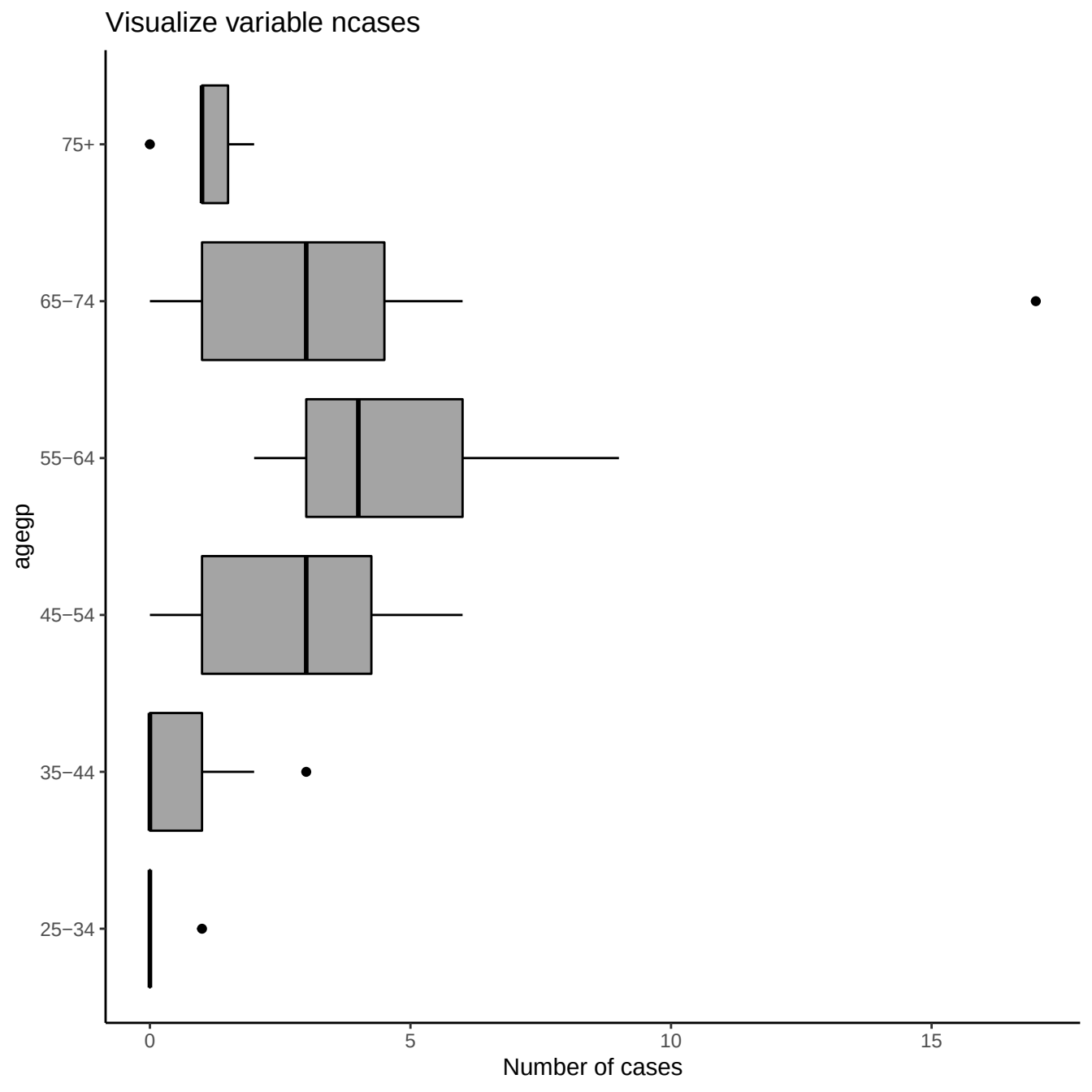
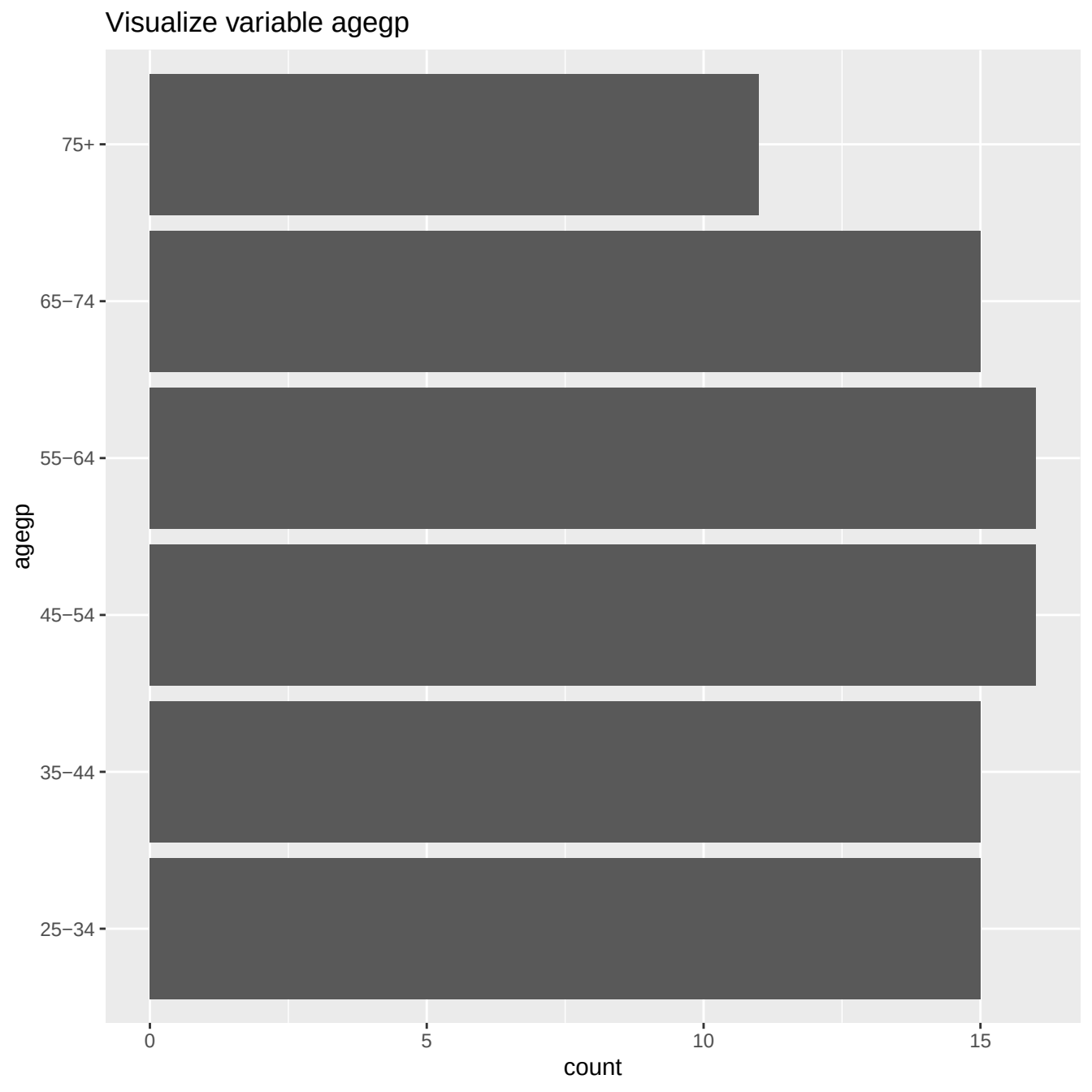
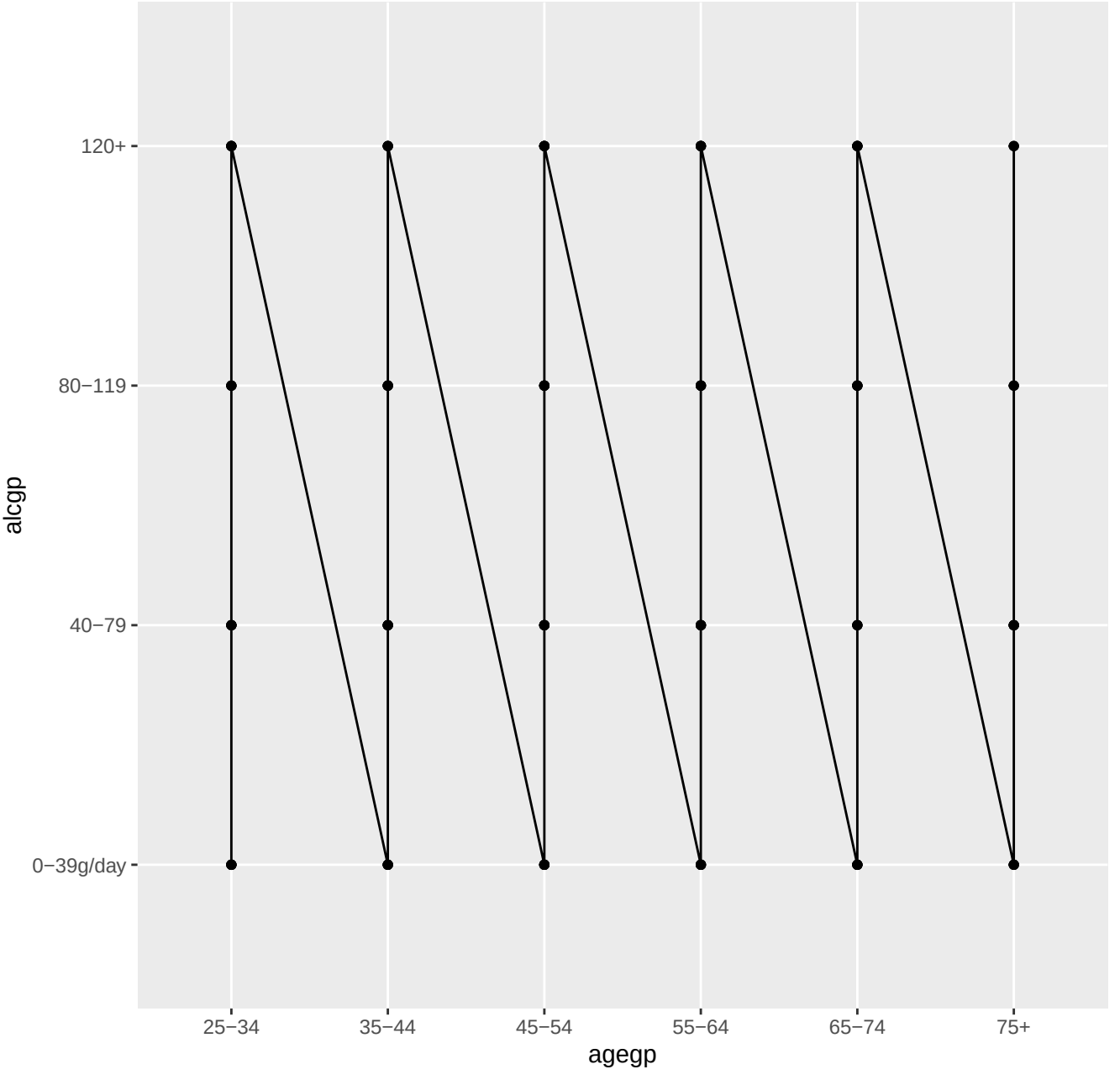
```
[1] "a. Does the dataset contain any NAs:  0"
'data.frame':   88 obs. of  5 variables:
 $ agegp    : Ord.factor w/ 6 levels "25-34"<"35-44"<..: 1 1 1 1 1 1 1 1 1 1 ...
 $ alcgp    : Ord.factor w/ 4 levels "0-39g/day"<"40-79"<..: 1 1 1 1 2 2 2 2 3 3 ...
 $ tobgp    : Ord.factor w/ 4 levels "0-9g/day"<"10-19"<..: 1 2 3 4 1 2 3 4 1 2 ...
 $ ncases   : num  0 0 0 0 0 0 0 0 0 0 ...
 $ ncontrols: num  40 10 6 5 27 7 4 7 2 1 ...
[1] "What is the type of variable tobgp: Ord.factor"

[1] "b. Does this variable contain outliers: Yes this variable contains outliers"
[1] "b. Do you think these values are really outliers or legitimate values?: These values seem to be legitimate values as they are not outside the scope."
Saving 7 x 7 in image
function (x, y, ...)
