Impacte del COVID-19 per estat de vacunació

Linear and Generalized Linear Models

Ulrik Danielsen, Alexander J Ohrt

29 desember, 2021

Contents

1 Introduction **2**

2 Objective **2**

3 Data Preparation **2**

4 Exploratory Data Analysis **3**
4.1 Critical Infections by Age . . . 3
4.2 Evolution of Critical Infections with Time . . . 3
4.3 Critical Infections by Vaccination Status and Age . . . 5
4.4 Cases by Vaccination Status per Age Group . . . 7
4.5 Critical Cases Relative to all Cases by Vaccination Status per Age Group . . . 7
4.6 Cases by Vaccination Status with Time (Months) . . . 8
4.7 Critical Cases by Vaccination Status with Time (Months) . . . 9

5 Model Fit **10**
5.1 Linear Model . . . 10
5.2 Generalized Linear Model . . . 11

6 Conclusions **23**

1 Introduction

We are analyzing a data set that contains information about the impact of COVID-19 by vaccination status in 2021 in Catalunya. The data comes from various information systems of the Department of Health and the Catalan Health Service, as detailed here. More specifically, the data shows the impact of COVID-19 (new cases, hospitalizations, critical hospitalizations and deaths) in the entirety of Catalunya, from January 1st 2021 until now, for the date of reference of each record, by gender, by age group, according to the vaccination status of the people. Keep in mind that a person is counted as many times as the number of events that has happened to the individual in question. This means that each person may be counted up to four times in the data set as a whole. More information about the data set can be found here. Note that the data we are using was downloaded December 17th at 12:00.

The data set contains 6 columns. They are explained shortly below.

Variable	SEXE	EDAT	DATA	ESDEVENIMENT	PAUTA	RECOMPTE
Explanation	Sex	Age Group	Date	Event (Severity of illness)	Vaccination status	Count
Factor Levels	Homme, Dona	0-9, 10-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80-,Not classified		Cas, Hospitalització, Critics, Defunció	No iniciada, Parcial, Completa	

2 Objective

The objective of the analysis is to answer the following question: *Do differences in the severity of COVID-19 infection exist between people that are unvaccinated, people that only have received one dose and people that are fully vaccinated?* More precisely, does the amount of people with a *critical* infection depend on their respective vaccination statuses?

The analysis will be performed separately for men and women. Moreover, we will only consider people that are older than 30 and only data from May 1st to December 12th. Additionally, we will only analyze differences in people that are reported with a critical event, disregarding people with new cases, people that are hospitalized (not critical) and people that have died.

3 Data Preparation

First we have to remove all people younger than 30 years old and cases before May 1st. The data is prepared according to the description above.

```
data <- read.csv("Impacte_del_COVID-19.csv") # Data updated 17.12.21.

clean <- function(data){
  data$DATA <- as.Date(data$DATA, format = "%d/%m/%Y")
  data$EDAT <- as.factor(data$EDAT)
  covid.data <- data %>% filter(DATA >= "2021-05-01" & DATA < "2021-12-13") %>%
        filter(EDAT %in% c("30 a 39", "40 a 49", "50 a 59", "60 a 69", "70 a 79", "80 o més"))
        # People with EDAT "No classificat" are not added
        # (these are the same people as "No classificat" in SEXE).

  covid.data$PAUTA <- factor(covid.data$PAUTA)
  covid.data$SEXE <- factor(covid.data$SEXE)

  # Create week and month variables.
  covid.data$Week <- as.factor(strftime(covid.data$DATA, format = "%W"))
  covid.data$Month <- as.factor(strftime(covid.data$DATA, format = "%m"))
```

```
  return(covid.data)
}
covid.data <- clean(data) %>% filter(ESDEVENIMENT == "Crítics")
```

Note also that there are 637 registries in the data set that are not classified into groups in `SEXE` and `EDAT`. These rows are disregarded in the analysis.

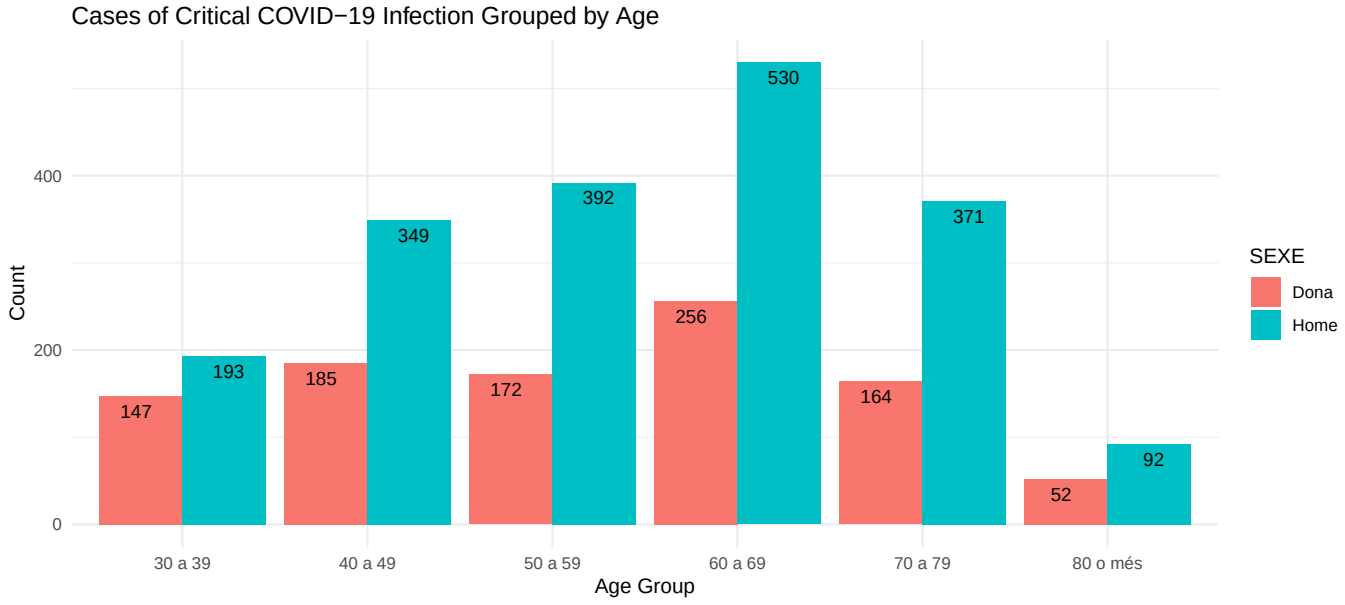
4 Exploratory Data Analysis

```
covid.data.women <- covid.data %>% filter(SEXE == "Dona")
covid.data.men <- covid.data %>% filter(SEXE == "Home")
```

The data set is split into men and women. There are 1927 critical cases for men and 976 critical cases for women in the data.

4.1 Critical Infections by Age

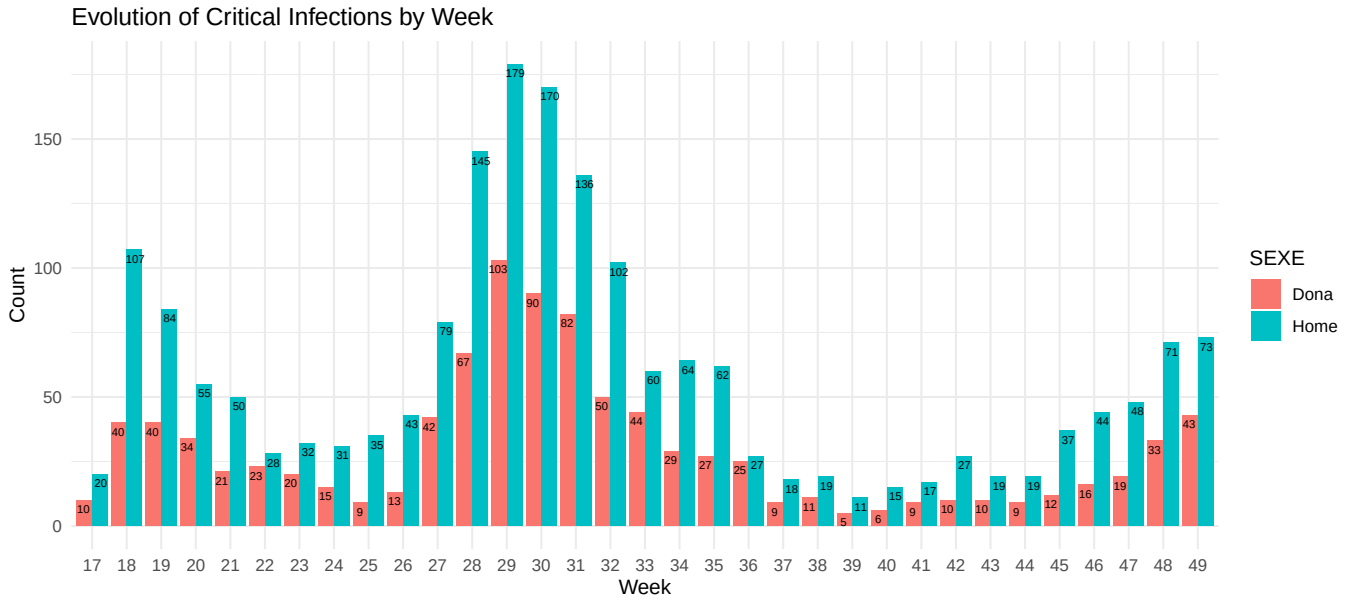
```
count.in.age.group <- function(data){
  data %>% group_by(SEXE, EDAT) %>% summarise(Count = sum(RECOMPTE), .groups = "keep")
}

all <- count.in.age.group(covid.data)
ggplot(all, aes(x = EDAT, y = Count)) +
  labs(x = "Age Group", y = "Count") +
  geom_col(aes(fill = SEXE), position = "dodge") +
  geom_text(aes(label=Count, group= SEXE), vjust = 1.6, color="black",
            size=3.5, position = position_dodge(width = 1)) +
  ggtitle("Cases of Critical COVID-19 Infection Grouped by Age") +
  theme_minimal()
```



