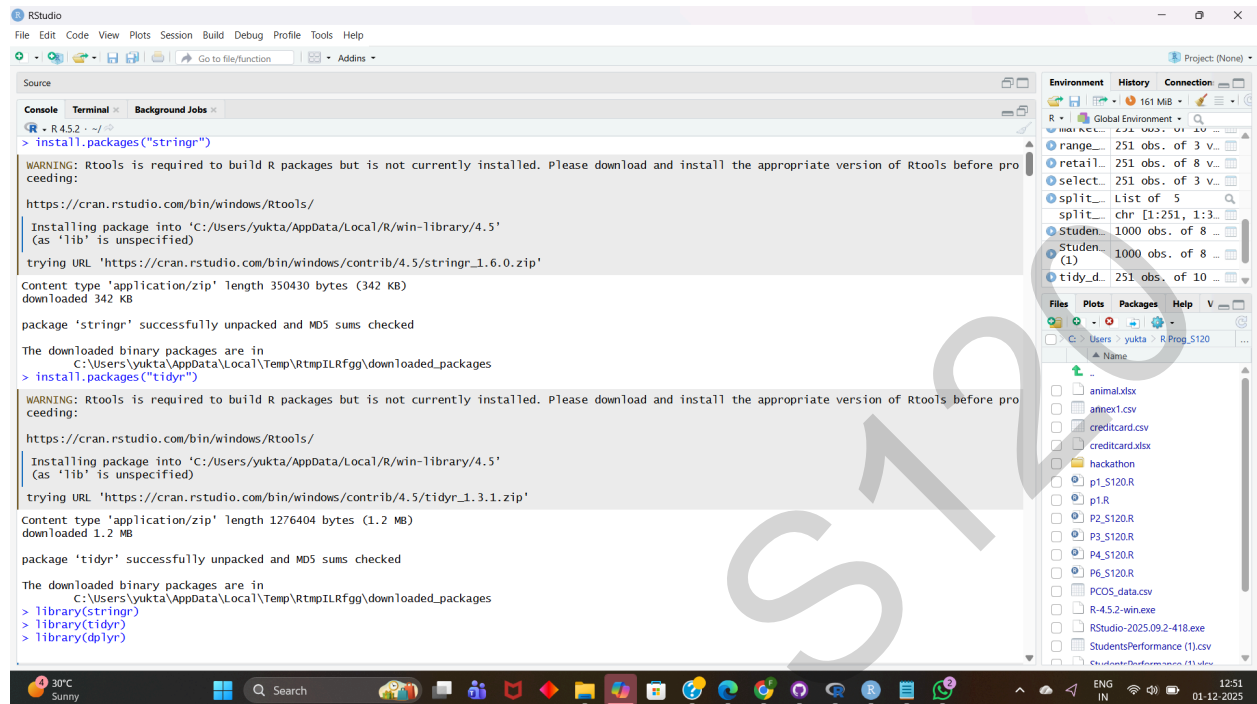


SHETH L.U.J AND SIR M.V COLLEGE

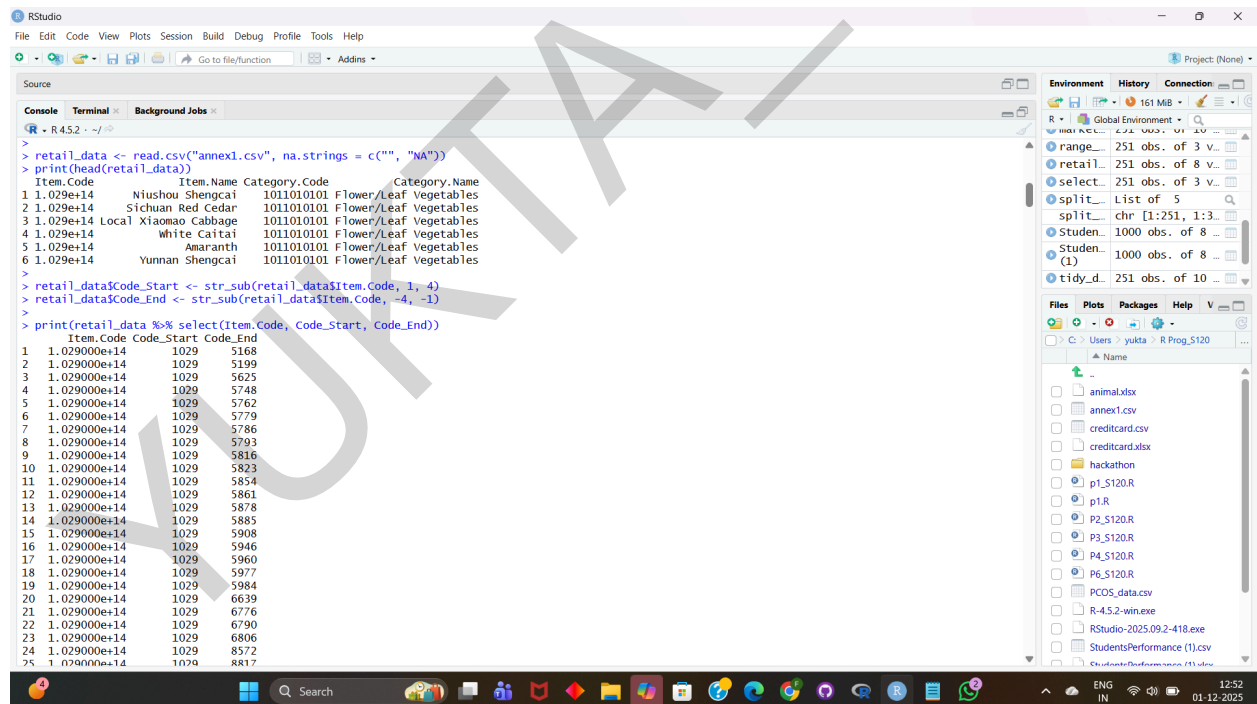
Aim: Performing text manipulation using `str_sub()`, `str_split()` (R). import dataset.



