Day 3 Pset

Cecilia Sui and all other TAs

Day 3 Outline:

1. Packages and Help Pages
2. Data Wrangling Part 2
3. R markdown Tips

dplyr Basics

We will use the resources dataset to do some exercises with dplyr

1. Select the first 10 rows and last 3 rows. Output the rows together as the same dataframe.

2. Select all rows where Year is 1951 and regime is non-negative.

3. Order the rows by Year in descending order. Break ties using the column oil in ascending order.

4. Select the 100th to 110th rows, and find the row with the highest value for metal.

5. Randomly select 10 rows with replacement, and standardize the columns oil and metal.

6. Select columns: cty_name, year, regime, oil, illit, and life and store them into a new dataframe.

7. Add a new column named new_oil using where the values are the product of standardized oil and regime type.

8. Calculate the **average** oil level for **each year** from 1966 to 1997.

9. Calculate the **median** metal value for **each** regime type from -10 to +10.