Summarize and Group Data

Step 1: Load your data.

```
library(tidyverse)
library(utf8)
library(ellipsis)
Credit <- read.csv("credit_cards.csv")
```

Step 2: Calculate the income range for students versus non-students.

First, use the group_by() function to group the Credit data set according to whether an individual is a student. Then, create a summary statistic with the summarise() function. Here, the summary statistic is IncomeRange = max(Income) - min(Income):

```
Credit %>%
  group_by(Student) %>%
  summarise(IncomeRange = max(Income)- min(Income))
```

```
## # A tibble: 2 x 2
##   Student IncomeRange
##   <chr>         <dbl>
## 1 No             176.
## 2 Yes            170.
```

Note that if the pipe operators were not being used here, you'd need to include the data set as the first argument in these functions.

Step 3: Calculate the income range for groups of multiple variables.

The setup to use group_by() and summarise() is the same as the previous step, but instead of only summarizing by students, you are also going to summarize by married people. You can do that by including both Student and Married in the group_by() function:

```
Credit %>%
 group_by(Student, Married) %>%
 summarise(MaxLimit = max(Limit))
```

```
## `summarise()` has grouped output by 'Student'. You can override using the
## `.groups` argument.
```

```
## # A tibble: 4 x 3
## # Groups:   Student [2]
##   Student Married MaxLimit
##   <chr>   <chr>      <int>
## 1 No      No         10748
## 2 No      Yes        13913
## 3 Yes     No          9560
## 4 Yes     Yes         9310
```