UseMethod("plot")
<bytecode: 0x000000001bfacbb0>
<environment: namespace:base>
```

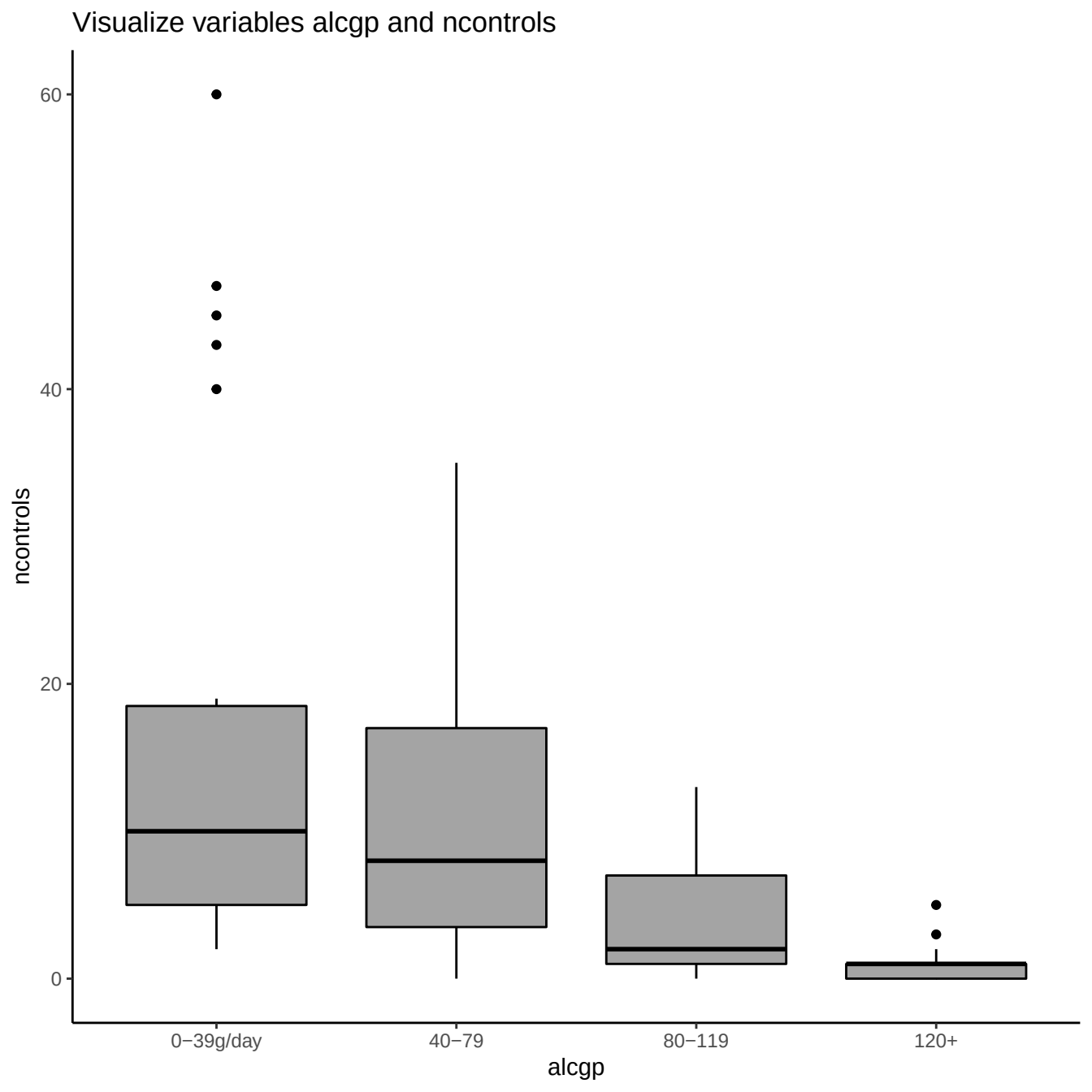
```r
library(ggplot2)

data("esoph")

naCount <- sum(is.na(esoph))

print(paste("a. Does the dataset contain any NAs: ", naCount))
str(esoph)
print("What is the type of variable tobgp: Ord.factor")
cat("\n")

print("b. Does this variable contain outliers: Yes this variable contains
        outliers")
print("b. Do you think these values are really outliers or legitimate values?:
These values seem to be legitimate values as they are not outside the scope.")


ggplot(esoph, aes(x=ncases, y=agegp)) +
  geom_boxplot(fill='#A4A4A4', color="black")
  + ggtitle('Visualize variable ncases')
  + theme_classic() + xlab('Number of cases')

ggplot(esoph) + geom_bar(aes(y = agegp)) + ggtitle('Visualize variable agegp')

ggplot(data=esoph, aes(x=agegp, y=alcgp, group=1)) +
  geom_line()+
  geom_point() + ggtitle('Visualize variables agegp and alcgp.')

