STA 326 2.0 Programming and Data Analysis with R *

23 March 2020

Week 1: Answers

Question 1

We are in the midst of a medical crisis! The deadly corona virus that originated in China has infected hundreds of people and is now spreading across the globe at an alarming rate. World Health Organization (WHO) alerted the world about the Novel Corona virus(2019-nCoV) in January, 2020. After issuance of the global alert, a formal reporting of Corona cases was put in place, and WHO published daily reports on the number of cases on their website here. Use WHO: Situation Report-21 for this question.

Table 1 reports the confirmed cases of 2019-nCoV reported by provinces, regions and cities in China (see slide 47).

   i) Enter confirmed cases in table 1 to a vector.

```
confirmed_cases <- c(29631, 1151, 1104, 1073, 879, 830, 771, 492, 468, 459, 405,
                     337, 331, 295, 261, 218, 213, 210, 141, 136, 119, 109, 107,
                     91, 85, 80, 58, 49, 49, 36, 18, 18, 10, 1)
sum(confirmed_cases) # This is to check if the Total match with the Total given in Table 1
```

```
[1] 40235
```

```
confirmed_cases
```

```
 [1] 29631  1151  1104  1073   879   830   771   492   468   459   405   337
[13]   331   295   261   218   213   210   141   136   119   109   107    91
[25]    85    80    58    49    49    36    18    18    10     1
```

   ii) Name the elements by province/regions/cities in China.

```
names(confirmed_cases) <- c("Hubei", "Guangdong", "Zhejiang", "Henan", "Hunan", "Anhui",
                            "Jiangxi", "Jiangsu", "Chongqing", "Shandong", "Sichuan",
                            "Beijing", "Heilongjiang", "Shanghai", "Fujian", "Hebei",
                            "Shaanxi", "Guangxi", "Yunnan", "Hainan", "Shanxi",
                            "Guizhou", "Liaoning", "Tianjin", "Gansu", "Jilin",
                            "Inner Mongolia", "Ningxia",
                            "Xinjiang", "Hong Kong SAR", "Qinghai",
                            "Taipei and environs",
                            "Macao SAR", "Xizang")
```

   iii) Write R codes to answer the following questions.

     • Which province/region/city has the highest number of confirmed cases?

```
# Method 1
which.max(confirmed_cases) # Returns the name of the district and index in the confirmed cases
```

```
Hubei
    1
```

```r
# Method 2
confirmed_cases[confirmed_cases == max(confirmed_cases)] # Returns the name of the district
```

```
Hubei
29631
```

If it is difficult to understand how the method 2 work, here is the step by step illustration of Method 2.

```r
max(confirmed_cases)
```

```
[1] 29631
```

```r
confirmed_cases == max(confirmed_cases)
```

Hubei	Guangdong	Zhejiang	Henan
TRUE	FALSE	FALSE	FALSE
Hunan	Anhui	Jiangxi	Jiangsu
FALSE	FALSE	FALSE	FALSE
Chongqing	Shandong	Sichuan	Beijing
FALSE	FALSE	FALSE	FALSE
Heilongjiang	Shanghai	Fujian	Hebei
FALSE	FALSE	FALSE	FALSE
Shaanxi	Guangxi	Yunnan	Hainan
FALSE	FALSE	FALSE	FALSE
Shanxi	Guizhou	Liaoning	Tianjin
FALSE	FALSE	FALSE	FALSE
Gansu	Jilin	Inner Mongolia	Ningxia
FALSE	FALSE	FALSE	FALSE
Xinjiang	Hong Kong SAR	Qinghai	Taipei and environs
FALSE	FALSE	FALSE	FALSE
Macao SAR	Xizang		
FALSE	FALSE		

```r
confirmed_cases[confirmed_cases == max(confirmed_cases)]
```

```
Hubei
29631
```

If you do not understand a big line of code, it is always a good idea to break it down to small pieces and check the output.

- Number of confirmed cases reported in Hebei, China.

```r
confirmed_cases["Hebei"]
```

```
Hebei
  218
```

- Total number of confirmed cases reported in China

```r
sum(confirmed_cases)
```

```
[1] 40235
```

- Number of cases reported in the capital of China

```r
confirmed_cases["Beijing"]
```

```
Beijing
    337
```

- Number of cases reported in Inner Mongolia

```
confirmed_cases["Inner Mongolia"]
```

```
Inner Mongolia
           58
```

Question 2

Table 2 reports the confirmed 2019-nCoV cases and deaths in China, Singapore, Republic of Korea, Japan, Malaysia, Australia, Viet Nam, Philippines, Cambodia, Thailand, India, Nepal, Sri Lanka, United States of America, Canada, Germany, France, The United Kingdom, Italy, Russian Federation, Spain , Belgium, Finland, Sweden, UAE as a <- c(40235, 43, 27, 26, 18, 15, 14, 3, 1, 32, 3, 1, 1, 12, 7, 14, 11, 4, 3, 2, 2, 1, 1, 100, 7).

