Comics

rishabh

GGPLOT Practice Excercise

1.Print the first rows of the data

```
r1 <- read.csv("Comics_Data.csv")
r1[1,]
```

```
##   X                      name     id align        eye       hair gender
## 1 1 Spider-Man (Peter Parker) Secret  Good Hazel Eyes Brown Hair   Male
##    gsm             alive appearances first_appear publisher
## 1 <NA> Living Characters        4043       Aug-62    marvel
```

```
names(r1)
```

```
##  [1] "X"            "name"         "id"           "align"
##  [5] "eye"          "hair"         "gender"       "gsm"
##  [9] "alive"        "appearances"  "first_appear" "publisher"
```

2.Check Levels of Align Variable

```
r2 <- r1[ ,c("align")]
levels(r2)
```

```
## [1] "Bad"                "Good"               "Neutral"
## [4] "Reformed Criminals"
```

3.Check Levels of Gender

```
r3 <- r1[ ,c("gender")]
levels(r3)
```

```
## [1] "Female" "Male"   "Other"
```

4.Create a 2 way table using align and gender

```
table(r1$align, r1$gender)
```

```
##
##                      Female Male Other
##   Bad                  1573 7561    32
##   Good                 2490 4809    17
##   Neutral               836 1799    17
##   Reformed Criminals      1    2     0
```

5.Remove 'Reformed Criminals' level from align variable using dplyr

```
library(dplyr)
```

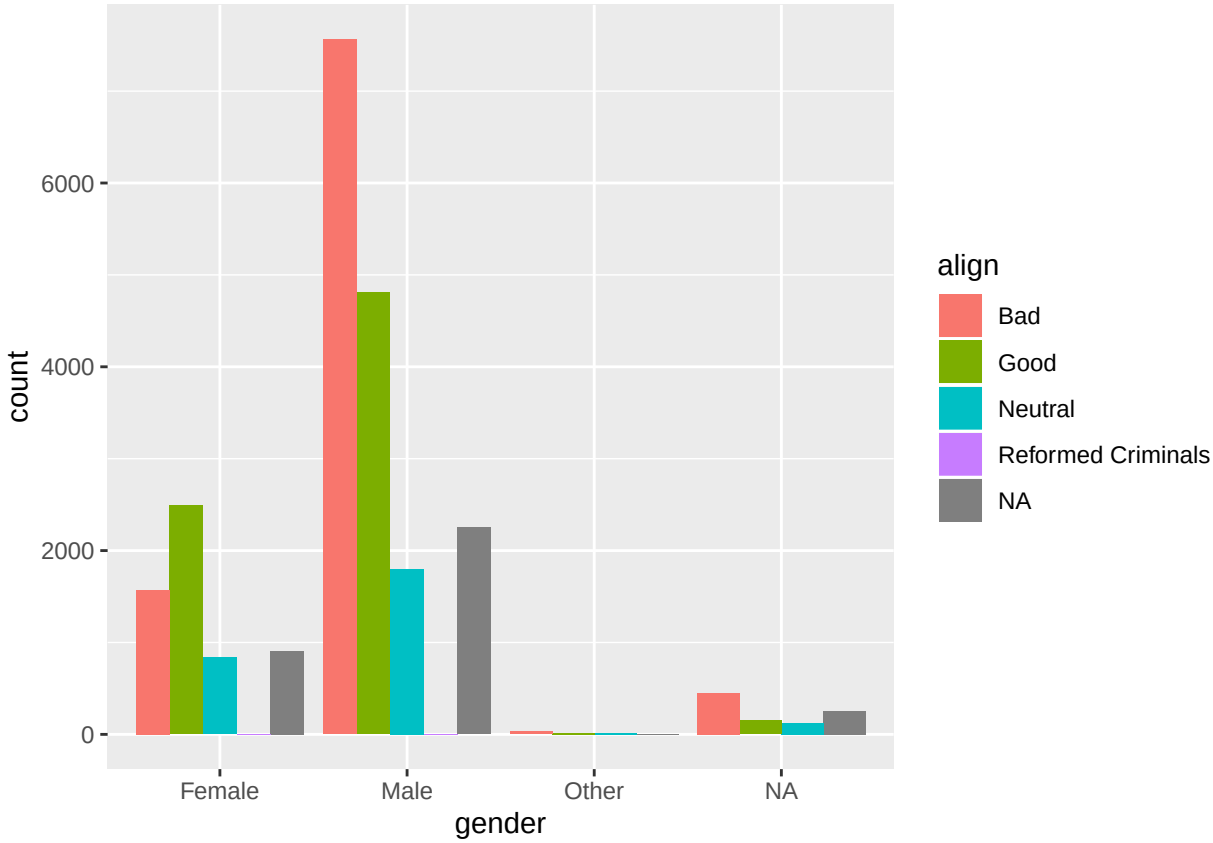
```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
r4 <- factor(droplevels(r2, "Reformed Criminals"))
head(r4)
```

```
## [1] Good    Good    Neutral Good    Good    Good
## Levels: Bad Good Neutral
```

