MIDTERM

MATTHEW LANGENDORFER

PART I

FREQUENCY OF TEXT FILE

```
import json
import matplotlib.pyplot as plt

numlist = []
dict = {}

f = open("/Users/ml/Desktop/midterm-project/numbers.txt", "r")
text = f.read()

for num in text:
    if num.isdigit():
        numlist.append(int(num))

for num in numlist:
    if num in dict.keys():
        dict[num] += 1
    else:
        dict[num] = 1

for key, value in dict.items():
    print(f"{key} : {value}")

plt.bar(list(dict.keys()), dict.values(), color='g')
plt.xlabel("Value")
plt.ylabel("Frequency")
plt.title('Frequency of Text File')
plt.show()

y = json.dumps(dict)
print(y)
with open("out.json", "w") as outfile:
    json.dump(y, outfile)
```

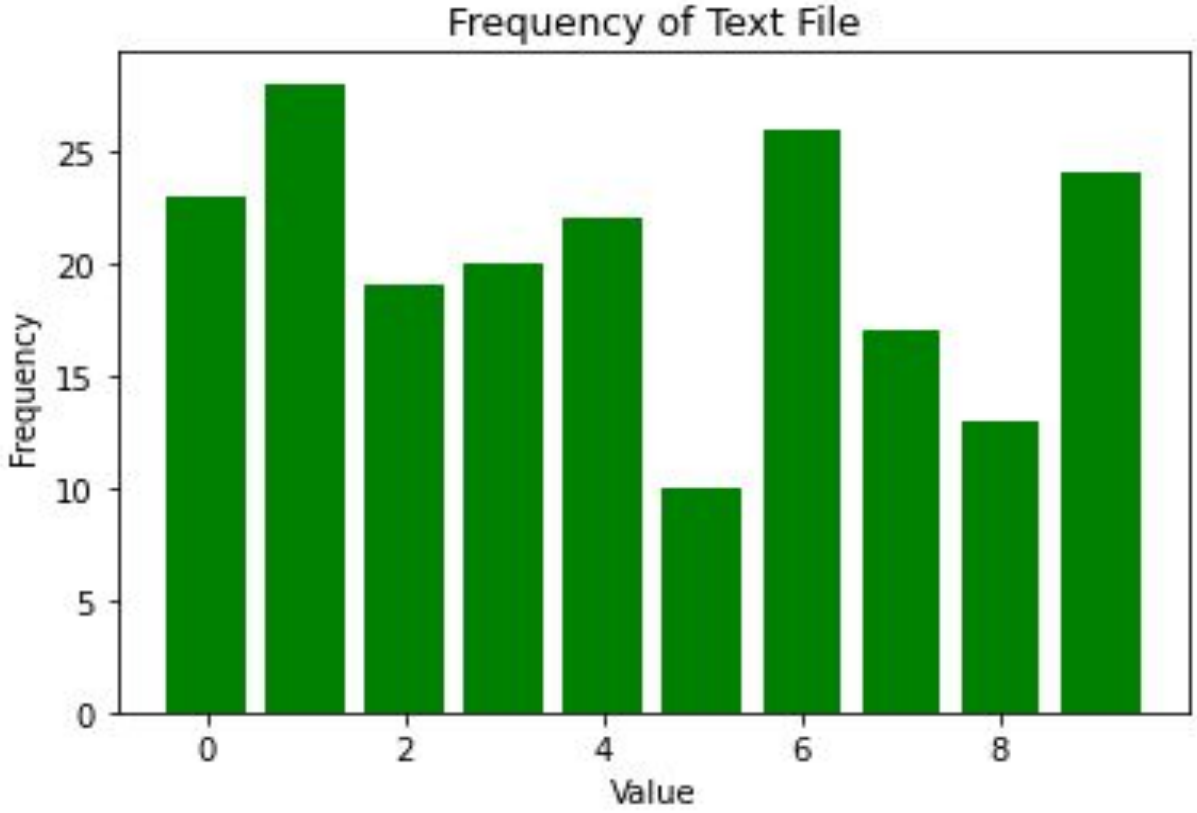
INPUT

```
import json
import matplotlib.pyplot as plt

numlist = []
dict = {}

f = open("/Users/ml/Desktop/midterm-project/numbers.txt", "r")
text = f.read()

for num in text:
    if num.isdigit():
        numlist.append(int(num))

for num in numlist:
    if num in dict.keys():
        dict[num] += 1
    else:
        dict[num] = 1

for key, value in dict.items():
    print(f"{key} : {value}")

plt.bar(list(dict.keys()), dict.values(), color='g')
plt.xlabel("Value")
plt.ylabel("Frequency")
plt.title('Frequency of Text File')
plt.show()

y = json.dumps(dict)
print(y)
with open("out.json", "w") as outfile:
    json.dump(y, outfile)
```

- First I imported the modules that I needed.
- Isdigit checks to see if the iterator is a number, so that it only takes the numbers out of the text file.
- The first loop takes all the numbers in the text and assigns them to a key, while the second loop adds to the value for each of the keys already established for each number in what is now the list of all the numbers.

OUTPUT

2 : 19
4 : 22
6 : 26
8 : 13
5 : 10
1 : 28
9 : 24
0 : 23
7 : 17
3 : 20
{"2": 19, "4": 22, "6": 26, "8": 13, "5": 10, "1": 28, "9": 24, "0": 23, "7": 17, "3": 20}



PART II

NETFLIX VIEWING HABITS

- USED THE SAMPLE NETFLIX .CSV FILE
- FOCUSED ON STAR TREK VIEWING HABITS

GRAPHS / STATS

Total Watchtime:

Timedelta('0 days 20:00:57')

aka 20 hours 57 seconds