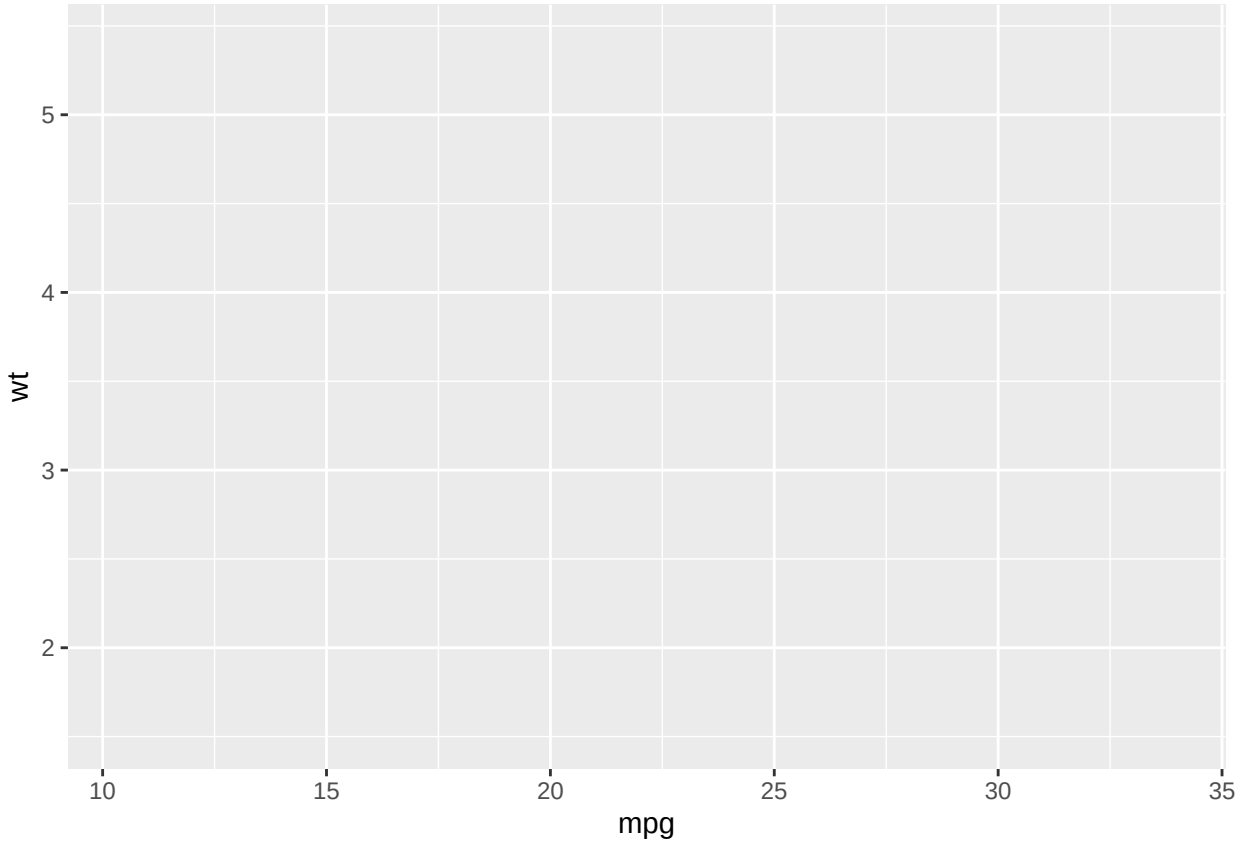
Workbook 3

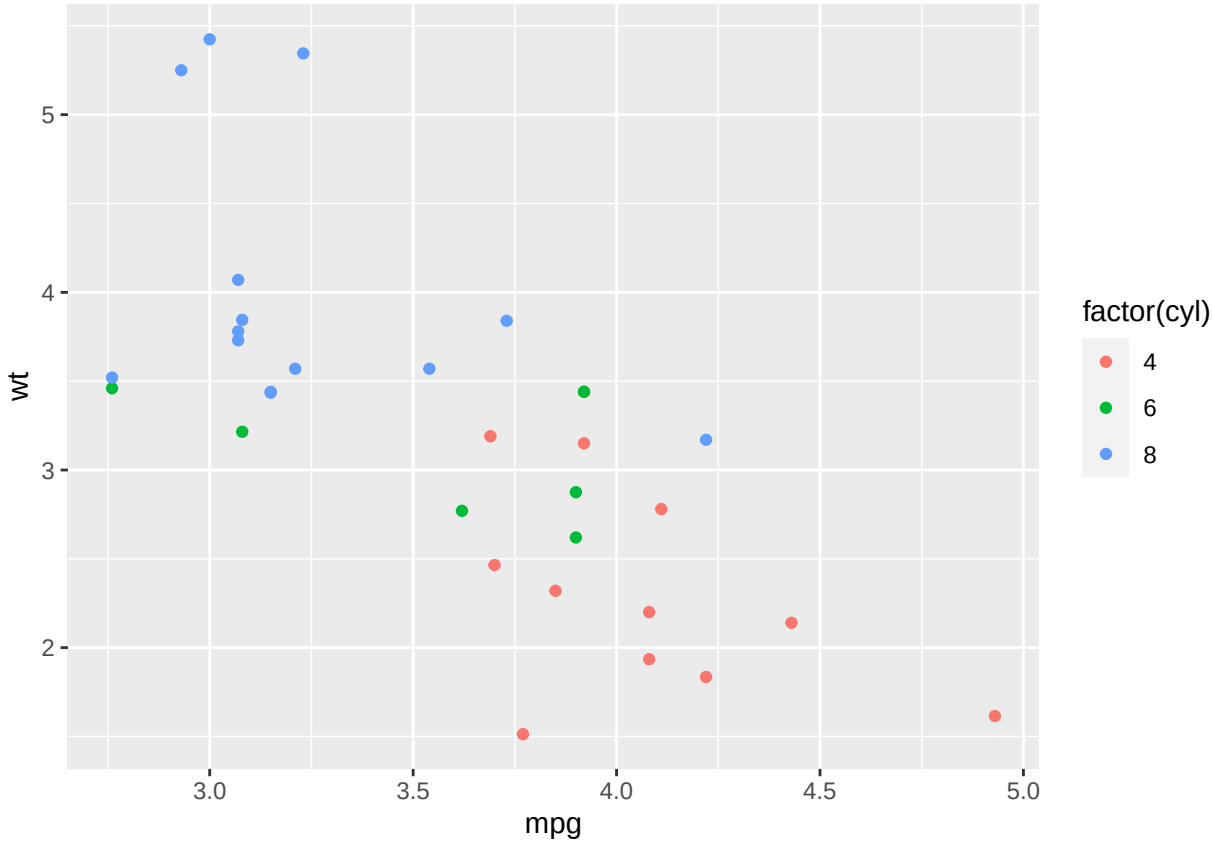
Emma

10/03/2021

