

Worldwide Ocean Temperature Report and EDA

Xiaoqian Xue, Sibo Zhu, Li Liu, Danni Fu

2017/11/6

Region name : Subcontinent West

To Run the Code

Create a directory called “data” in this project folder

Download the data text files from the URL we will be provided below. Save the data in the “data” directory. In this project, I have already included the data, so no need for further downloading.

Open and run “data_cleaning.Rmd”, all the cleaned data will be stored in an automatically created directory called “Cleaned_Data”

Open and run “Report and EDA.Rmd”, here stores all the codes for EDA analysis and our discussions about the data.

Summary and EDA of reported data.

The Condition of the Data and Interesting Detail

Without considering the location, lowest temperature is usually between December and January, while the highest temperatures around August. According to the geography, subcontinent west is close to equator so the temperature difference through the year should not be noticeable. Since water has bigger specific heat capacity, from the equation $\Delta T = Q/mc$ (Δ =temperature change, Q =heat, m =mass, c =specific heat capacity), we can predict that the sea temperature change has a smaller range than that of air temperature, which is consistence with our plots. Also, in our plots, there???s no pattern showing increase or decrease of temperature from 2001 to 2016.

The Data Acquisition and Density of Data.

For our analysis purpose, ship data is used instead of buoy data, since buoy data only contains the measurement of water depth. Ship data is dependent on the fact that ships needed to be passing through our assigned area. Data from VOS website is divided per month from 2001 to 2016. we make sure the range is in sub-continent west by limit the range of latitude from 6 to 20 and longitude from 60 to 80. In order to make data fit into preferred region while remaining consistent through years, we measure our data within a 6-hour range of noon. We dropped all the data with “NA” value in order to keep our data logically and easy for analysis. By dropping extreme measurement that out of range between (0.01,0.99), we kept the consistency of our data.

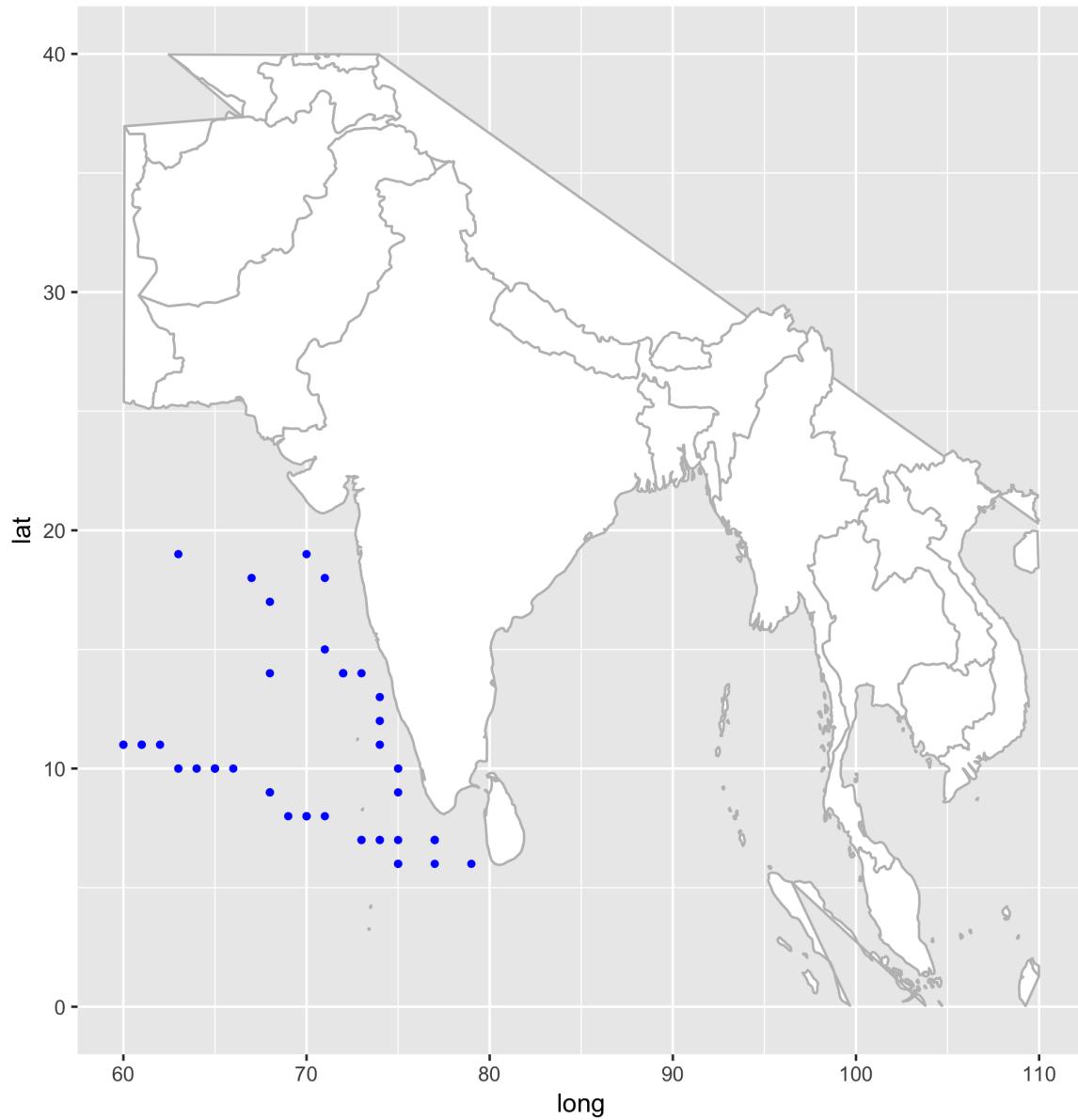
List of sources with URLs

2001 data: <https://www1.ncdc.noaa.gov/pub/data/vosclim/2001/>
2002 data: <https://www1.ncdc.noaa.gov/pub/data/vosclim/2002/>
2003 data : <https://www1.ncdc.noaa.gov/pub/data/vosclim/2003/>
2004 data: <https://www1.ncdc.noaa.gov/pub/data/vosclim/2004/>
2005 data: <https://www1.ncdc.noaa.gov/pub/data/vosclim/2005/>
2006 data : <https://www1.ncdc.noaa.gov/pub/data/vosclim/2006/>
2007 data: <https://www1.ncdc.noaa.gov/pub/data/vosclim/2007/>
2008 data: <https://www1.ncdc.noaa.gov/pub/data/vosclim/2008/>
2009 data : <https://www1.ncdc.noaa.gov/pub/data/vosclim/2009/>
2010 data: <https://www1.ncdc.noaa.gov/pub/data/vosclim/2010/>
2011 data: <https://www1.ncdc.noaa.gov/pub/data/vosclim/2011/>
2012 data : <https://www1.ncdc.noaa.gov/pub/data/vosclim/2012/>
2013 data: <https://www1.ncdc.noaa.gov/pub/data/vosclim/2013/>
2014 data: <https://www1.ncdc.noaa.gov/pub/data/vosclim/2014/>
2015 data : <https://www1.ncdc.noaa.gov/pub/data/vosclim/2015/>
2016 data: <https://www1.ncdc.noaa.gov/pub/data/vosclim/2016/>

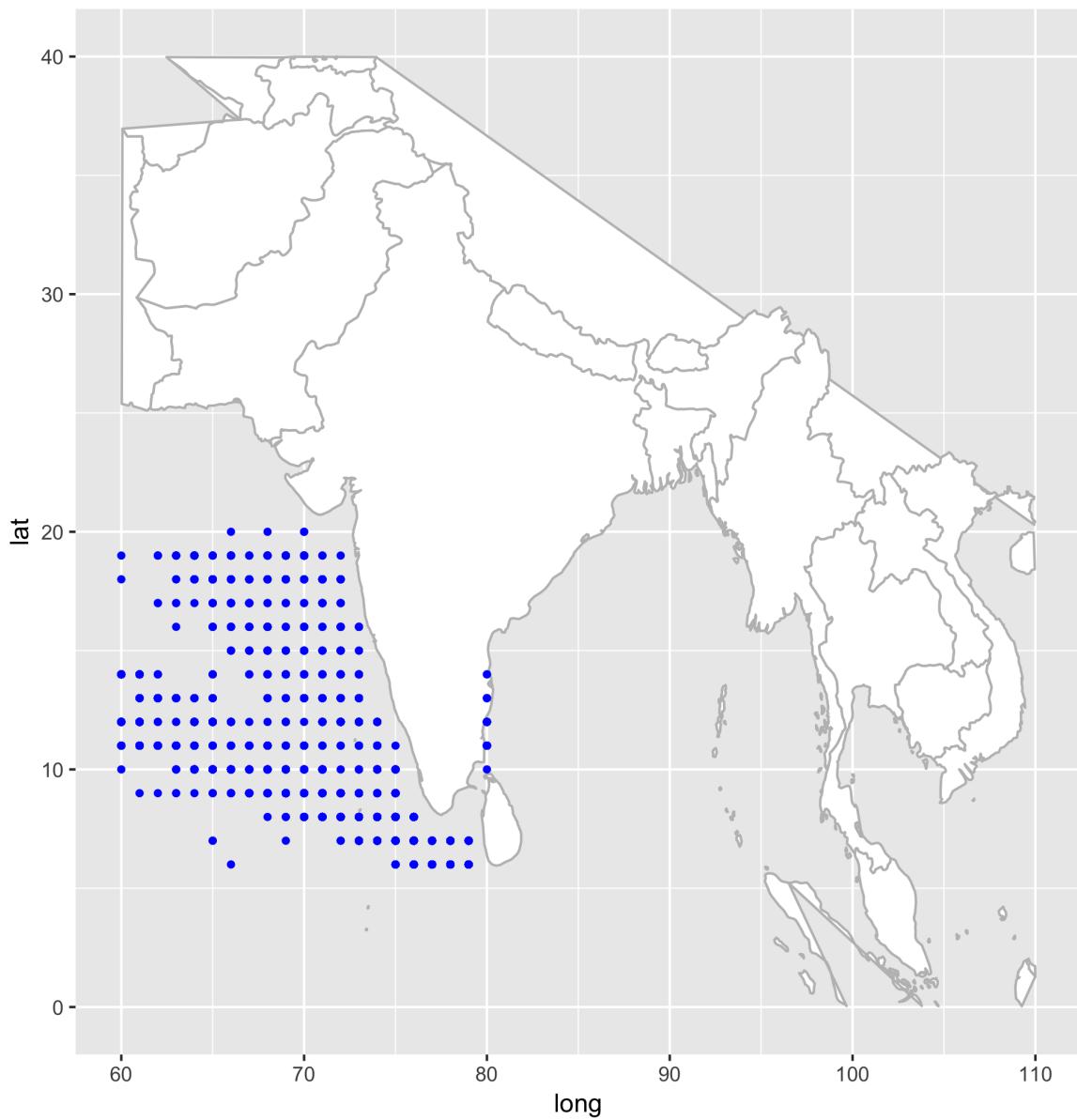
Maps

We created maps per year which show the data of Sub-Continent West region from 2001 to 2016. As we can see, when times goes on, records showing that there are more and more ship lanes are travel through this area and more data is collected in Sub-Continent West.

Subcontinent West 2001



Subcontinent West 2016

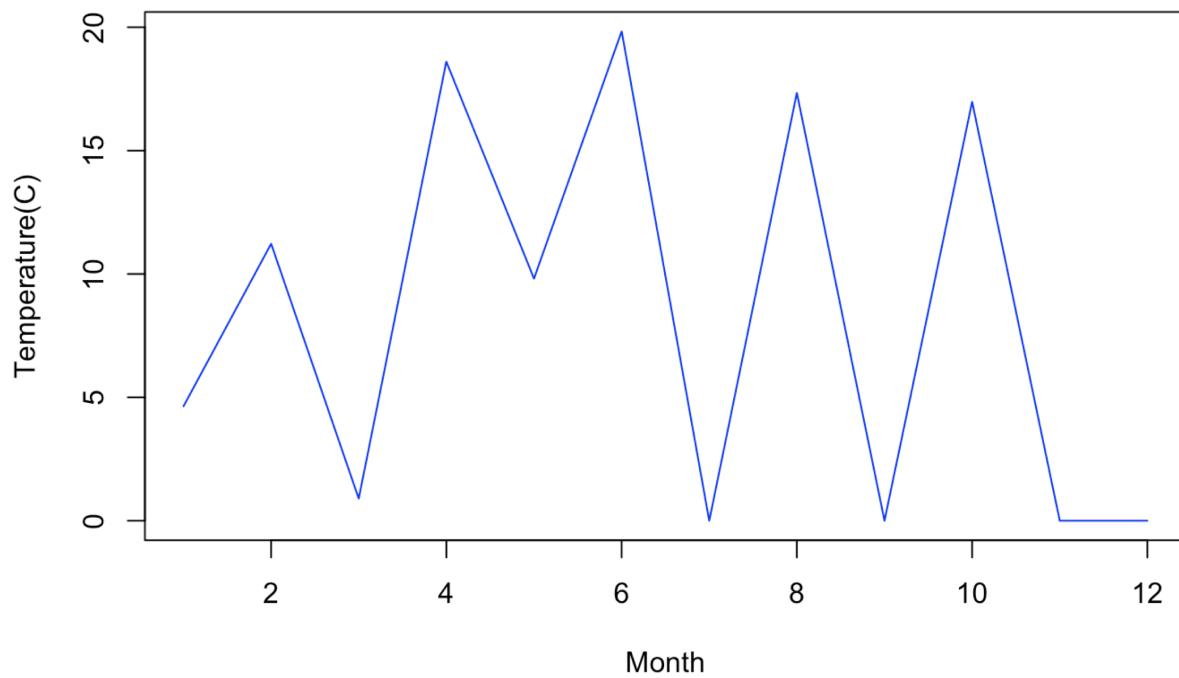


Discussion

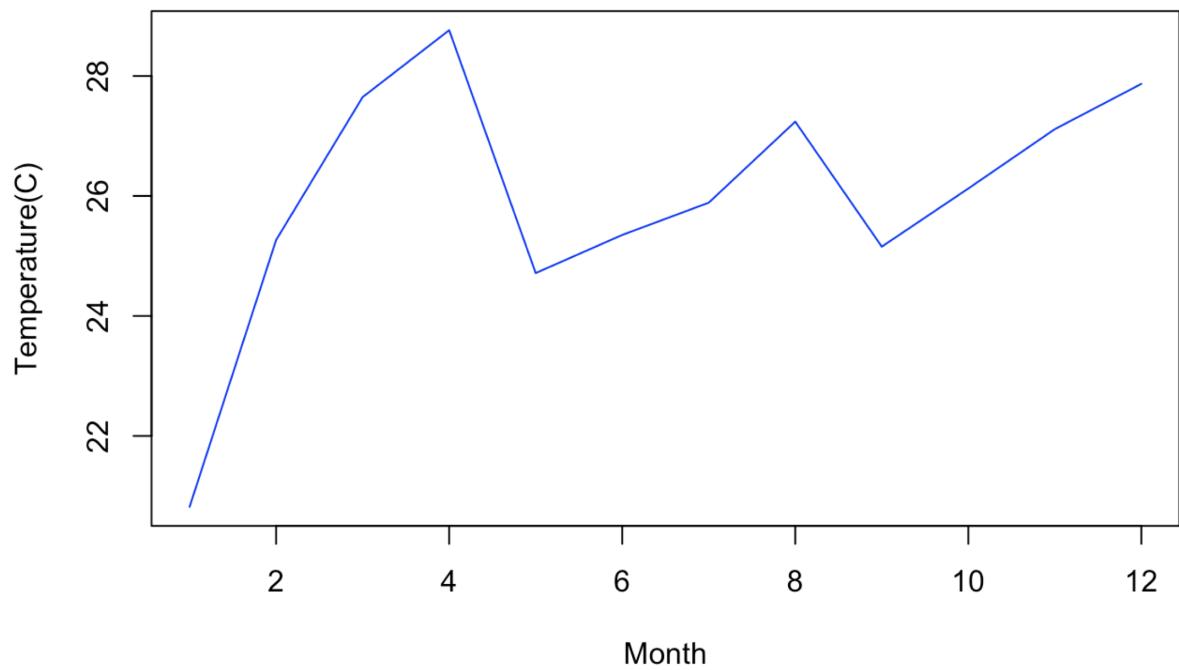
In subcontinent west, the average sea temperature has raised significantly from 2001 (around 10 degrees) to 2016 (around 27 degrees). The biggest change occurs during the period of 2001-2003 and after 2003 the average sea temperatures are mainly around 25 degrees.

Therefore, we expect that there could be some significant climate change or human-caused change happened during that period.