The screenshot shows the RStudio interface with the Console pane active. The user has installed the 'stringr' and 'tidyr' packages. The console output shows the following steps:

```
> install.packages("stringr")
WARNING: Rtools is required to build R packages but is not currently installed. Please download and install the appropriate version of Rtools before proceeding:
https://cran.rstudio.com/bin/windows/Rtools/
Installing package into 'C:/Users/yukta/AppData/Local/R/win-library/4.5'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.5/stringr_1.6.0.zip'
Content type 'application/zip' length 350430 bytes (342 KB)
downloaded 342 KB
package 'stringr' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
C:/Users/yukta/AppData/Local/Temp/RtmpILRfgg/downloaded_packages
> install.packages("tidyr")
WARNING: Rtools is required to build R packages but is not currently installed. Please download and install the appropriate version of Rtools before proceeding:
https://cran.rstudio.com/bin/windows/Rtools/
Installing package into 'C:/Users/yukta/AppData/Local/R/win-library/4.5'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.5/tidyr_1.3.1.zip'
Content type 'application/zip' length 1276404 bytes (1.2 MB)
downloaded 1.2 MB
package 'tidyr' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
C:/Users/yukta/AppData/Local/Temp/RtmpILRfgg/downloaded_packages
> library(stringr)
> library(tidyr)
```

The Environment pane on the right shows the loaded packages: 'stringr' and 'tidyr'.



