

# Untitled

*Xun Zhang*

*12/2/2019*

```
#install.packages("readr")  
require(readr)
```

```
## Loading required package: readr
```

```
a = read_csv("population density.csv")
```

```
## Parsed with column specification:  
## cols(  
##   .default = col_double(),  
##   `Country Name` = col_character(),  
##   `Country Code` = col_character(),  
##   `Indicator Name` = col_character(),  
##   `Indicator Code` = col_character(),  
##   `1960` = col_logical(),  
##   `2019` = col_logical()  
## )  
  
## See spec(...) for full column specifications.
```

```
b = read_csv("secure internet server.csv")
```

```
## Parsed with column specification:  
## cols(  
##   .default = col_double(),  
##   `Country Name` = col_character(),  
##   `Country Code` = col_character(),  
##   `Indicator Name` = col_character(),  
##   `Indicator Code` = col_character(),  
##   `2019` = col_logical()  
## )  
  
## See spec(...) for full column specifications.
```

```
c = read_csv("rural population.csv")
```

```
## Parsed with column specification:  
## cols(  
##   .default = col_double(),  
##   `Country Name` = col_character(),  
##   `Country Code` = col_character(),  
##   `Indicator Name` = col_character(),  
##   `Indicator Code` = col_character(),  
##   `2019` = col_logical()  
## )
```

```
## See spec(...) for full column specifications.
```

```
d = read_csv("GDP.csv")
```

```
## Parsed with column specification:
## cols(
##   .default = col_logical(),
##   `Country Name` = col_character(),
##   `Country Code` = col_character(),
##   `Indicator Name` = col_character(),
##   `Indicator Code` = col_character(),
##   `1991` = col_double(),
##   `1992` = col_double(),
##   `1993` = col_double(),
##   `1994` = col_double(),
##   `1995` = col_double(),
##   `1996` = col_double(),
##   `1997` = col_double(),
##   `1998` = col_double(),
##   `1999` = col_double(),
##   `2000` = col_double(),
##   `2001` = col_double(),
##   `2002` = col_double(),
##   `2003` = col_double(),
##   `2004` = col_double(),
##   `2005` = col_double(),
##   `2006` = col_double()
##   # ... with 13 more columns
## )
```

```
## See spec(...) for full column specifications.
```

```
e = read_csv("international tourism.csv")
```

```
## Parsed with column specification:
## cols(
##   .default = col_logical(),
##   `Country Name` = col_character(),
##   `Country Code` = col_character(),
##   `Indicator Name` = col_character(),
##   `Indicator Code` = col_character(),
##   `1995` = col_double(),
##   `1996` = col_double(),
##   `1997` = col_double(),
##   `1998` = col_double(),
##   `1999` = col_double(),
##   `2000` = col_double(),
##   `2001` = col_double(),
##   `2002` = col_double(),
##   `2003` = col_double(),
##   `2004` = col_double(),
##   `2005` = col_double(),

```

```
## `2006` = col_double(),
## `2007` = col_double(),
## `2008` = col_double(),
## `2009` = col_double(),
## `2010` = col_double()
## # ... with 7 more columns
## )

## See spec(...) for full column specifications.
```

```
f = read_csv("employment in service.csv")
```

```
## Parsed with column specification:
## cols(
##   .default = col_double(),
##   `Country Name` = col_character(),
##   `Country Code` = col_character(),
##   `Indicator Name` = col_character(),
##   `Indicator Code` = col_character(),
##   `2019` = col_logical()
## )

## See spec(...) for full column specifications.
```

```
require(tidyverse)
```

```
## Loading required package: tidyverse
```

```
## -- Attaching packages ----- tidyverse 1.2.1 --
```

```
## v ggplot2 3.2.1    v purrr  0.3.2
## v tibble  2.1.3    v dplyr  0.8.3
## v tidyr   1.0.0    v stringr 1.4.0
## v ggplot2 3.2.1    v forcats 0.4.0
```

```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
a1 <- select(a, "Country Name", "2013": "2017")
a1
```