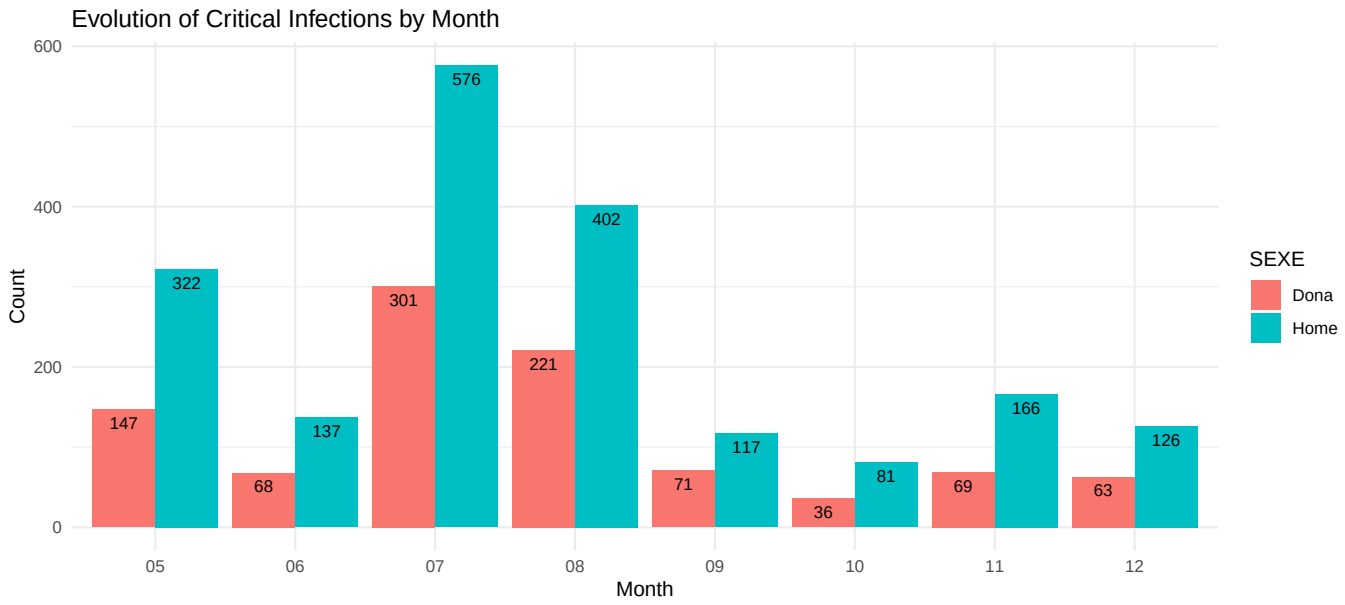
4.2 Evolution of Critical Infections with Time

```
count.each.week <- function(data){
  data %>% group_by(SEXE, Week) %>% summarise(Count = sum(RECOMPTE), .groups = "keep")
}
```

```
week.data <- count.each.week(covid.data)
ggplot(data = week.data, aes(x = as.factor(Week), y = Count)) +
  labs(x = "Week", y = "Count") +
  geom_col(aes(fill = SEXE), position = "dodge") +
  geom_text(aes(label=Count, group= SEXE), vjust = 1.6, color="black",
            size=2, position = position_dodge(width = 1)) +
  ggtitle("Evolution of Critical Infections by Week") +
  theme_minimal()
```

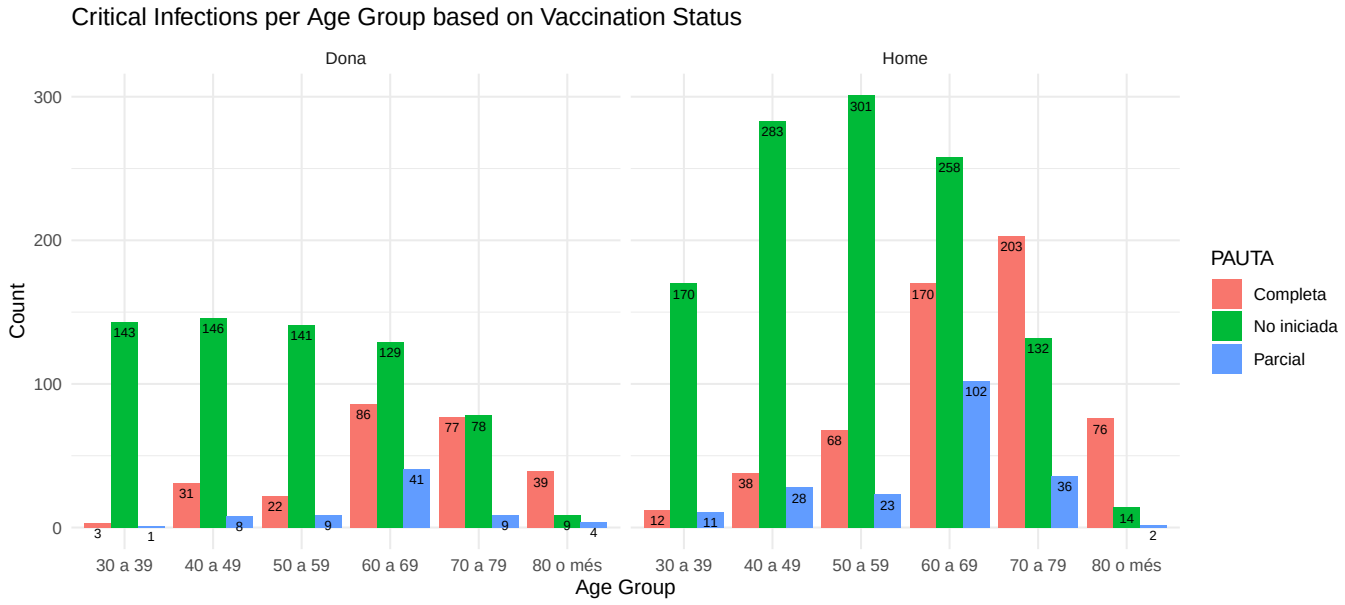



```
count.each.month <- function(data){
  data %>% group_by(SEXE, Month) %>% summarise(Count = sum(RECOMPTE), .groups = "keep")
}

month.data <- count.each.month(covid.data)
ggplot(data = month.data, aes(x = as.factor(Month), y = Count)) +
  labs(x = "Month", y = "Count") +
  geom_col(aes(fill = SEXE), position = "dodge") +
  geom_text(aes(label=Count, group= SEXE), vjust = 1.6, color="black",
            size=3, position = position_dodge(width = 0.9)) +
  ggtitle("Evolution of Critical Infections by Month") +
  theme_minimal()
```



4.3 Critical Infections by Vaccination Status and Age

```
group.by.vacc.age.gender <- function(data){
  data %>% group_by(SEXE, EDAT, PAUTA) %>% summarise(Count = sum(RECOMPTE), .groups = "keep")
}
d5 <- group.by.vacc.age.gender(covid.data)

ggplot(d5, aes(x = EDAT, y = Count)) +
  labs(x = "Age Group", y = "Count") +
  geom_col(aes(fill = PAUTA), position = "dodge") +
  geom_text(aes(label=Count, group= PAUTA), vjust = 1.6, color="black",
            size=2.5, position = position_dodge(width = 0.9)) +
  ggtitle("Critical Infections per Age Group based on Vaccination Status") +
  facet_wrap(~ SEXE) +
  theme_minimal()
```



```r
group.by.vacc.age <- function(data){
  data %>% group_by(EDAT, PAUTA) %>% summarise(Count = sum(RECOMPTE), .groups = "keep")
}
d4 <- group.by.vacc.age(covid.data.women)
knitr::kable(d4, caption = "Women with Critical Infection as
             Function of Vaccination Status and Age")
```

Table 2: Women with Critical Infection as Function of Vaccination Status and Age

EDAT	PAUTA	Count
30 a 39	Completa	3
30 a 39	No iniciada	143
30 a 39	Parcial	1
40 a 49	Completa	31
40 a 49	No iniciada	146
40 a 49	Parcial	8
50 a 59	Completa	22
50 a 59	No iniciada	141
50 a 59	Parcial	9
60 a 69	Completa	86
60 a 69	No iniciada	129
60 a 69	Parcial	41
70 a 79	Completa	77
70 a 79	No iniciada	78
70 a 79	Parcial	9
80 o més	Completa	39
80 o més	No iniciada	9
80 o més	Parcial	4

```r
d3 <- group.by.vacc.age(covid.data.men)
knitr::kable(d3, caption = "Men with Critical Infection as
             Function of Vaccination Status and Age")
```

Table 3: Men with Critical Infection as Function of Vaccination Status and Age

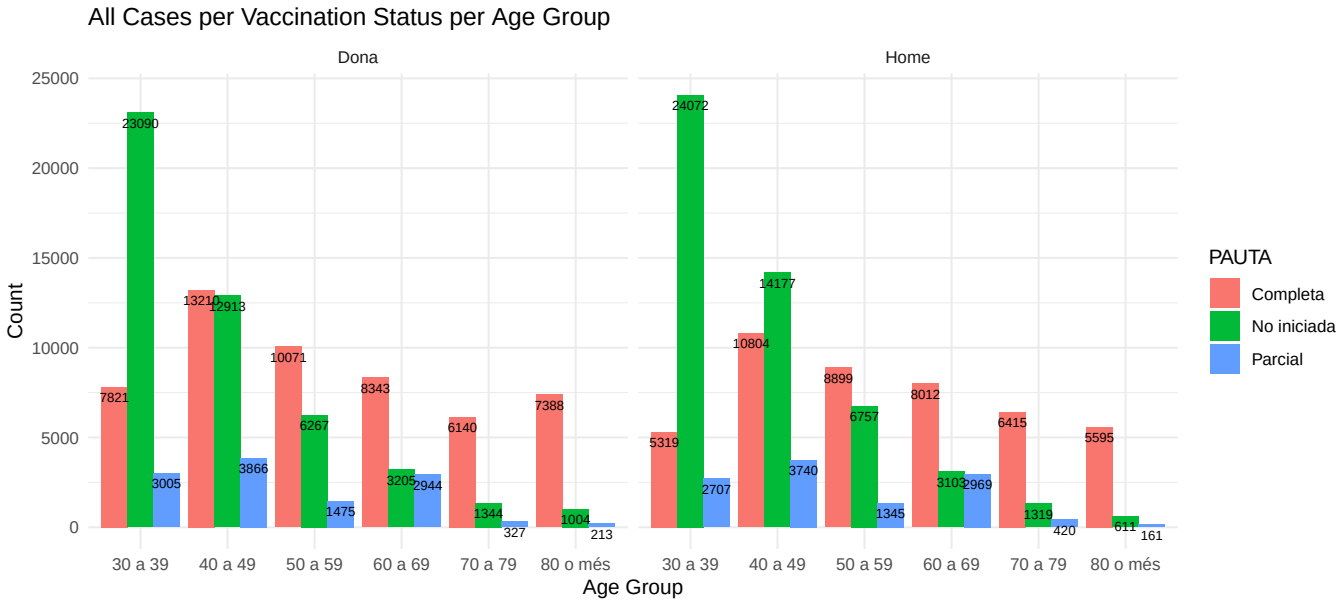
EDAT	PAUTA	Count
30 a 39	Completa	12
30 a 39	No iniciada	170
30 a 39	Parcial	11
40 a 49	Completa	38
40 a 49	No iniciada	283
40 a 49	Parcial	28
50 a 59	Completa	68
50 a 59	No iniciada	301
50 a 59	Parcial	23
60 a 69	Completa	170
60 a 69	No iniciada	258
60 a 69	Parcial	102
70 a 79	Completa	203
70 a 79	No iniciada	132
70 a 79	Parcial	36
80 o més	Completa	76