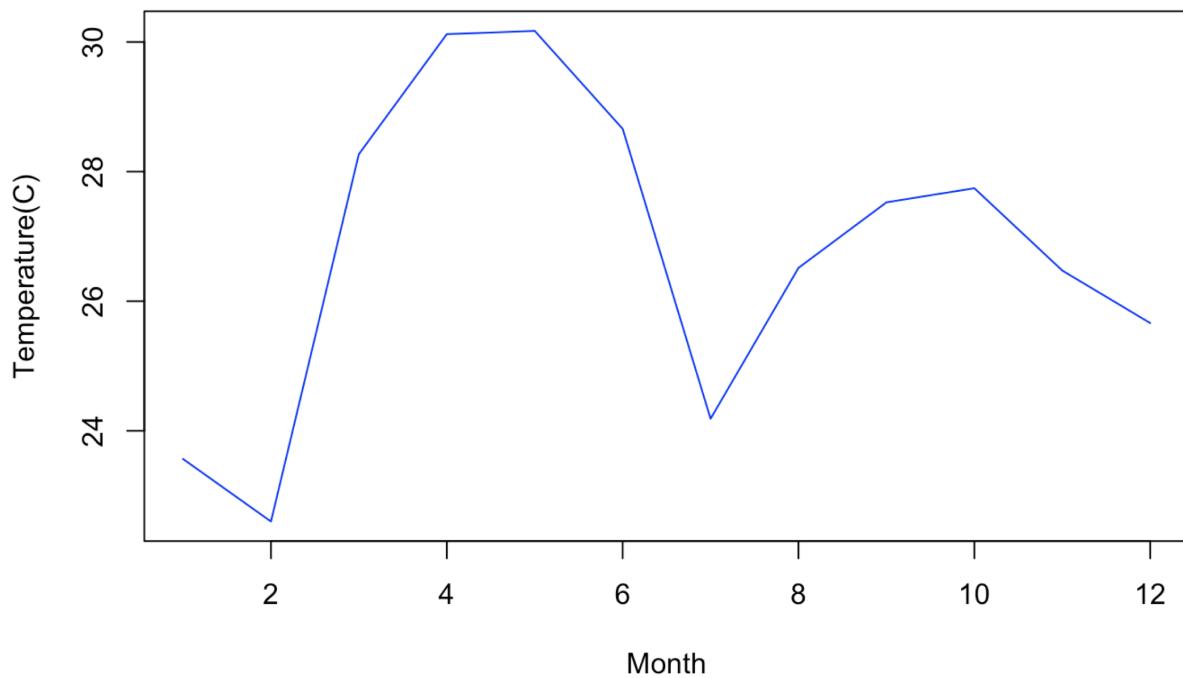
Average Sea Temperature of 2001



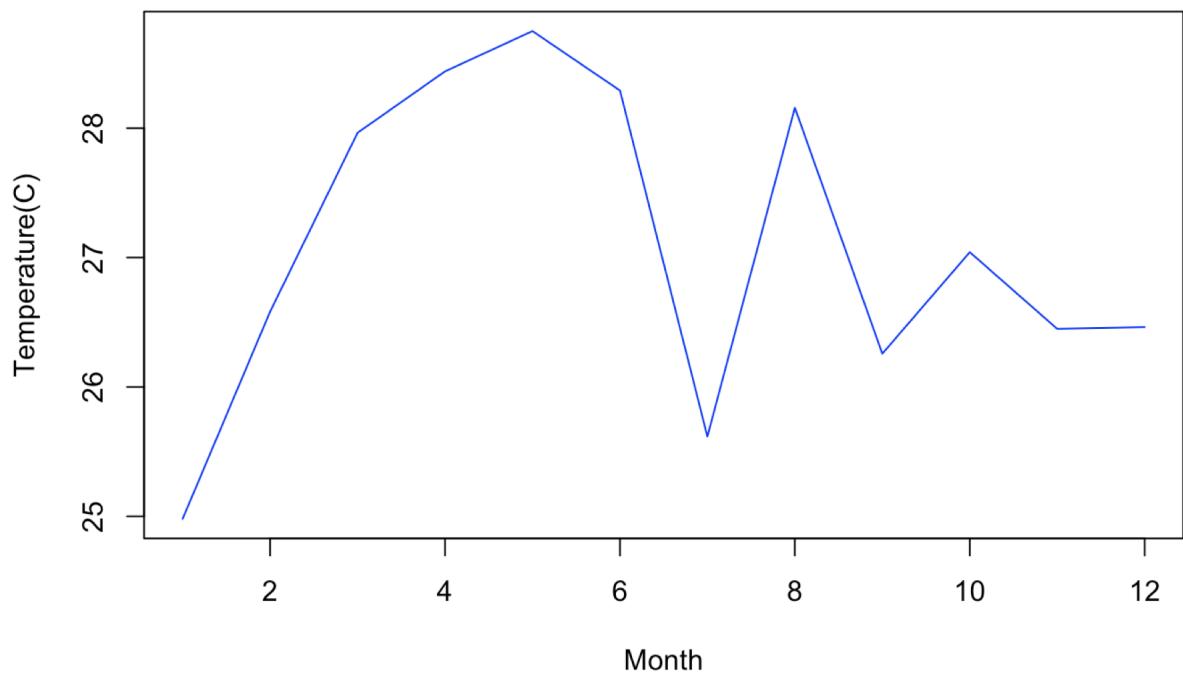
Average Sea Temperature of 2003



Average Sea Temperature of 2010



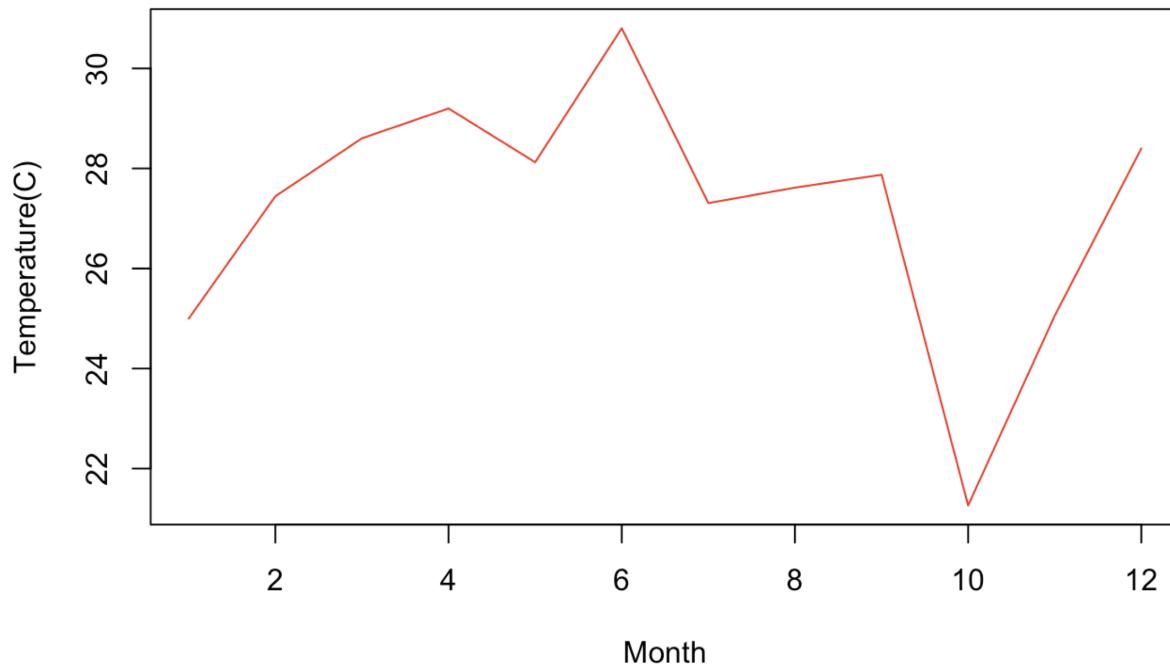
Average Sea Temperature of 2016



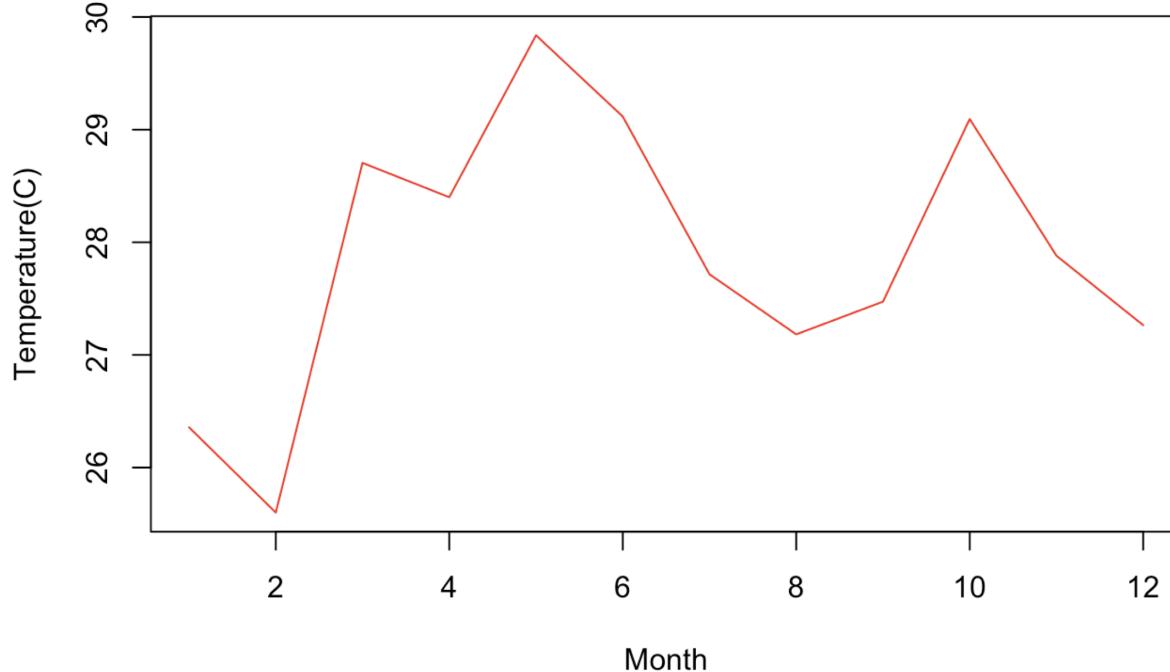
There is not so much difference between average air temperatures from 2001 to 2016 (mainly around 28 degrees). Clearly, the difference between average sea and air temperature has decreased a lot from 2001, 18 degrees' difference to 2016, approximately no difference.

Therefore, further investigation can be done to find out the main reason for the increase in sea temperature.

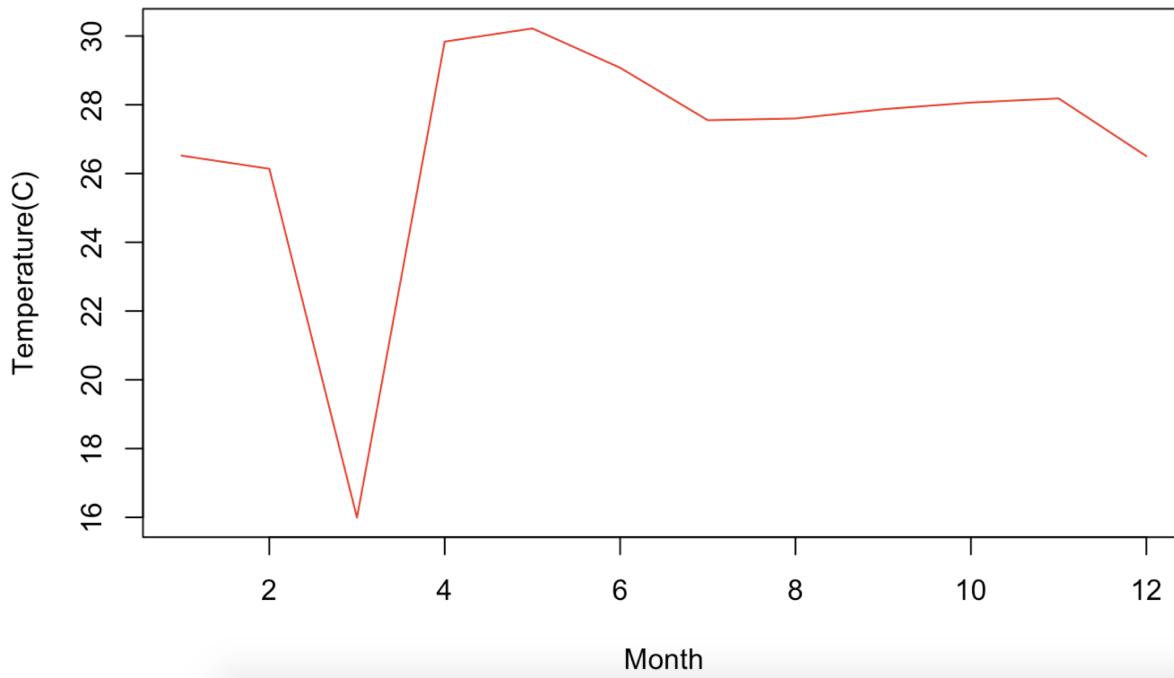
Average Air Temperture of 2001



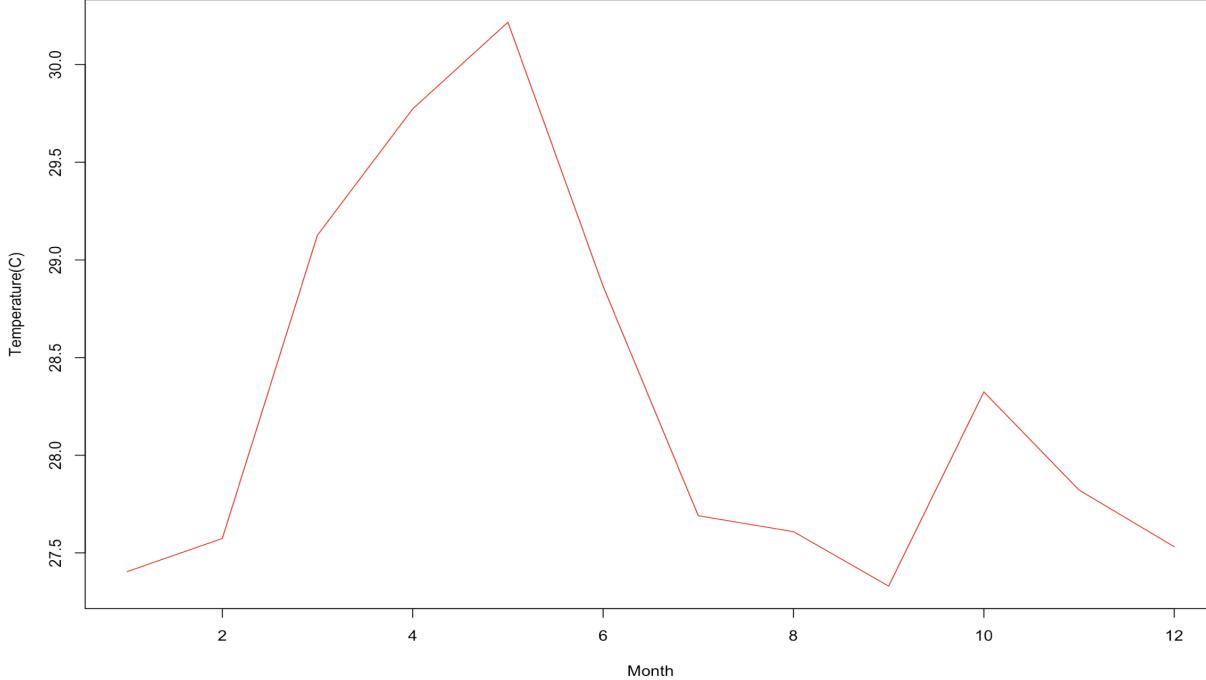
Average Air Temperture of 2002



Average Air Temperture of 2010

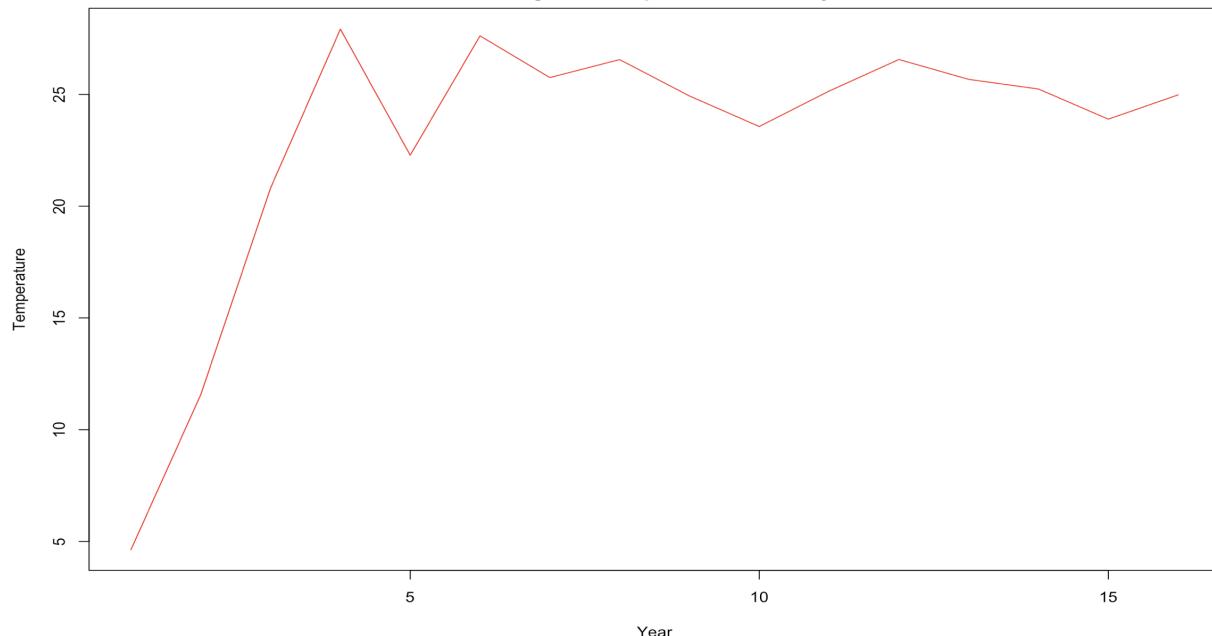


Average Air Temperture of 2016

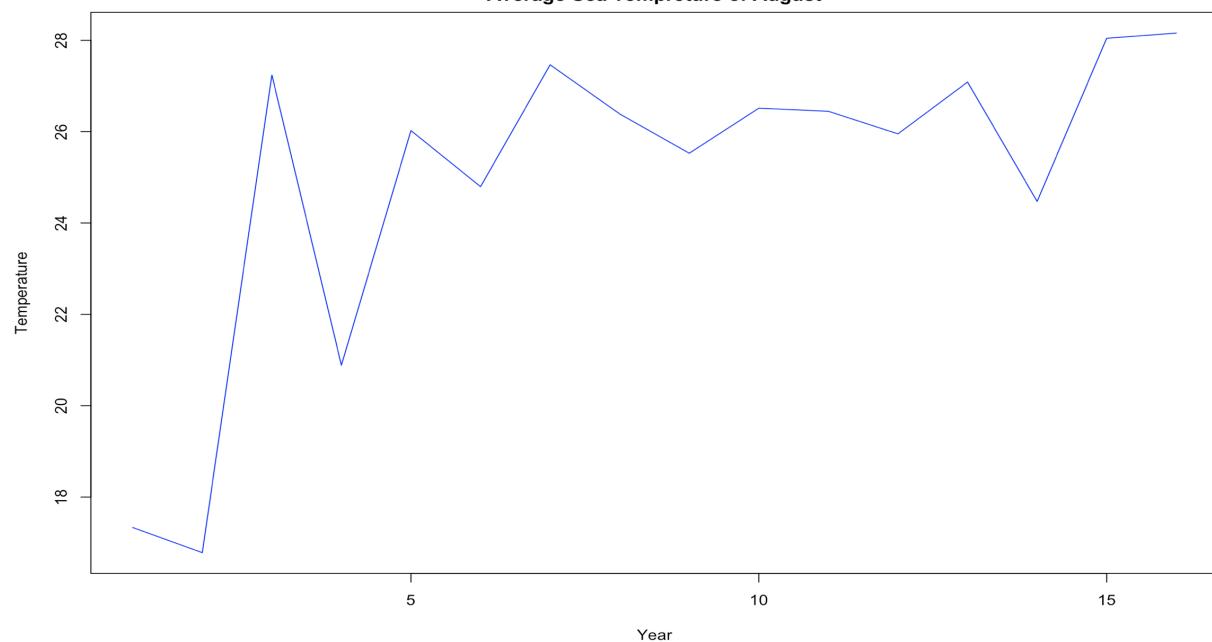


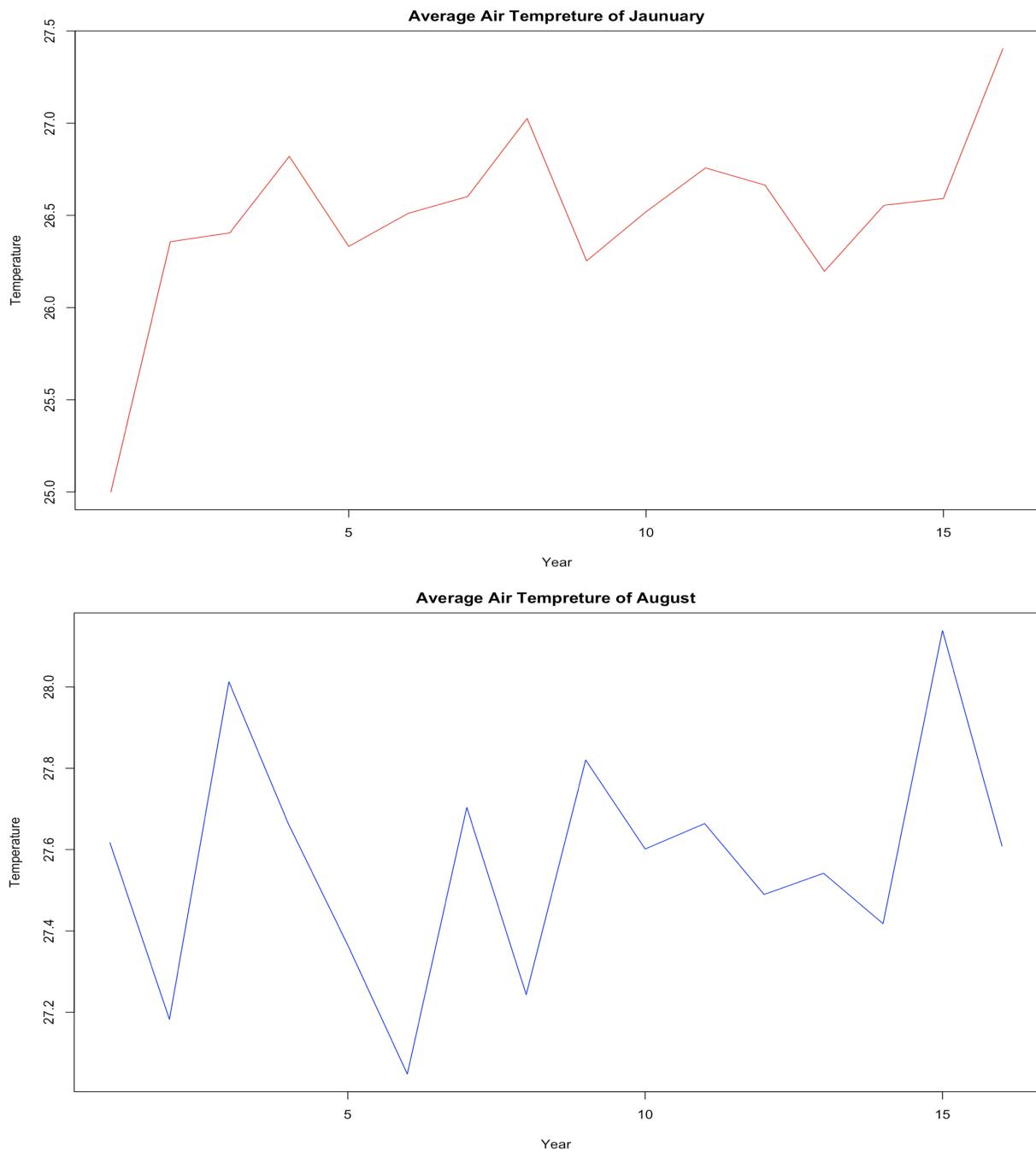
To further confirm our previous findings, we plot the average air and sea temperature for the same month (January and August) across 16 years. The overall trend of average sea temperature is the same across those 16 years with a sharp increase in the first five years and then remain almost the same in the next ten years. On the other hand, the trends of average air temperature are different for the January and August. The sharp increase in the air temperature occurs in 2002 of January but in 2003, 2006, 2008 and 2014 of August.