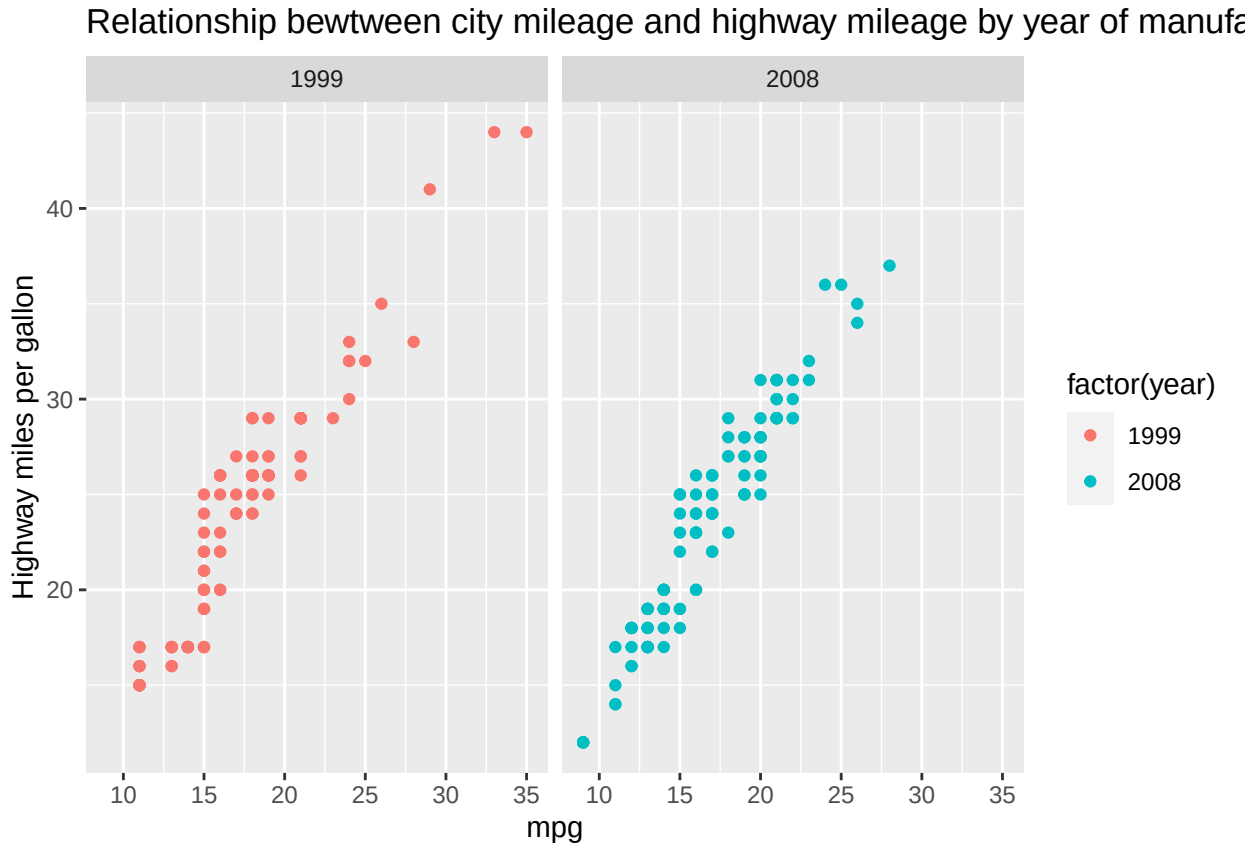
```
# Q1. error with the below equation
ggplot(data = mtcars) +
  aes(mpg, wt, colour=factor(cyl))
```



