

# حل تمرین جلسه اول

فواد اسماعیلی

2023 — 11 — 03

## فهرست مطالب

۱	۱ معرفی داده
۱	۲ فراخوانی داده‌ها
۲	۳ سوالات
۲	۱.۳ سوال اول:
۲	۲.۳ سوال دوم
۲	۳.۳ سوال سوم
۲	۴ حل سوالات:
۲	۱.۴ پیش پردازش داده
۳	۲.۴ سوال اول

(این جزوه با استفاده از بسته های knitr, rmarkdown, RTLknitr نوشته شده است.)

## ۱ معرفی داده

داده‌های ارائه شده در این تمرین، از سایت worldometers گرفته شده است، که آخرین آمار جمعیتی سازمان ملل متحد نسبت به هر کشور را ارائه کرده است.

## ۲ فراخوانی داده‌ها

فایل داده ها را از آدرس گیت‌هاب دانلود کرده و به صورت زیر فراخوانی کنید.

```
load("world-data.Rdata")
```

در آن سه شیء زیر در دسترس هستند. که شش مشاهده اول هر یک نمایش داده شده است.

```
head(Country)
```

```
## [1] "India" "China" "United States" "Indonesia"
## [5] "Pakistan" "Nigeria"
```

```
head(Migrants)
```

```
## [1] "-486,136" "-310,220" "999,700" "-49,997" "-165,988" "-59,996"
```

```
head(Population)
```

```
## [1] "1,428,627,663" "1,425,671,352" "339,996,563" "277,534,122"
```

```
## [5] "240,485,658" "223,804,632"
```

## ۳ سوالات

### ۱.۳ سوال اول:

چه کشور هایی بیشترین مهاجر پذیری را دارند؟

### ۲.۳ سوال دوم

از چه کشورهایی بیشترین مهاجرت ها صورت می گیرد؟

### ۳.۳ سوال سوم

آیا کشوری در سال ۲۰۲۳ هست که نه مهاجر پذیرفته باشد و نه کسی از آن مهاجرت کرده باشد؟

## ۴ حل سوالات:

### ۱.۴ پیش پردازش داده

از آنجایی که داده ها متاسفانه ساختار عددی ندارند، و به صورت رشته ذخیره شده اند، همچنین اعداد بین خود علامت '،' دارند، که یک داده اضافی است و باید آن را حذف کنیم، نیاز داریم تا این مراحل انجام شوند.

با استفاده از تابع `gsub` می توانیم ویرگول ها را برداریم.

```
migrants <- gsub(",", "", Migrants)
```

```
head(migrants)
```

```
## [1] "-486136" "-310220" "999700" "-49997" "-165988" "-59996"
```

```
population <- gsub(",", "", Population)
```

```
head(population)
```

```
## [1] "1428627663" "1425671352" "339996563" "277534122" "240485658"
```

```
## [6] "223804632"
```

بعد از برآشتن ویرگول ها، حال می توانیم اعدادی که به صورت رشته هستند را به عددی تبدیل کنیم

```
migrants <- as.numeric(migrants)
```

```
head(migrants)
```

```
## [1] -486136 -310220 999700 -49997 -165988 -59996
```

```
population <- as.numeric(population)
head(population)
```

```
## [1] 1428627663 1425671352 339996563 277534122 240485658 223804632
```

حال می‌توان به سوالات پاسخ داد.

## ۲.۴ سوال اول

دستور زیر ده کشوری که بیشترین مهاجرت پذیری را داشته اند، را نشان می‌دهد.

```
mig <- migrants[order(migrants,decreasing = TRUE)]
co <- Country[order(migrants,decreasing = TRUE)]
mig_co_mat <- cbind(mig,co)
head(mig_co_mat,10)
```

```
##      mig      co
## [1,] "1784718" "Ukraine"
## [2,] "999700"  "United States"
## [3,] "757103"  "Syria"
## [4,] "321106"  "Venezuela"
## [5,] "249746"  "Canada"
## [6,] "165790"  "United Kingdom"
## [7,] "155751"  "Germany"
## [8,] "139991"  "Australia"
## [9,] "99994"   "Japan"
## [10,] "67761"  "France"
```

دستوری زیر ده کشوری که بیشترین مهاجرت از آنها صورت پذیرفته را نشان می‌دهد.

```
tail(mig_co_mat,10)
```

```
##      mig      co
## [225,] "-157392" "Jordan"
## [226,] "-165988" "Pakistan"
## [227,] "-175051" "Colombia"
## [228,] "-177331" "Lebanon"
## [229,] "-254616" "Romania"
## [230,] "-309977" "Bangladesh"
## [231,] "-310220" "China"
## [232,] "-318067" "Turkey"
## [233,] "-486136" "India"
## [234,] "-910475" "Poland"
```

کشوری هایی با مهاجرت پذیری و خروج مهاجر برابر صفر عبارتند از:

```
Country[migrants==0]
```

```
## [1] "Kazakhstan"      "Azerbaijan"
## [3] "United Arab Emirates" "Oman"
## [5] "Qatar"           "Bahrain"
## [7] "Mauritius"       "Brunei"
## [9] "Mayotte"         "Vanuatu"
## [11] "Saint Lucia"     "Antigua and Barbuda"
## [13] "Bermuda"         "Faeroe Islands"
## [15] "Sint Maarten"    "Marshall Islands"
```

## [17]	"Saint Martin"	"Anguilla"
## [19]	"Saint Barthelemy"	"Saint Pierre & Miquelon"
## [21]	"Saint Helena"	"Montserrat"
## [23]	"Falkland Islands"	"Niue"
## [25]	"Tokelau"	"Holy See"