

**Part 1.1**, basic info:

File	Identity
1294_S1_L008_R1_001.fastq.gz	Read 1
1294_S1_L008_R2_001.fastq.gz	Index 1
1294_S1_L008_R3_001.fastq.gz	Index 2
1294_S1_L008_R4_001.fastq.gz	Read 2

Command: (\$ zcat 1294\*R1\* | wc -l) 1452986940

Read count per file:  $1452986940/4 = 363,246,735$

Sequence read length: 101

Index length: 8

**1.2b:** Quality score cutoff for sequence reads: 30

Quality score cutoff for index: 35, because this is more critical.

**1.2c:** Indexes with (N) base calls: Index 1- 3,976,613 , Index 2- 3,328,051

zcat \*R2\*.fastq.gz | sed -n '2~4p' | grep -c "N"

zcat \*R3\*.fastq.gz | sed -n '2~4p' | grep -c "N"

\*above commands were run on Talapas interactive node