Project Journal

Anna Repesh Belle Mahmoud

12/7/2021

#Reading in dataset  
birth <- read\_csv("NationalAndStatePregnancy\_PublicUse.csv")

## Rows: 912 Columns: 103

## -- Column specification --------------------------------------------------------  
## Delimiter: ","  
## chr (3): state, notes, versiondate  
## dbl (100): year, pregnancyratelt15, pregnancyrate1517, pregnancyrate1819, pr...

##   
## i Use `spec()` to retrieve the full column specification for this data.  
## i Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

#Select and filtering  
birth2 <- birth %>%  
 select("state", "year", "abortionrate2024" : "abortionrate40plus") %>%  
 filter(year >= 2000, state != "US" & state != "DC")   
  
  
pivot\_birth <- birth2 %>%  
 pivot\_longer(!state:year, names\_to = "groups", values\_to = "rates")  
  
  
state2 <- c("ME", "MA", "RI", "CT", "NH", "VT", "NY", "PA", "NJ", "DE", "MD", "WV", "VA", "KY", "TN", "NC", "SC", "GA", "AL", "MS", "AR", "LA", "FL", "OH", "IN", "MI", "IL", "MO", "WI", "MN", "IA", "KS", "NE", "SD", "ND", "TX", "OK", "NM", "AZ", "CO", "WY", "MT", "ID", "WA", "OR", "UT", "NV", "CA", "AK", "HI")  
region <- c("Northeast", "Northeast", "Northeast", "Northeast", "Northeast", "Northeast", "Northeast", "Northeast", "Northeast", "Northeast", "Northeast",   
 "Southeast", "Southeast", "Southeast", "Southeast", "Southeast", "Southeast", "Southeast", "Southeast", "Southeast", "Southeast", "Southeast", "Southeast",   
 "Midwest", "Midwest", "Midwest", "Midwest", "Midwest", "Midwest", "Midwest", "Midwest", "Midwest", "Midwest", "Midwest", "Midwest",  
 "Southwest", "Southwest", "Southwest", "Southwest",  
 "West", "West", "West", "West", "West", "West", "West", "West", "West", "West", "West")  
stateregion <- data.frame(state2, region)  
  
  
is.element(pivot\_birth$state, stateregion$state2) %>%  
 all()

## [1] TRUE

pivot\_birth\_region <- merge(pivot\_birth, stateregion, by.x = c("state"), by.y = c("state2"), all.x = FALSE)

#Southwest abortion rates  
Southwest <- function(x){  
 filter(pivot\_birth\_region, year==x, region == "Southwest") %>%  
 group\_by(groups, year) %>%  
 summarise(Rate\_Avg = mean(rates))}  
Southwest(2000)

## `summarise()` has grouped output by 'groups'. You can override using the `.groups` argument.

## # A tibble: 5 x 3  
## # Groups: groups [5]  
## groups year Rate\_Avg  
## <chr> <dbl> <dbl>  
## 1 abortionrate2024 2000 33.6   
## 2 abortionrate2529 2000 22.6   
## 3 abortionrate3034 2000 13.0   
## 4 abortionrate3539 2000 7.02  
## 5 abortionrate40plus 2000 2.55

#West Abortion rates  
West <- function(x){  
 filter(pivot\_birth\_region, year==x, region == "West") %>%  
 group\_by(groups, year) %>%  
 summarise(Rate\_Avg = mean(rates))}  
West(2000)

## `summarise()` has grouped output by 'groups'. You can override using the `.groups` argument.

## # A tibble: 5 x 3  
## # Groups: groups [5]  
## groups year Rate\_Avg  
## <chr> <dbl> <dbl>  
## 1 abortionrate2024 2000 37.8   
## 2 abortionrate2529 2000 26.1   
## 3 abortionrate3034 2000 14.9   
## 4 abortionrate3539 2000 9.41  
## 5 abortionrate40plus 2000 3.15

#Midwest abortion rates  
Midwest <- function(x){  
 filter(pivot\_birth\_region, year==x, region == "Midwest") %>%  
 group\_by(groups, year) %>%  
 summarise(Rate\_Avg = mean(rates))}  
Midwest(2000)

## `summarise()` has grouped output by 'groups'. You can override using the `.groups` argument.

## # A tibble: 5 x 3  
## # Groups: groups [5]  
## groups year Rate\_Avg  
## <chr> <dbl> <dbl>  
## 1 abortionrate2024 2000 27.2   
## 2 abortionrate2529 2000 19.4   
## 3 abortionrate3034 2000 11.0   
## 4 abortionrate3539 2000 5.78  
## 5 abortionrate40plus 2000 1.98