Average Sea Temperature of January



Average Sea Temperature of August





EDA Codes

```
matrix.jan.temp = NULL
matrix.aug.temp = NULL

#Graphing
for (i in 2001:2016) { #function used for graphing
  YEAR = i

  load(paste0("./cleaned_data/ave_temp_",YEAR,".Rdata")) #loading data by loop

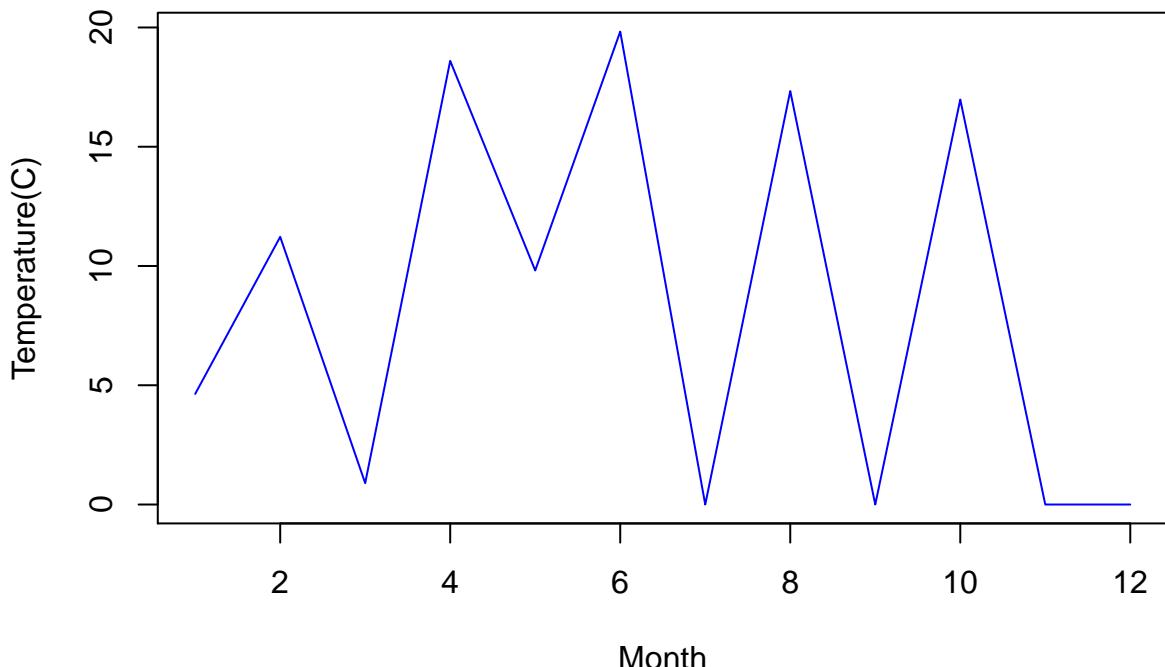
  Temp_Sea <- plot(EDA.year[,2], type = "l", main = paste("Average Sea Temperture of", YEAR, sep = " "))

  Temp_Air <- plot(EDA.year[,3], type = "l", main = paste("Average Air Temperture of", YEAR, sep = " "))

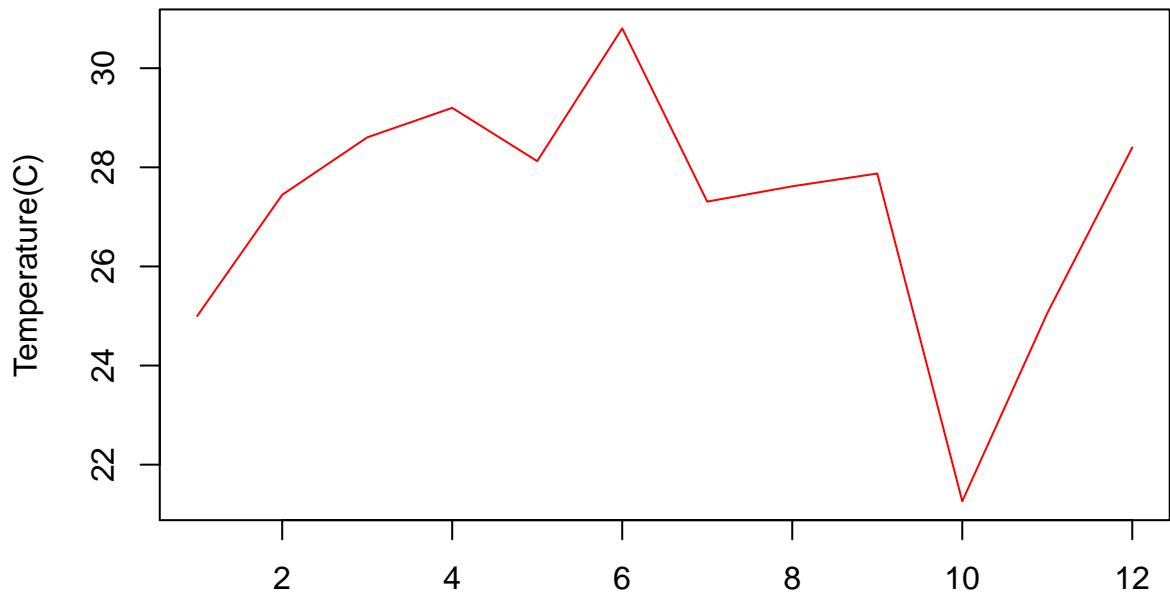
  jan.temp = EDA.year[1,2:3] #take out the record of Januarray of each year
  aug.temp = EDA.year[8,2:3] #take out the record of August of each year

  matrix.aug.temp = rbind(matrix.aug.temp , aug.temp) #build matrix to track average temperature of Aug
  matrix.jan.temp = rbind(matrix.jan.temp , jan.temp) #build matrix to track average temperature of Jan
}
```

