

Your script should use the file **nycSchools.txt**, that includes information about public elementary schools in New York City. This is a tab-delimited text file that has six columns with data for each school:

**a unique identifier**

**borough** (MN – Manhattan, BX – Bronx, BK – Brooklyn, QN – Queens, and SI – Staten Island)

**school name**

**school enrollment**

**x - coordinate**

**y - coordinate**

The answer to the first four questions below should be a separate function that takes as an argument the **nycSchools.txt** file:

1. Define a function that create a **dictionary** with the borough names as keys and the average school enrollment per each borough as values, and then write them into a report file. The report should be a nicely formatted text file. (20 points)

```
SI          668.65
...         ...
...         ...

BK          692.50
```

2. Define a function that returns the **proportion** of the schools located in Manhattan. (20 points)

3. Define a function that creates a **dictionary** with the borough names as keys and the number of schools with enrollment over 700 students as values, and then write them into a nicely formatted text file. (20 points)

```
SI          10
...         ...
...         ...

BK          43
```

4. Define a function that returns the **bounding box** (extent) of the area that all the schools fall in (minimum and maximum x and y). (20 points)

5. Define a function that reads the file **question\_5.txt**, counts the number of words that start with the same letter, and creates a dictionary. The letter should be the key and the count of the words that start with that letter should be the value. Use the dictionary values to print a histogram

ordered **alphabetically** where each line will contain the corresponding number of letters. (20 points)

```
D
...

U
aaaaaa
...

wwwwwww
```

**Bonus:** Define a function that takes as an argument the **nycSchools.txt** file and returns the **average** nearest neighbor distance between schools. (10 points)

### **Deliverables:**

- **Your Python script (100 points)**