# Coronavirus (COVID-19) vaccination in the UK

Elena Basargina

14/02/2022

# Introduction

## What do I know about Bristol?

- This city is a part of the UK, England, and South West.
- There are two universities.

# Data

I use the dataset from the website about Coronavirus (COVID-19) in the UK.

So, I am interested in data about the UK, England, South West, and Bristol.

**Bristol**

| age | date | First | Second | Third |
|---|---|---|---|---|
| 12_15 | 2022-02-09 | 28 | 45 | 0 |

**South West**

| date | First | Second | Third |
|---|---|---|---|
| 2022-01-26 | 986 | 2520 | 4034 |

# What do I know about the data?

- There may be incorrect data
- There are missing values
- The counts of doses per day spread out.
- There are sets of small and large counts.

# Questions 1

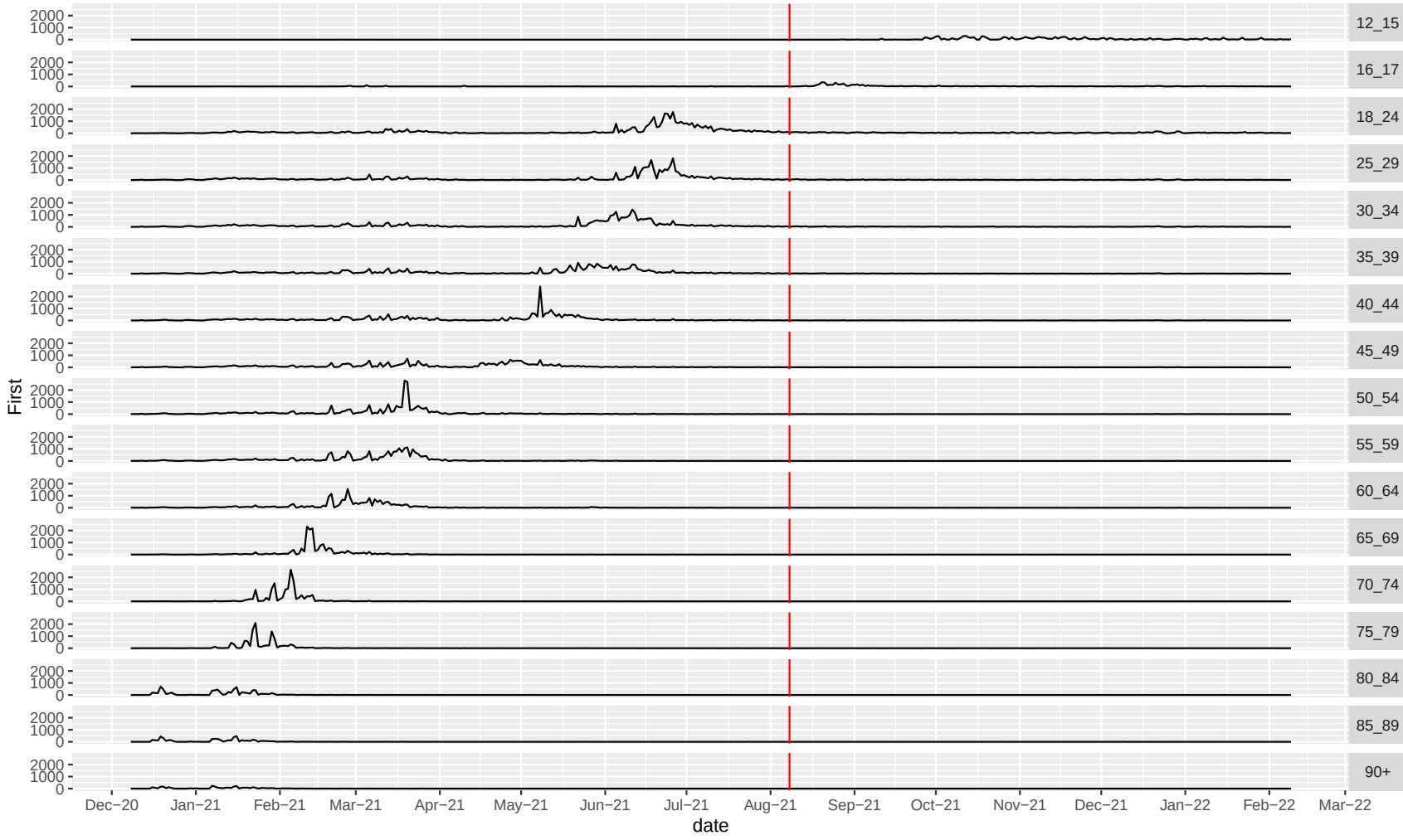
I was vaccinated by

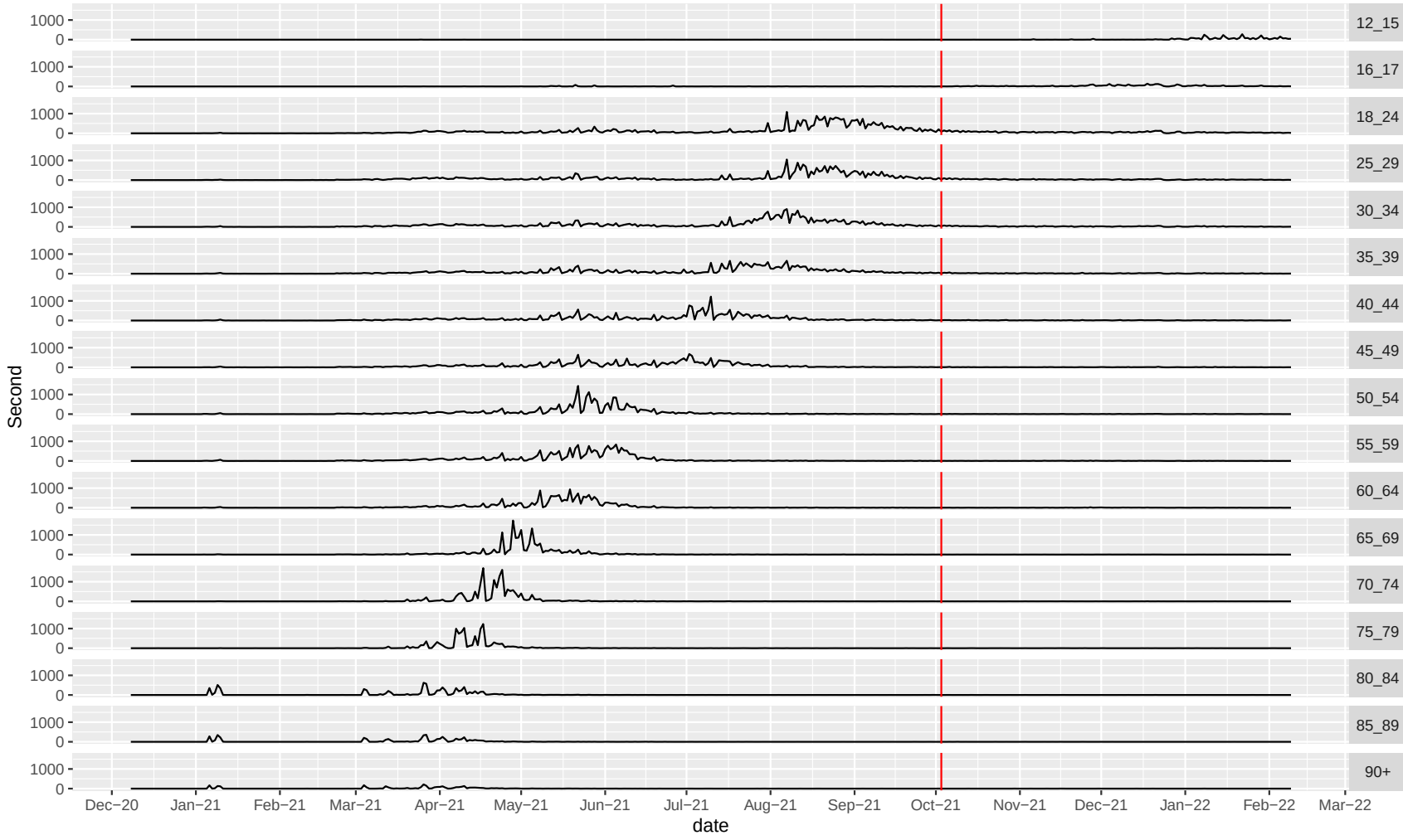
- the first dose on $2021-08-08$,
- the second dose on $2021-10-03$,
- the booster dose on $2022-01-08$.

