

**SHETH L.U.J. & SIR M.V. COLLEGE OF SCIENCE**  
**SUBJECT - Data Analysis with SAS / SPSS / R**

```
> # Load the dataset
> tips <- read_csv("tips.csv")
Rows: 244 Columns: 7
— Column specification —
Delimiter: ","
chr (4): sex, smoker, day, time
dbl (3): total_bill, tip, size

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
> # Inspect dataset
> head(tips)
# A tibble: 6 x 7
  total_bill tip sex    smoker day    time    size
    <dbl> <dbl> <chr> <chr> <chr> <chr> <dbl>
1     17.0  1.01 Female No     Sun    Dinner     2
2     10.3  1.66 Male   No     Sun    Dinner     3
3     21.0  3.5  Male   No     Sun    Dinner     3
4     23.7  3.31 Male   No     Sun    Dinner     2
5     24.6  3.61 Female No     Sun    Dinner     4
6     25.3  4.71 Male   No     Sun    Dinner     4
> glimpse(tips)
Rows: 244
Columns: 7
$ total_bill <dbl> 16.99, 10.34, 21.01, 23.68, 24.59, 25.29, 8.77, 26.88, 15...
$ tip        <dbl> 1.01, 1.66, 3.50, 3.31, 3.61, 4.71, 2.00, 3.12, 1.96, 3.2...
$ sex        <chr> "Female", "Male", "Male", "Male", "Female", "Male", "Male...
$ smoker     <chr> "No", "No", "No", "No", "No", "No", "No", "No", "No", "No...
$ day        <chr> "Sun", "Sun", "Sun", "Sun", "Sun", "Sun", "Sun", "Sun", "Sun", "...
$ time       <chr> "Dinner", "Dinner", "Dinner", "Dinner", "Dinner", "Dinner", "Dinner"
```

**SHETH L.U.J. & SIR M.V. COLLEGE OF SCIENCE**  
**SUBJECT - Data Analysis with SAS / SPSS / R**

```
> # Example 1: Single condition
> # Filter for bills greater than $30
> high_bill_subset <- subset(tips, total_bill > 30)
> cat("Number of large bills (total_bill > 30):", nrow(high_bill_subset), "\n")
Number of large bills (total_bill > 30): 32
> summary(high_bill_subset$total_bill)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 30.06  32.61   35.05   37.36  40.71   50.81
> # Example 2: Multiple conditions (AND)
> # Bills > 30 AND table size > 3 people
> large_group_large_bill <- subset(tips, total_bill > 30 & size > 3)
> cat("Large bill + large group:", nrow(large_group_large_bill), "\n")
Large bill + large group: 19
> head(large_group_large_bill)
# A tibble: 6 x 7
  total_bill tip sex  smoker day   time size
    <dbl> <dbl> <chr> <chr> <chr> <chr> <dbl>
1     35.3   5 Female No    Sun  Dinner     4
2     39.4  7.58 Male  No    Sat  Dinner     4
3     30.4   5.6 Male  No    Sun  Dinner     4
4     32.4   6   Male  No    Sun  Dinner     4
5     34.8   5.2 Female No    Sun  Dinner     4
6     38.0   3   Male  Yes   Sat  Dinner     4
> # Example 3: OR condition
> # Customers who are smokers OR gave a tip greater than $6
> smoker_or_big_tip <- subset(tips, smoker == "Yes" | tip > 6)
> cat("Smokers or big tippers:", nrow(smoker_or_big_tip), "\n")
Smokers or big tippers: 97
> head(smoker_or_big_tip)
# A tibble: 6 x 7
```

**SHETH L.U.J. & SIR M.V. COLLEGE OF SCIENCE**  
**SUBJECT - Data Analysis with SAS / SPSS / R**

```
3.070 7.510 8.510 8.083 9.550 9.940
> # Example 2: Filter females who are non-smokers
> female_non_smokers <- tips |>
+   filter(sex == "Female", smoker == "No")
> cat("Female non-smokers:", nrow(female_non_smokers), "\n")
Female non-smokers: 54
> head(female_non_smokers)
# A tibble: 6 x 7
  total_bill tip sex smoker day time size
  <dbl> <dbl> <chr> <chr> <chr> <chr> <dbl>
1    17.0  1.01 Female No    Sun Dinner 2
2    24.6  3.61 Female No    Sun Dinner 4
3    35.3  5     Female No    Sun Dinner 4
4    14.8  3.02 Female No    Sun Dinner 2
5    10.3  1.67 Female No    Sun Dinner 3
6    17.0  3.5   Female No    Sun Dinner 3
> # Example 3: Filter by set of days using %in%
> # Only Friday and Saturday entries
> fri_sat_tips <- tips |>
+   filter(day %in% c("Fri", "Sat"))
> cat("Tips on Fri or Sat:", nrow(fri_sat_tips), "\n")
Tips on Fri or Sat: 106
> table(fri_sat_tips$day)

Fri Sat
19  87
>
>
> |
```

**SHETH L.U.J. & SIR M.V. COLLEGE OF SCIENCE**  
**SUBJECT - Data Analysis with SAS / SPSS / R**

```

total_bill  tip sex  smoker day  time  size
<dbl> <dbl> <chr> <chr> <chr> <chr> <dbl>
1      39.4  7.58 Male  No    Sat   Dinner    4
2      38.0   3.00 Male  Yes   Sat   Dinner    4
3      11.2  1.76 Male  Yes   Sat   Dinner    2
4      48.3  6.73 Male  No    Sat   Dinner    4
5      20.3  3.21 Male  Yes   Sat   Dinner    2
6      13.8   2.00 Male  Yes   Sat   Dinner    2
> # Example 1: Filter bills under $10
> small_bills <- tips |>
+   filter(total_bill < 10)
> cat("Number of small bills (< 10):", nrow(small_bills), "\n")
Number of small bills (< 10): 17
> summary(small_bills$total_bill)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 3.070   7.510   8.510   8.083   9.550   9.940
> # Example 2: Filter females who are non-smokers
> female_non_smokers <- tips |>
+   filter(sex == "Female", smoker == "No")
> cat("Female non-smokers:", nrow(female_non_smokers), "\n")
Female non-smokers: 54
> head(female_non_smokers)
# A tibble: 6 x 7
  total_bill  tip sex  smoker day  time  size
    <dbl> <dbl> <chr> <chr> <chr> <chr> <dbl>
1     17.0  1.01 Female No    Sun   Dinner    2
2     24.6  3.61 Female No    Sun   Dinner    4
3     35.3   5.00 Female No    Sun   Dinner    4
4     14.8  3.02 Female No    Sun   Dinner    2
5     10.3  1.67 Female No    Sun   Dinner    2

```