An Overview of Brazil's weather conditions from 1991 to 2020

Tiago de Almeida Silva

23/03/2022

Personal project of weather conditions in Brazil from 1991 to 2020. In this study case was used an official document recently released by the Brazilian National Institute of Meteorology (INMET), this document presents several weather and climate conditions in Brazil from a 30-year range. In this project, I will analyze and present my findings concerning precipitation, minimum and maximum temperature in the Brazilian states during this period. I am using the average of the variables used in this study to get an overall overview of how the climate in Brazil has been behaving during the last 3 decades.

Part 1

In the first part of my project, I will work with the precipitation dataset that presents how much rain, in "mm", has been fallen monthly in the hundreds of INMET weather stations spread around the Brazilian territory. All my findings are related to the states in Brazil and for this reason, I will group the weather stations according to their respective states.

It is important to mention that the state of Rondônia, located at the north of Brazil, does not have any available data concerning monthly precipitation during this period, that's why the state is not mentioned in this first part of the project.

Loading the libraries used in this project

```
pacotes <- c("tidyverse","sf","tmap","rgdal","rgeos","adehabitatHR","knitr",
             "kableExtra", "tidyr", "ggplot2", "readxl", "maptools", "broom",
             "kableExtra", "RColorBrewer", "dplyr", "readr","zoo", "sqldf")

if(sum(as.numeric(!pacotes %in% installed.packages())) != 0){
  instalador <- pacotes[!pacotes %in% installed.packages()]
  for(i in 1:length(instalador)) {
    install.packages(instalador, dependencies = T)
    break()}
  sapply(pacotes, require, character = T)
} else {
  sapply(pacotes, require, character = T)
}
```

Loading the precipitation data frame to an object named "df" and replacing "-" for "NA"

```
df <- read_xlsx("Normal-Climatologica-PREC.xlsx", skip = 2,col_names = T)

df[df=="-"] <- NA
```

Having a quick overview of the "df" dataset

```
str(df)
```

```
## tibble [221 x 16] (S3: tbl_df/tbl/data.frame)
##  $ Código         : num [1:221] 82989 83249 82353 83007 82970 ...
##  $ Nome da Estação: chr [1:221] "AGUA BRANCA" "ALAGOINHAS" "ALTAMIRA" "ALTO DA BOA VISTA" ...
##  $ UF             : chr [1:221] "AL" "BA" "PA" "RJ" ...
##  $ Janeiro        : chr [1:221] "48.6" "54.6" "284.7" "228.6" ...
##  $ Fevereiro      : chr [1:221] NA "60" "330.3" "177.8" ...
##  $ Março          : chr [1:221] "85.7" "61.7" "399.5" "252" ...
##  $ Abril          : chr [1:221] "80" "121.5" "365.6" "201.1" ...
##  $ Maio           : chr [1:221] "133.19999999999999" "157.4" "232.9" "180.8" ...
##  $ Junho          : chr [1:221] "160.4" "144.30000000000001" "102.9" "154.1" ...
##  $ Julho          : chr [1:221] "170.1" "110.4" NA "179.9" ...
##  $ Agosto         : chr [1:221] "96" "91.5" "34.200000000000003" "150.5" ...
##  $ Setembro       : chr [1:221] "52.2" "66.3" "46.2" "219.5" ...
##  $ Outubro        : chr [1:221] "35.799999999999997" "51.3" "58" "194.7" ...
##  $ Novembro       : chr [1:221] "27.7" "60.2" "90.5" "239" ...
##  $ Dezembro       : chr [1:221] "39.9" "53.2" "164" "236.8" ...
##  $ Ano            : chr [1:221] NA "1032.4000000000001" NA "2414.8000000000002" ...
```

Converting the character data, related to the precipitation by month, to numeric one

```
class(df$Janeiro) <- "numeric"
class(df$Fevereiro) <- "numeric"
class(df$Março) <- "numeric"
class(df$Abril) <- "numeric"
class(df$Maio) <- "numeric"
class(df$Junho) <- "numeric"
class(df$Julho) <- "numeric"
class(df$Agosto) <- "numeric"
class(df$Setembro) <- "numeric"
class(df$Outubro) <- "numeric"
class(df$Novembro) <- "numeric"
class(df$Dezembro) <- "numeric"
class(df$Código) <- "character"
```

Selecting and gathering necessary data for the project

```
df <- df[ ,1:15]
```

Creating a variable named “annual_mean” for the mean of annual precipitation in every single observation (weather stations)

```
df_mean <- df %>%
mutate(annual_mean = rowMeans(df[ ,4:15], na.rm = T))
```

Grouping and summarizing the annual mean according to states in Brazil

```
df_state <- df_mean %>%
  group_by(UF) %>%
  dplyr::summarise(total = mean(annual_mean))
```

Loading the shapefile of Brazil

```
shp_brazil <- readOGR(dsn = "shp_br", layer = "UFEBRASIL", encoding = "UTF-8", use_iconv = T)
```

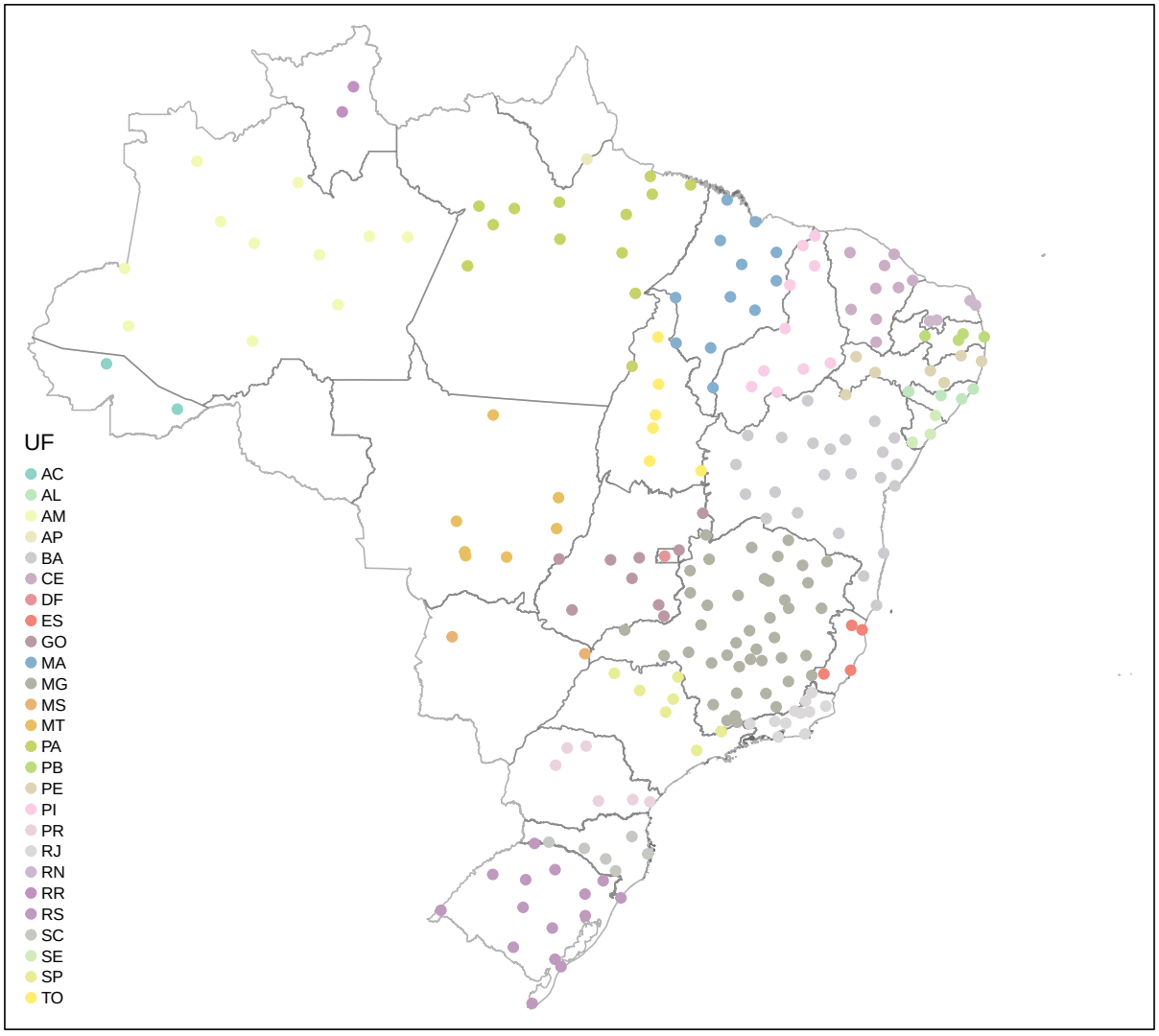
Taking a look at the shapefile

```
plot(shp_brazil)
```

Plotting the data points of the weather stations into the Shapefile to get an overview of their covering around the country.

```
class(df$Código) <- "numeric"

weather_st <- read_xlsx("Normal-Climatologica-ESTAÇÕES.xlsx", skip = 2, col_names = T)

stat_prec <- left_join(df,
                       weather_st,
                       by = c("Código", "UF", "Nome da Estação"))

stat_prec <- stat_prec[ ,c(1:3,17:19)]

sf_stat <- st_as_sf(x = stat_prec,
                         coords = c("Longitude", "Latitude"),
                         crs = 4326)
```

