Step 2 Look at the datasets

Region

As we can see on the website, Region metrics are available for regions of England. I am interested in the South West and metrics that start with "New":

```
## [1] "areaCode"
## [2] "areaName"
## [3] "areaType"
## [4] "date"
## [5] "newPeopleVaccinatedFirstDoseByVaccinationDate"
## [6] "newPeopleVaccinatedSecondDoseByVaccinationDate"
## [7] "newPeopleVaccinatedThirdInjectionByVaccinationDate"
```

We have additional columns. Let's look at them.

areaCode

```
## [1] "E12000009"
```

areaName

```
## [1] "South West"
```

areaType

```
## [1] "region"
```

So, we do not need to look at them in the future because these columns are used for filtering that we have already done on the website.

Let's prepare data for the plotting.

- Rename columns and columns

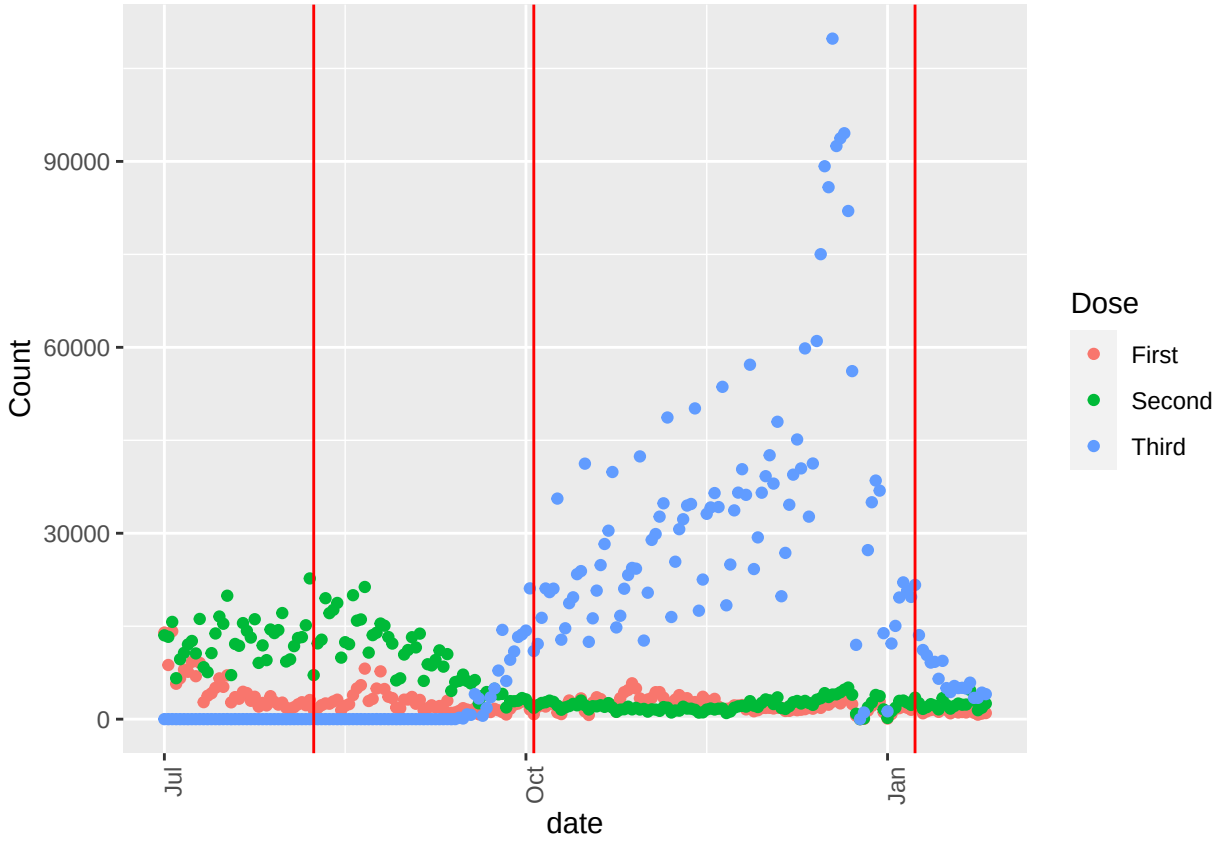
- Create long table

areaCode	areaName	areaType	date	First	Second	Third	MonthYear
E12000009	South West	region	2022-01-26	986	2520	4034	1.2022
E12000009	South West	region	2022-01-25	899	1845	4283	1.2022
E12000009	South West	region	2022-01-24	723	1445	3441	1.2022
E12000009	South West	region	2022-01-23	1035	3007	3439	1.2022
E12000009	South West	region	2022-01-22	1822	4709	5896	1.2022
E12000009	South West	region	2022-01-21	1085	2362	4944	1.2022

date	MonthYear	Dose	Count
2022-01-26	1.2022	First	986
2022-01-25	1.2022	First	899
2022-01-24	1.2022	First	723
2022-01-23	1.2022	First	1035
2022-01-22	1.2022	First	1822
2022-01-21	1.2022	First	1085

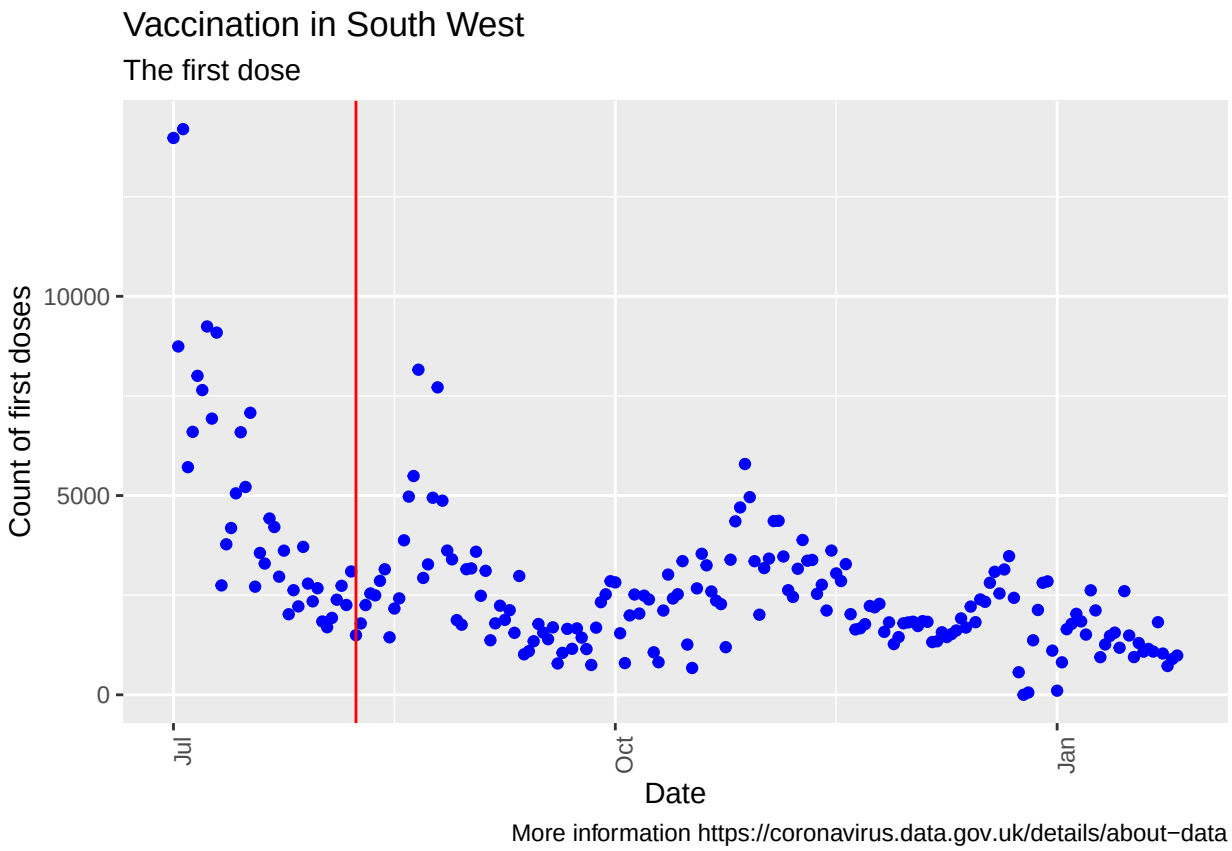
Let's plot something.