#Northeast abortion rates  
  
Northeast <- function(x){  
 filter(pivot\_birth\_region, year==x, region == "Northeast") %>%  
 group\_by(groups, year) %>%  
 summarise(Rate\_Avg = mean(rates))}  
Northeast(2000)

## `summarise()` has grouped output by 'groups'. You can override using the `.groups` argument.

## # A tibble: 5 x 3  
## # Groups: groups [5]  
## groups year Rate\_Avg  
## <chr> <dbl> <dbl>  
## 1 abortionrate2024 2000 51.0  
## 2 abortionrate2529 2000 34.1  
## 3 abortionrate3034 2000 18.9  
## 4 abortionrate3539 2000 10.2  
## 5 abortionrate40plus 2000 3.5

#Southeast Abortion rates  
  
Southeast <- function(x){  
 filter(pivot\_birth\_region, year==x, region == "Southeast") %>%  
 group\_by(groups, year) %>%  
 summarise(Rate\_Avg = mean(rates))}  
Southeast(2000)

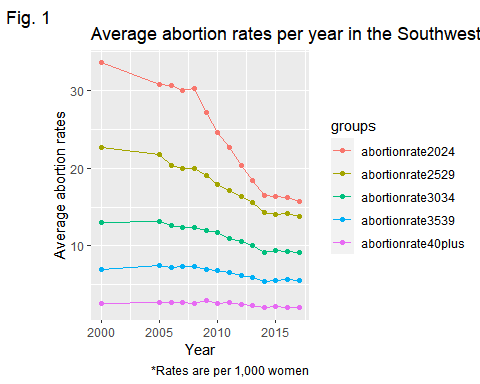
## `summarise()` has grouped output by 'groups'. You can override using the `.groups` argument.

## # A tibble: 5 x 3  
## # Groups: groups [5]  
## groups year Rate\_Avg  
## <chr> <dbl> <dbl>  
## 1 abortionrate2024 2000 33.1   
## 2 abortionrate2529 2000 22.1   
## 3 abortionrate3034 2000 11.9   
## 4 abortionrate3539 2000 6.12  
## 5 abortionrate40plus 2000 1.88

#Line graph for Southwest  
Southwest1 <- pivot\_birth\_region %>%  
 filter(region == "Southwest") %>%  
 group\_by(groups, year) %>%  
 summarise(Rate\_Avg = mean(rates)) %>%  
 ggplot(mapping=aes(x=year, y=Rate\_Avg, color=groups)) +  
 geom\_point() +  
 geom\_line() +  
 labs(x = "Year",  
 y= "Average abortion rates",  
 title = "Average abortion rates per year in the Southwest region from 2000-2017",  
 caption = "\*Rates are per 1,000 women",  
 tag = "Fig. 1")

## `summarise()` has grouped output by 'groups'. You can override using the `.groups` argument.

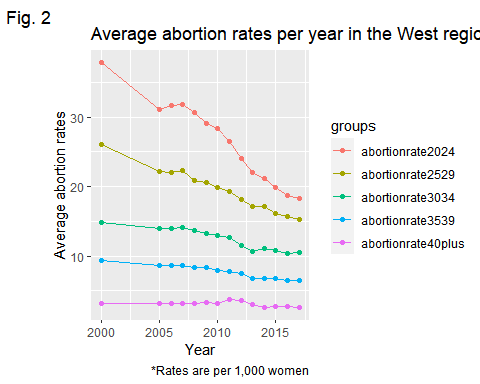
Southwest1



#line graph for west  
West1 <- pivot\_birth\_region %>%  
 filter(region == "West") %>%  
 group\_by(groups, year) %>%  
 summarise(Rate\_Avg = mean(rates)) %>%  
 ggplot(mapping=aes(x=year, y=Rate\_Avg, color=groups)) +  
 geom\_point() +  
 geom\_line() +  
 labs(x = "Year",  
 y= "Average abortion rates",  
 title = "Average abortion rates per year in the West region from 2000-2017",  
 caption = "\*Rates are per 1,000 women",  
 tag = "Fig. 2")