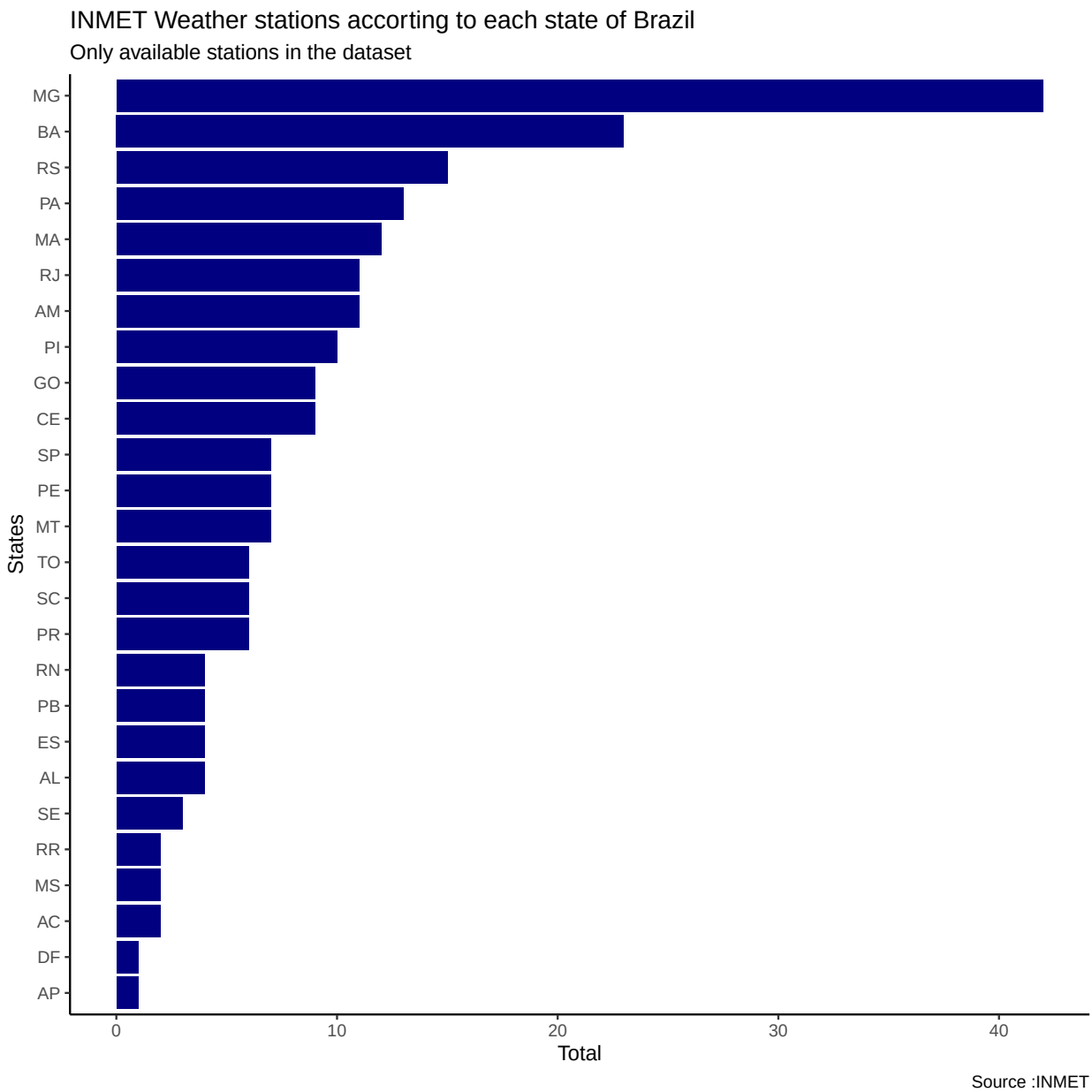
```
#plotting the map

tm_shape(shp = shp_brazil) +
  tm_borders(alpha = 0.5) +
  tm_shape(shp = sf_stat) +
  tm_dots(col = "UF",
          size = 0.2) +
  tmap_options(check.and.fix = TRUE)
```



Making a graph about the weather stations available in the dataset

```
station_plot <- sf_stat %>%
  group_by(UF) %>%
  summarise(Total = n())
```

```
ggplot(station_plot, aes(reorder(UF,(Total)), Total)) +
  geom_col(fill = "navy") +
  coord_flip() +
  labs(
    title = "INMET Weather stations accorting to each state of Brazil",
    subtitle = "Only available stations in the dataset",
    x = "States",
    caption = "Source :INMET"
  ) +
  theme_classic()
```



Adding the official state codes to each observation. This will help us to match every state Shapefile to its

respective mean of annual precipitation. Also renamed the "SIGLA" variable to "UF" to make the joining process easier.

```
state_codes <- read.csv("estados.csv", header = T, encoding = "UTF-8")

state_codes <- state_codes %>%
 rename("UF" = "SIGLA") %>%
dplyr::select(everything(), -"NOME")

state_codes$UF <- str_replace_all(state_codes$UF, fixed(" "), "")
```

Merging the state_codes object to the df_state one.

```
df_prec <- left_join(x = df_state, y = state_codes, by = "UF") %>%
  rename("CD_GEOCODU" = "COD")
```

Now I am merging the df_prec object to the dataset from the shp_brazil shapefile.

```
shp_complete <- merge(x = shp_brazil,
                      y = df_prec,
                      by = "CD_GEOCODU")
```

Quick overview of the complete Shapefile data frame.

```
shp_complete@data
```

```
##    CD_GEOCODU ID           NM_ESTADO    NM_REGIAO   UF     total
## 1          11  1            RONDÔNIA        NORTE <NA>        NA
## 2          12  2                ACRE        NORTE   AC 179.86250
## 3          13  3            AMAZONAS        NORTE   AM 215.36384
## 4          14  4             RORAIMA        NORTE   RR 163.12424
## 5          15  5                PARÁ        NORTE   PA 190.30379
## 6          16  6               AMAPÁ        NORTE   AP 210.47500
## 7          17  1           TOCANTINS        NORTE   TO 134.96667
## 8          21  1            MARANHÃO     NORDESTE   MA 132.18542
## 9          22  1               PIAUÍ     NORDESTE   PI  79.72432
## 10         23  2               CEARÁ     NORDESTE   CE  80.97562
## 11         24  2 RIO GRANDE DO NORTE     NORDESTE   RN  87.16042
## 12         25  3             PARAÍBA     NORDESTE   PB  97.78394
## 13         26  3           PERNAMBUCO     NORDESTE   PE  69.67208
## 14         27  1              ALAGOAS     NORDESTE   AL 114.77436
## 15         28  2              SERGIPE     NORDESTE   SE  89.04722
## 16         29  4                BAHIA     NORDESTE   BA  75.66268
## 17         31  4        MINAS GERAIS      SUDESTE   MG 103.67999
## 18         32  5      ESPIRITO SANTO      SUDESTE   ES 109.10000
## 19         33  5      RIO DE JANEIRO      SUDESTE   RJ 113.53623
## 20         35  6           SÃO PAULO      SUDESTE   SP 120.56894
## 21         41  3              PARANÁ          SUL   PR 147.70467
## 22         42  4      SANTA CATARINA          SUL   SC 160.12917
## 23         43  5   RIO GRANDE DO SUL          SUL   RS 139.67399
## 24         50  2  MATO GROSSO DO SUL CENTRO-OESTE   MS 101.62083
## 25         51  3         MATO GROSSO CENTRO-OESTE   MT 138.05764
## 26         52  4               GOIÁS CENTRO-OESTE   GO 125.98056
## 27         53  5    DISTRITO FEDERAL CENTRO-OESTE   DF 123.23333
```

Saving the updated Shapefile into the project folder.

```
# writeOGR(obj = shp_complete,
#          layer = "brazil_precipitation",
#         driver = "ESRI Shapefile",
#         dsn = "shp_precipitation")
```

Plotting our findings into the shapefile of Brazil. As I mentioned before, the state of Rondônia (RO) do not have data related to precipitation during this period and for this reason the state is in gray.

```
tm_shape(shp = shp_complete) +
  tm_polygons("total",
              title = "Monthly precipitation (mm)",
              style = "fixed",
              palette = c("sienna1", "khaki", "lightblue", "steelblue1", "blue1", "blue3", "blue4"),
              breaks = c(-Inf, 100, 120, 140, 160, 180, 200, Inf),
              colorNA = "gray",
              textNA = "No available data for RO") +
  tm_layout(main.title = "Average of monthly precipitation in the brazilian states from 1991 to 2020",
                        title.size = 0.5) +
  tm_compass(type = "8star",
             show.labels = 3,
             size = 3,
             position = c(0.9, 0.05)) +
  tm_credits("Source:INMET",
             position = 0.9) +
  tmap_options(check.and.fix = T) +
  tm_layout(legend.title.size = 1.5,
            legend.text.size = 1)
```

