Abinash Satapathy

Reg. no.: 16BCE0081

CSE3020 – Data Visualisation

Slot: C1

Faculty: Lydia Jane G

Visualisation of Global Temperature using R

1. Dataset details

a. URL: https://www.kaggle.com/berkeleyearth/climate-change-earth-surface-temperature-data/downloads/climate-change-earth-surface-temperature-data.zip/2

b. No. of attributes = 7
No. of rows = 239, 178

c. Attributes definition:

Date	The date whose temperature is included in the dataset
Average temperature	The mean temperature for that day
Uncertainty in calculation	The error amount in calculation of the temperature
City	Place whose temperature is taken into consideration
Country	The country where the place is situated
Latitude	The angular distance of a place north or south of the earth's equator
longitude	The angular distance of a place east or west of the Greenwich meridian

2. Data Abstraction

Date	Quantitative
Average temperature	Quantitative
Uncertainty in calculation	Quantitative
City	Qualitative (nominal)
Country	Qualitative (nominal)
Latitude	Quantitative
longitude	Quantitative

3. Task Abstraction

a. Potential questions that can be asked by looking at the dataset

(i) What is the average temperature of a month of all the cities of a country during a particular year range?

(ii) What is the year-wise average of different cities of a country?

(iii) What are the peak temperatures of a set of cities?
(iv) What is the temperature rise/fall over a wide range of years (like 50 to 100 years range)?
(v) Which country has the highest mean temperature for a specific month?
(vi) Which country has the highest range of mean temperatures?
(vii) What is the range of latitudes where the mean temperature is greater than some 'x' value?
(viii) Is climate change really happening?
(ix) What is the annual increase/decrease of the mean temperature of a country?
(x) What is the difference pattern of temperature differences between mean temperature of a city and human body, where 37°C is the normal temperature of a human body?

b. Attributes required for answering the above questions and others as well

(i) The Average Temperature attribute is the most important numerical data in the dataset. So, usually all the queries will include the temperature data.
(ii) The city attribute is important for questions targeted to a particular city or a set of cities.
(iii) The country attribute is of equal importance as the city, as it helps in getting patterns on a global scale.

The actions that would be performed mostly are:

- Plotting the temperature
- Calculating the range of mean temperatures by Highest – Lowest formula
- Standard deviation formula might be used in case a query regarding that is asked.

$$SD = \sqrt{\frac{\sum |x - \bar{x}|^2}{n}}$$

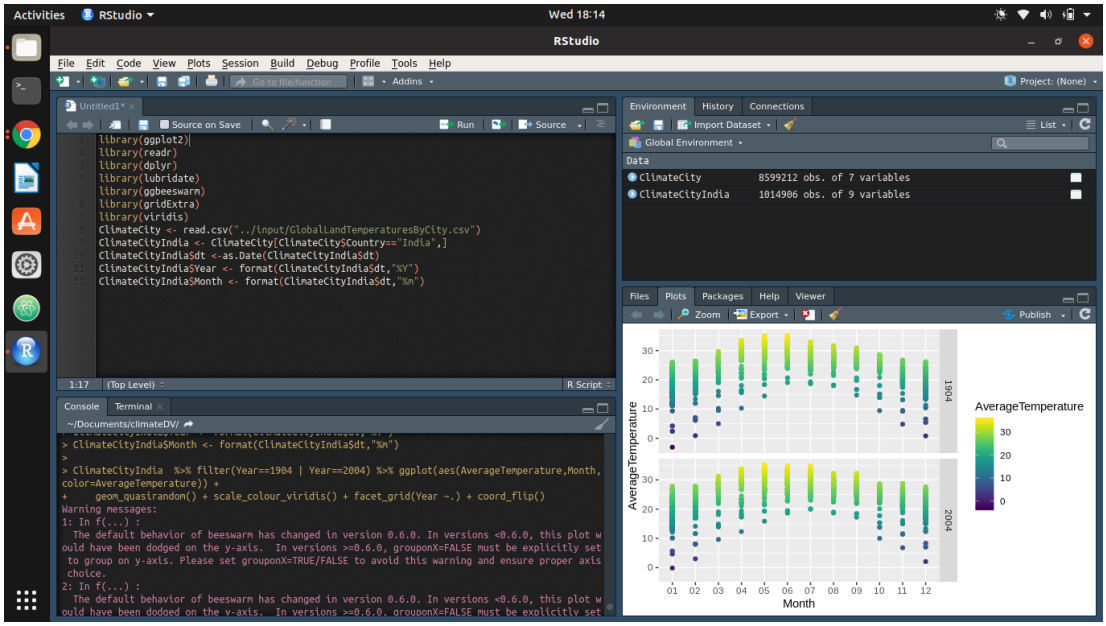
4. Encoding

a. Appropriate encoding techniques

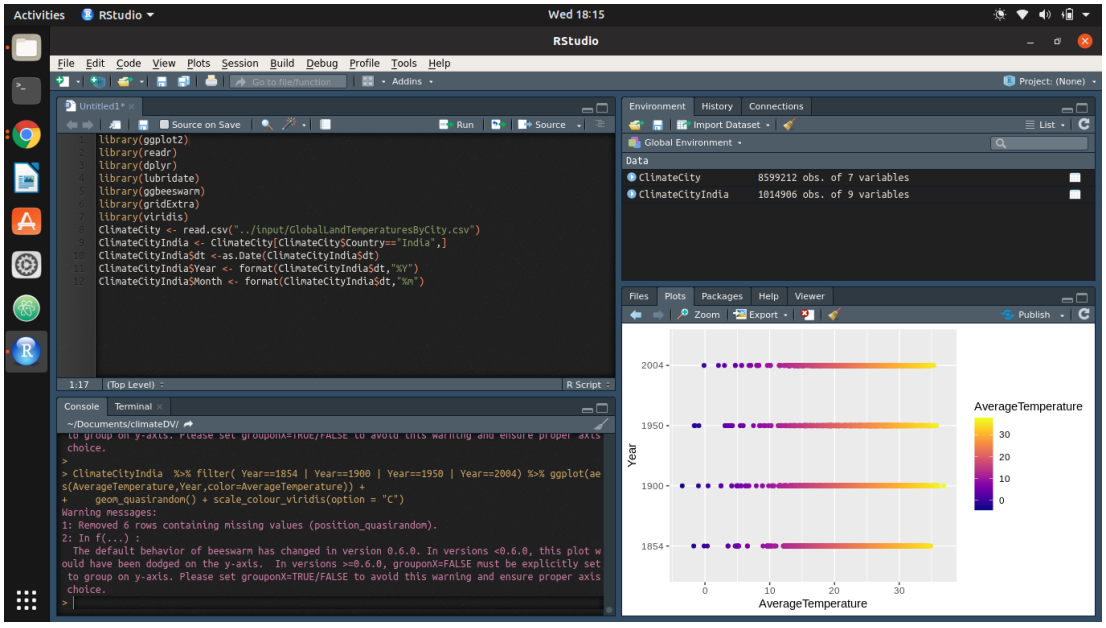
(i) Jitter plot
(ii) Line graph
(iii) Heat map
(iv) Density plot
(v) Scatterplot

b. Example graphs for analysis of certain questions

(i) ClimateCityIndia %>% filter(Year==1904 | Year==2004) %>% ggplot(aes(AverageTemperature,Month,color=AverageTemperature)) + geom_quasirandom() + scale_colour_viridis() + facet_grid(Year ~.) + coord_flip()

