPLY  RESIDENTIAL BUILDING
10458      2704 DECATUR AVENUE     DECATUR AVENUE  EAST 195 STREET
EAST 197 STREET                   NaN                   NaN
ADDRESS     BRONX      NaN           NaN      Open
NaN         10/21/2013 12:00:00 AM        07 BRONX      BRONX
1015067.0                    254449.0        Unspecified        BRONX
Unspecified   Unspecified   Unspecified  Unspecified
Unspecified    Unspecified  Unspecified  Unspecified  Unspecified
NaN                           NaN          NaN                  NaN
NaN                 NaN                      NaN       NaN
NaN             NaN             NaN                 NaN  40.865025 -
73.888584   (40.86502456816568, -73.88858414414646)
49996    26524479  10/21/2013 12:00:00 AM                     NaN
HPD  Department of Housing Preservation and Develop...          PAINT
- PLASTER                         WALLS  RESIDENTIAL BUILDING
11373  51-55 VAN KLEECK STREET  VAN KLEECK STREET    CODWISE PLACE
KNEELAND AVENUE                   NaN                   NaN
ADDRESS  Elmhurst      NaN           NaN      Open
NaN         10/21/2013 12:00:00 AM       04 QUEENS     QUEENS
1016917.0                    207448.0        Unspecified       QUEENS
Unspecified   Unspecified   Unspecified  Unspecified
Unspecified    Unspecified  Unspecified  Unspecified  Unspecified
NaN                           NaN          NaN                  NaN
NaN                 NaN                      NaN       NaN
NaN             NaN             NaN                 NaN  40.736013 -
73.882124   (40.73601315955848, -73.88212432130575)
49997    26523160  10/21/2013 12:00:00 AM                     NaN
HPD  Department of Housing Preservation and Develop...     GENERAL
CONSTRUCTION                  CERAMIC-TILE  RESIDENTIAL BUILDING
10468      2600 CRESTON AVENUE     CRESTON AVENUE  EAST 192 STREET
EAST 193 STREET                   NaN                   NaN
ADDRESS     BRONX      NaN           NaN      Open
NaN         10/21/2013 12:00:00 AM        07 BRONX      BRONX
1012886.0                    254495.0        Unspecified        BRONX
Unspecified   Unspecified   Unspecified  Unspecified
Unspecified    Unspecified  Unspecified  Unspecified  Unspecified
NaN                           NaN          NaN                  NaN
NaN                 NaN                      NaN       NaN
NaN             NaN             NaN                 NaN  40.865158 -
73.896469  (40.865158168593744, -73.89646913260533)
49998    26523190  10/21/2013 12:00:00 AM  10/22/2013 12:00:00 AM
HPD  Department of Housing Preservation and Develop...
```

```
NONCONST                        VERMIN  RESIDENTIAL BUILDING
10458     3184 GRAND CONCOURSE    GRAND CONCOURSE  EAST 206 STREET
ST GEORGES CRESCENT                   NaN                   NaN
ADDRESS     BRONX      NaN           NaN    Closed
NaN         10/22/2013 12:00:00 AM        07 BRONX      BRONX
1015863.0                    258627.0        Unspecified        BRONX
Unspecified   Unspecified   Unspecified  Unspecified
Unspecified    Unspecified  Unspecified  Unspecified  Unspecified
NaN                          NaN          NaN                    NaN
NaN                 NaN                      NaN       NaN
NaN              NaN             NaN                 NaN  40.876489 -
73.885687   (40.87648907534718, -73.88568657501374)

[49999 rows x 52 columns]
```

1.2 Selecting columns and rows

To select a column, we index with the name of the column, like this:

```
complaints['Complaint Type']

0        Noise - Street/Sidewalk
1                Illegal Parking
2             Noise - Commercial
3                Noise - Vehicle
4                         Rodent
                  ...
49994            PAINT - PLASTER
49995                   ELECTRIC
49996            PAINT - PLASTER
49997       GENERAL CONSTRUCTION
49998                   NONCONST
Name: Complaint Type, Length: 49999, dtype: object
```

To get the first 5 rows of a dataframe, we can use a slice: `df[:5]`.

This is a great way to get a sense for what kind of information is in the dataframe -- take a minute to look at the contents and get a feel for this dataset.