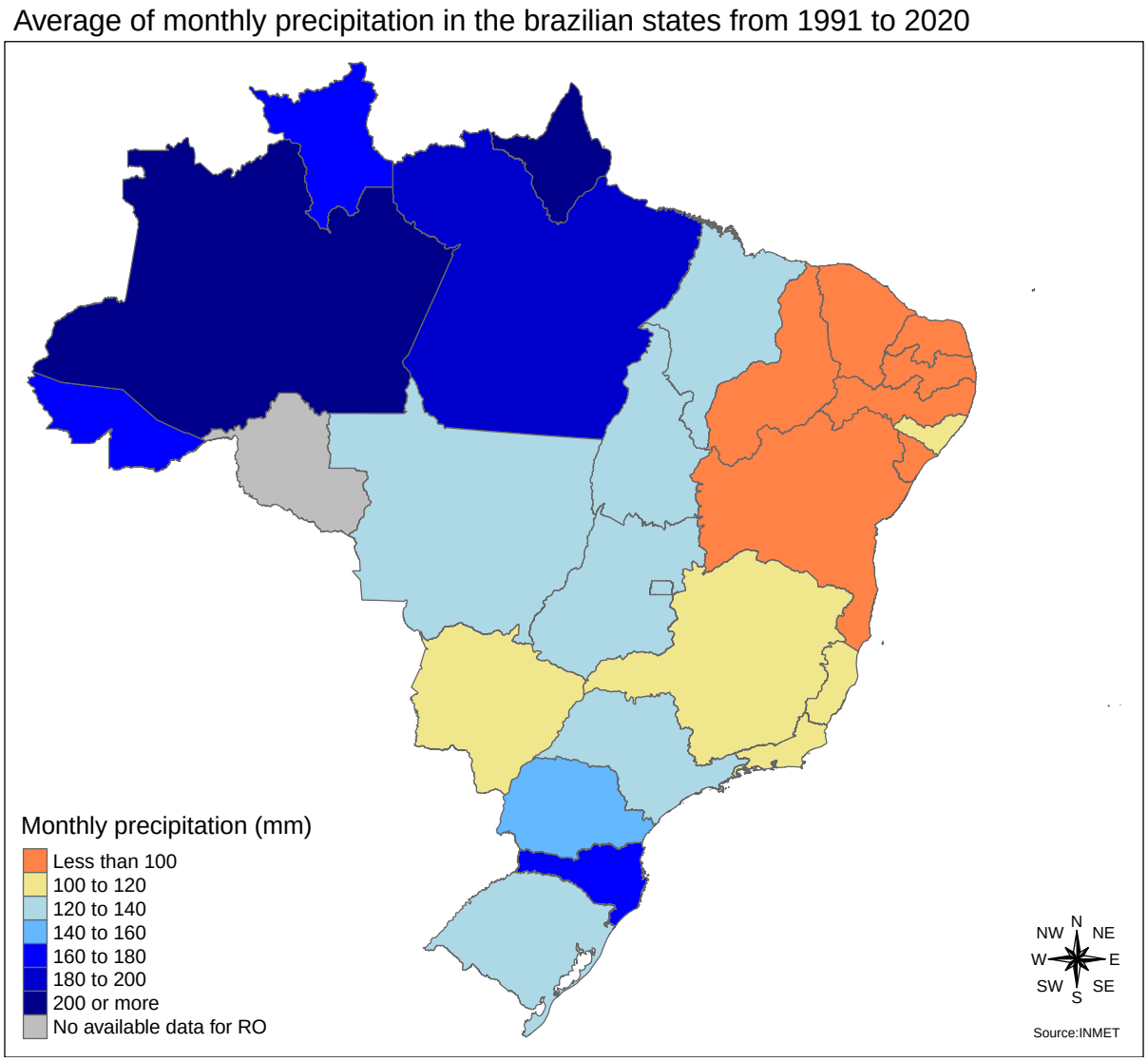
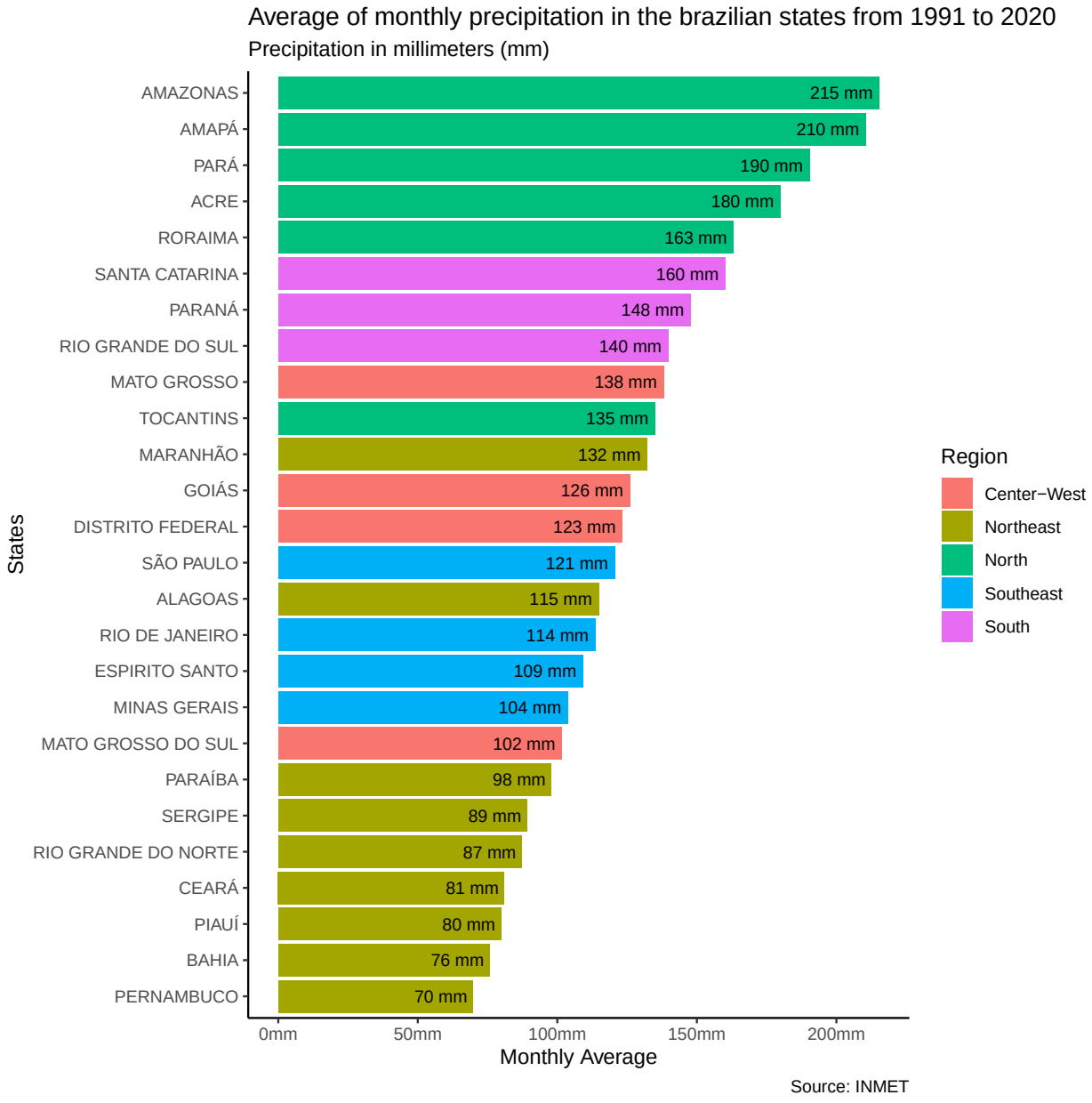


Table showing the average of monthly precipitation in the brazilian states during a 30-year period. The country regions are also labeled in the table.

```
df_complete <- shp_complete@data[2:27, ]

ggplot(data = df_complete, aes(x = reorder(NM_ESTADO, (total)), y = total, fill =  NM_REGIAO)) +
  geom_col() +
  coord_flip() +
  labs(
    title = "Average of monthly precipitation in the brazilian states from 1991 to 2020",
```

```
    subtitle = "Precipitation in millimeters (mm)",
    x = "States",
    y = "Monthly Average",
    caption = ("Source: INMET")) +
geom_text(label = paste(round(df_complete$total), "mm"), size = 3, hjust = 1.1) +
scale_fill_discrete(name = "Region",
                    labels = c("Center-West", "Northeast","North", "Southeast","South")) +
scale_y_continuous(breaks = seq(0,250, by = 50),
                   labels = paste0(seq(0,250, by = 50), "mm")) +
   theme_classic()
```



It is interesting to notice, in the map and graph, the high volume of monthly rain in northern Brazil (darker blue region in the map) and that's due to the Amazon rainforest that covers most of these states. Santa Catarina state, in the south of Brazil, also presents a high volume, and the state is known to be the only

place in Brazil to suffer the consequences of cyclones and tornados, perhaps that's one of the explanations for this federal unit shows higher precipitation than the neighbor states. The northeast states are drier (orange region in the map) than the rest of Brazil because there is a large semiarid region, known as "Caatinga", throughout this region.

Part 2

In the second part of the project, I will analyze what was the lowest temperature registered in the available INMET weather stations throughout the country, and after that, I will gather the data showing the lowest temperature registered within each state of Brazil from 1991 to 2020.

There are no available data for the states of Rondônia (RO) e Mato Grosso do Sul (MS), for this reason, both states are not mentioned in this part of the project.

Loading, tidying and cleaning the dataset.

```
df_tmin <- read_xlsx("Normal-Climatologica-TMINABS.xlsx", skip = 2, col_names = T)

df_tmin <- na.omit(df_tmin)

df_tmin[df_tmin == "-"] <- NA
```

Turning the character variables regarding the temperatures registered by month into numeric ones.

```
class(df_tmin$...5) <- "numeric"
class(df_tmin$...7) <- "numeric"
class(df_tmin$...9) <- "numeric"
class(df_tmin$...11) <- "numeric"
class(df_tmin$...13) <- "numeric"
class(df_tmin$...15) <- "numeric"
class(df_tmin$...17) <- "numeric"
class(df_tmin$...19) <- "numeric"
class(df_tmin$...21) <- "numeric"
class(df_tmin$...23) <- "numeric"
class(df_tmin$...25) <- "numeric"
class(df_tmin$...27) <- "numeric"
```

Selecting the useful variables for this project

```
df_tmin <-  df_tmin[ ,c(1:3, 5,7,9,11,13,15,17,19,21,23,25,27)]
```

Finding the lowest temperature to each weather station and assigning it to a new column named "lowest_temp" and selecting the variables that will be used in the following steps.

```
df_tmin2 <- df_tmin %>%
  mutate(lowest_temp = apply(df_tmin[ ,4:15], 1, FUN = min, na.rm = TRUE))

df_tmin2 <- df_tmin2[ ,c(1:3, 16)]
```

Loading the dataset related to the geospatial data of the weather stations from INMET. That's important to plot the data points concerning the weather station into a map. After loading it, I will clean and tidy the dataset to prepare it for merging.

```
weather_st <- read_xlsx("Normal-Climatologica-ESTAÇÕES.xlsx", skip = 2, col_names = T)

weather_st <- weather_st[ ,c(2,5:7)]
```

Merging the weather_st object to the df_tmin one to get all the variables needed to get the desired outcome. The merging will be added to a new object named "df_tmin_c"