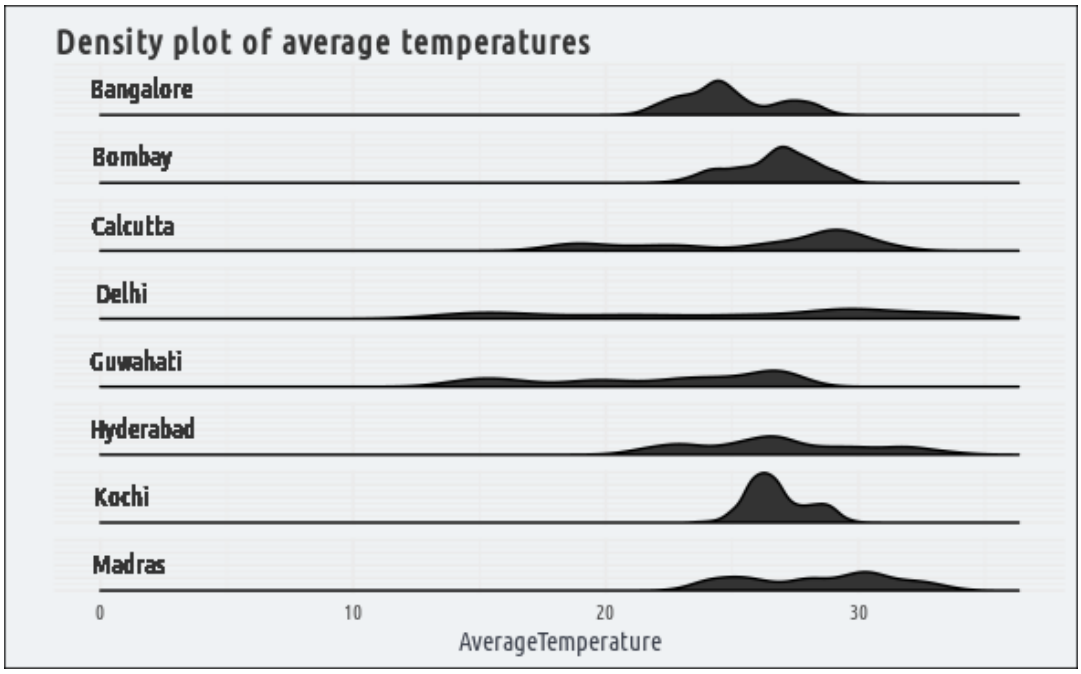


(ii) ClimateCityIndia %>% filter(Year==1854 | Year==1900 | Year==1950 | Year==2004) %>% ggplot(aes(AverageTemperature,Year,color=AverageTemperature)) + geom_quasirandom() + scale_colour_viridis(option = "C")



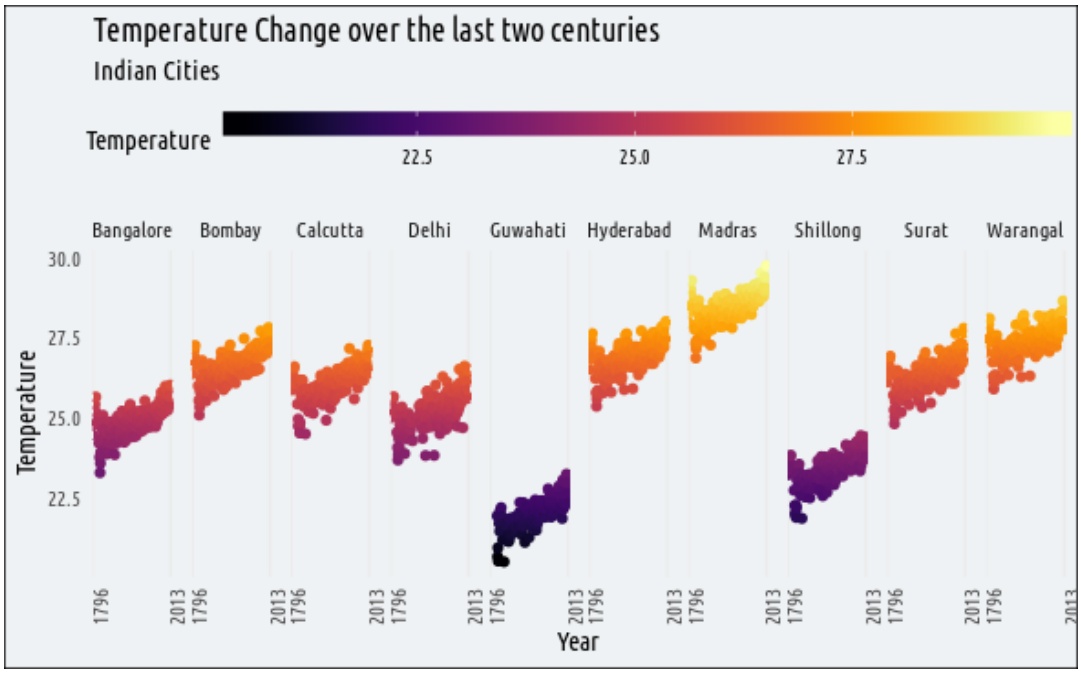
(iii) Major_Cities<- c("Madras","Hyderabad", "Bombay", "Bangalore" , "Delhi" ,"Calcutta","Guwahati","Kochi")
df <- ClimateCityIndia %>% filter(City %in% Major_Cities)
ggplot(df,aes(AverageTemperature,label=paste(City,""))) + geom_density(fill="gray20") + facet_grid(City ~.) + ylab("") + geom_text(aes(0,0.2),family="Ubuntu Condensed",fontface="plain",hjust=0.1,color="gray20") + theme_minimal() + ggtitle("Density plot of average temperatures") + theme(text = element_text(family = 'Ubuntu Condensed', face = 'plain', color = '#3A3F4A'), axis.text.y = element_blank(), axis.ticks.y = element_blank(), axis.line.y = element_blank(), strip.background =

element_blank(), strip.text.y = element_blank(), axis.line.x = element_blank(), plot.background = element_rect(fill = "#EFF2F4"), plot.title = element_text(size = 14, face = "bold", colour = "gray20", vjust = -1))

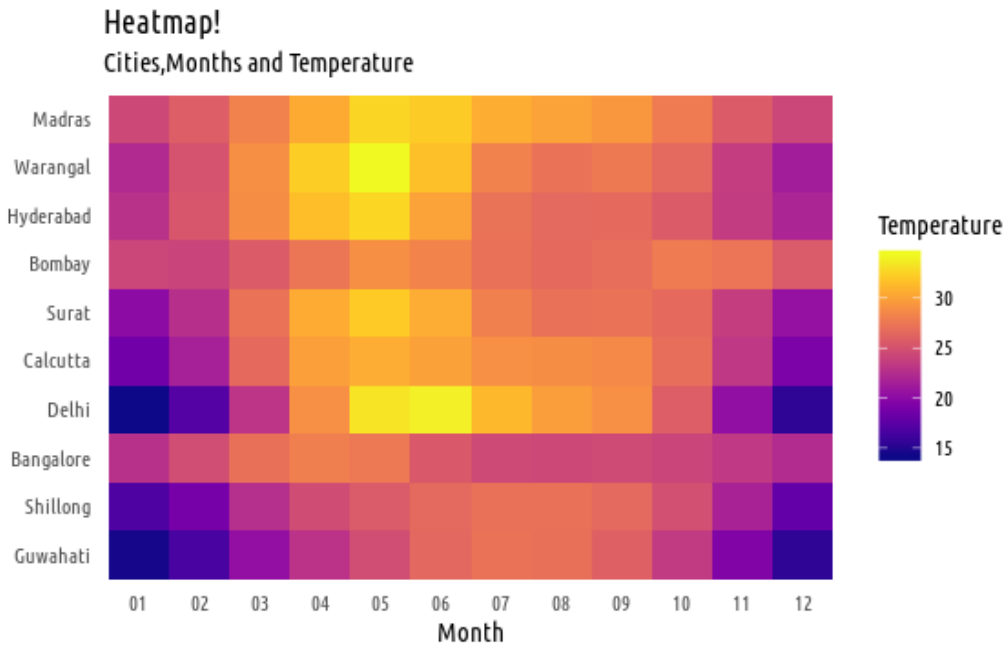


(iv) Major_Cities<- c("Madras","Hyderabad", "Bombay", "Bangalore" , "Delhi" ,"Calcutta","Guwahati","Shillong","Surat","Warangal")