Table 2-please see slide 50.

   i) Rename the vector a as confirmed_cases_countries

```
a <- c(40235, 43, 27, 26, 18, 15, 14, 3, 1, 32, 3, 1, 1, 12, 7, 14, 11, 4, 3, 2, 2, 1, 1, 100, 7)
confirmed_cases_countries <- a
confirmed_cases_countries
```

```
 [1] 40235    43    27    26    18    15    14     3     1    32     3     1
[13]     1    12     7    14    11     4     3     2     2     1     1   100
[25]     7
```

   ii) Name elements according to the associated country

```
names(confirmed_cases_countries) <- c("China", "Singapore", "Republic of Korea", "Japan",
                                      "Malaysia", "Australia", "Viet Nam", "Philippines",
                                      "Cambodia", "Thailand", "India", "Nepal", "Sri Lanka",
                                      "United States of America", "Canada", "Germany", "France",
                                      "The United Kingdom", "Italy", "Russian Federation", "Spain",
                                      "Belgium", "Finland", "Sweden", "UAE")
confirmed_cases_countries
```

```
                   China                Singapore        Republic of Korea
                   40235                       43                       27
                   Japan                 Malaysia                Australia
                      26                       18                       15
                Viet Nam              Philippines                 Cambodia
                      14                        3                        1
                Thailand                    India                    Nepal
                      32                        3                        1
               Sri Lanka United States of America                   Canada
                       1                       12                        7
                 Germany                   France       The United Kingdom
                      14                       11                        4
                   Italy       Russian Federation                    Spain
                       3                        2                        2
                 Belgium                  Finland                   Sweden
                       1                        1                      100
                     UAE
                       7
```

   iii) Mistakenly 100 cases were recorded to Sweden, correct it.

```
confirmed_cases_countries["Sweden"] <- 1
# Let's check the vector again. Now Sweden reads as 1.
confirmed_cases_countries
```

```
                   China                Singapore        Republic of Korea
```

3

```
     40235                      43                        27
     Japan                   Malaysia                 Australia
        26                      18                        15
  Viet Nam                 Philippines                Cambodia
        14                       3                         1
  Thailand                    India                      Nepal
        32                       3                         1
 Sri Lanka United States of America              Canada
         1                      12                         7
   Germany                    France       The United Kingdom
        14                      11                         4
     Italy       Russian Federation                    Spain
         3                       2                         2
   Belgium                   Finland                   Sweden
         1                       1                         1
       UAE
         7
```

iv) Add the record of "other" category into your vector.

```r
confirmed_cases_countries <- c(confirmed_cases_countries, 70)
names(confirmed_cases_countries)[length(confirmed_cases_countries)] <-
  "International conveyance (Japan)"
confirmed_cases_countries
```

```
                    China                     Singapore
                    40235                            43
        Republic of Korea                         Japan
                       27                            26
                 Malaysia                     Australia
                       18                            15
                 Viet Nam                   Philippines
                       14                             3
                 Cambodia                      Thailand
                        1                            32
                    India                         Nepal
                        3                             1
                Sri Lanka      United States of America
                        1                            12
                   Canada                       Germany
                        7                            14
                   France            The United Kingdom
                       11                             4
                    Italy            Russian Federation
                        3                             2
                    Spain                       Belgium
                        2                             1
                  Finland                        Sweden
                        1                             1
                      UAE International conveyance (Japan)
                        7                            70
```

v) Create a new vector to enter WHO regions

```r
WHO_Region <- c(rep("Western Pacific Region", 9),
           rep("South-East Asia Region", 4),
           rep("Region of Americas", 2),
           rep("Europe Region", 9),
           rep("Eastern Mediterranean Region", 1),
```

```
                  rep("Other", 1))
WHO_Region
```

```
 [1] "Western Pacific Region"       "Western Pacific Region"
 [3] "Western Pacific Region"       "Western Pacific Region"
 [5] "Western Pacific Region"       "Western Pacific Region"
 [7] "Western Pacific Region"       "Western Pacific Region"
 [9] "Western Pacific Region"       "South-East Asia Region"
[11] "South-East Asia Region"       "South-East Asia Region"
[13] "South-East Asia Region"       "Region of Americas"
[15] "Region of Americas"           "Europe Region"
[17] "Europe Region"                "Europe Region"
[19] "Europe Region"                "Europe Region"
[21] "Europe Region"                "Europe Region"
[23] "Europe Region"                "Europe Region"
[25] "Eastern Mediterranean Region" "Other"
```

vi) New cases have been reported in China, Singapore, Malaysia, The United Kingdom, Spain. Create a new vector to code these countries as TRUE and the rest as FALSE

```
country_names_new_cases <- c("China", "Singapore", "Malaysia", "The United Kingdom", "Spain")
new_cases <- names(confirmed_cases_countries) %in% country_names_new_cases
new_cases
```

```
 [1]  TRUE  TRUE FALSE FALSE  TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
[13] FALSE FALSE FALSE FALSE FALSE  TRUE FALSE FALSE  TRUE FALSE FALSE FALSE
[25] FALSE FALSE
```