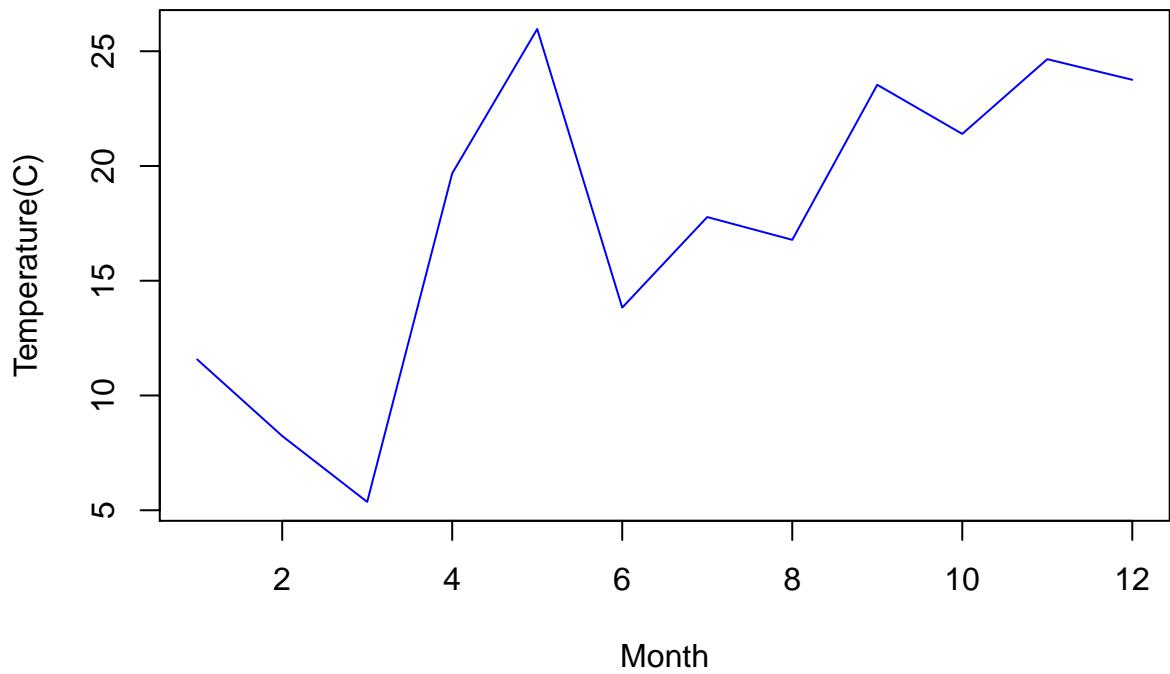
Average Sea Temperture of 2001



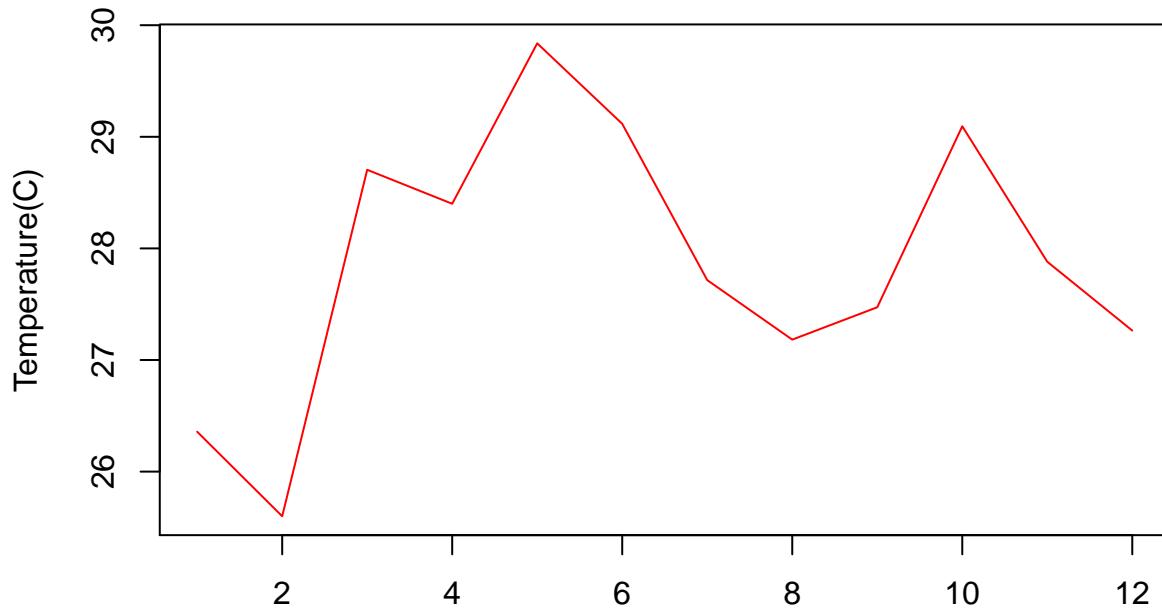
Average Air Temperature of 2001



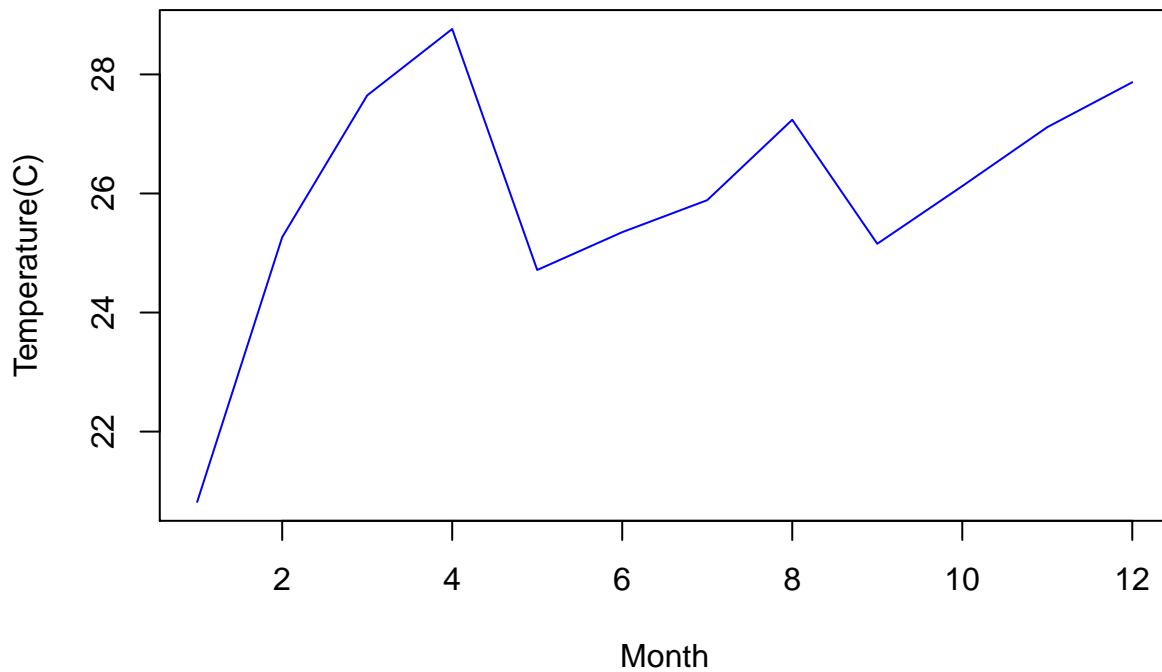
Average Sea Temperature of 2002



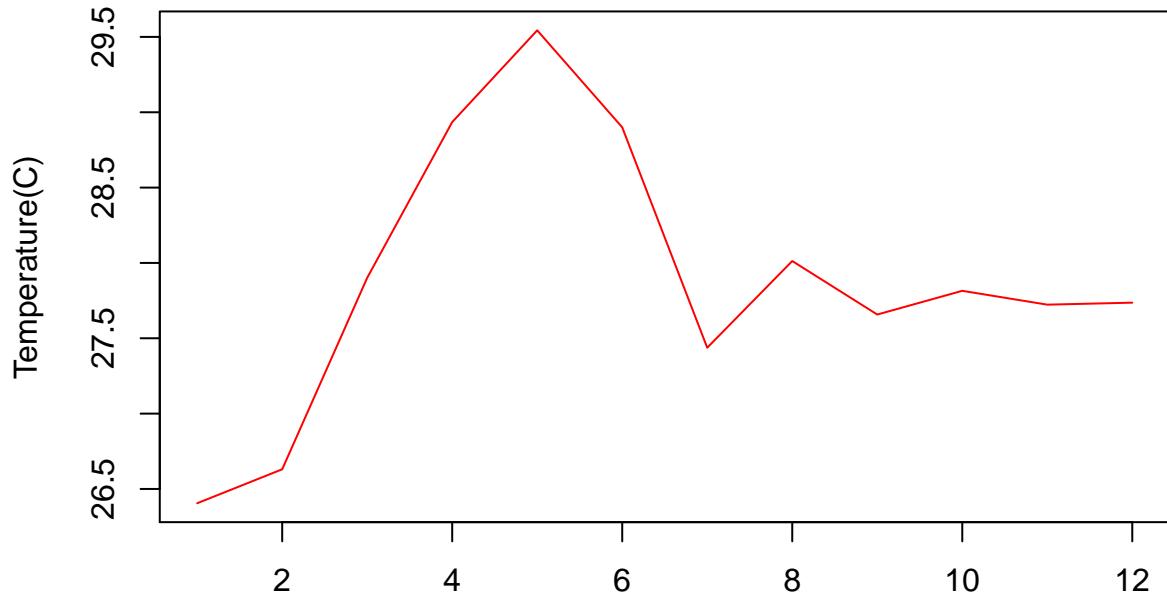
Average Air Temperature of 2002



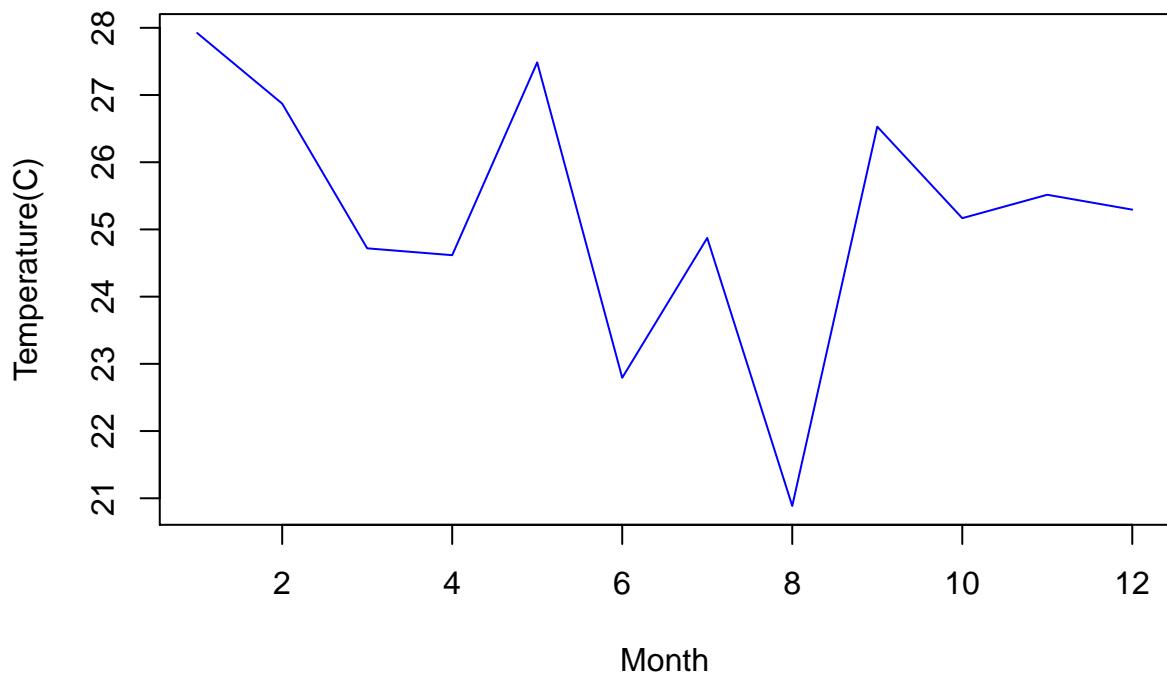
Average Sea Temperature of 2003



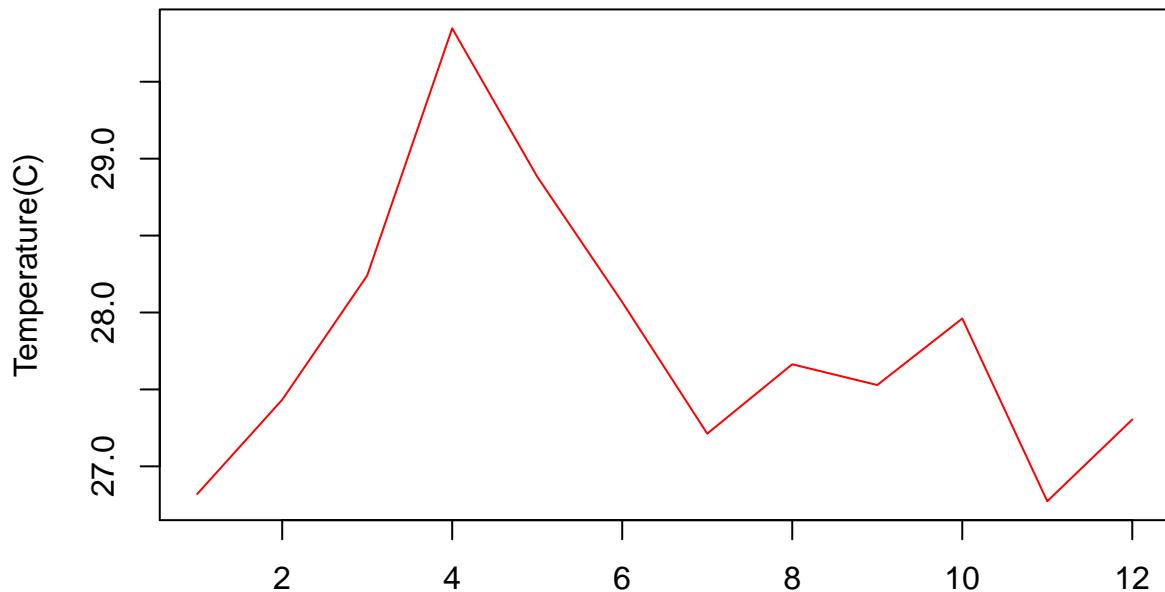
Average Air Temperture of 2003



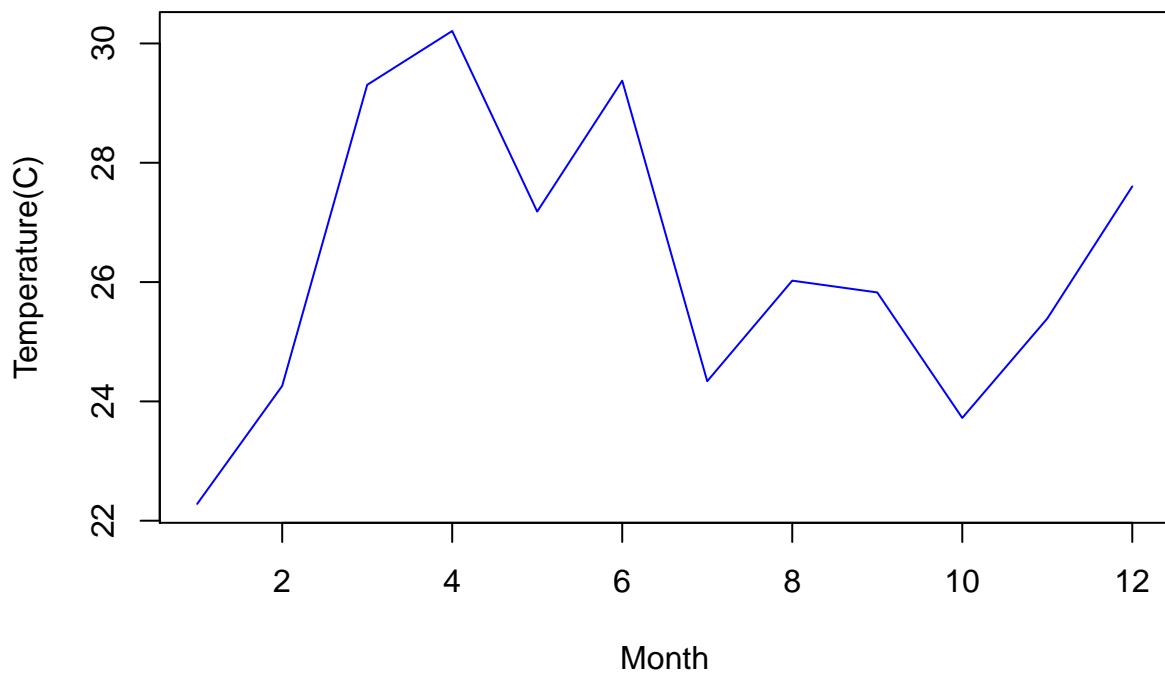
Average Sea Temperture of 2004



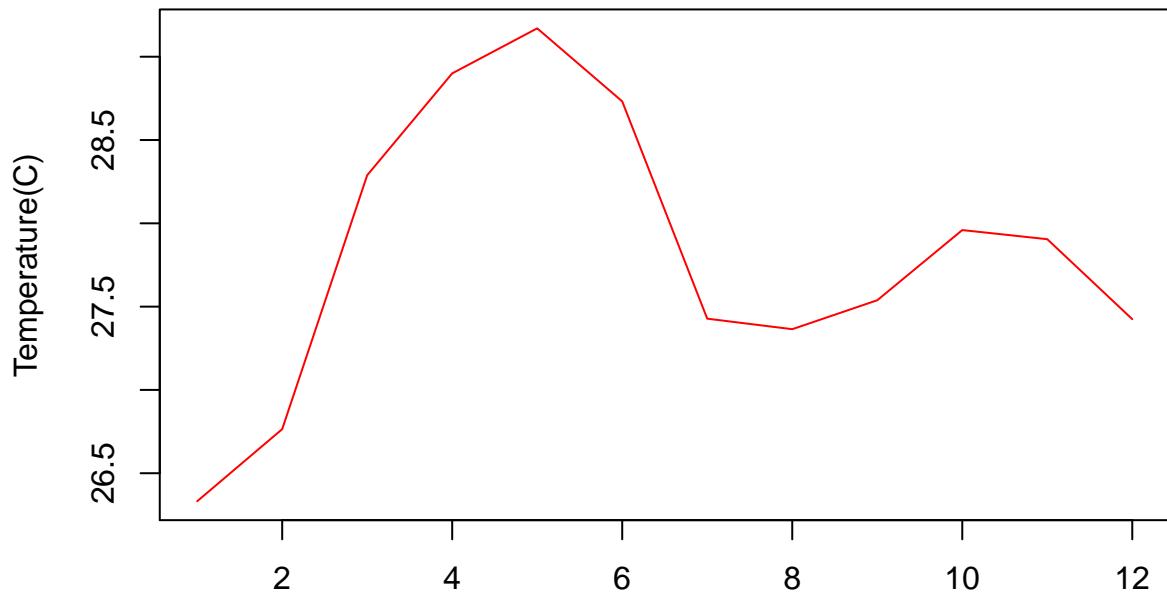
Average Air Temperature of 2004



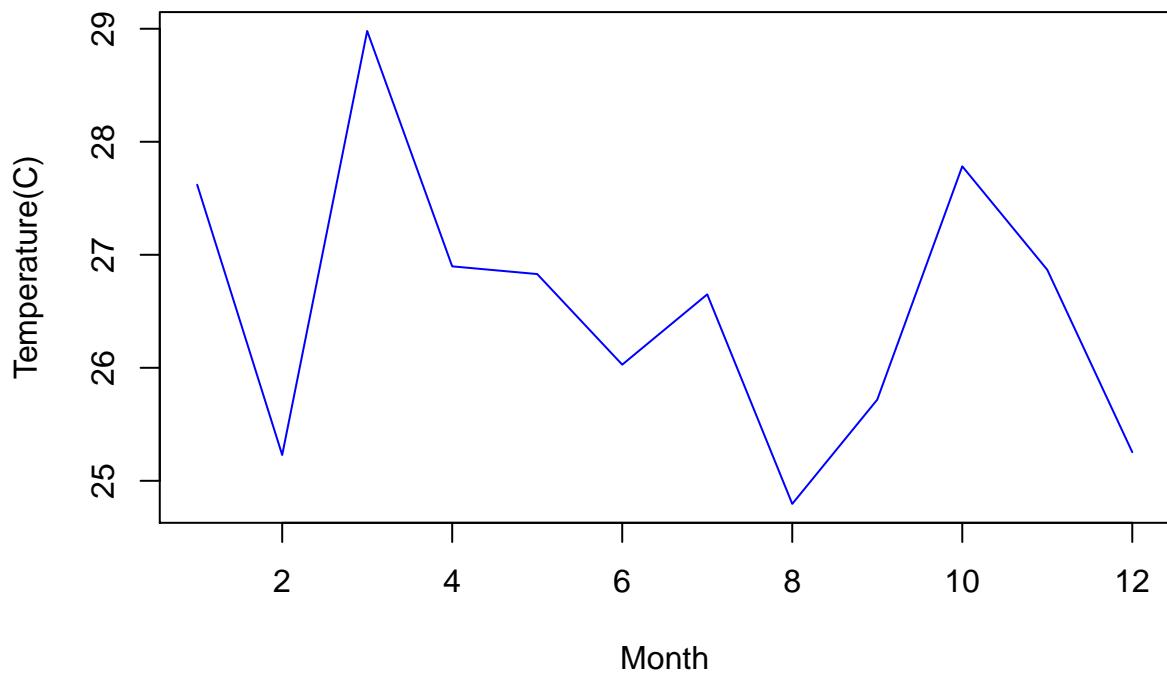
Average Sea Temperature of 2005



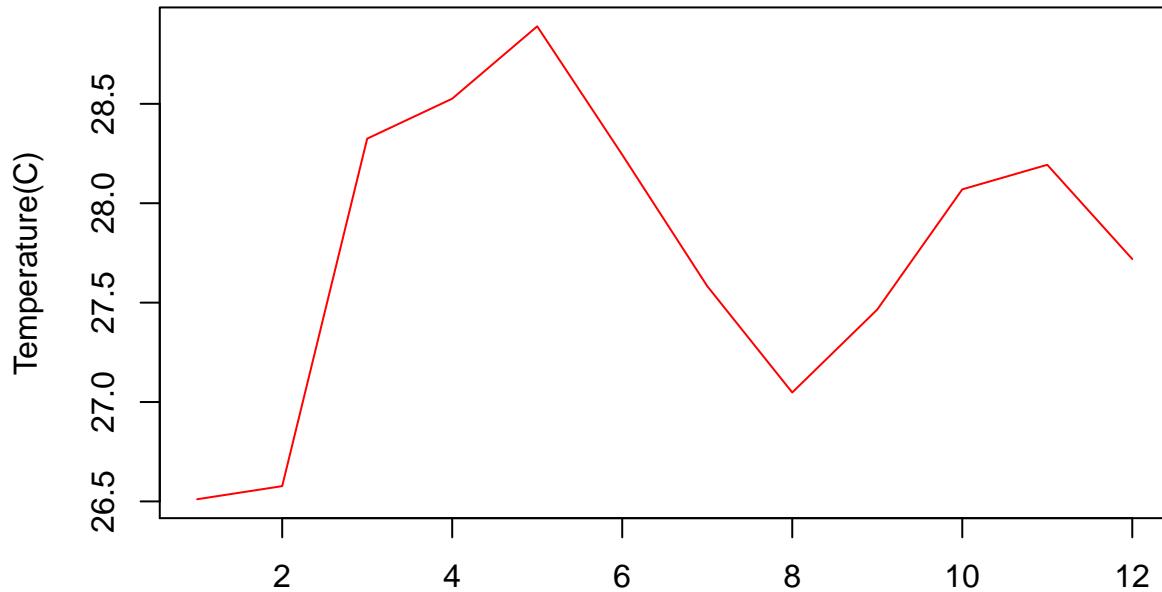
Average Air Temperature of 2005



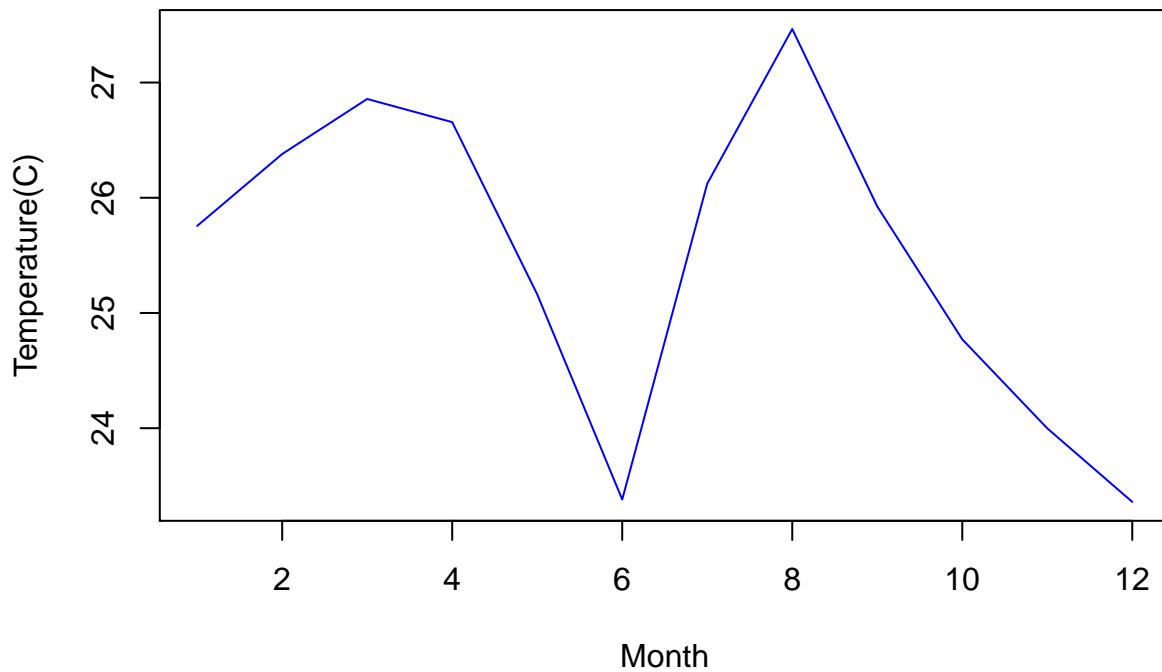
Average Sea Temperature of 2006



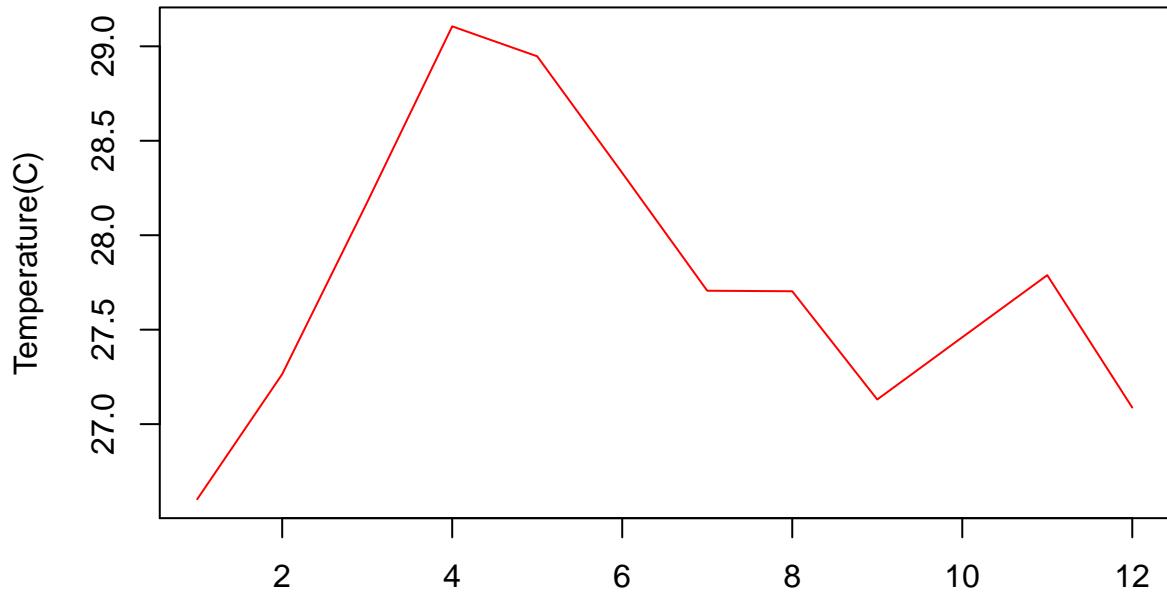
Average Air Temperature of 2006



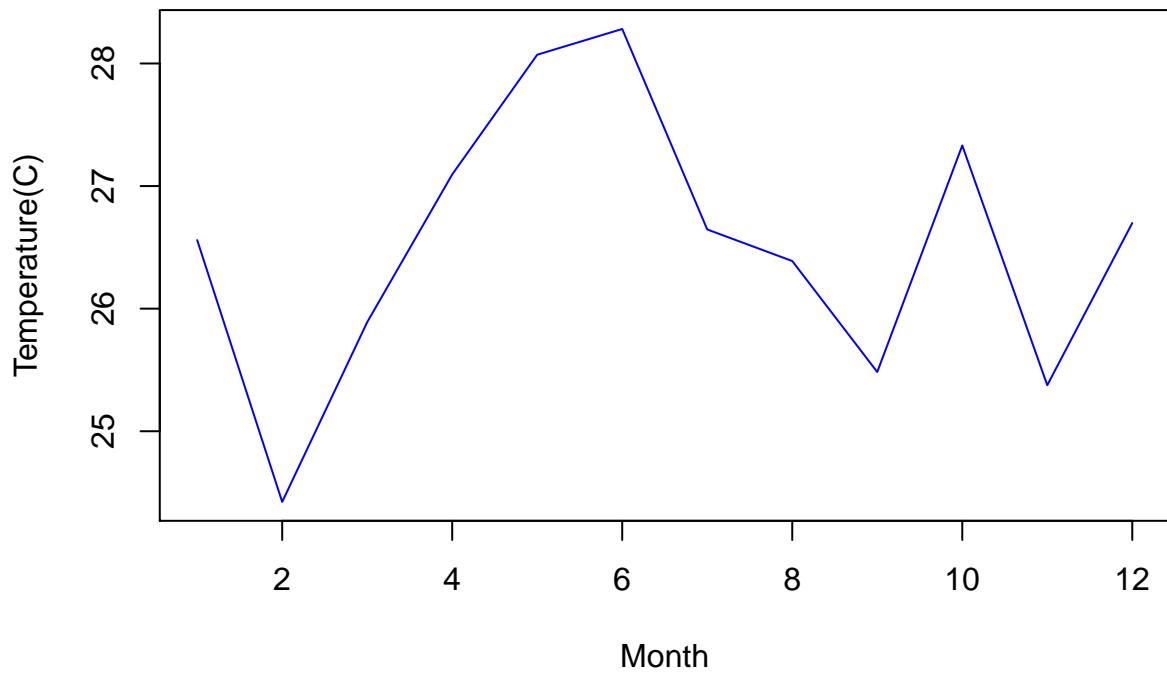
Average Sea Temperature of 2007



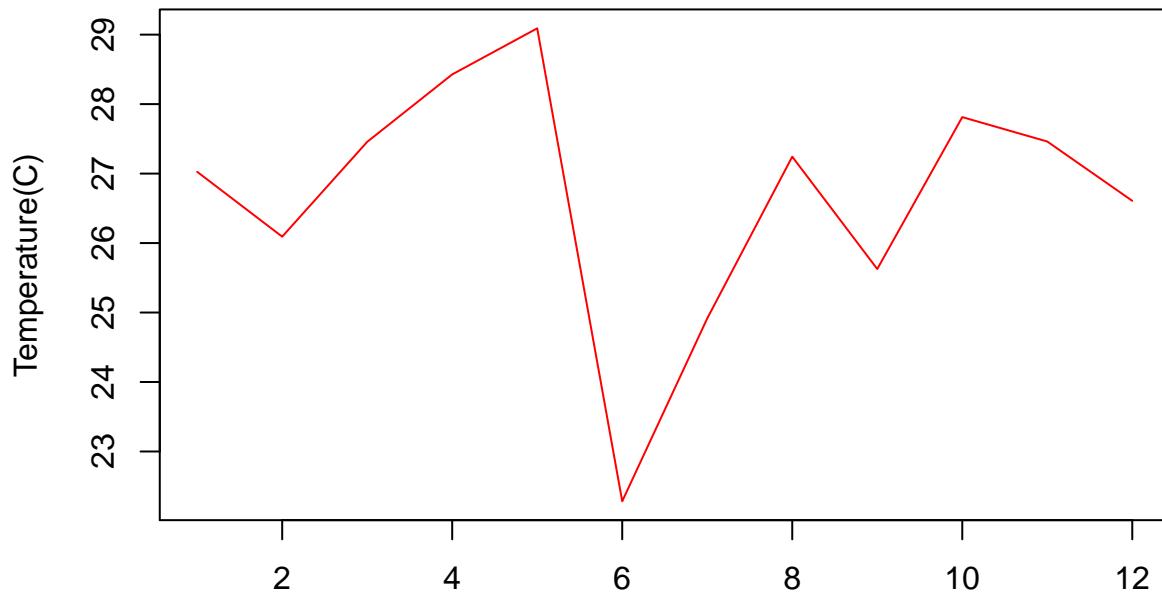
Average Air Temperature of 2007



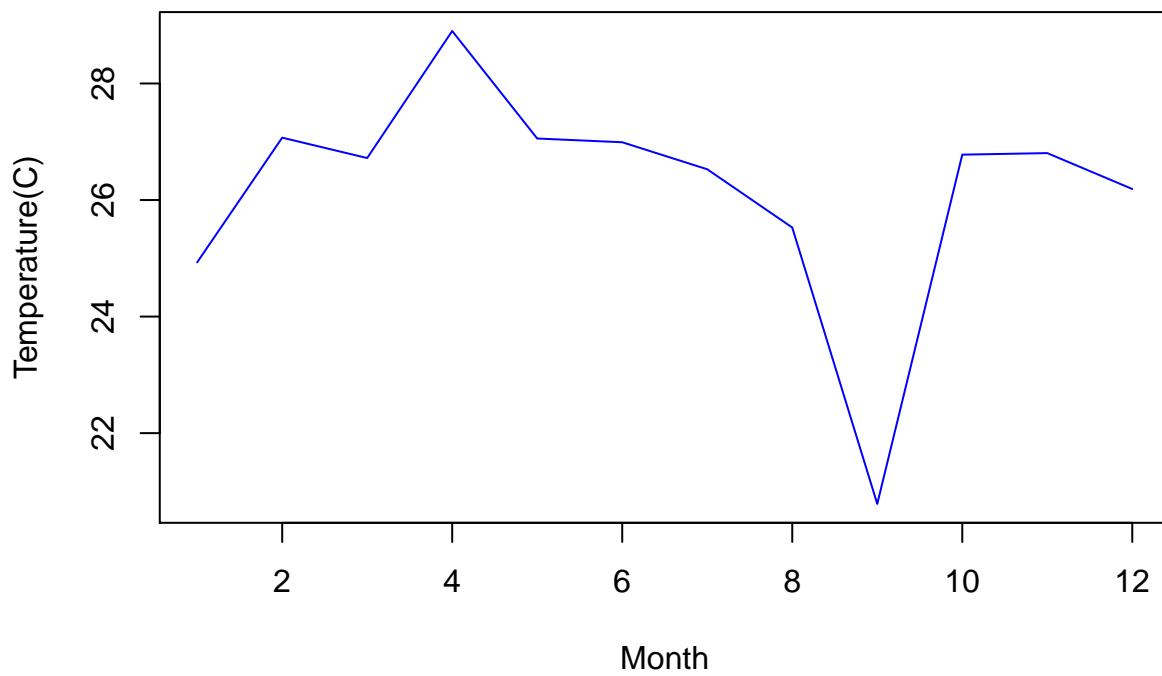
Average Sea Temperature of 2008



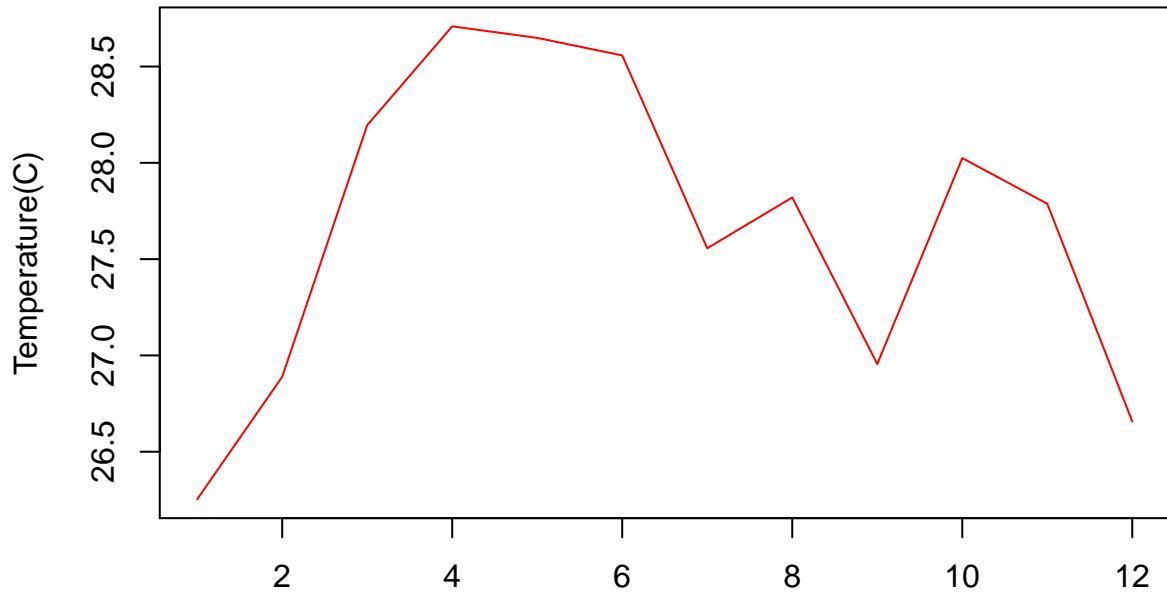
Average Air Temperature of 2008



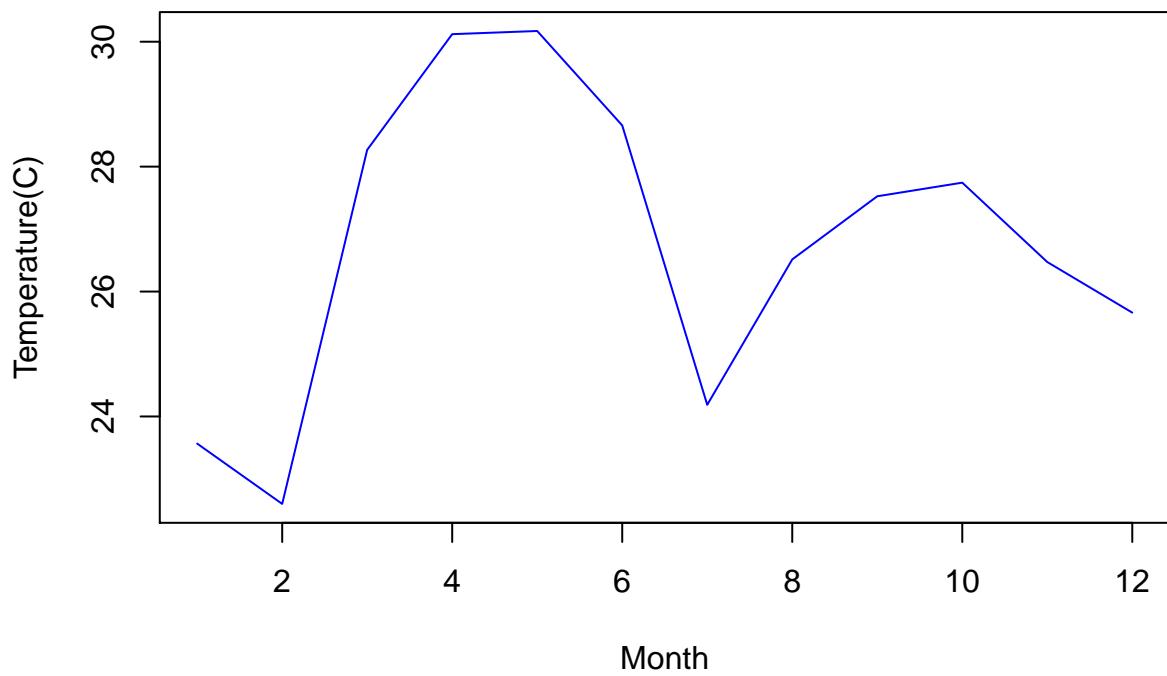
