Week 11
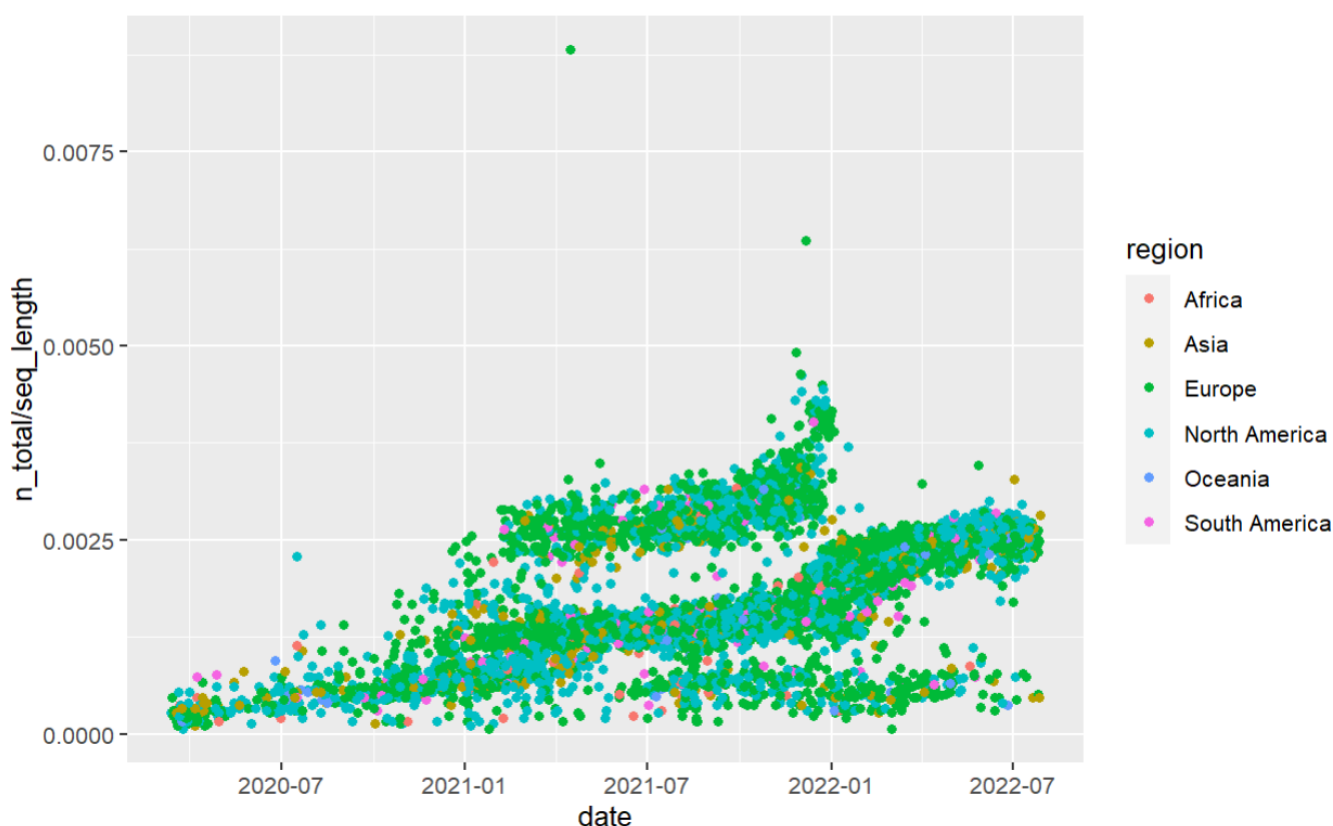
1. What happened with the stratification of data?