ggplot(esoph, aes(x=alcgp, y=ncontrols)) +
  geom_boxplot(fill='#A4A4A4', color="black")+
  ggtitle('Visualize variables alcgp and ncontrols') +
  theme_classic()

ggplot(esoph, aes(x=ncases , y=ncontrols)) + geom_point()
+ ggtitle('Visualize variables ncases and ncontrols')

ggplot(esoph, aes(x=ncases , y=ncontrols, color=alcgp))
+ geom_point()
+ ggtitle('Visualize variables ncases and ncontrols and alcgp')
ggsave("homework#3.pdf")

print(plot)
```

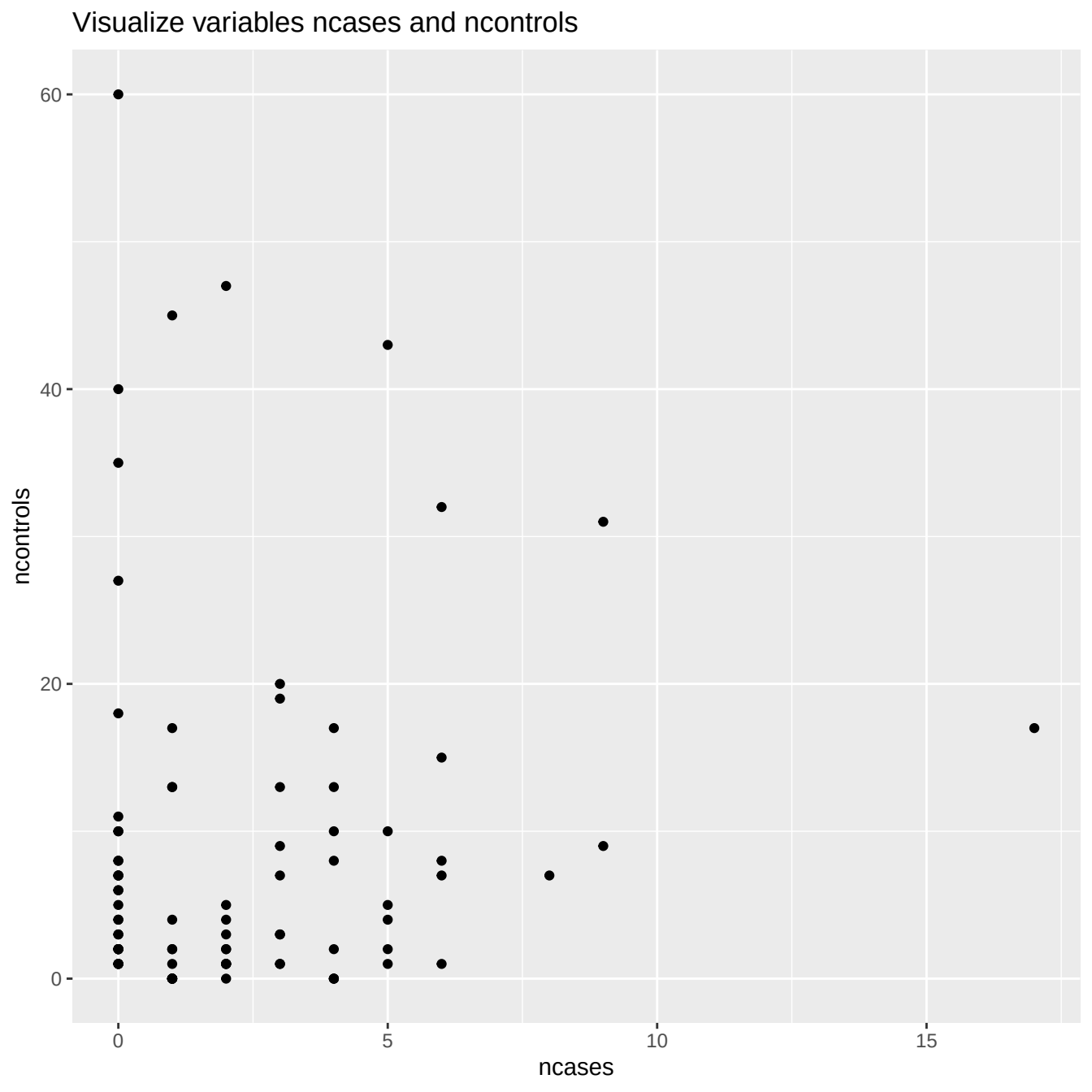

Visualize variable ncases
75+