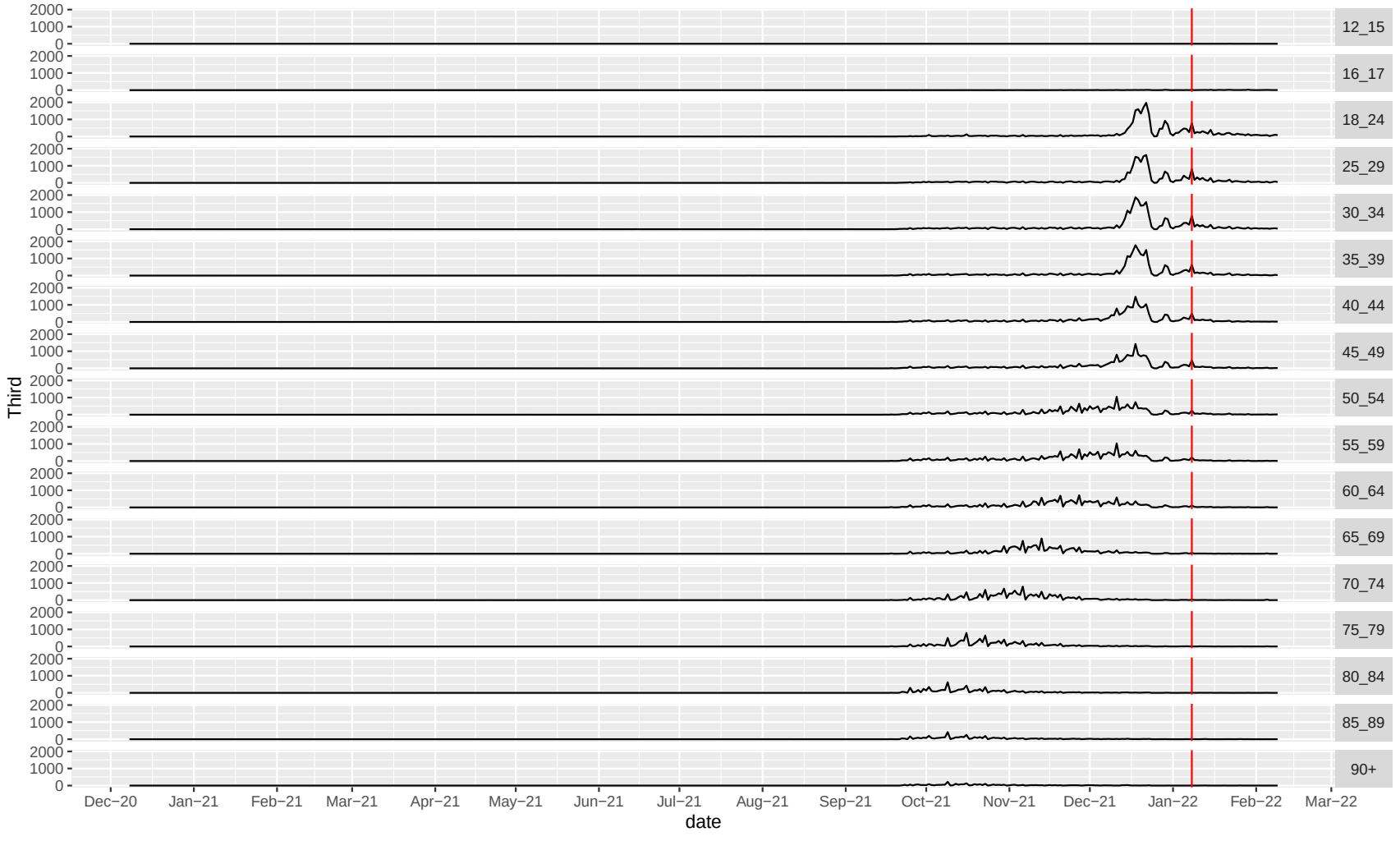
How many people got their jabs with me?

## Bristol

| date | age | Dose | Count |
|---|---|---|---|
| 2021-08-08 | 25_29 | First | 16 |
| 2021-10-03 | 25_29 | Second | 37 |
| 2022-01-08 | 25_29 | Third | 806 |