Average Sea Temperature of 2009



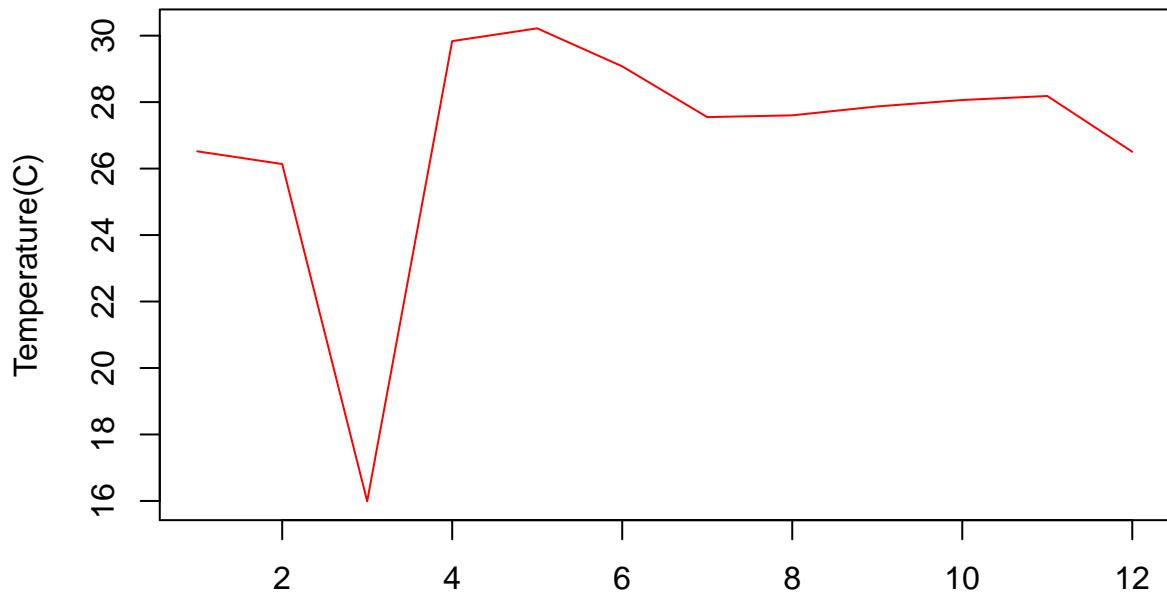
Average Air Temperture of 2009



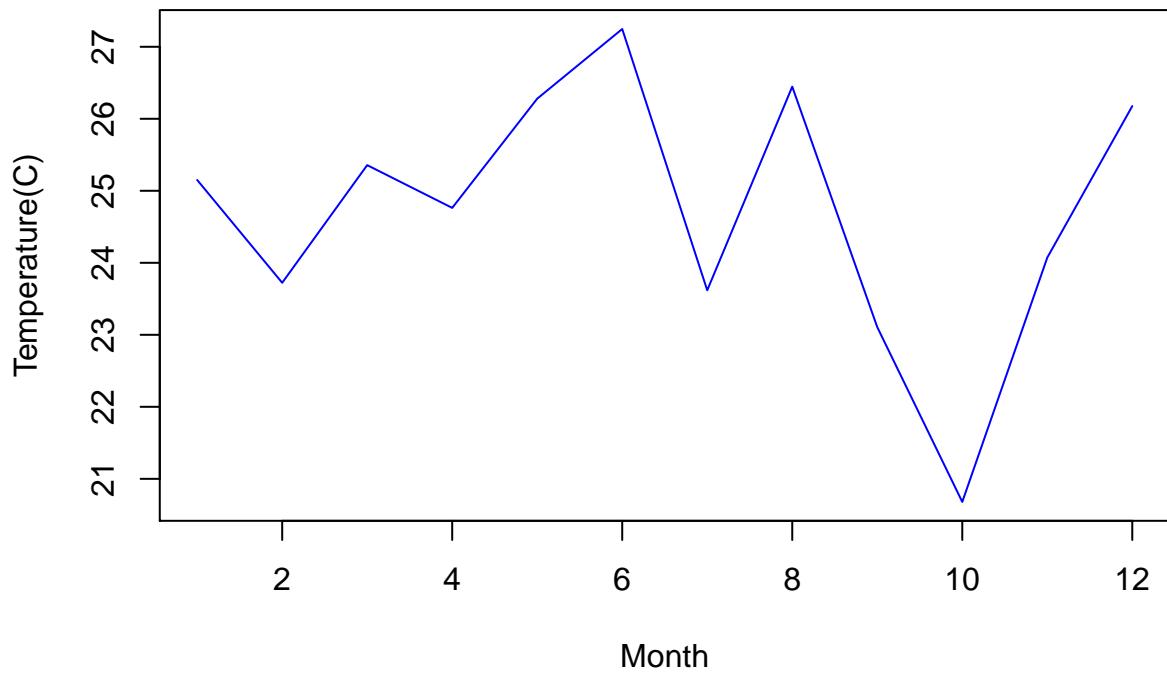
Average Sea Temperture of 2010



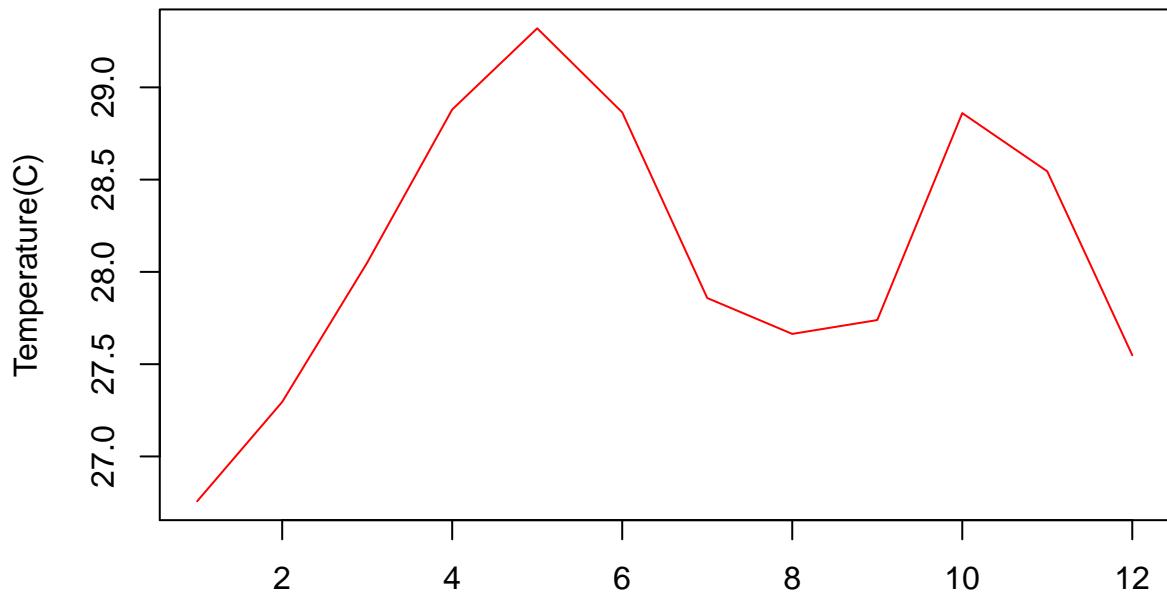
Average Air Temperature of 2010



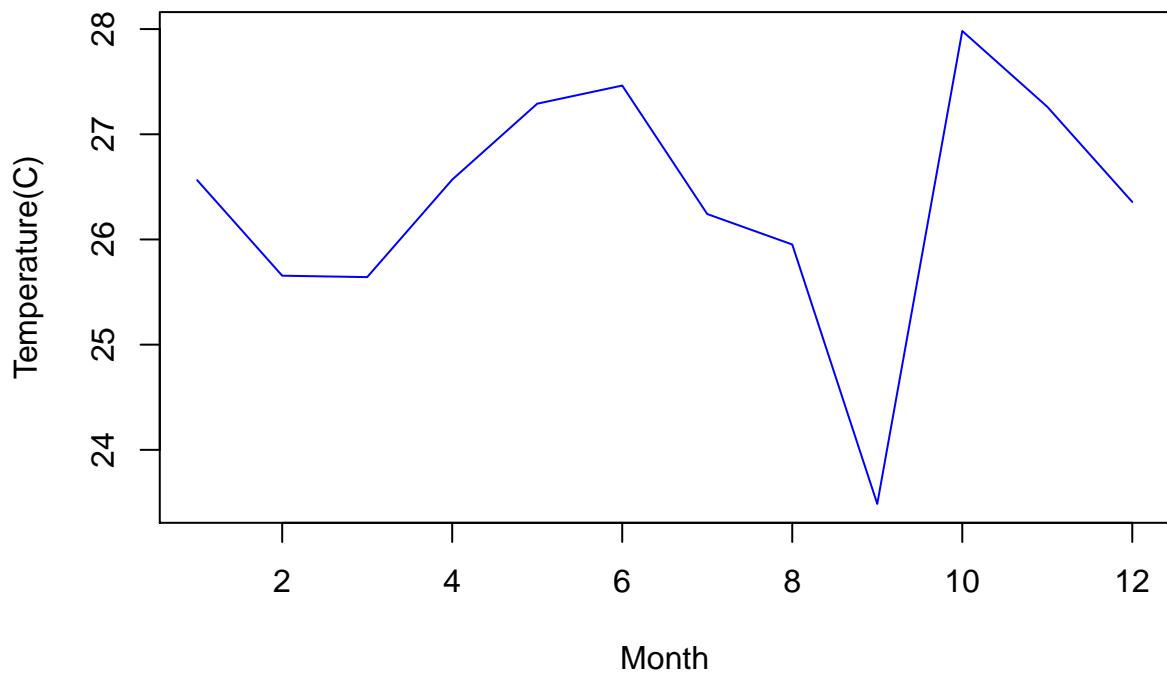
Average Sea Temperature of 2011



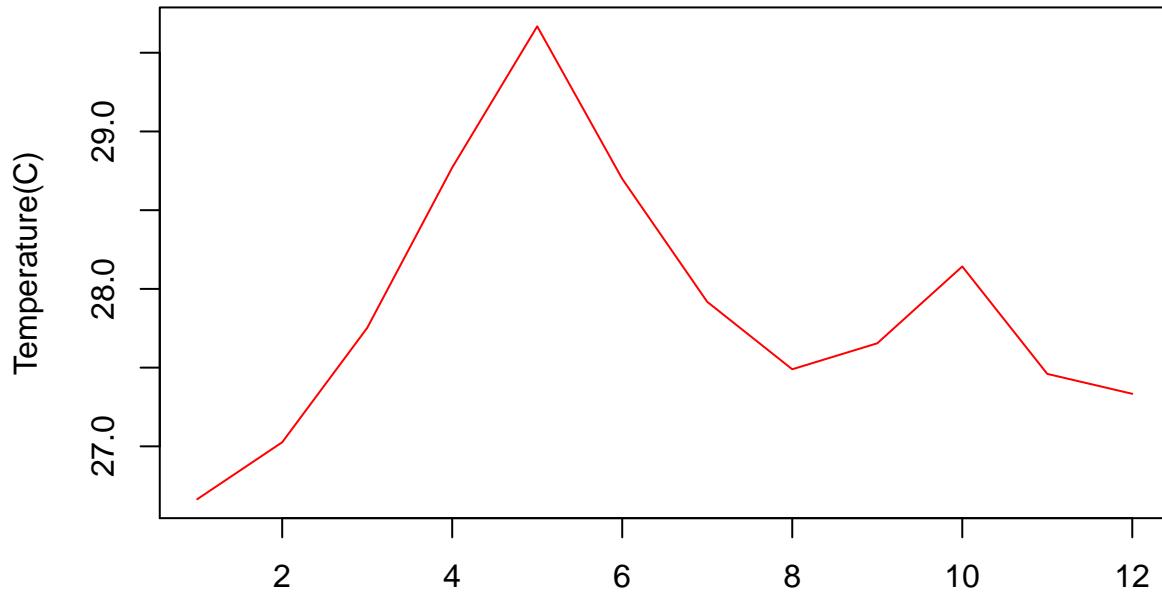
Average Air Temperature of 2011



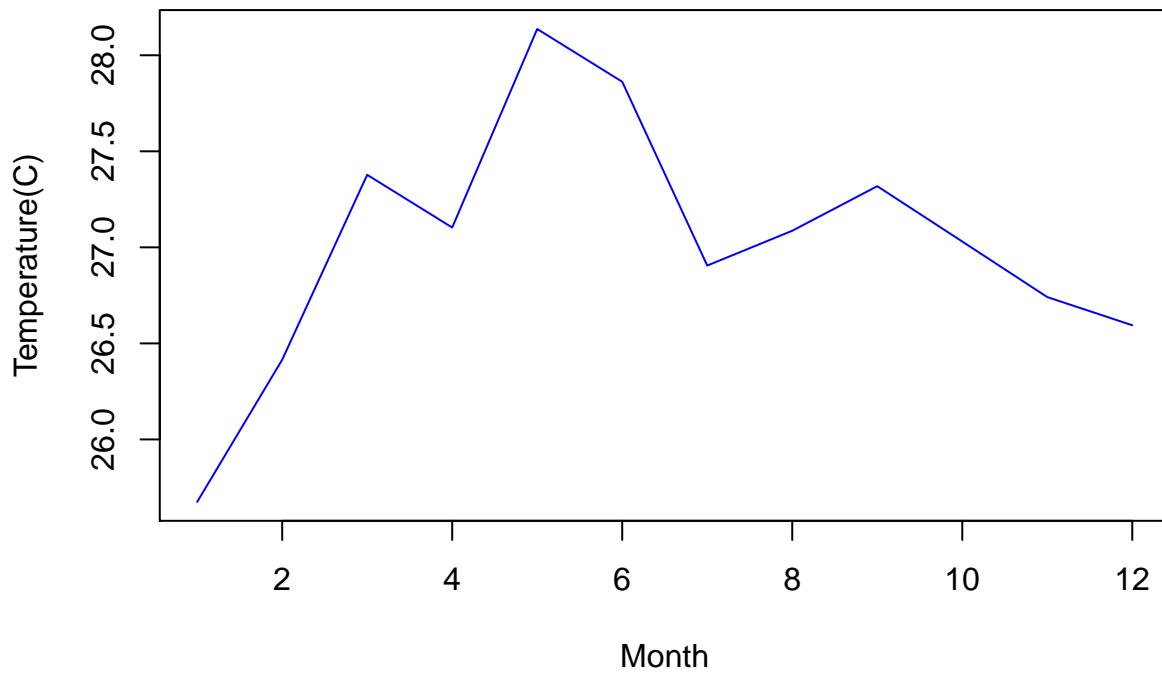
Average Sea Temperature of 2012



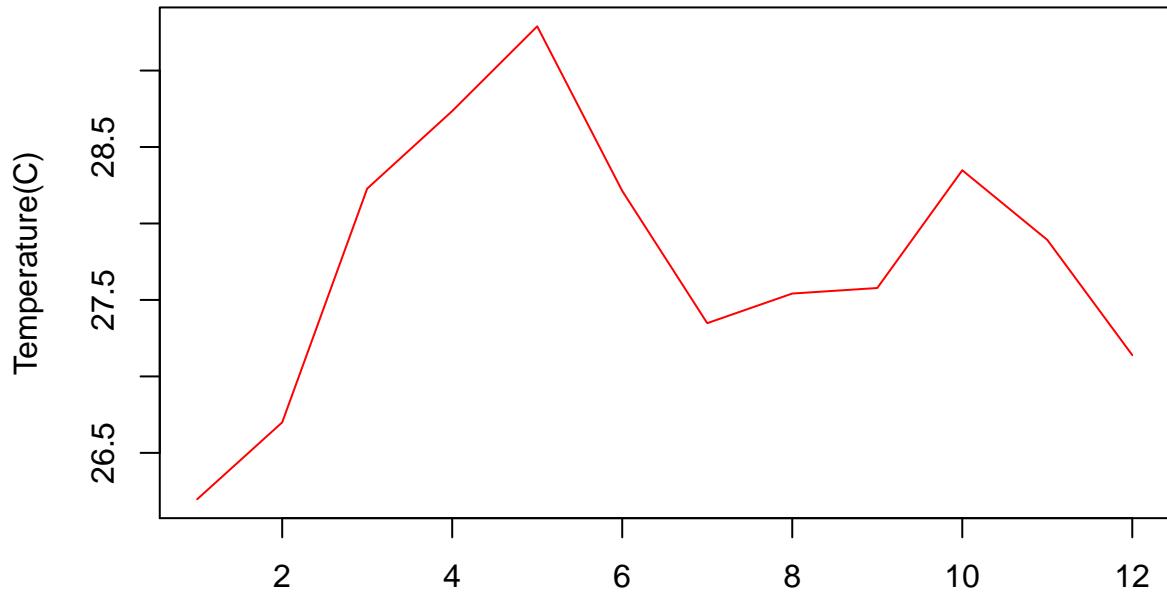
Average Air Temperature of 2012



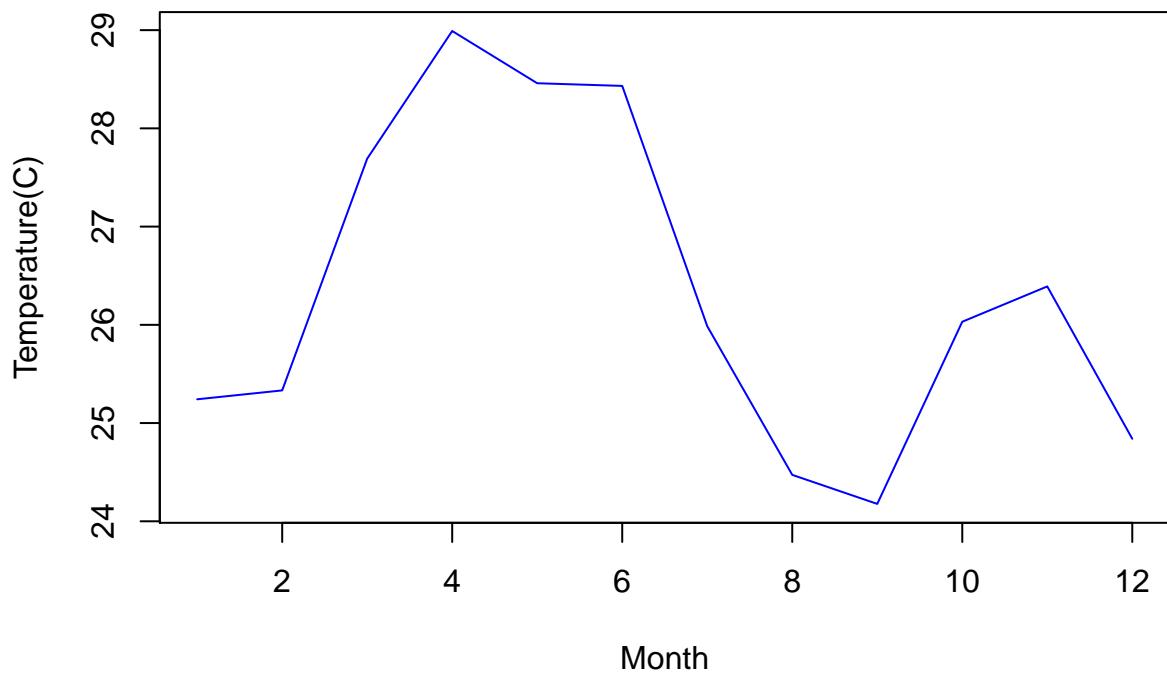
Average Sea Temperature of 2013



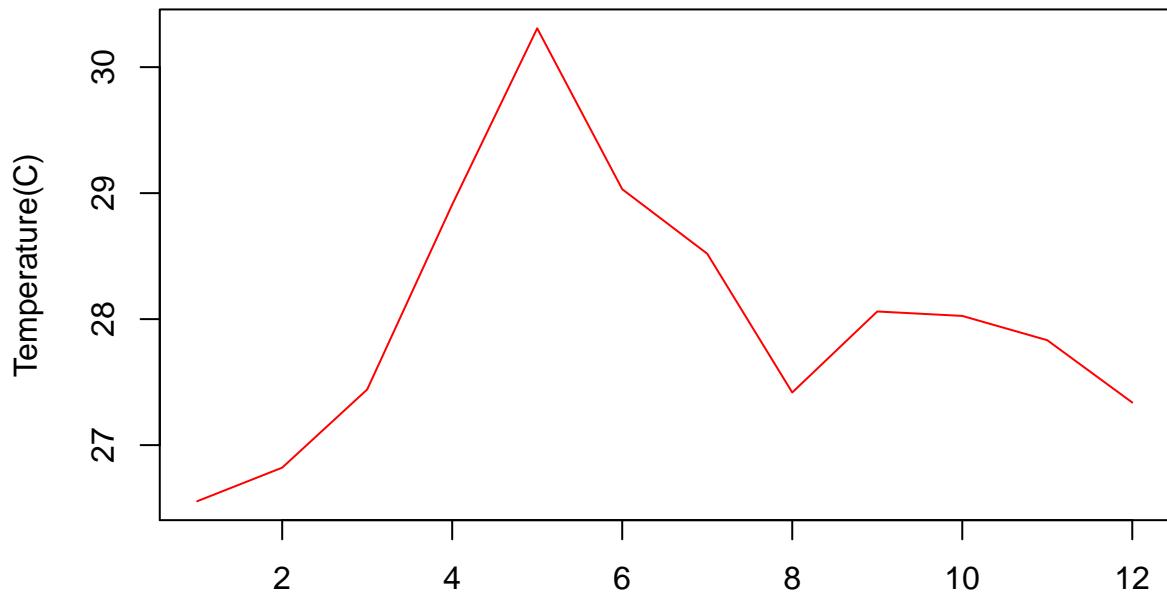
Average Air Temperature of 2013



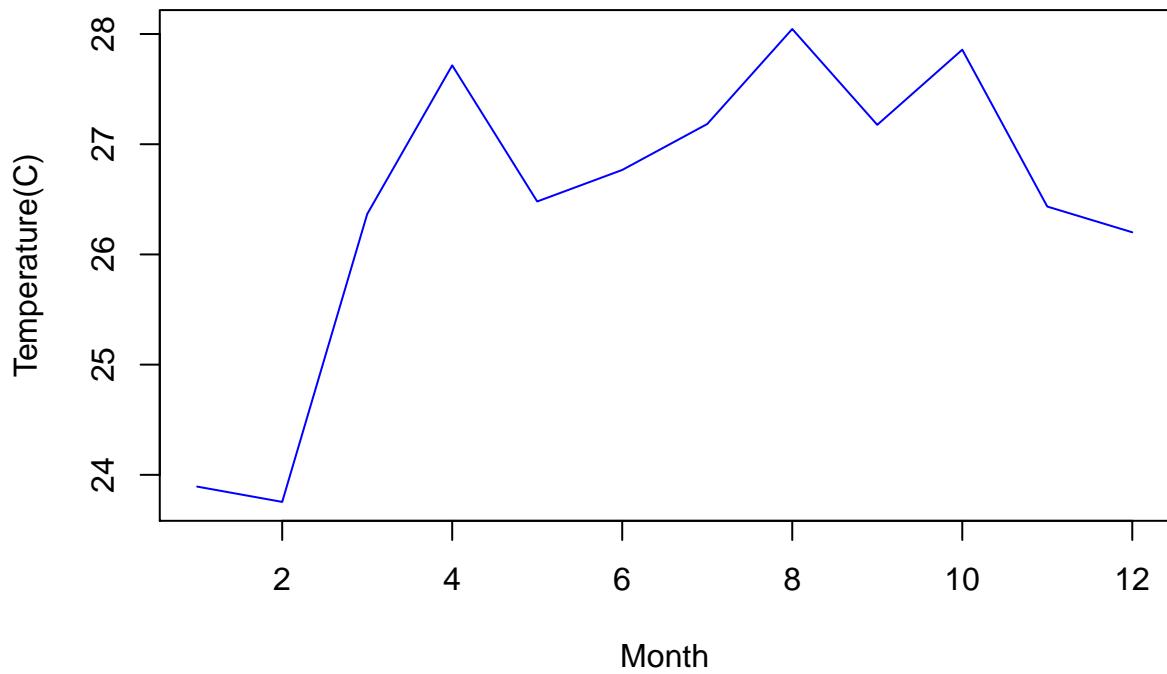
Average Sea Temperature of 2014



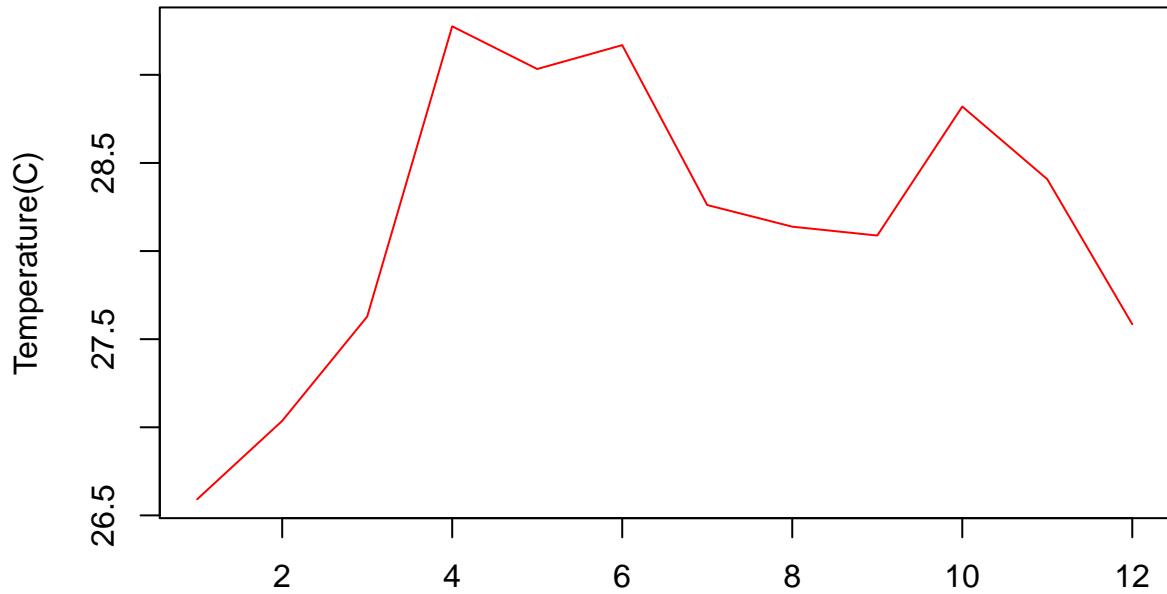
Average Air Temperature of 2014



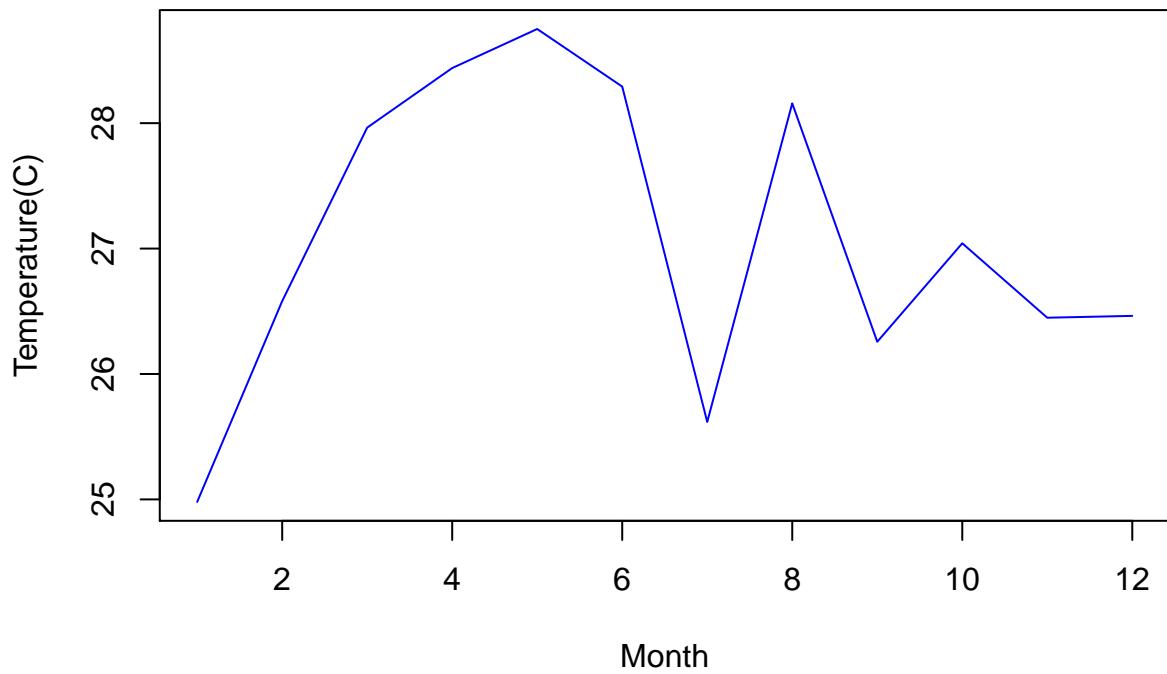
Average Sea Temperature of 2015



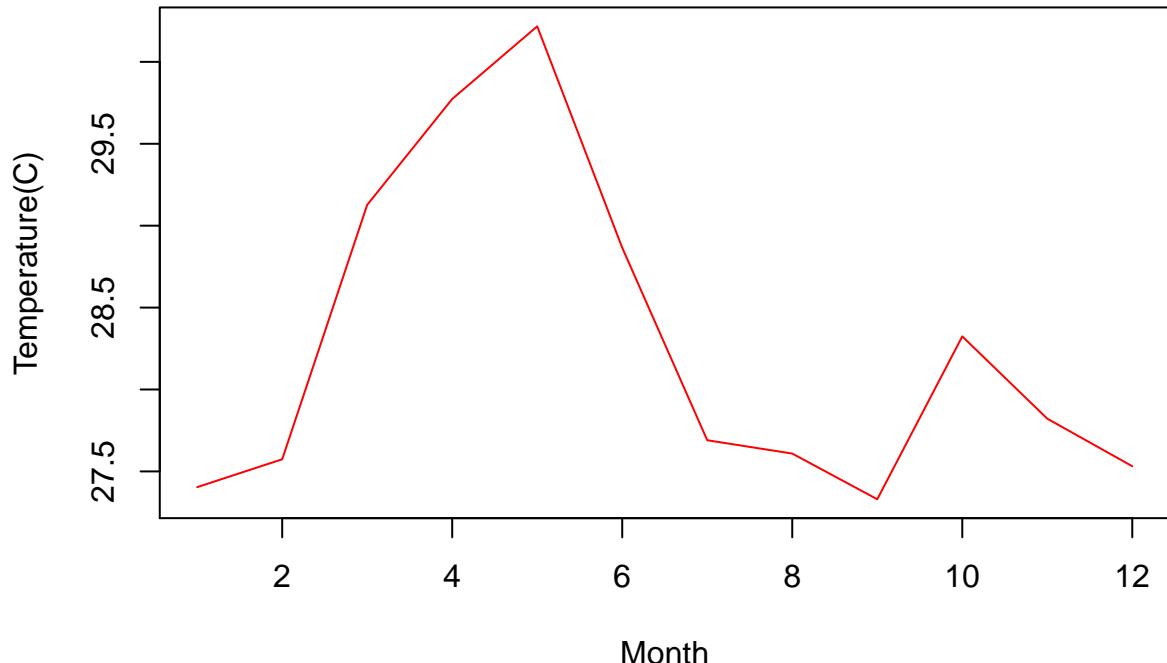
Average Air Temperature of 2015



Average Sea Temperature of 2016



Average Air Temperture of 2016



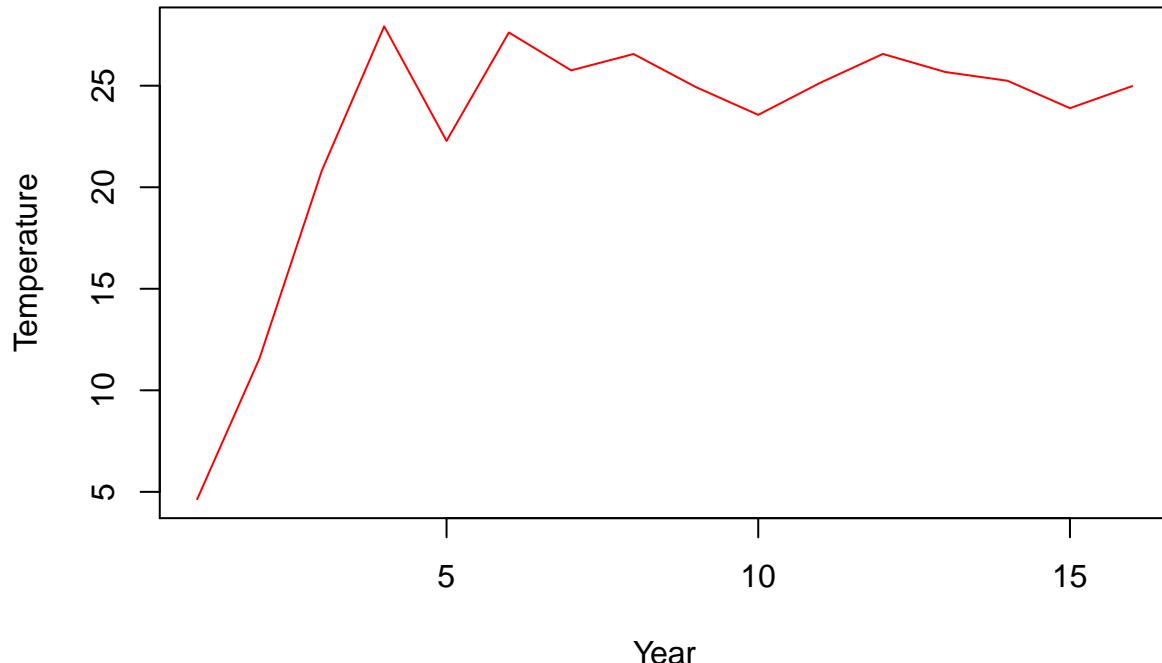
```
colnames(matrix.jan.temp) = c("Sea Temp" , "Air Temp") #rename each column