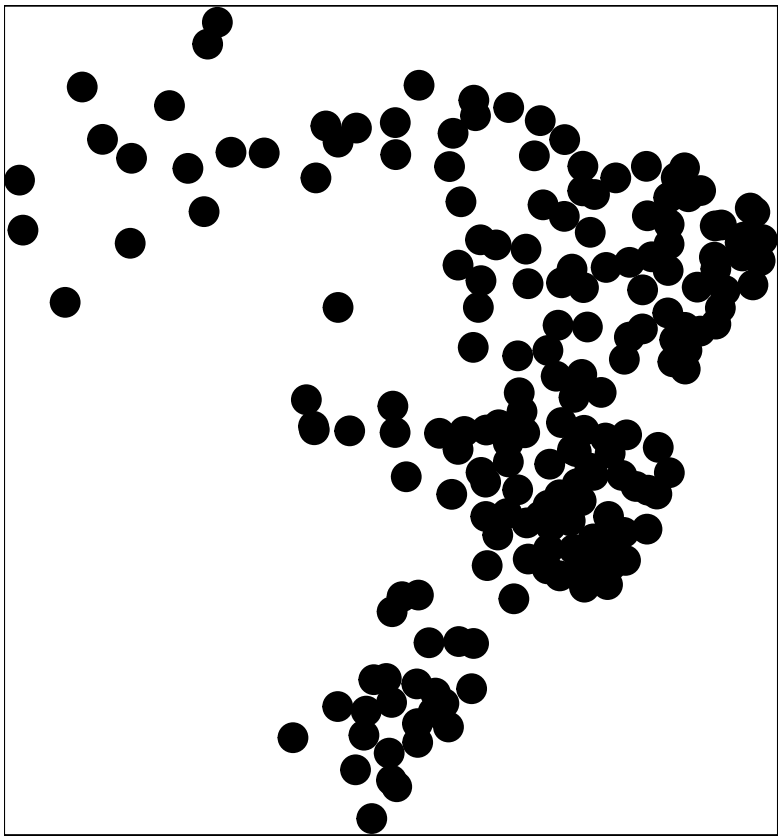
```
df_tmin_c <- left_join(df_tmin2,
          weather_st,
          by = "Código")
```

Creating a spatial feature (SF) object with the new object. Sf objects allow us to plot data points into a real map of the earth's surface, in this case, the Brazilian one.

```
sf_tmin <- st_as_sf(x = df_tmin_c,
                         coords = c("Longitude", "Latitude"),
                         crs = 4326)
```

Plotting the sf_tmin spatially to get a better overview of the data points.

```
tm_shape(shp = sf_tmin) +
  tm_dots(size = 1)
```

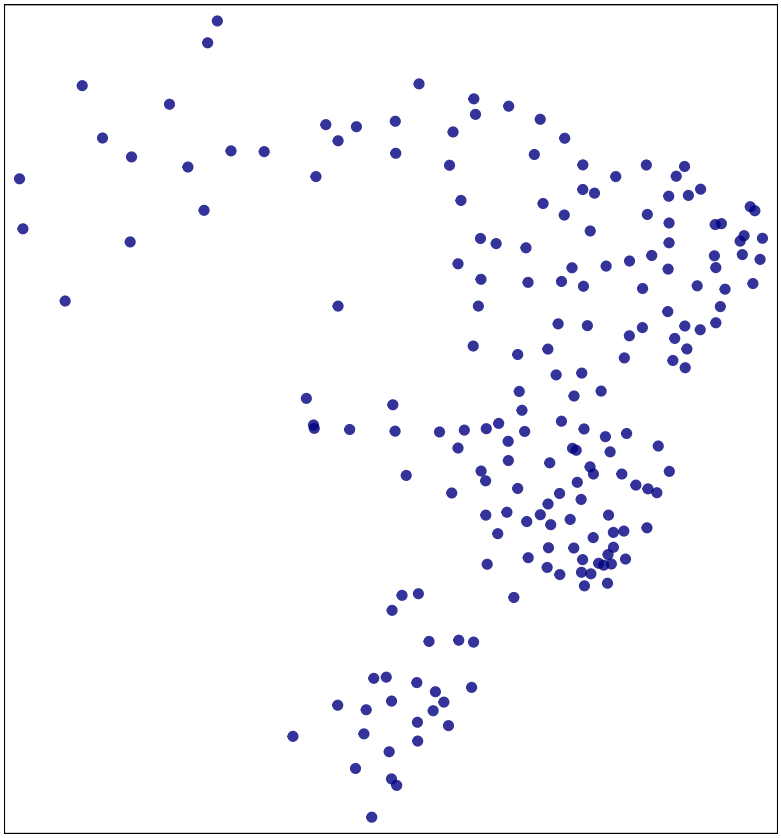


Viewing the same data points ploted into the Brazil surface.

```
#It is necessary to copy this code chunk to a script because view mode is not supported in the R markdo

tmap_mode("plot")
```

```
## tmap mode set to plotting
```

```
tm_shape(shp = sf_tmin) +
  tm_dots(col = "navy",
          border.col = "black",
          size = 0.1,
          alpha = 0.8)
```



Switching the view mode to “plot”

```
tmap_mode("plot")
```

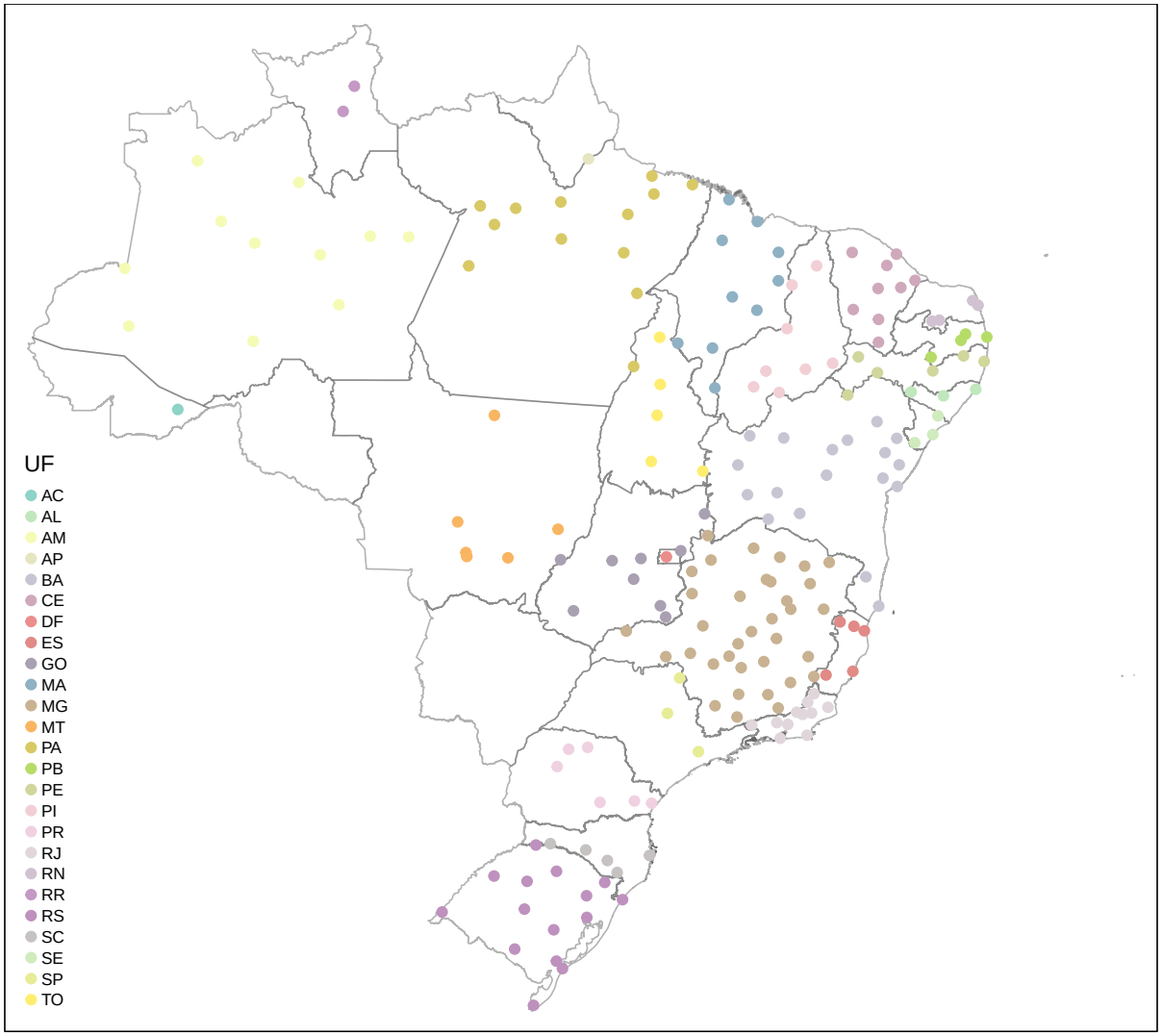
```
## tmap mode set to plotting
```

Loading the Brazil shapefile to plot the data points in it.

```
shp_tmin <- readOGR(dsn = "shp_br", layer = "UFEBRASIL", encoding = "UTF-8", use_iconv = T)
```

Matching the sf_tmin to the shp_tmin to plot both objects together

```
tm_shape(shp = shp_tmin) +
  tm_borders(alpha = 0.5) +
  tm_shape(shp = sf_tmin) +
  tm_dots(col = "UF",
          size = 0.2) +
  tmap_options(check.and.fix = TRUE)
```



Loading and cleaning the state_codes object

```
state_codes <- read.csv("estados.csv", header = T, encoding = "UTF-8") %>%
   rename("UF" = "SIGLA")

state_codes$UF <- str_replace_all(state_codes$UF, fixed(" "), "")
```

Making a SQL query to get the lowest temperature for each state and getting all the useful information about it at the same time