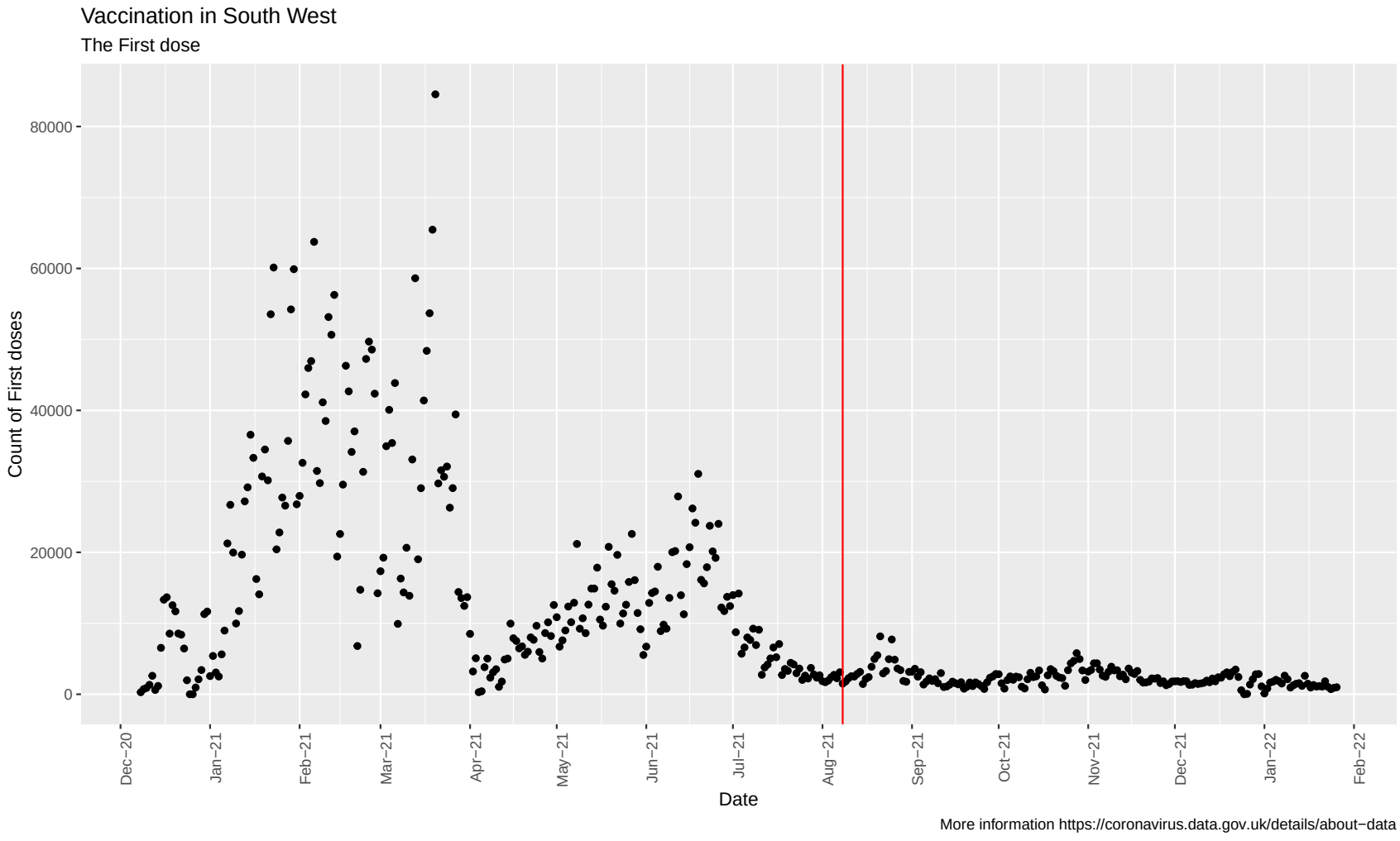

First
2000
1000
0
12_15
16_17
18_24
25_29
30_34
35_39
40_44
45_49
50_54
55_59
60_64
65_69
70_74
75_79
80_84
85_89
90+
Dec−20
Jan−21
Feb−21
Mar−21
Apr−21
May−21
Jun−21
Jul−21
Aug−21
Sep−21
Oct−21
Nov−21
Dec−21
Jan−22
Feb−22
Mar−22
date

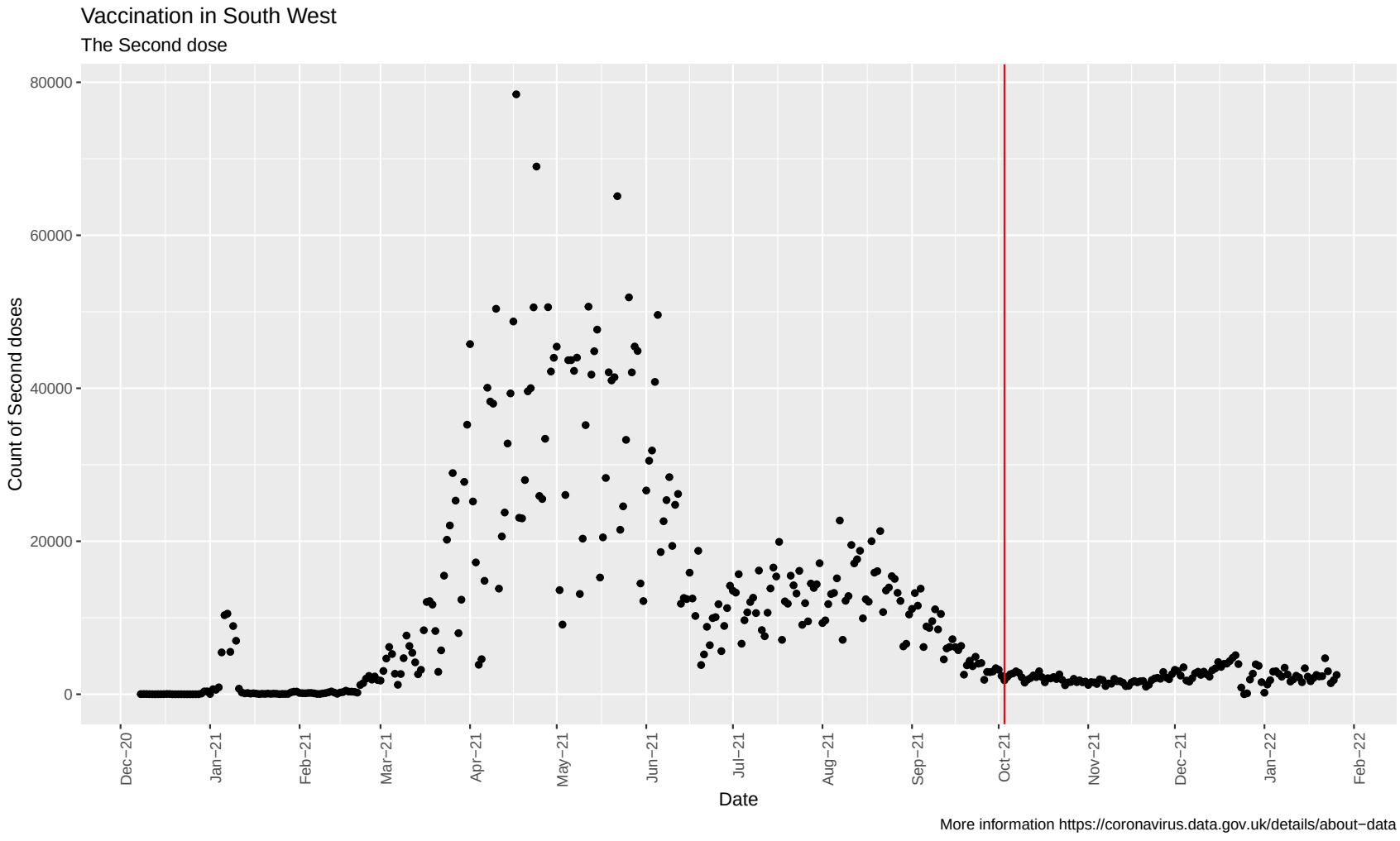

Second
1000
0
12_15
16_17
18_24
25_29
30_34
35_39
40_44
45_49
50_54
55_59
60_64
65_69
70_74
75_79
80_84
85_89
90+
Dec−20
Jan−21
Feb−21
Mar−21
Apr−21
May−21
Jun−21
Jul−21
Aug−21
Sep−21
Oct−21
Nov−21
Dec−21
Jan−22
Feb−22
Mar−22
date

