Lab 12 - Statistics, Coordinates, Facets, and Themes

Ramin Jabbarialghanab

November 14, 2017

Complete the following exercises below. Knit together the PDF document and commit both the Lab 12 RMDfile and the PDF document to Git. Push the changes to GitHub so both documents are visible in your public GitHub repository.

1. Choose one or more graphics you created for Lab 11 and either experiment with the underlying statistical layer if it already has one (i.e. if you made a histogram experiement with different bin widths) or add a separate statistical layer to your plot (i.e. a smooothing curve). Choose something you think will offer meaningful insight and describe why you made the choice you did. What additional information does this provide viewers of your graphic?

I used geom point to show the exact place of the value for each year in the scale. Then in order to know the general inclination of unemployment rate for both gender, I added geom smooth with se= FALSE.I coded se= FALSE because in my data it is not meaninful to use it. If I had a spreaded data, I could use se = TRUE to inclue the values. Here it is not the case. In the second table, I used only geom_smooth se = FALSE function to show general tendency without showing the exact value for the years.

```r
# funtion geom point with smooth se= FALSE

install.packages("ggplot2")
```

```
## Installing package into '/Users/Ramin/Desktop/Autumn 2017/Statistics 321/unemployment rate/unemployme
## (as 'lib' is unspecified)
```

```
##
## The downloaded binary packages are in
##   /var/folders/z9/b9hh4hpj6hl6x9r9dxjctv240000gn/T//RtmpsbS4CZ/downloaded_packages
```

```r
library(ggplot2)
library(readxl)
joint_unemployment_total <- read_excel("~/Desktop/Autumn 2017/Statistics 321/unemployment rate/joint_une
ggplot(joint_unemployment_total,
       aes(x = year,
           y = value,
           col = gender)) +
  geom_point(size = 3) +
  scale_colour_hue(h = c(0, 90)) +
  geom_smooth(se = FALSE)
```
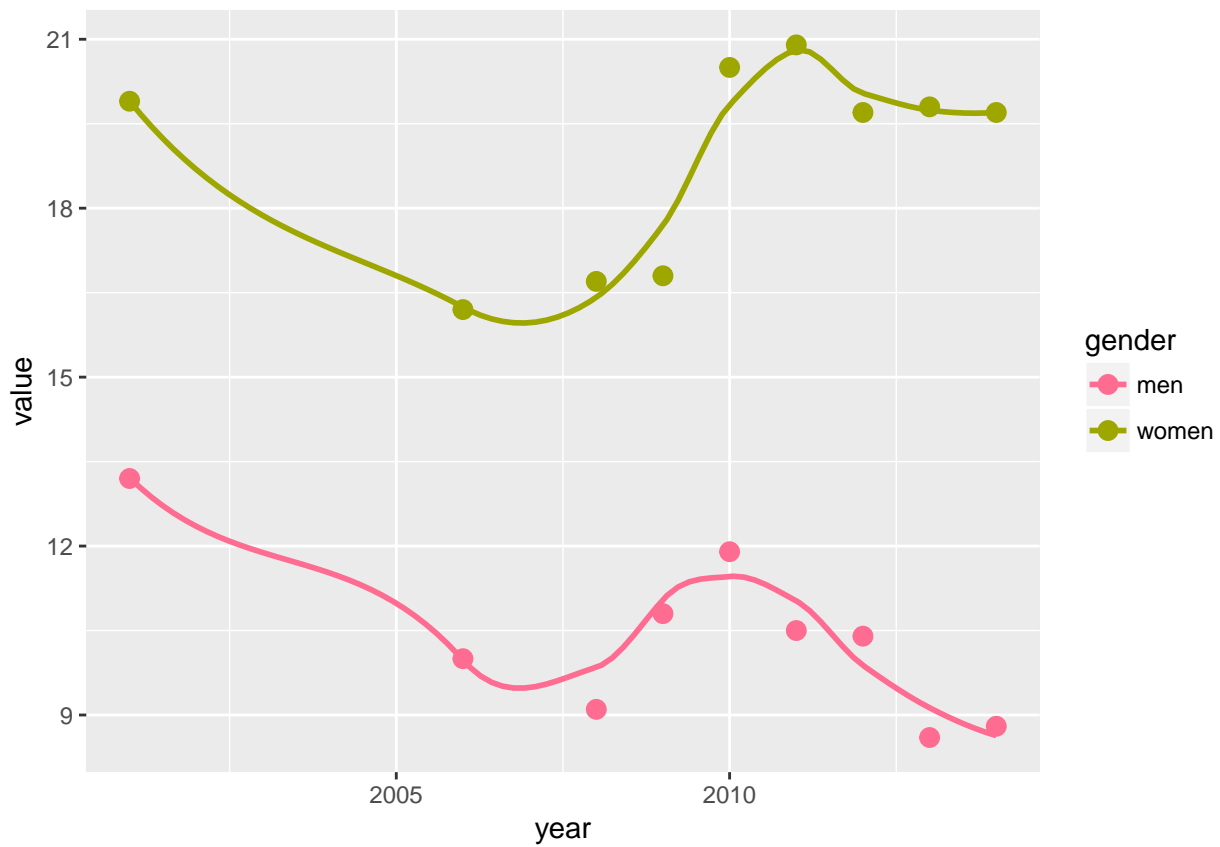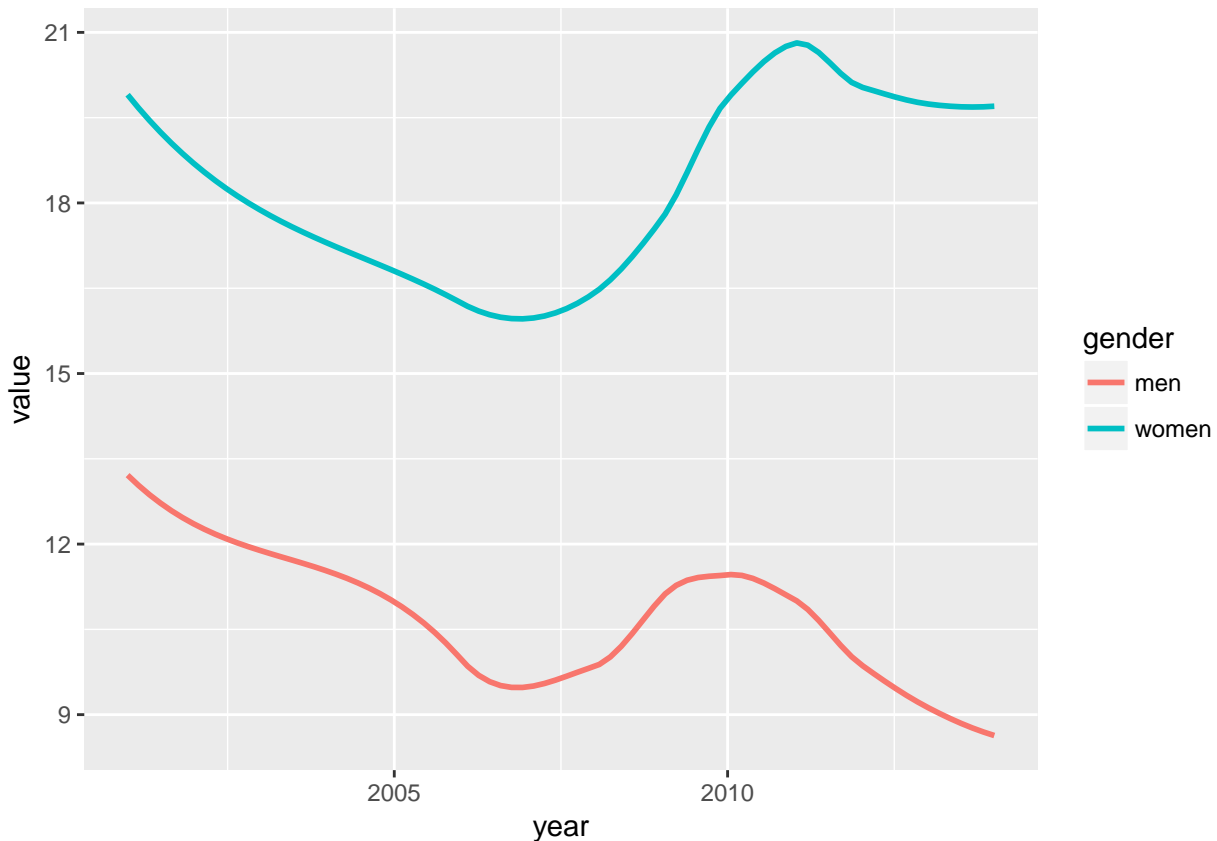
```
## `geom_smooth()` using method = 'loess'
```

```
ggplot(joint_unemployment_total,
       aes(x = year,
           y = value,
           color = gender)) +
  geom_smooth(se = FALSE)
```

```
## `geom_smooth()` using method = 'loess'
```

