Class 16

Joseph Girgiss (PID: A17388247)

Table of contents

Plot my results

```
library(readr)
b <- read_tsv("my_results.tsv", col_names = c("qseqid", "sseqid", "pident", "length", "misma
```

```
Rows: 39830 Columns: 12
-- Column specification --------------------------------------------------------
Delimiter: "\t"
chr  (2): qseqid, sseqid
dbl (10): pident, length, mismatch, gapopen, qstart, qend, sstart, send, eva...

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```
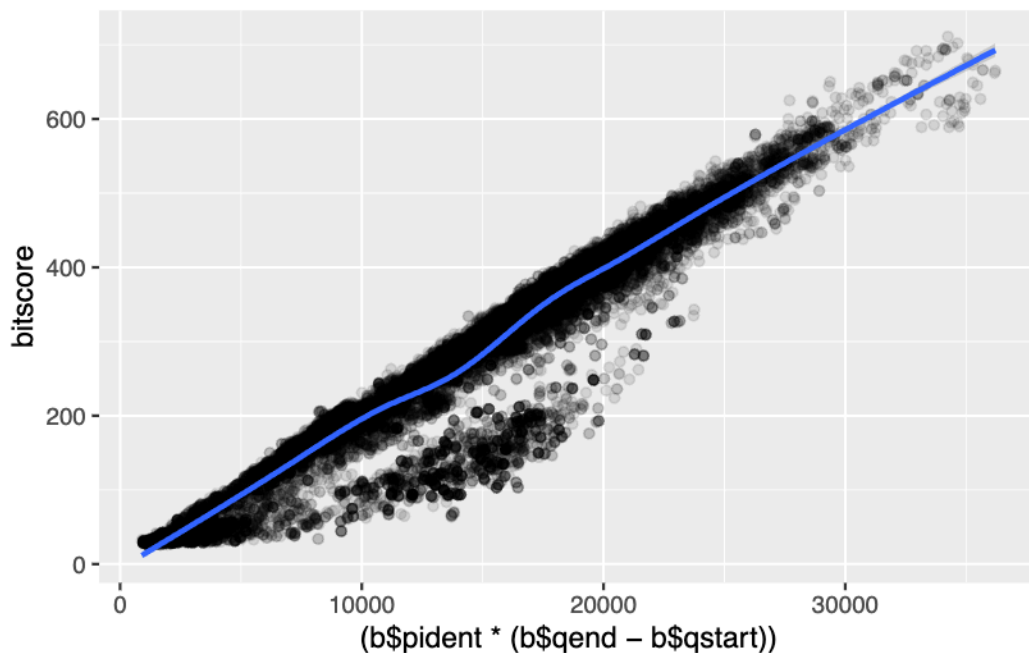
```
library(ggplot2)

ggplot(b, aes((b$pident * (b$qend - b$qstart)), bitscore)) +
  geom_point(alpha=0.1) +
  geom_smooth()
```

```
`geom_smooth()` using method = 'gam' and formula = 'y ~ s(x, bs = "cs")'
```

Lab 16 Questions for Gradescope

Q1. [6pt] List the UNIX bash shell commands to:

1.1 open a secure shell on a remote machine: ssh josephgirgiss@biglabcluster.ucsd.edu

1.2 make a new folder in your home area called "test": mkdir ~/test

1.3 download this file "https://files.rcsb.org/download/5P21.pdb.gz": curl -O https://files.rcsb.org/download/5F

1.4 unzip/decompress it the file: gunzip 5P21.pdb.gz

1.5 print to screen the first 6 lines: head -6 5P21.pdb

1.6 print to lines beginning with ATOM to a new file called "coords.pdb": grep "^ATOM" 5P21.pdb > coords.pdb

Q2. [3pt] List the UNIX commands to copy securely the file "myaln.fa" in your current working directory to your home area on the remote machine "biglabcluster.ucsd.edu":

scp myaln.fa josephgirgiss@biglabcluster.ucsd.edu:~

Q3. [1pt] The alignment file "myaln.fa" is not in your current working directory but it is in your "Downloads" directory. Write the R code to import this alignment to the named object "aln" using a function from the bio3d package.

```
library(bio3d) aln <- read.fasta("~/Downloads/myaln.fa")
```