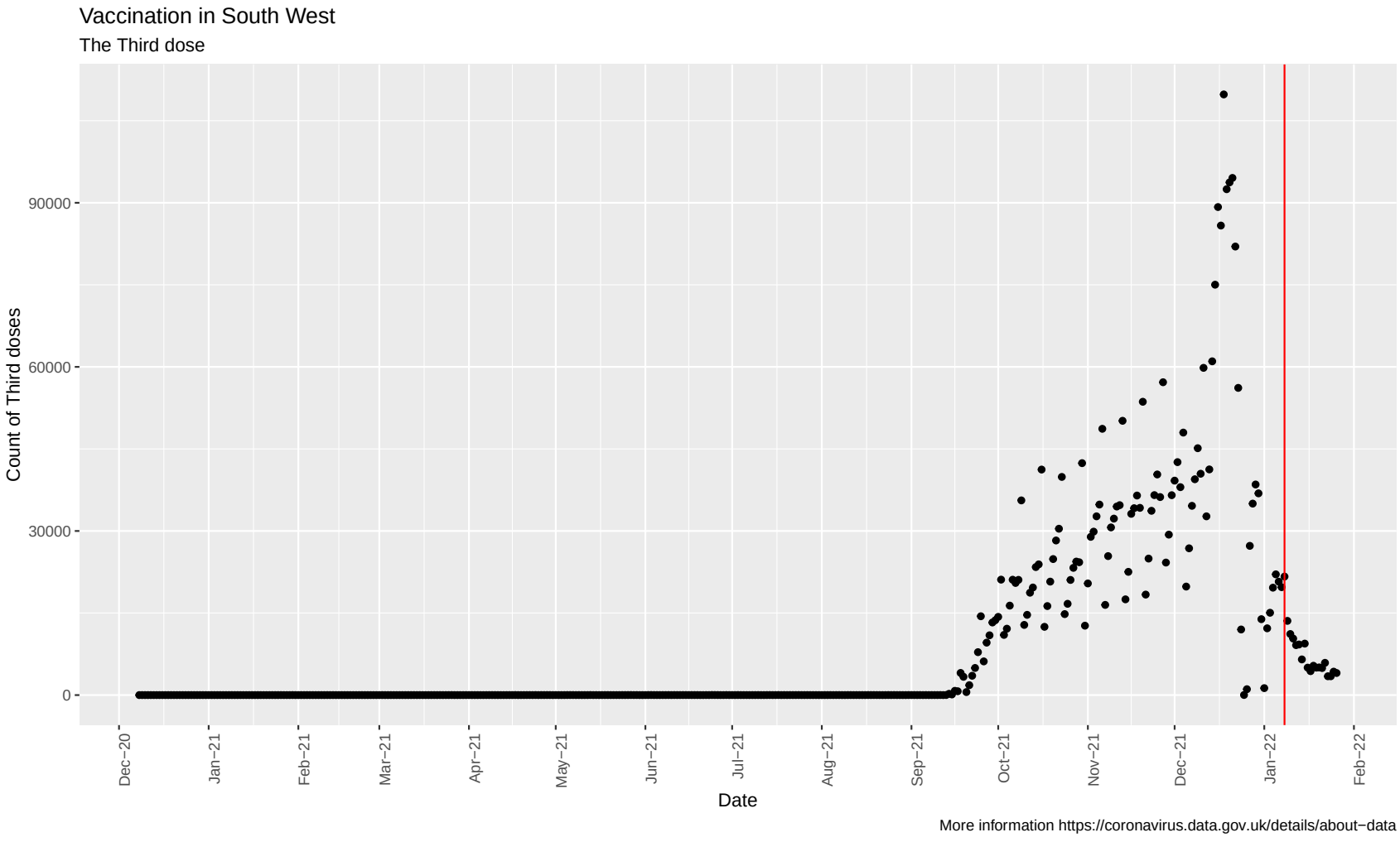

Third
2000
1000
0
12_15
16_17
18_24
25_29
30_34
35_39
40_44
45_49
50_54
55_59
60_64
65_69
70_74
75_79
80_84
85_89
90+
Dec−20
Jan−21
Feb−21
Mar−21
Apr−21
May−21
Jun−21
Jul−21
Aug−21
Sep−21
Oct−21
Nov−21
Dec−21
Jan−22
Feb−22
Mar−22
date

## South West

| date | Dose | Count |
|---|---|---|
| 2021-08-08 | First | 1496 |
| 2021-10-03 | Second | 1828 |
| 2022-01-08 | Third | 21664 |


Vaccination in South West
The First dose
Count of First doses
80000
60000
40000
20000
0
Dec−20
Jan−21
Feb−21
Mar−21
Apr−21
May−21
Jun−21
Jul−21
Aug−21
Sep−21
Oct−21
Nov−21
Dec−21
Jan−22
Feb−22
Date
More information https://coronavirus.data.gov.uk/details/about−data


Vaccination in South West
The Second dose
Count of Second doses
80000
60000
40000
20000
0
Dec−20
Jan−21
Feb−21
Mar−21
Apr−21
May−21
Jun−21
Jul−21
Aug−21
Sep−21
Oct−21
Nov−21
Dec−21
Jan−22
Feb−22
Date
More information https://coronavirus.data.gov.uk/details/about−data


Vaccination in South West
The Third dose
Count of Third doses
90000
60000
30000
0
Dec−20
Jan−21
Feb−21
Mar−21
Apr−21
May−21
Jun−21
Jul−21
Aug−21
Sep−21
Oct−21
Nov−21
Dec−21
Jan−22
Feb−22
Date
More information https://coronavirus.data.gov.uk/details/about−data