2. With the same or a different plot created in Lab 11, experiment with zooming in on specific areas of your graphic and changing the aspect ratio. Are their any benefits/drawbacks with either or both of these approaches for the visualizations you've created? What are they?

Actually, for my data the aspect ratio did not change my tables significantly. Although in other case could be more useful, for my tables I can seen big difference. But zooming is helpful to look to the value differences in the chosen years. That seems that applying both zooming and aspect ratio at the same time could make hard to read the table easily and found what would convey the data.

```
# Zooming on x between 2010 to 2014
ggplot(joint_unemployment_total,
      aes(x = year,
          y = value,
          col = gender)) +
  geom_point(size = 3) +
  scale_colour_hue(h = c(0, 90)) +
  geom_smooth() +
  scale_x_continuous(limits = c(2010, 2014))
```

```
## `geom_smooth()` using method = 'loess'

## Warning: Removed 8 rows containing non-finite values (stat_smooth).

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 2010

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
```

```
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : span too small.
## fewer data values than degrees of freedom.

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : pseudoinverse used
## at 2010

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : neighborhood radius
## 2.02

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : reciprocal
## condition number 0

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : There are other
## near singularities as well. 4.0804

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 2010

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : span too small.
## fewer data values than degrees of freedom.

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : pseudoinverse used
## at 2010

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : neighborhood radius
```
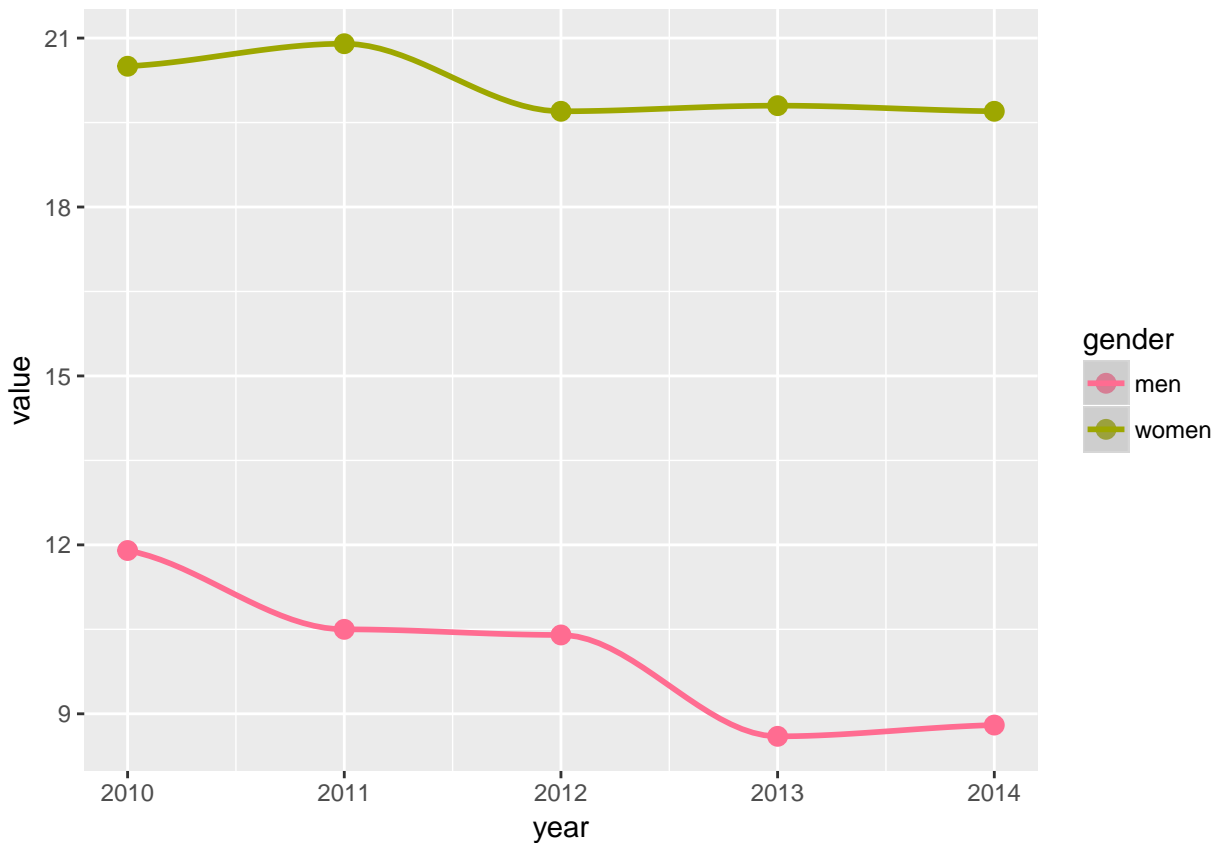
```
## 2.02
```

```
## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : reciprocal
## condition number 0
```

```
## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : There are other
## near singularities as well. 4.0804
```

```
## Warning: Removed 8 rows containing missing values (geom_point).
```



```r
#Changing aspect ratio
ggplot(joint_unemployment_total,
       aes(x = year,
           y = value,
           col = gender)) +
  geom_point(size = 3) +
  scale_colour_hue(h = c(0, 90)) +
  geom_smooth() +
  coord_equal()
```

```
## `geom_smooth()` using method = 'loess'
```
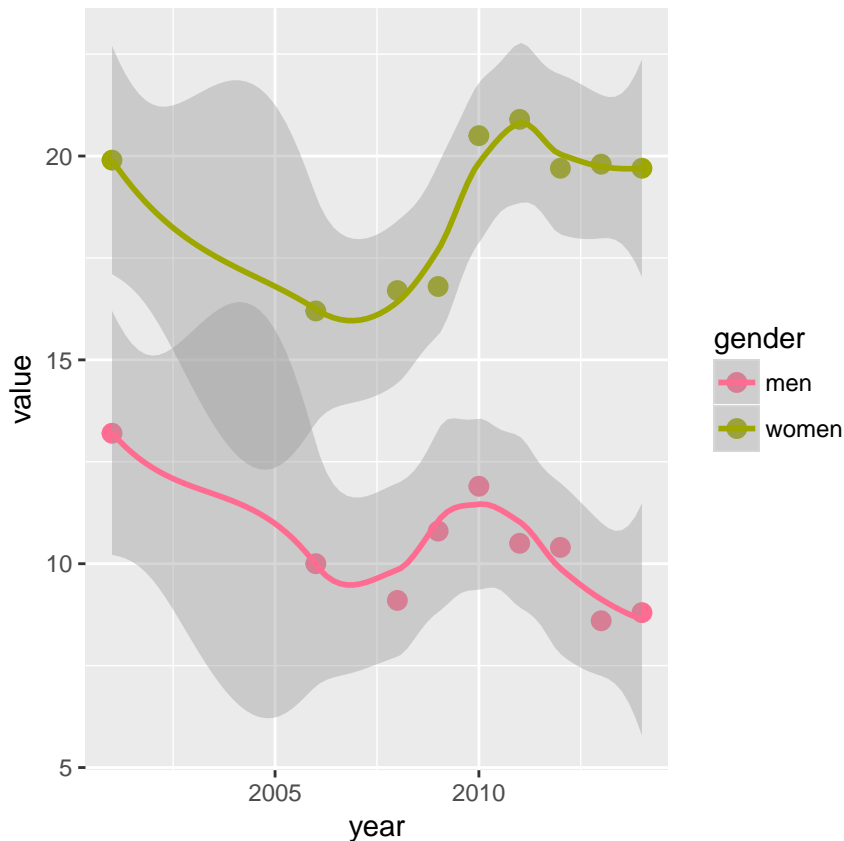
```
#Changing aspect ratio and zooming
ggplot(joint_unemployment_total,
       aes(x = year,
           y = value,
           col = gender)) +
  geom_point(size = 3) +
  scale_colour_hue(h = c(0, 90)) +
  geom_smooth() +
  scale_x_continuous(limits = c(2010, 2014)) +
 coord_equal()
```

`geom_smooth()` using method = 'loess'

Warning: Removed 8 rows containing non-finite values (stat_smooth).

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : span too small. fewer data values than degrees of freedom.

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : pseudoinverse used at 2010

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : neighborhood radius 2.02

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : reciprocal condition number 0

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : There are other near singularities as well. 4.0804