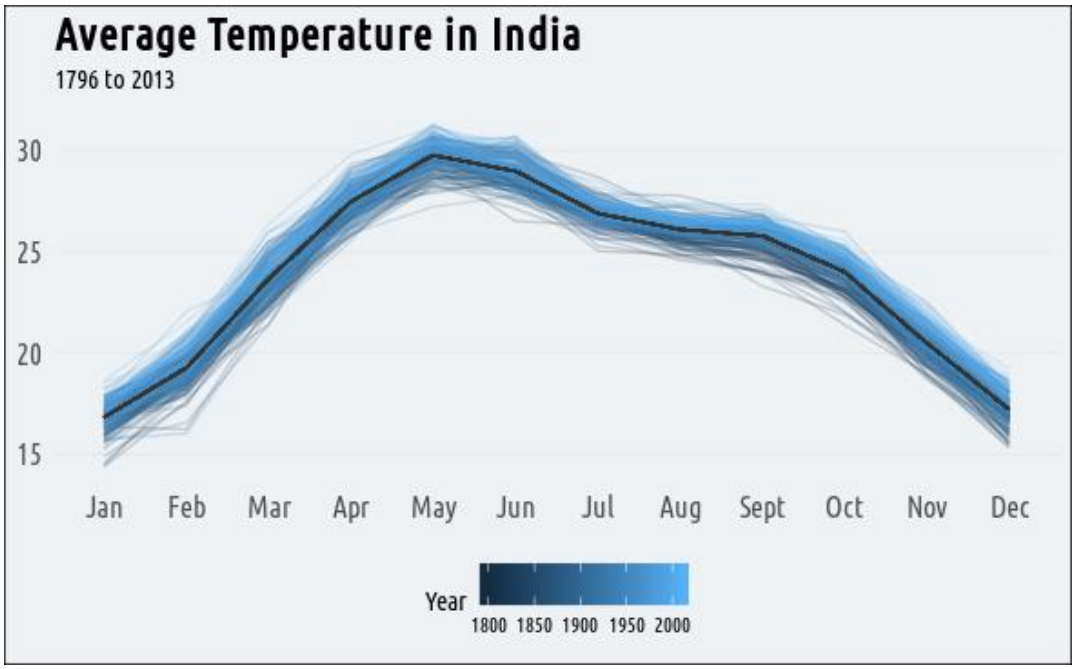
ClimateCityIndia %>% filter(City %in% Major_Cities) %>% group_by(City,Year) %>% summarise(avg_temp= mean(AverageTemperature)) %>% ggplot(aes(Year,avg_temp,color=avg_temp)) + geom_point() + scale_x_discrete(breaks = c(1796,2013)) + theme_minimal(base_family = "Ubuntu Condensed") + scale_color_viridis(option="B") + theme(axis.text.x = element_text(angle = 90, hjust = 1), panel.spacing = unit(0.6, "lines"), legend.position = "top", legend.text=element_text(size=8), panel.grid.major.y = element_blank(), panel.grid.minor.y = element_blank(), axis.ticks.y = element_blank(), axis.ticks.x = element_blank(), plot.background = element_rect(fill = "#EFF2F4"), legend.key.height = unit(10, "pt"), legend.key.width = unit(70, "pt")) + ggtitle("Temperature Change over the last two centuries",subtitle = "Indian Cities") + ylab("Temperature")+labs(color="Temperature") + facet_grid(~City)



(v) ClimateCityIndia %>% filter(City %in% Major_Cities) %>% group_by(City,Month) %>% summarise(avg_temp = mean(AverageTemperature, na.rm=T)) %>% ggplot(aes(Month, reorder(City, avg_temp), fill=avg_temp)) + geom_tile() + scale_fill_viridis(option = "C") + theme_minimal(base_family = "Ubuntu Condensed")+ labs(fill= "Temperature ")+ ylab("") + theme(panel.grid= element_blank()) + ggtitle("Heatmap!",subtitle = "Cities,Months and Temperature")



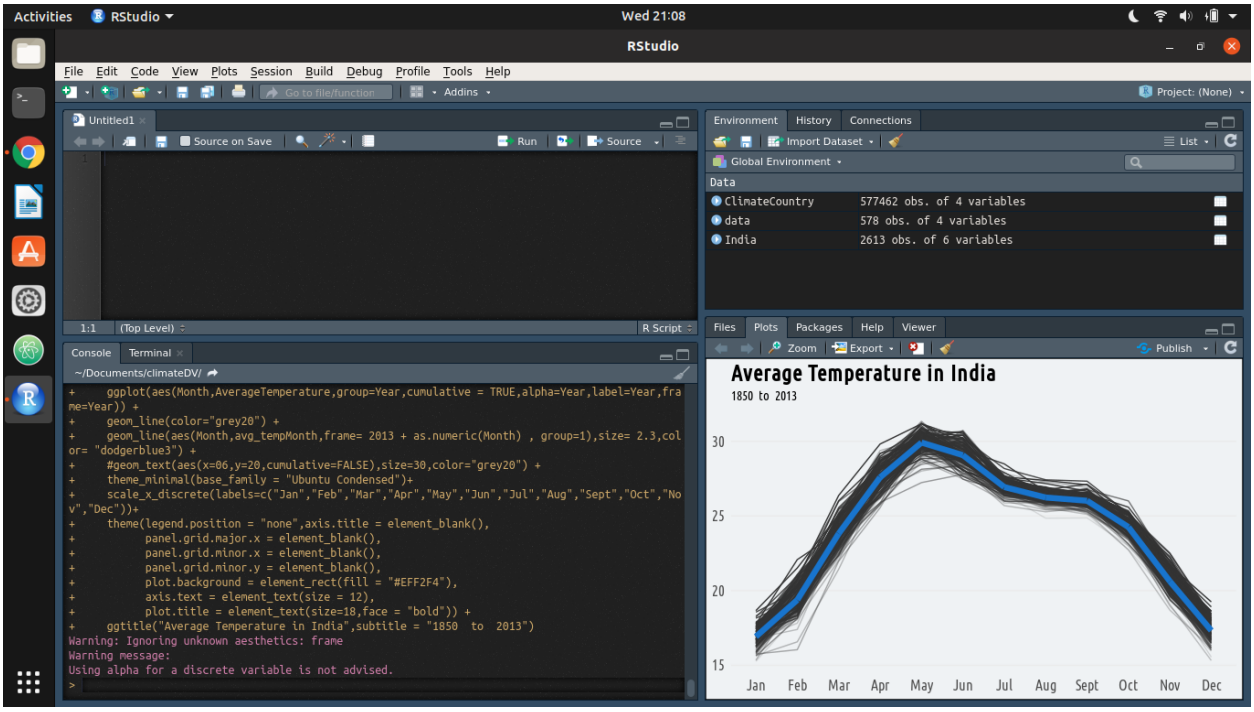
(vi) Average Temperature in India




```
library(ggplot2)
library(readr)
library(dplyr)
library(RColorBrewer)
library(gganimate)
India <- ClimateCountry %>% filter(Country=="India")
India$dt <-as.Date(India$dt)
India$Year <- format(India$dt,"%Y")
India$Month <- format(India$dt,"%m")



India %>% filter(!is.na(AverageTemperature))  %>%
  filter(Year > 1850) %>%
  group_by(Year) %>% mutate(no_of_cases= length(Year)) %>%
  group_by(Month) %>%
  mutate(avg_tempMonth= mean(AverageTemperature)) %>%
  filter(no_of_cases==12) %>%
  ggplot(aes(Month,AverageTemperature,group=Year,cumulative =
TRUE,alpha=Year,label=Year,frame=Year)) +
```

```
  geom_line(color="grey20") +
  geom_line(aes(Month,avg_tempMonth,frame= 2013 + as.numeric(Month) , group=1),size= 2.3,color=
"dodgerblue3") +
  #geom_text(aes(x=06,y=20,cumulative=FALSE),size=30,color="grey20") +
  theme_minimal(base_family = "Ubuntu Condensed")+
  scale_x_discrete(labels=c("Jan","Feb","Mar","Apr","May","Jun","Jul","Aug","Sept","Oct","Nov","Dec"))+
  theme(legend.position = "none",axis.title = element_blank(),
        panel.grid.major.x = element_blank(),
        panel.grid.minor.x = element_blank(),
        panel.grid.minor.y = element_blank(),
        plot.background = element_rect(fill = "#EFF2F4"),
        axis.text = element_text(size = 12),
        plot.title = element_text(size=18,face = "bold")) +
  ggtitle("Average Temperature in India",subtitle = "1850  to  2013") +
labs(caption= "Vamsi Krishna \t Source: Kaggle")



p<-
  India %>% filter(!is.na(AverageTemperature))  %>%
  filter(Year > 1850) %>%
  group_by(Year) %>% mutate(no_of_cases= length(Year)) %>%
  group_by(Month) %>%
  mutate(avg_tempMonth= mean(AverageTemperature)) %>%
  filter(no_of_cases==12) %>%
  ggplot(aes(Month,AverageTemperature,group=Year,cumulative =
TRUE,alpha=Year,label=Year,frame=Year)) +
  geom_line(color="grey20") +
  geom_line(aes(Month,avg_tempMonth,frame= 2013 + as.numeric(Month) , group=1),size= 2.3,color=
"dodgerblue3") +
  geom_text(aes(x=06,y=20,cumulative=FALSE),size=30,color="grey20") +
  theme_minimal(base_family = "Ubuntu Condensed")+
```

```
  scale_x_discrete(labels=c("Jan","Feb","Mar","Apr","May","Jun","Jul","Aug","Sept","Oct","Nov","Dec"))+
  theme(legend.position = "none",axis.title = element_blank(),
        panel.grid.major.x = element_blank(),
        panel.grid.minor.x = element_blank(),
        panel.grid.minor.y = element_blank(),
        plot.background = element_rect(fill = "#EFF2F4"),
        axis.text = element_text(size = 12),
        plot.title = element_text(size=18,face = "bold")) +
  ggtitle("Average Temperature in India",subtitle = "1850 to 2013") +
labs(caption= "Vamsi Krishna \t Source: Kaggle")



gganimate(p,"Output.gif",ani.width=810, ani.height=520, interval=0.1,title_frame = F)
```



