SHETH L.U.J AND SIR M.V. COLLEGE

SUBJECT NAME: **Data Analysis with SAS / SPSS /R**

Practical No. 14

Aim: Extracting date components using lubridate: functions (R).

Output:

```
> library(lubridate)
> library(dplyr)
>
> # ================================================================================
> # 1. SETUP: Import Data (Manual Selection)
> # ================================================================================
>
> # Select your 'CL_in_15_minute_new.csv' file
> print("--- ACTION: Please select your FINANCE CSV file ---")
[1] "--- ACTION: Please select your FINANCE CSV file ---"
> finance_df <- read.csv(file.choose())
>
> print("--- 1. Raw Data (First few rows) ---")
[1] "--- 1. Raw Data (First few rows) ---"
> print(head(finance_df$datetime))
[1] "2025-07-31 22:00:00" "2025-07-31 22:15:00" "2025-07-31 22:30:00" "2025-07-31 22:45:00"
[5] "2025-07-31 23:00:00" "2025-07-31 23:15:00"
>
> # ================================================================================
> # 2. PARSE AND EXTRACT
> # ================================================================================
>
> processed_data <- finance_df %>%
+   mutate(
+     # A. Parsing: Convert text to a real Date-Time object
+     Actual_Time = ymd_hms(datetime),
+
+     # B. Extraction Functions
+     Year_Num     = year(Actual_Time),
+     Month_Name   = month(Actual_Time, label = TRUE),
+     Day_Num      = day(Actual_Time),
+     Weekday_Name = wday(Actual_Time, label = TRUE, abbr = FALSE),
+
+     # C. Time Specific Extractions (Crucial for 15-min data)
+     Hour_Num     = hour(Actual_Time),
+     Minute_Num   = minute(Actual_Time)
+   ) %>%
+   # Let's select just the new columns to verify
+   select(datetime, Actual_Time, Year_Num, Weekday_Name, Hour_Num, Minute_Num)
>
> print("--- 2. Data with Extracted Components ---")
[1] "--- 2. Data with Extracted Components ---"
> print(head(processed_data))
             datetime         Actual_Time Year_Num Weekday_Name Hour_Num Minute_Num
1 2025-07-31 22:00:00 2025-07-31 22:00:00     2025     Thursday       22          0
2 2025-07-31 22:15:00 2025-07-31 22:15:00     2025     Thursday       22         15
3 2025-07-31 22:30:00 2025-07-31 22:30:00     2025     Thursday       22         30
4 2025-07-31 22:45:00 2025-07-31 22:45:00     2025     Thursday       22         45
5 2025-07-31 23:00:00 2025-07-31 23:00:00     2025     Thursday       23          0
6 2025-07-31 23:15:00 2025-07-31 23:15:00     2025     Thursday       23         15
>
> # ================================================================================
> # 3. ANALYSIS EXAMPLE: Count Trades by Hour
> # ================================================================================
>
>
> hourly_activity <- processed_data %>%
+   group_by(Hour_Num) %>%
+   count()
>
> print("--- 3. Activity by Hour ---")
[1] "--- 3. Activity by Hour ---"
> print(hourly_activity)
# A tibble: 23 x 2
# Groups:   Hour_Num [23]
   Hour_Num     n
      <int> <int>
 1        0   264
 2        1   264
 3        2   264
 4        3   264
 5        4   264
 6        5   264
 7        6   264
 8        7   264
 9        8   264
10        9   264
```