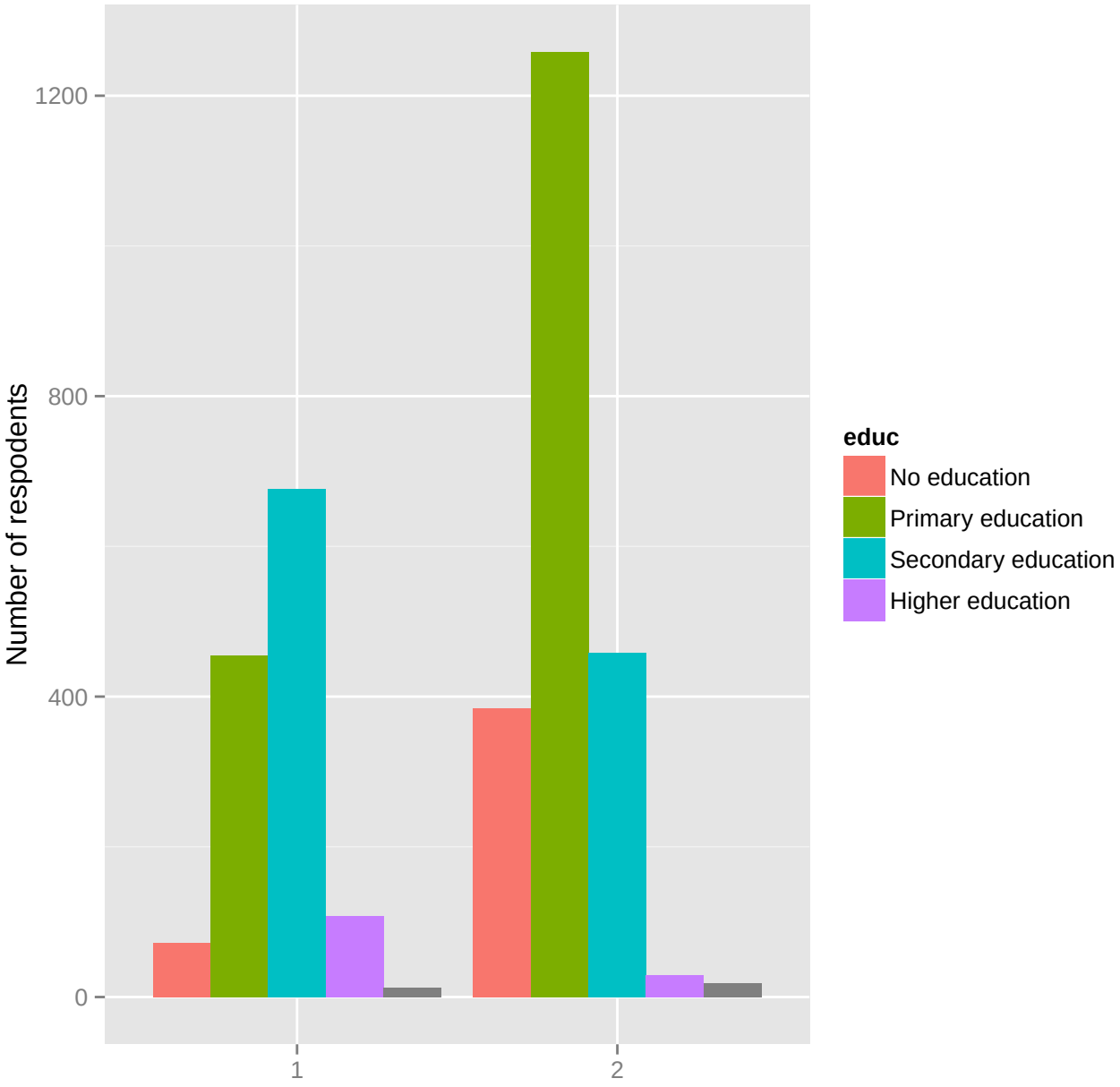
Solution to s3.6-graph-exercise

December 11, 2013

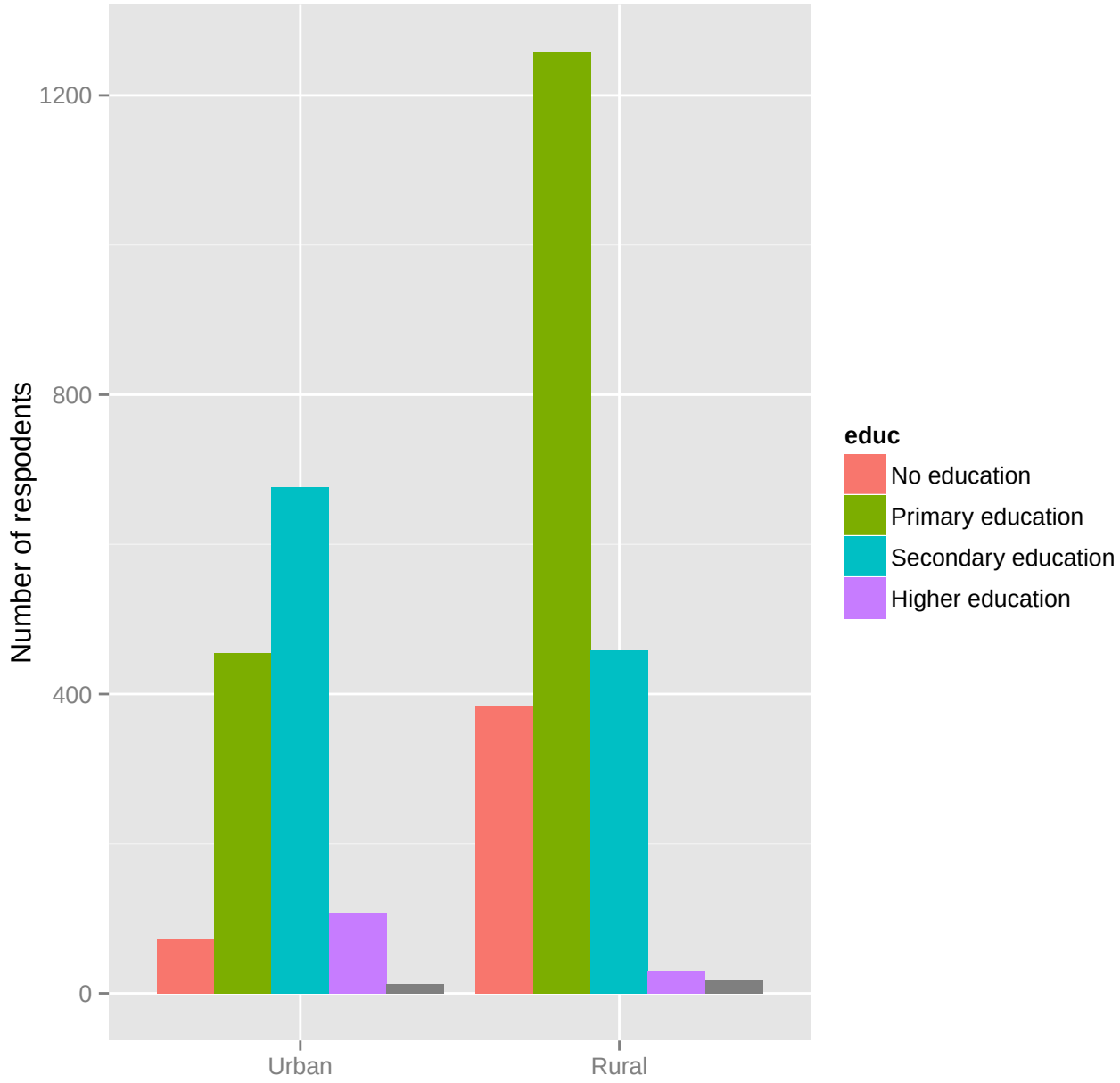
```r
# setting the working directory


# load the packages
library(ggplot2)  #used for plotting
library(foreign)  #allow reading of dta files
# library(reshape2)
library(doBy)  #provide summary of data

## Loading required package:  survival
## Loading required package:  splines
## Loading required package:  MASS
```

******** QUESTION 2 SOLUTION *******

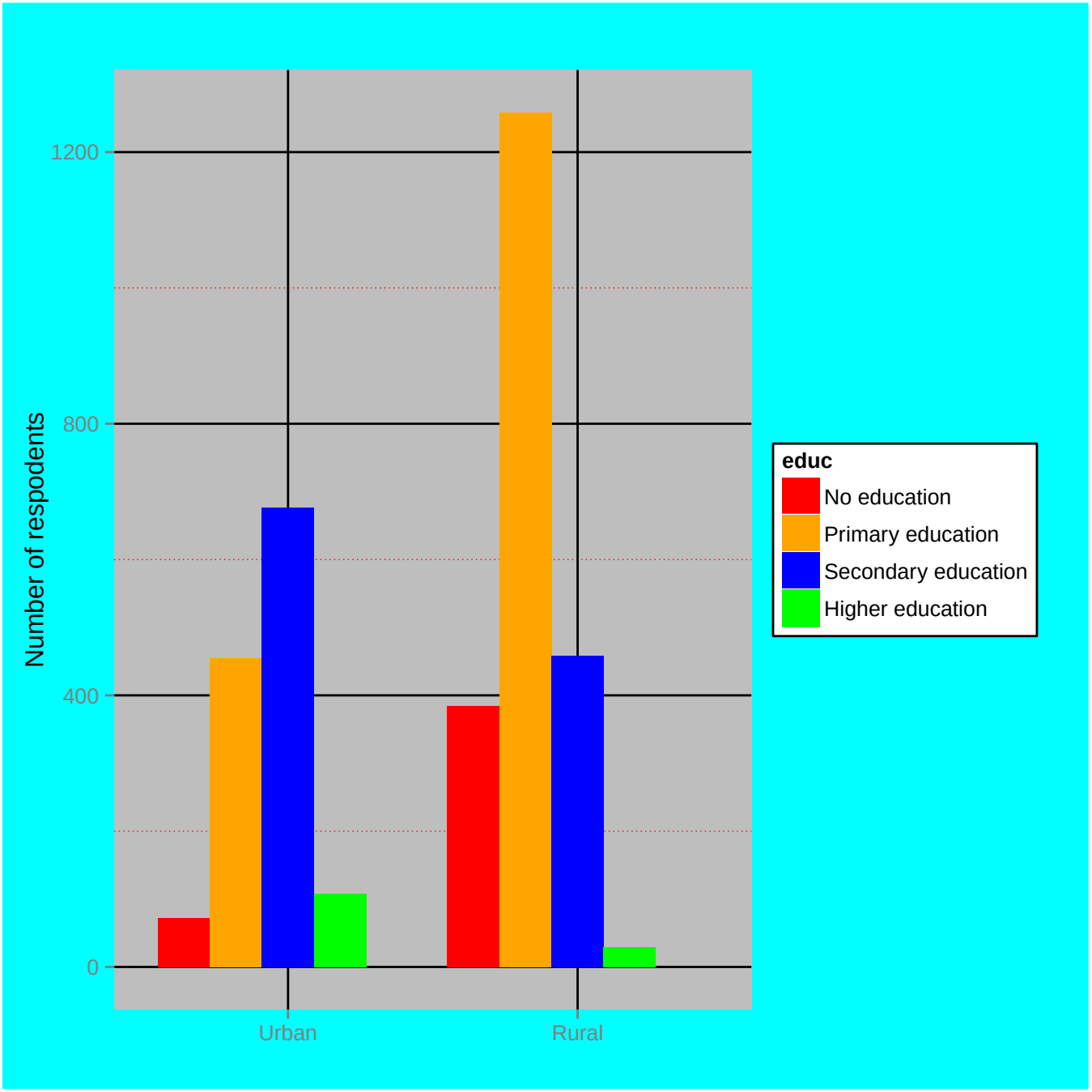
```r
# load the data set
zambia3 <- as.data.frame(read.dta("zambia3.dta", convert.dates = TRUE))