EDAT	PAUTA	Count
80 o més	No iniciada	14
80 o més	Parcial	2

4.4 Cases by Vaccination Status per Age Group

When considering the critical cases by vaccination status and age for each gender, it is good to keep all cases per vaccination status per age group in the entire data set in mind. This data is shown below, where all four different classifications of severity of infection are counted in the data set.

```
vacc.status.per.age <- group.by.vacc.age.gender(clean(data))
ggplot(vacc.status.per.age, aes(x = EDAT, y = Count)) +
  labs(x = "Age Group", y = "Count") +
  geom_col(aes(fill = PAUTA), position = "dodge") +
  geom_text(aes(label=Count, group= PAUTA), vjust = 1.6, color="black",
            size=2.5, position = position_dodge(width = 0.9)) +
  ggtitle("All Cases per Vaccination Status per Age Group") +
  facet_wrap(~ SEXE) +
  theme_minimal()
```

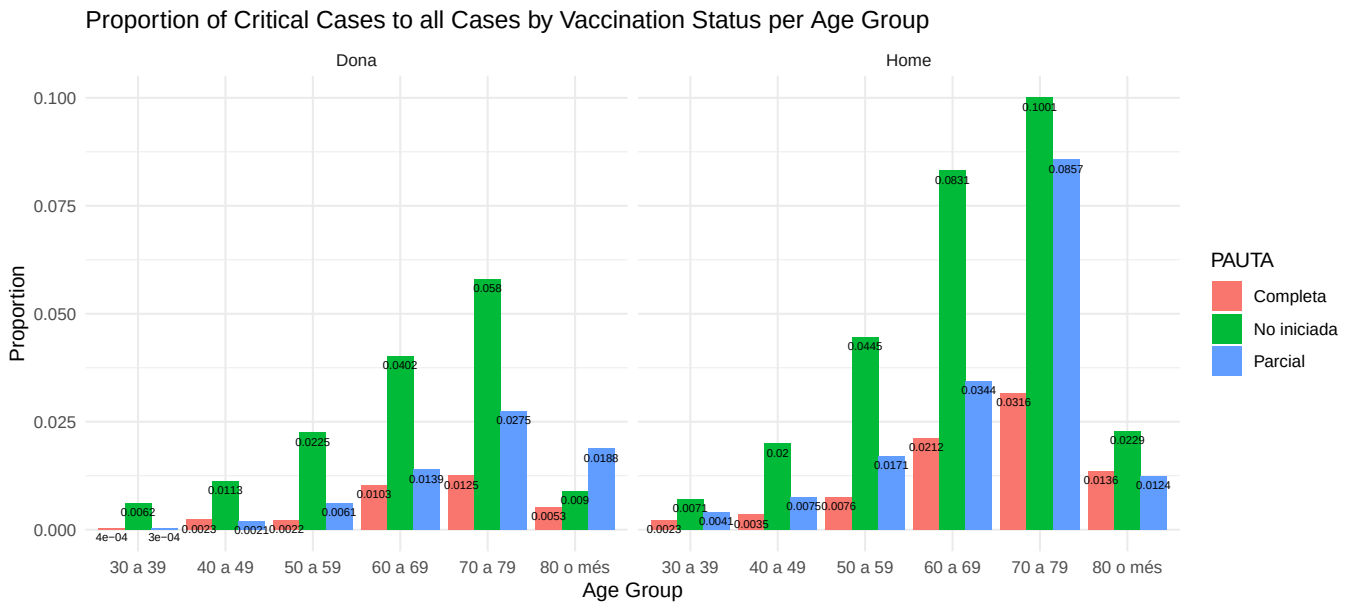


4.5 Critical Cases Relative to all Cases by Vaccination Status per Age Group

The next plot shows the critical cases of COVID-19 relative to all cases by vaccination status in the entire data set per age group. Essentially, this shows the proportion of the counts of the two earlier plots. Thus, for example, even though plot 4.3 shows that there are more critical cases for completely vaccinated individuals in the ages greater than 70 (except for women aged 70 to 79), plot 4.5 shows that the proportion of cases is still greatest for the non-vaccinated (except for women aged 80 or more). This is because most of the population in these ages are in fact completely vaccinated.

```
d6 <- d5
d6$Count <- d5$Count/vacc.status.per.age$Count
ggplot(d6, aes(x = EDAT, y = Count)) +
  labs(x = "Age Group", y = "Proportion") +
  geom_col(aes(fill = PAUTA), position = "dodge") +
  geom_text(aes(label=round(Count,4), group= PAUTA), vjust = 1.6, color="black",
            size=2, position = position_dodge(width = 0.9)) +
  ggtitle("Proportion of Critical Cases to all Cases by Vaccination Status per Age Group") +
```

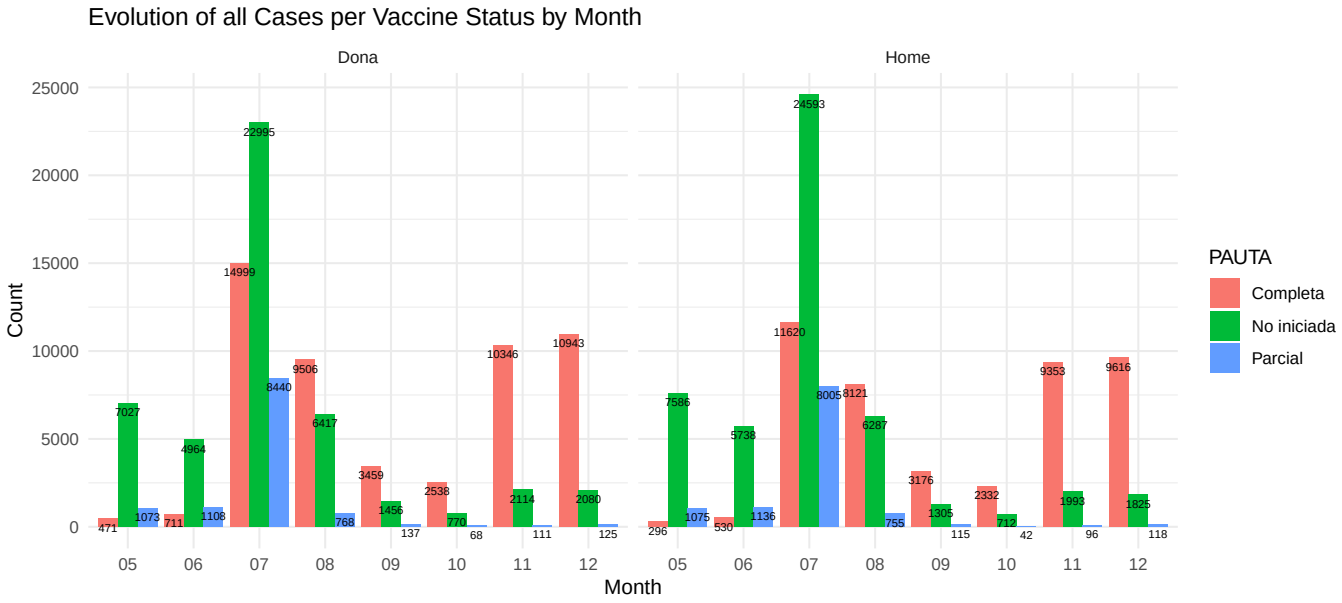
```
  facet_wrap(~ SEXE) +
  theme_minimal()
```



4.6 Cases by Vaccination Status with Time (Months)

All cases per vaccination status in the entire data set is shown by month in the plot below. It is apparent that the later in the year, the more cases are registered for the fully vaccinated. This is most likely a result of the fact that more and more people are getting vaccinated. i.e. the proportion of vaccinated people in the population is growing throughout the year.

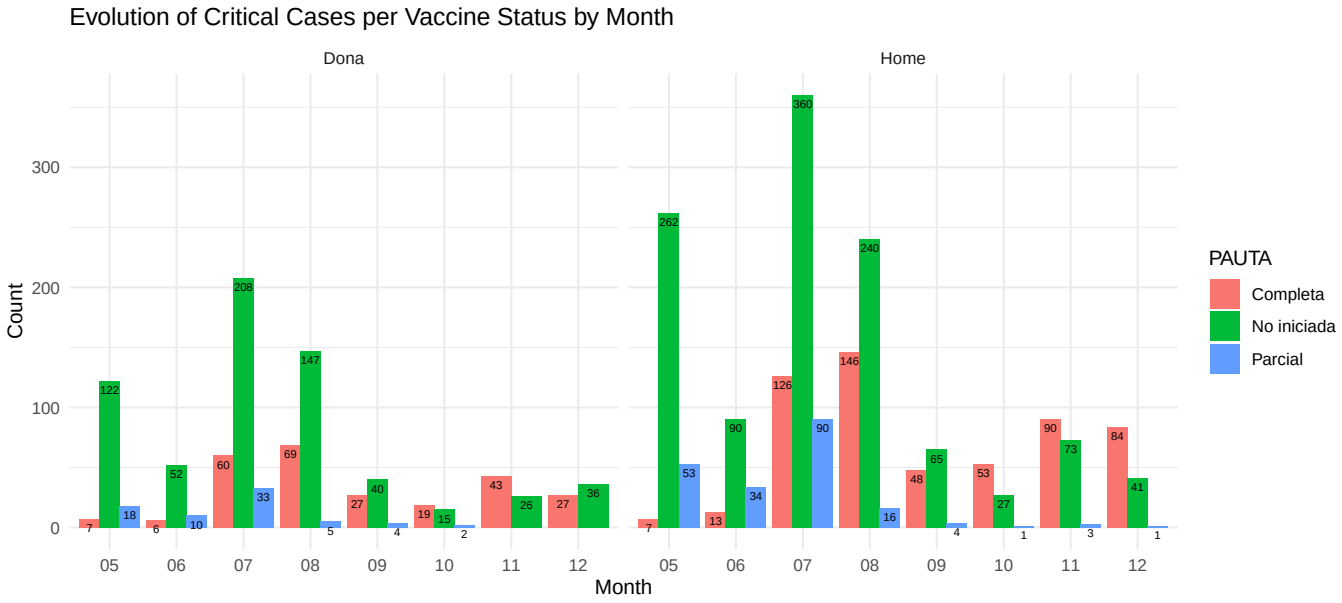
```
vacc.status.with.time <- function(data){
  data %>% group_by(SEXE, PAUTA, Month) %>% summarise(Count = sum(RECOMPTE), .groups = "keep")
}
vacc.with.time <- vacc.status.with.time(clean(data))
ggplot(data = vacc.with.time, aes(x = as.factor(Month), y = Count)) +
  labs(x = "Month", y = "Count") +
  geom_col(aes(fill = PAUTA), position = "dodge") +
  geom_text(aes(label=Count, group= PAUTA), vjust = 1.6, color="black",
            size=2, position = position_dodge(width = 0.9)) +
  ggtitle("Evolution of all Cases per Vaccine Status by Month") +
  facet_wrap(~ SEXE) +
  theme_minimal()
```



4.7 Critical Cases by Vaccination Status with Time (Months)

Only the critical cases per vaccination status is shown by month below.

```
vacc.status.with.time <- function(data){
  data %>% group_by(SEXE, PAUTA, Month) %>% summarise(Count = sum(RECOMPTE), .groups = "keep")
}
vacc.with.time.critical <- vacc.status.with.time(covid.data)
ggplot(data = vacc.with.time.critical, aes(x = as.factor(Month), y = Count)) +
  labs(x = "Month", y = "Count") +
  geom_col(aes(fill = PAUTA), position = "dodge") +
  geom_text(aes(label=Count, group= PAUTA), vjust = 1.6, color="black",
            size=2, position = position_dodge(width = 0.9)) +
  ggtitle("Evolution of Critical Cases per Vaccine Status by Month") +
  facet_wrap(~ SEXE) +
  theme_minimal()
```



Already, the exploratory data analysis hints at the fact that the amount of critical infections does depend on the vaccination status. This should be proven statistically however, an issue that is discussed in the following.