It is a gif.

library(sp)

library(maps)

library(raster)

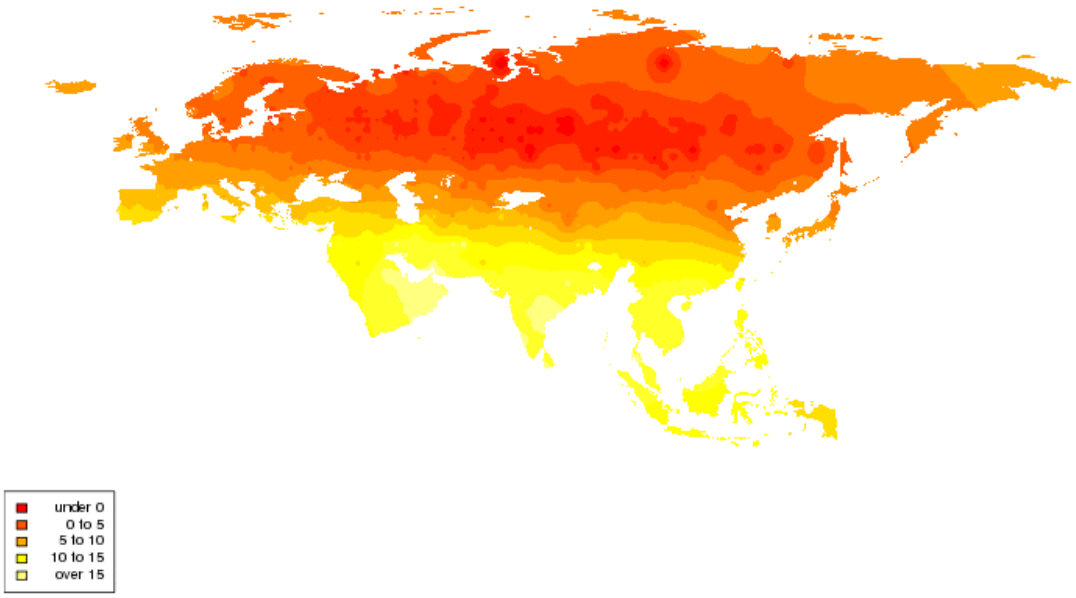
```
library(gstat)
library(geoR)
library(maptools)
library(GISTools)
library(animation)
#library(readr)

glt$O1 <- substr(glt$Longitude, nchar(glt$Longitude), nchar(glt$Longitude))
glt$O2 <- substr(glt$Latitude, nchar(glt$Latitude), nchar(glt$Latitude))
glt$Longitude <- ifelse(glt$O1 == "W",
                        -as.numeric(substr(glt$Longitude, 1, nchar(glt$Longitude)-1)),
                        as.numeric(substr(glt$Longitude, 1, nchar(glt$Longitude)-1)))

glt$Latitude <- ifelse(glt$O2 == "S",
                        -as.numeric(substr(glt$Latitude, 1, nchar(glt$Latitude)-1)),
                        as.numeric(substr(glt$Latitude, 1, nchar(glt$Latitude)-1)))

glt$Year <- substr(glt$dt, 1, 4)

glt.continent = data.frame(
Country = c(
"Algeria","Angola","Benin","Botswana","Burkina Faso","Burundi","Cameroon","Central African
Republic","Chad","Congo","Congo (Democratic Republic Of The)","Côte
D'Ivoire","Djibouti","Egypt","Equatorial
Guinea","Eritrea","Ethiopia","Gabon","Gambia","Ghana","Guinea","Guinea
Bissau","Kenya","Lesotho","Liberia","Libya","Madagascar","Malawi","Mali","Mauritania","Mauritius","Moro
cco","Mozambique","Namibia","Niger","Nigeria","Reunion","Rwanda","Senegal","Sierra
Leone","Somalia","South
Africa","Sudan","Swaziland","Tanzania","Togo","Tunisia","Uganda","Zambia","Zimbabwe","Afghanistan","B
ahrain","Bangladesh","Burma","Cambodia","China","Hong
Kong","India","Indonesia","Iran","Iraq","Israel","Japan","Jordan","Kazakhstan","Laos","Lebanon","Malaysia",
"Mongolia","Nepal","Oman","Pakistan","Philippines","Qatar","Russia","Saudi Arabia","Singapore","South
Korea","Sri Lanka","Syria","Taiwan","Tajikistan","Thailand","Turkey","Turkmenistan","United Arab
Emirates","Uzbekistan","Vietnam","Yemen","Albania","Armenia","Austria","Azerbaijan","Belarus","Belgium"
,"Bosnia And Herzegovina","Bulgaria","Croatia","Cyprus","Czech
Republic","Denmark","Estonia","Finland","France","Georgia","Germany","Greece","Hungary","Iceland","Irela
```

```
nd","Italy","Latvia","Lithuania","Macedonia","Moldova","Montenegro","Netherlands","Norway","Poland","Po
rtugal","Romania","Serbia","Slovakia","Slovenia","Spain","Sweden","Switzerland","Ukraine",

   #"United Kingdom",

   "UK",

   "Bahamas","Canada","Costa Rica","Cuba","Dominican Republic","El
Salvador","Guatemala","Haiti","Honduras","Jamaica","Mexico","Nicaragua","Panama","Puerto Rico",

   #"United States",

   "USA",

   "Australia","New Zealand","Papua New
Guinea","Argentina","Bolivia","Brazil","Chile","Colombia","Ecuador","Guyana","Paraguay","Peru","Suriname
","Uruguay","Venezuela"

),

Continent = c(

"Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Af
rica","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa
","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa","
Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Africa","Afr
ica","Asia","Asia","Asia","Asia","Asia","Asia","Asia","Asia","Asia","Asia","Asia","Asia","Asia","Asia","Asia
","Asia","Asia","Asia","Asia","Asia","Asia","Asia","Asia","Asia","Asia","Asia","Asia","Asia","Asia","Asia","
Asia","Asia","Asia","Asia","Asia","Asia","Asia","Asia","Asia","Europe","Europe","Europe","Europe","Europe
","Europe","Europe","Europe","Europe","Europe","Europe","Europe","Europe","Europe","Europe","Europe","
Europe","Europe","Europe","Europe","Europe","Europe","Europe","Europe","Europe","Europe","Europe","Eur
ope","Europe","Europe","Europe","Europe","Europe","Europe","Europe","Europe","Europe","Europe","Europe
","Europe","North America","North America","North America","North America","North America","North
America","North America","North America","North America","North America","North America","North
America","North America","North America","North America","Oceania","Oceania","Oceania","South
America","South America","South America","South America","South America","South America","South
America","South America","South America","South America","South America","South America America"

))

glt.y <- aggregate(list(

   AvgTemp    = glt$AverageTemperature,

   AvgTempUnc = glt$AverageTemperatureUncertainty

),

by = list(

   Year      = round(as.numeric(glt$Year) / 2) * 2,

   Latitude  = round(glt$Latitude  / 2) * 2,

   Longitude = round(glt$Longitude / 2) * 2),

FUN = mean, na.rm = TRUE
```

```
)

glt.y <- subset(glt.y, !is.na(glt.y$AvgTemp))
coordinates(glt.y) <- ~ Longitude + Latitude



get.countries <- function(continents, resolution) {

   countries <- as.character(glt.continent[glt.continent$Continent %in% continents, ]$Country)

   # Chose map
   w <- map("world", fill = TRUE, plot = FALSE)
   IDs <- sapply(strsplit(w$names, ":"), function(x) x[1])
   w <- map2SpatialPolygons(w, IDs = IDs)

   # Subsetting by countries
   w <- w[names(w) %in% countries, ]

   ext.bez  <- extent(w)
   xy       <- abs(apply(as.matrix(bbox(ext.bez)), 1, diff))
   r        <- raster(ext.bez, ncol=xy[1], nrow=xy[2])
   res(r)   <- resolution

   # Rasterize
   ras <-rasterize(w, r)
   ras <- as(ras, "SpatialPixels")

   years <- unique(subset(glt.y@data, !is.na(AvgTemp))$Year)  # No data for 1746?
   ani.options(interval = 0.1)
```

