Constructing a line graph

The goal here is to use the data from the Stata file `FCEFD.dta` to create a grouped line graph comparing three series (average, small firms, and large firms) over some years. At the year 2008, a vertical line should be added.

The most important part of this figure is to draw a vertical line at 2008 to compare pre- and post-crisis trends.

To accomplish this, we will use the following packages:

```
here::i_am("material/Session-03-Example-Lineplot.qmd") # Adjust to your case
library(here)
library(haven) # to import stata files
library(dplyr) # to do data manipulation
library(tidyr) # to use do data wrangling
library(ggplot2) # for visualization
library(scales) # for scaling
```

We first have a look at the data set:

```
firm_raw <- haven::read_dta(here("material/data/FCEFD.dta"))
firm_raw <- as_tibble(firm_raw)
head(firm_raw)
```

```
# A tibble: 6 x 4
   year     avr L_firms  H_firms
  <dbl>   <dbl>   <dbl>    <dbl>
1  2003 0.00380 0.00359 0.00406
2  2004 0.00183 0.00315 0
3  2005 0.00512 0.00340 0.00752
4  2006 0.00442 0.00595 0.00240
5  2007 0.00400 0.00613 0.000991
6  2008 0.00383 0.00493 0.00231
```

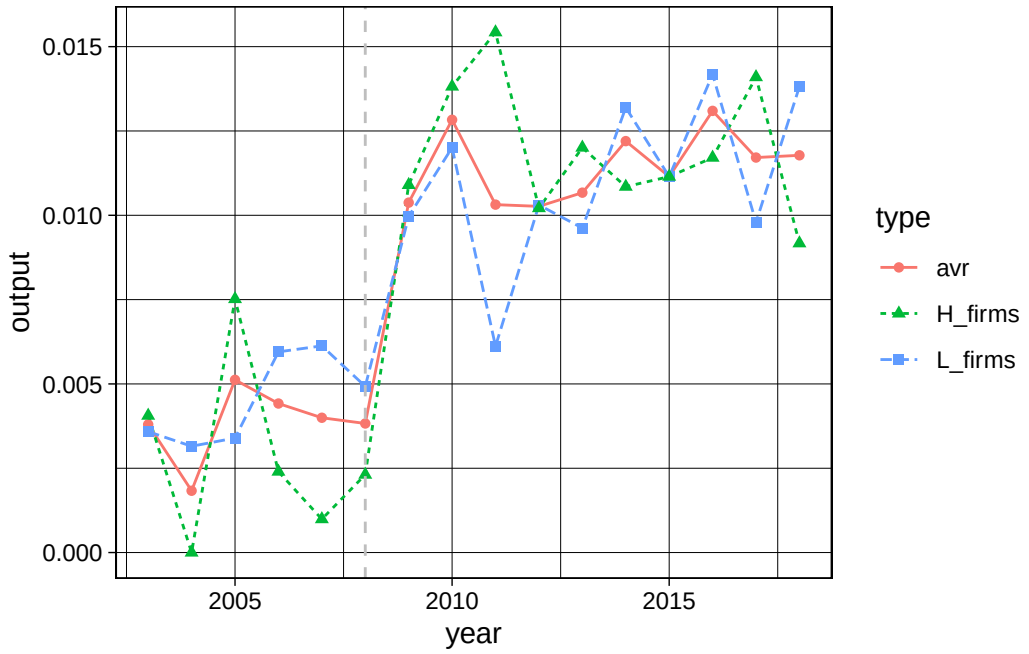
This data is almost ready to use. We just need to make it tidy:

```
firm_tidy <- firm_raw %>%
  pivot_longer(cols = -"year", names_to = "type", values_to = "output")
head(firm_tidy)
```

```
# A tibble: 6 x 3
   year type     output
  <dbl> <chr>     <dbl>
1  2003 avr     0.00380
2  2003 L_firms 0.00359
3  2003 H_firms 0.00406
4  2004 avr     0.00183
5  2004 L_firms 0.00315
6  2004 H_firms 0
```

Now we can have a first preview of the plot:

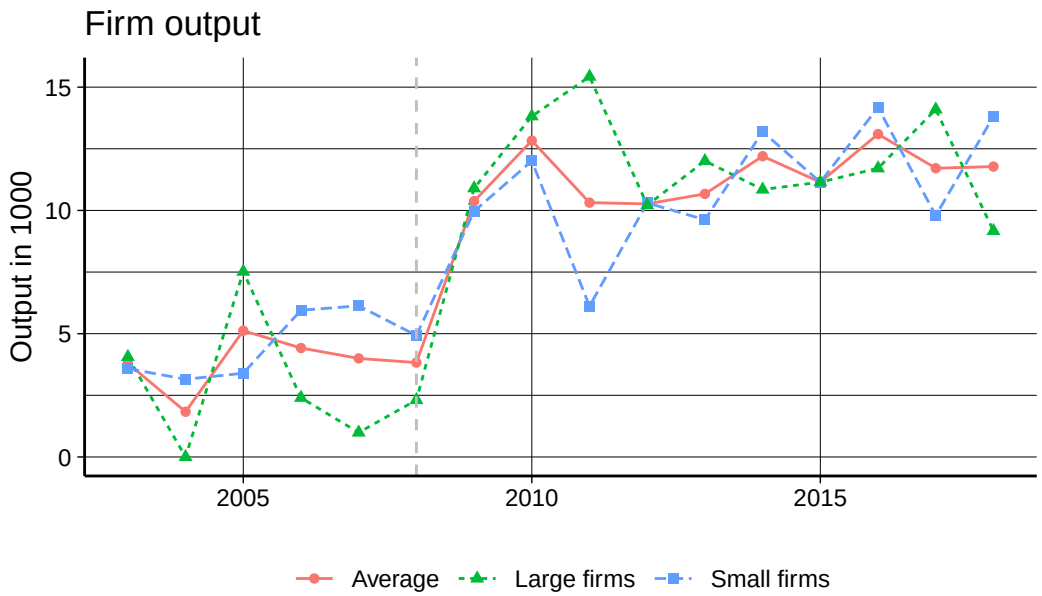
```
ggplot(
  data = firm_tidy,
  mapping = aes(
    x=year, y=output, color=type,
    linetype = type, shape=type)
  ) +
  geom_line() + geom_point() +
  geom_vline(xintercept = 2008, linetype = "dashed", color = "grey") +
  theme_linedraw()
```



This version here has received some visual refinements, and we changed the names for the firms beforehand to make the legend more clear:

```
firm_plot <- firm_tidy %>%
  mutate(type=case_match(
    type,
    "avr" ~ "Average",
    "L_firms" ~ "Small firms",
    "H_firms" ~ "Large firms"
  )) %>%
  ggplot(
    data = .,
    mapping = aes(
      x=year, y=output, color=type,
      linetype = type, shape=type)
    ) +
  geom_line() + geom_point() +
  geom_vline(xintercept = 2008, linetype = "dashed", color = "grey") +
  scale_y_continuous(labels = label_number(scale = 1000)) +
  labs(y = "Output in 1000", title = "Firm output",
       caption = "Data: FCEFD (2024).") +
  theme_linedraw() +
  theme(
```

```
    legend.position = "bottom",
    legend.title = element_blank(),
    panel.grid.minor.x = element_blank(),
    axis.title.x = element_blank(),
    panel.border = element_blank(),
    axis.line = element_line(color="black"))
firm_plot
```



Data: FCEFD (2024).

And, finally, save the plot:

```
ggsave(
  plot = firm_plot,
  filename = here("material/Day3-Line-plot.pdf"),
  width = 6, height = 4)
```