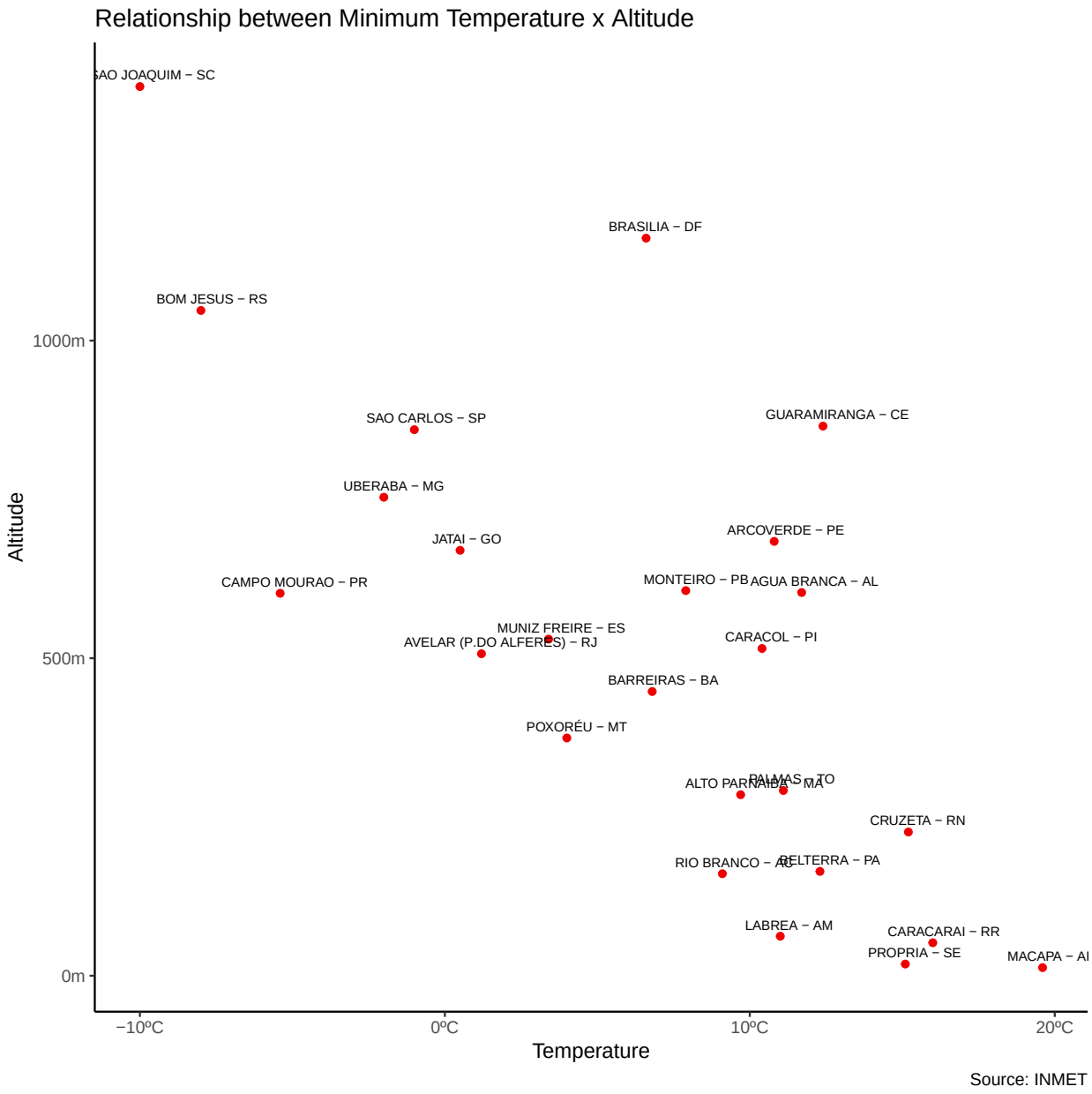
```
df_tmin_c1 <- df_tmin_c %>%
  rename("Code" = "Código",
         "Station" = "Nome da Estação")

df_tmin_complete <- sqldf("SELECT Code, Station, UF, Latitude, Longitude, Atitude,
                           MIN(lowest_temp) as Lowest
                           FROM df_tmin_c1
                           GROUP BY UF") %>%
  rename("Altitude" = "Atitude")
```

Merging the df_tmin_complete object to state_codes one to get all the necessary information needed for the project

```
df_tmin_complete <- left_join(df_tmin_complete,
                          state_codes,
                          by = "UF")
```

Plotting a scatter plot to check if there is any relationship between minimun temperature and altitude.

```
ggplot(df_tmin_complete, aes(Lowest ,Altitude)) +
  geom_point(color = "Red2") +
  geom_text(label = paste(df_tmin_complete$Station, df_tmin_complete$UF, sep = " - "),
            size = 2.4, vjust = -0.7, hjust = 0.4) +
  labs(
    title = "Relationship between Minimum Temperature x Altitude",
    x = "Temperature",
    caption = "Source: INMET"
  ) +
   scale_y_continuous(labels = paste0(seq(0, 1500, by = 500), "m")) +
  scale_x_continuous(labels = paste0(seq(-10,20, by = 10), "ºC")) +
  theme_classic()
```



As we can see above, there is a strong relationship between temperature and altitude when it comes to the lowest temperature registered by state. The lowest temperature comes from the higher weather station and the highest temperature comes from a low-altitude weather station.

Plotting the previous dataset into the map

```
#Preparing the dataset

sf_tmin_complete <- st_as_sf(x = df_tmin_complete,
                         coords = c("Longitude", "Latitude"),
                         crs = 4326)

df_tmin_complete2 <- df_tmin_complete %>%
  rename("CD_GEOCODU" = "COD")
```

```r
shp_brazil2 <- merge(shp_brazil,
                     df_tmin_complete2,
                     by = c("CD_GEOCODU"))

#Plotting the map

tm_shape(shp = shp_brazil2) +
  tm_borders(alpha = 0.5) +
  tm_polygons("Lowest",
              title = "Lowest Temperature by State (ºC)",
              style = "fixed",
              palette = c("blue4", "blue3", "royalblue1", "steelblue1",
                          "lightblue2", "cadetblue1", "khaki","darkorange1"),
              breaks = c(-Inf, -8, -4, 0, 4, 8, 12, 16, Inf),
              colorNA = "gray",
              textNA = "No available data for RO and MS") +
  tm_layout(main.title = "Weather stations that presented the
            lowest Temperatures by state from 1991 to 2020",
          title = "Only INMET Weather Stations available during this period",
                      title.size = 0.7) +
  tm_compass(type = "8star",
             show.labels = 3,
             size = 3,
             position = c(0.9, 0.05)) +
  tm_credits("Source:INMET",
             position = 0.9) +
  tmap_options(check.and.fix = T) +
  tm_layout(legend.title.size = 1,
            legend.text.size = 0.8) +
   tm_shape(shp = sf_tmin_complete) +
  tm_dots(col = "red",
          size = 0.4) +
  tm_text("Station", size = 0.5, col = "black", just = "bottom", ymod = 0.5)
```

Weather stations that presented the
lowest Temperatures by state from 1991 to 2020

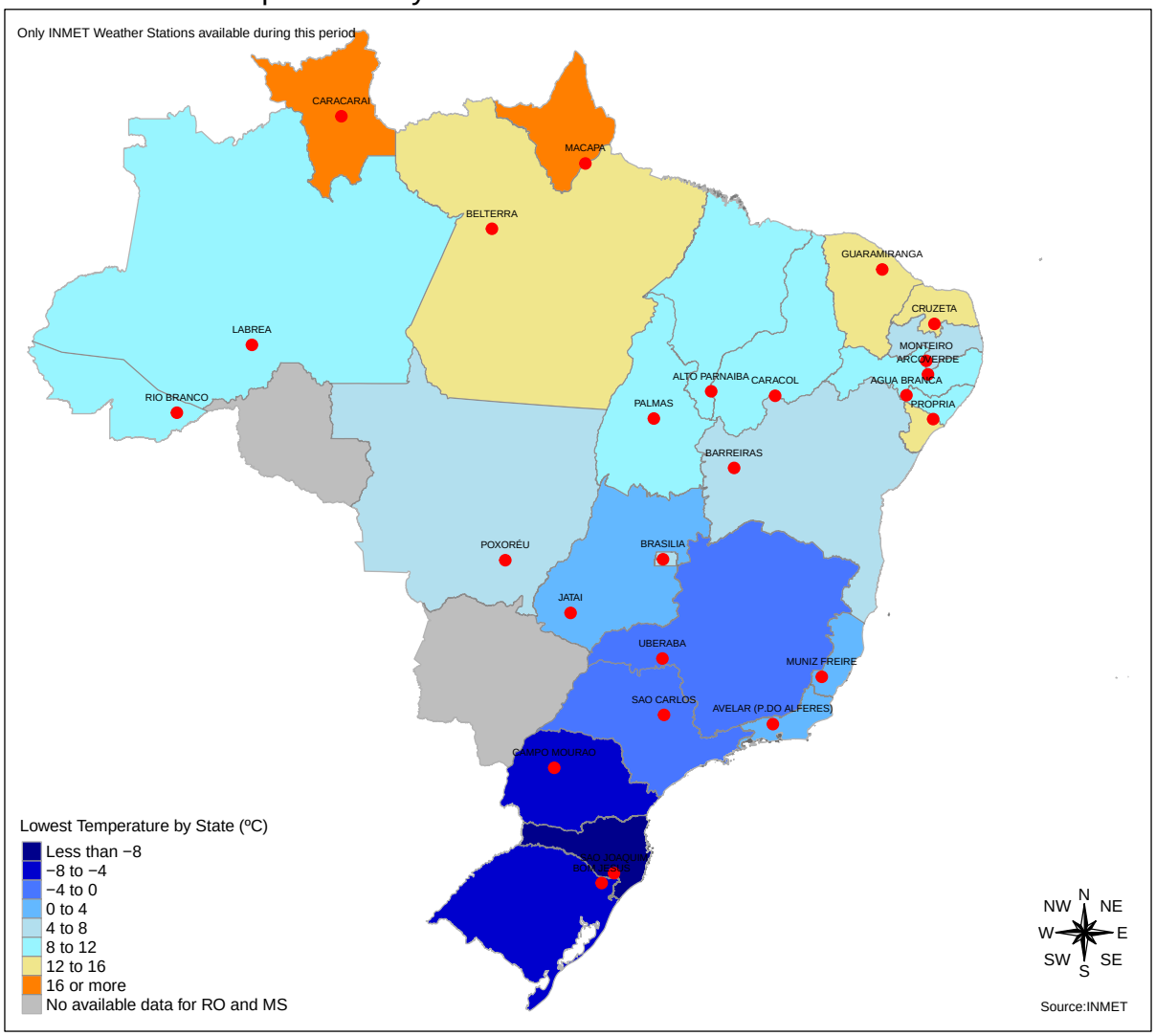
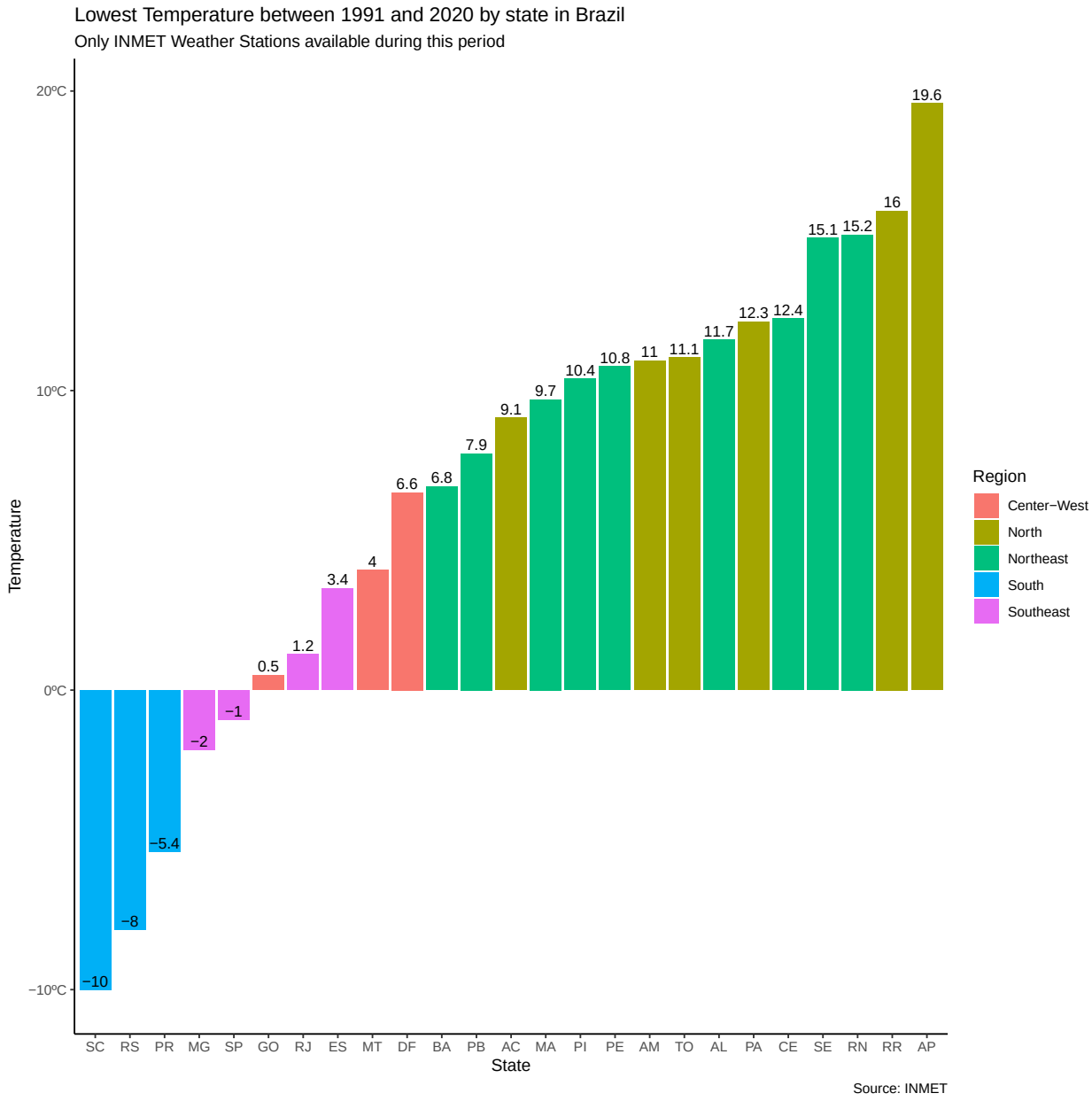