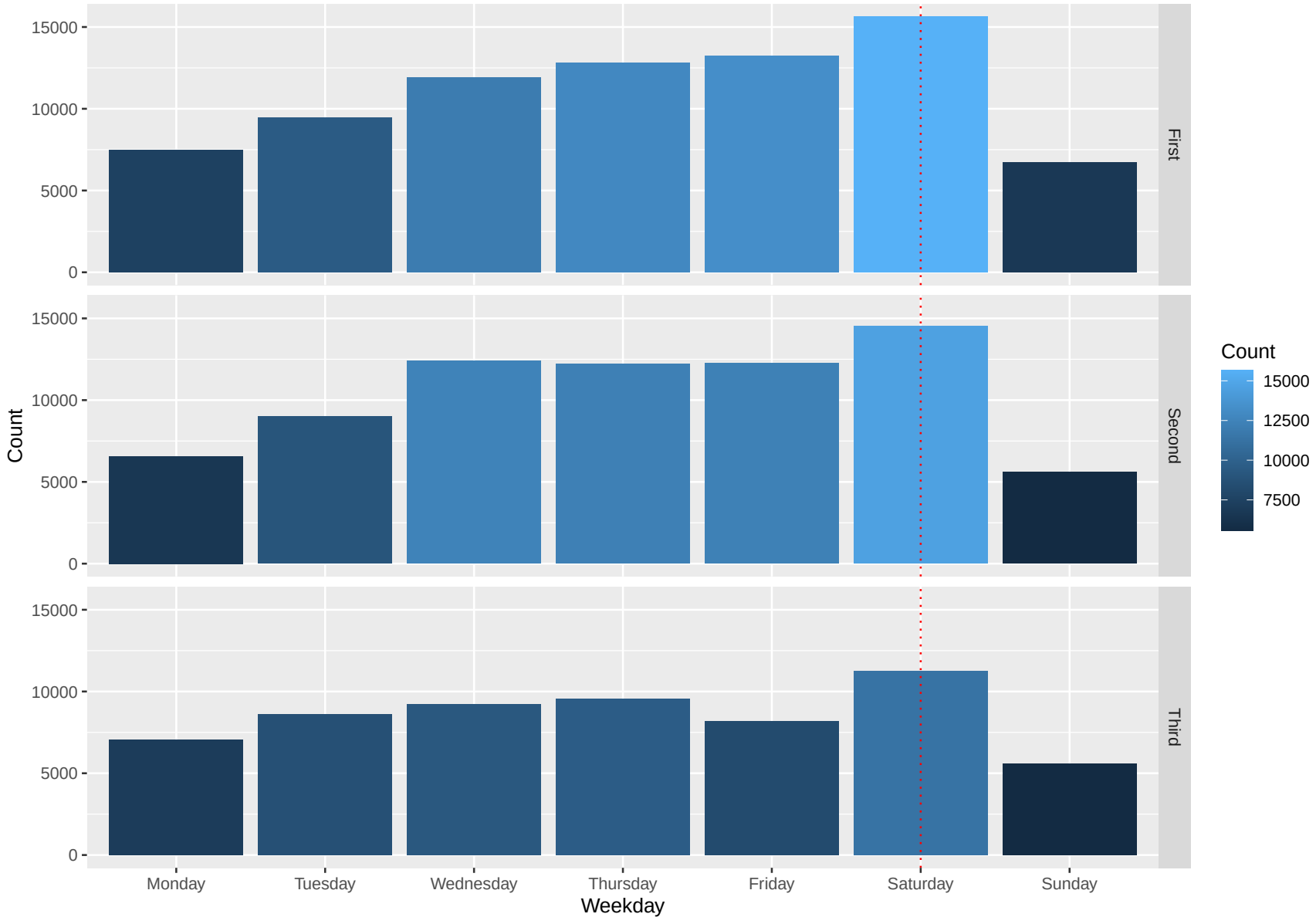
# Conclusion

I got my jabs in the not active phases of the vaccination, but I got the third one with a lot of people.

Also, I can say that there are waves. So, the count of jabs depends on dates. Keep this fact in mind.

# Question 2

When do people prefer to get a jab: weekdays or weekends/Saturdays or Sundays?


15000
10000
5000
0
First
Second
Third
Count
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
Weekday
Count
15000
12500
10000
7500

Are there other dependencies?

First of all, I am going to use Regression Machine Learning models:

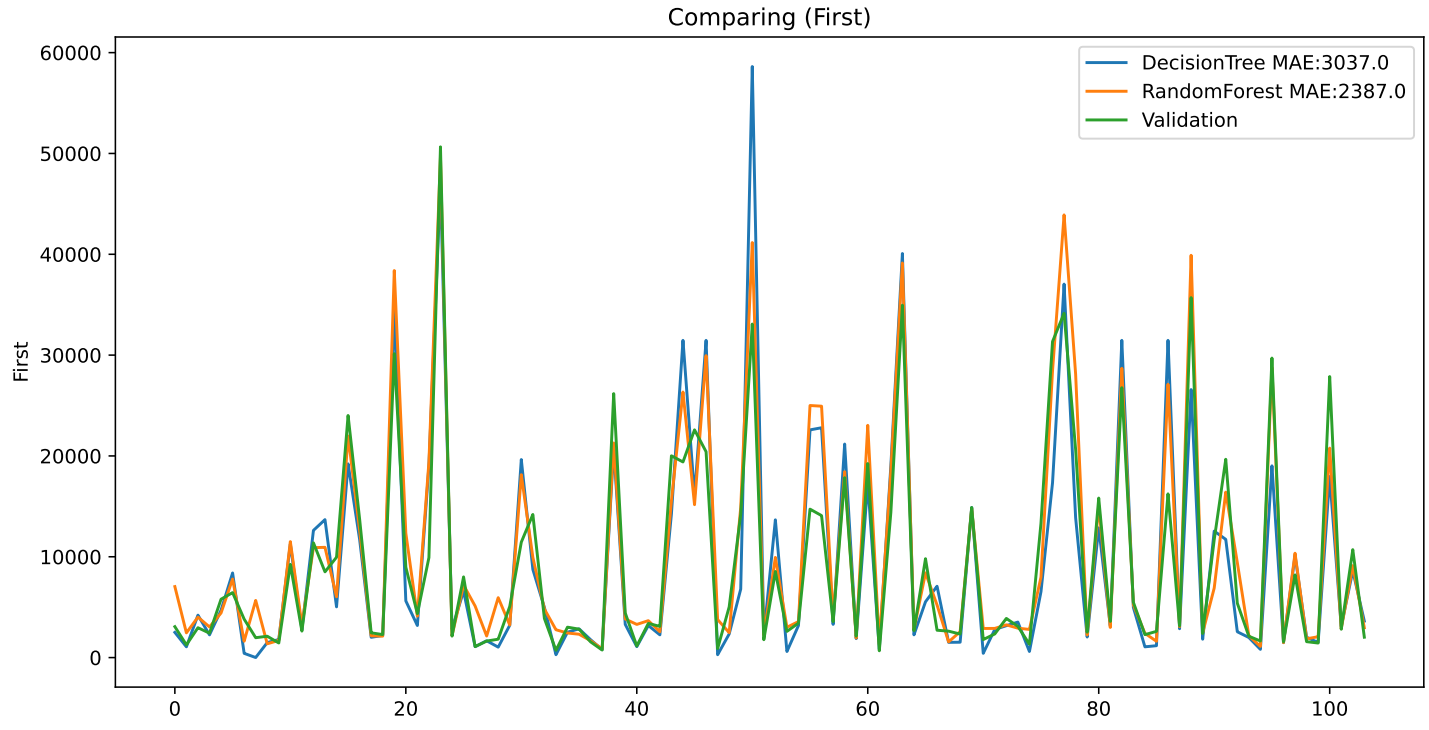
- Decision Tree
- Random Forest.

Which columns should I choose?

- “Weekday”
- “Year”
- “DayOfYear”

