**Task 1: (1.2)**

1. Open the general help and look at the Introduction to R section.
2. Install the vcd package.
3. Load the package and read the description of the dataset Arthritis.
4. List the available dataset.
5. Print out the Arthritis dataset.
6. Run the example that comes with the Arthritis dataset.

**Task 2: (3.1)**

Let’s summarize the dataframe **energy** by calculating max, min, and average values of **ele\_total** [Access to electricity (% of total population)] in each year.

**Task 3: (4.1)**

* Let’s left\_join dataframes **energy** and **classf** by common column **ccode**.
* Summarize by calculating average values of **ele\_total** [Access to electricity (% of total population)] for each **year** and **country** **group** (i.e., H, UM, LM, L).
* Save the summarized dataframe as **wb\_energy**.

**Task 4: (4.1)**

* Load **"data/008-nlfs2.dta"** dataset and store it in **nlfs2** dataframe.
* From **nlfs2,** create a dataframe of family size named **fsize.**
* Merge dataframes **nlfs2** and **fsize**. Then replace the **nlfs2** dataframe with the merged dataframe.
* Keep columns **psu, hhid, family\_size, q10, q09.**
* Rename **q10** to **age**, **q09** to **gender**.

**Task 5: (6.3)**

Suppose there is no built-in function in R to calculate mean and standard deviation. Write a user defined functions ***func\_mean*** and ***func\_sd*** to calculate mean and standard deviation of a given vector.

**Task 6 : (7.1)**

Web-scrape the Historical ranking table from <https://en.wikipedia.org/wiki/ICC_Men%27s_T20I_Team_Rankings>

**Task 7: (8.4)**

Using NMICS6 data (009-hl.sav), conduct a hypothesis test whether average age between male and female is statistically different.

**Task 8: (10.1)**

Using 011-Affairs.RData, complete the following tasks.

1. Load the **011-Affairs.RData**
2. Tabulate the frequency of **affairs** variable from **Affairs** dataframe.
3. Create a variable **ynaffairs** in **Affairs** dataframe such that the variable takes value 0 if no affairs and 1 if the person is involved in affairs.
4. Set **ynaffairs** and **rating** variables as factor variables.
5. Set ‘0’ as reference for **ynaffairs** variable, ‘5’ for **rating**, ‘no’ for **children**, and ‘female’ for **gender** variables.
6. Fit a logistic regression model with **ynaffairs** as dependent variable and **gender, age, yearsmarried, children, rating** as independent variable.
7. Calculate average marginal effect for each variables using the margins() function.