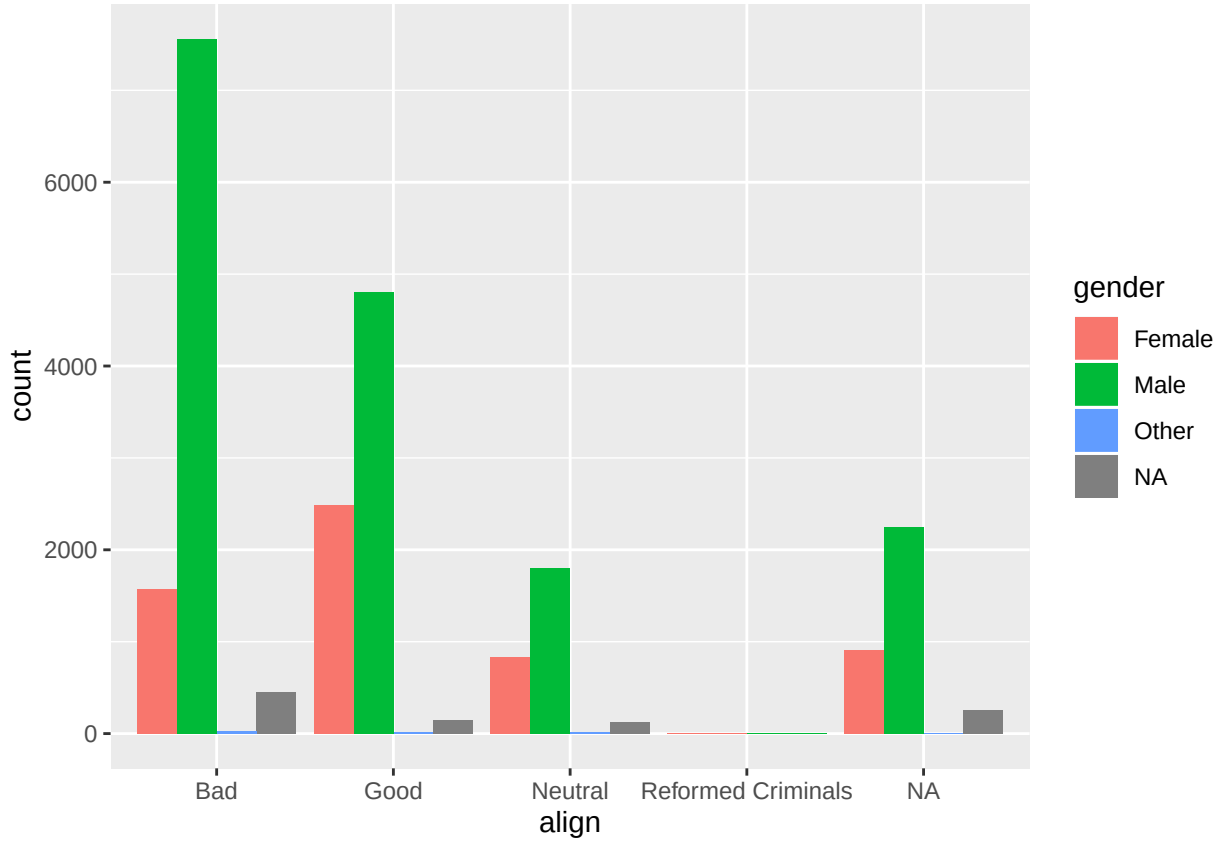
6.Create a side by side bar chart of gender by align variable using ggplot

```
library(ggplot2)
ggplot(r1, aes(x = gender, fill = align)) + geom_bar(position = "dodge")
```



7.Create a side by side bar chart of alignment by gender using ggplot

```
ggplot(r1, aes(x = align, fill = gender)) + geom_bar(position = "dodge")
```



8.Write the interpretation of bar charts

1. There is an association between Gender and Alignment.
2. There are less number of female characters than that of male characters.
3. bad alignment is less in female.
4. The Characters with Neutral alignment, Males are the most common.