Warning in predLoess(object$y, object$x, newx = if

```
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : span too small.
## fewer data values than degrees of freedom.

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : pseudoinverse used
## at 2010

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : neighborhood radius
## 2.02

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : reciprocal
## condition number 0

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : There are other
## near singularities as well. 4.0804

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 2010

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : span too small.
## fewer data values than degrees of freedom.

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : pseudoinverse used
## at 2010

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : neighborhood radius
## 2.02

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : reciprocal
## condition number 0

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
```
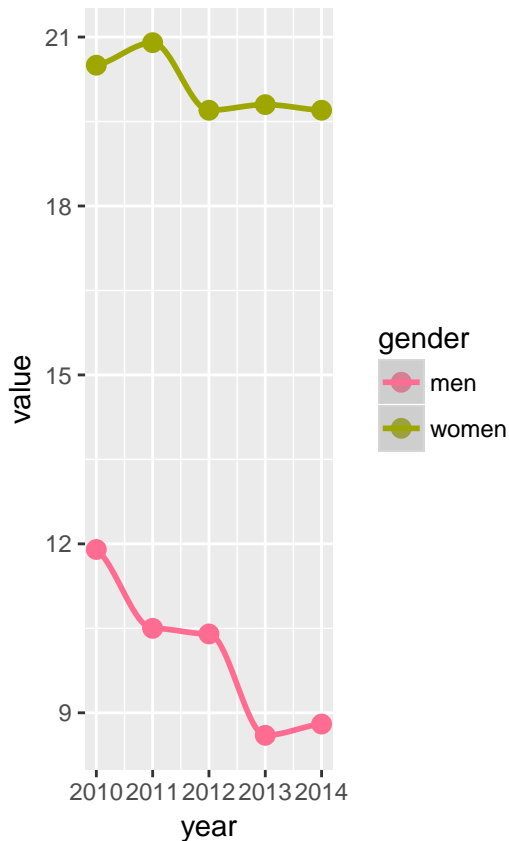
```
## as.matrix(model.frame(delete.response(terms(object)), : There are other
## near singularities as well. 4.0804
```

```
## Warning: Removed 8 rows containing missing values (geom_point).
```



```
# zooming on x 2010 to 2014
ggplot(joint_unemployment_total,
       aes(x = year,
           y = value,
           color = gender)) +
  geom_smooth() +
scale_x_continuous(limits = c(2010, 2014))
```

```
## `geom_smooth()` using method = 'loess'
```

```
## Warning: Removed 8 rows containing non-finite values (stat_smooth).
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 2010
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0
```

```
## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804
```

```
## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : span too small.
## fewer data values than degrees of freedom.

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : pseudoinverse used
## at 2010

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : neighborhood radius
## 2.02

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : reciprocal
## condition number 0

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : There are other
## near singularities as well. 4.0804

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 2010

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : span too small.
## fewer data values than degrees of freedom.

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : pseudoinverse used
## at 2010

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : neighborhood radius
## 2.02

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : reciprocal
## condition number 0

## Warning in predLoess(object$y, object$x, newx = if
```
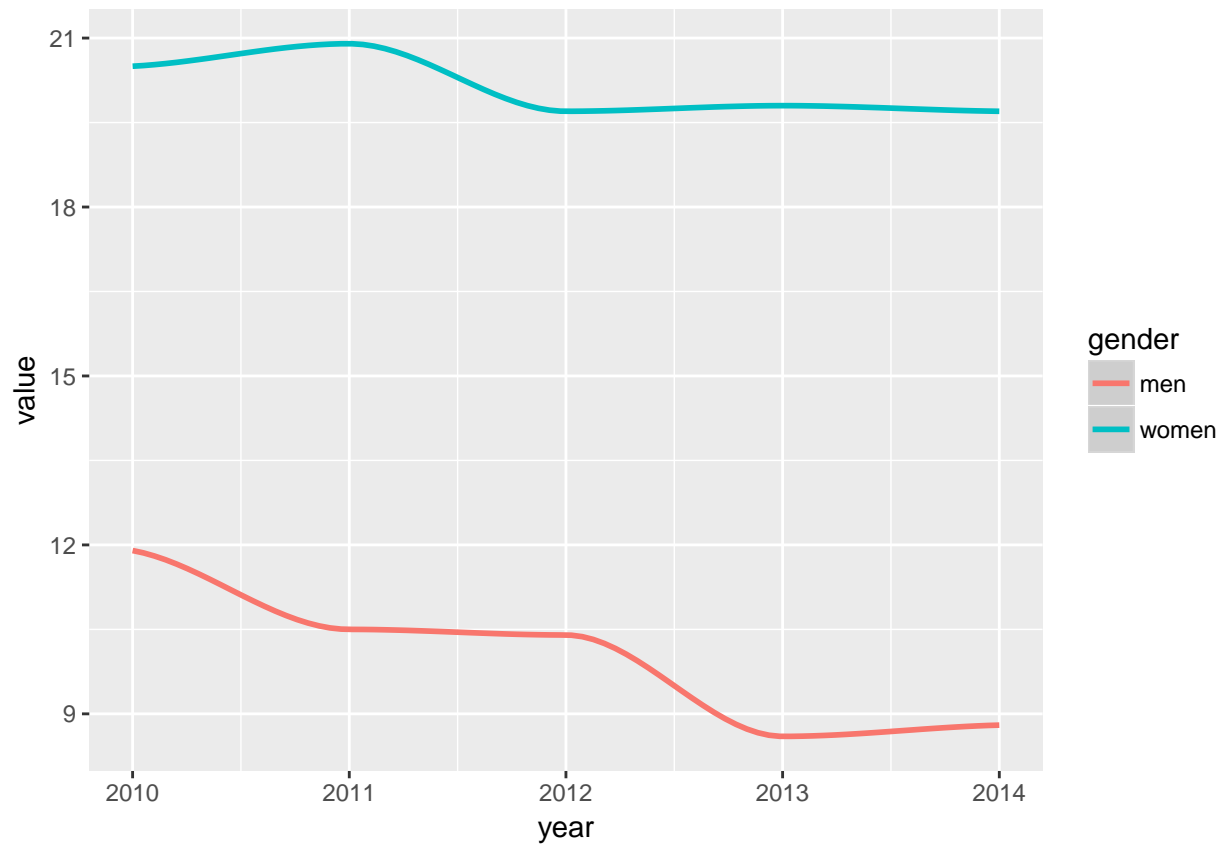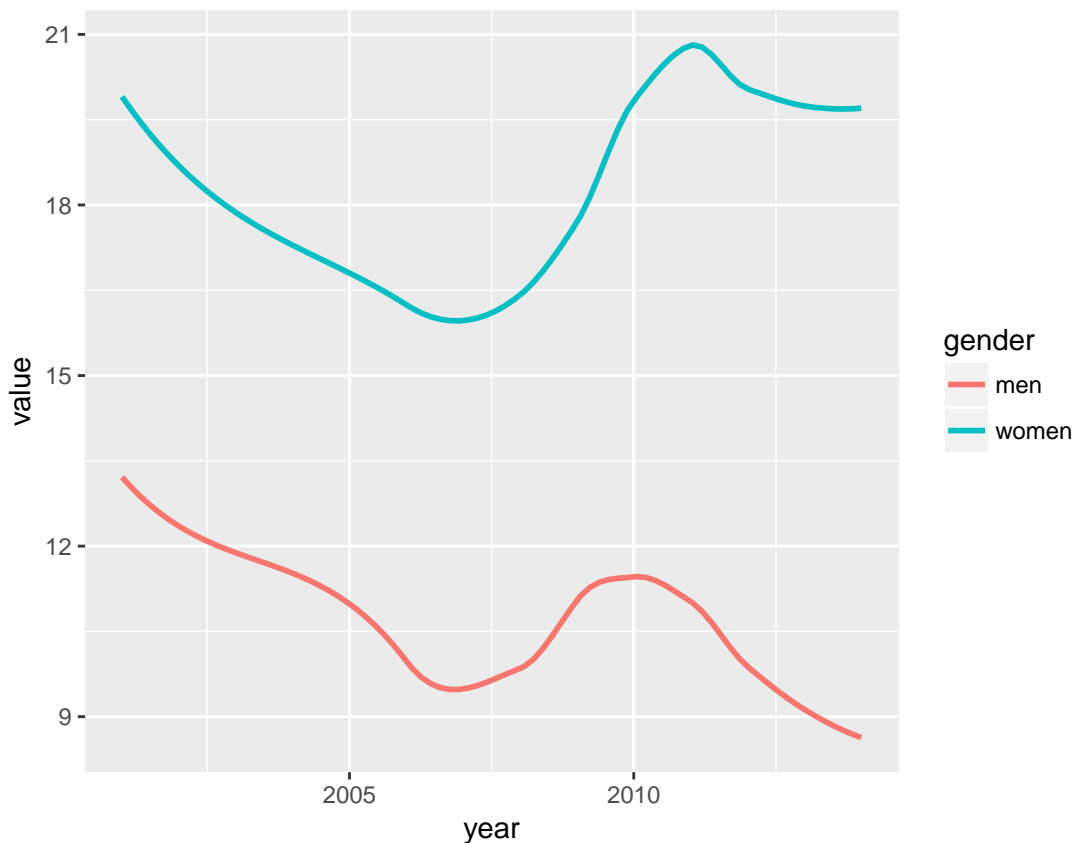
```
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : There are other
## near singularities as well. 4.0804
```



```r
# changin aspect ratio

ggplot(joint_unemployment_total,
       aes(x = year,
           y = value,
           color = gender)) +
  geom_smooth(se = FALSE) +
  coord_equal()
```

```
## `geom_smooth()` using method = 'loess'
```

```
# zooming on x 2010 to 2014 and changing aspect ratio
ggplot(joint_unemployment_total,
       aes(x = year,
           y = value,
           color = gender)) +
  geom_smooth() +
  scale_x_continuous(limits = c(2010, 2014)) +
  coord_equal()
```