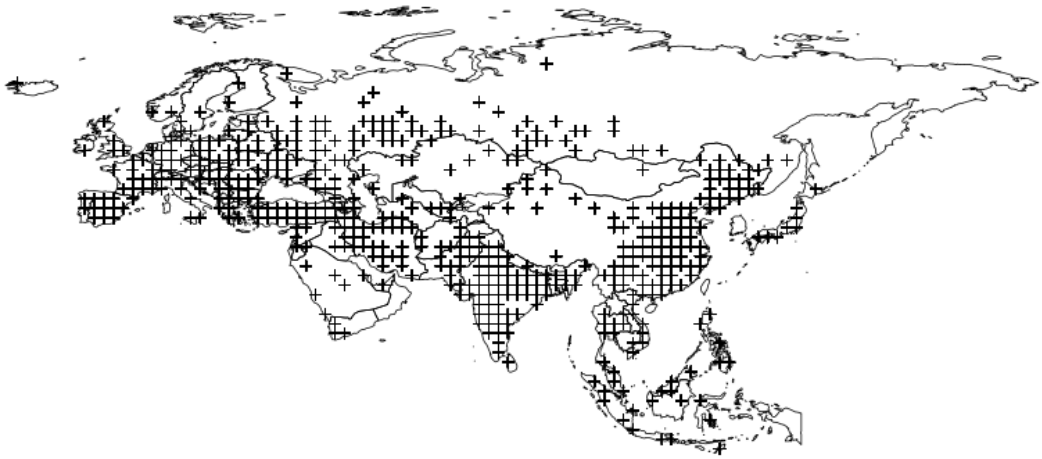
```
  saveGIF( {
    sapply(years, function(x) {

    d <- subset(glt.y, Year==x)
    m <- "Average Land Temperature in Celsius"
    #s <- seq(-5, 30, by=2.5)

    idw <- krige(AvgTemp ~ 1, d, ras, block= c(100))

    image(idw, xlab=x, asp=1, main=m)

    sh <- shading(breaks = seq(0, 15, by=5), cols = heat.colors(5))
    choro.legend(px = "bottomleft", sh = sh, bg = "white", cex = 0.75)

      ani.pause()
    })
  },
  movie.name = paste(paste(continents, collapse="_"), ".gif", sep = ""),
  img.name = "Rplot",
  convert = "convert",
  ani.width = 800,
  ani.height = 600)

  plot(w, main = "The Grid")
  points( glt.y[!is.na(over(glt.y, w)), ], pch=3)
}

get.countries(c("Europe", "Asia"), 0.3)
```