9.Create a table of number of comics using align variable

```
r5 <- table(r1$align)
r5
```

```
##
##                Bad               Good            Neutral
##               9615               7468               2773
## Reformed Criminals
##                  3
```

10.Create a propotional table using the above table using prop.table(), what is the proportion of Bad, Good and Neutral

```
r6 <- table(r1$align)
prop.table(r6)
```

```
##
##                 Bad                Good             Neutral
##         0.484163352         0.376051161         0.139634423
## Reformed Criminals
##         0.000151065
```

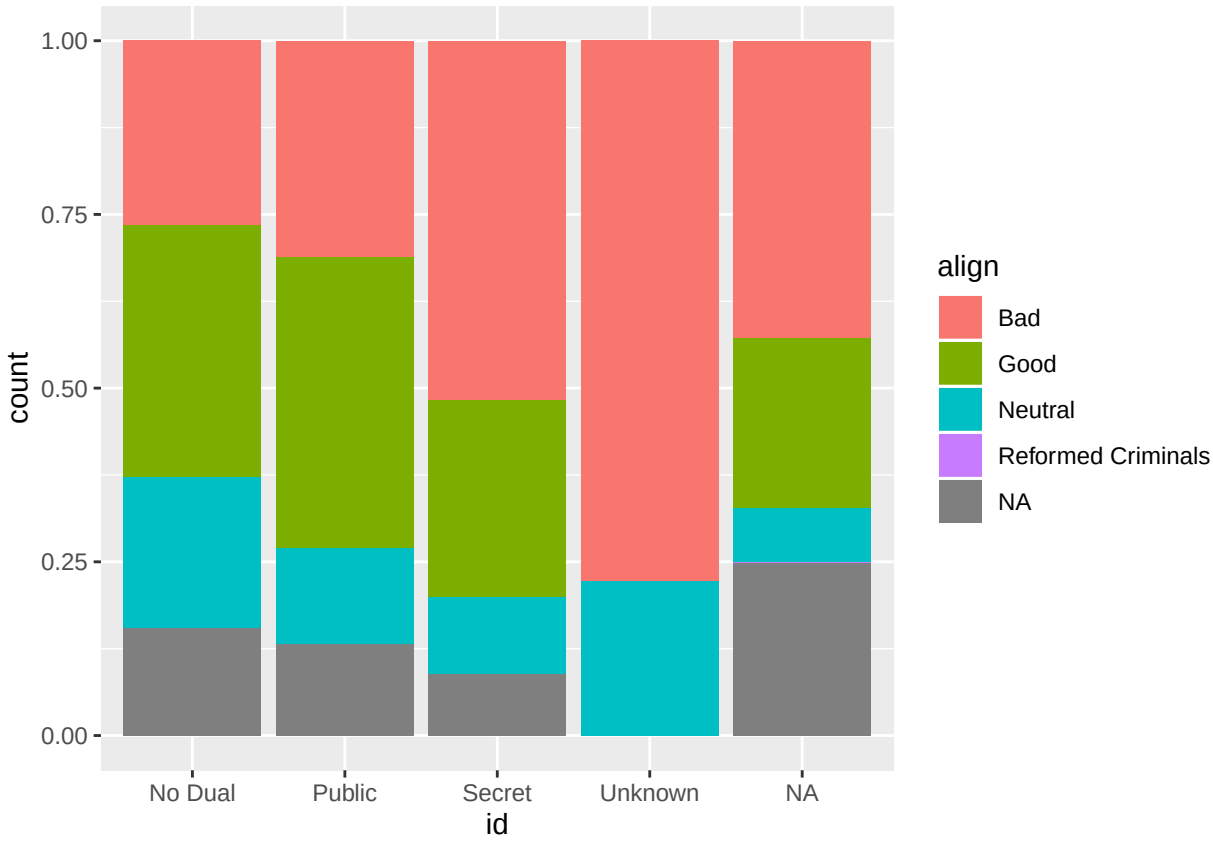
11.Create a propotional table using the above table using prop.table(), what is the proportion of No Dual, Public, Secret and Unknown

```
r7 <- table(r1$id)
prop.table(r7)
```

```
##
##      No Dual       Public       Secret      Unknown
## 0.1022356910 0.3999085139 0.4973411859 0.0005146092
```

