SPEC REU R Resources: Visualizing Regression Results with dot-and-whisker Plots – Homework

Yeiyoung Choo and Claudia Salas Gimenez

Summer 2024

In this homework, we will practice generating dot-and-whisker plots using regression results from tidy dataframes. These plots will visually represent the coefficient estimates and confidence intervals from our regression models.

As the final module of our online training, this assignment also aims to refine your skills in data management, regression analysis, and data visualization, equipping you with the tools needed for publication-ready quantitative analysis.

Initial Setup

Begin by setting your working directory to the location of your data files, and load the necessary libraries and dataset. We'll be working with the `.rds` file you saved in the groupwork assignment (`reg_aslaksen2010.rds`).

Note that for this module will be using the `dotwhisker` package, and as of April 2024, we need to install its dependencies `prediction` and `margins` from GitHub in order to load the `dotwhisker` package.

Exercise 1: Run Linear Regressions

Let's run the regression models previously specified in the groupwork assignment and store them as objects (`n1`, `n2` and `n3`).

- **Model 1 (n1):** This model predicts political rights in the upcoming year (`pr_lead`) using the predictors `pr` (political rights) and `oilshare` (share of oil in income).
- **Model 2 (n2):** Expanding on Model n1, this multiple linear regression includes additional predictors: `lrgdppc` (logged GDP per capita), `lpop` (logged population), and `educ` (education level).
- **Model 3 (n3):** Further extending Model n2, this model also incorporates `open` (openness to trade) to assess a broader range of factors influencing political rights.

Check the results for each model using the `summary()` function.

Exercise 2: Generate a Dot-and-Whisker Plot Using `ggplot2`

Use object `n2` to create a dot-and-whisker plot using the `ggplot2` package. Before you start plotting, make sure to create a dataframe to store the regression results and calculate the 95% confidence interval for each coefficient.

Exercise 3: Generate a Dot-and-Whisker Plot Using `dwplot()`

Now, create of a dot-and-whisker plot using the `dotwhisker` package with `n3` as input. Remember to tidy the dataframe before generating the dot-and-whisker plot.

Exercise 4: Create a Dot-and-Whisker Plot for Three Models with `dwplot()`

For this last exercise, create a single dot-and-whisker plot to visualize results from multiple regression models (`n1`, `n2` and `n3`), specifically focusing on the variables `oilshare`, `lpop`, and `educ`. Remember to tidy up the models `n1`, `n2`, and `n3`, filter for the variables `oilshare`, `lpop`, and `educ`, and label each model with a new column called `Model (model number)` before combining the dataframes with `rbind()`. Then, combine these dataframes using `rbind()`. Also, customize the plot to generate a clear and visually appealing graph. Save your final plot as a `.png` file using `ggsave()`.

Helpful Hint: The `ggsave()` structure is: `ggsave("your filepath/your filename.png"...)`