Department of Physics, IIT ISM, Dhanbad
Email: sudipto@iitism.ac.in
Course: PHC517 - Computation & Simulation
Instructor: Dr. Sudipto Singha Roy

Practical - 8
Course Credit: 2
Term: Monsoon 2024
TA: A. Majumder, S. Ghosh

Aim of this Practical: Write a program for loading data created in an external program.

Task 1

Here we have some data of covid 19 cases across different states of India.

	Name of State / UT	**Total Confirmed cases**	**Cured /Discharged /Migrated**	**Death**	**Active Cases**
0	Maharashtra	1985	217	149	1619
1	Delhi	1154	27	24	1103
2	Tamil Nadu	1043	50	11	982
3	Rajasthan	804	21	3	780
4	Madhya Pradesh	564	0	36	528
5	Gujarat	516	44	25	447
6	Telengana	504	43	9	452
7	Uttar Pradesh	483	46	5	432
8	Andhra Pradesh	427	11	7	409
9	Kerala	376	179	2	195
10	Jammu and Kashmir	245	6	4	235
11	Karnataka	232	57	6	169
12	Haryana	185	29	3	153
13	West Bengal	152	29	7	116
14	Punjab	151	5	11	135
15	Bihar	64	19	1	44
16	Odisha	54	12	1	41
17	Uttarakhand	35	5	0	30
18	Himachal Pradesh	32	13	1	18
19	Chhattisgarh	31	10	0	21
20	Assam	29	0	1	28
21	Chandigarh	21	7	0	14
22	Jharkhand	19	0	2	17
23	Ladakh	15	10	0	5
24	Andaman and Nicobar Islands	11	10	0	1
25	Puducherry	7	1	0	6
26	Goa	7	5	0	2
27	Manipur	2	1	0	1
28	Tripura	2	0	0	2
29	Arunachal Pradesh	1	0	0	1
30	Mizoram	1	0	0	1

Table 1: Data of covid-19 cases across different parts of India taken from Github

Download the data set in *.csv* format from Github and import it in MATLAB. Manipulate the data in MATLAB to get the following tasks done:-

1. Plot bar graph of total confirmed cases for different states
2. Plot bar graph of combined active and death cases for different states
3. Find the total number of confirmed cases in India.

Solution: MATLAB script will be provided after the class.

Task 2

In any given number, The leading digit, or first significant digit, of a number is the first non-zero digit when you ignore any negative signs or decimal points. This means we're interested in the most significant digit that represents the magnitude of the number.

For example:

- For the number -42.7, the leading digit is 4.
- For 0.00567, the leading digit is 5

Benford's Law states that in many naturally occurring datasets, the leading digit is more likely to be small. Specifically, the distribution of leading digits follows a logarithmic pattern, where the digit d (from 1 to 9) appears with a probability given by the formula:

$$P(d) = \log_{10}(d+1) - \log_{10}(d)$$

This implies that the number 1 appears as the leading digit about 30.1% of the time, while the number 9 appears only about 4.6% of the time.

Task

You have been provided with a dataset. Your task is to check whether this dataset satisfies Benford's Law using MATLAB.

1. **Datasets**: Load the the given dataset into MATLAB.
2. **Calculate Leading Digits**: For the given dataset, extract the leading digits from the numerical values.
3. **Count Frequencies**: Count the frequency of each leading digit (1-9) in the dataset. For instance, how many times you are getting the number x, $x \in [1, 9]$ as the fist significant digit in the data.
4. **Compare with Benford's Distribution**: Plot the observed frequencies (noramlized) against the expected frequencies from Benford's Law.

Download this dataset from here and import it to MATLAB.

Solution: MATLAB script will be provided after the class.