In this plot, we can see that there are many n>50 before 2022-01, which looks weird. If this happened due to the limitation of detection condition?
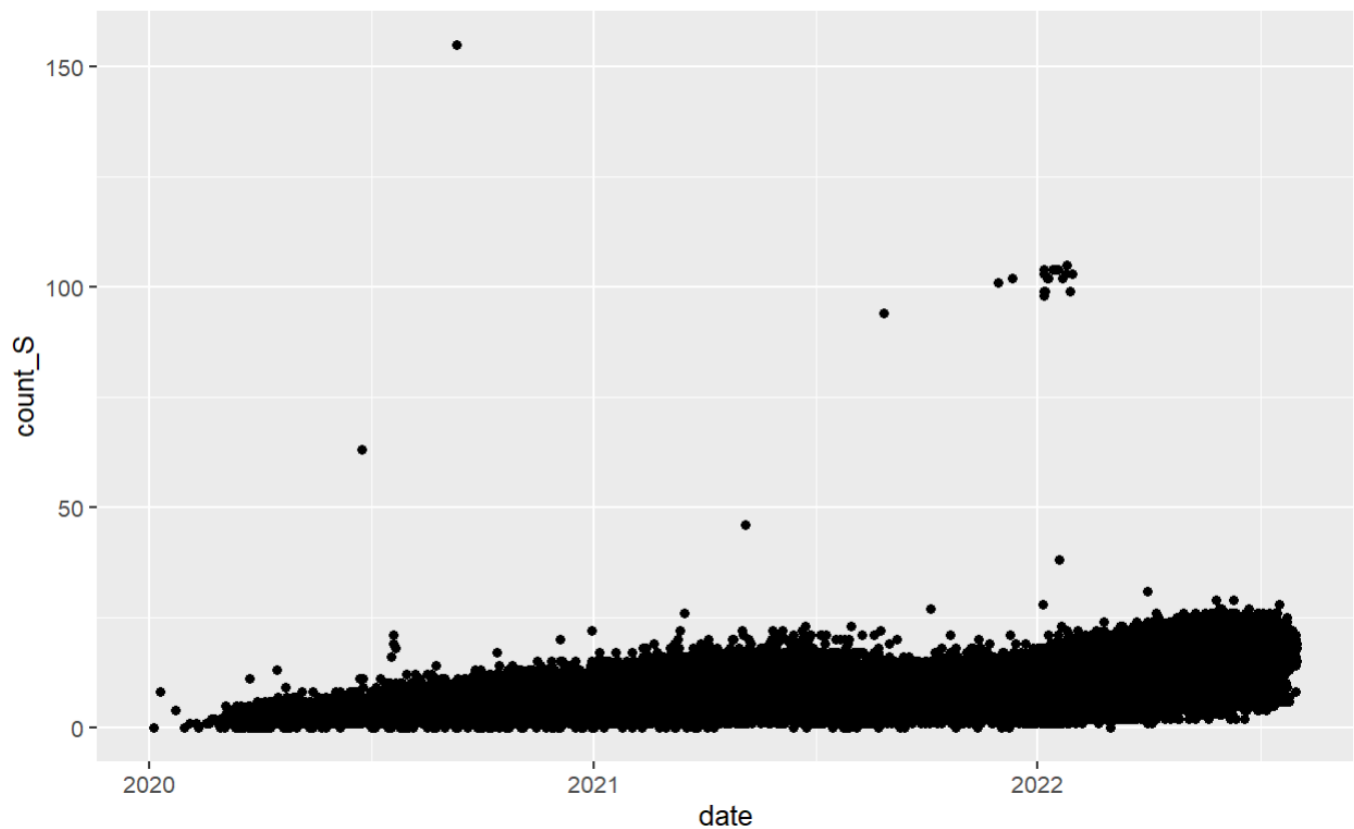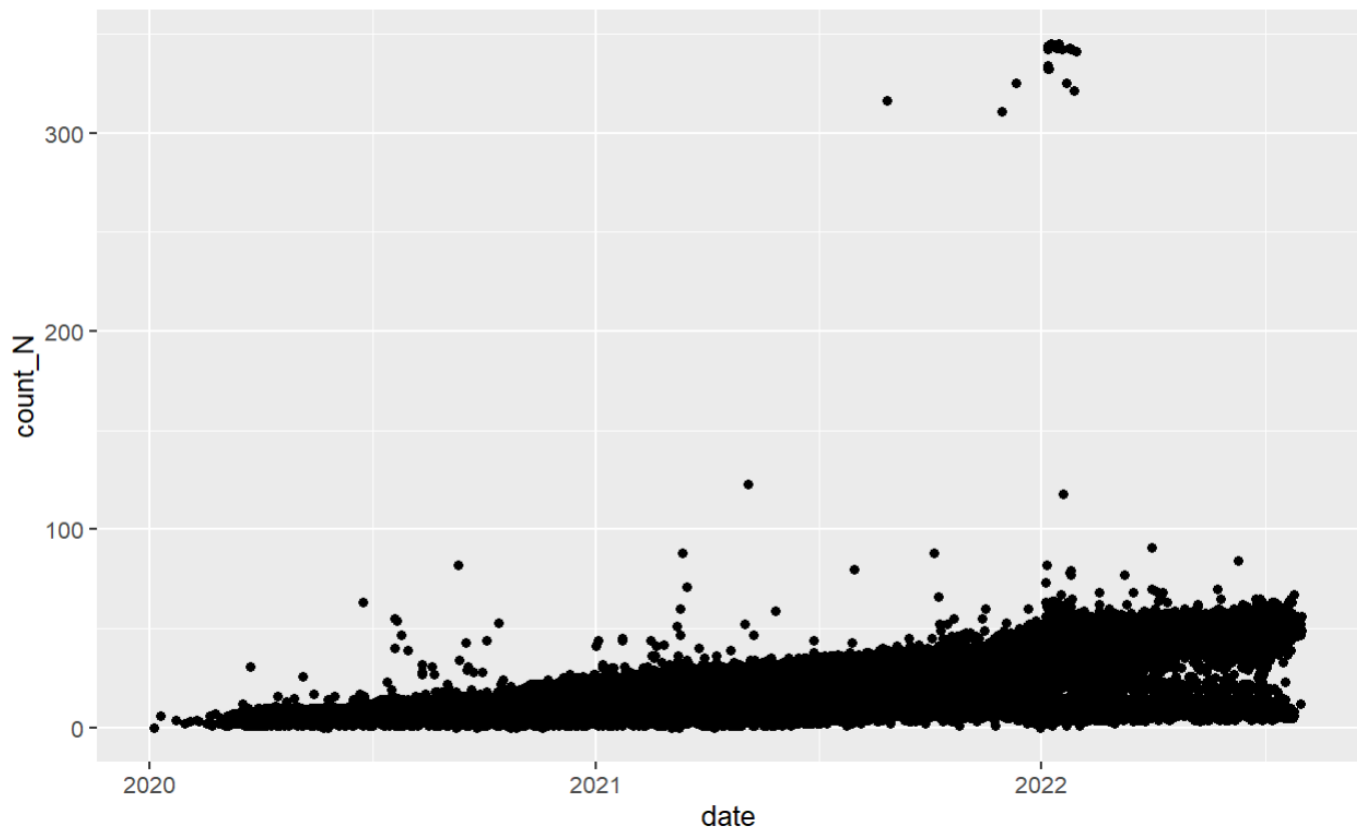
country <chr>	n <int>
USA	8019
United Kingdom	6524
Denmark	1370
Germany	1245
France	724
Canada	523
Japan	486
Switzerland	450
Italy	372
Sweden	359
1-10 of 119 rows	Previous 1 2 3 4 5 6 ... 12 Next