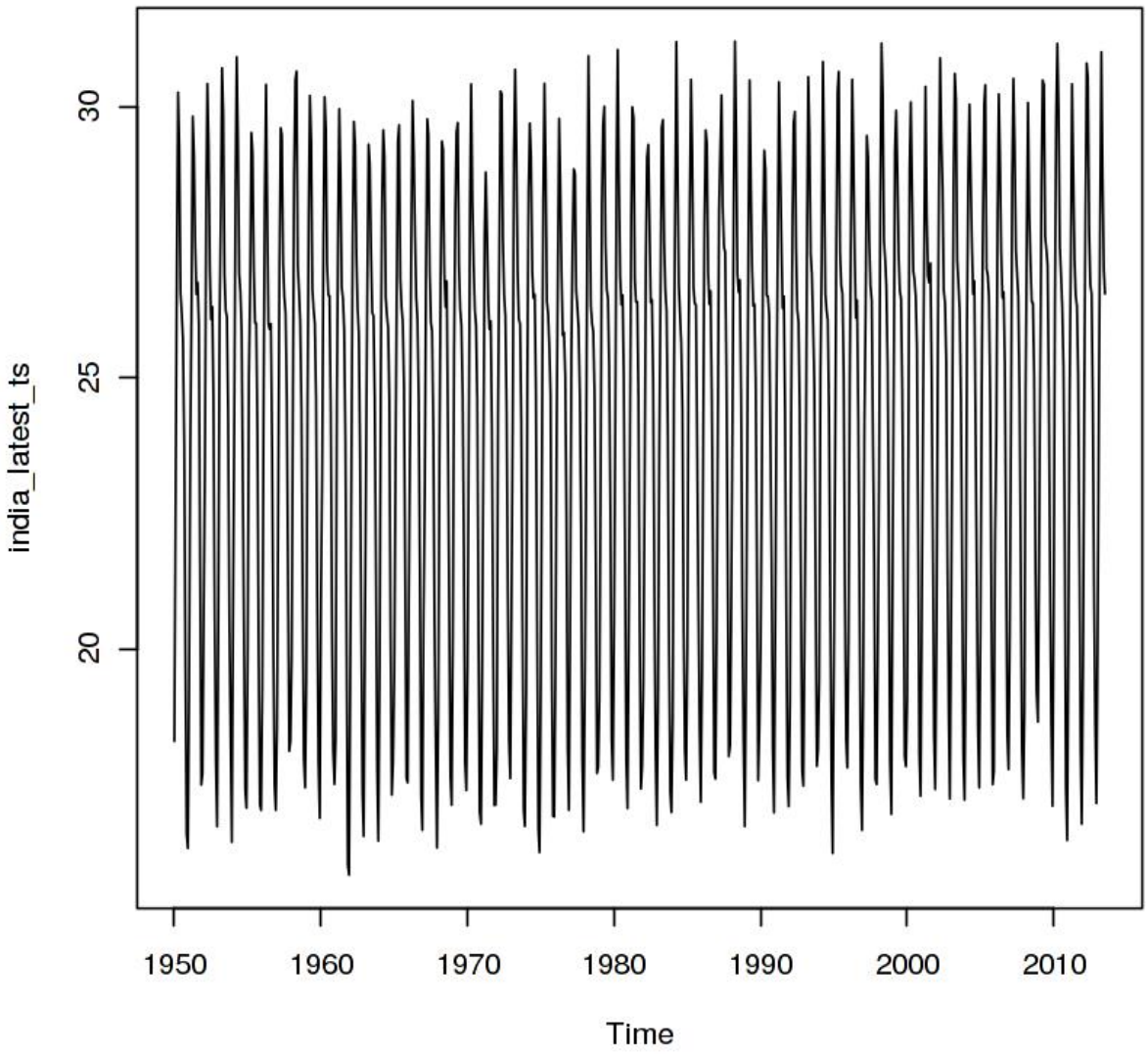
Average Land Temperature in Celsius

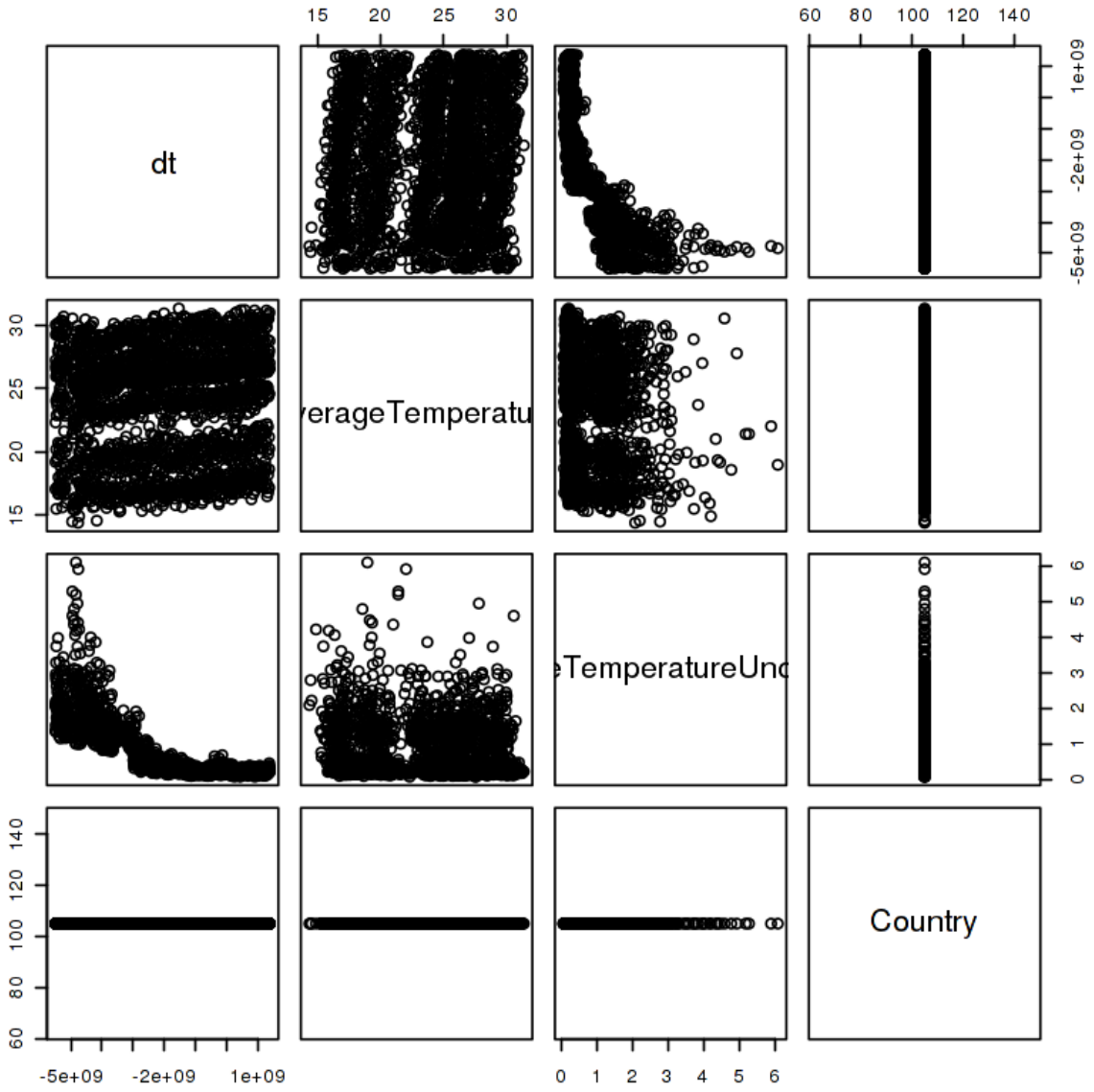


1828

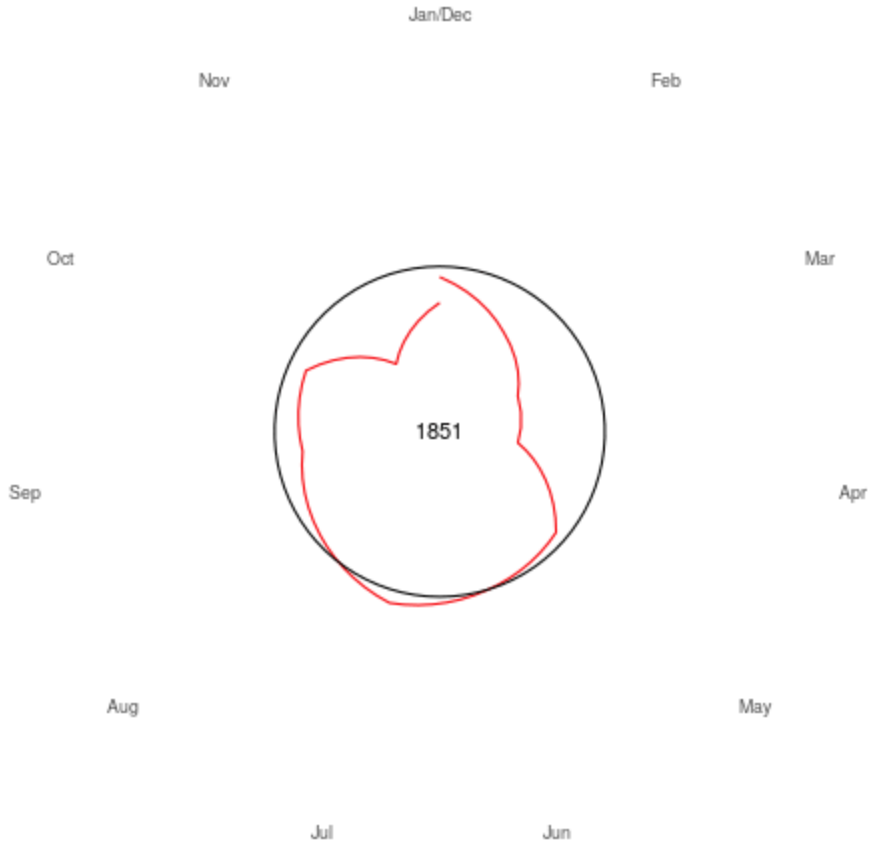
The Grid



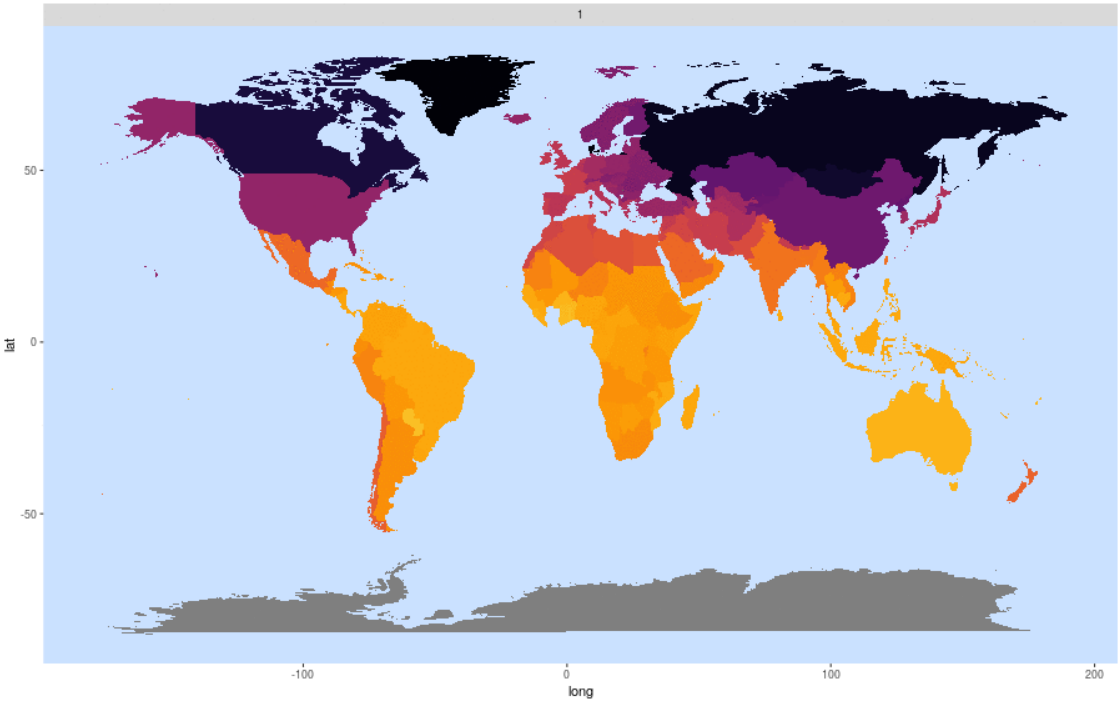