## `summarise()` has grouped output by 'groups'. You can override using the `.groups` argument.

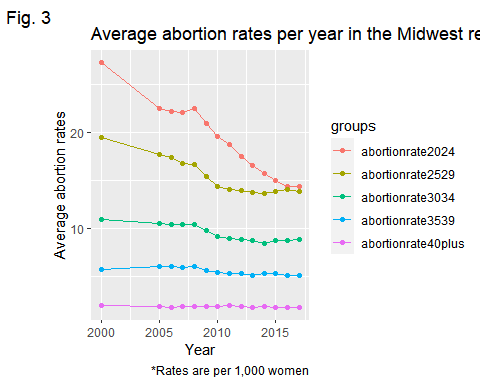
West1



#Line graph for Midwest  
Midwest1 <- pivot\_birth\_region %>%  
 filter(region == "Midwest") %>%  
 group\_by(groups, year) %>%  
 summarise(Rate\_Avg = mean(rates)) %>%  
 ggplot(mapping=aes(x=year, y=Rate\_Avg, color=groups)) +  
 geom\_point() +  
 geom\_line() +  
 labs(x = "Year",  
 y= "Average abortion rates",  
 title = "Average abortion rates per year in the Midwest region from 2000-2017",  
 caption = "\*Rates are per 1,000 women",  
 tag = "Fig. 3")

## `summarise()` has grouped output by 'groups'. You can override using the `.groups` argument.

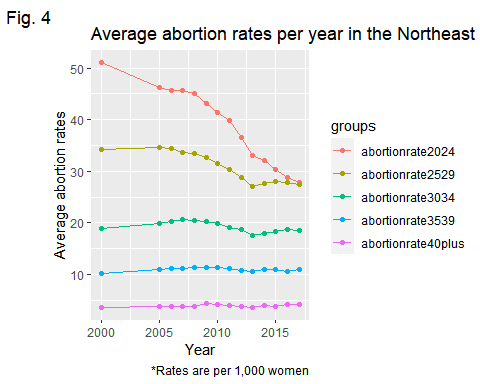
Midwest1



#Line graph for Northeast  
Northeast1 <- pivot\_birth\_region %>%  
 filter(region == "Northeast") %>%  
 group\_by(groups, year) %>%  
 summarise(Rate\_Avg = mean(rates)) %>%  
 ggplot(mapping=aes(x=year, y=Rate\_Avg, color=groups)) +  
 geom\_point() +  
 geom\_line() +  
 labs(x = "Year",  
 y= "Average abortion rates",  
 title = "Average abortion rates per year in the Northeast region from 2000-2017",  
 caption = "\*Rates are per 1,000 women",  
 tag = "Fig. 4")

## `summarise()` has grouped output by 'groups'. You can override using the `.groups` argument.

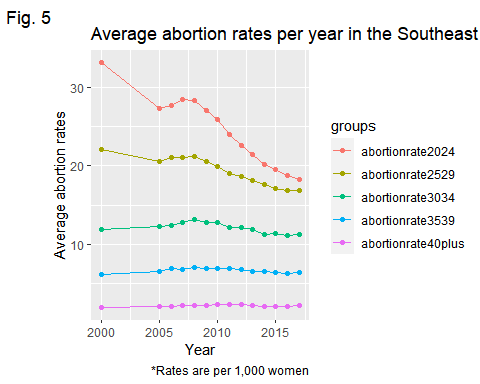
Northeast1



#Line graph for Southeast  
Southeast1 <- pivot\_birth\_region %>%  
 filter(region == "Southeast") %>%  
 group\_by(groups, year) %>%  
 summarise(Rate\_Avg = mean(rates)) %>%  
 ggplot(mapping=aes(x=year, y=Rate\_Avg, color=groups)) +  
 geom\_point() +  
 geom\_line() +  
 labs(x = "Year",  
 y= "Average abortion rates",  
 title = "Average abortion rates per year in the Southeast region from 2000-2017",  
 caption = "\*Rates are per 1,000 women",  
 tag = "Fig. 5")

## `summarise()` has grouped output by 'groups'. You can override using the `.groups` argument.

Southeast1



#Simple Southwest  
simplesouthw <- pivot\_birth\_region %>%  
 filter(region == "Southwest") %>%  
 group\_by(groups, year) %>%  
 summarise(Rate\_Avg = mean(rates))