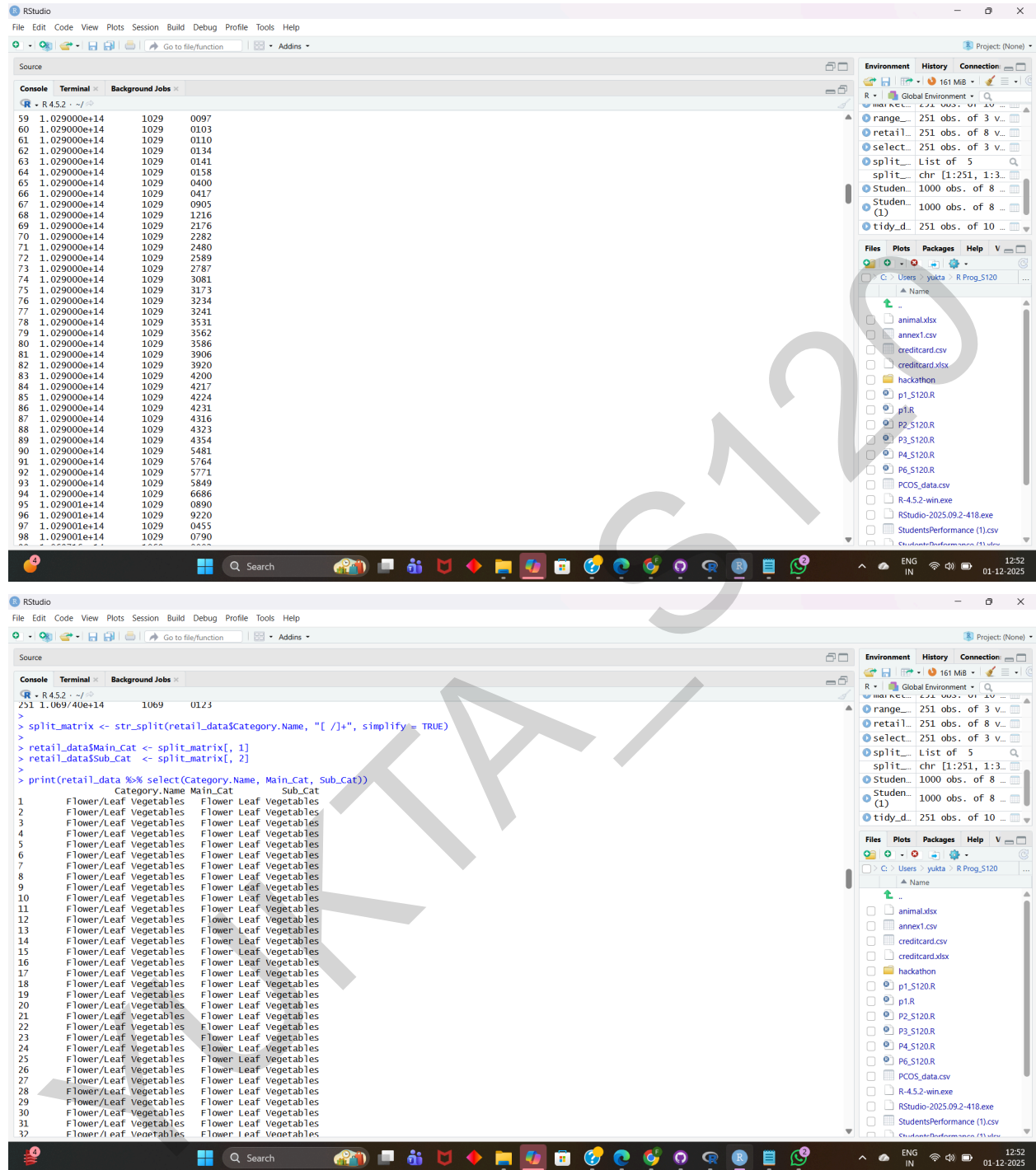
The screenshot shows the RStudio interface with the Console pane active. The user has loaded the 'annex1.csv' dataset and performed text manipulation using `str_sub()` and `str_split()`. The console output shows the following steps:

```
> retail_data <- read.csv("annex1.csv", na.strings = c("", "NA"))
> print(head(retail_data))
  Item.Code Item.Name Category.Code Category.Name
1 1.029e+14 Niuoshou Shengcai 1011010101 Flower/Leaf Vegetables
2 1.029e+14 Sichuan Red Cedar 1011010101 Flower/Leaf Vegetables
3 1.029e+14 Local Xiaomao Cabbage 1011010101 Flower/Leaf Vegetables
4 1.029e+14 White Caitai 1011010101 Flower/Leaf Vegetables
5 1.029e+14 Amaranth 1011010101 Flower/Leaf Vegetables
6 1.029e+14 Yunnan Shengcai 1011010101 Flower/Leaf Vegetables
> retail_data$Code_Start <- str_sub(retail_data$Item.Code, 1, 4)
> retail_data$Code_End <- str_sub(retail_data$Item.Code, -4, -1)
> print(retail_data %>% select(Item.Code, Code_Start, Code_End))
  Item.Code Code_Start Code_End
1 1.029000e+14 1029 5168
2 1.029000e+14 1029 5199
3 1.029000e+14 1029 5625
4 1.029000e+14 1029 5748
5 1.029000e+14 1029 5762
6 1.029000e+14 1029 5779
7 1.029000e+14 1029 5786
8 1.029000e+14 1029 5793
9 1.029000e+14 1029 5816
10 1.029000e+14 1029 5823
11 1.029000e+14 1029 5854
12 1.029000e+14 1029 5861
13 1.029000e+14 1029 5878
14 1.029000e+14 1029 5885
15 1.029000e+14 1029 5908
16 1.029000e+14 1029 5946
17 1.029000e+14 1029 5960
18 1.029000e+14 1029 5977
19 1.029000e+14 1029 5984
20 1.029000e+14 1029 6639
21 1.029000e+14 1029 6776
22 1.029000e+14 1029 6790
23 1.029000e+14 1029 6806
24 1.029000e+14 1029 8572
25 1.029000e+14 1029 8817
```

The Environment pane on the right shows the loaded dataset: 'retail_data'.

YUKTA SONAWANE S120
R PROGRAMMING
PRACTICAL NO: 9

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



YUKTA SONAWANE S120
R PROGRAMMING
PRACTICAL NO: 9

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

The image displays two screenshots of the RStudio interface, demonstrating data manipulation using the tidyverse package.

Top Screenshot: The console shows the initial data structure and the execution of the following code:

```
R - R 4.5.2 ~ / ~  
> tidy_data <- retail_data %>%  
+ separate(Item.Code, into = c("Part1", "Part2"), sep = 4, remove = FALSE)  
> print(tidy_data %>% select(Item.Code, Part1, Part2))
```

The output shows a table with columns Item.Code, Part1, and Part2, containing 32 rows of data.

Bottom Screenshot: The console shows the execution of the following code:

```
R - R 4.5.2 ~ / ~  
> print("Yukta Sonawane S120")  
[1] "Yukta Sonawane S120"  
>
```

The output shows the string "Yukta Sonawane S120".

YUKTA SONAWANE S120
R PROGRAMMING
PRACTICAL NO: 9