I got my jabs on 8 August 2021, 3 October 2021, 8 January 2022. I want to plot the graph that will show the count of jabs after 1 July 2021.

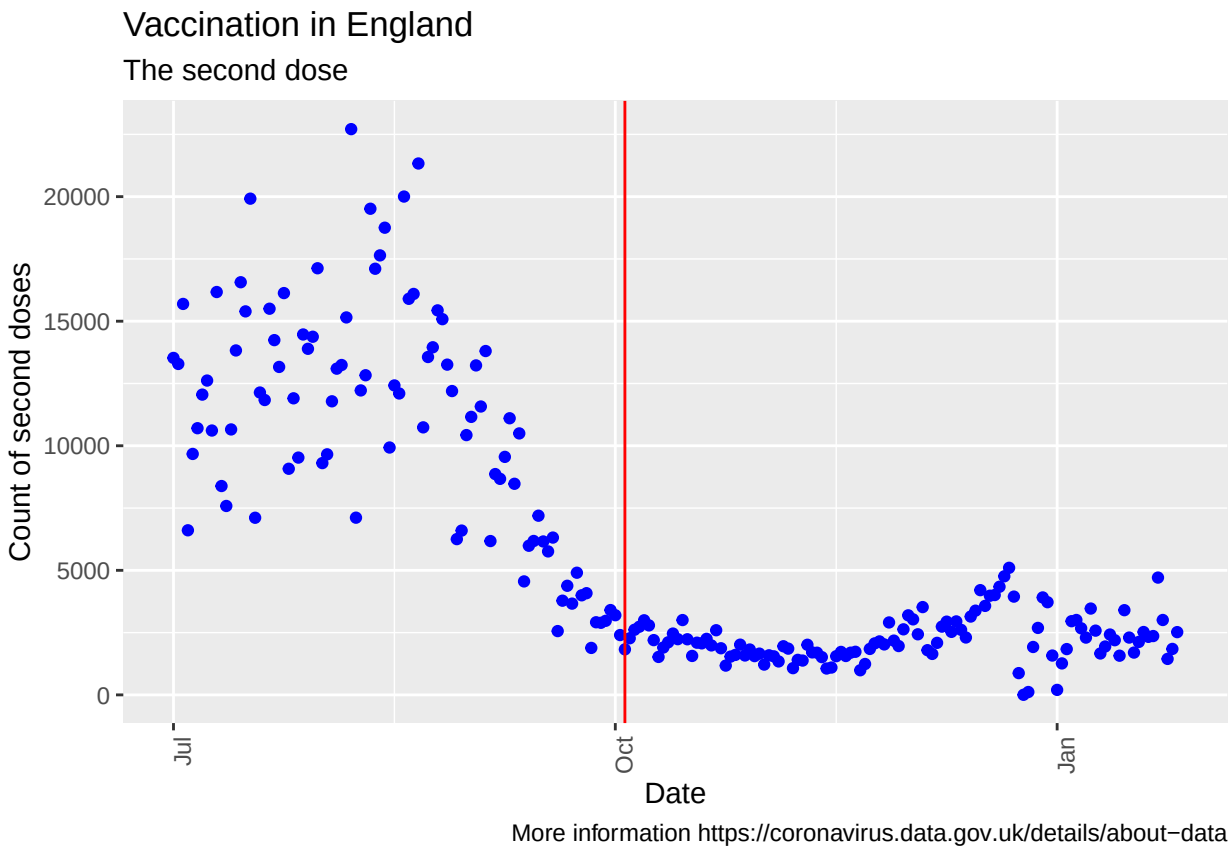