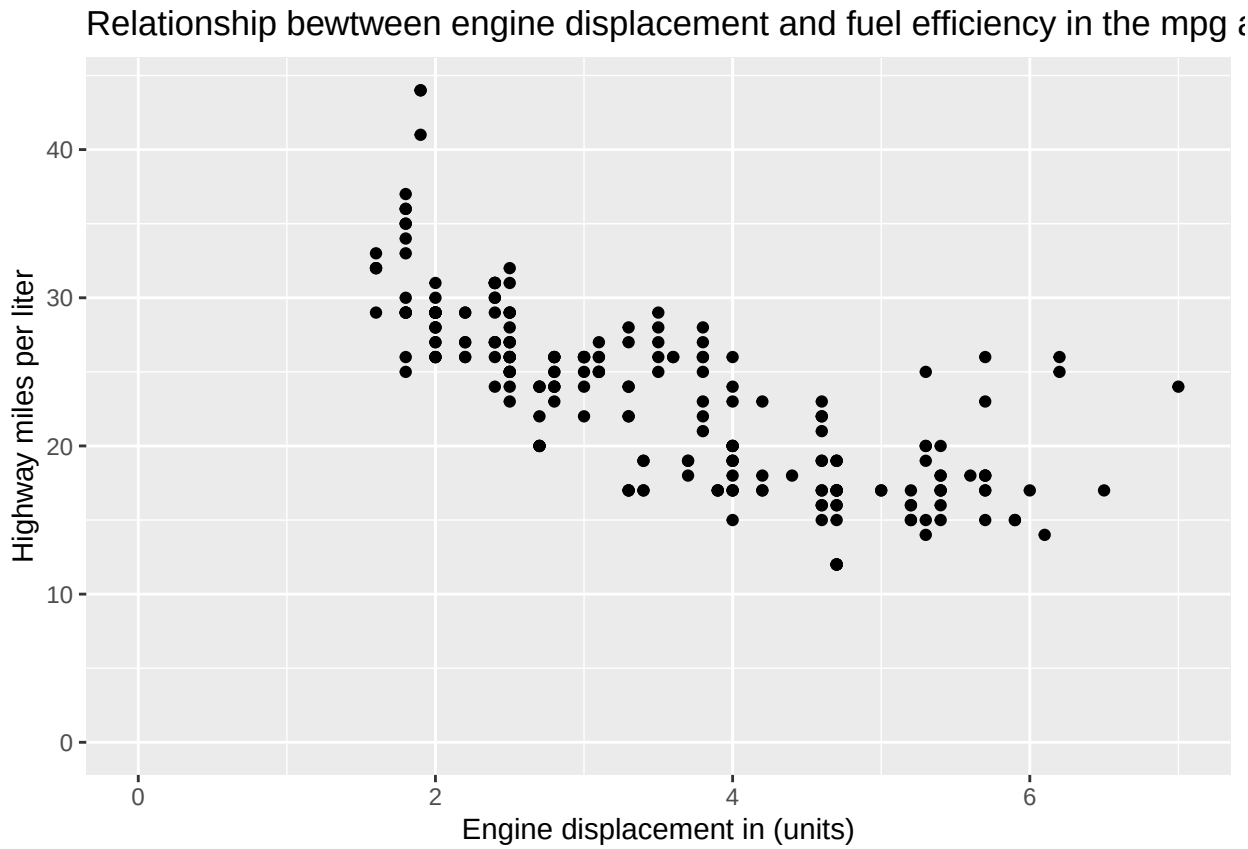

```
# the x and y axis are not named, so no data is pulled for the graph,
#an example of how it should look is below:

ggplot(data = mtcars) +
  geom_point(mapping = aes(x = drat, y = wt)) +
  aes(mpg, wt, colour=factor(cyl))
```




```
# Q2. Using the mpg dataset, graph the relationship between city milage (cty)
#and highway mileage (hwy) by year manufacture (year)

ggplot(data = mpg) +
  geom_point(mapping = aes(x = cty, y = hwy)) +
  labs(title = "Relationship bewtween city mileage and highway mileage by year of manufacture in the mpg
  xlab("City Mileage per gallon") +
  ylab("Highway miles per gallon") +
  facet_wrap(~ year) +
  aes(mpg, colour = factor(year))
```




```
# Q3. edit code so that both axis start at 0.

ggplot(data = mpg) +
  geom_point(mapping = aes(x = displ, y = hwy)) +
  labs(title = "Relationship bewtween engine displacement and fuel efficiency in the mpg automobile data
  xlab("Engine displacement in (units)") +
  ylab("Highway miles per liter") +
  expand_limits(x = 0, y = 0)
```



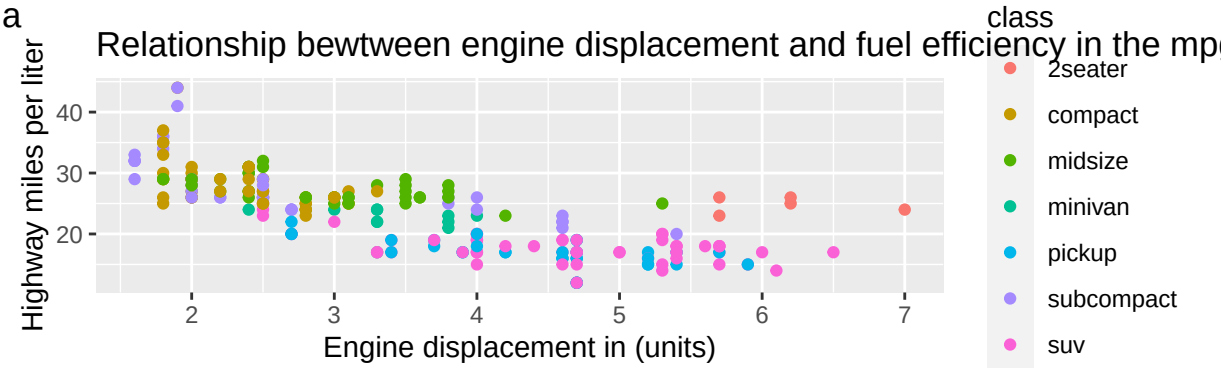
Q4. What what is one benefit and one limitation for this graph (in which the x and y values start at 0?)

Benefit - Data is easier to read as both axis start at 0 and it is easier to tell where individual values are. Limitation - The data is squished to a little to the top right corner, when it didn't start at 0, the data fit the graph better.

