

Before the conversion:

[illegible]

After, used a script to convert the Common Log Format information into a .json format. Technique used to parse the data was regex as the data in Common Log Format was all the same for each line of information. The regex implementation separated the information into 8 different categories/fields, "IP\_of\_requesting\_host", "Remote\_user", "Timestamp", "Request\_from\_client", "HTTP\_response\_code", "Size\_of\_bytes\_returned", "Http\_referer", and "Http\_user\_agent":

```
PS C:\Users\fireb\Desktop\Coding Projects\Python Folder> log2json dataset.log -t json -o "C:\Users\fireb\Desktop\Coding Projects\Python Folder\vscode\Lab8\final.json" -e "%(P?IP of requesting_host>.*?) - (P?Remote_user>.*?) \[(P?Timestamp>.*?)\] \"(P?Request_from_client>.*?)\" (P?HTTP_response_code>\\d+) (P?Size_of_bytes_returned>\\d+)\" \"(P?Http_referer>.*?)\" \"(P?Http_user_agent>.*?)\" \"$"
```

After parsing dataset.log we export it as a .json file as stated on the assignment:

[illegible]

Created a script that reads .json and prints out a desired field

Coded a python script that will take the .json file that was created in the previous step and output the information for a desired field by the user.

```

1  from pathlib import Path
2  import json
3  target = Path.home() / "Desktop" / "Coding Projects" / "Python Folder" / ".vscode" / "Lab8" / "final.json"
4
5  with target.open("r") as f:
6      data = json.load(f)
7
8  usr_input = input("Please enter a field of your choice: 1) IP_of_requesting_host 2) Remote_user 3) Remote_user 4) Request_from_client 5) HTTP_response_c
9  final_str = ""
10 for x in data:
11     for key in x.items():
12         if usr_input == key[0]:
13             final_str = final_str + ", " + key[1]
14
15 print("\nAll the results of the field " + usr_input + ": " + final_str[2:])

```

