

class16homework

Laurie Chang A16891192

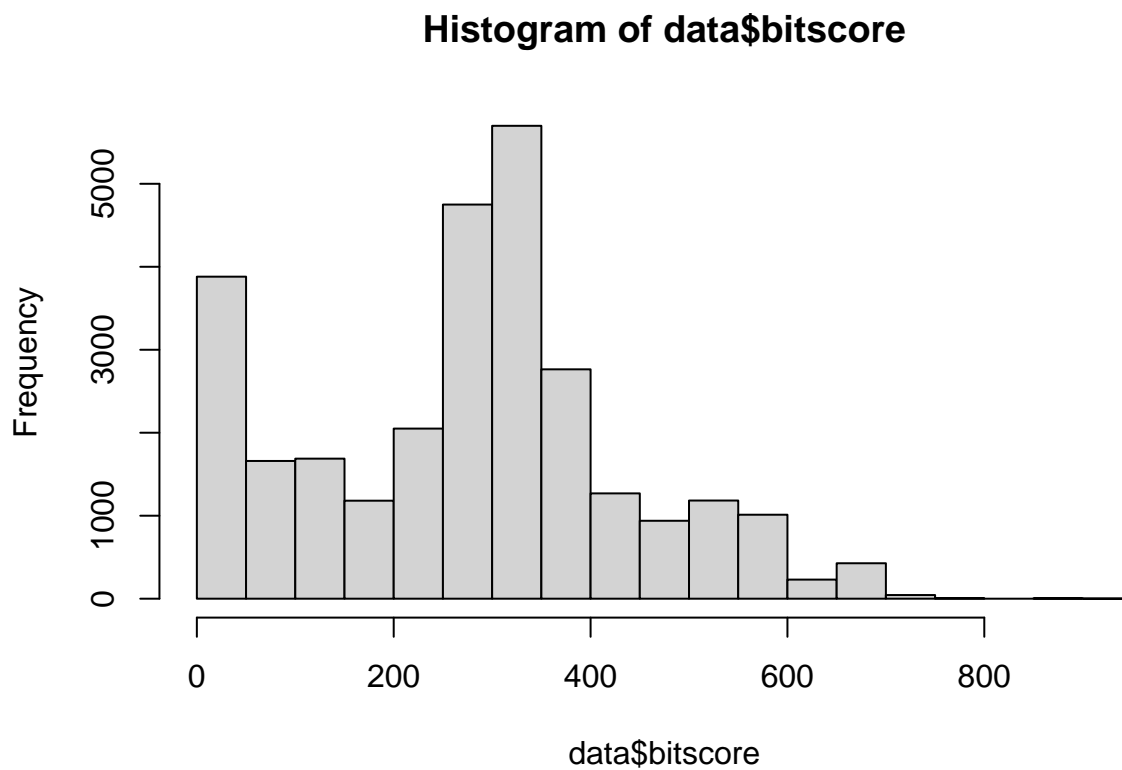
2024-03-01

Reading the tsv file

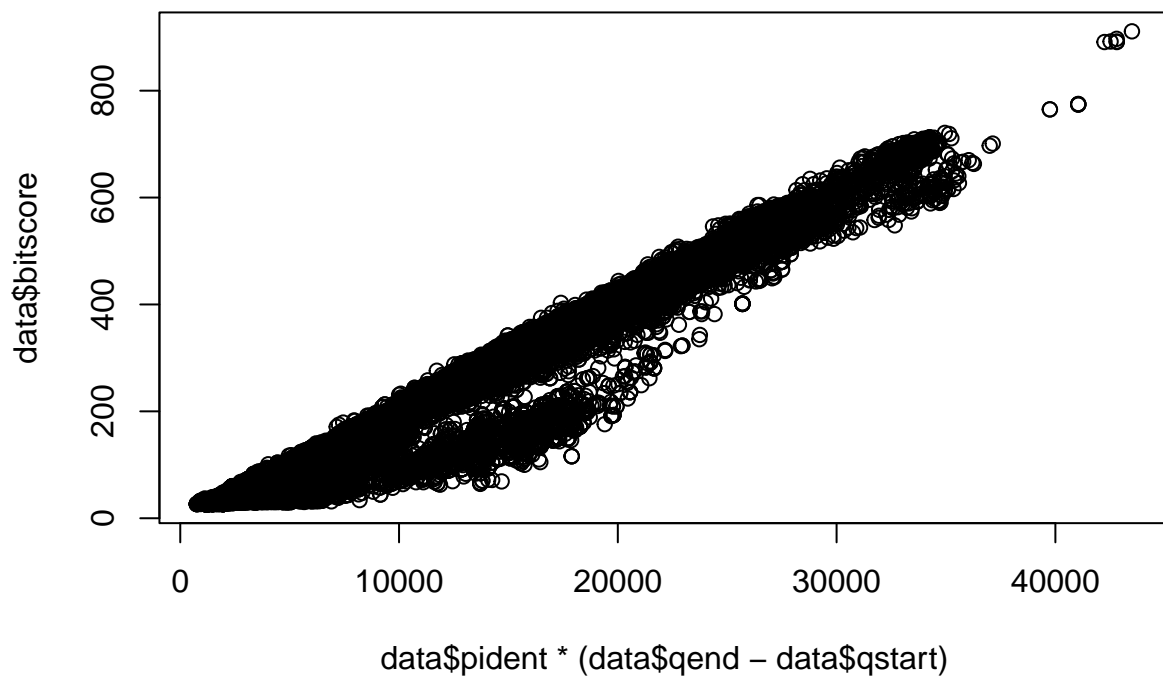
```
data <- read.table("mm-second.x.zebrafish.tsv", col.names = c("qseqid", "sseqid", "pident", "length", "i"
```