india_latest_ts
30
25
20
1950
1960
1970
1980
1990
2000
2010
Time

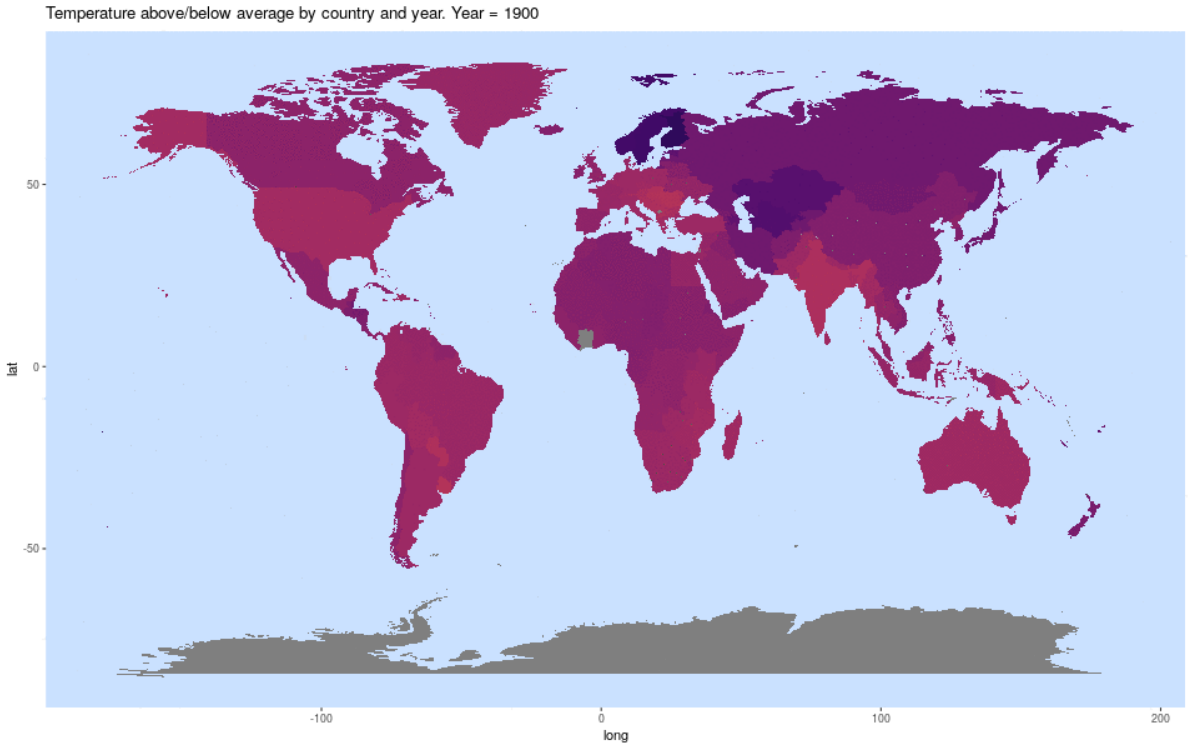

dt
verageTemperatu
eTemperatureUnc
Country
15
20
25
30
60
80
100
120
140
-5e+09
-2e+09
1e+09
0
1
2
3
4
5
6