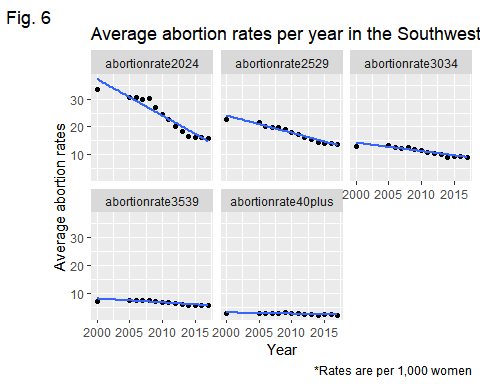
## `summarise()` has grouped output by 'groups'. You can override using the `.groups` argument.

simplesouthw

## # A tibble: 70 x 3  
## # Groups: groups [5]  
## groups year Rate\_Avg  
## <chr> <dbl> <dbl>  
## 1 abortionrate2024 2000 33.6  
## 2 abortionrate2024 2005 30.8  
## 3 abortionrate2024 2006 30.6  
## 4 abortionrate2024 2007 30.0  
## 5 abortionrate2024 2008 30.2  
## 6 abortionrate2024 2009 27.2  
## 7 abortionrate2024 2010 24.6  
## 8 abortionrate2024 2011 22.7  
## 9 abortionrate2024 2012 20.3  
## 10 abortionrate2024 2013 18.4  
## # ... with 60 more rows

#SLR for Simple Southwest  
ggplot(data=simplesouthw, aes(x=year, y=Rate\_Avg)) +  
 geom\_point() +  
 geom\_smooth(method = lm, se= FALSE) +  
 facet\_wrap(vars(groups))+  
 labs(x = "Year",  
 y= "Average abortion rates",  
 title = "Average abortion rates per year in the Southwest region from 2000-2017",  
 caption = "\*Rates are per 1,000 women",  
 tag = "Fig. 6")

## `geom\_smooth()` using formula 'y ~ x'



#Model for Simple Southwest  
model <- lm(Rate\_Avg ~ year + groups, data=simplesouthw)  
model %>%  
 broom::tidy()

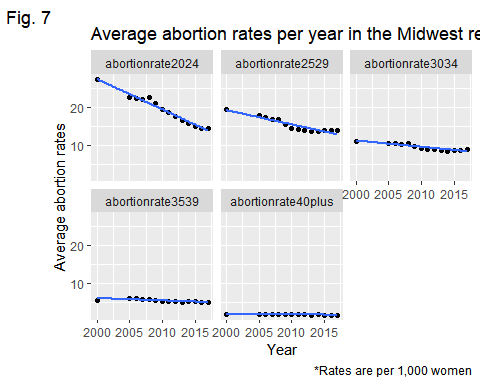
## # A tibble: 6 x 5  
## term estimate std.error statistic p.value  
## <chr> <dbl> <dbl> <dbl> <dbl>  
## 1 (Intercept) 999. 126. 7.94 4.06e-11  
## 2 year -0.485 0.0625 -7.75 8.75e-11  
## 3 groupsabortionrate2529 -6.15 0.907 -6.78 4.50e- 9  
## 4 groupsabortionrate3034 -12.7 0.907 -14.0 4.37e-21  
## 5 groupsabortionrate3539 -17.3 0.907 -19.1 4.38e-28  
## 6 groupsabortionrate40plus -21.3 0.907 -23.5 3.55e-33

#Simple Midwest  
simplemid <- pivot\_birth\_region %>%  
 filter(region == "Midwest") %>%  
 group\_by(groups, year) %>%  
 summarise(Rate\_Avg = mean(rates))

## `summarise()` has grouped output by 'groups'. You can override using the `.groups` argument.

#SLR for simple Midwest  
ggplot(data=simplemid, mapping=aes(x=year, y=Rate\_Avg)) +  
 geom\_point() +  
 geom\_smooth(method = lm, se=F) +  
 facet\_wrap(vars(groups)) +  
 labs(x = "Year",  
 y= "Average abortion rates",  
 title = "Average abortion rates per year in the Midwest region from 2000-2017",  
 caption = "\*Rates are per 1,000 women",  
 tag = "Fig. 7")

## `geom\_smooth()` using formula 'y ~ x'



#Model for simple Midwest  
model <- lm(Rate\_Avg ~ year + groups, data=simplemid)  
summary(model)

