CS 213: Python Assignment

1. Create a function `randGen()` to generate a 10,000-line file `dataset.txt`, each line of which is as per the following format

   <age>, <gender>, <state>, <phone number>, <height>, <weight>

and the corresponding data is to be generated as described below.

- Age (uniformly distributed integer between 1 and 100)
- Gender (randomly marked as Male or Female)
- State (randomly chosen from the 28 states of India)
- Phone number (random 10-digit numbers starting with 6, 7, 8 or 9)
- Height (Gaussian distributed real number with mean 160 cm and deviation 10 cm)
- Weight (Gaussian distributed real number with mean 70 kg and deviation 5 kg)

2. Create a `Person` class, with the following attributes: `Age`, `Gender`, `State`, `Phone number`, `Height`, `Weight`.

3. Generate 10000 instances of the Person class with data read from `dataset.txt`.

4. Calculate the average height and weight of the dataset and append to `dataset.txt`.

5. Based on the dataset, create the following charts and save them as follows.

- `height.jpg` including two subplots, namely, histogram of male and female heights
- `weight.jpg` including two subplots, namely, histogram of male and female weights
- `gender.jpg` – pie chart of male and female gender
- `phone.jpg` – pie chart of numbers starting with 6, 7, 8 and 9
- `age.jpg` – two line plots (with legend) of cumulative distribution function of male age and female age
- `state.jpg` – bar plot with state name on x-axis and number of people in that state (based on the dataset) as the bar height

What to submit?

- *<roll-no>*.py, where *<roll-no>* should be replaced with your IIT Dharwad roll number.
- It should include all function definitions, class definitions, and the main program.