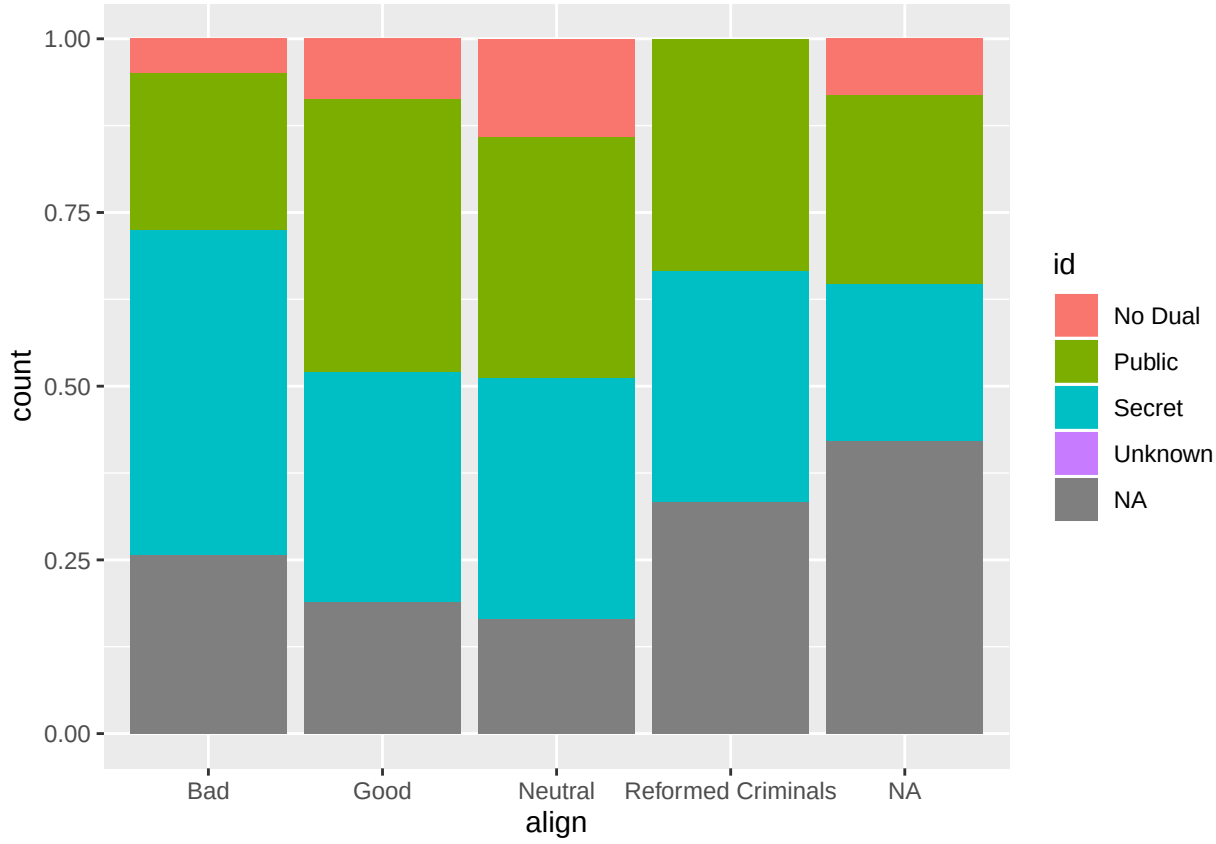
12. Plot the proportions using ggplot(use bar plot); x = id, y = proportion, fill = align

```
ggplot(data = r1, aes(x = id, fill = align)) + geom_bar(position = "fill")
```



13. Plot same as above, swap the id and align variable

```
ggplot(data = r1, aes(x = align, fill = id)) + geom_bar(position = "fill")
```