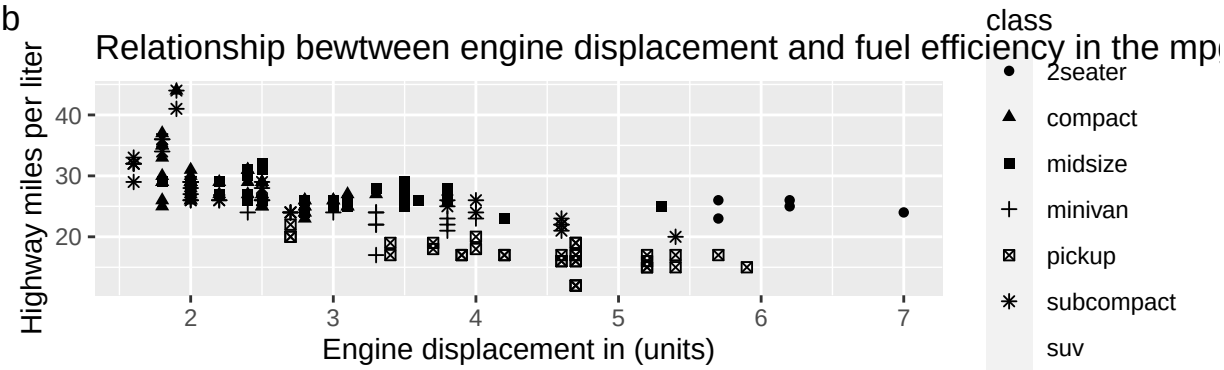
```
# Q5. Which of these two graphs do you prefer and why?

g1 <-ggplot(data = mpg) +
  geom_point(mapping = aes(x = displ, y = hwy, colour =  class )) +
  labs(title = "Relationship bewtween engine displacement and fuel efficiency in the mpg automobile data
  xlab("Engine displacement in (units)") +
  ylab("Highway miles per liter")

g2 <-ggplot(data = mpg) +
  geom_point(mapping = aes(x = displ, y = hwy, shape =  class )) +
  labs(title = "Relationship bewtween engine displacement and fuel efficiency in the mpg automobile data
  xlab("Engine displacement in (units)") +
  ylab("Highway miles per liter")

library("patchwork")

g1 / g2 + plot_annotation(title = "Which plot do you prefer and why?", tag_levels = 'a')
```

Which plot do you prefer and why?

a

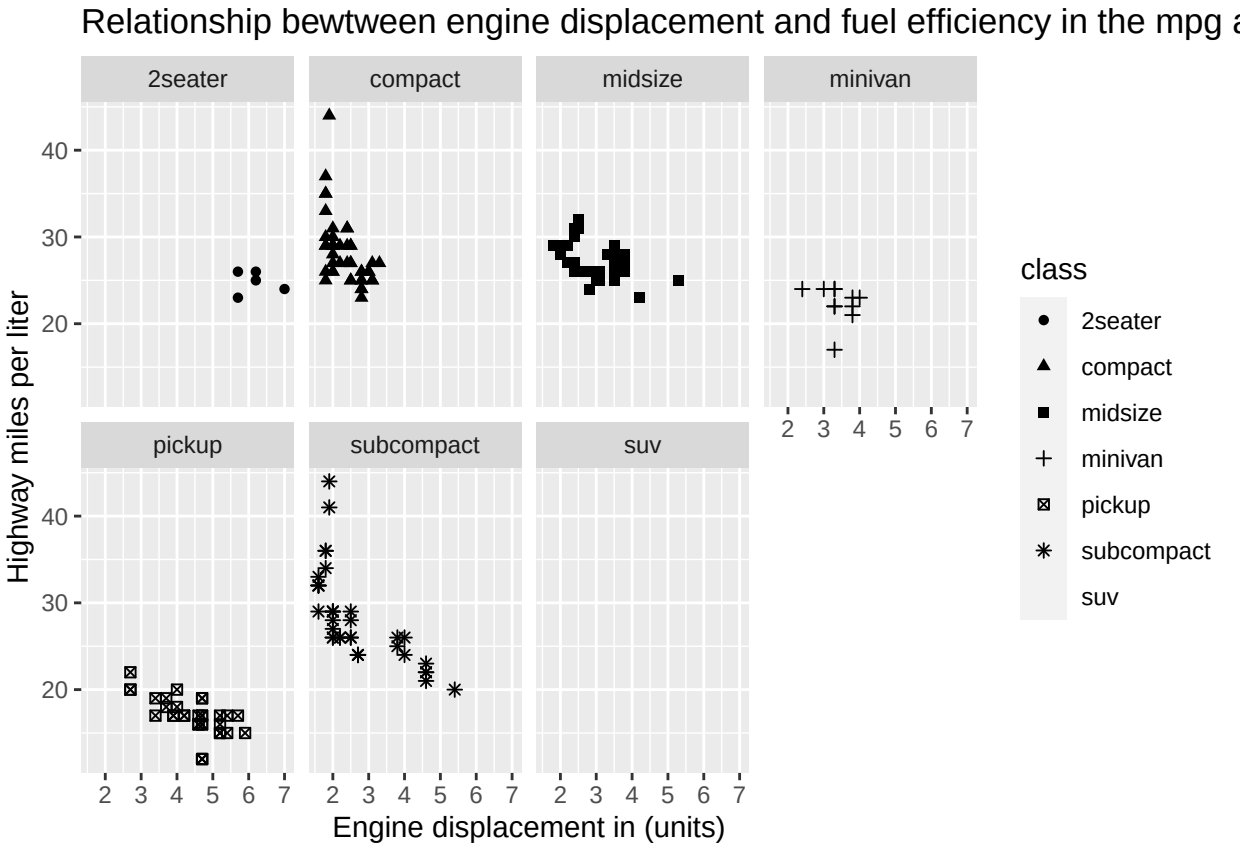


b



Answer - I prefer the graph with coloured dots rather than shapes because I find it easier to understand where each car class fits on the graph when glancing at it. Both have merits however, for example if anyone in the audience is colour-blind then the shape graph would be preferable.