# summarize the count of data by education and urban/rural
total.id <- summaryBy(id ~ educ + urban, data = zambia3, FUN = function(x) c(count = length(x)))
as.data.frame(total.id)

##                    educ urban id.count
## 1          No education     1       72
## 2          No education     2      385
## 3     Primary education     1      455
## 4     Primary education     2     1258
## 5     Primary education    NA        1
## 6   Secondary education     1      676
## 7   Secondary education     2      458
## 8   Secondary education    NA        1
## 9      Higher education     1      108
## 10     Higher education     2       29
## 11                 <NA>     1       13
## 12                 <NA>     2       18

# *DISTRIBUTION OF EDUCATION IN URBAN AND RURAL AREAS
ggplot(total.id[!is.na(total.id$urban), ], aes(as.factor(urban), fill = educ,
    weight = id.count)) + geom_bar(position = "dodge") + ylab("Number of respodents") +
    xlab("") + ggtitle("")
```




```
# *YOU PROBABLY WANT TO ADD A TEXT LABEL TO THE CATEGORY AXIS. *EITHER LABEL
# THE URBAN VARIABLE OR:
total.id$urban <- factor(total.id$urban, levels = c(1, 2), labels = c("Urban",
    "Rural"))
ggplot(total.id[!is.na(total.id$urban), ], aes(as.factor(urban), fill = educ,
    weight = id.count)) + geom_bar(position = "dodge") + ylab("Number of respodents") +
    xlab("") + ggtitle("")
```

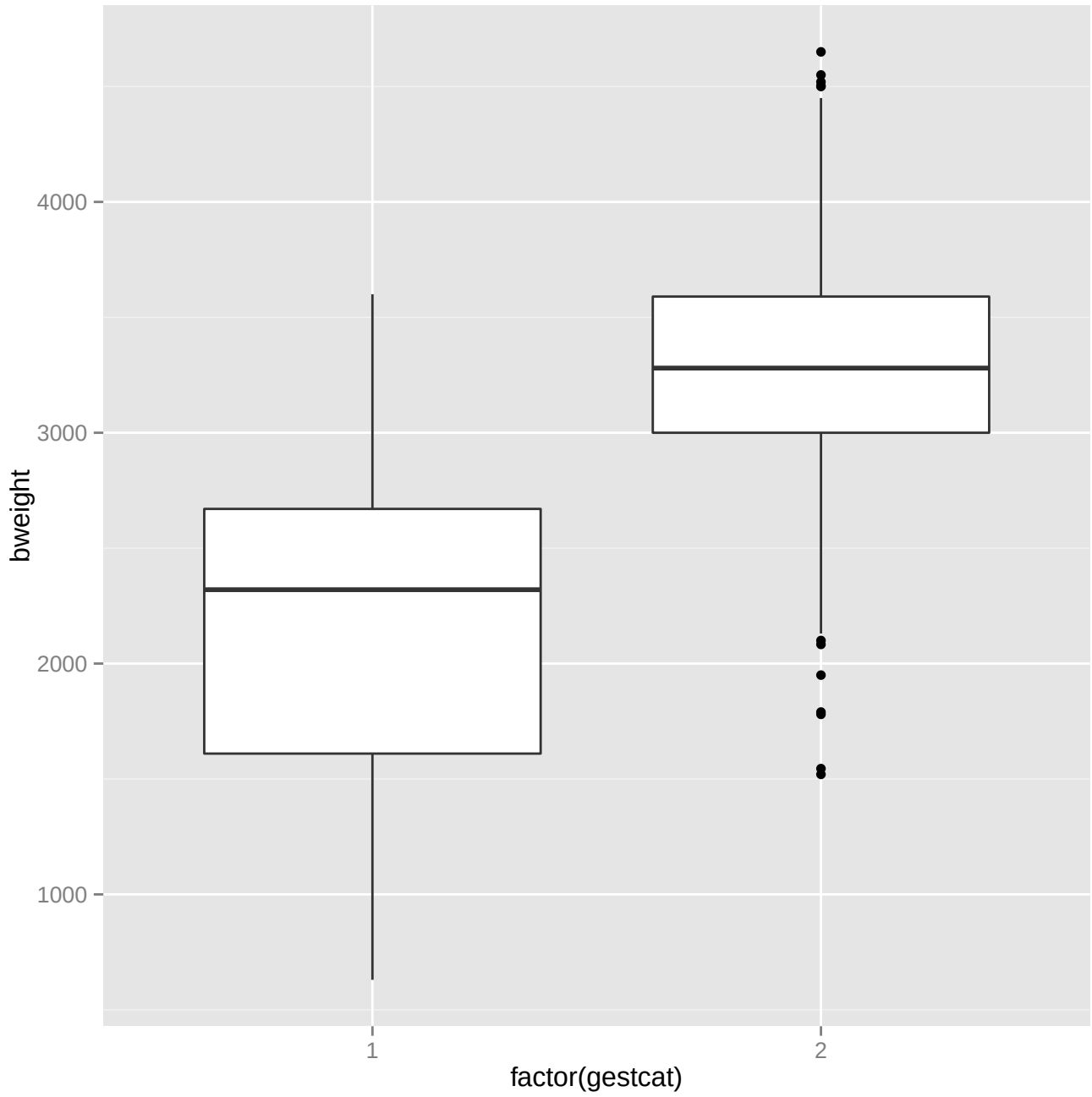



```
ggsave(filename = "gr2b4.wmf")

## Saving 7 x 7 in image

# *WITH MOST OF THE COLOURS CHANGED
ggplot(total.id[!is.na(total.id$urban), ], aes(as.factor(urban), fill = educ,
    weight = id.count)) + geom_bar(position = "dodge") + ylab("Number of respodents") +
    xlab("") + ggtitle("") + scale_fill_manual(values = c(`No education` = "red",