14. Using proportions, answer approximately what proportion of all female characters are good

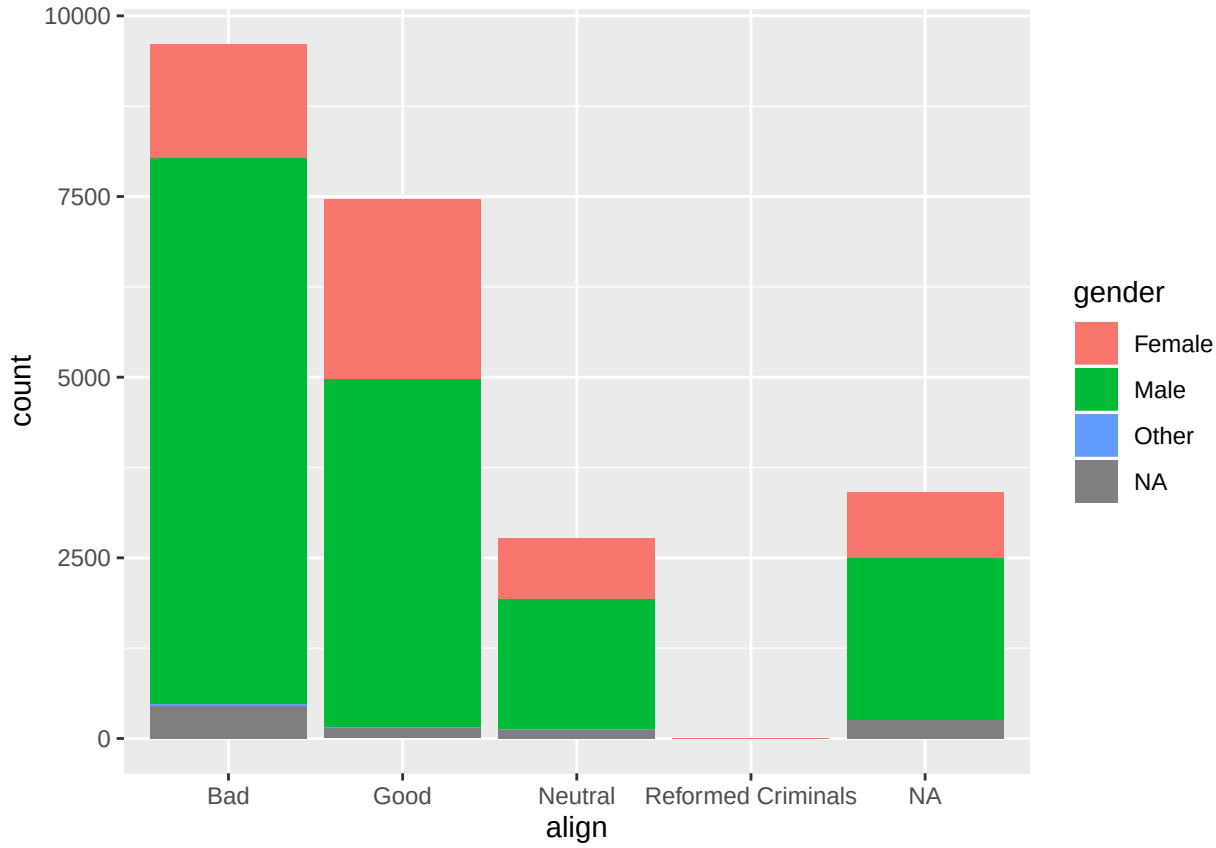
```
r8 <- table(r1$align, r1$gender)
prop.table(r8, 2)
```

```
##
##                          Female         Male        Other
##   Bad                0.3210204082 0.5335544422 0.4848484848
##   Good               0.5081632653 0.3393550208 0.2575757576
##   Neutral            0.1706122449 0.1269494037 0.2575757576
##   Reformed Criminals 0.0002040816 0.0001411333 0.0000000000
```