`geom_smooth()` using method = 'loess'

Warning: Removed 8 rows containing non-finite values (stat_smooth).

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : span too small. fewer data values than degrees of freedom.

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : pseudoinverse used at 2010

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : neighborhood radius 2.02

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : reciprocal condition number 0

Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
parametric, : There are other near singularities as well. 4.0804

Warning in predLoess(object$y, object$x, newx = if
(is.null(newdata)) object$x else if (is.data.frame(newdata))
as.matrix(model.frame(delete.response(terms(object)), : span too small.

```
## fewer data values than degrees of freedom.

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : pseudoinverse used
## at 2010

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : neighborhood radius
## 2.02

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : reciprocal
## condition number 0

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : There are other
## near singularities as well. 4.0804

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : span too small. fewer data values than degrees of freedom.

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : pseudoinverse used at 2010

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : neighborhood radius 2.02

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : reciprocal condition number 0

## Warning in simpleLoess(y, x, w, span, degree = degree, parametric =
## parametric, : There are other near singularities as well. 4.0804

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : span too small.
## fewer data values than degrees of freedom.

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : pseudoinverse used
## at 2010

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : neighborhood radius
## 2.02

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : reciprocal
## condition number 0

## Warning in predLoess(object$y, object$x, newx = if
## (is.null(newdata)) object$x else if (is.data.frame(newdata))
## as.matrix(model.frame(delete.response(terms(object)), : There are other
## near singularities as well. 4.0804
```
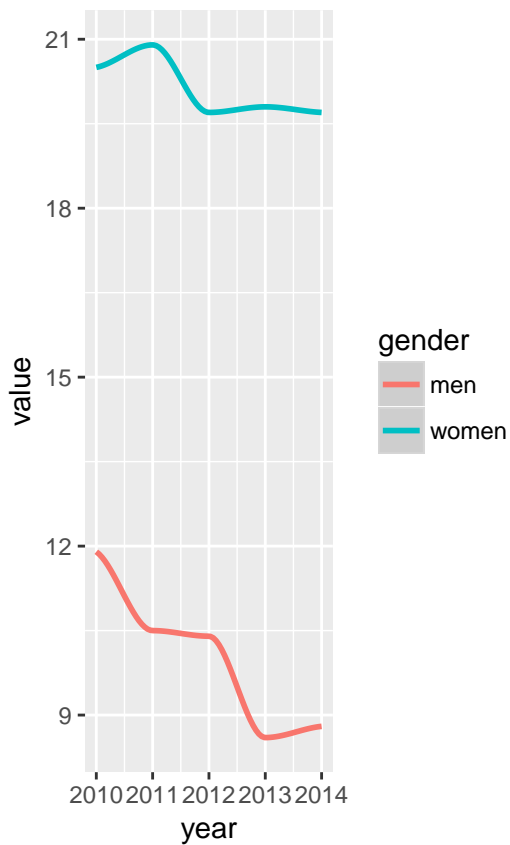
3. Try facetting a plot you have made by another categorical variable in your data (this can even be as simple as Male/Female). What is the difference between `facet_wrap()` and `facet_grid()`? How might facetting be useful in data visualization?