##   
## Call:  
## lm(formula = Rate\_Avg ~ year + groups, data = simplemid)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -3.2250 -0.6589 -0.0351 0.8937 5.1146   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 584.89180 75.80619 7.716 1.02e-10 \*\*\*  
## year -0.28138 0.03771 -7.462 2.87e-10 \*\*\*  
## groupsabortionrate2529 -3.88274 0.54680 -7.101 1.24e-09 \*\*\*  
## groupsabortionrate3034 -9.73214 0.54680 -17.798 < 2e-16 \*\*\*  
## groupsabortionrate3539 -13.68690 0.54680 -25.031 < 2e-16 \*\*\*  
## groupsabortionrate40plus -17.35714 0.54680 -31.743 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.447 on 64 degrees of freedom  
## Multiple R-squared: 0.956, Adjusted R-squared: 0.9526   
## F-statistic: 278 on 5 and 64 DF, p-value: < 2.2e-16

model %>%  
 broom::tidy()

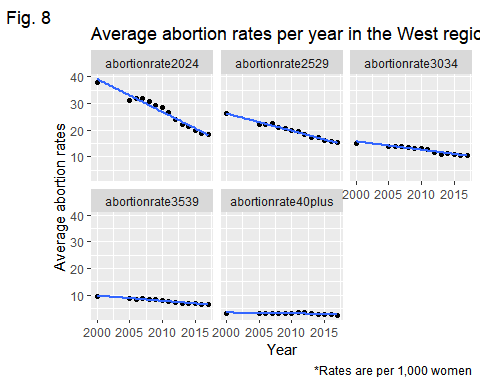
## # A tibble: 6 x 5  
## term estimate std.error statistic p.value  
## <chr> <dbl> <dbl> <dbl> <dbl>  
## 1 (Intercept) 585. 75.8 7.72 1.02e-10  
## 2 year -0.281 0.0377 -7.46 2.87e-10  
## 3 groupsabortionrate2529 -3.88 0.547 -7.10 1.24e- 9  
## 4 groupsabortionrate3034 -9.73 0.547 -17.8 1.79e-26  
## 5 groupsabortionrate3539 -13.7 0.547 -25.0 9.15e-35  
## 6 groupsabortionrate40plus -17.4 0.547 -31.7 7.02e-41

#Simple West  
simplewest <- pivot\_birth\_region %>%  
 filter(region == "West") %>%  
 group\_by(groups, year) %>%  
 summarise(Rate\_Avg = mean(rates))

## `summarise()` has grouped output by 'groups'. You can override using the `.groups` argument.

#SLR for simple West  
ggplot(data=simplewest, mapping=aes(x=year, y=Rate\_Avg)) +  
 geom\_point() +  
 geom\_smooth(method = lm, se=F) +  
 facet\_wrap(vars(groups)) +  
 labs(x = "Year",  
 y= "Average abortion rates",  
 title = "Average abortion rates per year in the West region from 2000-2017",  
 caption = "\*Rates are per 1,000 women",  
 tag = "Fig. 8")

## `geom\_smooth()` using formula 'y ~ x'



#Model for simple West  
model <- lm(Rate\_Avg ~ year + groups, data=simplewest)  
summary(model)

##   
## Call:  
## lm(formula = Rate\_Avg ~ year + groups, data = simplewest)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -5.0493 -1.0044 0.1694 1.3050 6.3908   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 998.02382 110.87042 9.002 5.62e-13 \*\*\*  
## year -0.48329 0.05515 -8.763 1.47e-12 \*\*\*  
## groupsabortionrate2529 -7.01623 0.79973 -8.773 1.41e-12 \*\*\*  
## groupsabortionrate3034 -14.02208 0.79973 -17.534 < 2e-16 \*\*\*  
## groupsabortionrate3539 -18.73506 0.79973 -23.427 < 2e-16 \*\*\*  
## groupsabortionrate40plus -23.41818 0.79973 -29.283 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 2.116 on 64 degrees of freedom  
## Multiple R-squared: 0.9477, Adjusted R-squared: 0.9437   
## F-statistic: 232.1 on 5 and 64 DF, p-value: < 2.2e-16

model %>%  
 broom::tidy()

