

# Mervar - Lab 1

Alexander Mervar

2022-08-24

1. Install R on your machine by visiting <https://www.r-project.org/>

**Done**

2. Most likely, you'll want to install R studio as well. It is a useful interactive environment for R programming that includes windows for code, instantaneously executed commands, variables, and plots. Download R Studio at <https://www.rstudio.com/products/rstudio/download2/>

**Done**

3. Read in the data file using the command: `cities<-read.table("/Users/rgoldsto/Desktop/top200cities.txt", header = TRUE)` replacing the path with the one that leads to where you put the file.

```
cities<-read.table("top200cities.txt", header = TRUE)
```

4. Cities has the following variables: rank, city [the city's name], state, pop2015 [population in 2015], pop2010, pctchange [% change to population from 2010 to 2015], landsqmi [area of city in square miles], landsqkm, popdensity [population density in square miles], popdensitysqkm, latitude [in degrees; larger values are more to the North], longitude [in degrees; larger values are to the West]. Report these names with the command: `names(cities)`

```
names(cities)
```

```
## [1] "rank"          "city"          "state"         "pop2015"
## [5] "pop2010"       "pctchange"     "landsqmi"      "landsqkm"
## [9] "popdensitysqmi" "popdensitysqkm" "latitude"      "longitude"
```

5. Are there more cities to the East or West of the midpoint of the continental U.S., which is located at 98.35 degrees? Show the R commands you used to answer this question. You shouldn't have to count them by hand.

```
midpointLong<-98.35
numOfWest<-nrow(cities[cities$longitude>midpointLong,])
numOfEast<-nrow(cities[cities$longitude<midpointLong,])
numOfWest
```

```
## [1] 81
```

```
numOfEast
```

```
## [1] 119
```

Looking at the data, there are more cities on the East side of the midpoint of the continental U.S., which is located at 98.35 degrees.

6. Are there more cities to the South or North of the midpoint of the continental U.S., which is located at 39.50 degrees? Show the R commands you used to answer this question.

```
midpointLat<-39.5
numOfNorth<-nrow(cities[cities$latitude>midpointLat,])
numOfSouth<-nrow(cities[cities$latitude<midpointLat,])
numOfNorth
```

```
## [1] 55
```

```
numOfSouth
```

```
## [1] 145
```

Looking at the data, there are more cities to the South side of the midpoint of the continental U.S., which is located at 39.50 degrees.

7. Show the R command that let's you answer the question: what is the population density of Miami?

```
cities$popdensqmi[cities$city=="Miami"]
```

```
## [1] 11539
```

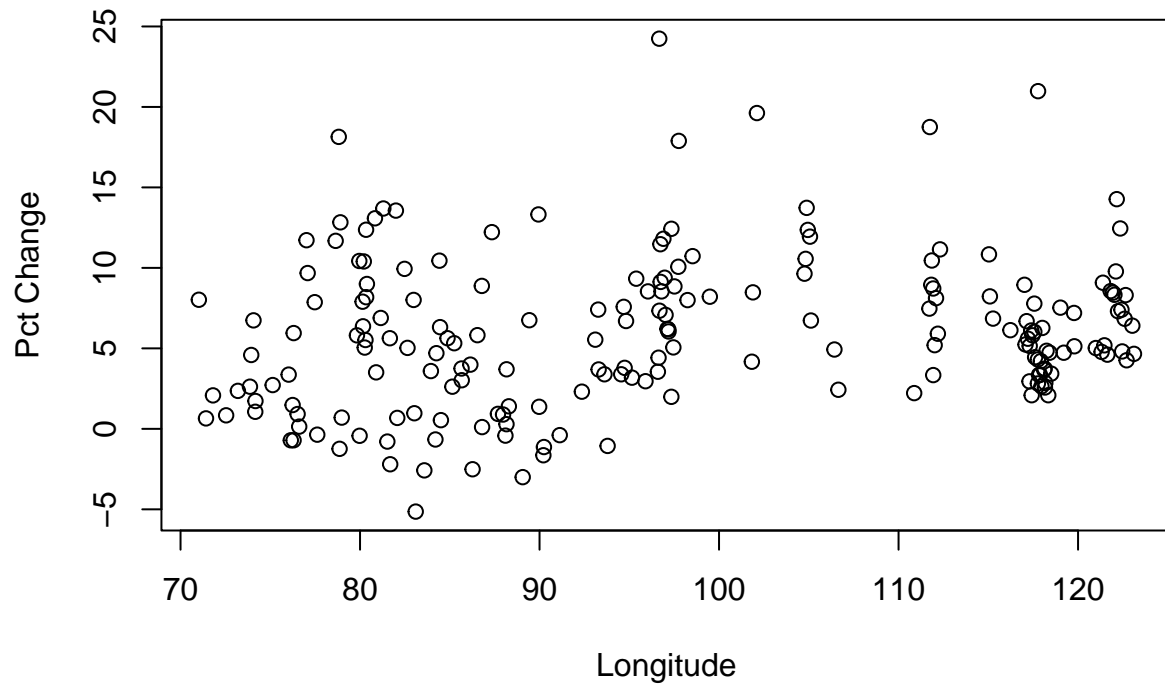
8. There's one city that has a percentage growth more than 50% from 2010 to 2015. Show the R command that finds and prints the name of this city.

```
cities$city[cities$pctchange>50]
```

```
## [1] "Macon"
```

9. Make a new data frame called `mostcities` that is the same as `cities` but excluding cities with growth more than 30% or longitudes > 140 degrees (Hawaii and Alaska). Now show the scattergram that plots (using "plot") the percentage change in `mostcities` as a function of longitude. The plot should have one circle for each of the remaining 196 cities. Looking at this plot, describe how recent growth of cities in America varies with their East-West location.

```
mostcities<-cities[cities$pctchange<30,]
mostcities<-mostcities[mostcities$longitude<140,]
plot(mostcities$longitude,mostcities$pctchange,xlab = "Longitude",ylab = "Pct Change")
```



Looking at the plot, it seems that the West side of the country has a major packet of cities at longitude 118 that have strong amount of growth ~3-7%. The East side of the country does not seem to have this kind of strong packet of cities. Instead, it has a lot of smaller packets with a cluster of a handful of cities.