Answer = 50%

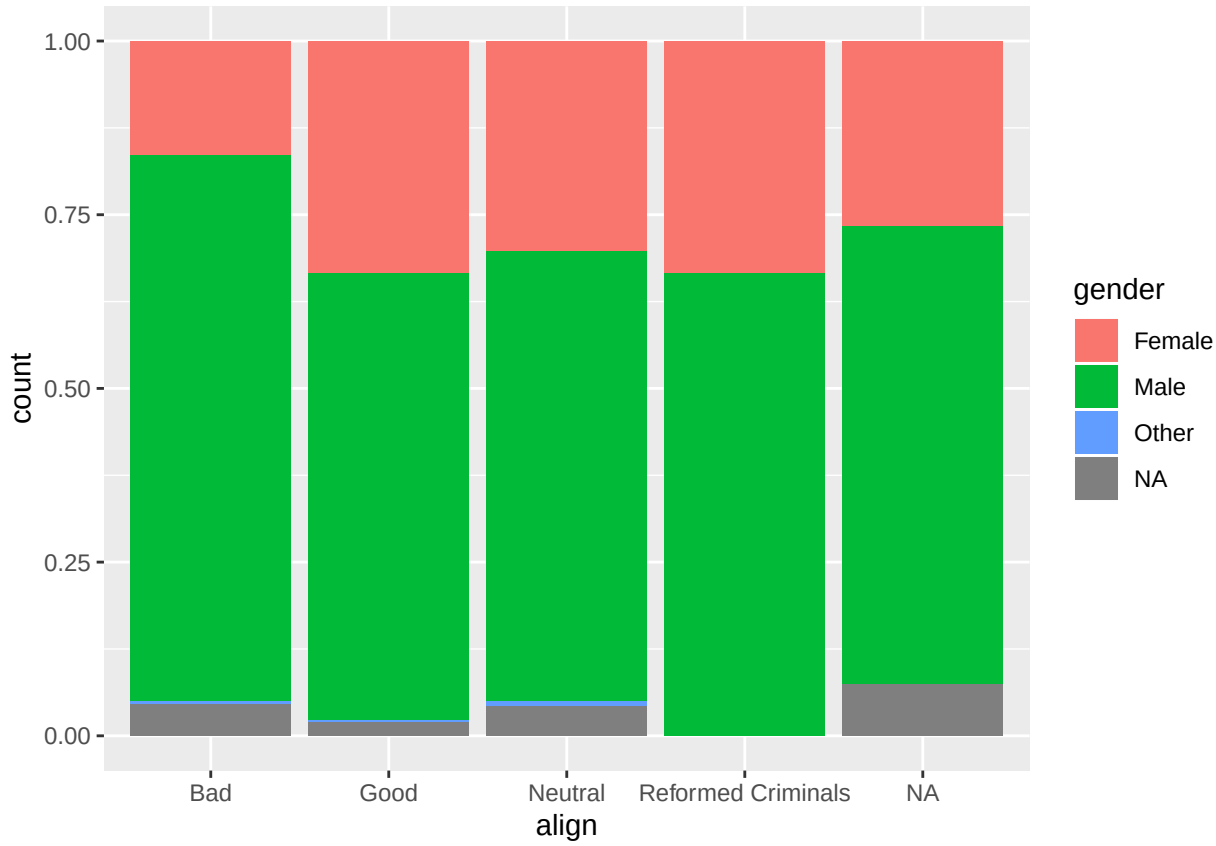
15. Plot the count of gender using align (use ggplot with bar)

```
ggplot(r1, aes(x = align, fill = gender)) + geom_bar()
```



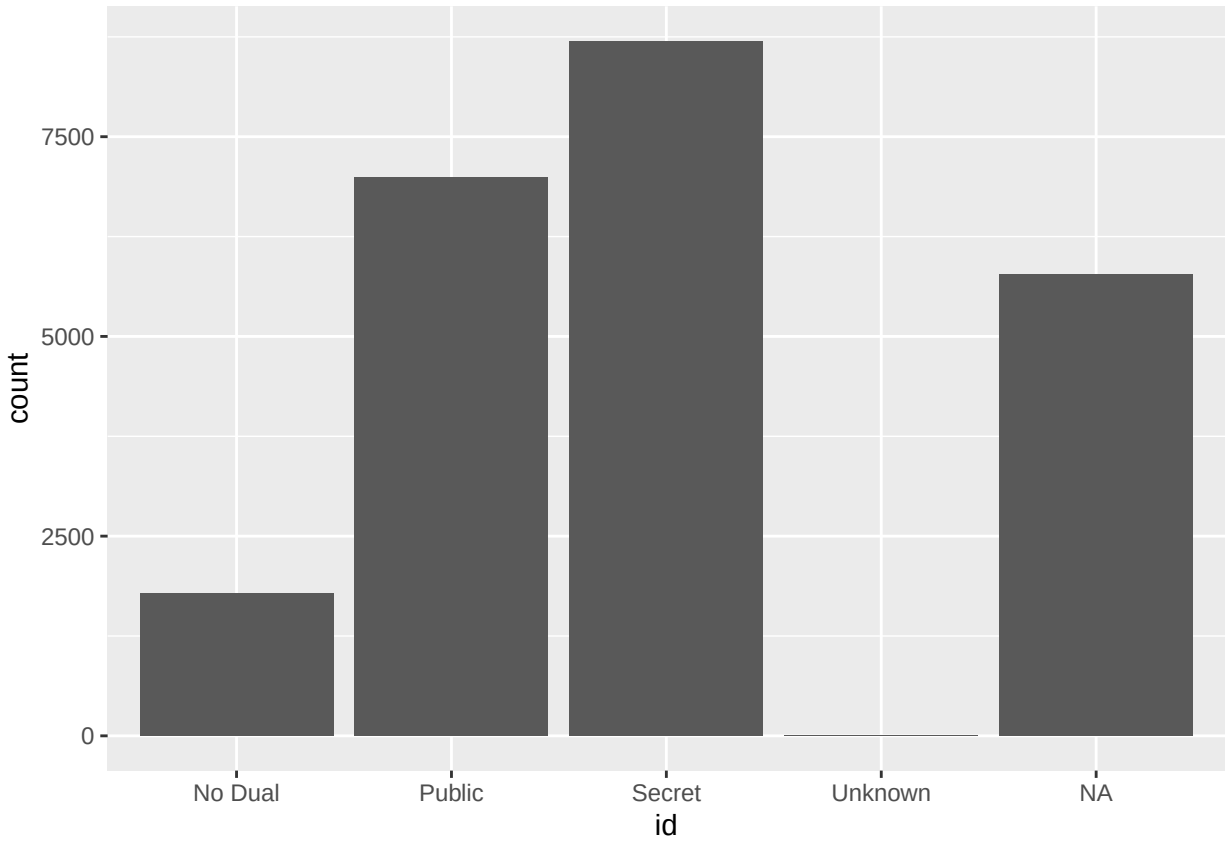
16. Plot the above plot using proportions

```
ggplot(r1, aes(x = align, fill = gender)) + geom_bar(position = "fill")
```