rownames(matrix.jan.temp) = c("2001","2002","2003","2004","2005","2006","2007","2008","2009","2010","2011","2012","2013","2014","2015","2016")

colnames(matrix.aug.temp) = c("Sea Temp" , "Air Temp") #rename each column
rownames(matrix.aug.temp) = c("2001","2002","2003","2004","2005","2006","2007","2008","2009","2010","2011","2012","2013","2014","2015","2016")
```

Graphing the records of average temperatures of January and August in 16 years.

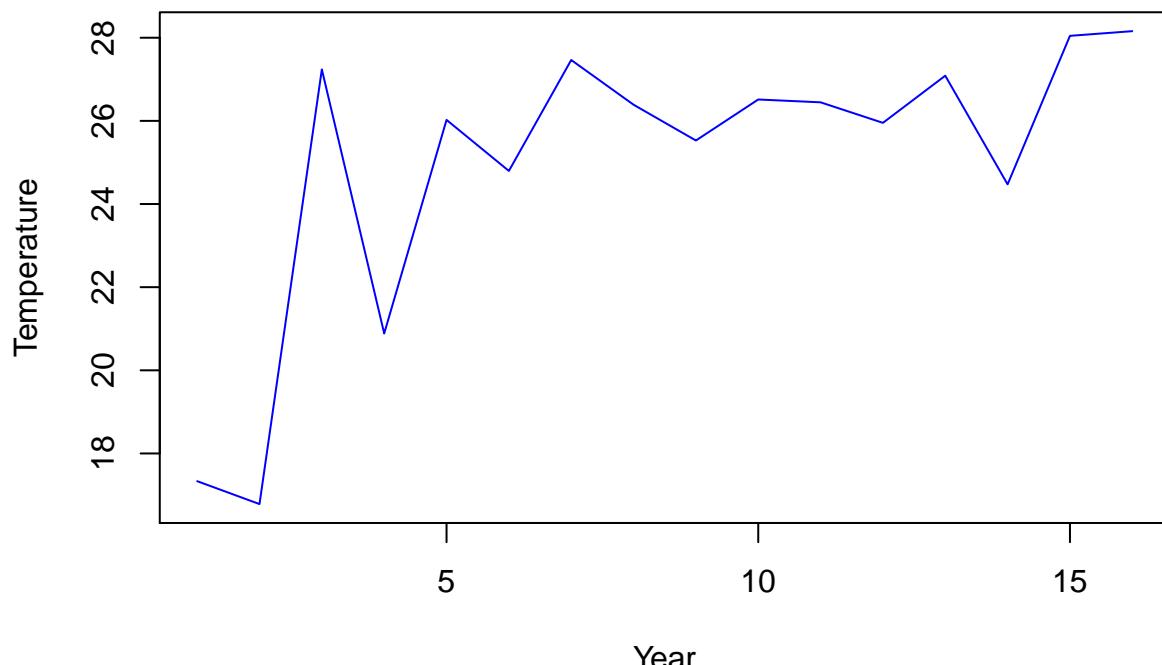
```
#####graphing the records of average temperatures of January and August in 16 years.
plot(matrix.jan.temp[,1], type = "l", main = "Average Sea Temperture of Jaunuary" , xlab = "Year", yla
```

Average Sea Temperture of Jaunuary



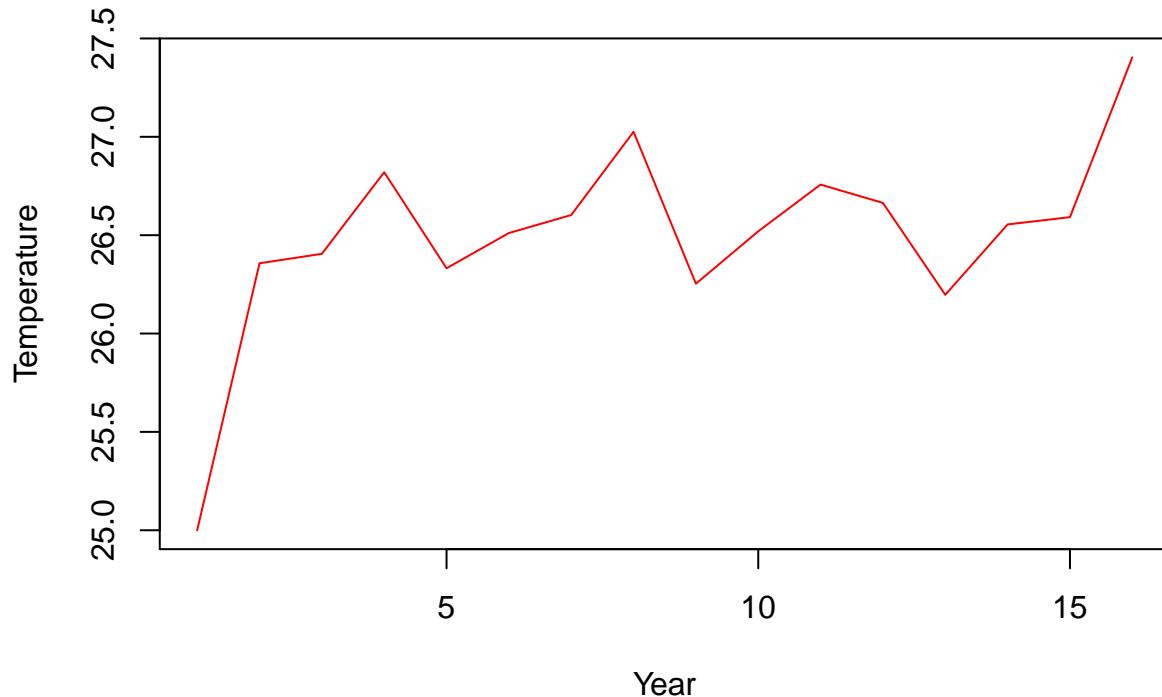
```
plot(matrix.aug.temp[,1], type = "l", main = "Average Sea Temperture of August" , xlab = "Year", ylab = "Temperature")
```

Average Sea Temperture of August



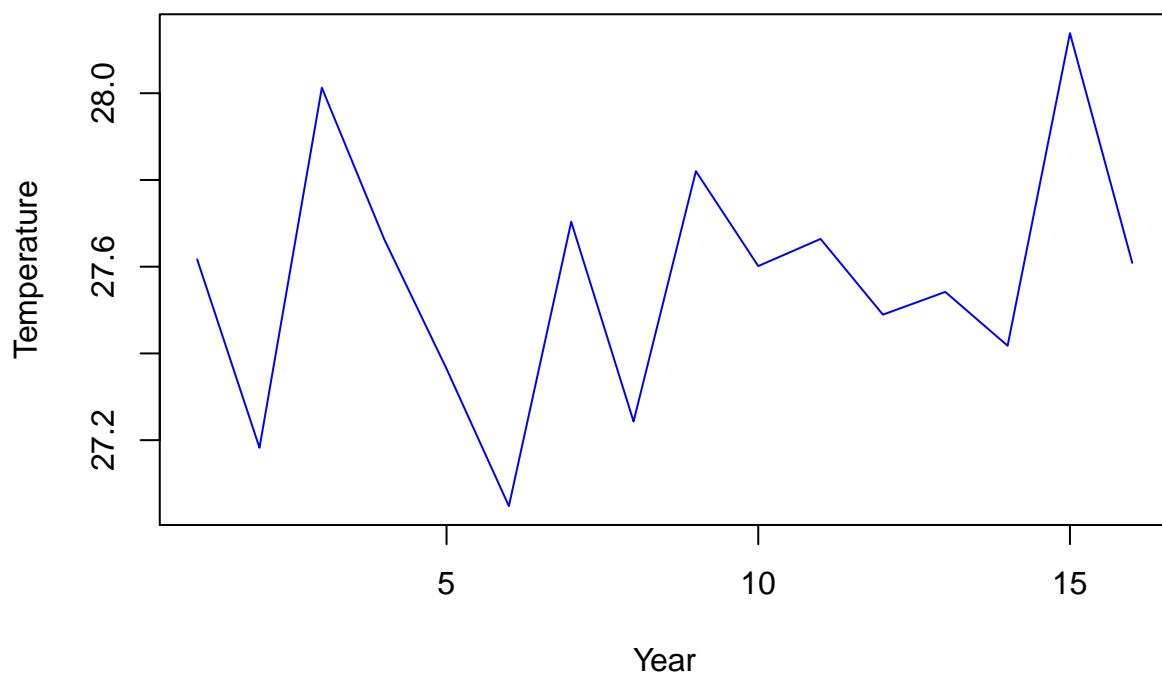
```
plot(matrix.jan.temp[,2], type = "l", main = "Average Air Temperture of Jaunuary" , xlab = "Year", ylab = "Temperature")
```

Average Air Temperture of Jaunuary

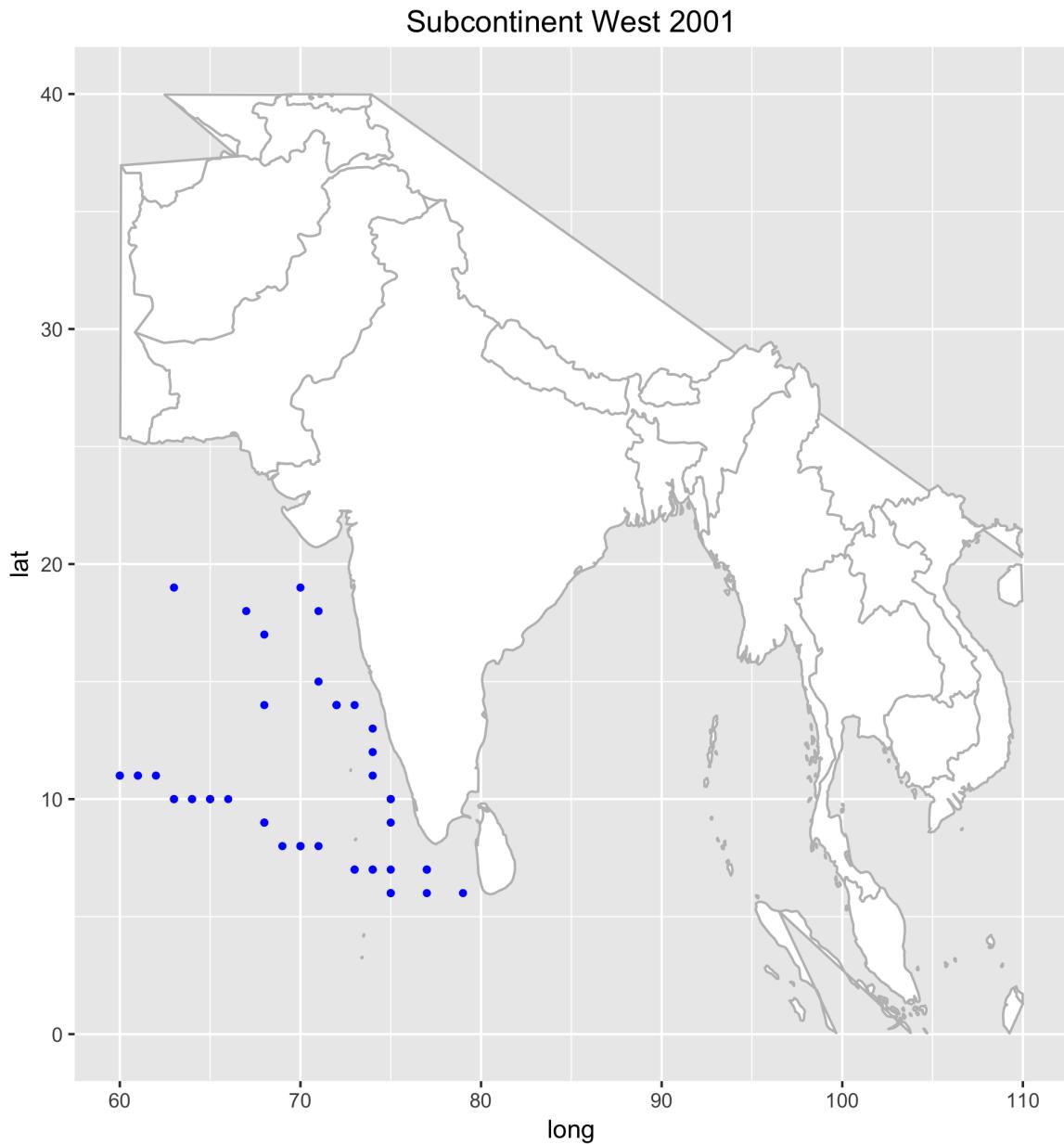


```
plot(matrix.aug.temp[,2], type = "l", main = "Average Air Temperture of August" , xlab = "Year", ylab = "Temperature")
```