In my data there is no different between facet_wrap and face_grid.

```r
# facetting with facet grid
ggplot(joint_unemployment_total,
       aes(x = year,
           y = value,
           col = gender)) +
  geom_point(size = 3) +
  scale_colour_hue(h = c(0, 90)) +
  facet_grid(~gender)
```
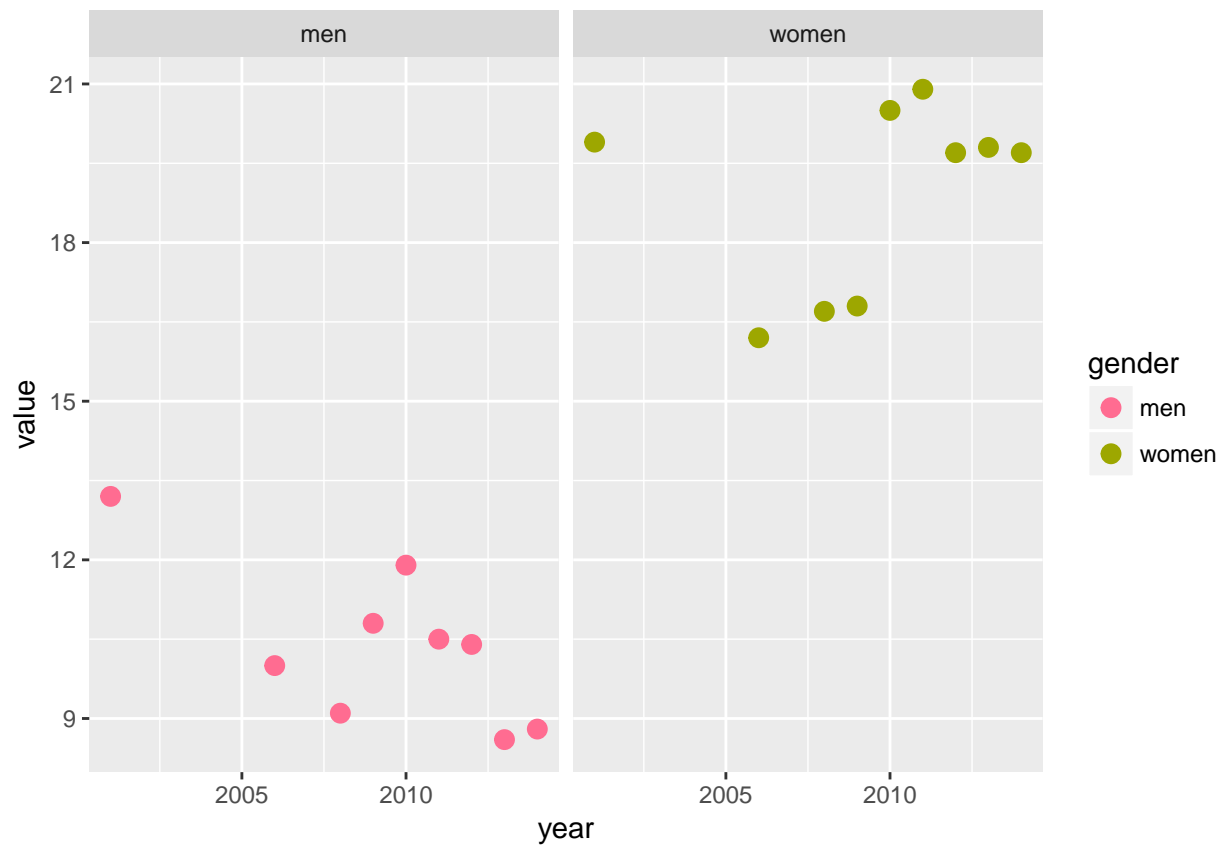
```
# facettting with facet wrap

ggplot(joint_unemployment_total,
       aes(x = year,
           y = value,
           col = gender)) +
  geom_point(size = 3) +
  scale_colour_hue(h = c(0, 90)) +
  facet_wrap(~gender)
```
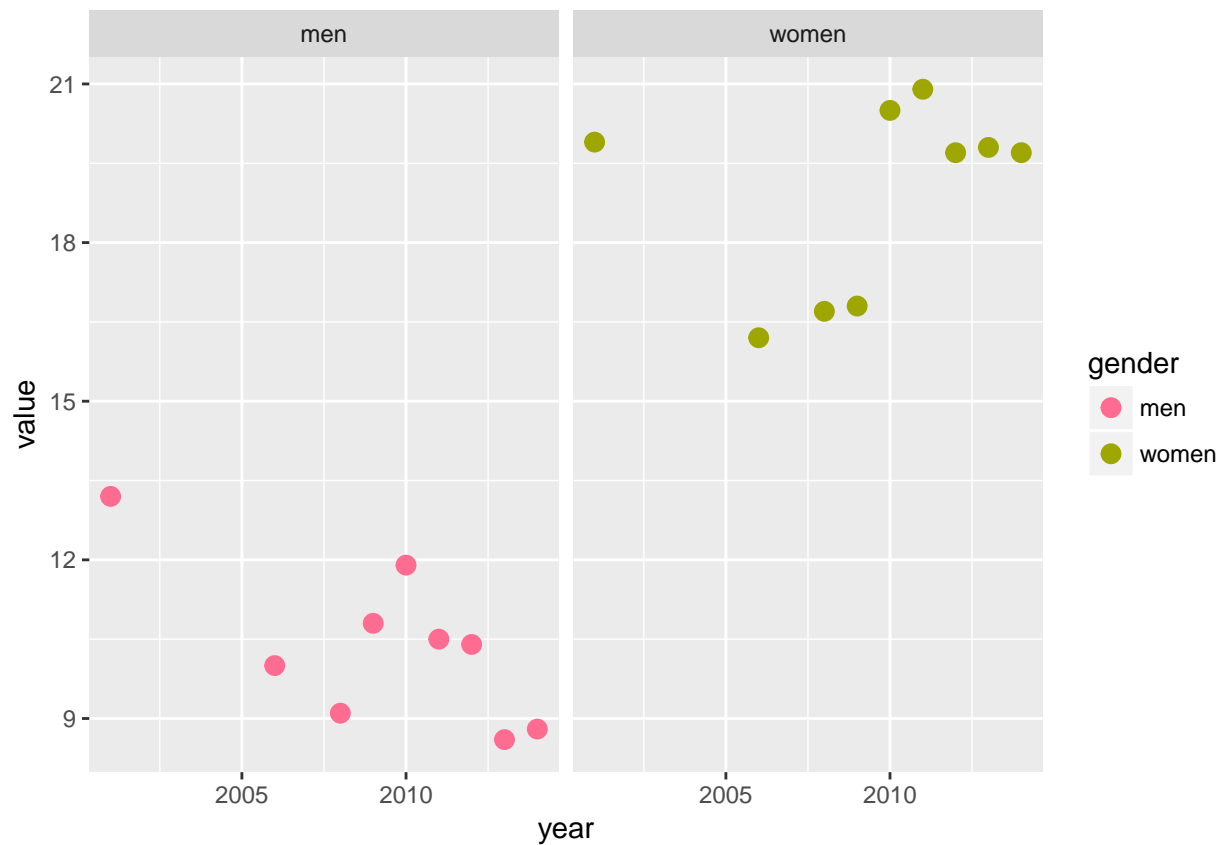
4. Use the `theme()` layer to change the appearance of a plot of your choice including the