65−74
55−64
45−54
35−44
25−34
agegp
0
5
10
15
Number of cases

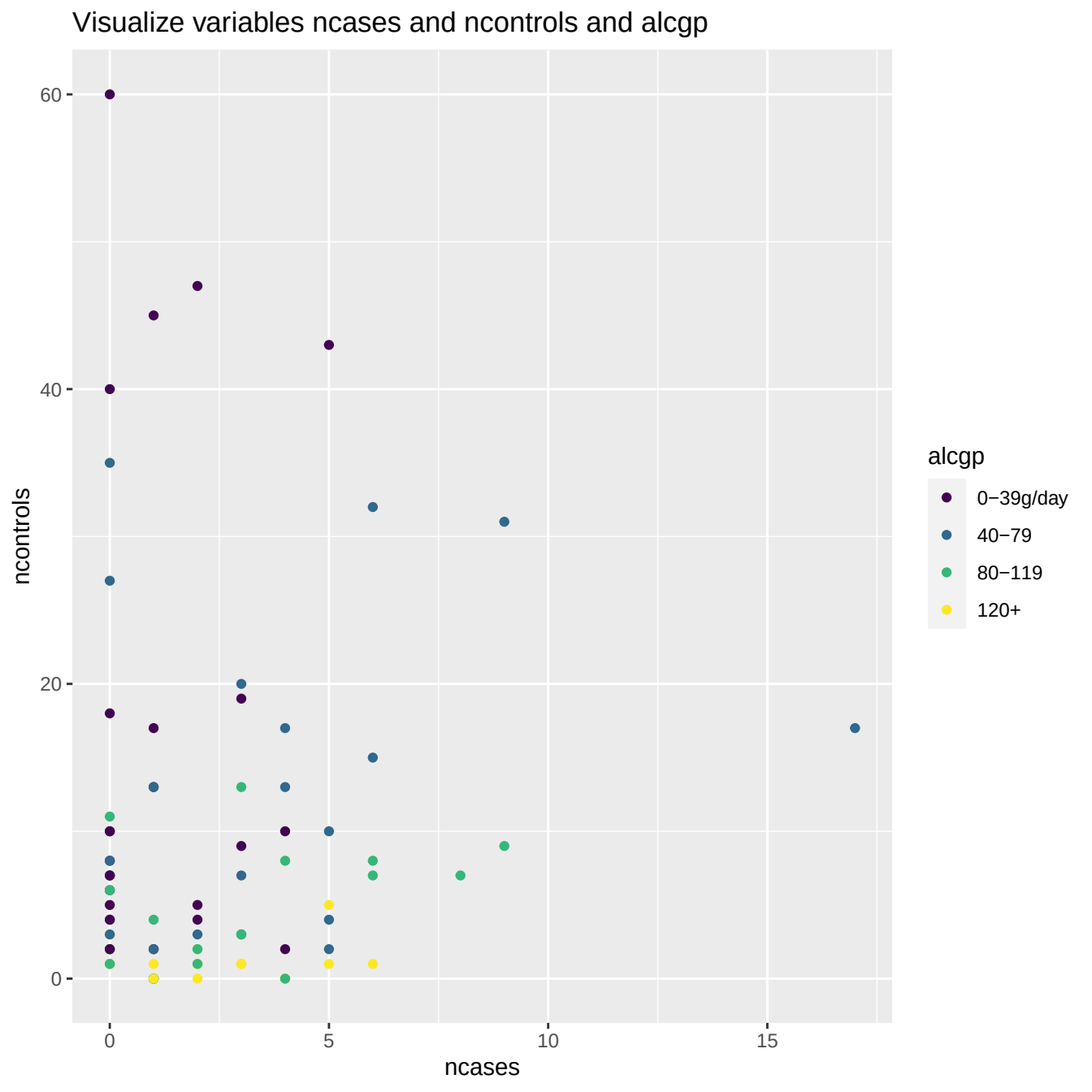

Visualize variable agegp
75+
65−74
55−64
45−54
35−44
25−34
agegp
0
5
10
15
count

Visualize variables agegp and alcgp.




Visualize variables alcgp and ncontrols
60
40
20
0
ncontrols
0−39g/day
40−79
80−119
120+
alcgp


Visualize variables ncases and ncontrols
60
40
20
0
ncontrols
0
5
10
15
ncases


Visualize variables ncases and ncontrols and alcgp
ncontrols
ncases
0
20
40
60
0
5
10
15
alcgp
0−39g/day
40−79
80−119
120+