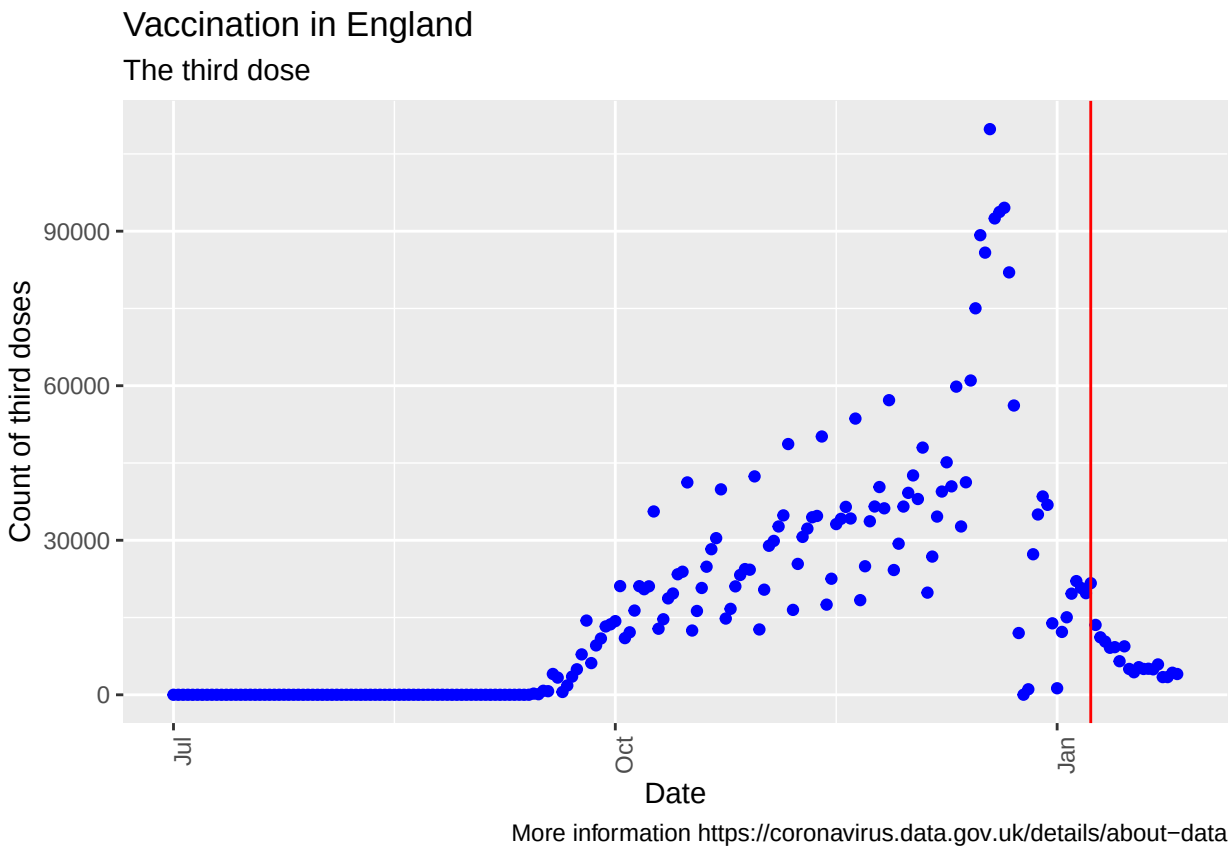
The result is not beautiful because of the active growth of the third jabs count at the end of 2021.

Let's plot them separately.



It is so interesting why the graph is wavy.