5 Model Fit

We wish to examine how the vaccination status affects the number of critical patients in this period. Since we observe that the number of critical patients vary very much from month to month, we choose to aggregate the number of patients each month, and include it as a categorical covariate in our model.

```
covid.data.aggr <- aggregate(RECOMPTE ~ EDAT + SEXE + PAUTA + Month,
                             covid.data, FUN = sum)
head(covid.data.aggr)
```

```
##       EDAT SEXE    PAUTA Month RECOMPTE
## 1  30 a 39 Dona Completa    05        1
## 2  50 a 59 Dona Completa    05        1
## 3  60 a 69 Dona Completa    05        1
## 4  70 a 79 Dona Completa    05        1
## 5 80 o més Dona Completa    05        3
## 6  30 a 39 Home Completa    05        1
```

```
dim(covid.data.aggr)
```

```
## [1] 221   5
```

5.1 Linear Model

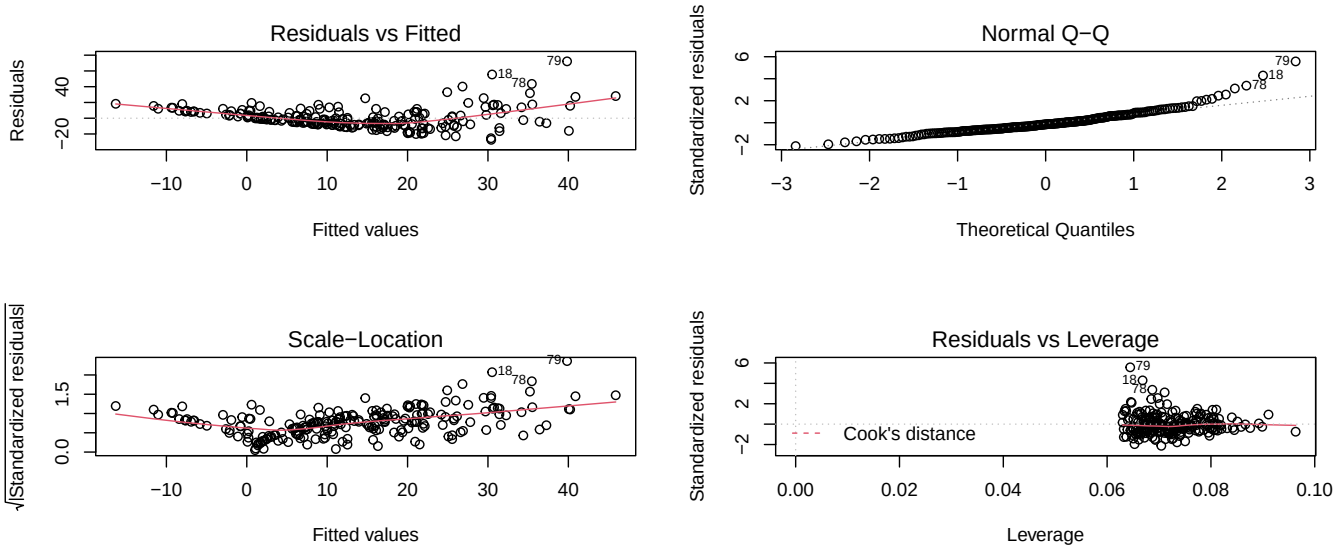
First we try a simple linear model with all covariates included. As the adjusted R-squared is low, and the residuals do not fit the normal assumption, we reject this model.

```
model1 <- lm(RECOMPTE ~ EDAT + SEXE + PAUTA + Month, data = covid.data.aggr)
summary(model1)
```

```
##
## Call:
## lm(formula = RECOMPTE ~ EDAT + SEXE + PAUTA + Month, data = covid.data.aggr)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -27.405  -7.925  -1.933   6.400  72.158
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)          6.996      3.835   1.824 0.069541 .
## EDAT40 a 49          4.384      3.410   1.286 0.200031
## EDAT50 a 59          5.431      3.379   1.607 0.109507
## EDAT60 a 69         10.434      3.322   3.140 0.001937 **
## EDAT70 a 79          4.608      3.332   1.383 0.168151
## EDAT80 o més        -5.053      3.529  -1.432 0.153718
## SEXEHome             8.624      1.811   4.762 3.63e-06 ***
## PAUTANo iniciada     9.465      2.044   4.631 6.45e-06 ***
## PAUTAParcial        -8.874      2.535  -3.501 0.000568 ***
## Month06             -9.372      3.463  -2.706 0.007376 **
## Month07             10.373      3.358   3.089 0.002288 **
## Month08              4.706      3.534   1.331 0.184526
## Month09            -11.145      3.567  -3.124 0.002042 **
## Month10            -13.528      3.656  -3.700 0.000277 ***
## Month11             -9.971      3.714  -2.685 0.007850 **
## Month12            -11.317      3.768  -3.003 0.003003 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 13.39 on 205 degrees of freedom
## Multiple R-squared:  0.4554, Adjusted R-squared:  0.4155
## F-statistic: 11.43 on 15 and 205 DF,  p-value: < 2.2e-16
```

```
par(mfrow = c(2,2))
plot(model1, ask = F)
```



Since we have aggregated the number of critical patients by month, we can interpret this data collection as a counting process which starts each month.

5.2 Generalized Linear Model

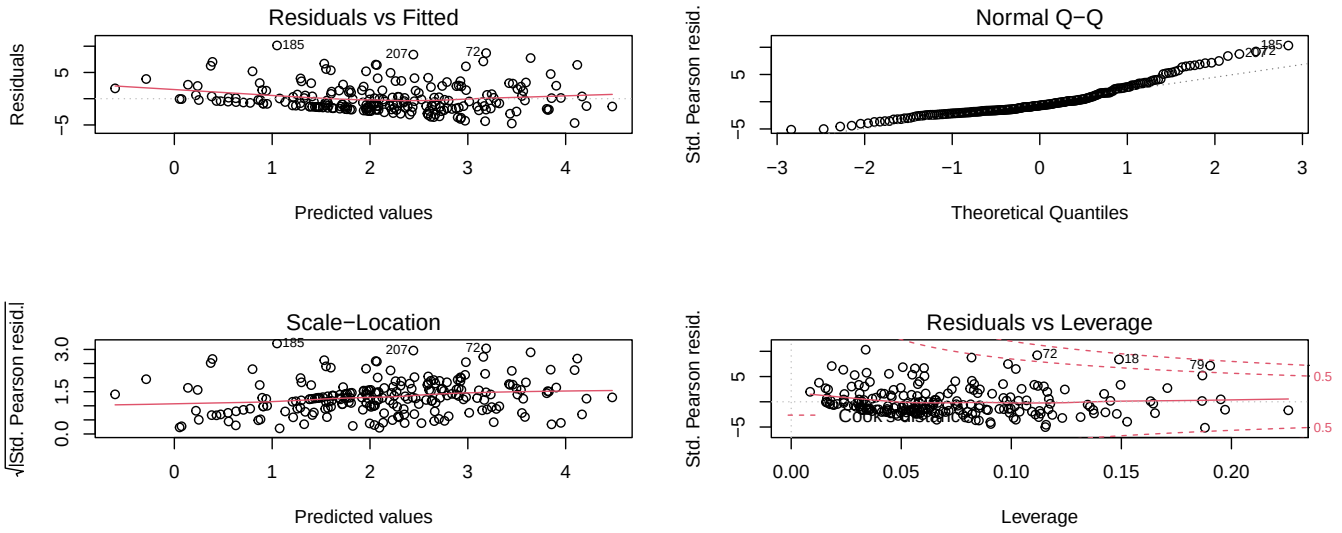
The described counting process is a poisson process, which we can fit as a generalized linear model from the poisson family. The model is fitted without interaction, and includes all covariates. The residuals still look bad, and the model has a dispersion parameter above seven.

```
model2 <- glm(RECOMPTE ~ EDAT + SEXE + PAUTA + Month, data = covid.data.aggr,
              family = poisson(link = "log"))
summary(model2)
```

```
##
## Call:
## glm(formula = RECOMPTE ~ EDAT + SEXE + PAUTA + Month, family = poisson(link = "log"),
##     data = covid.data.aggr)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -5.8785  -1.9662  -0.7042   1.2364   7.1265
##
## Coefficients:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)        1.99219    0.08075  24.670  < 2e-16 ***
## EDAT40 a 49        0.27202    0.06958   3.909 9.25e-05 ***
## EDAT50 a 59        0.32011    0.06884   4.650 3.32e-06 ***
## EDAT60 a 69        0.62929    0.06523   9.647  < 2e-16 ***
## EDAT70 a 79        0.24453    0.06954   3.516 0.000438 ***
## EDAT80 o més      -0.91594    0.09961  -9.195  < 2e-16 ***
## SEXEHome           0.65979    0.03933  16.774  < 2e-16 ***
## PAUTANo iniciada   0.66973    0.04215  15.889  < 2e-16 ***
## PAUTAParcial      -0.82644    0.07029 -11.757  < 2e-16 ***
```

```
## Month06          -0.85545    0.08379 -10.209  < 2e-16 ***
## Month07           0.52527    0.05729   9.169  < 2e-16 ***
## Month08           0.26199    0.06126   4.276 1.90e-05 ***
## Month09          -0.93686    0.08640 -10.843  < 2e-16 ***
## Month10          -1.36409    0.10349 -13.181  < 2e-16 ***
## Month11          -0.68628    0.08013  -8.565  < 2e-16 ***
## Month12          -0.83792    0.08650  -9.686  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 3746.1  on 220  degrees of freedom
## Residual deviance: 1400.5  on 205  degrees of freedom
## AIC: 2248.7
##
## Number of Fisher Scoring iterations: 5
```

```
par(mfrow = c(2,2))
plot(model2, ask = F)
```




```
# Checking for overdispersion.
PS <- sum(residuals(model2, type="pearson")^2)
PS
```

```
## [1] 1575.38
```

```
phi <- PS / model2$df.res
phi
```

```
## [1] 7.684779
```