    `Primary education` = "orange", `Secondary education` = "blue", `Higher education` = "green")
    theme(panel.background = element_rect(fill = "gray"), panel.grid.major = element_line(colour
        panel.grid.minor = element_line(colour = "red", linetype = "dotted"),
        plot.background = element_rect(fill = "cyan"), legend.background = element_rect(colour =
        legend.key = element_rect(fill = "yellow"))
```

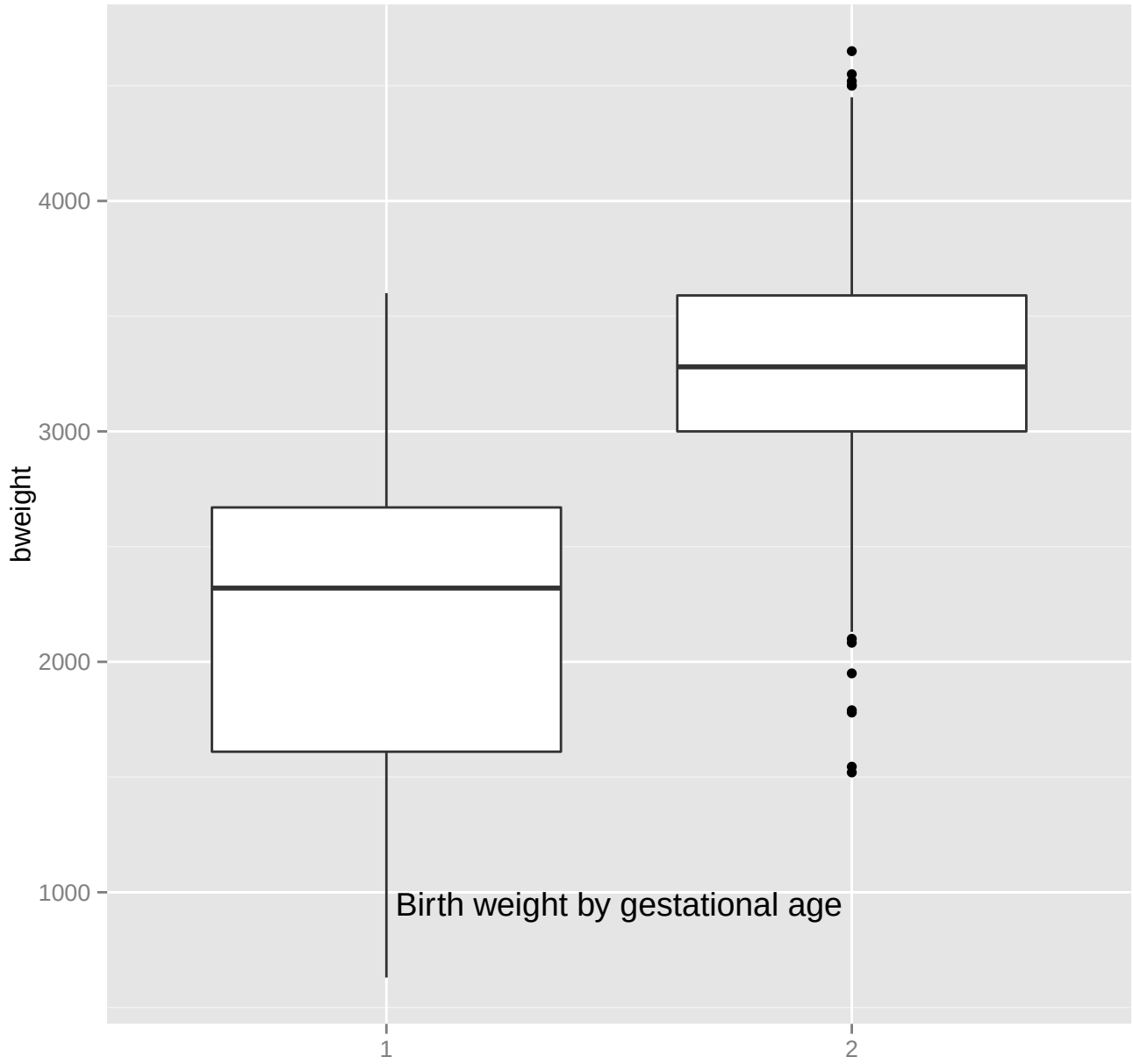


******** QUESTION 3 SOLUTION *******

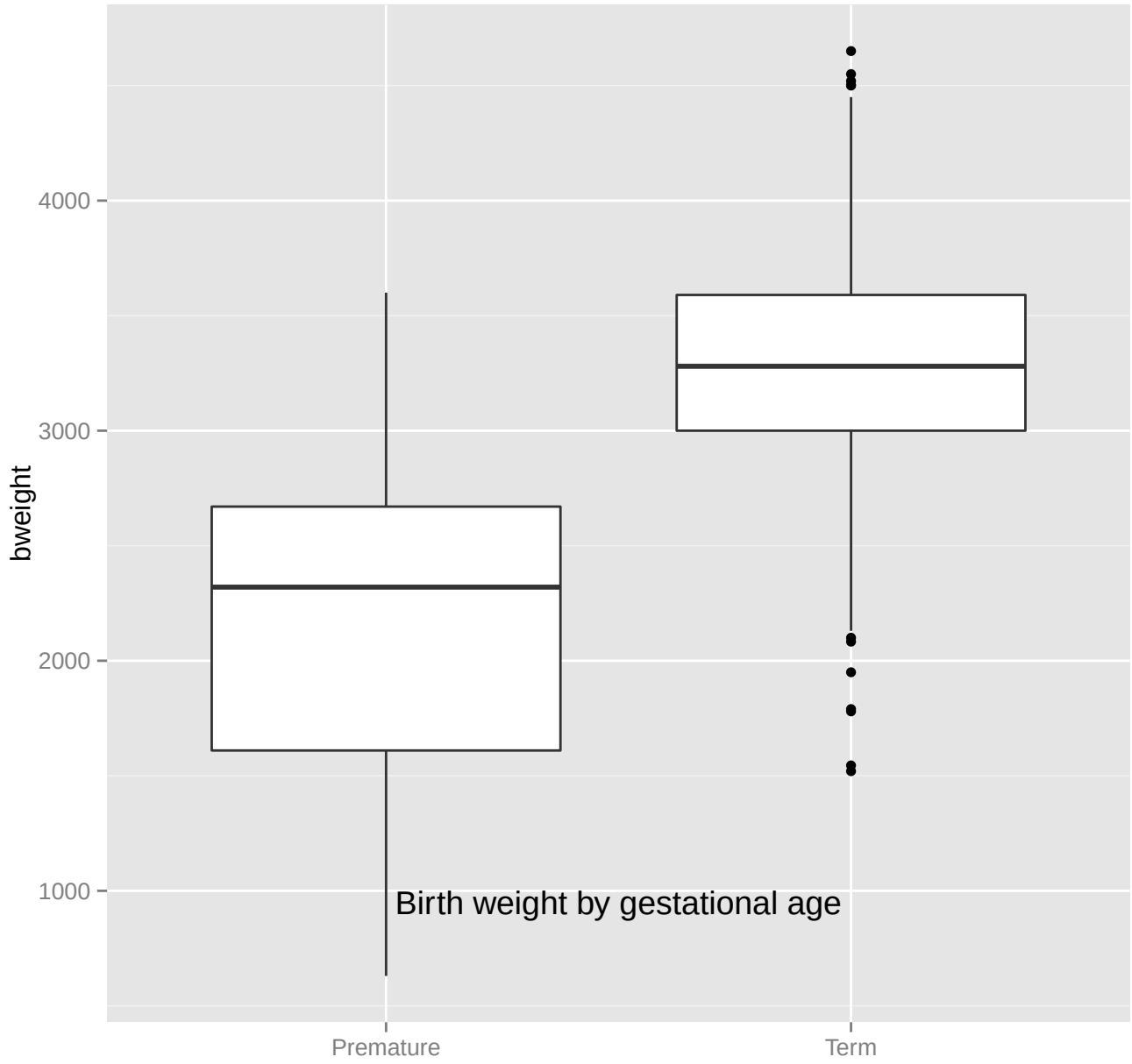
```
# load the data set
bab9 <- as.data.frame(read.dta("bab9.dta", convert.dates = TRUE))