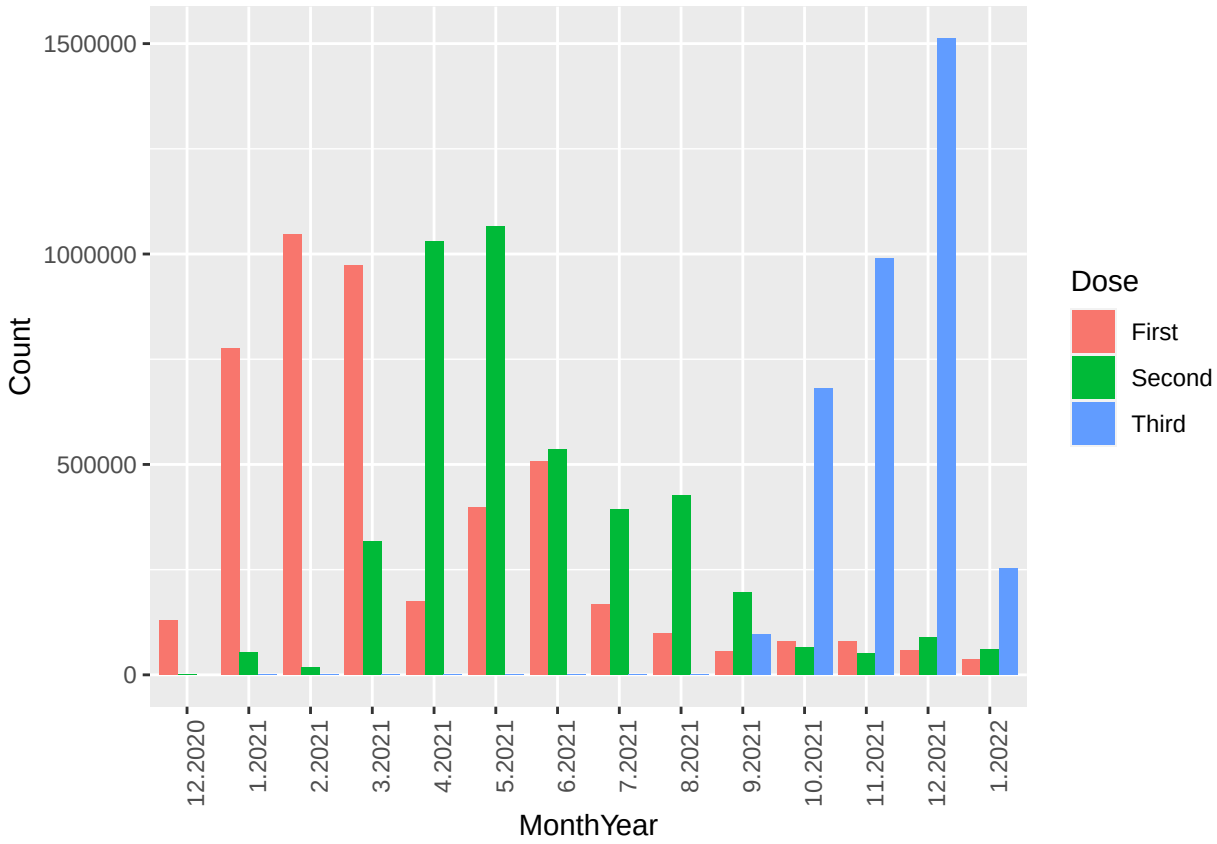

Vaccination in England
The second dose
Count of second doses
20000
15000
10000
5000
0
Jul
Oct
Jan
Date
More information https://coronavirus.data.gov.uk/details/about−data



We can see when the active phase of vaccination by the third dose started.

Let's calculate the date.

```
## Warning: Removed 1 rows containing missing values (geom_col).
```


1500000
1000000
500000
0
Count
12.2020
1.2021
2.2021
3.2021
4.2021
5.2021
6.2021
7.2021
8.2021
9.2021
10.2021
11.2021
12.2021
1.2022
MonthYear
Dose
First
Second
Third