Making a Histogram

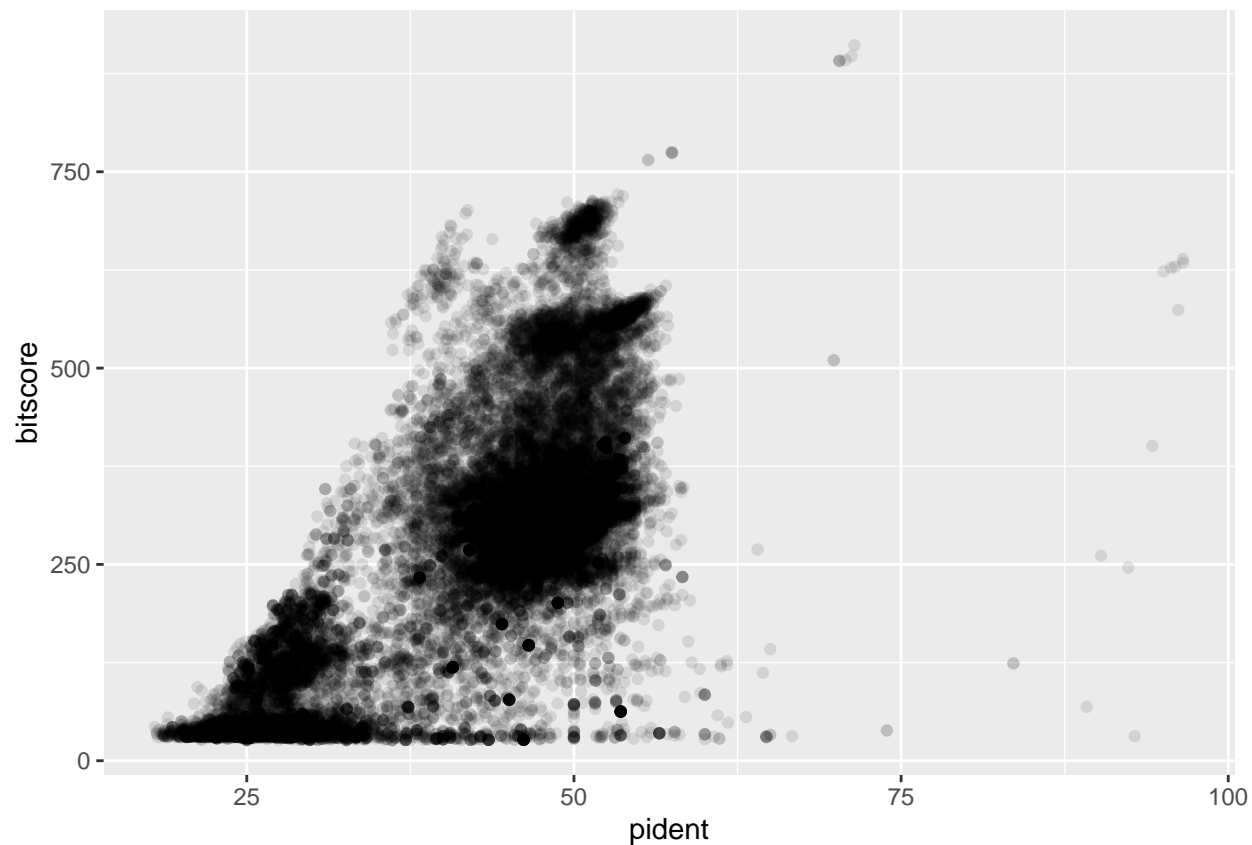
```
hist(data$bitscore, breaks = 30)
```



```
plot(data$pident * (data$qend - data$qstart), data$bitscore)
```



```
library(ggplot2)
ggplot(data, aes(pident, bitscore)) + geom_point(alpha=0.1)
```



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

```
## Warning: Use of 'data$pident' is discouraged.  
## i Use 'pident' instead.
```

```
## Warning: Use of 'data$qend' is discouraged.  
## i Use 'qend' instead.
```

```
## Warning: Use of 'data$qstart' is discouraged.  
## i Use 'qstart' instead.
```

```
## Warning: Use of 'data$pident' is discouraged.  
## i Use 'pident' instead.
```

```
## Warning: Use of 'data$qend' is discouraged.  
## i Use 'qend' instead.
```

```
## Warning: Use of 'data$qstart' is discouraged.  
## i Use 'qstart' instead.
```

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