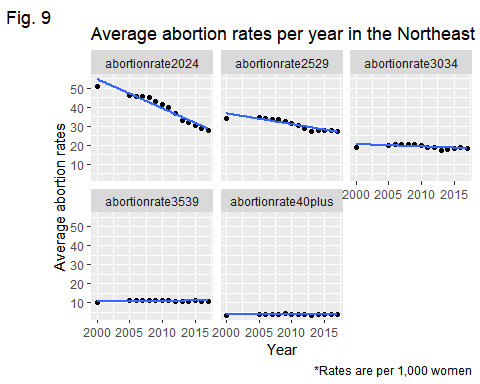
## # A tibble: 6 x 5  
## term estimate std.error statistic p.value  
## <chr> <dbl> <dbl> <dbl> <dbl>  
## 1 (Intercept) 998. 111. 9.00 5.62e-13  
## 2 year -0.483 0.0552 -8.76 1.47e-12  
## 3 groupsabortionrate2529 -7.02 0.800 -8.77 1.41e-12  
## 4 groupsabortionrate3034 -14.0 0.800 -17.5 3.97e-26  
## 5 groupsabortionrate3539 -18.7 0.800 -23.4 4.20e-33  
## 6 groupsabortionrate40plus -23.4 0.800 -29.3 8.84e-39

#Simple Northeast  
simplenorth <- pivot\_birth\_region %>%  
 filter(region == "Northeast") %>%  
 group\_by(groups, year) %>%  
 summarise(Rate\_Avg = mean(rates))

## `summarise()` has grouped output by 'groups'. You can override using the `.groups` argument.

#SLR for simple Northeast  
ggplot(data=simplenorth, mapping=aes(x=year, y=Rate\_Avg)) +  
 geom\_point() +  
 geom\_smooth(method = lm, se=F) +  
 facet\_wrap(vars(groups)) +  
 labs(x = "Year",  
 y= "Average abortion rates",  
 title = "Average abortion rates per year in the Northeast region from 2000-2017",  
 caption = "\*Rates are per 1,000 women",  
 tag = "Fig. 9")

## `geom\_smooth()` using formula 'y ~ x'



#Model for simple Northeast  
model <- lm(Rate\_Avg ~ year + groups, data=simplenorth)  
summary(model)

##   
## Call:  
## lm(formula = Rate\_Avg ~ year + groups, data = simplenorth)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -8.180 -1.206 0.280 1.615 7.500   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 919.76579 158.25345 5.812 2.12e-07 \*\*\*  
## year -0.43812 0.07872 -5.565 5.52e-07 \*\*\*  
## groupsabortionrate2529 -8.21818 1.14151 -7.199 8.30e-10 \*\*\*  
## groupsabortionrate3034 -19.82597 1.14151 -17.368 < 2e-16 \*\*\*  
## groupsabortionrate3539 -28.07597 1.14151 -24.596 < 2e-16 \*\*\*  
## groupsabortionrate40plus -35.15909 1.14151 -30.801 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 3.02 on 64 degrees of freedom  
## Multiple R-squared: 0.9526, Adjusted R-squared: 0.9489   
## F-statistic: 257.5 on 5 and 64 DF, p-value: < 2.2e-16

model %>%  
 broom::tidy()

## # A tibble: 6 x 5  
## term estimate std.error statistic p.value  
## <chr> <dbl> <dbl> <dbl> <dbl>  
## 1 (Intercept) 920. 158. 5.81 2.12e- 7  
## 2 year -0.438 0.0787 -5.57 5.52e- 7  
## 3 groupsabortionrate2529 -8.22 1.14 -7.20 8.30e-10  
## 4 groupsabortionrate3034 -19.8 1.14 -17.4 6.56e-26  
## 5 groupsabortionrate3539 -28.1 1.14 -24.6 2.53e-34  
## 6 groupsabortionrate40plus -35.2 1.14 -30.8 4.30e-40

#Simple Southeast  
simplesouthe <- pivot\_birth\_region %>%  
 filter(region == "Southeast") %>%  
 group\_by(groups, year) %>%  
 summarise(Rate\_Avg = mean(rates))

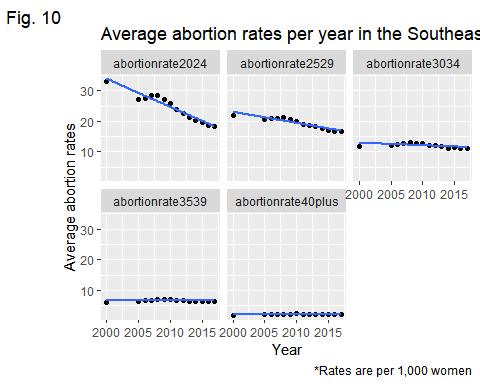
## `summarise()` has grouped output by 'groups'. You can override using the `.groups` argument.

