Bioinformatics programming assignment (that should be doable in a couple of hours):

Write a NGS read simulator that randomly picks out reads from a genome and outputs them as a fastq file (use dummy quality values). (Use read length of 50 bp and generate 100,000 reads from the human genome)
Add a uniform error rate of 0.01 (1% of the time a base is randomly replaced with another base) to the fastq file.
Align the resulting fastq file with bwa and find the error rate (read aligned to a part of the genome other than where it originated from)

Checkin your code to Github and send the link to aditya@indna.in and aditya@pieriandx.com