Python Challenge

**AIM:** To create a program that will extract a given set of IPs from a .txt file of unstructured data and perform Geo IP and RDAP lookups. We collect data obtained from these lookups and create a custom query language to filter the results as required.

**METHODOLOGY:**

1. Extracting IPS – We parse the .txt file in the main function and filter out the IPs and store them in a List.
2. GEO IP lookup – We use the website <http://ipinfodb.com/ip_locator.php> for GEO IP lookups as it provides free service and the html source code is quite easy to understand. We extract the IP, Country, State/Province, City and Zip Code from the results by using Regular Expression for data manipulation in the source code.
3. Database – We build a database using SQL lite3 to store the results of the 5000 IPs
4. Custom Query Language – We allow filtering to be done on Country, State and City. So our query language asks us to enter the details.

Country – Two-character UPPER case code for country. Example – US, CN, IN, IT

State – First Character Upper Case. Example – Colorado, California

City – First character Upper case. Example – Fort Collins, Denver, Los Angeles.

1. Filtering – Based on the filtering parameters entered, we create a SQL query and access our database to filter out the entries and display all of them and the number of entries too.

**NOTES:**

1. Each module has been tested individually and are isolated from each other in their functioning.
2. Major difficulty faced in the challenge was to perform the GEO IP lookup.