Subsetting tables in R

tibble() as_tibble()

```
#tibble:简单的 dataframe  更加易于打印
```

select()

```
 df <- data.frame(
   ID = c(1, 2, 3),
   Name = c("Alice", "Bob", "Charlie"),
   Age = c(25, 30, 22)
)

# 选择其中的列
selected_df <- select(df, ID, Age)
selected_df1<-select(df, NewID=ID, Newage=Age)
```

relocate()

```
# 示例数据框
df <- data.frame(
   ID = c(1, 2, 3),
   Name = c("Alice", "Bob", "Charlie"),
   Age = c(25, 30, 22)
)

# 重新排列列的顺序
relocated_df <- relocate(df, c("Age", "ID", "Name"))
```

rename()

```
data <- data.frame(
   A = c(1, 2, 3),
   B = c(4, 5, 6),
   C = c(7, 8, 9)
)

# 使用 rename() 修改列名
data <- rename(data, new_A = A, new_B = B)
```

filter()

```
# 使用 filter() 筛选符合条件的行
filtered_data <- filter(data, Age > 25)
```

slice()

```
# 使用 slice() 选择特定的行
selected_rows <- slice(data, 2:3)
```

Modifying tables in R

mutate()

```
# 使用 mutate() 创建新的列或修改现有列
modified_data <- mutate(data, BMI = Weight / ((Height/100)^2))
```

transmute()

```
# 使用 transmute() 创建新的列
new_data <- transmute(data, BMI = Weight / ((Height/100)^2))
仅返回 BMI 列
arrange()
# 使用 arrange() 对数据框的行进行排序
sorted_data <- arrange(data, Age)
默认升序排列 如果改为降序则 desc(Age)
group_by()和 summarize()
grouped_data <- group_by(data, Gender)
# 使用 summarize() 对分组后的数据进行汇总
summary_data <- summarize(grouped_data,
                          Mean_Age = mean(Age, na.rm = TRUE),
                          Median_Height = median(Height, na.rm = TRUE))
# A tibble: 2 × 3
  Gender Mean_Age Median_Height
  <chr>     <dbl>         <dbl>
1 Female     25            162.5
2 Male       27.3          175
```

```
# 创建一个示例数据框
df <- data.frame(
  group = c("A", "A", "B", "B", "B"),
  value = c(1, 2, 3, 4, 5)
)

# 按照group变量进行分组
df_grouped <- group_by(df, group)

# 对每个组计算平均值
df_avg <- summarise(df_grouped, avg_value = mean(value))

# A tibble: 2 x 2
  group avg_value
  <chr>     <dbl>
1 A           1.5
2 B           4.0
```