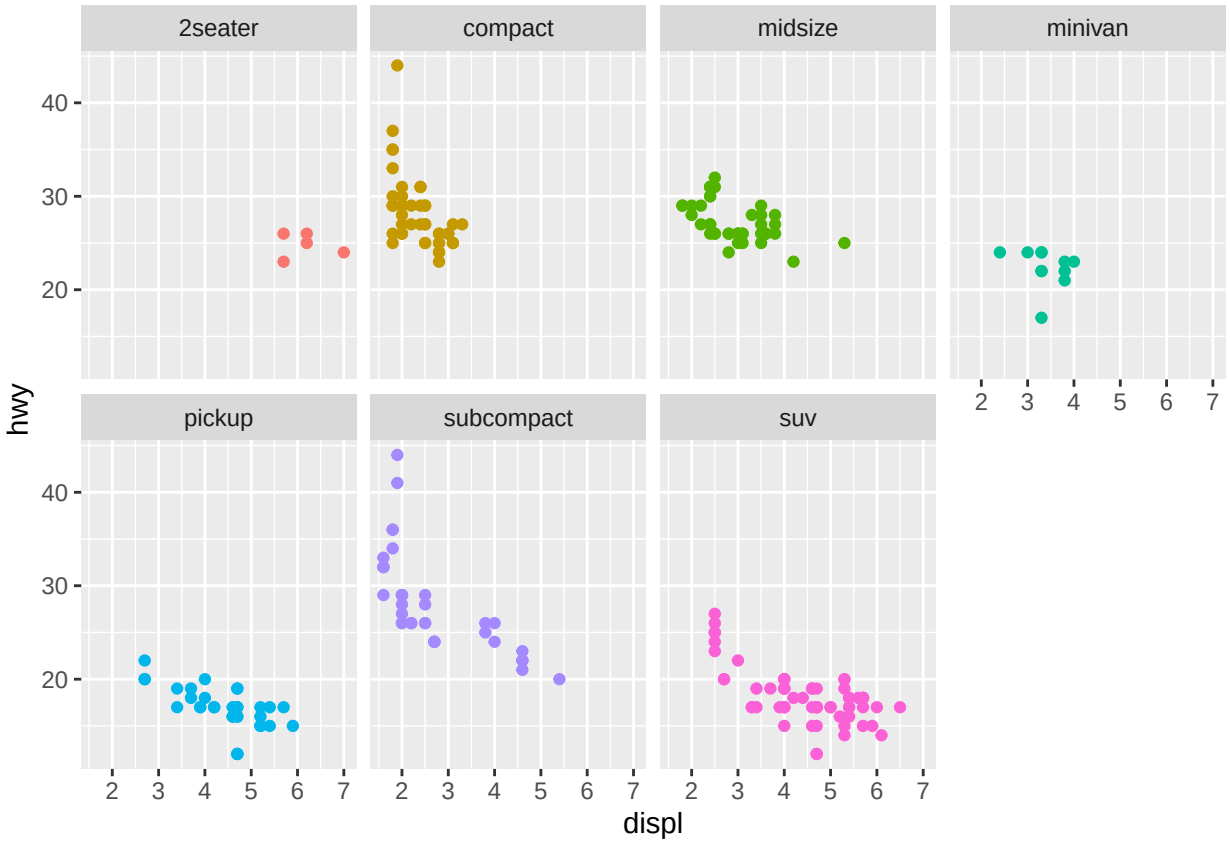
```
# Q6. add a facet to this graph for the "class" variable

ggplot(data = mpg) +
  geom_point(mapping = aes(x = displ, y = hwy, shape =  class )) +
  labs(title = "Relationship bewtween engine displacement and fuel efficiency in the mpg automobile data
  xlab("Engine displacement in (units)") +
  ylab("Highway miles per liter") +
  facet_wrap( ~ class, nrow = 2)
```




```
# Q7. Which graph is more informative and why?
```

Both graphs are informative in their own way depending on what you want to see in the data. The first graph shows the interaction between all classes, the second allows you to see the data at class level and the interactions between highway miles per gallon and engine displacement per car class.

```
# Q8. remove the legend from the facet graph

ggplot(data = mpg) +
  geom_point(mapping = aes(x = displ, y = hwy, color = class)) +
  facet_wrap( ~ class, nrow = 2)+
  theme(legend.position = "none")
```




```
#Q9. Code

issp <- readr::read_csv2(url("https://raw.githubusercontent.com/go-bayes/psych-447/main/data/issp.csv"))
```

```
## i Using ',' as decimal and '.' as grouping mark. Use `read_delim()` for more control.

##
## -- Column specification --------------------------------------------------------
## cols(
##   .default = col_character(),
##   id = col_double(),
##   age = col_double(),
##   eduyears = col_double(),
##   nzeuro = col_double(),
##   rightwing = col_double(),
##   thr_ath = col_double(),
##   wave = col_double(),
##   religiosity = col_double()
## )
## i Use `spec()` for the full column specifications.
```

```
head(issp)
```

```
## # A tibble: 6 x 21
```

```
##      id   age male   eduyears nzeuro rightwing neg_ath   neg_bd  neg_ch  neg_hd
##   <dbl> <dbl> <chr>     <dbl>  <dbl>     <dbl> <chr>     <chr>   <chr>   <chr>
## 1     1    52 Not M~       16      1         4 Neither ~ Somewh~ Somewh~ Somewh~
## 2     1    53 Not M~       13      1         3 Neither ~ Neithe~ Neithe~ Somewh~
## 3     2    63 Not M~       10      1         4 Neither ~ Neithe~ Somewh~ Neithe~
## 4     2    64 Not M~       12      1         1 Neither ~ Neithe~ Somewh~ Neithe~
## 5     3    64 Male         NA      1         8 Neither ~ Neithe~ Neithe~ Neithe~
## 6     3    65 Male         16      1         7 Neither ~ Neithe~ Neithe~ Neithe~
## # ... with 11 more variables: neg_jw <chr>, neg_ms <chr>, thr_ath <dbl>,
## #   thr_bd <chr>, thr_ch <chr>, thr_hd <chr>, thr_jw <chr>, rural <chr>,
## #   thr_ms <chr>, wave <dbl>, religiosity <dbl>
```

