Was a URL

Analysis Steps

- Downloaded file from url
    - Note due to folder / file length issue, renames file “house_power.txt”
    - No other edits to the file made other than viewing in text editor
- Opened in Text Editor
    - Verified that there were no commas in the file
    - Verified it was semi-colon delineated
    - Verified Rows
        - 2,075,261
    - Verified Data Range
        - 12/16/2006 - 11/26/2010
- Noted from question
    - Desired Date Range
        - 2/1/2007 – 2/2/2007
- In R Studio
    - Did Date Diff on Start Date / End Date
        - 1,441 days
    - Did estimate of data points per day
        - 2,075,261 / 1,441 ~= 1,440
    - Did estimate of how many readings per day per problem
        - matches 60 * 24 for per minute measurements ~= 1,440
    - Estimated per R calcs
        - Pre-Subset ~= 67,687 rows
        - Subset ~= 2,880 rows
        - Post-Subset ~= 2,006,133 rows
    - Reading in subset (using above estimates as a guide / starting point and trial / error to get exact range
        - skip = 67,636 rows, with nrows = 2,880 rows
- Used read.table to read into R Studio
    - data.housePower <- read.table(
      file = "house_power.txt", header = TRUE, sep = ";", na.strings = "?",
      skip = 66636, nrows = 2880,
      col.names = c("ReadDate", "ReadTime", "Global_active_power",
      "Global_reactive_power", "Voltage", "Global_intensity", "Sub_metering_1",
      "Sub_metering_2", "Sub_metering_3")
)
- viewed the resulting data.table
    - double-check the resulting data to be processed
- added columns with Date & Time format respectively to data.table per assignment recommendations