simplesouthe

## # A tibble: 70 x 3  
## # Groups: groups [5]  
## groups year Rate\_Avg  
## <chr> <dbl> <dbl>  
## 1 abortionrate2024 2000 33.1  
## 2 abortionrate2024 2005 27.3  
## 3 abortionrate2024 2006 27.6  
## 4 abortionrate2024 2007 28.5  
## 5 abortionrate2024 2008 28.4  
## 6 abortionrate2024 2009 27.1  
## 7 abortionrate2024 2010 25.9  
## 8 abortionrate2024 2011 24.0  
## 9 abortionrate2024 2012 22.6  
## 10 abortionrate2024 2013 21.4  
## # ... with 60 more rows

#SLR for simple Southeast  
ggplot(data=simplesouthe, mapping=aes(x=year, y=Rate\_Avg)) +  
 geom\_point() +  
 geom\_smooth(method = lm, se=F) +  
 facet\_wrap(vars(groups)) +  
 labs(x = "Year",  
 y= "Average abortion rates",  
 title = "Average abortion rates per year in the Southeast region from 2000-2017",  
 caption = "\*Rates are per 1,000 women",  
 tag = "Fig. 10")

## `geom\_smooth()` using formula 'y ~ x'



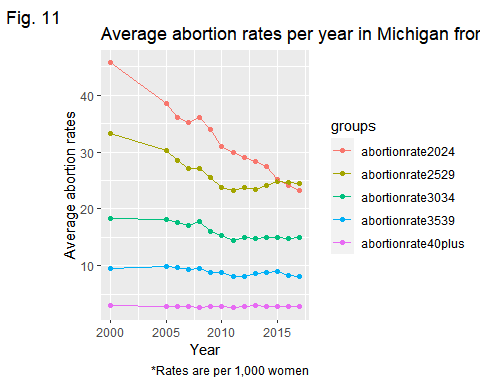
#Model for simple Southeast  
model <- lm(Rate\_Avg ~ year + groups, data=simplesouthe)  
summary(model)

##   
## Call:  
## lm(formula = Rate\_Avg ~ year + groups, data = simplesouthe)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -4.3636 -0.7904 0.2513 0.8725 5.8896   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 569.01673 92.78278 6.133 6.00e-08 \*\*\*  
## year -0.27089 0.04616 -5.869 1.69e-07 \*\*\*  
## groupsabortionrate2529 -5.14702 0.66926 -7.691 1.13e-10 \*\*\*  
## groupsabortionrate3034 -12.40476 0.66926 -18.535 < 2e-16 \*\*\*  
## groupsabortionrate3539 -17.86429 0.66926 -26.693 < 2e-16 \*\*\*  
## groupsabortionrate40plus -22.34167 0.66926 -33.383 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.771 on 64 degrees of freedom  
## Multiple R-squared: 0.9595, Adjusted R-squared: 0.9563   
## F-statistic: 302.9 on 5 and 64 DF, p-value: < 2.2e-16

model %>%  
 broom::tidy()

## # A tibble: 6 x 5  
## term estimate std.error statistic p.value  
## <chr> <dbl> <dbl> <dbl> <dbl>  
## 1 (Intercept) 569. 92.8 6.13 6.00e- 8  
## 2 year -0.271 0.0462 -5.87 1.69e- 7  
## 3 groupsabortionrate2529 -5.15 0.669 -7.69 1.13e-10  
## 4 groupsabortionrate3034 -12.4 0.669 -18.5 2.02e-27  
## 5 groupsabortionrate3539 -17.9 0.669 -26.7 2.13e-36  
## 6 groupsabortionrate40plus -22.3 0.669 -33.4 3.35e-42

#Average abortion rates for Michigan  
MI <- pivot\_birth %>%  
 filter(state=="MI") %>%  
 group\_by(groups, year) %>%  
 ggplot(mapping=aes(x=year, y=rates, color=groups)) +  
 geom\_point() +  
 geom\_line() +  
 labs(x = "Year",  
 y= "Average abortion rates",  
 title = "Average abortion rates per year in Michigan from 2000-2017",  
 caption = "\*Rates are per 1,000 women",  
 tag = "Fig. 11")  
  