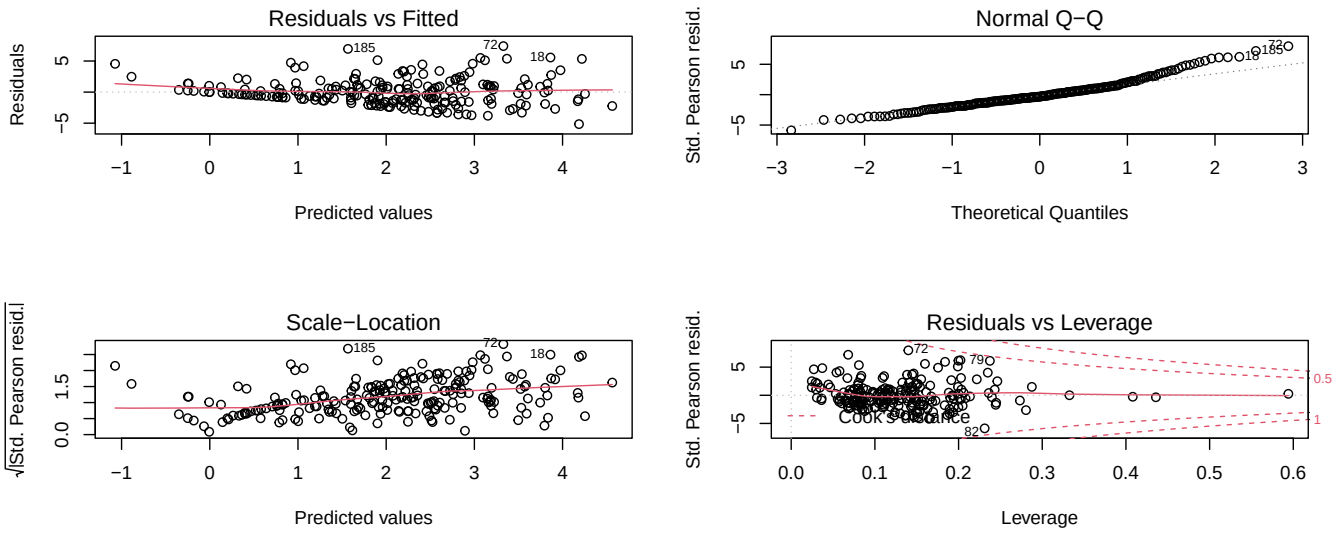
Next we include an interaction between the vaccination status and month. This makes sense because the vaccination status should become more skewed towards complete as time goes and more people are vaccinated.

```
model3 <- glm(RECOMPTE ~ EDAT + SEXE + PAUTA + Month + PAUTA:Month, data = covid.data.aggr,
              family = poisson(link = "log"))
summary(model3)
```

```
##
## Call:
```

```
## glm(formula = RECOMPTE ~ EDAT + SEXE + PAUTA + Month + PAUTA:Month,
##     family = poisson(link = "log"), data = covid.data.aggr)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -5.9298  -1.4289  -0.2761   0.9461   6.2597
##
## Coefficients:
##                           Estimate Std. Error z value Pr(>|z|)
## (Intercept)               -0.06542    0.27267  -0.240 0.810389
## EDAT40 a 49                0.24289    0.06986   3.477 0.000508 ***
## EDAT50 a 59                0.27829    0.06928   4.017 5.89e-05 ***
## EDAT60 a 69                0.58723    0.06559   8.952  < 2e-16 ***
## EDAT70 a 79                0.21008    0.07001   3.001 0.002693 **
## EDAT80 o més              -1.00809    0.10001 -10.080  < 2e-16 ***
## SEXEHome                   0.64748    0.03937  16.444  < 2e-16 ***
## PAUTANo iniciada           3.00268    0.27219  11.031  < 2e-16 ***
## PAUTAParcial               1.61943    0.29278   5.531 3.18e-08 ***
## Month06                    0.18656    0.35242   0.529 0.596550
## Month07                    2.32816    0.27729   8.396  < 2e-16 ***
## Month08                    2.57699    0.27615   9.332  < 2e-16 ***
## Month09                    1.49058    0.29135   5.116 3.12e-07 ***
## Month10                    1.39339    0.29221   4.769 1.86e-06 ***
## Month11                    1.99276    0.28113   7.088 1.36e-12 ***
## Month12                    1.81194    0.28377   6.385 1.71e-10 ***
## PAUTANo iniciada:Month06 -1.16329    0.36587  -3.180 0.001475 **
## PAUTAParcial:Month06     -0.73956    0.40127  -1.843 0.065322 .
## PAUTANo iniciada:Month07 -1.93668    0.28505  -6.794 1.09e-11 ***
## PAUTAParcial:Month07     -2.04851    0.31517  -6.500 8.04e-11 ***
## PAUTANo iniciada:Month08 -2.56920    0.28539  -9.003  < 2e-16 ***
## PAUTAParcial:Month08     -3.58618    0.37203  -9.639  < 2e-16 ***
## PAUTANo iniciada:Month09 -2.76918    0.31148  -8.890  < 2e-16 ***
## PAUTAParcial:Month09     -3.50456    0.47335  -7.404 1.32e-13 ***
## PAUTANo iniciada:Month10 -3.57151    0.33440 -10.680  < 2e-16 ***
## PAUTAParcial:Month10     -3.54097    0.65818  -5.380 7.45e-08 ***
## PAUTANo iniciada:Month11 -3.33020    0.30290 -10.995  < 2e-16 ***
## PAUTAParcial:Month11     -4.20512    0.65332  -6.437 1.22e-10 ***
## PAUTANo iniciada:Month12 -3.32022    0.30999 -10.711  < 2e-16 ***
## PAUTAParcial:Month12     -4.60066    1.04655  -4.396 1.10e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 3746.1  on 220  degrees of freedom
## Residual deviance:  954.5  on 191  degrees of freedom
## AIC: 1830.6
##
## Number of Fisher Scoring iterations: 5
```

```
par(mfrow = c(2,2))
plot(model3, ask = F)
```




```
# Checking for overdispersion.
PS <- sum(residuals(model3, type="pearson")^2)
PS
```

```
## [1] 969.4384
```

```
phi <- PS / model3$df.res
phi
```

```
## [1] 5.075594
```

The model is improved, but the residuals still show clear patterns, and the model is overdispersed.

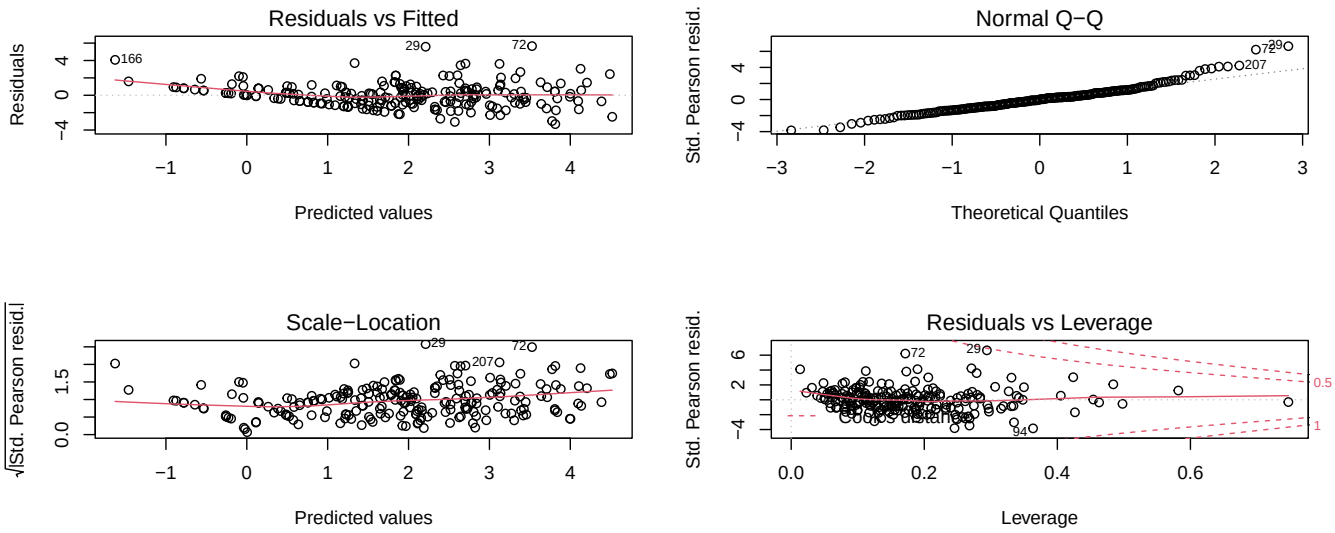
Now we include an interaction between vaccination status and age group. This makes sense since the older age group likely was vaccinated earlier than the younger age group.

```
model4 <- glm(RECOMPTE ~ EDAT + SEXE + PAUTA + Month + PAUTA:Month + EDAT:PAUTA, data = covid.data.aggr,
              family = poisson(link = "log"))
summary(model4)
```

```
##
## Call:
## glm(formula = RECOMPTE ~ EDAT + SEXE + PAUTA + Month + PAUTA:Month +
##     EDAT:PAUTA, family = poisson(link = "log"), data = covid.data.aggr)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.8305  -0.9572   0.0000   0.6922   4.9875
##
## Coefficients:
##                                Estimate Std. Error z value Pr(>|z|)
## (Intercept)                   -1.461571   0.366753  -3.985 6.74e-05 ***
## EDAT40 a 49                    0.925016   0.286089   3.233 0.001224 **
## EDAT50 a 59                    1.201044   0.280038   4.289 1.80e-05 ***
## EDAT60 a 69                    2.214167   0.266924   8.295  < 2e-16 ***
## EDAT70 a 79                    2.287555   0.266153   8.595  < 2e-16 ***
## EDAT80 o més                   1.413922   0.275749   5.128 2.94e-07 ***
## SEXEHome                       0.679437   0.039378  17.254  < 2e-16 ***
## PAUTANo iniciada               4.596782   0.372534  12.339  < 2e-16 ***
## PAUTAParcial                   2.350647   0.489088   4.806 1.54e-06 ***
## Month06                        0.005237   0.352897   0.015 0.988160
```

```
## Month07                         2.209180   0.277733   7.954 1.80e-15 ***
## Month08                         2.376726   0.276635   8.592  < 2e-16 ***
## Month09                         1.339389   0.291830   4.590 4.44e-06 ***
## Month10                         1.277531   0.292624   4.366 1.27e-05 ***
## Month11                         1.873782   0.281564   6.655 2.83e-11 ***
## Month12                         1.692963   0.284198   5.957 2.57e-09 ***
## PAUTANo iniciada:Month06       -0.995412   0.366310  -2.717 0.006580 **
## PAUTAParcial:Month06           -0.547188   0.401796  -1.362 0.173244
## PAUTANo iniciada:Month07       -1.817701   0.285483  -6.367 1.93e-10 ***
## PAUTAParcial:Month07           -1.848283   0.317081  -5.829 5.57e-09 ***
## PAUTANo iniciada:Month08       -2.368943   0.285858  -8.287  < 2e-16 ***
## PAUTAParcial:Month08           -3.460160   0.375512  -9.215  < 2e-16 ***
## PAUTANo iniciada:Month09       -2.631431   0.311917  -8.436  < 2e-16 ***
## PAUTAParcial:Month09           -3.408012   0.474021  -7.190 6.50e-13 ***
## PAUTANo iniciada:Month10       -3.481329   0.334735 -10.400  < 2e-16 ***
## PAUTAParcial:Month10           -3.484864   0.659551  -5.284 1.27e-07 ***
## PAUTANo iniciada:Month11       -3.224665   0.303287 -10.632  < 2e-16 ***
## PAUTAParcial:Month11           -4.344508   0.654287  -6.640 3.14e-11 ***
## PAUTANo iniciada:Month12       -3.174093   0.310531 -10.221  < 2e-16 ***
## PAUTAParcial:Month12           -4.965018   1.047933  -4.738 2.16e-06 ***
## EDAT40 a 49:PAUTANo iniciada   -0.642183   0.295628  -2.172 0.029836 *
## EDAT50 a 59:PAUTANo iniciada   -0.888359   0.289659  -3.067 0.002163 **
## EDAT60 a 69:PAUTANo iniciada   -2.034367   0.277582  -7.329 2.32e-13 ***
## EDAT70 a 79:PAUTANo iniciada   -2.719072   0.280746  -9.685  < 2e-16 ***
## EDAT80 o més:PAUTANo iniciada -3.973817   0.350421 -11.340  < 2e-16 ***
## EDAT40 a 49:PAUTAParcial       -0.249233   0.441089  -0.565 0.572046
## EDAT50 a 59:PAUTAParcial       -0.894509   0.444376  -2.013 0.044119 *
## EDAT60 a 69:PAUTAParcial       -0.510626   0.407023  -1.255 0.209647
## EDAT70 a 79:PAUTAParcial       -1.644763   0.426127  -3.860 0.000113 ***
## EDAT80 o més:PAUTAParcial      -2.325473   0.579507  -4.013 6.00e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 3746.09  on 220  degrees of freedom
## Residual deviance:  398.57  on 181  degrees of freedom
## AIC: 1294.7
##
## Number of Fisher Scoring iterations: 5
```

```
par(mfrow = c(2,2))
plot(model4, ask = F)
```




```
# Checking for overdispersion
PS <- sum(residuals(model4, type="pearson")^2)
PS
```

```
## [1] 415.0215
```

```
phi <- PS / model4$df.res
phi
```

```
## [1] 2.292937
```

```
Anova(model4)
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: RECOMPTE
##              LR Chisq Df Pr(>Chisq)
## EDAT          445.54  5  < 2.2e-16 ***
## SEXE          314.91  1  < 2.2e-16 ***
## PAUTA         778.80  2  < 2.2e-16 ***
## Month        1154.92  7  < 2.2e-16 ***
## PAUTA:Month   435.27 14  < 2.2e-16 ***
## EDAT:PAUTA    555.94 10  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Now the model shows clear improvement, and the dispersion parameter is reduced to right above two.

Including an interaction between all three covariates provides us with an even better model, as shown below. The graphical plots are much better, and the dispersion parameter 1.36, which is still slightly overdispersed, but better. The Pearson residual for the model is 130.4, which is slightly below the critical value for the χ^2_{n-p} distribution. Thus, we can conclude that this model is a good fit to the data.