bab9 <- as.data.frame(bab9)
ggplot(bab9, aes(factor(gestcat), bweight)) + geom_boxplot()
```




```
# *WITH A TITLE BELOW THE GRAPH
ggplot(bab9, aes(factor(gestcat), bweight)) + geom_boxplot() + ggtitle("Birth weight by gestation
    theme(plot.title = element_text(vjust = -54)) + xlab("")
```

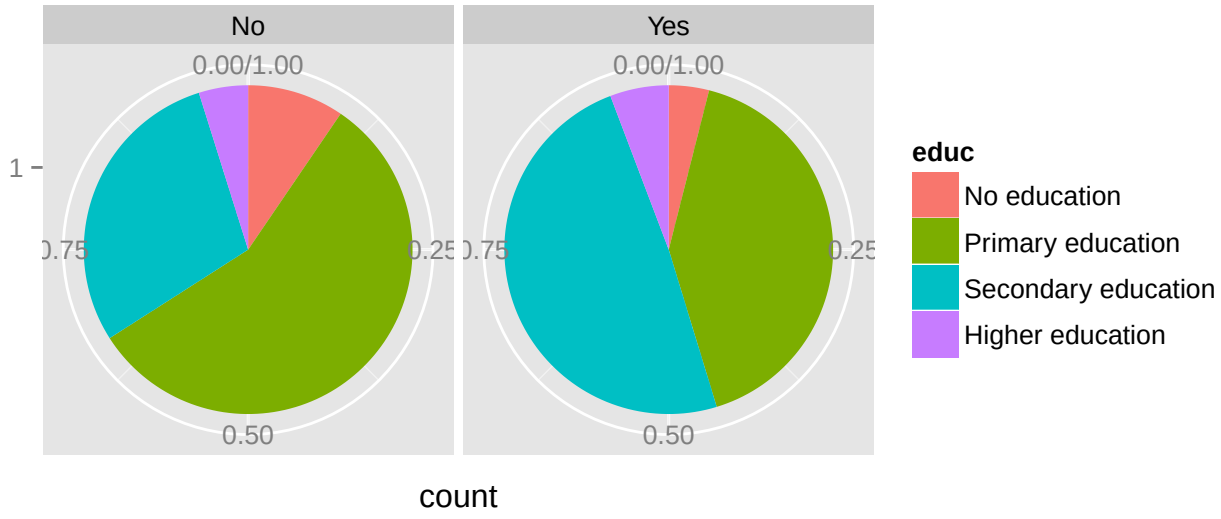



```
# *AGAIN YOU CAN LABEL THE gestcat VARIABLE OR WRITE YOUR *OWN LABELS ON THE
# X-AXIS
bab9$gestcat <- factor(bab9$gestcat, levels = c(1, 2), labels = c("Premature",
    "Term"))
