**BIOS 7659 HW4**

**Guannan Shen**

**(a) Read Information**

The .fastq format contains 4 lines per entry.

The line 1 begins with a “@”

Line 2 is the raw sequence, containing “A”, “T”, “C”, “G”

The first entry is:

@SRR390924.1.1 COLUMBO:1:1:1:1926 length=36

AAAAAAAANAAAAAAAAAAAAAAAAAAAAAAAAAAA

+SRR390924.1.1 COLUMBO:1:1:1:1926 length=36

####################################

This is a low quality read, the “#” represents the third lowest quality level. There are dozens of quality levels. Hence, the third lowest level is very poor quality. The length is 36, single-end (by the paper).

The total number of reads in the file is 3614610.

**(b) Summary Statistics**

Columns returned by “FASTQ Summary Statistics”