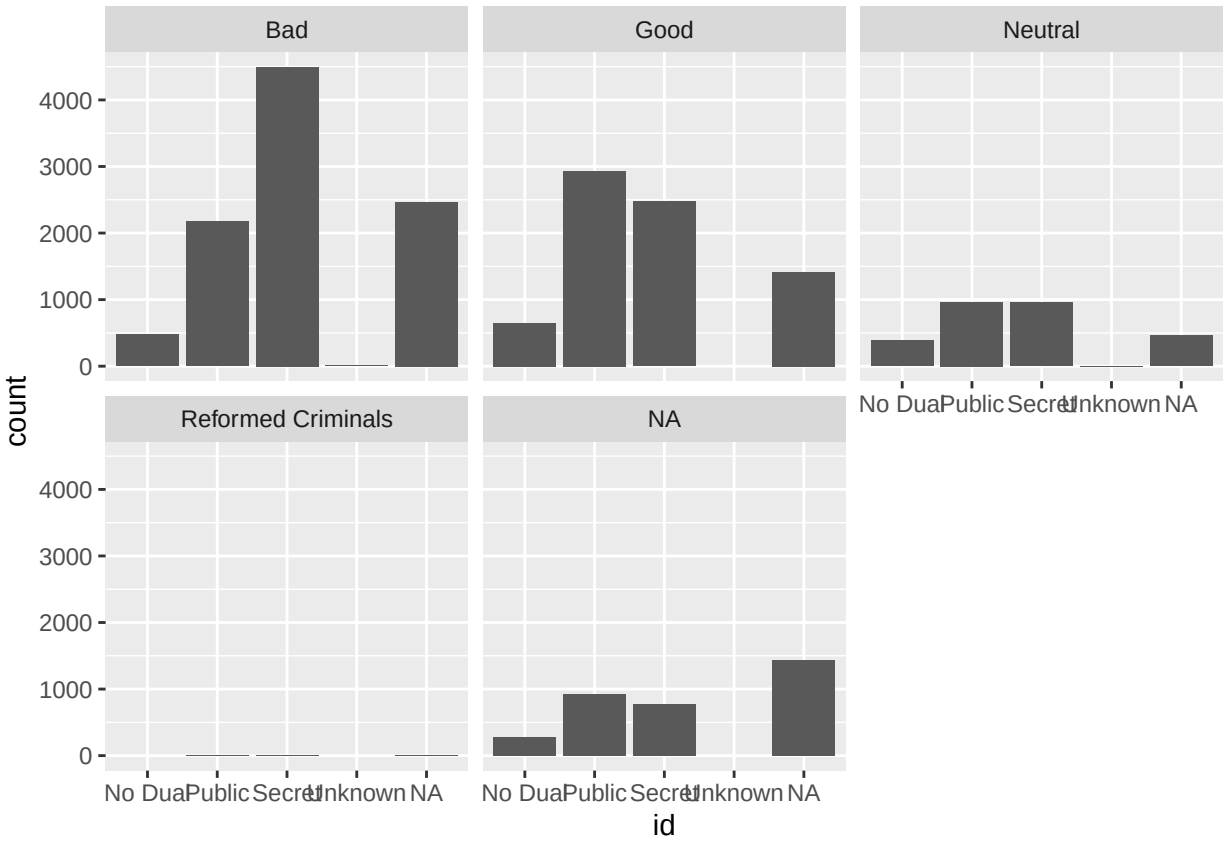
17. Plot the distribution of id variable

```
ggplot(r1, aes(x = id)) + geom_bar()
```