ggplot(bab9, aes(factor(gestcat), bweight)) + geom_boxplot() + ggtitle("Birth weight by gestation
    theme(plot.title = element_text(vjust = -54)) + xlab("")
```

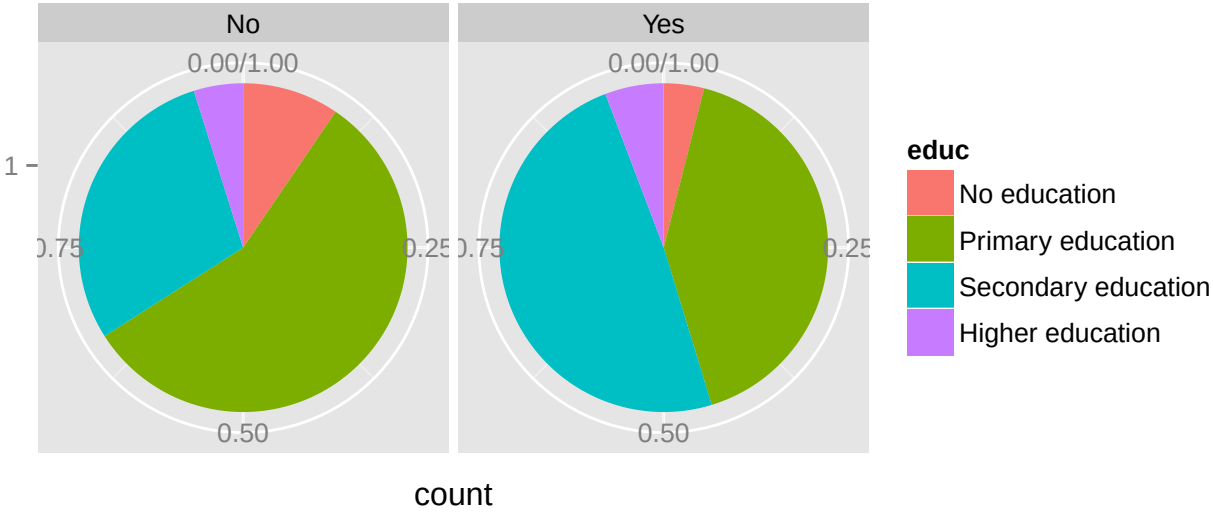


******** QUESTION 4 SOLUTION *******

```
# load the data set
zambia4 <- as.data.frame(read.dta("zambia4_isingo.dta", convert.dates = TRUE))

## Warning:  value labels ('currmarrlbl') for 'married' are missing

zambia4 <- as.data.frame(zambia4)
# *PIE CHART FOR EDUCATIONAL LEVEL OF WOMEN WHO USED A CONDOM *AT LAST SEX
# AND THOSE WHO DID NOT summarize the count of data by education and
# urban/rural
ggplot(zambia4[!is.na(zambia4$clastsex), ], aes(x = factor(1), fill = educ,
    weight = weight)) + coord_polar(theta = "y") + scale_x_discrete("") + facet_grid(facets = .
    clastsex) + geom_bar(width = 1, position = "fill")
```




```
# *WITH ALTERED CAPTION
ggplot(zambia4[!is.na(zambia4$clastsex), ], aes(x = factor(1), fill = educ,
    weight = weight)) + coord_polar(theta = "y") + scale_x_discrete("") + facet_grid(facets = .
    clastsex) + geom_bar(width = 1, position = "fill") + ggtitle("Educational level of women by
    theme(plot.title = element_text(vjust = -30))
```



Educational level of women by condom use at last sex