Table with some extra information about my findings

```
df_tmin_complete2$Region[df_tmin_complete2$UF %in% c("RS", "SC", "PR")] <- "South"
df_tmin_complete2$Region[df_tmin_complete2$UF %in% c("RJ", "SP", "MG", "ES")] <- "Southeast"
df_tmin_complete2$Region[df_tmin_complete2$UF %in% c("MT", "GO", "DF")] <- "Center-West"
df_tmin_complete2$Region[df_tmin_complete2$UF %in% c("AM", "RR", "AP", "PA", "TO", "AC")] <- "North"
df_tmin_complete2$Region[df_tmin_complete2$UF %in% c("AL", "SE", "BA", "PE",
                                                     "PB", "CE", "MA", "PI", "RN")] <- "Northeast"

ggplot(df_tmin_complete2, aes(reorder(UF, (Lowest)), Lowest, fill = Region)) +
  geom_col() +
  geom_text(label = df_tmin_complete2$Lowest, vjust = -0.3, size = 3.5) +
  scale_y_continuous(labels = paste0(seq(-10, 20, by = 10), "ºC")) +
  labs(
```

```
    title = "Lowest Temperature between 1991 and 2020 by state in Brazil",
    subtitle = "Only INMET Weather Stations available during this period",
    x = "State",
    y = "Temperature",
    caption = "Source: INMET"
  ) +
  theme_classic()
```



Part 3

In the third part of the project, I will analyze what was the highest temperature registered in the available INMET weather stations throughout the country, and after that, I will gather the data showing the highest temperature registered within each state of Brazil from 1991 to 2020.

There is no available data for the state of Rondônia (RO), for this reason, it is not mentioned in this part of

the project.

Loading, tidying and cleaning the dataset.

```
df_tmax <- read_xlsx("Normal-Climatologica-TMAXABS.xlsx", skip = 2, col_names = T)
```

```
## New names:
## * `` -> ...5
## * `` -> ...7
## * `` -> ...9
## * `` -> ...11
## * `` -> ...13
## * ...
```

```
df_tmax <- na.omit(df_tmax)
df_tmax <- df_tmax[ ,-29]

df_tmax[df_tmax == "-"] <- NA
```

Finding the highest temperature for each weather station and assigning it to a new column named "highest_temp" and selecting the variables that will be used in the following steps.

```
df_tmax2 <- df_tmax %>%
  mutate(highest_temp = apply(df_tmax[ ,4:15], 1, FUN = max, na.rm = TRUE))

df_tmax2 <- df_tmax2[ ,c(1:3, 29)]
```

Merging the weather_st object to the df_tmax2 one to get all the variables needed to get the desired outcome. The merging will be added to a new object named "df_tmax_c"

```
df_tmax_c <- left_join(df_tmax2,
          weather_st,
          by = "Código")
```

Creating a spatial feature (SF) object with the new object created above. SF objects allow us to plot data points into a real map of the earth's surface, in this case, the Brazilian one.

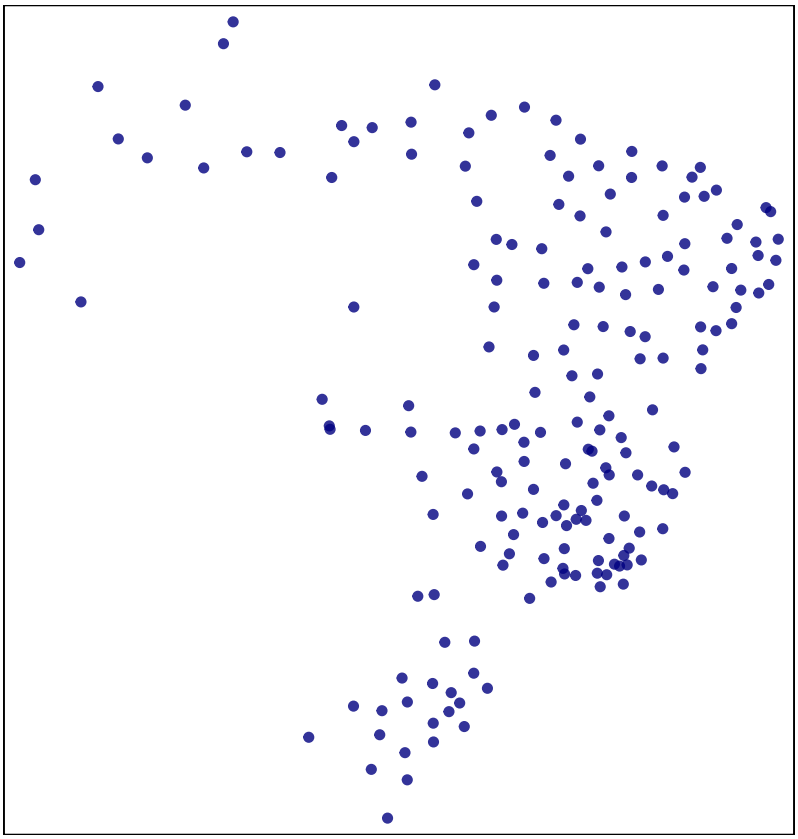
```
sf_tmax <- st_as_sf(x = df_tmax_c,
                         coords = c("Longitude", "Latitude"),
                         crs = 4326)
```

Viewing the same data points plotted on the Brazil surface.

```
tmap_mode("plot")
```

```
## tmap mode set to plotting
```

```
tm_shape(shp = sf_tmax) +
  tm_dots(col = "navy",
          border.col = "black",
          size = 0.1,
          alpha = 0.8)
```