```
complaints[:5]

   Unique Key            Created Date             Closed Date Agency
Agency Name           Complaint Type                  Descriptor
Location Type Incident Zip  Incident Address      Street Name   Cross
Street 1                    Cross Street 2 Intersection Street 1
Intersection Street 2 Address Type      City Landmark Facility Type
Status                Due Date Resolution Action Updated Date
Community Board    Borough  X Coordinate (State Plane)  Y Coordinate
```

```
(State Plane) Park Facility Name Park Borough  School Name School
Number School Region  School Code School Phone Number School Address
School City School State   School Zip School Not Found  School or
Citywide Complaint Vehicle Type Taxi Company Borough Taxi Pick Up
Location Bridge Highway Name Bridge Highway Direction Road Ramp Bridge
Highway Segment Garage Lot Name Ferry Direction Ferry Terminal Name
Latitude  Longitude                                  Location
0    26589651  10/31/2013 02:08:41 AM                     NaN   NYPD
New York City Police Department  Noise - Street/Sidewalk
Loud Talking      Street/Sidewalk        11432  90-03 169 STREET
169 STREET        90 AVENUE                      91 AVENUE
NaN                   NaN      ADDRESS   JAMAICA      NaN
Precinct  Assigned  10/31/2013 10:08:41 AM         10/31/2013 02:35:17
AM       12 QUEENS     QUEENS                   1042027.0
197389.0        Unspecified       QUEENS  Unspecified   Unspecified
Unspecified  Unspecified         Unspecified    Unspecified
Unspecified  Unspecified  Unspecified               N
NaN          NaN                NaN                   NaN
NaN                      NaN       NaN                   NaN
NaN             NaN                NaN  40.708275 -73.791604
(40.70827532593202, -73.79160395779721)
1    26593698  10/31/2013 02:01:04 AM                     NaN   NYPD
New York City Police Department          Illegal Parking  Commercial
Overnight Parking      Street/Sidewalk        11378        58 AVENUE
58 AVENUE         58 PLACE                        59 STREET
NaN                   NaN    BLOCKFACE   MASPETH      NaN
Precinct      Open  10/31/2013 10:01:04 AM
NaN       05 QUEENS     QUEENS                   1009349.0
201984.0        Unspecified       QUEENS  Unspecified   Unspecified
Unspecified  Unspecified         Unspecified    Unspecified
Unspecified  Unspecified  Unspecified               N
NaN          NaN                NaN                   NaN
NaN                      NaN       NaN                   NaN
NaN             NaN                NaN  40.721041 -73.909453
(40.721040535628305, -73.90945306791765)
2    26594139  10/31/2013 02:00:24 AM  10/31/2013 02:40:32 AM   NYPD
New York City Police Department       Noise - Commercial
Loud Music/Party  Club/Bar/Restaurant        10032     4060 BROADWAY
BROADWAY  WEST 171 STREET                  WEST 172 STREET
NaN                   NaN      ADDRESS  NEW YORK      NaN
Precinct    Closed  10/31/2013 10:00:24 AM         10/31/2013 02:39:42
AM    12 MANHATTAN  MANHATTAN                  1001088.0
246531.0        Unspecified    MANHATTAN  Unspecified   Unspecified
Unspecified  Unspecified         Unspecified    Unspecified
Unspecified  Unspecified  Unspecified               N
NaN          NaN                NaN                   NaN
NaN                      NaN       NaN                   NaN
NaN             NaN                NaN  40.843330 -73.939144
(40.84332975466513, -73.93914371913482)
```