It seems no? But we don't know why then.

2. Non-syn and Syn number change during time

It seems that both mutate gradually?
Then do dn/ds test!!!

3. alternative of dn/ds

Since we don't have the gene sequencial data, we cannot make count on number of syn or non-syn sites.

ZLY formula(rdNdS)

$$\frac{dN}{dS} = \frac{\#N/\$N}{\#S/\$S}$$

$$dNdSe = \frac{\#Ne/\$N}{\#Se/\$S}$$

$$dNdSy = \frac{\#Ny/\$N}{\#Sy/\$S} = \frac{\#Ne/\#Se}{\#Ny/\#Sy}$$

s.t. same species in diff env.

$$\Rightarrow \$y = \frac{\#N/\$S}{}$$

$$\$ye = \#Ne/\#Se$$

$$\$yy = \#Ny/\#Sy$$

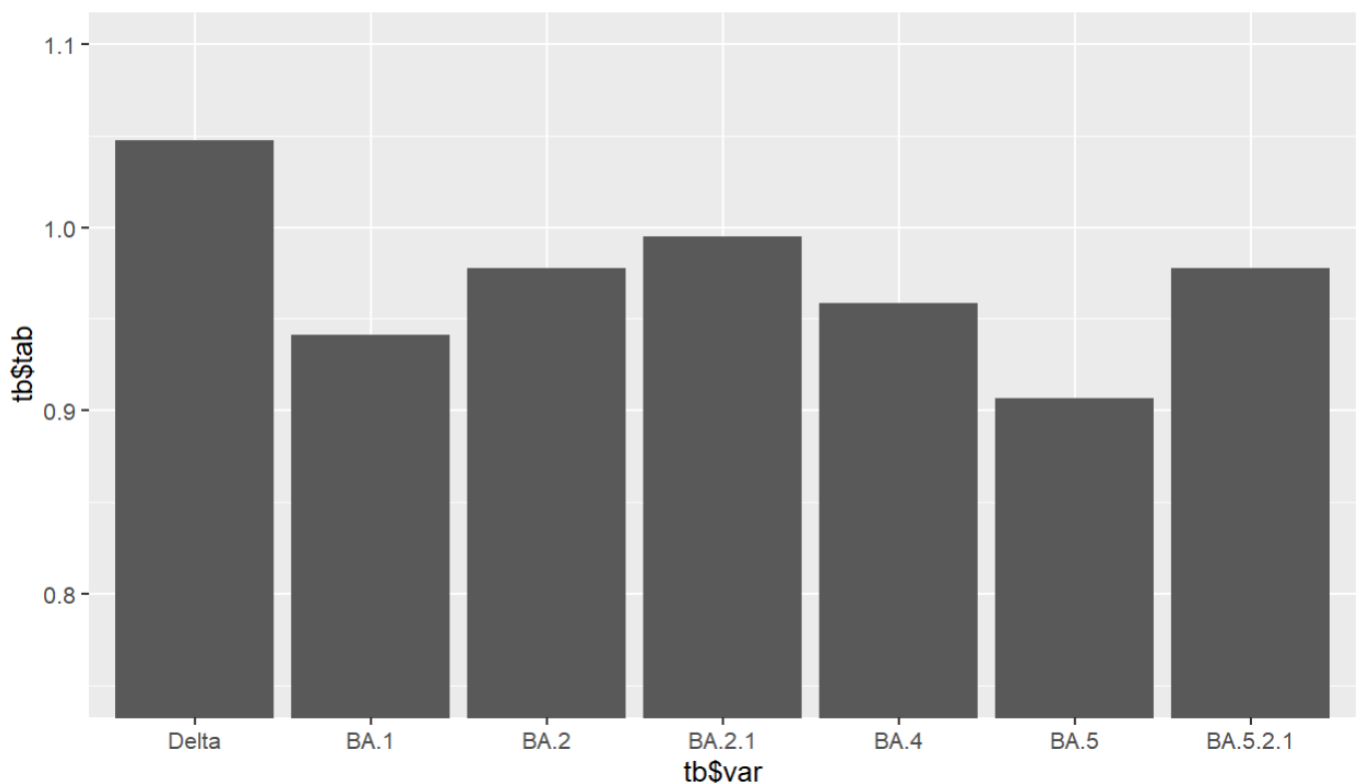$$\Rightarrow rdNds = \frac{\$ye}{\$yy} = \frac{\#Ne/\#Se}{\#Ny/\#Sy}$$

rdNdS: ratio of dn/ds

	Delta	BA.1	BA.2	BA.2.1	BA.4	BA.5	BA.5.2.1
rdNdS	1.047233	0.941485	0.977622	0.994798	0.958456	0.906679	0.977549

rdNdS in different lineages and variants. **Numerator represents the elder group, and Denominator**

represents the youth group. It shows that Delta variants were better adapted to elder group, while Omicron are inverted.

Grouping by different age, it's because generally we assume that elder group has weaker immune system than youth, thus we can regard them as two different immune environments.
But it's very surprising to see Omicron is better adapted to the youth.

Q1: In 2., count_S decreased in a period? why not gradually?

Q2: About rdNdS, why prefer youth group now?

plausible reasons:

1. in both groups, it is negative selected, dnds < 1, while new variants get stronger adaptation in **youth group**.
2. dnds(youth) > 1, dnds(elder) < 1, but why?
3. both > 1, least possible, due to the selection in nature