```
## # A tibble: 264 x 6
##   `Country Name`   `2013` `2014` `2015` `2016` `2017`
##   <chr>           <dbl>  <dbl>  <dbl>  <dbl>  <dbl>
## 1 Aruba           573.   577.   580.   583.   585.
## 2 Afghanistan     49.4   51.1   52.7   54.2   55.6
## 3 Angola           20.9   21.6   22.4   23.1   23.9
## 4 Albania          106.   105.   105.   105.   105.
## 5 Andorra          172.   169.   166.   164.   164.
```

```
## 6 Arab World          33.8  34.5  35.3  36.0  36.7
## 7 United Arab Emirates 130.   130.  130.  132.  134.
## 8 Argentina           15.4  15.6  15.8  15.9  16.1
## 9 Armenia              102.   102.  103.  103.  103.
## 10 American Samoa      279.   279.  279.  279.  278.
## # ... with 254 more rows
```

```
a2 <- a1 %>% gather("year", "population density", -"Country Name")
a2
```

```
## # A tibble: 1,320 x 3
##   `Country Name`      year `population density`
##   <chr>              <chr>          <dbl>
## 1 Aruba              2013              573.
## 2 Afghanistan        2013              49.4
## 3 Angola              2013              20.9
## 4 Albania             2013              106.
## 5 Andorra            2013              172.
## 6 Arab World          2013              33.8
## 7 United Arab Emirates 2013              130.
## 8 Argentina           2013              15.4
## 9 Armenia             2013              102.
## 10 American Samoa     2013              279.
## # ... with 1,310 more rows
```

```
b1 <- select(b, "Country Name", "2013":"2017")
b1
```

```
## # A tibble: 264 x 6
##   `Country Name`      `2013` `2014` `2015` `2016` `2017`
##   <chr>              <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 Aruba              427.   569.   757.  1173.   978.
## 2 Afghanistan        1.43   1.71   2.15   12.0   43.4
## 3 Angola              4.57   5.94   8.14   8.32  10.8
## 4 Albania             35.6   52.3   68.4   145.   443.
## 5 Andorra            978.  1098.  2025.  2833.  3416.
## 6 Arab World          20.5   26.7   36.5   55.8   127.
## 7 United Arab Emirates 275.   367.   587.   939.  1285.
## 8 Argentina           73.9   89.9  123.   735.  1629.
## 9 Armenia             54.5   62.5   76.6   190.   271.
## 10 American Samoa      0      430.   17.9   179.   989.
## # ... with 254 more rows
```

```
b2 <- b1 %>% gather("year", "secure internet server", -"Country Name")
b2
```

```
## # A tibble: 1,320 x 3
##   `Country Name`      year `secure internet server`
##   <chr>              <chr>          <dbl>
## 1 Aruba              2013              427.
## 2 Afghanistan        2013              1.43
## 3 Angola              2013              4.57
```