MI



#adding variables to the for the sas code  
abortion2014 <- birth2 %>%  
 filter(year == 2014)  
  
  
state <- c("WA", "OR", "CA", "NV", "MT", "WV", "CO", "NM", "MN", "DE", "IA", "IL", "WY", "MD", "NJ", "NY", "CT", "MA", "VT", "NH", "ME", "AK", "HI", "RI", "PA", "MI", "KY", "TN", "NC", "SC", "GA","ID", "UT", "AZ", "ND", "SD", "NE", "KS", "OK", "TX", "WI", "MO", "AR", "LA", "MS", "AL", "FL", "IN", "OH", "VA")  
rank <- c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,2,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3)  
staterank <- data.frame(state, rank)  
  
  
is.element(abortion2014$state, staterank$state) %>%  
 all()

## [1] TRUE

Abortionrateswithrank <- merge(abortion2014, staterank, by = 'state', all.x = FALSE)  
  
is.element(Abortionrateswithrank$state, stateregion$state2) %>%  
 all()

## [1] TRUE

Abortionrateswithrankandregion <- merge(Abortionrateswithrank, stateregion, by.x = c("state"), by.y = c("state2"), all.x = FALSE)  
  
#getting average abortion rate and rank  
AveragesperRegion <- Abortionrateswithrankandregion %>%  
 group\_by(region) %>%  
 summarise(Avg\_Rate2024 = mean(abortionrate2024),   
 Avg\_Rate2529 = mean(abortionrate2529),   
 Avg\_Rate3034 = mean(abortionrate3034),  
 Avg\_Rate3539 = mean(abortionrate3539),  
 Avg\_Rate40plus = mean(abortionrate40plus),  
 Avg\_Rank = mean(rank),  
 Avg\_Rank\_Round = round(Avg\_Rank))  
  
AveragesperRegion

## # A tibble: 5 x 8  
## region Avg\_Rate2024 Avg\_Rate2529 Avg\_Rate3034 Avg\_Rate3539 Avg\_Rate40plus  
## <chr> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 Midwest 15.7 13.7 8.48 5.4 1.89  
## 2 Northeast 32.1 27.6 17.9 10.9 4   
## 3 Southeast 20.2 17.6 11.3 6.52 2.1   
## 4 Southwest 16.6 14.4 9.15 5.38 2.1   
## 5 West 21.2 17.2 11.1 6.77 2.66  
## # ... with 2 more variables: Avg\_Rank <dbl>, Avg\_Rank\_Round <dbl>

#These are commented out but they download the charts to my computer  
  
#write.csv(Abortionrateswithrank, "Abortion\_rates.csv")  
#write.csv(Abortionrateswithrankandregion, "Abortion\_rates\_Regions.csv")

#SAS output

%let path=/home/u58831290/CIS661;

Libname CIS661 "&Path.";

**proc** **import** datafile="&path./Abortion\_rates\_Regions.csv"

out=abortion

dbms=csv

replace;

getnames=yes;

**run**;

\*2024;

title2 'Comparison of Abortion Rates to Laws 20-24';

**proc** **glm** data = abortion;

class rank;

model abortionrate2024= rank;

means rank / bon lines;

lsmeans rank;

**run**;

Chart, box and whisker chart

Description automatically generated

Graphical user interface, application, table

Description automatically generated

\*2529;

title 'Comparison of Abortion Rates to Laws for 25-29';

**proc** **glm** data = abortion;

class rank;

model abortionrate2529= rank;

means rank / bon lines;

lsmeans rank;

Chart, box and whisker chart

Description automatically generated**run**;

Graphical user interface

Description automatically generated

\*3034;

title'Comparison of Abortion Rates to Laws for 30-34';

**proc** **glm** data = abortion;

class rank;

model abortionrate3034= rank;

means rank / bon lines;

lsmeans rank;

Chart, box and whisker chart

Description automatically generated**run**;

Graphical user interface, application

Description automatically generated

\*3539;

title'Comparison of Abortion Rates to Laws for 35-39';

**proc** **glm** data = abortion;

class rank;

model abortionrate3539= rank;

means rank / bon lines;

lsmeans rank;

Chart, box and whisker chart

Description automatically generated**run**;

Chart

Description automatically generated with low confidence

\*40plus;

title'Comparison of Abortion Rates to Laws for 40+';

**proc** **glm** data = abortion;

class rank;

model abortionrate40plus= rank;

means rank / bon lines;

lsmeans rank;

Chart, box and whisker chart

Description automatically generated**run**;

Graphical user interface

Description automatically generated with medium confidence