- plot, axes, and legend titles
- axes tick marks
- text size
- legend position

```
# Changing apprearance of the plot
plot_grid <- ggplot(joint_unemployment_total,
      aes(x = year,
          y = value,
          col = gender)) +
  geom_point(size = 3) +
  scale_colour_hue(h = c(0, 90)) +
  facet_grid(~gender)

plot_grid +
theme(legend.title = element_text(colour = "brown", size = "10"), axis.title = element_text(colour = "b)
)
```
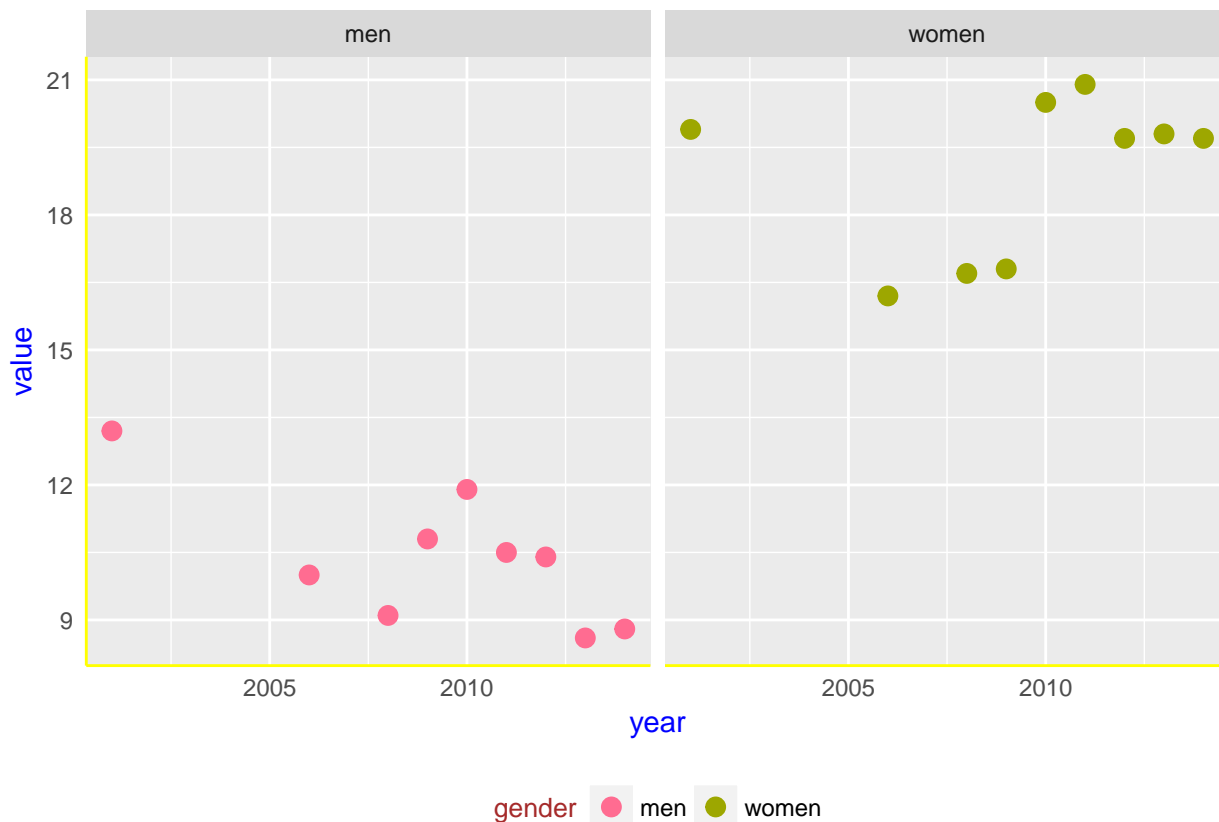
5. Create three versions of a graphic of your choice using different built-in themes or a theme created from `ggthemes`. Which ones do you think are best for presenting in an academic journal? A poster session? What are the qualities of the themes that you choice that you think make them more appropriate for presentation?

```
install.packages('ggthemes', dependencies = TRUE)
```

```
## Installing package into '/Users/Ramin/Desktop/Autumn 2017/Statistics 321/unemployment rate/unemployme
## (as 'lib' is unspecified)
##
## The downloaded binary packages are in
##   /var/folders/z9/b9hh4hpj6hl6x9r9dxjctv240000gn/T//RtmpsbS4CZ/downloaded_packages
```