```
model5 <- glm(RECOMPTE ~ SEXE + EDAT*PAUTA*Month,
              data = covid.data.aggr,
              family = poisson(link = "log"))
summary(model5)
```

```
##
## Call:
## glm(formula = RECOMPTE ~ SEXE + EDAT * PAUTA * Month, family = poisson(link = "log"),
##     data = covid.data.aggr)
```

```
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.0144  -0.4083   0.0000   0.3728   2.4252
##
## Coefficients: (20 not defined because of singularities)
##                                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)                         -3.974e-01  7.076e-01  -0.562 0.574346
## SEXEHome                             6.811e-01  3.965e-02  17.177  < 2e-16
## EDAT40 a 49                          4.095e-01  1.000e+00   0.409 0.682194
## EDAT50 a 59                         -5.806e-15  1.000e+00   0.000 1.000000
## EDAT60 a 69                          3.974e-01  1.225e+00   0.324 0.745616
## EDAT70 a 79                          6.931e-01  8.660e-01   0.800 0.423492
## EDAT80 o més                         1.496e+00  9.133e-01   1.638 0.101390
## PAUTANo iniciada                     2.862e+00  7.270e-01   3.937 8.26e-05
## PAUTAParcial                         3.927e+00  1.778e+00   2.209 0.027174
## Month06                             -2.836e-01  1.225e+00  -0.232 0.816866
## Month07                              8.150e-01  9.130e-01   0.893 0.372038
## Month08                              8.805e-01  1.014e+00   0.869 0.385038
## Month09                              8.150e-01  9.130e-01   0.893 0.372038
## Month10                              6.931e-01  8.660e-01   0.800 0.423492
## Month11                             -2.836e-01  1.225e+00  -0.232 0.816866
## Month12                             -2.836e-01  1.225e+00  -0.232 0.816866
## EDAT40 a 49:PAUTANo iniciada         4.660e-01  1.020e+00   0.457 0.647837
## EDAT50 a 59:PAUTANo iniciada         1.172e+00  1.019e+00   1.151 0.249853
## EDAT60 a 69:PAUTANo iniciada         3.780e-01  1.242e+00   0.304 0.760880
## EDAT70 a 79:PAUTANo iniciada         4.196e-02  8.901e-01   0.047 0.962397
## EDAT80 o més:PAUTANo iniciada       -3.953e+00  1.094e+00  -3.615 0.000301
## EDAT40 a 49:PAUTAParcial            -3.939e+00  2.159e+00  -1.824 0.068136
## EDAT50 a 59:PAUTAParcial            -3.927e+00  2.040e+00  -1.925 0.054198
## EDAT60 a 69:PAUTAParcial            -1.462e+00  1.460e+00  -1.001 0.316646
## EDAT70 a 79:PAUTAParcial            -1.879e+00  1.855e+00  -1.013 0.311177
## EDAT80 o més:PAUTAParcial           -5.423e+00  2.159e+00  -2.512 0.012002
## EDAT40 a 49:Month06                  2.796e-01  1.528e+00   0.183 0.854788
## EDAT50 a 59:Month06                  6.811e-01  1.733e+00   0.393 0.694236
## EDAT60 a 69:Month06                  5.794e-01  1.658e+00   0.349 0.726823
## EDAT70 a 79:Month06                  6.891e-01  1.385e+00   0.498 0.618676
## EDAT80 o més:Month06                -5.193e-01  1.443e+00  -0.360 0.719045
## EDAT40 a 49:Month07                  7.214e-01  1.186e+00   0.609 0.542834
## EDAT50 a 59:Month07                  1.824e+00  1.170e+00   1.559 0.119029
## EDAT60 a 69:Month07                  2.046e+00  1.361e+00   1.503 0.132861
## EDAT70 a 79:Month07                  1.926e+00  1.049e+00   1.837 0.066277
## EDAT80 o més:Month07                 2.917e-01  1.097e+00   0.266 0.790368
## EDAT40 a 49:Month08                  1.275e+00  1.252e+00   1.019 0.308434
## EDAT50 a 59:Month08                  1.604e+00  1.253e+00   1.281 0.200275
## EDAT60 a 69:Month08                  2.360e+00  1.428e+00   1.652 0.098519
## EDAT70 a 79:Month08                  2.037e+00  1.136e+00   1.793 0.072977
## EDAT80 o més:Month08                -3.617e-01  7.917e-01  -0.457 0.647813
## EDAT40 a 49:Month09                 -3.082e-01  1.239e+00  -0.249 0.803495
## EDAT50 a 59:Month09                  5.108e-01  1.238e+00   0.413 0.679952
## EDAT60 a 69:Month09                  9.277e-01  1.376e+00   0.674 0.500082
## EDAT70 a 79:Month09                  1.233e+00  1.056e+00   1.167 0.243203
## EDAT80 o més:Month09                -3.651e-01  1.113e+00  -0.328 0.742834
## EDAT40 a 49:Month10                  6.109e-15  1.225e+00   0.000 1.000000
## EDAT50 a 59:Month10                  8.109e-01  1.167e+00   0.695 0.487003
## EDAT60 a 69:Month10                  9.889e-01  1.347e+00   0.734 0.462724
```