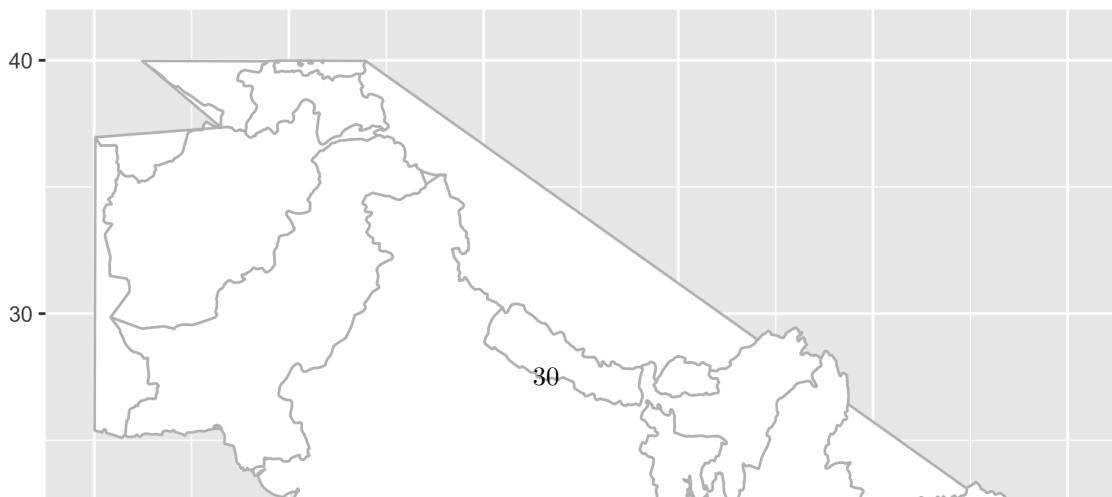
Average Air Temperture of August



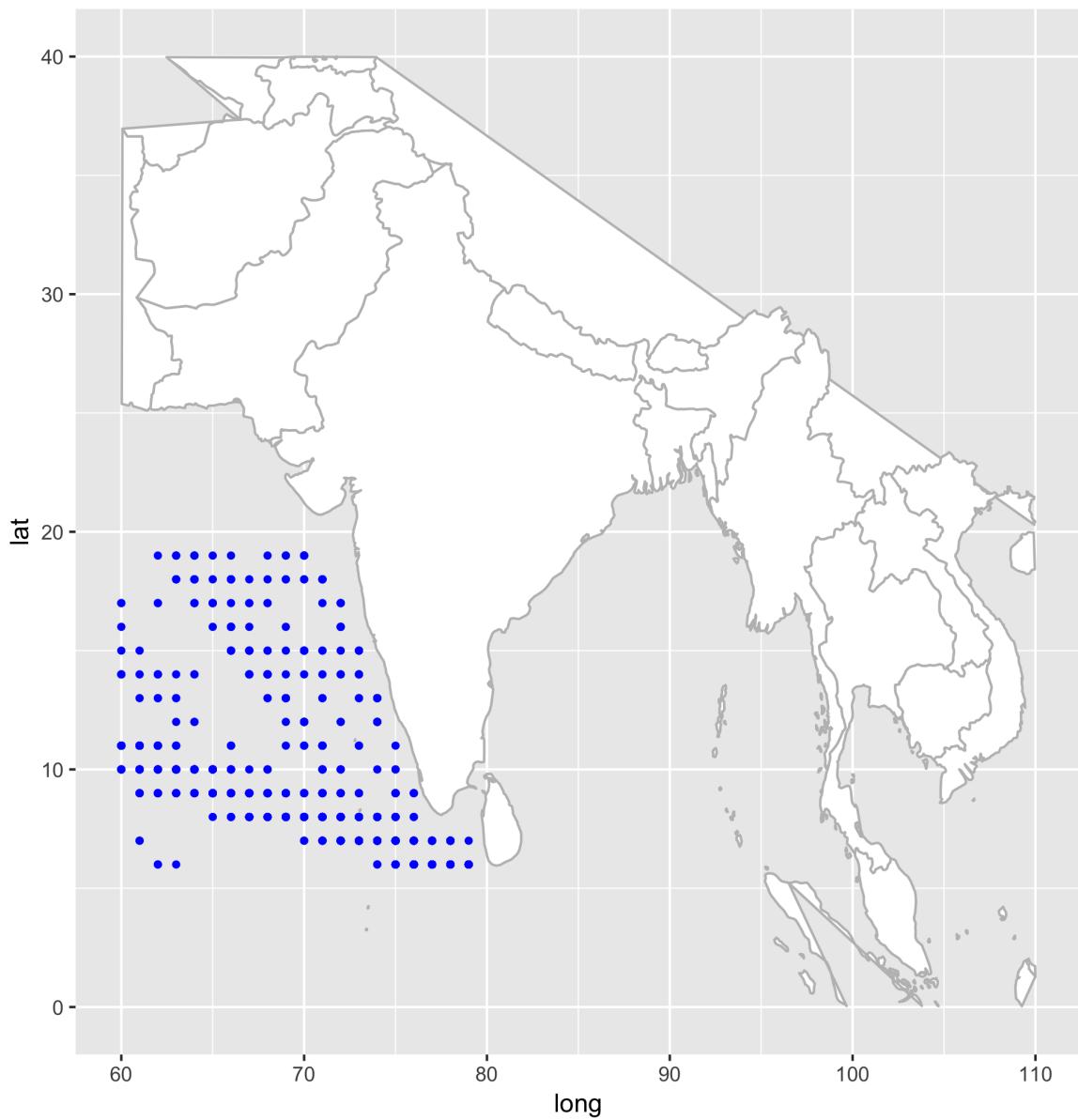
All Maps of Subcontinent West with data collection locations
(marked as blue spot) in year 2001 - 2016