```
3    26595721  10/31/2013 01:56:23 AM  10/31/2013 02:21:48 AM   NYPD
New York City Police Department          Noise - Vehicle
Car/Truck Horn      Street/Sidewalk        10023    WEST 72 STREET
WEST 72 STREET  COLUMBUS AVENUE                  AMSTERDAM AVENUE
NaN                   NaN    BLOCKFACE  NEW YORK      NaN
Precinct    Closed  10/31/2013 09:56:23 AM        10/31/2013 02:21:10
AM    07 MANHATTAN  MANHATTAN                    989730.0
222727.0        Unspecified    MANHATTAN  Unspecified   Unspecified
Unspecified  Unspecified         Unspecified    Unspecified
Unspecified  Unspecified  Unspecified               N
NaN         NaN                NaN                  NaN
NaN                      NaN       NaN                    NaN
NaN             NaN                NaN  40.778009 -73.980213
(40.7780087446372, -73.98021349023975)
4    26590930  10/31/2013 01:53:44 AM                     NaN  DOHMH
Department of Health and Mental Hygiene                   Rodent
Condition Attracting Rodents           Vacant Lot        10027   WEST
124 STREET  WEST 124 STREET     LENOX AVENUE  ADAM CLAYTON POWELL JR
BOULEVARD                   NaN                  NaN    BLOCKFACE
NEW YORK      NaN           NaN   Pending  11/30/2013 01:53:44 AM
10/31/2013 01:59:54 AM    10 MANHATTAN  MANHATTAN
998815.0                   233545.0        Unspecified    MANHATTAN
Unspecified   Unspecified   Unspecified  Unspecified
Unspecified    Unspecified  Unspecified  Unspecified  Unspecified
N                           NaN          NaN                 NaN
NaN                  NaN                      NaN       NaN
NaN             NaN             NaN                NaN  40.807691 -
73.947387   (40.80769092704951, -73.94738703491433)
```

We can combine these to get the first 5 rows of a column:

```
complaints['Complaint Type'][:5]
```

```
0    Noise - Street/Sidewalk
1            Illegal Parking
2         Noise - Commercial
3            Noise - Vehicle
4                     Rodent
Name: Complaint Type, dtype: object
```

and it doesn't matter which direction we do it in:

```
complaints[:5]['Complaint Type']
```

```
0    Noise - Street/Sidewalk
1            Illegal Parking
2         Noise - Commercial
3            Noise - Vehicle
```

```
4                      Rodent
Name: Complaint Type, dtype: object
```

1.3 Selecting multiple columns

What if we just want to know the complaint type and the borough, but not the rest of the information? Pandas makes it really easy to select a subset of the columns: just index with list of columns you want.

```
complaints[['Complaint Type', 'Borough']]

                Complaint Type    Borough
0      Noise - Street/Sidewalk     QUEENS
1              Illegal Parking     QUEENS
2           Noise - Commercial  MANHATTAN
3              Noise - Vehicle  MANHATTAN
4                       Rodent  MANHATTAN
...                        ...        ...
49994          PAINT - PLASTER   BROOKLYN
49995                 ELECTRIC      BRONX
49996          PAINT - PLASTER     QUEENS
49997     GENERAL CONSTRUCTION      BRONX
49998                 NONCONST      BRONX

[49999 rows x 2 columns]
```

That showed us a summary, and then we can look at the first 10 rows:

```
complaints[['Complaint Type', 'Borough']][:10]

            Complaint Type    Borough
0  Noise - Street/Sidewalk     QUEENS
1          Illegal Parking     QUEENS
2       Noise - Commercial  MANHATTAN
3          Noise - Vehicle  MANHATTAN
4                   Rodent  MANHATTAN
5       Noise - Commercial     QUEENS
6         Blocked Driveway     QUEENS
7       Noise - Commercial     QUEENS
8       Noise - Commercial  MANHATTAN
9       Noise - Commercial   BROOKLYN
```

1.4 What's the most common complaint type?

This is a really easy question to answer! There's a `.value_counts()` method that we can use:

```
complaints['Complaint Type'].value_counts()
```

```
Complaint Type
HEATING                       11512
Street Light Condition         2995
GENERAL CONSTRUCTION           2947
PLUMBING                       2148
DOF Literature Request         2093
                              ...
Poison Ivy                        1
Tunnel Condition                  1
Drinking Water                    1
Municipal Parking Facility        1
Trans Fat                         1
Name: count, Length: 158, dtype: int64
```

If we just wanted the top 10 most common complaints, we can do this:

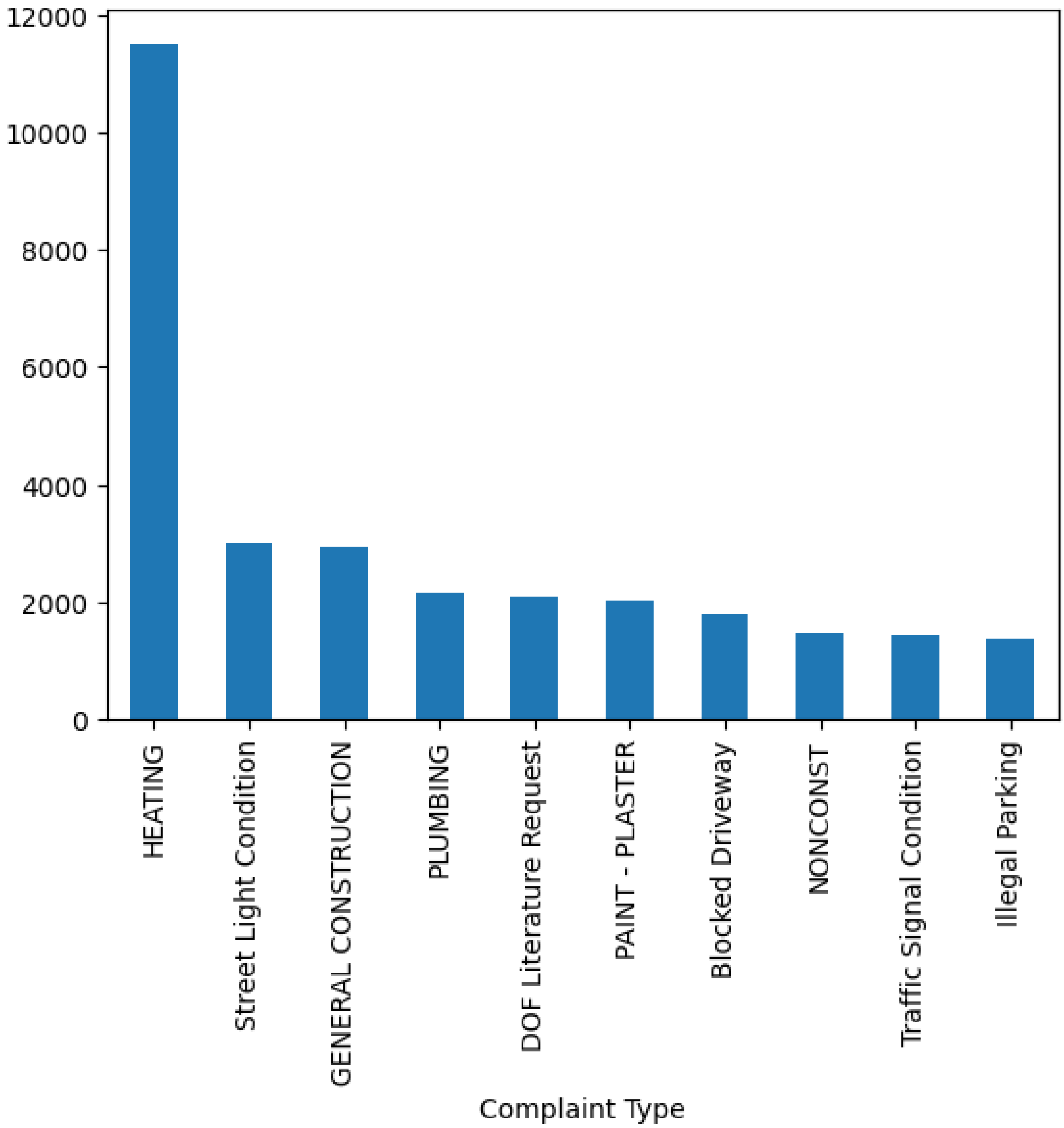
```
complaint_counts = complaints['Complaint Type'].value_counts()
complaint_counts[:10]
```

```
Complaint Type
HEATING                     11512
Street Light Condition       2995
GENERAL CONSTRUCTION         2947
PLUMBING                     2148
DOF Literature Request       2093
PAINT - PLASTER              2031
Blocked Driveway             1804
NONCONST                     1462
Traffic Signal Condition     1426
Illegal Parking              1354
Name: count, dtype: int64
```

But it gets better! We can plot them!

```
complaint_counts[:10].plot(kind='bar')
```

```
<Axes: xlabel='Complaint Type'>
```


12000
10000
8000
6000
4000
2000
0
HEATING
Street Light Condition
GENERAL CONSTRUCTION
PLUMBING
DOF Literature Request
PAINT - PLASTER
Blocked Driveway
NONCONST
Traffic Signal Condition
Illegal Parking
Complaint Type