Variation from monthly average temperature



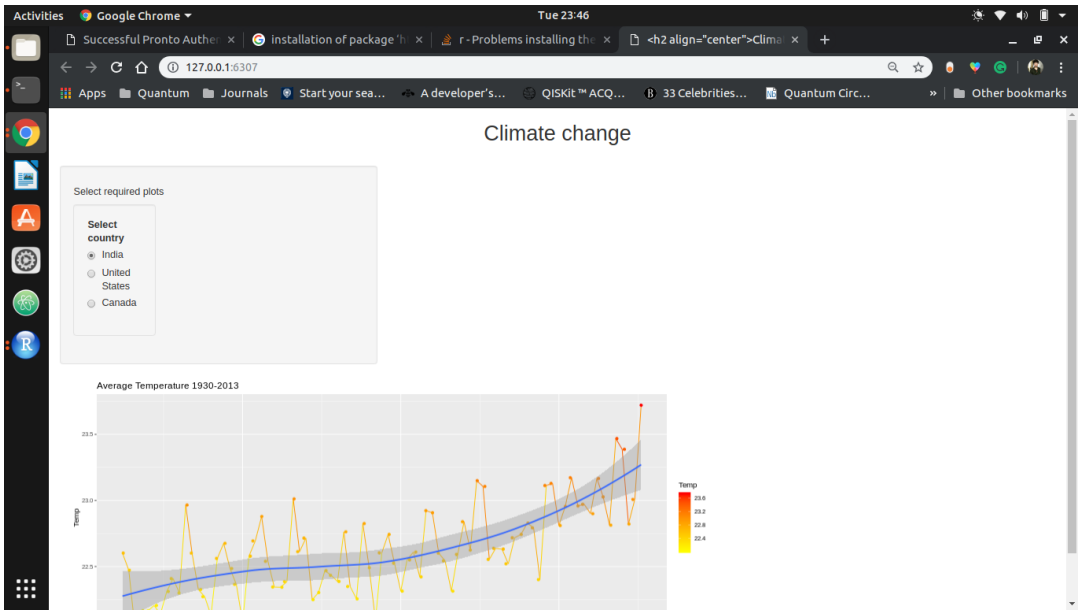
Temperature by country

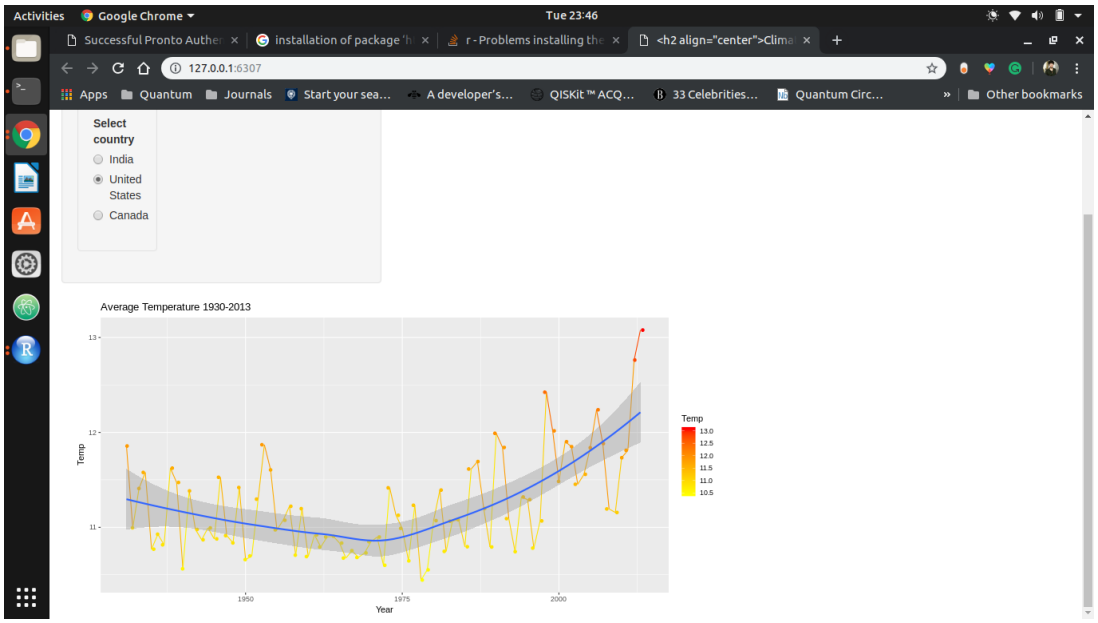


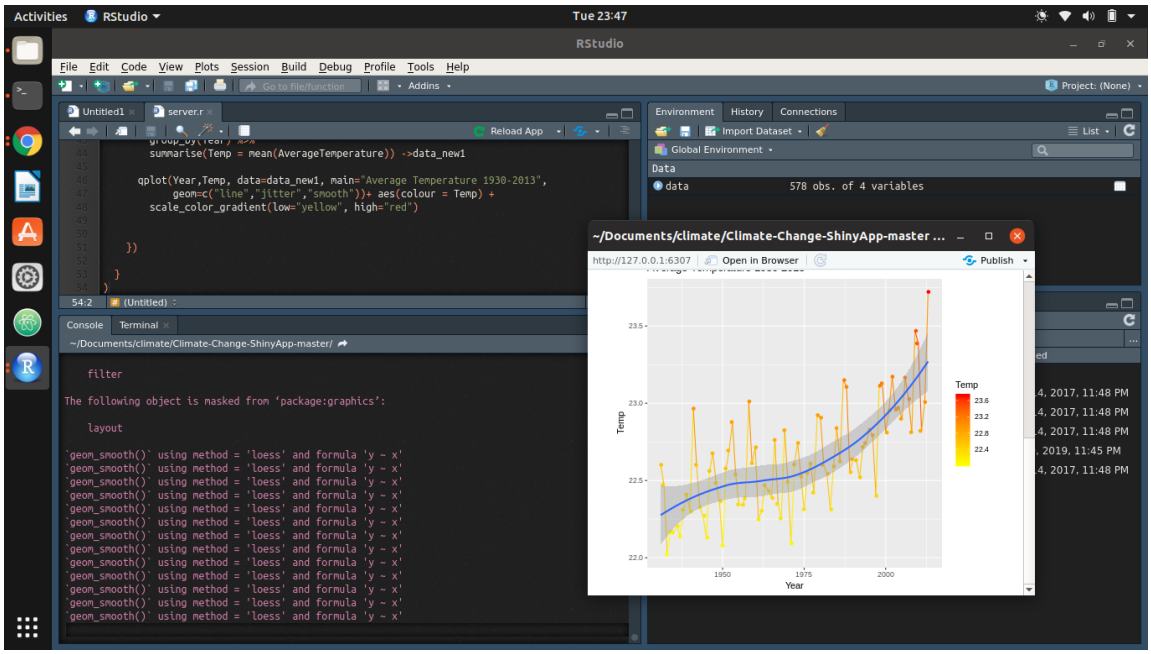



```
library(shiny)
library(plotly)
library(ggplot2)
library(ggthemes)
library(data.table) library(tidyr)
library('tidyr')

data <- read.csv('Data/GlobalLandTemperaturesByState.csv', TRUE,",")
head(data)

row.has.na <- apply(data, 1, function(x){any(is.na(x))})
sum(row.has.na)
data <- data[!row.has.na,]
data <- separate(data,col = dt, into = c("Year", "Month", "Day"), convert = TRUE)
data<- filter(data,Year>1930)

shinyServer(
```

```
function(input,output) {

    output$myplot <- renderPlot( {

  data_new <- filter(data,Country==input$plot_ctry)


  data_new %>%
    group_by(Year) %>%
    summarise(Temp = mean(AverageTemperature)) -> data_new1


 data_new <- filter(data,Country==input$plot_ctry)

 data_new %>%
  filter(Year>1930) %>%
  group_by(Year) %>%
  summarise(Temp = mean(AverageTemperature)) ->data_new1

  qplot(Year,Temp, data=data_new1, main="Average Temperature 1930-2013",
        geom=c("line","jitter","smooth"))+ aes(colour = Temp) +
    scale_color_gradient(low="yellow", high="red")


 })

}
)
```


Activities
Google Chrome
Tue 23:46
127.0.0.1:6307
Apps
Quantum
Journals
Other bookmarks
Climate change
Select required plots
Select country
India
United States
Canada
Average Temperature 1930-2013
Temp


Activities
Google Chrome
Tue 23:46
127.0.0.1:6307
Apps
Quantum
Journals
Other bookmarks
Select country
India
United States
Canada
Average Temperature 1930-2013
Temp
Year


Activities
RStudio
Tue 23:47
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Untitled1
server.r
summarise(Temp = mean(AverageTemperature)) ->data_new1
qplot(Year,Temp, data=data_new1, main="Average Temperature 1930-2013",
geom=c("line","jitter","smooth"))+ aes(colour = Temp) +
scale_color_gradient(low="yellow", high="red")
Reload App
Environment
History
Connections
Import Dataset
Global Environment
Data
data
578 obs. of 4 variables
Console
Terminal
~/Documents/climate/Climate-Change-ShinyApp-master/
filter
The following object is masked from 'package:graphics':
layout
`geom_smooth()` using method = 'loess' and formula 'y ~ x'
~/Documents/climate/Climate-Change-ShinyApp-master ...
http://127.0.0.1:6307
Open in Browser
Publish
Temp
Year
23.5
23.0
22.5
22.0
1950
1975
2000
23.6
23.2
22.8
22.4