```
## 4 Albania                2013                35.6
## 5 Andorra                 2013                978.
## 6 Arab World              2013                20.5
## 7 United Arab Emirates   2013                275.
## 8 Argentina               2013                73.9
## 9 Armenia                 2013                54.5
## 10 American Samoa        2013                 0
## # ... with 1,310 more rows
```

```
c1 <- select(c,"Country Name","2013":"2017")
c1
```

```
## # A tibble: 264 x 6
##   `Country Name`   `2013`   `2014`   `2015`   `2016`   `2017`
##   <chr>           <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
## 1 Aruba           2.58e 9  2.65e 9  2.69e 9  2.65e 9  2.70e 9
## 2 Afghanistan     2.06e10  2.05e10  1.99e10  1.94e10  2.02e10
## 3 Angola           1.37e11  1.46e11  1.16e11  1.01e11  1.22e11
## 4 Albania          1.28e10  1.32e10  1.14e10  1.19e10  1.30e10
## 5 Andorra          3.28e 9  3.35e 9  2.81e 9  2.88e 9  3.01e 9
## 6 Arab World       2.87e12  2.91e12  2.56e12  2.51e12  2.59e12
## 7 United Arab Emir~ 3.90e11  4.03e11  3.58e11  3.57e11  3.83e11
## 8 Argentina        5.52e11  5.26e11  5.95e11  5.58e11  6.43e11
## 9 Armenia          1.11e10  1.16e10  1.06e10  1.05e10  1.15e10
## 10 American Samoa   6.41e 8  6.43e 8  6.61e 8  6.53e 8  6.34e 8
## # ... with 254 more rows
```

```
c2 <- c1 %>% gather("year","rural population",- "Country Name")
c2
```

```
## # A tibble: 1,320 x 3
##   `Country Name`   year `rural population`
##   <chr>           <chr>   <dbl>
## 1 Aruba           2013    2581564246
## 2 Afghanistan     2013    20561054090
## 3 Angola           2013    136710000000
## 4 Albania          2013    12776280961
## 5 Andorra          2013     3281585236
## 6 Arab World       2013    2867270000000
## 7 United Arab Emirates 2013    390108000000
## 8 Argentina        2013    552025000000
## 9 Armenia          2013    11121465767
## 10 American Samoa   2013     641000000
## # ... with 1,310 more rows
```

```
d1 <- select(d,"Country Name","2013":"2017")
d1
```

```
## # A tibble: 264 x 6
##   `Country Name`   `2013` `2014` `2015` `2016` `2017`
##   <chr>           <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 Aruba           NA     NA     NA     NA     NA
```

```
## 2 Afghanistan      27.6  27.7  27.8  28.0  28.2
## 3 Angola            41.8  42.5  42.9  42.9  43.1
## 4 Albania           36.6  37.9  38.3  40.3  42.4
## 5 Andorra           NA     NA     NA     NA     NA
## 6 Arab World        58.4  58.7  60.0  60.5  60.8
## 7 United Arab Emirates 93.6  93.7  93.7  93.8  93.8
## 8 Argentina         90.9  91.5  91.3  91.1  90.9
## 9 Armenia           48.8  51.4  52.4  55.2  55.5
## 10 American Samoa   NA     NA     NA     NA     NA
## # ... with 254 more rows
```

```
d2 <- d1 %>% gather("year", "GDP", -"Country Name")
d2
```

```
## # A tibble: 1,320 x 3
##   `Country Name`      year    GDP
##   <chr>             <chr> <dbl>
## 1 Aruba             2013    NA
## 2 Afghanistan       2013  27.6
## 3 Angola             2013  41.8
## 4 Albania            2013  36.6
## 5 Andorra            2013    NA
## 6 Arab World         2013  58.4
## 7 United Arab Emirates 2013  93.6
## 8 Argentina          2013  90.9
## 9 Armenia            2013  48.8
## 10 American Samoa    2013    NA
## # ... with 1,310 more rows
```

```
e1 <- select(e, "Country Name", "2013":"2017")
e1
```

```