```
## EDAT70 a 79:Month10                    1.179e+00  1.019e+00   1.157 0.247427
## EDAT80 o més:Month10                  -3.174e-01  1.077e+00  -0.295 0.768325
## EDAT40 a 49:Month11                    1.260e+00  1.458e+00   0.865 0.387312
## EDAT50 a 59:Month11                    2.075e+00  1.443e+00   1.438 0.150486
## EDAT60 a 69:Month11                    2.907e+00  1.589e+00   1.829 0.067347
## EDAT70 a 79:Month11                    2.611e+00  1.332e+00   1.960 0.050001
## EDAT80 o més:Month11                   1.496e+00  1.366e+00   1.095 0.273669
## EDAT40 a 49:Month12                    1.127e+00  1.464e+00   0.770 0.441484
## EDAT50 a 59:Month12                    1.788e+00  1.453e+00   1.230 0.218570
## EDAT60 a 69:Month12                    3.085e+00  1.588e+00   1.943 0.052008
## EDAT70 a 79:Month12                    2.481e+00  1.333e+00   1.861 0.062807
## EDAT80 o més:Month12                   2.917e-01  1.394e+00   0.209 0.834322
## PAUTANo iniciada:Month06              -1.361e-02  1.252e+00  -0.011 0.991325
## PAUTAParcial:Month06                   1.208e+00  1.191e+00   1.015 0.310117
## PAUTANo iniciada:Month07               6.133e-01  9.322e-01   0.658 0.510592
## PAUTAParcial:Month07                  -3.132e+00  1.896e+00  -1.652 0.098480
## PAUTANo iniciada:Month08              -2.312e-01  9.920e-01  -0.233 0.815725
## PAUTAParcial:Month08                  -4.398e+00  1.065e+00  -4.131 3.61e-05
## PAUTANo iniciada:Month09              -1.480e+00  9.579e-01  -1.545 0.122361
## PAUTAParcial:Month09                  -4.383e+00  8.049e-01  -5.446 5.16e-08
## PAUTANo iniciada:Month10              -3.150e+00  1.054e+00  -2.987 0.002816
## PAUTAParcial:Month10                  -4.896e+00  1.149e+00  -4.260 2.04e-05
## PAUTANo iniciada:Month11              -7.868e-01  1.270e+00  -0.620 0.535465
## PAUTAParcial:Month11                  -5.352e+00  1.143e+00  -4.682 2.85e-06
## PAUTANo iniciada:Month12              -3.894e-01  1.303e+00  -0.299 0.765005
## PAUTAParcial:Month12                  -5.947e+00  1.431e+00  -4.154 3.26e-05
## EDAT40 a 49:PAUTANo iniciada:Month06  -6.290e-01  1.561e+00  -0.403 0.686938
## EDAT50 a 59:PAUTANo iniciada:Month06  -1.374e+00  1.761e+00  -0.780 0.435416
## EDAT60 a 69:PAUTANo iniciada:Month06  -1.568e+00  1.696e+00  -0.924 0.355245
## EDAT70 a 79:PAUTANo iniciada:Month06  -2.736e+00  1.463e+00  -1.870 0.061431
## EDAT80 o més:PAUTANo iniciada:Month06  8.205e-01  1.727e+00   0.475 0.634761
## EDAT40 a 49:PAUTAParcial:Month06      -1.885e+00  2.063e+00  -0.914 0.360784
## EDAT50 a 59:PAUTAParcial:Month06       1.860e-01  1.872e+00   0.099 0.920848
## EDAT60 a 69:PAUTAParcial:Month06      -1.968e+00  1.656e+00  -1.189 0.234442
## EDAT70 a 79:PAUTAParcial:Month06      -3.256e+00  1.426e+00  -2.284 0.022393
## EDAT80 o més:PAUTAParcial:Month06             NA         NA      NA       NA
## EDAT40 a 49:PAUTANo iniciada:Month07  -1.505e+00  1.208e+00  -1.246 0.212680
## EDAT50 a 59:PAUTANo iniciada:Month07  -3.261e+00  1.193e+00  -2.734 0.006249
## EDAT60 a 69:PAUTANo iniciada:Month07  -3.151e+00  1.382e+00  -2.279 0.022650
## EDAT70 a 79:PAUTANo iniciada:Month07  -3.981e+00  1.084e+00  -3.674 0.000239
## EDAT80 o més:PAUTANo iniciada:Month07 -1.027e+00  1.319e+00  -0.779 0.436237
## EDAT40 a 49:PAUTAParcial:Month07       3.763e+00  2.282e+00   1.649 0.099113
## EDAT50 a 59:PAUTAParcial:Month07       2.508e+00  2.167e+00   1.157 0.247131
## EDAT60 a 69:PAUTAParcial:Month07       9.356e-01  1.612e+00   0.580 0.561610
## EDAT70 a 79:PAUTAParcial:Month07      -1.944e+00  2.056e+00  -0.946 0.344260
## EDAT80 o més:PAUTAParcial:Month07      2.423e+00  2.476e+00   0.979 0.327796
## EDAT40 a 49:PAUTANo iniciada:Month08  -2.024e+00  1.244e+00  -1.627 0.103775
## EDAT50 a 59:PAUTANo iniciada:Month08  -2.406e+00  1.243e+00  -1.936 0.052857
## EDAT60 a 69:PAUTANo iniciada:Month08  -2.796e+00  1.421e+00  -1.967 0.049136
## EDAT70 a 79:PAUTANo iniciada:Month08  -3.085e+00  1.132e+00  -2.725 0.006426
## EDAT80 o més:PAUTANo iniciada:Month08         NA         NA      NA       NA
## EDAT40 a 49:PAUTAParcial:Month08       3.097e+00  1.679e+00   1.845 0.065025
## EDAT50 a 59:PAUTAParcial:Month08       2.310e+00  1.782e+00   1.297 0.194762
## EDAT60 a 69:PAUTAParcial:Month08              NA         NA      NA       NA
## EDAT70 a 79:PAUTAParcial:Month08              NA         NA      NA       NA
## EDAT80 o més:PAUTAParcial:Month08             NA         NA      NA       NA
```

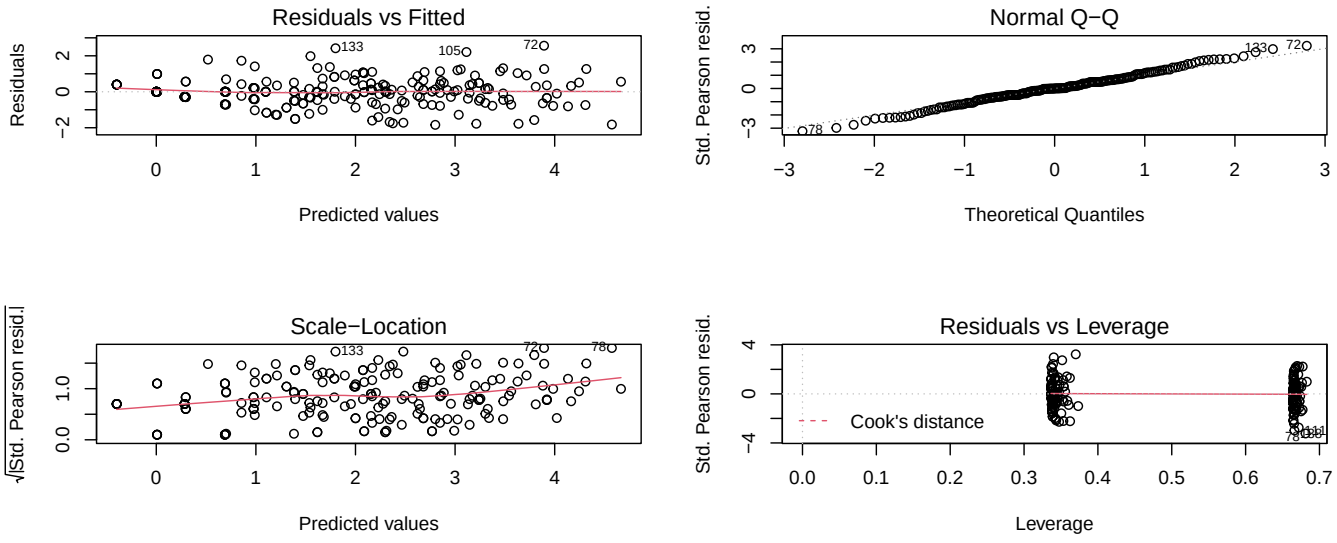
```
## EDAT40 a 49:PAUTANo iniciada:Month09  -3.666e-01  1.294e+00  -0.283 0.777006
## EDAT50 a 59:PAUTANo iniciada:Month09  -1.395e+00  1.292e+00  -1.080 0.280018
## EDAT60 a 69:PAUTANo iniciada:Month09  -1.226e+00  1.423e+00  -0.862 0.388783
## EDAT70 a 79:PAUTANo iniciada:Month09  -2.460e+00  1.142e+00  -2.154 0.031232
## EDAT80 o més:PAUTANo iniciada:Month09  3.410e-01  1.630e+00   0.209 0.834261
## EDAT40 a 49:PAUTAParcial:Month09              NA         NA      NA       NA
## EDAT50 a 59:PAUTAParcial:Month09       3.057e+00  1.532e+00   1.995 0.046023
## EDAT60 a 69:PAUTAParcial:Month09       1.837e-01  1.438e+00   0.128 0.898386
## EDAT70 a 79:PAUTAParcial:Month09              NA         NA      NA       NA
## EDAT80 o més:PAUTAParcial:Month09             NA         NA      NA       NA
## EDAT40 a 49:PAUTANo iniciada:Month10   5.909e-01  1.397e+00   0.423 0.672259
## EDAT50 a 59:PAUTANo iniciada:Month10  -1.002e+00  1.363e+00  -0.735 0.462081
## EDAT60 a 69:PAUTANo iniciada:Month10  -7.834e-01  1.521e+00  -0.515 0.606490
## EDAT70 a 79:PAUTANo iniciada:Month10  -9.329e-01  1.241e+00  -0.752 0.452045
## EDAT80 o més:PAUTANo iniciada:Month10  3.459e+00  1.536e+00   2.253 0.024285
## EDAT40 a 49:PAUTAParcial:Month10       4.203e+00  2.018e+00   2.083 0.037234
## EDAT50 a 59:PAUTAParcial:Month10              NA         NA      NA       NA
## EDAT60 a 69:PAUTAParcial:Month10       7.495e-01  1.848e+00   0.406 0.685101
## EDAT70 a 79:PAUTAParcial:Month10              NA         NA      NA       NA
## EDAT80 o més:PAUTAParcial:Month10             NA         NA      NA       NA
## EDAT40 a 49:PAUTANo iniciada:Month11  -1.913e+00  1.522e+00  -1.257 0.208820
## EDAT50 a 59:PAUTANo iniciada:Month11  -2.554e+00  1.499e+00  -1.704 0.088306
## EDAT60 a 69:PAUTANo iniciada:Month11  -2.701e+00  1.637e+00  -1.650 0.098995
## EDAT70 a 79:PAUTANo iniciada:Month11  -3.266e+00  1.406e+00  -2.323 0.020193
## EDAT80 o més:PAUTANo iniciada:Month11 -1.570e-02  1.626e+00  -0.010 0.992298
## EDAT40 a 49:PAUTAParcial:Month11              NA         NA      NA       NA
## EDAT50 a 59:PAUTAParcial:Month11              NA         NA      NA       NA
## EDAT60 a 69:PAUTAParcial:Month11       2.761e-01  1.691e+00   0.163 0.870334
## EDAT70 a 79:PAUTAParcial:Month11              NA         NA      NA       NA
## EDAT80 o més:PAUTAParcial:Month11             NA         NA      NA       NA
## EDAT40 a 49:PAUTANo iniciada:Month12  -2.177e+00  1.555e+00  -1.399 0.161706
## EDAT50 a 59:PAUTANo iniciada:Month12  -2.751e+00  1.537e+00  -1.790 0.073464
## EDAT60 a 69:PAUTANo iniciada:Month12  -3.652e+00  1.666e+00  -2.192 0.028361
## EDAT70 a 79:PAUTANo iniciada:Month12  -3.796e+00  1.445e+00  -2.627 0.008623
## EDAT80 o més:PAUTANo iniciada:Month12 -2.413e-02  1.724e+00  -0.014 0.988835
## EDAT40 a 49:PAUTAParcial:Month12              NA         NA      NA       NA
## EDAT50 a 59:PAUTAParcial:Month12              NA         NA      NA       NA
## EDAT60 a 69:PAUTAParcial:Month12              NA         NA      NA       NA
## EDAT70 a 79:PAUTAParcial:Month12              NA         NA      NA       NA
## EDAT80 o més:PAUTAParcial:Month12             NA         NA      NA       NA
##
## (Intercept)
## SEXEHome                              ***
## EDAT40 a 49
## EDAT50 a 59
## EDAT60 a 69
## EDAT70 a 79
## EDAT80 o més
## PAUTANo iniciada                      ***
## PAUTAParcial                          *
## Month06
## Month07
## Month08
## Month09
## Month10
## Month11
```