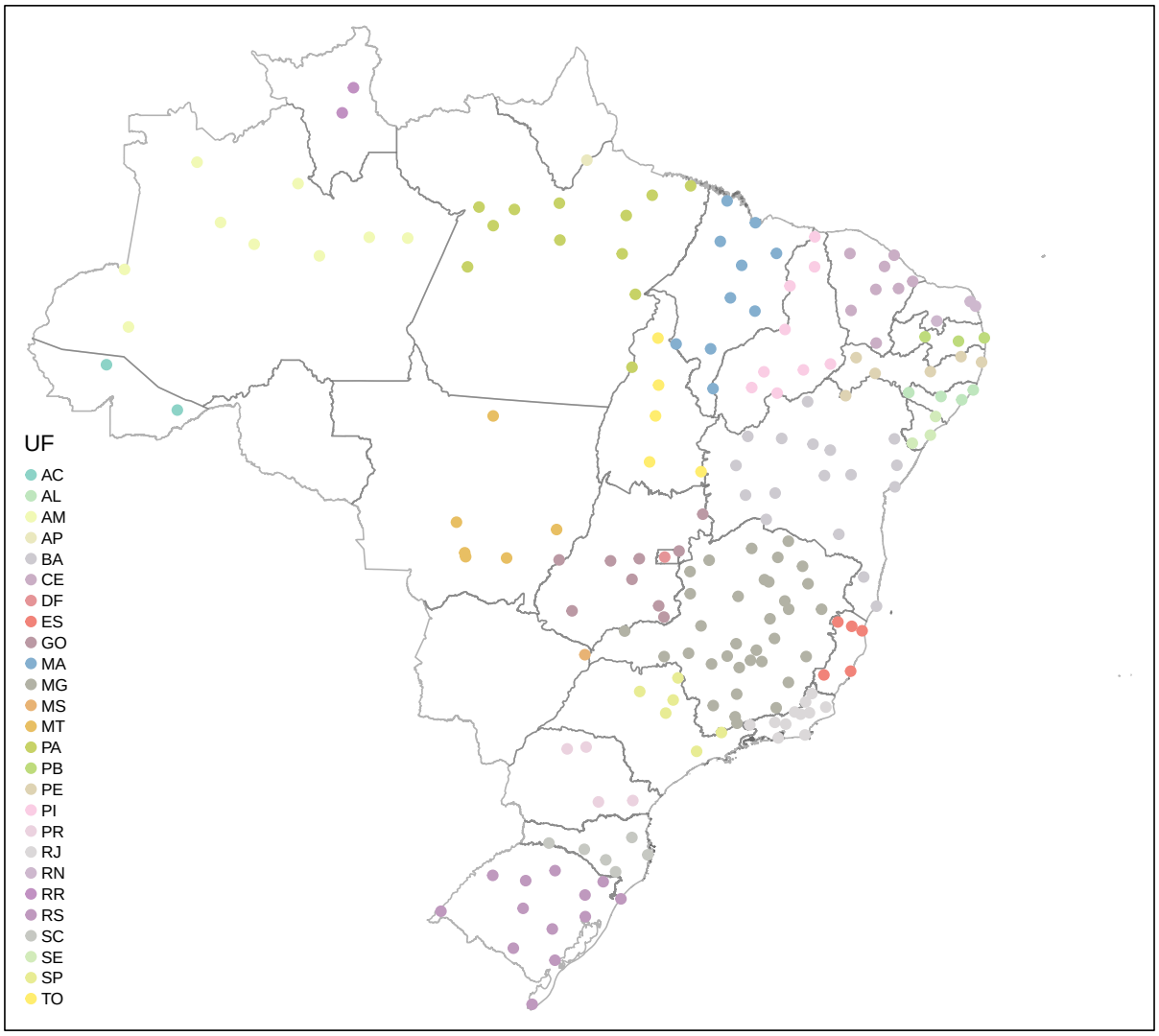
Changing the view mode to “plot”

```
tmap_mode("plot")
```

```
## tmap mode set to plotting
```

Matching the sf_tmax to the shp_tmax to plot both objects together

```
tm_shape(shp = shp_brazil) +
  tm_borders(alpha = 0.5) +
  tm_shape(shp = sf_tmax) +
  tm_dots(col = "UF",
          size = 0.2) +
  tmap_options(check.and.fix = TRUE)
```



Making a SQL query to get the lowest temperature for each state and getting all the useful information about it at the same time

```
df_tmax_c1 <- df_tmax_c %>%
  rename("Code" = "Código",
         "Station" = "Nome da Estação")

df_tmax_complete <- sqldf("SELECT Code, Station, UF, Latitude, Longitude,
                          Atitude, ROUND(MAX(highest_temp)) as Highest
                FROM df_tmax_c1
                GROUP BY UF") %>%
  rename("Altitude" = "Atitude")
```

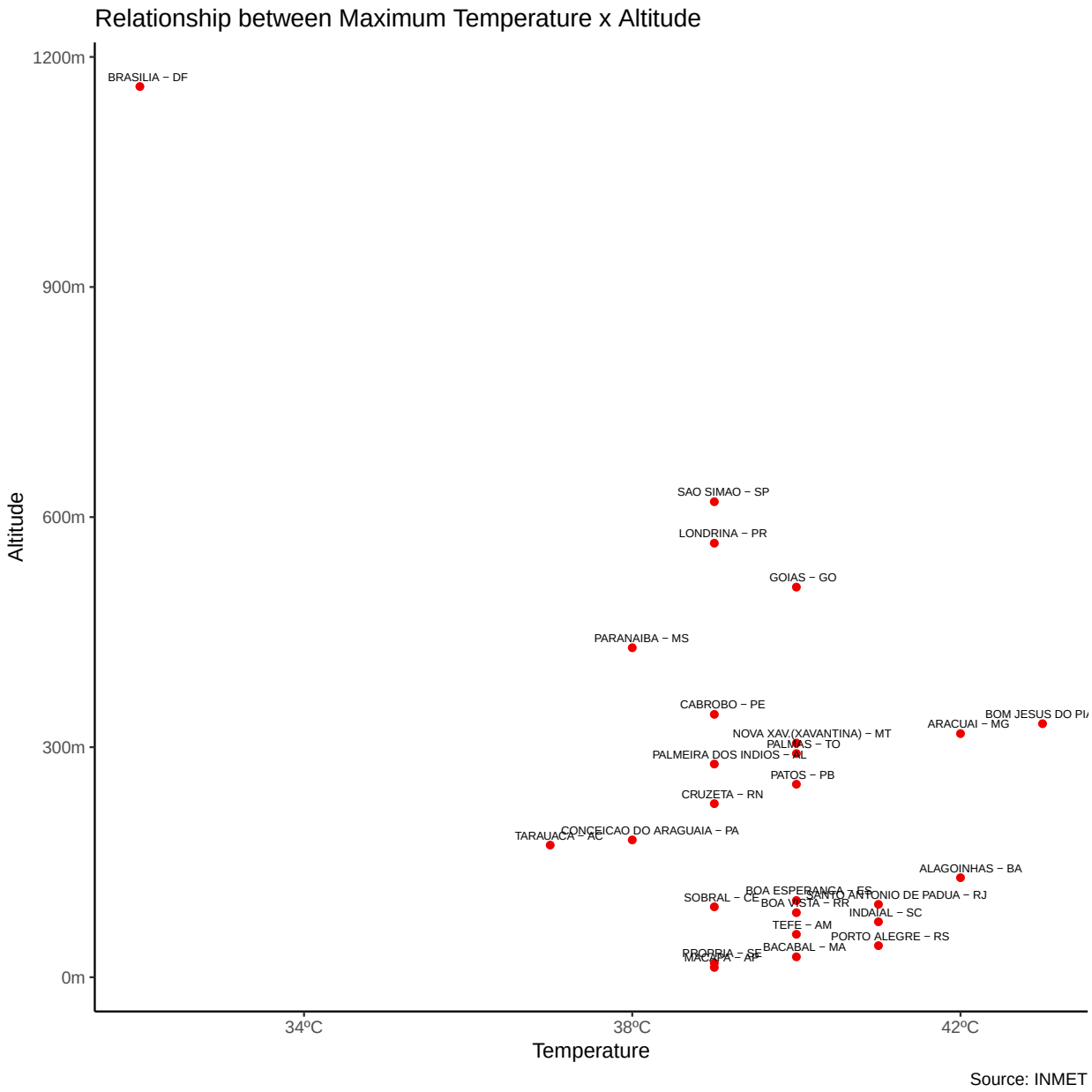
Merging the df_tmax_complete object to state_codes one to get all the necessary information needed for

the project

```
df_tmax_complete <- left_join(df_tmax_complete,
                          state_codes,
                          by = "UF")
```

Plotting a scatter plot to check if there is any relationship between maximum temperature and altitude.

```
ggplot(df_tmax_complete, aes(Highest ,Altitude)) +
  geom_point(color = "Red2") +
  geom_text(label = paste(df_tmax_complete$Station, df_tmax_complete$UF, sep = " - "),
            size = 2.1, vjust = -0.7, hjust = 0.4) +
  labs(
    title = "Relationship between Maximum Temperature x Altitude",
    x = "Temperature",
    caption = "Source: INMET"
  ) +
   scale_y_continuous(breaks = seq(0, 1200, by = 300),
                      labels = paste0(seq(0, 1200, by = 300), "m")) +
  scale_x_continuous(breaks = seq(30, 44, by = 4),
                      labels = paste0(seq(30, 44, by = 4), "ºC")) +
  theme_classic()
```



Again, as we can see above there is a strong relationship between high temperatures and altitude. The higher the altitude is, the less hot it will be. We can see Brasilia presented a very low maximum temperature during the last 30 years and that is thanks to the high altitude.