## # A tibble: 264 x 6
##   `Country Name`      `2013` `2014` `2015` `2016` `2017`
##   <chr>             <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 Aruba             15.1  15.6  16.9  16.5    NA
## 2 Afghanistan       1.35  1.67  1.75  1.19  1.62
## 3 Angola             0.640  0.943  1.02  3.21  4.30
## 4 Albania           26.1  27.1  25.8  24.6  24.0
## 5 Andorra           NA     NA     NA     NA     NA
## 6 Arab World         7.79  8.55  8.94  10.0  9.75
## 7 United Arab Emirates NA     NA     NA     NA     NA
## 8 Argentina          9.33  9.68  12.2  15.6  15.2
## 9 Armenia           18.4  20.1  23.9  24.8  23.2
## 10 American Samoa    NA     NA     NA     NA     NA
## # ... with 254 more rows
```

```
e2 <- e1 %>% gather("year", "international tourism", -"Country Name")
e2
```

```
## # A tibble: 1,320 x 3
##   `Country Name`      year `international tourism`
```

```
##      <chr>                <chr>                <dbl>
## 1 Aruba                  2013                  15.1
## 2 Afghanistan            2013                   1.35
## 3 Angola                 2013                   0.640
## 4 Albania                2013                   26.1
## 5 Andorra                2013                    NA
## 6 Arab World             2013                   7.79
## 7 United Arab Emirates  2013                    NA
## 8 Argentina              2013                   9.33
## 9 Armenia                2013                  18.4
## 10 American Samoa       2013                    NA
## # ... with 1,310 more rows
```

```
f1 <- select(f,"Country Name","2013":"2017")
f1
```

```
## # A tibble: 264 x 6
##   `Country Name`      `2013`      `2014`      `2015`      `2016`      `2017`
##   <chr>              <dbl>      <dbl>      <dbl>      <dbl>      <dbl>
## 1 Aruba              58811      59109      59362      59576      59750
## 2 Afghanistan        24404522   25165917   25877997   26530269   27131559
## 3 Angola              9885476    10040932    10192857    10340319    10483867
## 4 Albania            1291587     1258985     1226200     1195854     1167112
## 5 Andorra              9259        9156        9092        9084        9125
## 6 Arab World         161751073   164114145   166207258   168295267   170355303
## 7 United Arab Emirates 1373616     1347573     1327003     1313814     1304680
## 8 Argentina           3693179     3679391     3664923     3649822     3634137
## 9 Armenia             1067354     1074327     1079968     1083966     1086546
## 10 American Samoa      7049        7092        7123        7136        7136
## # ... with 254 more rows
```

```
f2 <- f1 %>% gather("year","employet in survice",- "Country Name")
f2
```

```
## # A tibble: 1,320 x 3
##   `Country Name`      year `employet in survice`
##   <chr>              <chr>                <dbl>
## 1 Aruba              2013                  58811
## 2 Afghanistan        2013                24404522
## 3 Angola              2013                9885476
## 4 Albania            2013                1291587
## 5 Andorra            2013                  9259
## 6 Arab World         2013                161751073
## 7 United Arab Emirates 2013                1373616
## 8 Argentina          2013                3693179
## 9 Armenia            2013                1067354
## 10 American Samoa     2013                  7049
## # ... with 1,310 more rows
```

```
dat1 <- full_join(a2,b2,
  by=c('Country Name'
    , 'year') )
dat1
```