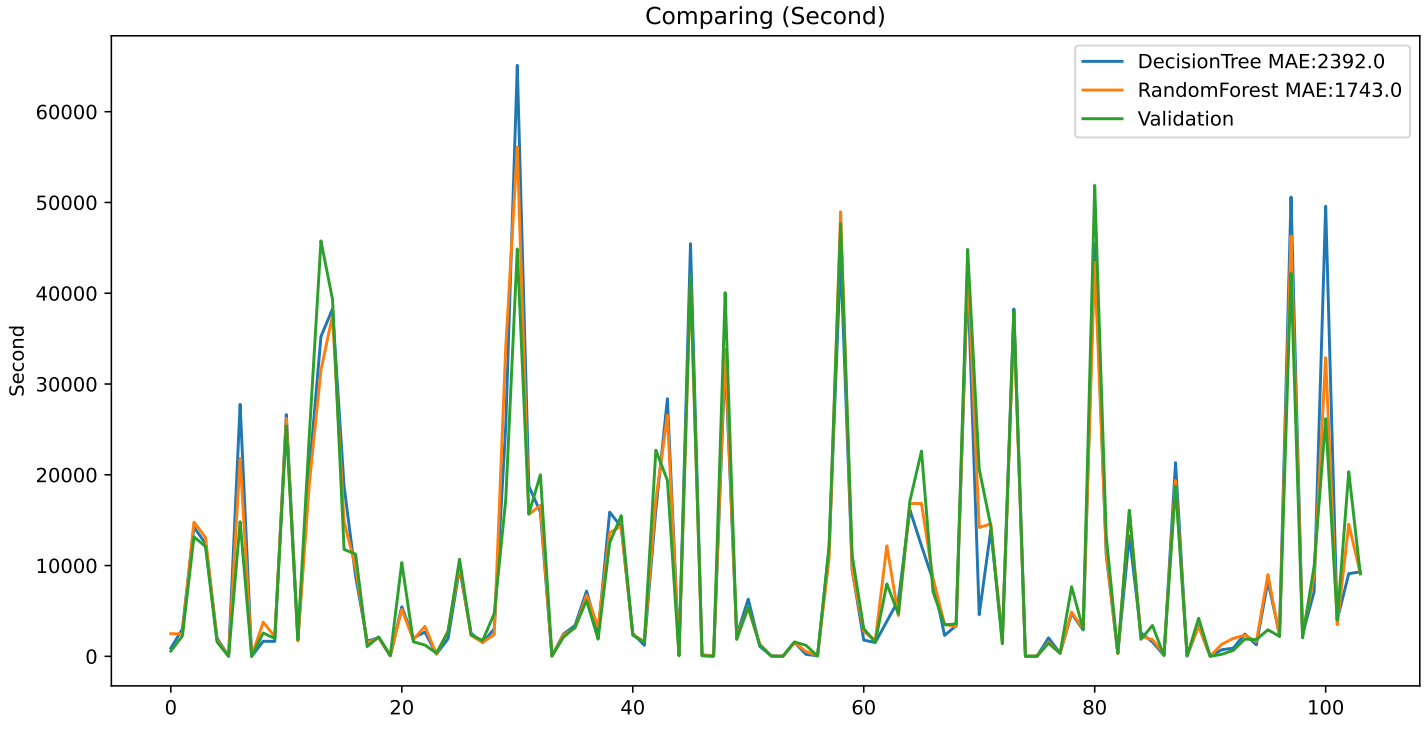
## DecisionTree:  0.7248630326024768

## RandomForest:  0.7837038702898657



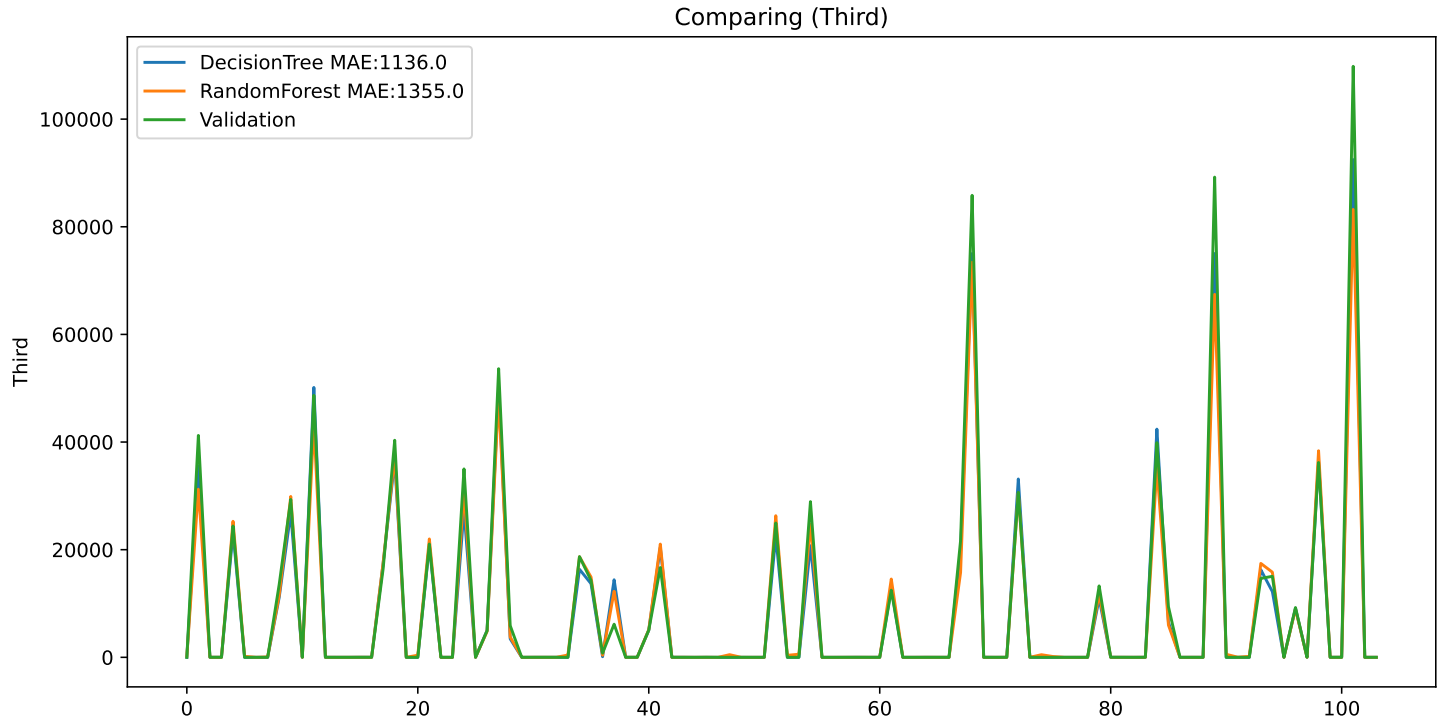
## DecisionTree:  0.7692636418171874

## RandomForest:  0.8318561653086721



## DecisionTree:  0.8664423499345815

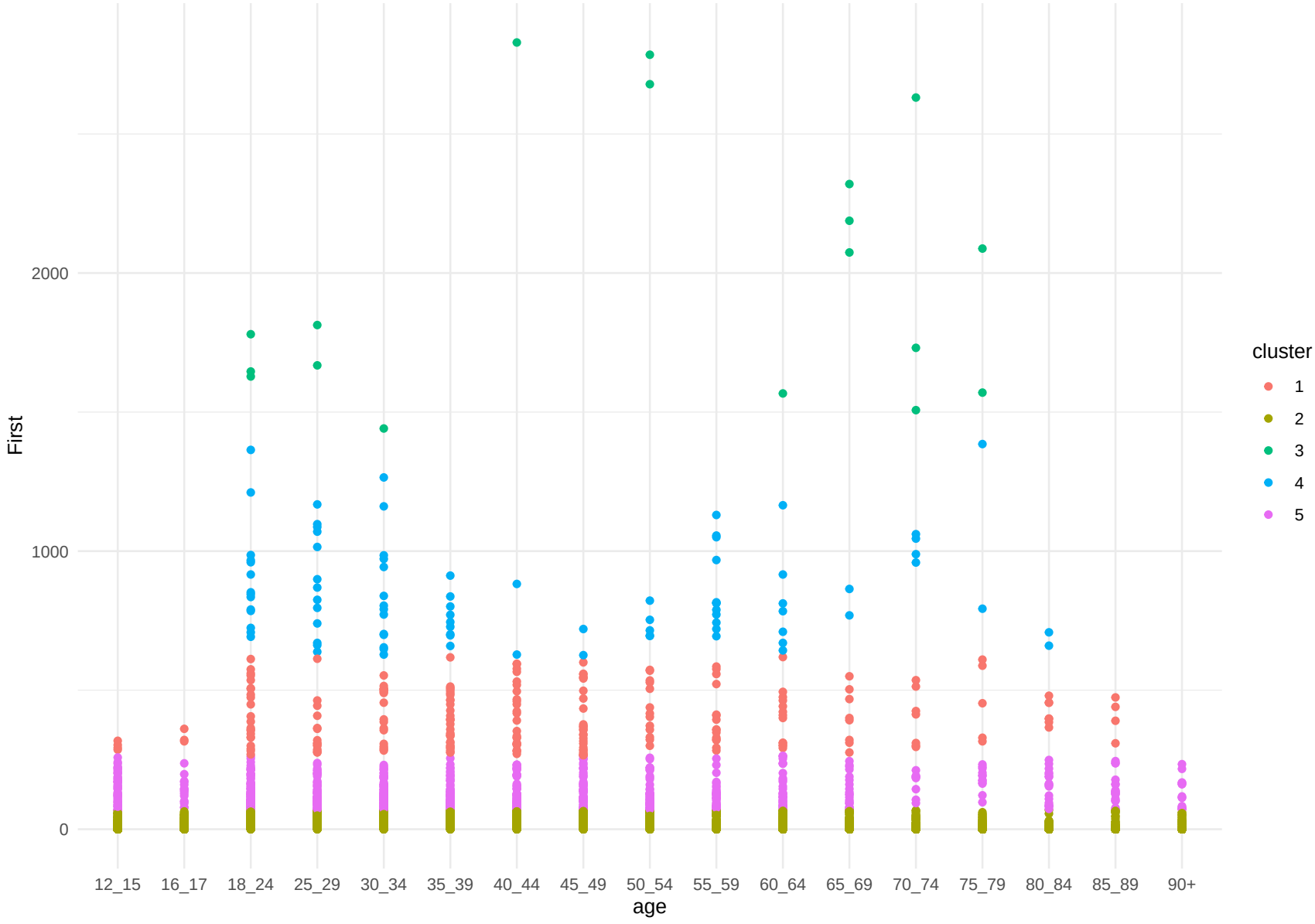
## RandomForest:  0.840745098066024

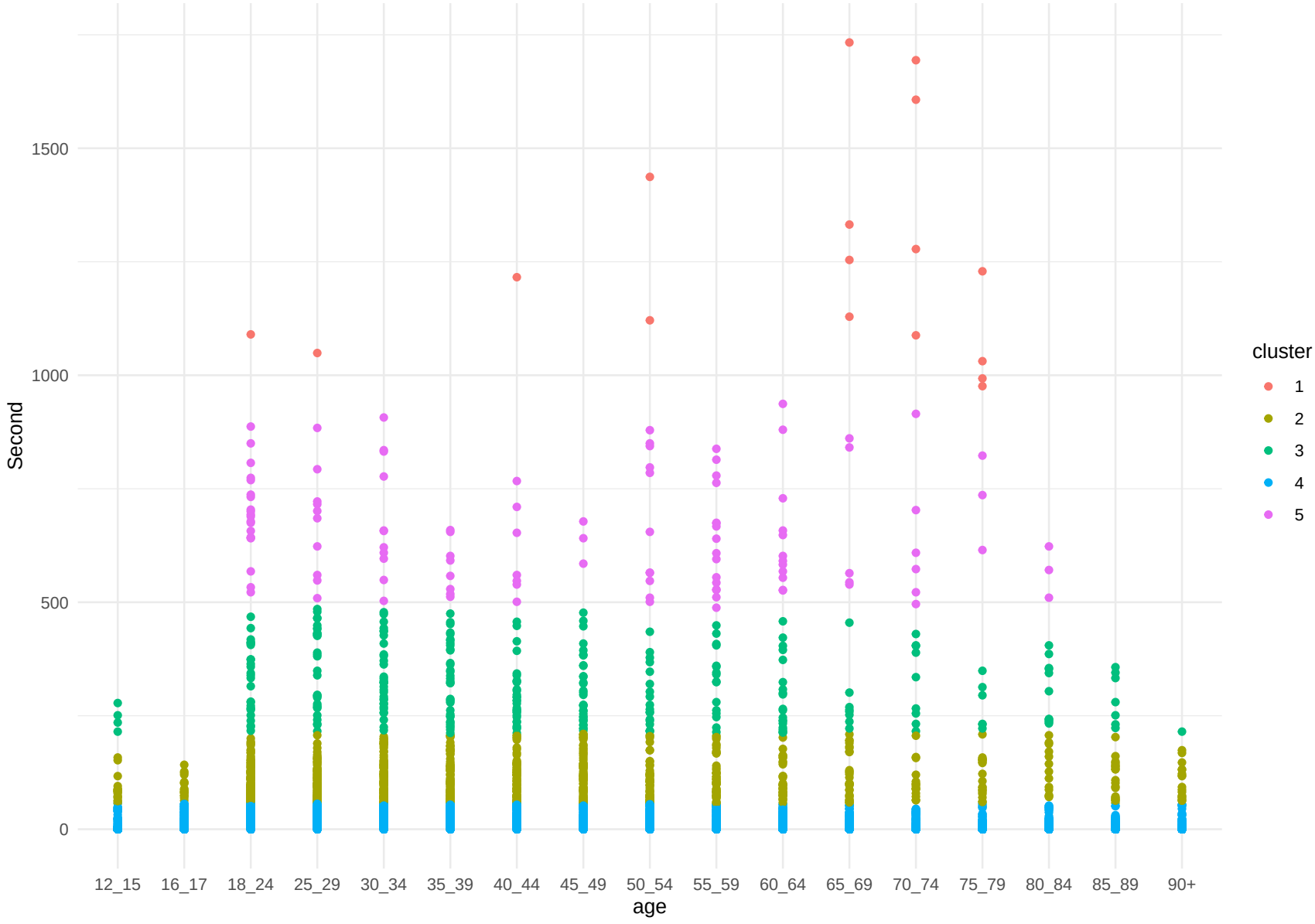


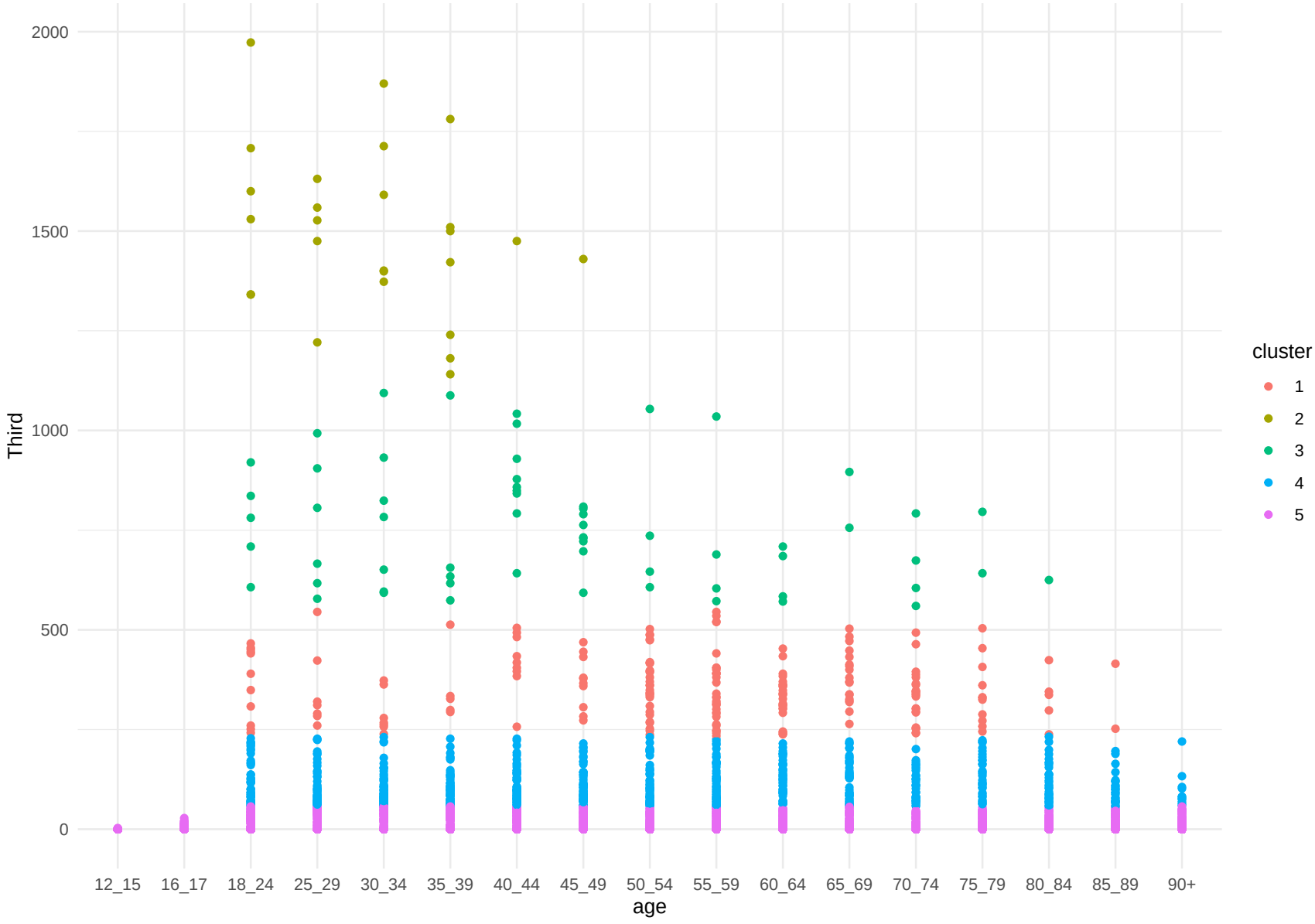
- The waves are recognisable.
- The extreme values are bigger than in real data.