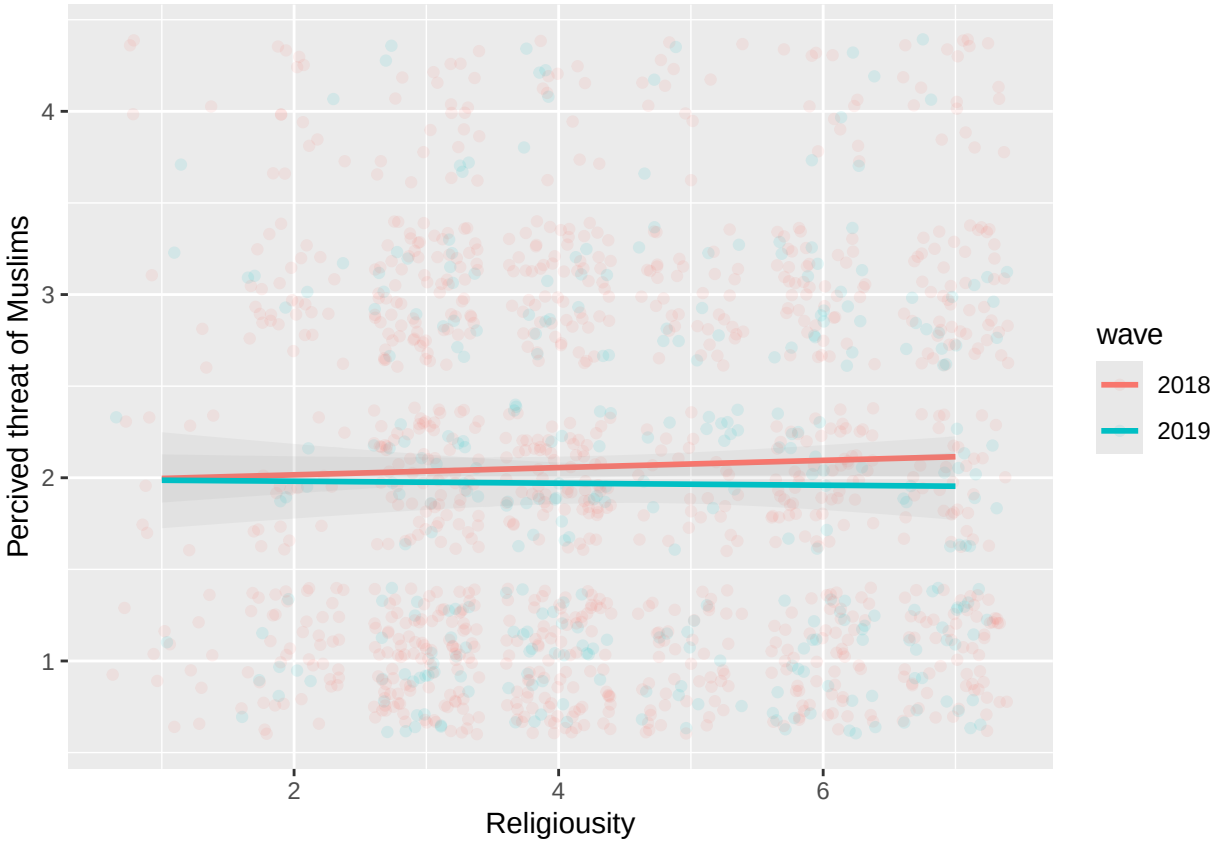
```
str(issp)
```

```
## spec_tbl_df [2,668 x 21] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
##  $ id         : num [1:2668] 1 1 2 2 3 3 4 4 5 5 ...
##  $ age        : num [1:2668] 52 53 63 64 64 65 29 30 41 42 ...
##  $ male       : chr [1:2668] "Not Male" "Not Male" "Not Male" "Not Male" ...
##  $ eduyears   : num [1:2668] 16 13 10 12 NA 16 11 13 11 14 ...
##  $ nzeuro     : num [1:2668] 1 1 1 1 1 1 NA 1 1 1 ...
##  $ rightwing  : num [1:2668] 4 3 4 1 8 7 8 8 6 7 ...
##  $ neg_ath    : chr [1:2668] "Neither negative nor positive" "Neither negative nor positive" "Neithe
##  $ neg_bd     : chr [1:2668] "Somewhat negative" "Neither negative nor positive" "Neither negative n
##  $ neg_ch     : chr [1:2668] "Somewhat negative" "Neither negative nor positive" "Somewhat negative"
##  $ neg_hd     : chr [1:2668] "Somewhat negative" "Somewhat negative" "Neither negative nor positive"
##  $ neg_jw     : chr [1:2668] "Somewhat negative" "Neither negative nor positive" "Neither negative n
##  $ neg_ms     : chr [1:2668] "Somewhat negative" "Somewhat negative" "Somewhat negative" "Somewhat n
##  $ thr_ath    : num [1:2668] 2 2 1 1 2 2 1 1 3 NA ...
##  $ thr_bd     : chr [1:2668] "Not very threatening" "Somewhat threatening" "Not threatening at all"
##  $ thr_ch     : chr [1:2668] "Not very threatening" "Somewhat threatening" "Somewhat threatening" "S
##  $ thr_hd     : chr [1:2668] "Not very threatening" "Somewhat threatening" "Not threatening at all"
##  $ thr_jw     : chr [1:2668] "Not very threatening" "Somewhat threatening" "Not threatening at all"
##  $ rural      : chr [1:2668] "Not Rural" "Not Rural" "Not Rural" "Not Rural" ...
##  $ thr_ms     : chr [1:2668] "Not very threatening" "Somewhat threatening" "Somewhat threatening" "N
##  $ wave       : num [1:2668] 2018 2019 2018 2019 2018 ...
##  $ religiosity: num [1:2668] 6 4 3 4 4 6 7 7 5 5 ...
##  - attr(*, "spec")=
##   .. cols(
##   ..   id = col_double(),
##   ..   age = col_double(),
##   ..   male = col_character(),
##   ..   eduyears = col_double(),
##   ..   nzeuro = col_double(),
##   ..   rightwing = col_double(),
##   ..   neg_ath = col_character(),
##   ..   neg_bd = col_character(),
##   ..   neg_ch = col_character(),
##   ..   neg_hd = col_character(),
##   ..   neg_jw = col_character(),
##   ..   neg_ms = col_character(),
##   ..   thr_ath = col_double(),
##   ..   thr_bd = col_character(),
##   ..   thr_ch = col_character(),
##   ..   thr_hd = col_character(),
```

```
##   ..   thr_jw = col_character(),
##   ..   rural = col_character(),
##   ..   thr_ms = col_character(),
##   ..   wave = col_double(),
##   ..   religiosity = col_double()
##   .. )
```