18. Using ggplot show the distribution of id variable, use facets to show the distribution for Bad, Good and Netural (Align Variable)

```
ggplot(r1, aes(x = id)) + geom_bar() + facet_wrap(~ align)
```



19. Realgin the levels of align to show the plot as Bad, Netural and Good instead of Bad, Good and Neutral

```
levels(r4)
```

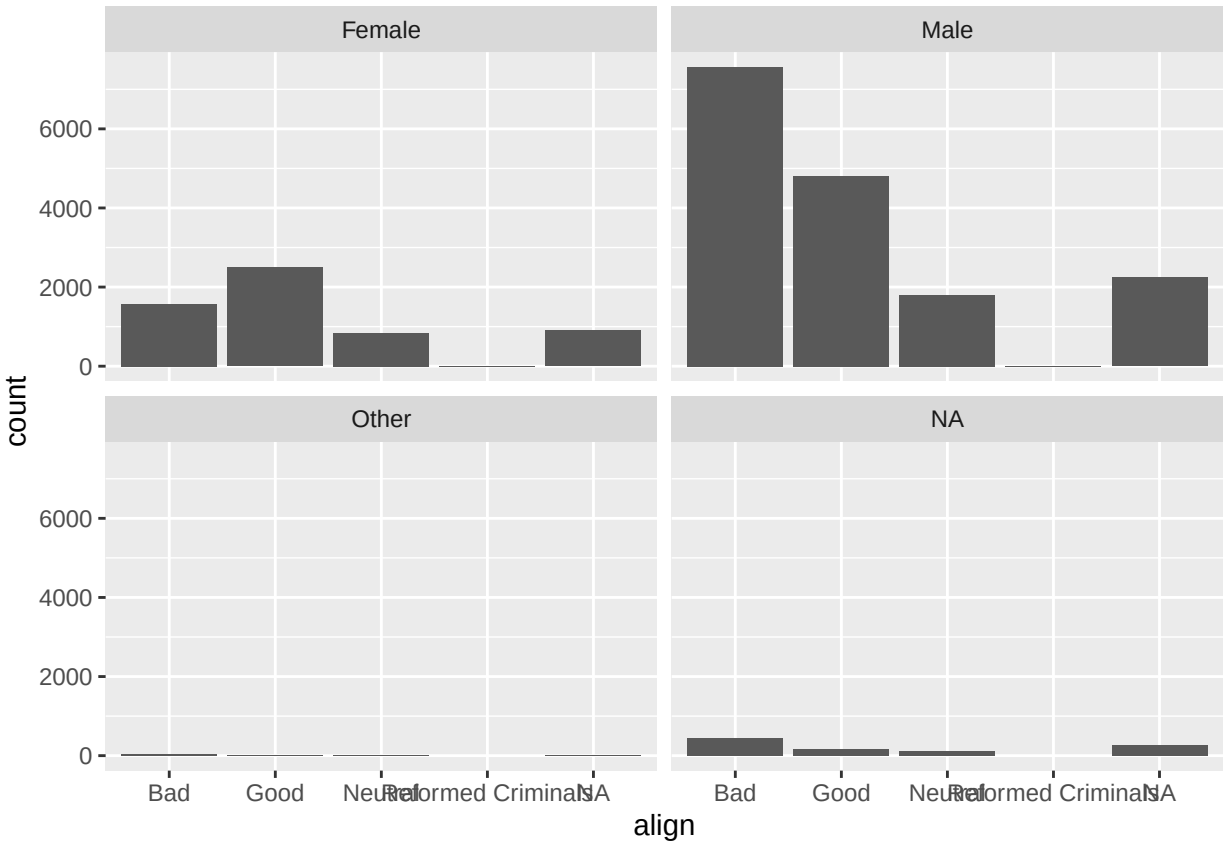
```
## [1] "Bad"     "Good"    "Neutral"
```

```
r9 <- factor(r1, levels = c("Bad", "Good", "Neutral"), labels = c("Bad", "Neutral", "Good"), exclude = N
head(r9)
```

```
##     X  name    id align   eye  hair
##  <NA>  <NA>  <NA>  <NA>  <NA>  <NA>
## Levels: Bad Neutral Good
```

20. Make a facet plot of alignment broken down by gender

```
ggplot(r1, aes(x = align)) + geom_bar() + facet_wrap(~ gender)
```


Female
Male
Other
NA
count
6000
4000
2000
0
Bad
Good
Neutral
Reformed Criminals
NA
align