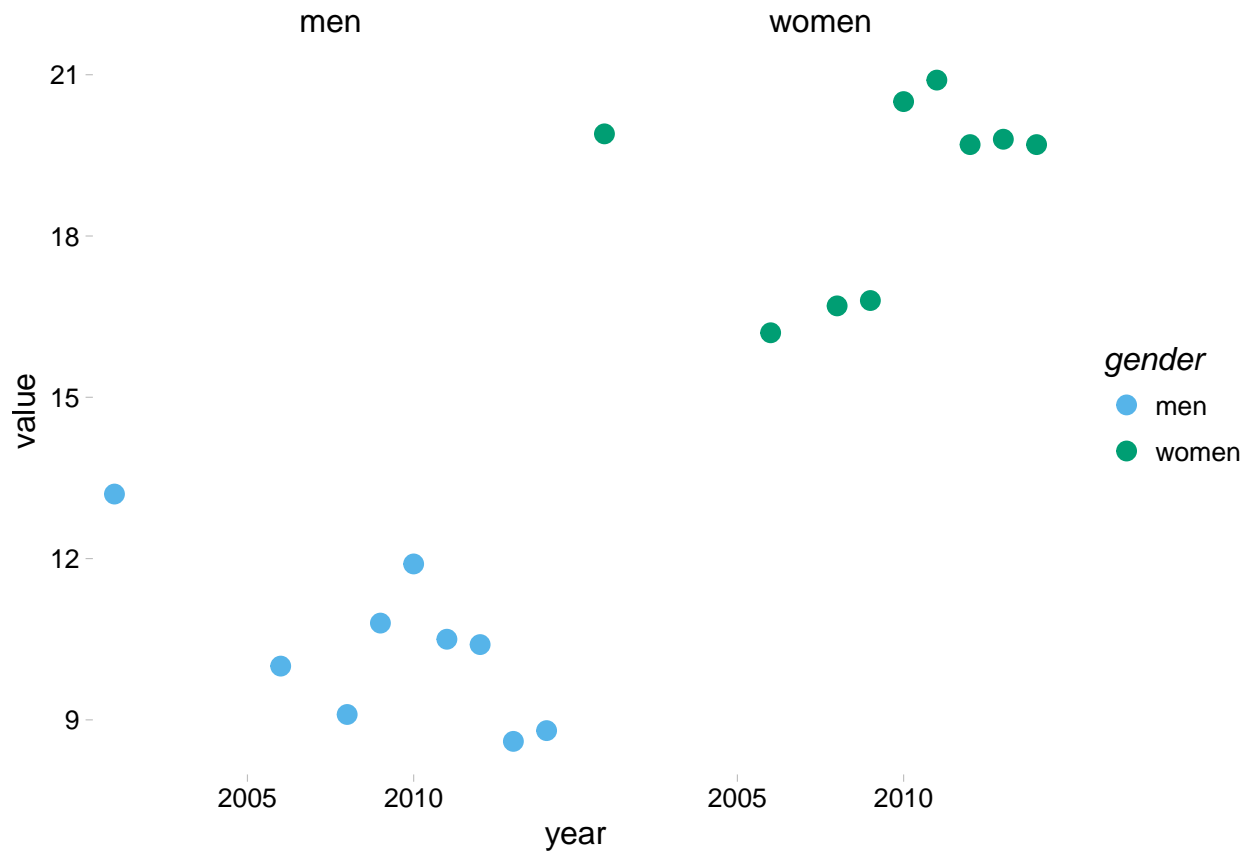
```
library(ggplot2)
library(ggthemes)
plot_grid +
theme_pander() + scale_colour_pander()
```
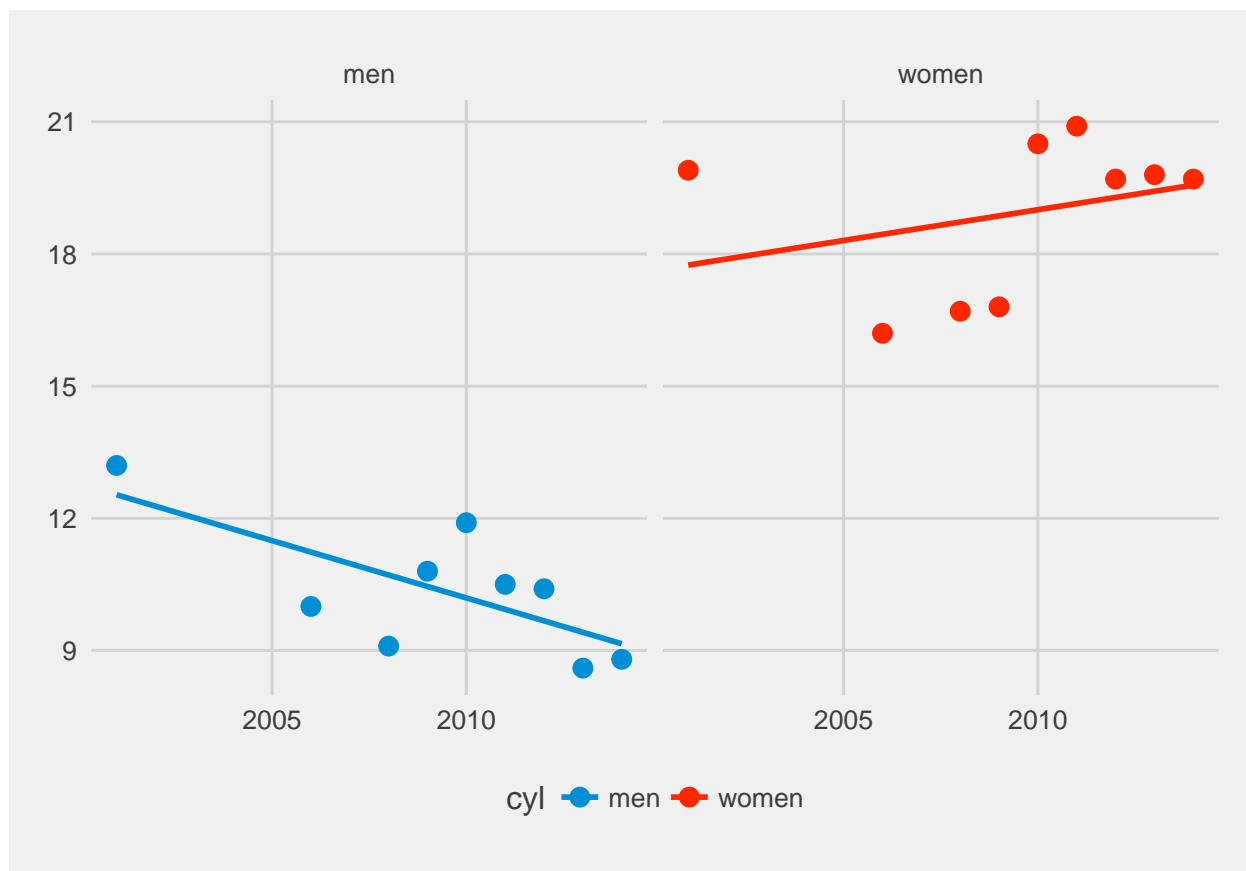
```
## Scale for 'colour' is already present. Adding another scale for
## 'colour', which will replace the existing scale.
```

```
plot_grid + geom_smooth(method = "lm", se = FALSE) +
  scale_color_fivethirtyeight("cyl") +
  theme_fivethirtyeight()
```

```
## Scale for 'colour' is already present. Adding another scale for
## 'colour', which will replace the existing scale.
```

I think this one at the bottom is the best one for presenttion and publishiing in academic journal since the visualization is simple and the variables are clearly shown, year, value, and gender(men and women). In addition the values for both gender are distingusihed with different colors which help to the audiance to to know that each line is drawn for which gender. The line also accurately shows the tendency of uneployment for both genders.

```
plot_grid +
  ggtitle("Unemployment Rate") +
  theme_hc() +
  scale_fill_hc("darkunica") +
  geom_line()
```

Unemployment Rate