Plotting the previous dataset into the map to get a better overview.

```
#Preparing the dataset

sf_tmax_complete <- st_as_sf(x = df_tmax_complete,
                         coords = c("Longitude", "Latitude"),
                         crs = 4326)

df_tmax_complete2 <- df_tmax_complete %>%
  rename("CD_GEOCODU" = "COD")
```

```r
shp_brazil3 <- merge(shp_brazil,
                     df_tmax_complete2,
                     by = c("CD_GEOCODU"))

#Plotting the map

tm_shape(shp = shp_brazil3) +
  tm_borders(alpha = 0.5) +
  tm_polygons("Highest",
              title = "Highest Temperature by State (ºC)",
              style = "fixed",
              palette = c("YlOrRd"),
              breaks = c(-Inf, 36, 38, 40, 42, Inf),
              colorNA = "gray",
              textNA = "No available data for RO") +
  tm_layout(main.title = "Weather Stations that presented the highest
            Temperatures  by state from 1991 to 2020",
          title = "Only INMET Weather Stations available during this period",
                       title.size = 0.7) +
  tm_compass(type = "8star",
             show.labels = 3,
             size = 3,
             position = c(0.9, 0.05)) +
  tm_credits("Source:INMET",
             position = 0.9) +
  tmap_options(check.and.fix = T) +
  tm_layout(legend.title.size = 1.3,
            legend.text.size = 0.9) +
   tm_shape(shp = sf_tmax_complete) +
  tm_dots(col = "blue",
          size = 0.4) +
  tm_text("Station", size = 0.5, col = "black", just = "bottom", ymod = 0.5)
```

Weather Stations that presented the highest
Temperatures by state from 1991 to 2020

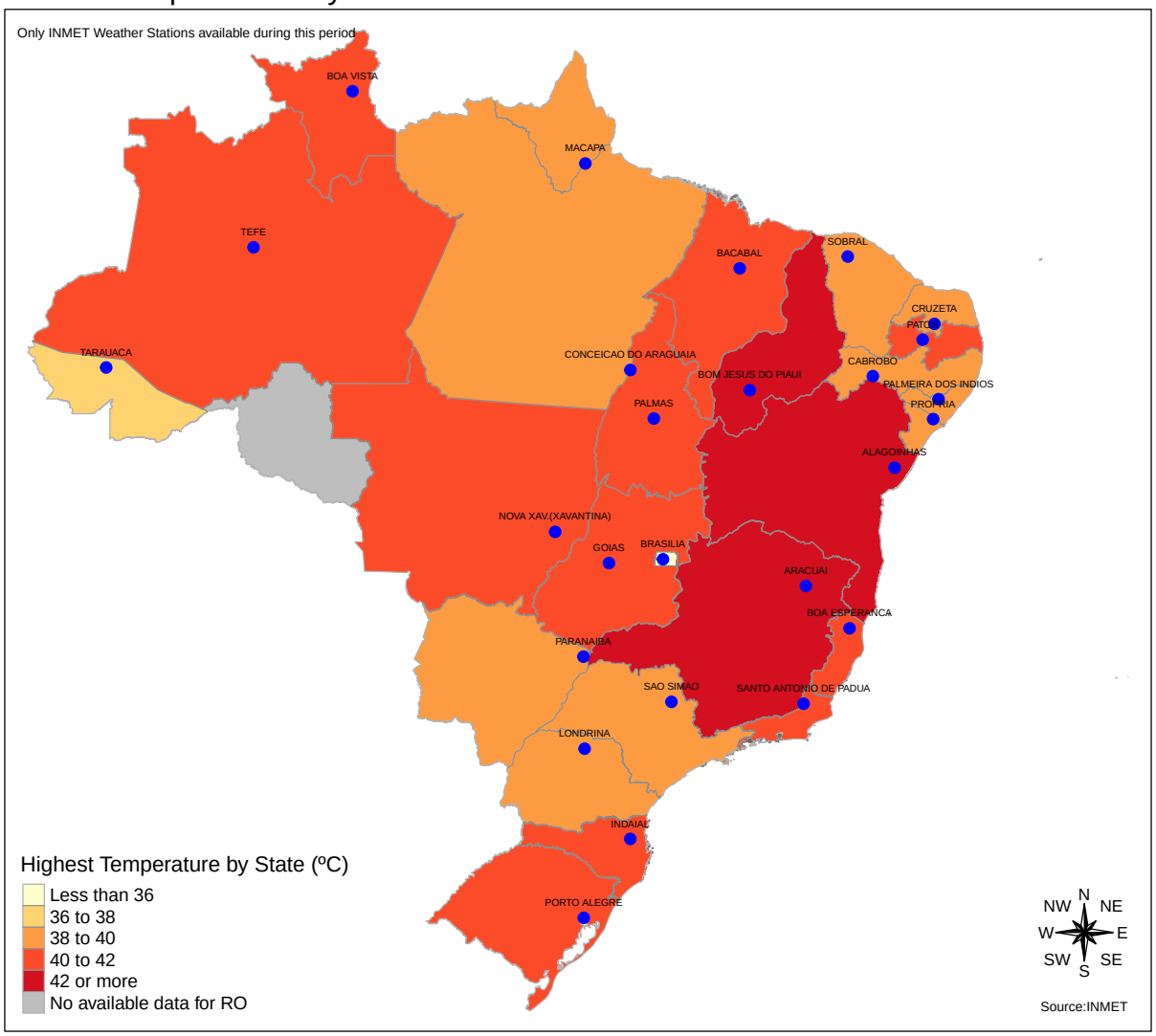
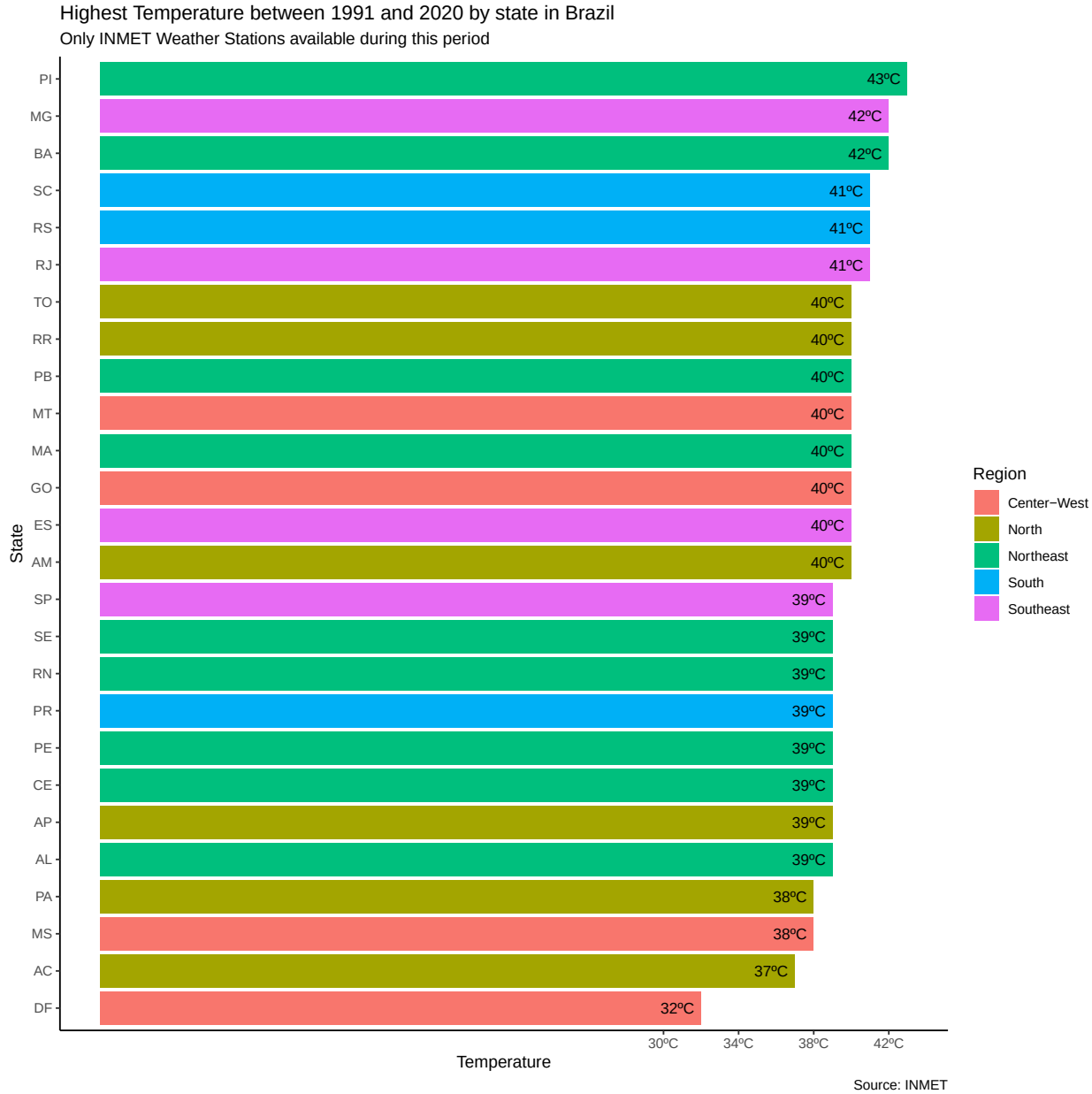


Table with some extra information about my findings

```
df_tmax_complete2$Region[df_tmax_complete2$UF %in% c("RS", "SC", "PR")] <- "South"
df_tmax_complete2$Region[df_tmax_complete2$UF %in% c("RJ", "SP", "MG", "ES")] <- "Southeast"
df_tmax_complete2$Region[df_tmax_complete2$UF %in% c("MT", "GO", "DF", "MS")] <- "Center-West"
df_tmax_complete2$Region[df_tmax_complete2$UF %in% c("AM", "RR", "AP", "PA", "TO", "AC")] <- "North"
df_tmax_complete2$Region[df_tmax_complete2$UF %in% c("AL", "SE", "BA", "PE",
                                                     "PB", "CE", "MA", "PI", "RN")] <- "Northeast"

ggplot(df_tmax_complete2, aes(reorder(UF, (Highest)), Highest, fill = Region)) +
  geom_col() +
  geom_text(label = paste0(df_tmax_complete2$Highest, "ºC"), hjust = 1.2, size = 3.5) +
  coord_flip() +
  labs(title = "Highest Temperature between 1991 and 2020 by state in Brazil",
```

```
  subtitle = "Only INMET Weather Stations available during this period",
  x = "State",
  y = "Temperature",
  caption = "Source: INMET") +
theme_classic() +
scale_y_continuous(breaks = seq(30, 44, by = 4),
                   labels = paste0(seq(30, 44, by = 4), "ºC"))
```



Conclusion

As it is shown in this project, the weather conditions in Brazil vary substantially according to the country's regions and states. The is more rain in the northern states than the northeast ones because the Amazon, which is a rainforest, is located in this region. On the other hand, there is a large semiarid area in the northeast of Brazil, that's why there is way less precipitation over there.

When it comes to temperature, Brazil is also diverse. The southern states, which have a subtropical climate, are considerably cold during autumn and winter seasons, and temperature can drop to -10^{o}C in this region. In the north part of the country, Northeast states, which have a tropical climate, present elevated temperatures that can surpass 42^{o}C during summer days.