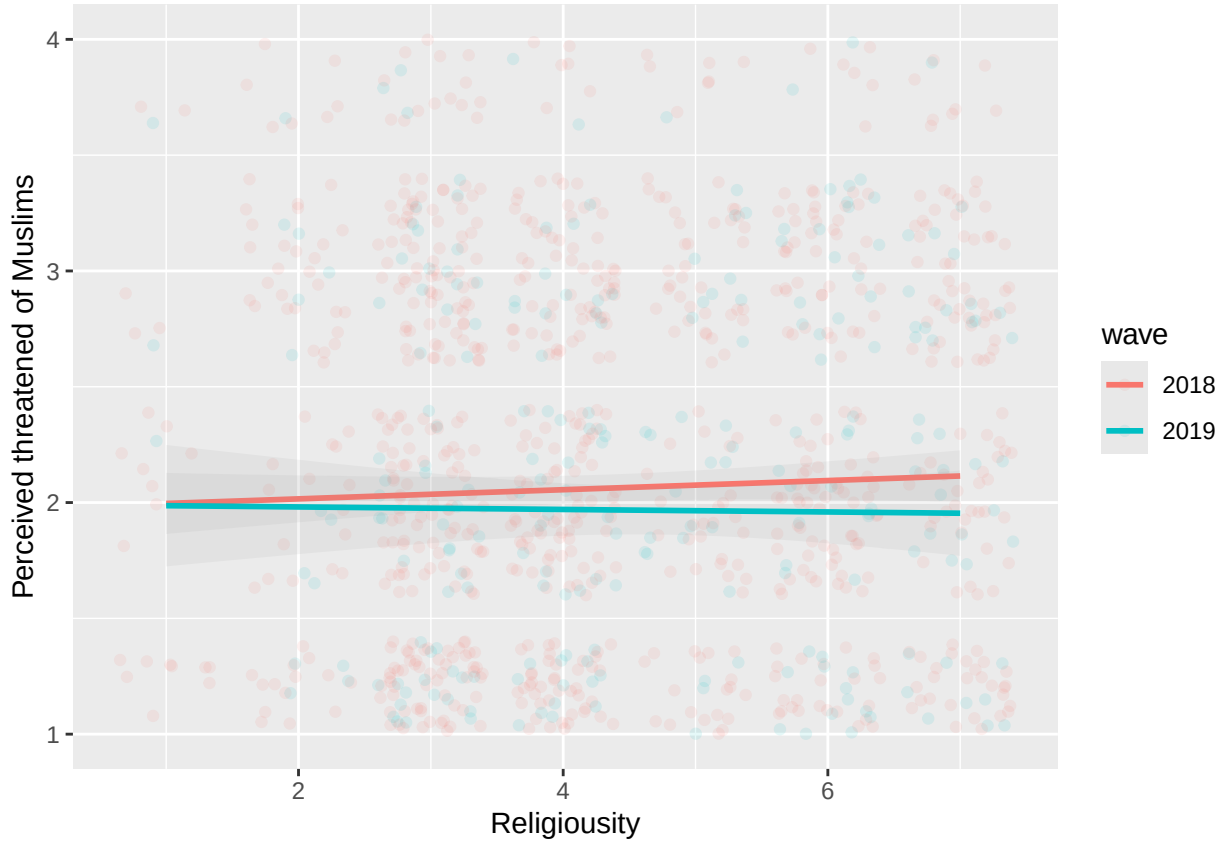
```
ip <- issp %>%
  mutate(
    id = factor(id),
    thr_ath = as.factor(thr_ath),
    thr_bd = as.factor(thr_bd),
    thr_ch = as.factor(thr_ch),
    thr_hd = as.factor(thr_hd),
    thr_jw = as.factor(thr_jw),
    thr_ms = as.factor(thr_ms),
    neg_ath = as.factor(neg_ath),
    neg_bd = as.factor(neg_bd),
    neg_ch = as.factor(neg_ch),
    neg_hd  = as.factor(neg_hd),
    neg_jw = as.factor(neg_jw),
    neg_ms = as.factor(neg_ms),
    wave  = as.factor(wave),
    nzeuro = as.factor(nzeuro),
    eduyears = as.numeric(eduyears),
    male = as.factor(male),
    age = as.numeric(age),
    rightwing = as.numeric(rightwing),
    rural = as.factor(rural),
    religiosity = as.numeric(religiosity)
  )

# Q9. Convert the y variable to "numeric" and graph the relationship betweeen
# religiousity (x-axis) and `thr_mus`` (y-axis) in the
# ISSP dataset. Create new axis labels

library(ggplot2)
ggplot(data = ip, aes(y = as.numeric(thr_ms), x = religiosity, colour = wave))  +
  geom_jitter(alpha = .1) +
  geom_smooth(method = lm, fullrange = FALSE, alpha = 0.1) +
  xlab("Religiousity") +
  ylab("Percived threat of Muslims") +
   scale_y_continuous()
```

```
## `geom_smooth()` using formula 'y ~ x'
```