```
## # A tibble: 1,320 x 4
##   `Country Name`   year `population density` `secure internet server`
##   <chr>           <chr>         <dbl>         <dbl>
## 1 Aruba           2013           573.           427.
## 2 Afghanistan     2013           49.4           1.43
## 3 Angola           2013           20.9           4.57
## 4 Albania          2013           106.           35.6
## 5 Andorra          2013           172.           978.
## 6 Arab World       2013           33.8           20.5
## 7 United Arab Emirates 2013          130.           275.
## 8 Argentina        2013           15.4           73.9
## 9 Armenia          2013           102.           54.5
## 10 American Samoa  2013           279.            0
## # ... with 1,310 more rows
```

```
dat2 <- full_join(c2,d2,
  by=c('Country Name'
    , 'year') )
dat2
```

```
## # A tibble: 1,320 x 4
##   `Country Name`   year `rural population` GDP
##   <chr>           <chr>         <dbl> <dbl>
## 1 Aruba           2013      2581564246 NA
## 2 Afghanistan     2013      20561054090 27.6
## 3 Angola           2013      136710000000 41.8
## 4 Albania          2013      12776280961 36.6
## 5 Andorra          2013       3281585236 NA
## 6 Arab World       2013      2867270000000 58.4
## 7 United Arab Emirates 2013      390108000000 93.6
## 8 Argentina        2013      552025000000 90.9
## 9 Armenia          2013      11121465767 48.8
## 10 American Samoa  2013       641000000 NA
## # ... with 1,310 more rows
```

```
dat3<- full_join(e2,f2,
  by=c('Country Name'
    , 'year') )
dat3
```

```
## # A tibble: 1,320 x 4
##   `Country Name`   year `international tourism` `employmet in survic~
##   <chr>           <chr>         <dbl>         <dbl>
## 1 Aruba           2013          15.1           58811
## 2 Afghanistan     2013          1.35          24404522
## 3 Angola           2013          0.640          9885476
## 4 Albania          2013          26.1           1291587
## 5 Andorra          2013          NA              9259
## 6 Arab World       2013          7.79          161751073
## 7 United Arab Emirates 2013          NA              1373616
## 8 Argentina        2013          9.33           3693179
## 9 Armenia          2013          18.4           1067354
## 10 American Samoa  2013          NA              7049
## # ... with 1,310 more rows
```



```
dat4 <- full_join(dat1,dat2,
  by=c('Country Name'
    , 'year') )
dat4
```

```
## # A tibble: 1,320 x 6
##   `Country Name` year `population den~` `secure interne~` `rural populati~`
##   <chr>          <chr>          <dbl>          <dbl>          <dbl>
## 1 Aruba          2013          573.          427.          2581564246
## 2 Afghanistan    2013          49.4          1.43         20561054090
## 3 Angola          2013          20.9          4.57         136710000000
## 4 Albania         2013          106.          35.6         12776280961
## 5 Andorra         2013          172.          978.          3281585236
## 6 Arab World      2013          33.8          20.5         2867270000000
## 7 United Arab E~  2013          130.          275.          390108000000
## 8 Argentina       2013          15.4          73.9         552025000000
## 9 Armenia         2013          102.          54.5         11121465767
## 10 American Samoa 2013          279.          0            641000000
## # ... with 1,310 more rows, and 1 more variable: GDP <dbl>
```

```
dat5 <- full_join(dat3,dat4,
  by=c('Country Name'
    , 'year') )
dat5
```

```
## # A tibble: 1,320 x 8
##   `Country Name` year `international ~` `employment in s~` `population den~`
##   <chr>          <chr>          <dbl>          <dbl>          <dbl>
## 1 Aruba          2013          15.1          58811          573.
## 2 Afghanistan    2013          1.35         24404522         49.4
## 3 Angola          2013          0.640        9885476         20.9
## 4 Albania         2013          26.1        1291587         106.
## 5 Andorra         2013          NA           9259          172.
## 6 Arab World      2013          7.79        161751073         33.8
## 7 United Arab E~  2013          NA          1373616         130.
## 8 Argentina       2013          9.33        3693179          15.4
## 9 Armenia         2013          18.4        1067354          102.
## 10 American Samoa 2013          NA           7049          279.
## # ... with 1,310 more rows, and 3 more variables: `secure internet
## #   server` <dbl>, `rural population` <dbl>, GDP <dbl>
```

```
dat6 <- drop_na(dat5)
dat6
```

```
## # A tibble: 1,020 x 8
##   `Country Name` year `international ~` `employment in s~` `population den~`
##   <chr>          <chr>          <dbl>          <dbl>          <dbl>
## 1 Afghanistan    2013          1.35         24404522         49.4
## 2 Angola          2013          0.640        9885476         20.9
## 3 Albania         2013          26.1        1291587         106.
## 4 Arab World      2013          7.79        161751073         33.8
## 5 Argentina       2013          9.33        3693179          15.4
```

```
## 6 Armenia      2013      18.4      1067354      102.
## 7 Australia    2013      11.2      3353116       3.01
## 8 Austria      2013       4.73      3617832      103.
## 9 Azerbaijan   2013      15.6      4318074      114.
## 10 Burundi     2013       3.72      8444873      372.
## # ... with 1,010 more rows, and 3 more variables: `secure internet
## #   server` <dbl>, `rural population` <dbl>, GDP <dbl>
```

```
dat7 <- filter(dat6, 'year' == "2014")
ggplot(data=dat7,
       aes(x='international tourism',y='rural population')) +
  geom_point()
```