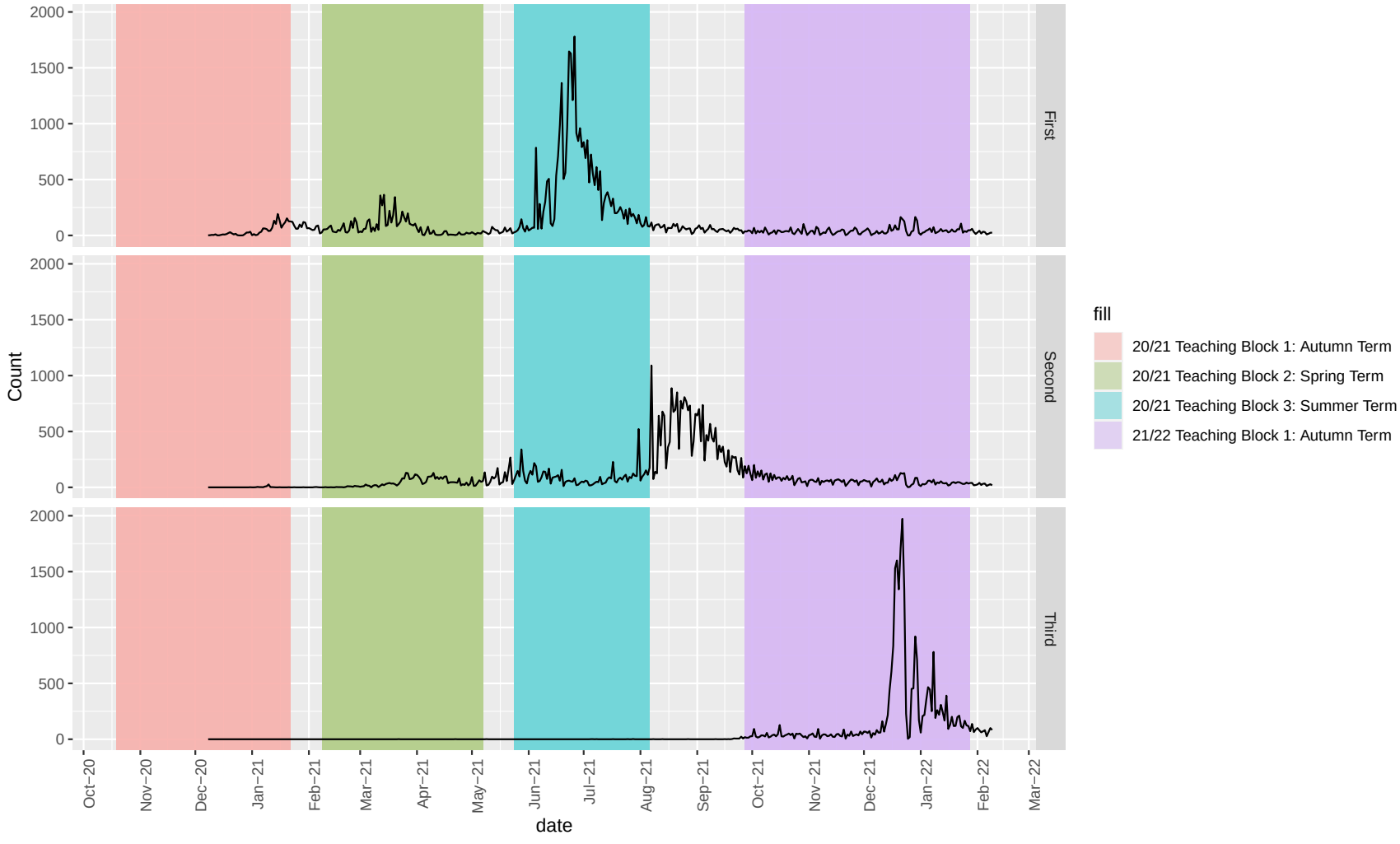
# Question 3

Are there dependencies between academic year events and vaccination waves?


First
2000
1000
0
12_15
16_17
18_24
25_29
30_34
35_39
40_44
45_49
50_54
55_59
60_64
65_69
70_74
75_79
80_84
85_89
90+
age
cluster
1
2
3
4
5


Second
1500
1000
500
0
12_15
16_17
18_24
25_29
30_34
35_39
40_44
45_49
50_54
55_59
60_64
65_69
70_74
75_79
80_84
85_89
90+
age
cluster
1
2
3
4
5


2000
1500
1000
500
0
Third
12_15
16_17
18_24
25_29
30_34
35_39
40_44
45_49
50_54
55_59
60_64
65_69
70_74
75_79
80_84
85_89
90+
age
cluster
1
2
3
4
5


Count
2000
1500
1000
500
0
First
Second
Third
Oct−20
Nov−20
Dec−20
Jan−21
Feb−21
Mar−21
Apr−21
May−21
Jun−21
Jul−21
Aug−21
Sep−21
Oct−21
Nov−21
Dec−21
Jan−22
Feb−22
Mar−22
date
fill
20/21 Teaching Block 1: Autumn Term
20/21 Teaching Block 2: Spring Term
20/21 Teaching Block 3: Summer Term
21/22 Teaching Block 1: Autumn Term

# What is next?

- the UK data
- England data
- deep learning and models considering time series.
- domain knowledge facts

# Thanks for your attention

To be continued. . . :)