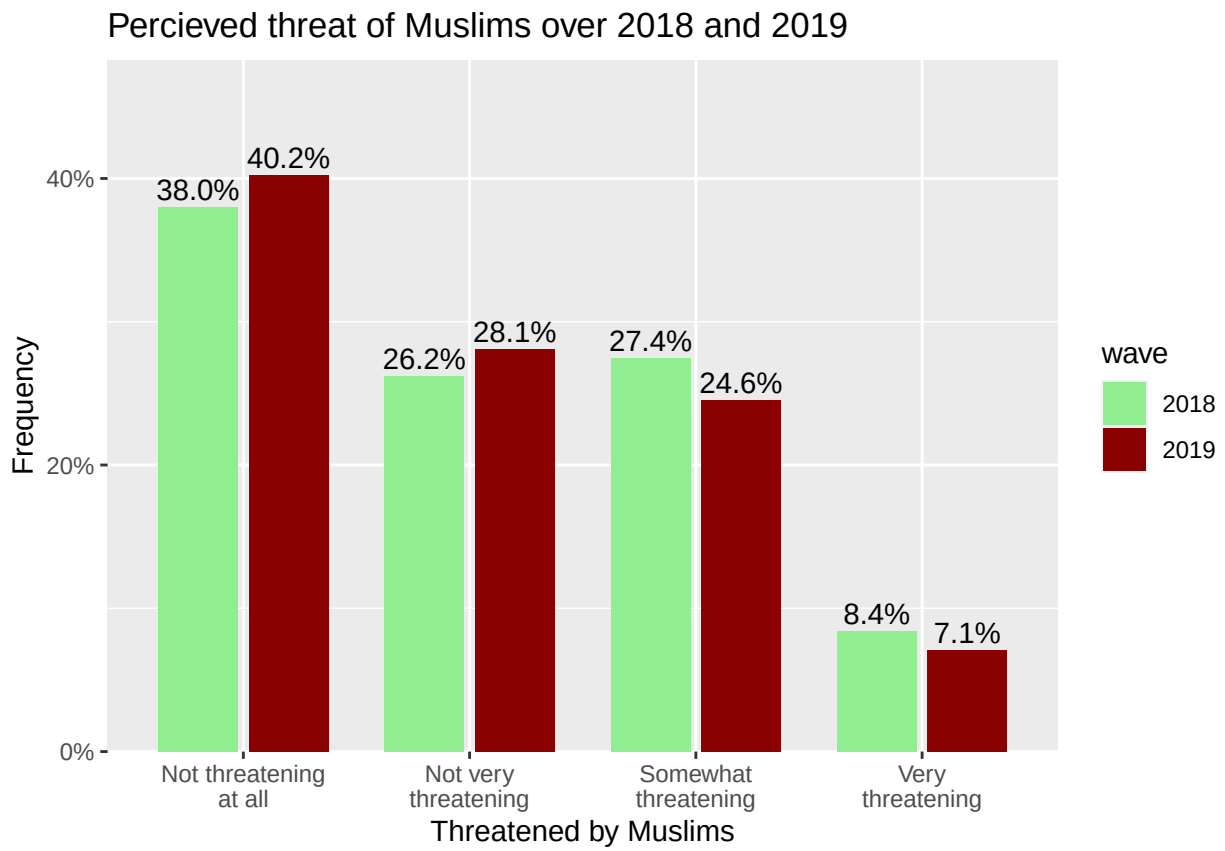



```
# Q10. Fix the graph to run from 1 to 4.

library(ggplot2)
ggplot(data = ip, aes(y = as.numeric(thr_ms), x = religiosity, colour = wave))  +  geom_jitter(alpha =
  geom_smooth(method = lm, fullrange = FALSE, alpha = 0.1) +
  xlab("Religiousity") +
  ylab("Perceived threatened of Muslims") +
   scale_y_continuous(limits = c(1,4))
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
#Q11. improve the code.

#Explanation: Added a title to show what the graph was looking at.

library(sjPlot)
plot_xtab(
    ip$thr_ms,
    ip$wave,
    show.total = F,
    show.n = F,
    geom.colors = c("lightgreen", "darkred")
  ) +
  xlab("Threatened by Muslims") +  ylab("Frequency") +
  labs(title = "Percieved threat of Muslims over 2018 and 2019")
```




```
  theme(axis.text.x = element_text(angle = 20, hjust = 1))
```

```
## List of 1
##  $ axis.text.x:List of 11
##   ..$ family       : NULL
##   ..$ face         : NULL
##   ..$ colour       : NULL
##   ..$ size         : NULL
##   ..$ hjust        : num 1
##   ..$ vjust        : NULL
##   ..$ angle        : num 20
##   ..$ lineheight   : NULL
##   ..$ margin       : NULL
##   ..$ debug        : NULL
##   ..$ inherit.blank: logi FALSE
##   ..- attr(*, "class")= chr [1:2] "element_text" "element"
##  - attr(*, "class")= chr [1:2] "theme" "gg"
##  - attr(*, "complete")= logi FALSE
##  - attr(*, "validate")= logi TRUE
```