```
## Month12
## EDAT40 a 49:PAUTANo iniciada
## EDAT50 a 59:PAUTANo iniciada
## EDAT60 a 69:PAUTANo iniciada
## EDAT70 a 79:PAUTANo iniciada
## EDAT80 o més:PAUTANo iniciada         ***
## EDAT40 a 49:PAUTAParcial              .
## EDAT50 a 59:PAUTAParcial              .
## EDAT60 a 69:PAUTAParcial
## EDAT70 a 79:PAUTAParcial
## EDAT80 o més:PAUTAParcial             *
## EDAT40 a 49:Month06
## EDAT50 a 59:Month06
## EDAT60 a 69:Month06
## EDAT70 a 79:Month06
## EDAT80 o més:Month06
## EDAT40 a 49:Month07
## EDAT50 a 59:Month07
## EDAT60 a 69:Month07
## EDAT70 a 79:Month07                   .
## EDAT80 o més:Month07
## EDAT40 a 49:Month08
## EDAT50 a 59:Month08
## EDAT60 a 69:Month08                   .
## EDAT70 a 79:Month08                   .
## EDAT80 o més:Month08
## EDAT40 a 49:Month09
## EDAT50 a 59:Month09
## EDAT60 a 69:Month09
## EDAT70 a 79:Month09
## EDAT80 o més:Month09
## EDAT40 a 49:Month10
## EDAT50 a 59:Month10
## EDAT60 a 69:Month10
## EDAT70 a 79:Month10
## EDAT80 o més:Month10
## EDAT40 a 49:Month11
## EDAT50 a 59:Month11
## EDAT60 a 69:Month11                   .
## EDAT70 a 79:Month11                   .
## EDAT80 o més:Month11
## EDAT40 a 49:Month12
## EDAT50 a 59:Month12
## EDAT60 a 69:Month12                   .
## EDAT70 a 79:Month12                   .
## EDAT80 o més:Month12
## PAUTANo iniciada:Month06
## PAUTAParcial:Month06
## PAUTANo iniciada:Month07
## PAUTAParcial:Month07                  .
## PAUTANo iniciada:Month08
## PAUTAParcial:Month08                  ***
## PAUTANo iniciada:Month09
## PAUTAParcial:Month09                  ***
## PAUTANo iniciada:Month10              **
## PAUTAParcial:Month10                  ***
```

```
## PAUTANo iniciada:Month11
## PAUTAParcial:Month11                  ***
## PAUTANo iniciada:Month12
## PAUTAParcial:Month12                  ***
## EDAT40 a 49:PAUTANo iniciada:Month06
## EDAT50 a 59:PAUTANo iniciada:Month06
## EDAT60 a 69:PAUTANo iniciada:Month06
## EDAT70 a 79:PAUTANo iniciada:Month06  .
## EDAT80 o més:PAUTANo iniciada:Month06
## EDAT40 a 49:PAUTAParcial:Month06
## EDAT50 a 59:PAUTAParcial:Month06
## EDAT60 a 69:PAUTAParcial:Month06
## EDAT70 a 79:PAUTAParcial:Month06      *
## EDAT80 o més:PAUTAParcial:Month06
## EDAT40 a 49:PAUTANo iniciada:Month07
## EDAT50 a 59:PAUTANo iniciada:Month07  **
## EDAT60 a 69:PAUTANo iniciada:Month07  *
## EDAT70 a 79:PAUTANo iniciada:Month07  ***
## EDAT80 o més:PAUTANo iniciada:Month07
## EDAT40 a 49:PAUTAParcial:Month07      .
## EDAT50 a 59:PAUTAParcial:Month07
## EDAT60 a 69:PAUTAParcial:Month07
## EDAT70 a 79:PAUTAParcial:Month07
## EDAT80 o més:PAUTAParcial:Month07
## EDAT40 a 49:PAUTANo iniciada:Month08
## EDAT50 a 59:PAUTANo iniciada:Month08  .
## EDAT60 a 69:PAUTANo iniciada:Month08  *
## EDAT70 a 79:PAUTANo iniciada:Month08  **
## EDAT80 o més:PAUTANo iniciada:Month08
## EDAT40 a 49:PAUTAParcial:Month08      .
## EDAT50 a 59:PAUTAParcial:Month08
## EDAT60 a 69:PAUTAParcial:Month08
## EDAT70 a 79:PAUTAParcial:Month08
## EDAT80 o més:PAUTAParcial:Month08
## EDAT40 a 49:PAUTANo iniciada:Month09
## EDAT50 a 59:PAUTANo iniciada:Month09
## EDAT60 a 69:PAUTANo iniciada:Month09
## EDAT70 a 79:PAUTANo iniciada:Month09  *
## EDAT80 o més:PAUTANo iniciada:Month09
## EDAT40 a 49:PAUTAParcial:Month09
## EDAT50 a 59:PAUTAParcial:Month09      *
## EDAT60 a 69:PAUTAParcial:Month09
## EDAT70 a 79:PAUTAParcial:Month09
## EDAT80 o més:PAUTAParcial:Month09
## EDAT40 a 49:PAUTANo iniciada:Month10
## EDAT50 a 59:PAUTANo iniciada:Month10
## EDAT60 a 69:PAUTANo iniciada:Month10
## EDAT70 a 79:PAUTANo iniciada:Month10
## EDAT80 o més:PAUTANo iniciada:Month10 *
## EDAT40 a 49:PAUTAParcial:Month10      *
## EDAT50 a 59:PAUTAParcial:Month10
## EDAT60 a 69:PAUTAParcial:Month10
## EDAT70 a 79:PAUTAParcial:Month10
## EDAT80 o més:PAUTAParcial:Month10
## EDAT40 a 49:PAUTANo iniciada:Month11
## EDAT50 a 59:PAUTANo iniciada:Month11  .
```

```
## EDAT60 a 69:PAUTANo iniciada:Month11  .
## EDAT70 a 79:PAUTANo iniciada:Month11  *
## EDAT80 o més:PAUTANo iniciada:Month11
## EDAT40 a 49:PAUTAParcial:Month11
## EDAT50 a 59:PAUTAParcial:Month11
## EDAT60 a 69:PAUTAParcial:Month11
## EDAT70 a 79:PAUTAParcial:Month11
## EDAT80 o més:PAUTAParcial:Month11
## EDAT40 a 49:PAUTANo iniciada:Month12
## EDAT50 a 59:PAUTANo iniciada:Month12  .
## EDAT60 a 69:PAUTANo iniciada:Month12  *
## EDAT70 a 79:PAUTANo iniciada:Month12  **
## EDAT80 o més:PAUTANo iniciada:Month12
## EDAT40 a 49:PAUTAParcial:Month12
## EDAT50 a 59:PAUTAParcial:Month12
## EDAT60 a 69:PAUTAParcial:Month12
## EDAT70 a 79:PAUTAParcial:Month12
## EDAT80 o més:PAUTAParcial:Month12
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for poisson family taken to be 1)
##
##     Null deviance: 3746.09  on 220  degrees of freedom
## Residual deviance:  133.53  on  96  degrees of freedom
## AIC: 1199.7
##
## Number of Fisher Scoring iterations: 5
```

```
par(mfrow = c(2,2))
plot(model5, ask = F)
```




```
# Checking for overdispersion
PS <- sum(residuals(model5, type="pearson")^2)
PS
```

```
## [1] 130.3739
```

```
phi <- PS / model5$df.res
phi
```

```
## [1] 1.358062
```

```
Anova(model5)
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: RECOMPTE
##                   LR Chisq Df Pr(>Chisq)
## SEXE                312.39  1  < 2.2e-16 ***
## EDAT                445.54  5  < 2.2e-16 ***
## PAUTA               791.26  2  < 2.2e-16 ***
## Month              1154.92  7  < 2.2e-16 ***
## EDAT:PAUTA          481.35 10  < 2.2e-16 ***
## EDAT:Month          164.37 35  < 2.2e-16 ***
## PAUTA:Month         383.01 14  < 2.2e-16 ***
## EDAT:PAUTA:Month    100.66 50  2.893e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
qchisq(0.99, model5$df.res, lower.tail=TRUE)
```

```
## [1] 131.1412
```

Also, note that the AIC decreases for each of the increasingly complicated generalized linear models that are fitted to the data. This shows that the increasingly more complicated models perform better compared to the earlier models. Moreover, the Pearson residuals decrease as well, which points to the same conclusion.

6 Conclusions

The data analysed in this report contains covid status of patients in Catalunya between 01.05.2021 and 12.12.2021. All critical patients are counted each month, and are modelled as a general linear model with Poisson response. The covariates gender, vaccination status, age group and month are included with interactions, and the model proved to fit the data well.

Returning to the main question: does the amount of people with a *critical* infection depend on their respective vaccination statuses? Based on the final, very complicated model, the answer is yes. The Anova table shows that all covariates containing `PAUTA` are significant to any reasonable significance level. Moreover, looking at the p-values from the summary of the final model, we can conclude the following: To a significance level of $\alpha = 0.05$

- `PAUTA: No iniciada` is significantly different from the baseline (`PAUTA: Completa`) ($p = 8.26 \cdot 10^{-5}$). The estimated parameter value is ≈ 2.82, which means that the amount of critical infections per month is increased when individuals have not received any vaccines, compared to when individuals are completely vaccinated, when the rest of the individual's profile is identical (the other covariates are kept fixed).
- `Pauta: Parcial` is significantly different from the baseline (`PAUTA: Completa`) ($p \approx 0.27$). The estimated parameter value is ≈ 3.93, which means that the amount of critical infections per month is increased when individuals are partially vaccinated, compared to when individuals are completely vaccinated, when the rest of the individual's profile is identical.
- The interactions between `EDAT: 80 o més` and `PAUTA: Parcial`, and `EDAT: 80 o més` and `PAUTA: No iniciada`, are statistically significant to α. Both parameter estimates are negative, which means that the model predicts the amount of critical infections in this age group to be larger when the individuals are fully vaccinated. This is a strange result, which goes against what one might think. When reflecting on this result, keep in mind the fact that most of the people in this age group are fully vaccinated, as shown in the exploratory data analysis. This might give rise to this slightly strange result.
- The interaction effect between `PAUTA: Parcial` and `Month` is statistically significant from month 8 onwards. All the parameter estimates in question are negative, which means that the model predicts a smaller amount of critical infections in the later months when the individuals are partially vaccinated, compared to when the individuals are completely vaccinated. This is the case when the rest of the patient profile is kept fixed. This is also a strange result, which might depend on different relationships in the data. Plot 4.6 might help to explain this, since a larger and larger proportion of the population are vaccinated towards the end of the year.

Interestingly, month 8 is the first month where the amount of cases (taking all cases into account), for both sexes, is larger for the fully vaccinated compared to the two other vaccination statuses.

Without being sure about the reasons for all the predictions listed above, noting that some of them are strange, we can conclude that the amount of people with a critical infection depends on the vaccination status.