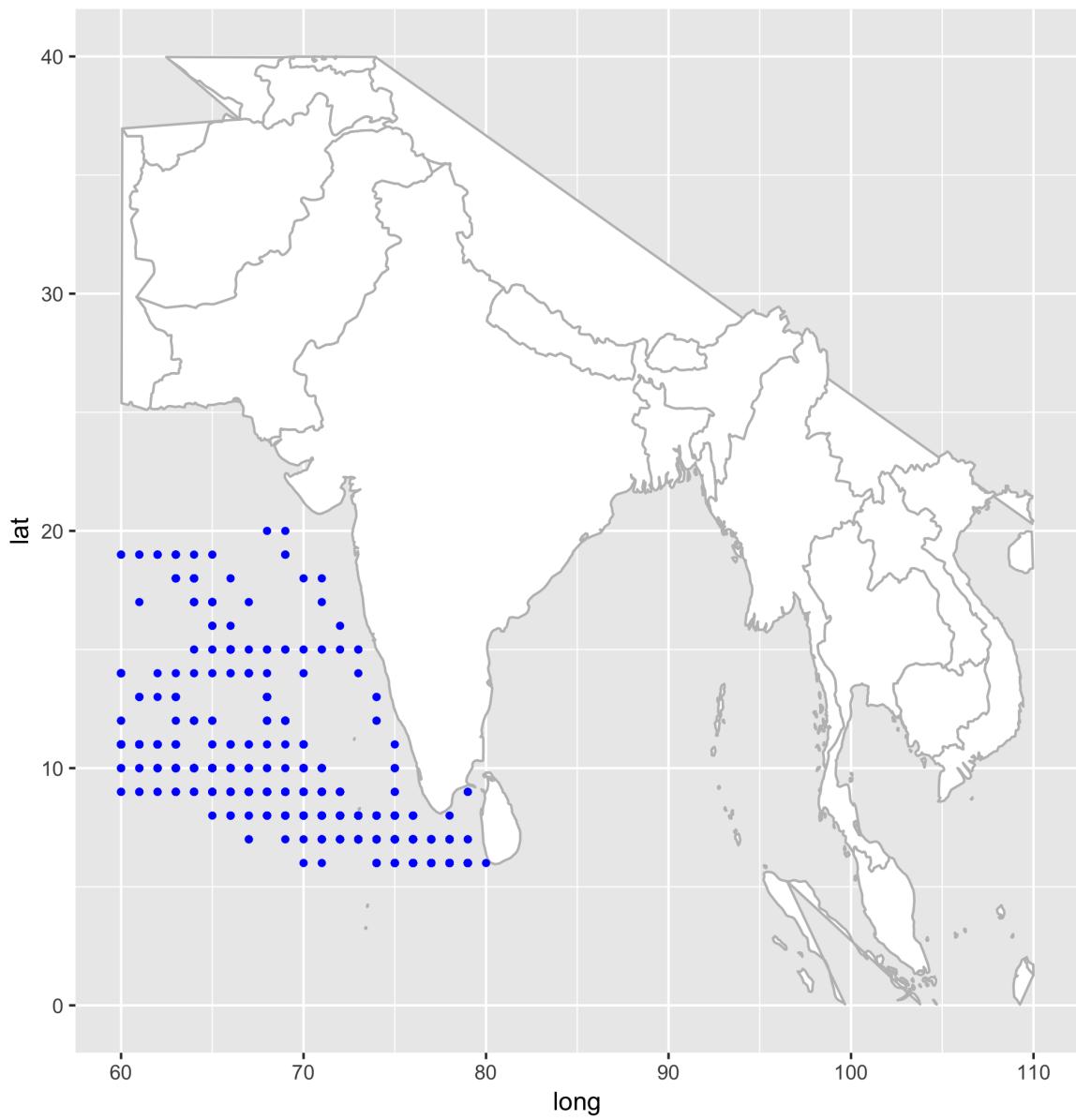
Subcontinent West 2002



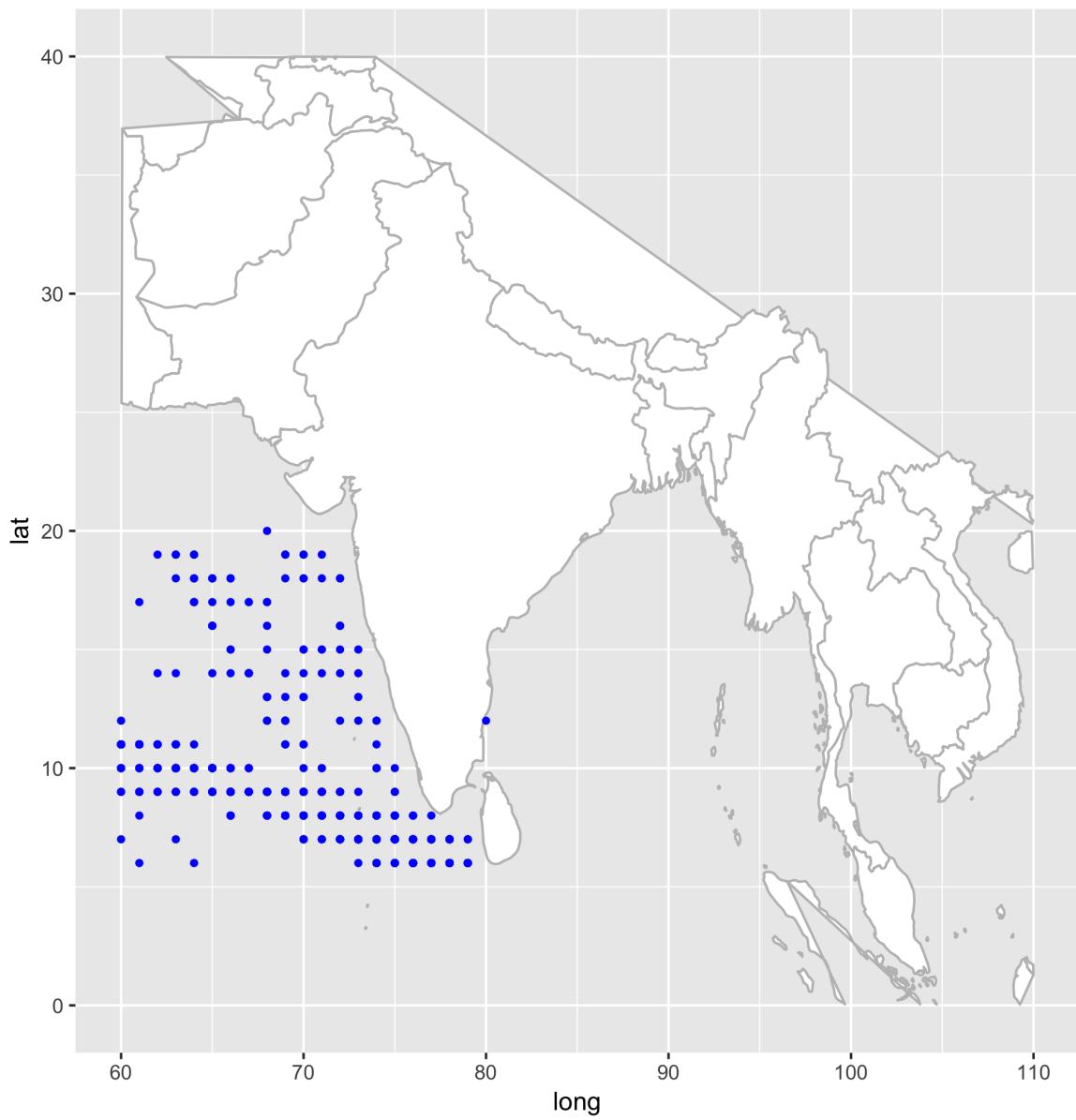
Subcontinent West 2003



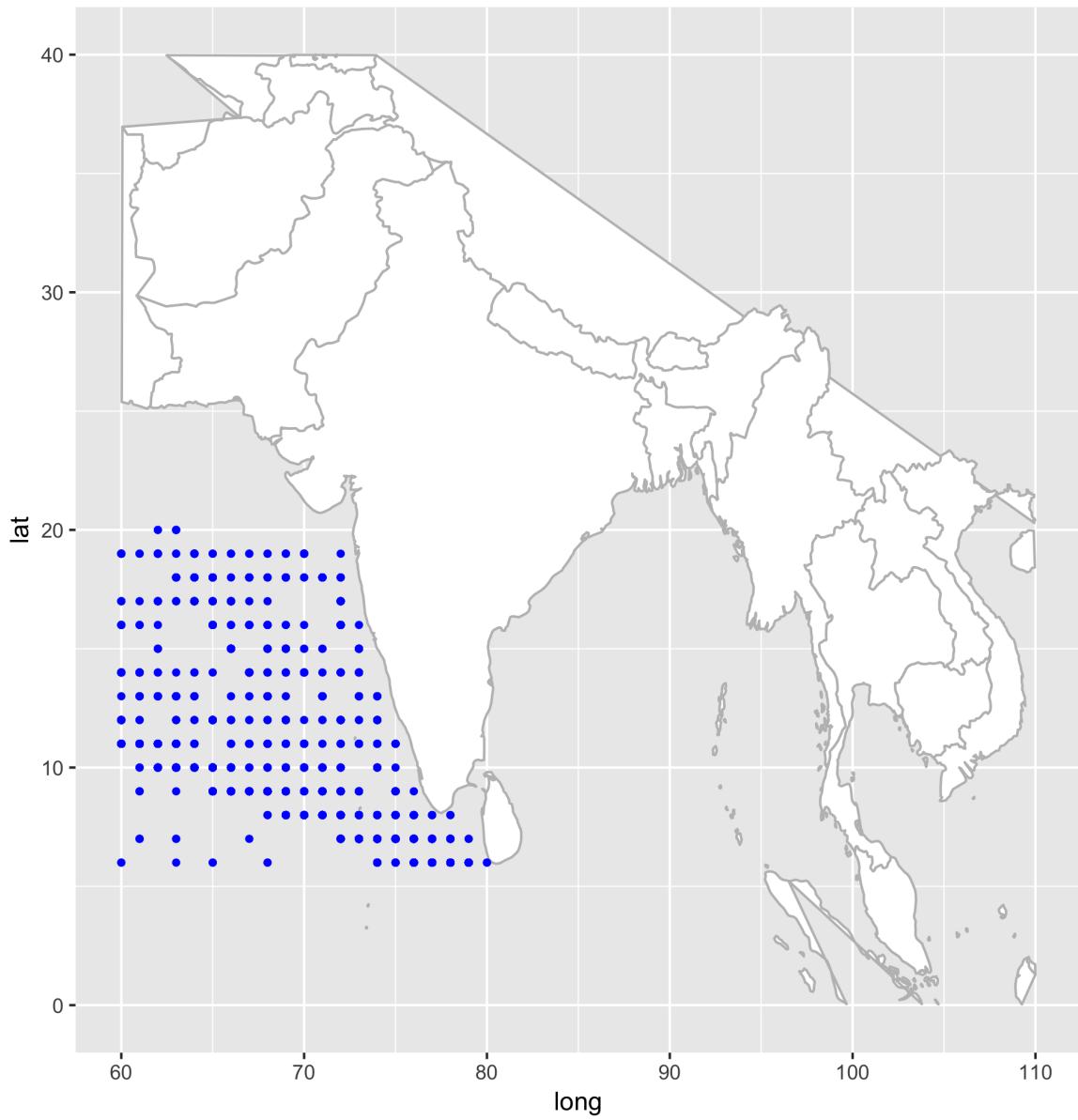
Subcontinent West 2004



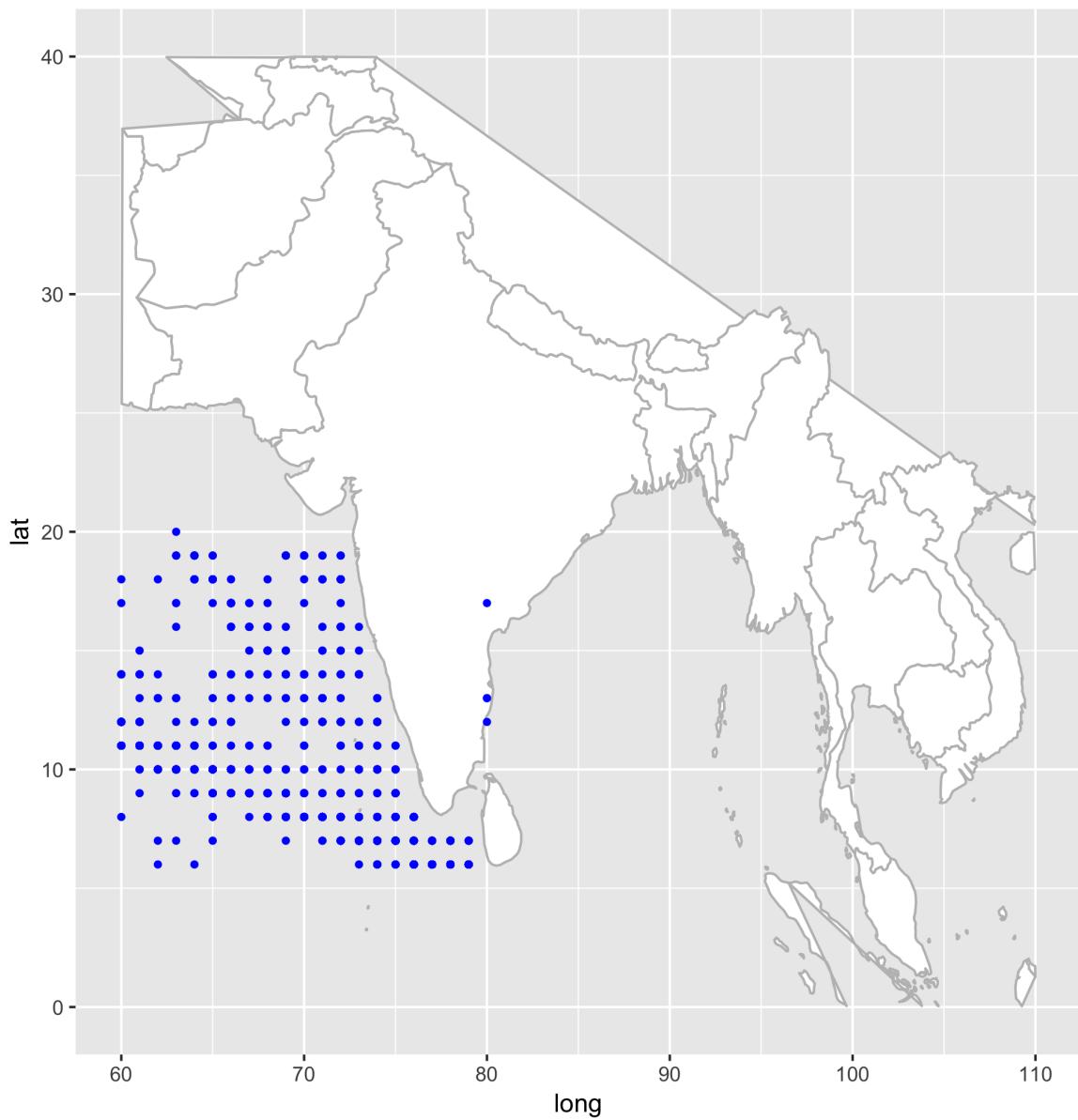
Subcontinent West 2005



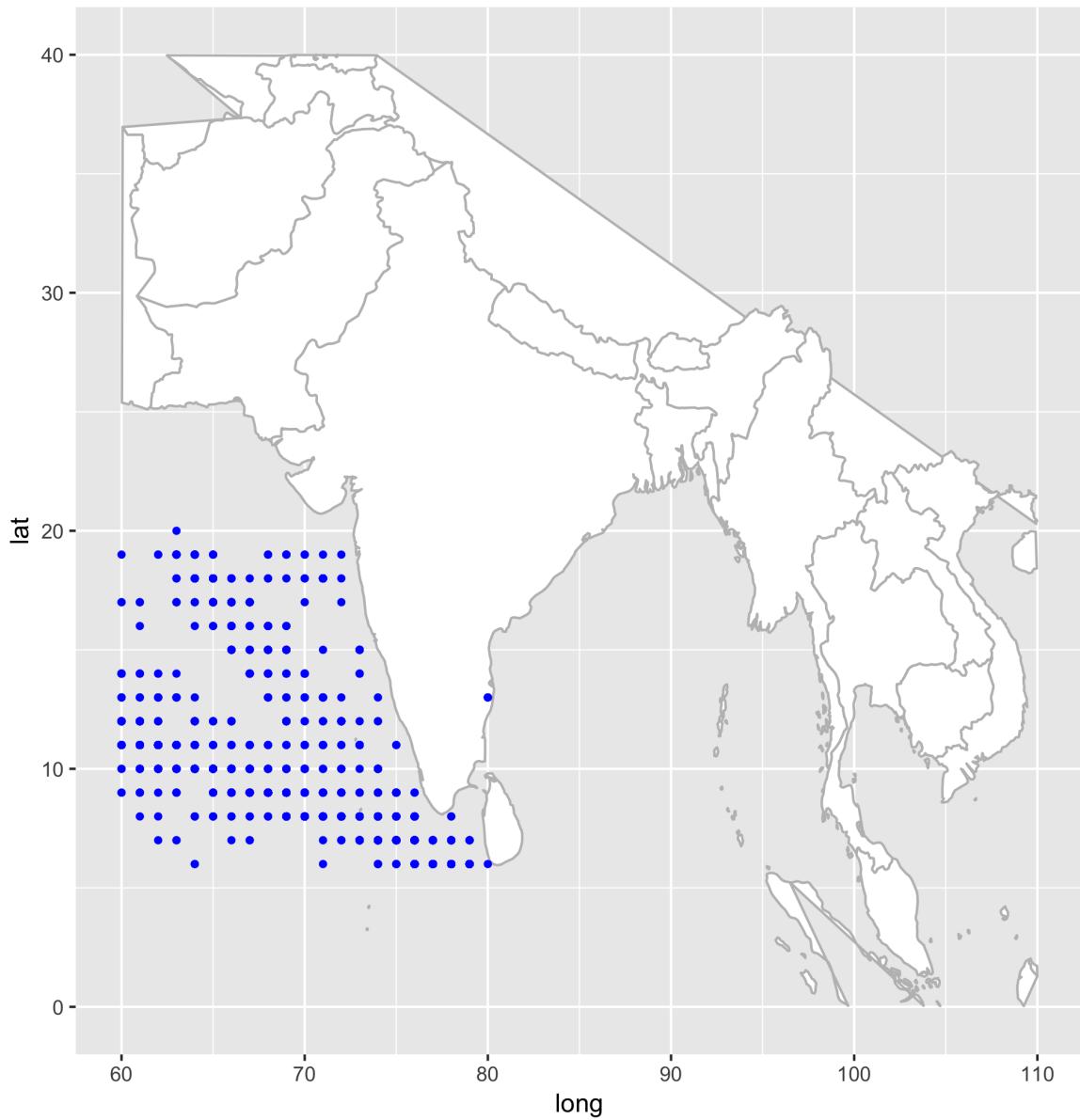
Subcontinent West 2006



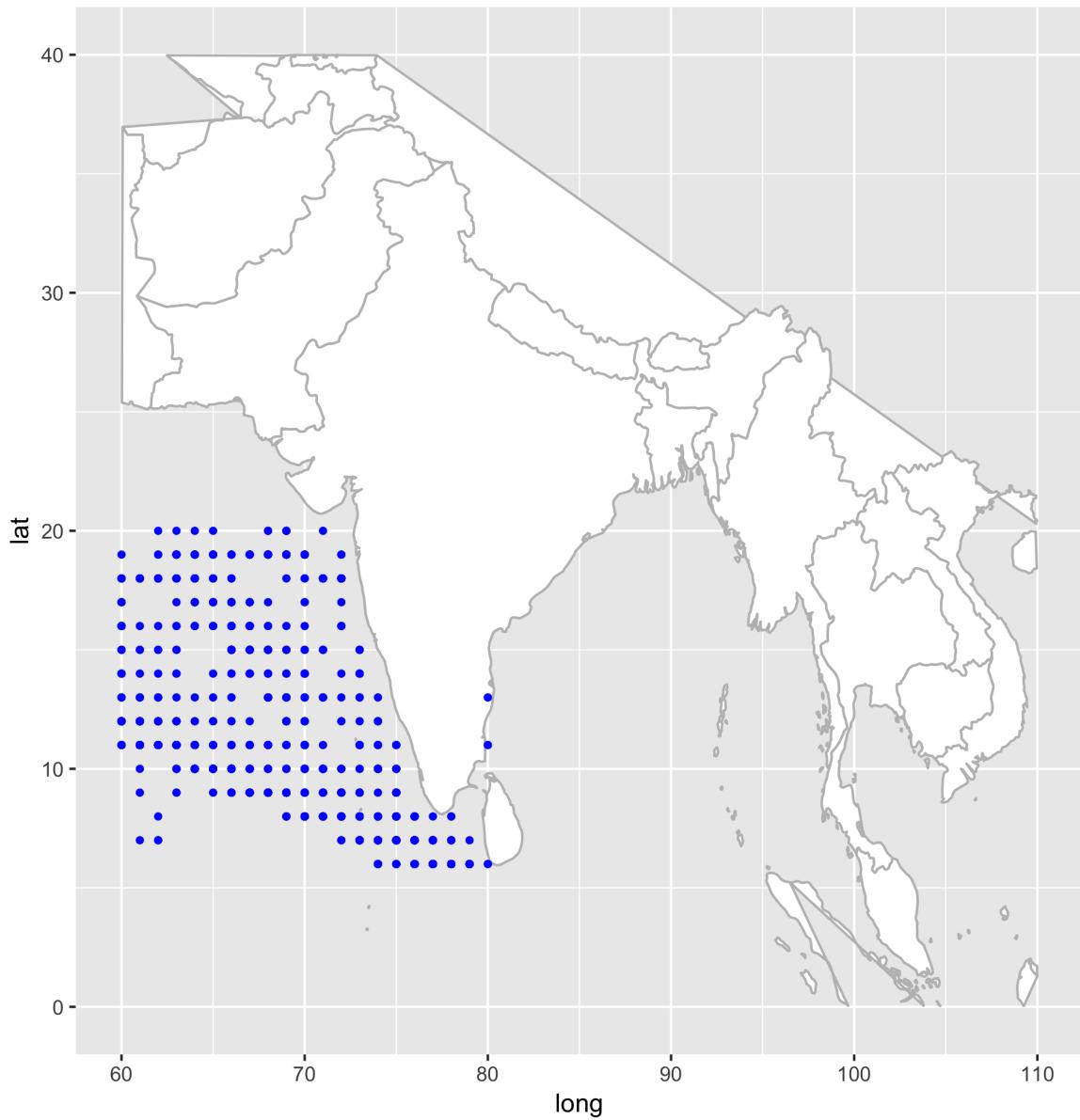
Subcontinent West 2007



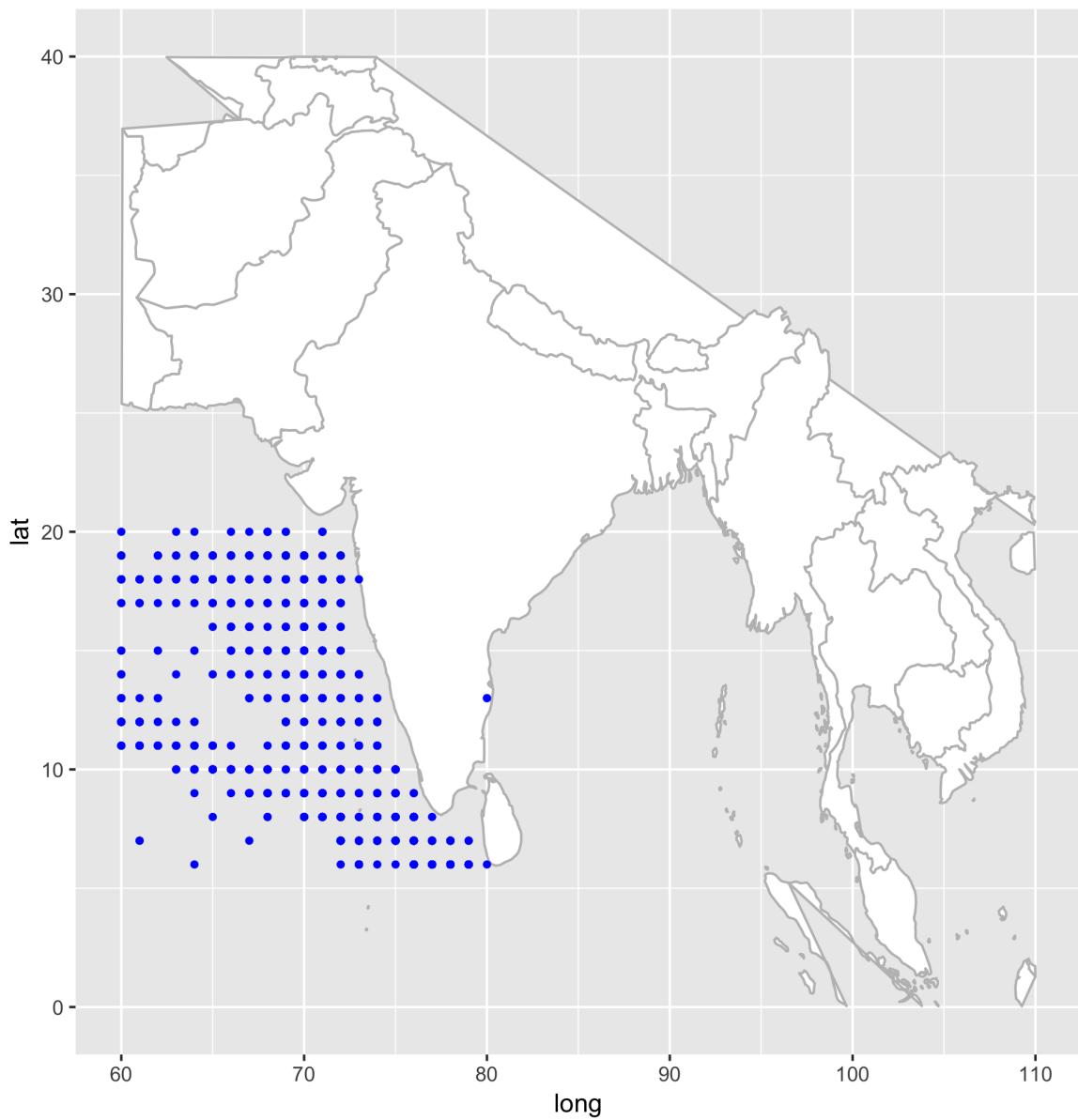
Subcontinent West 2008



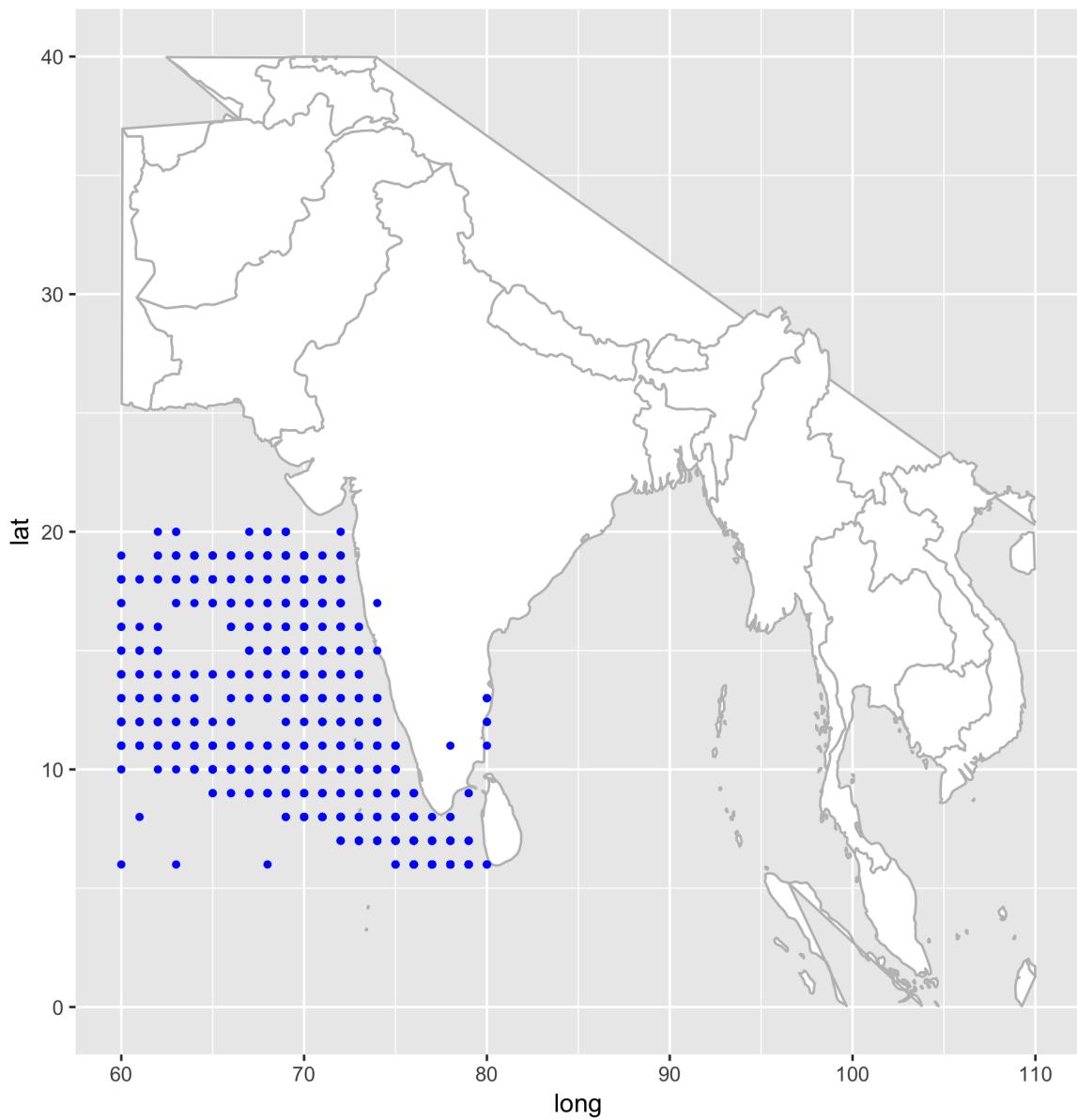
Subcontinent West 2009



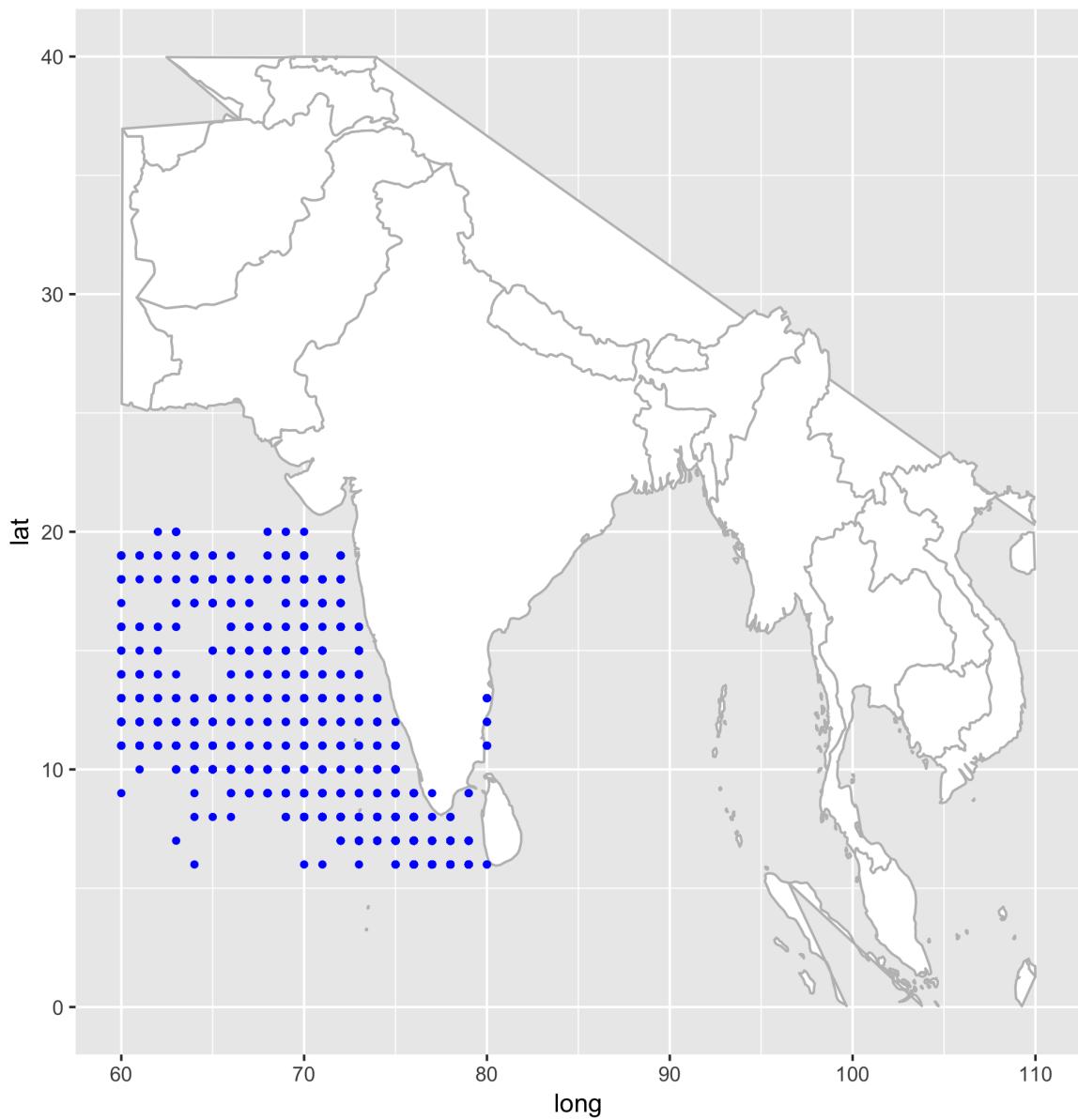
Subcontinent West 2010



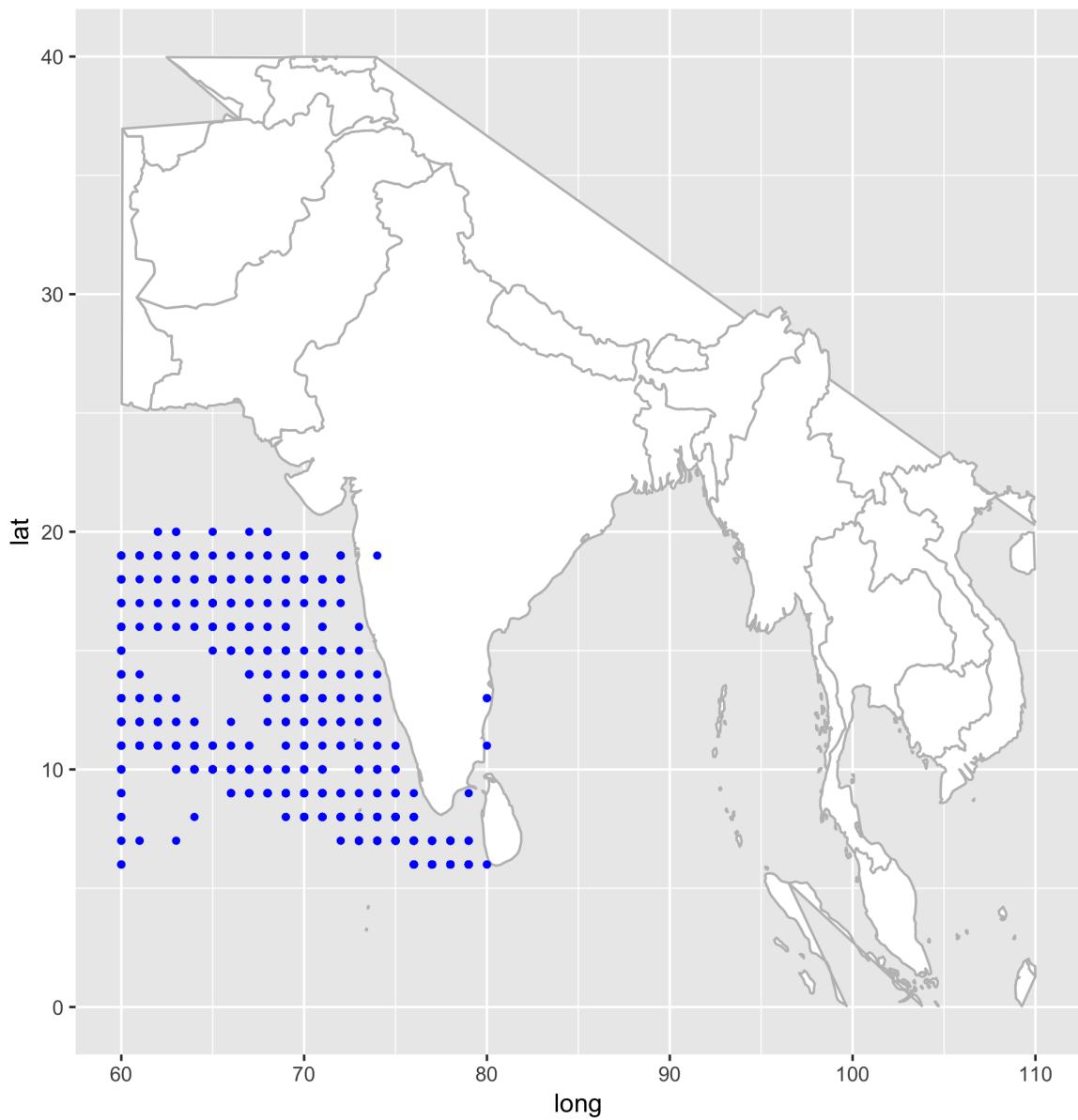
Subcontinent West 2011



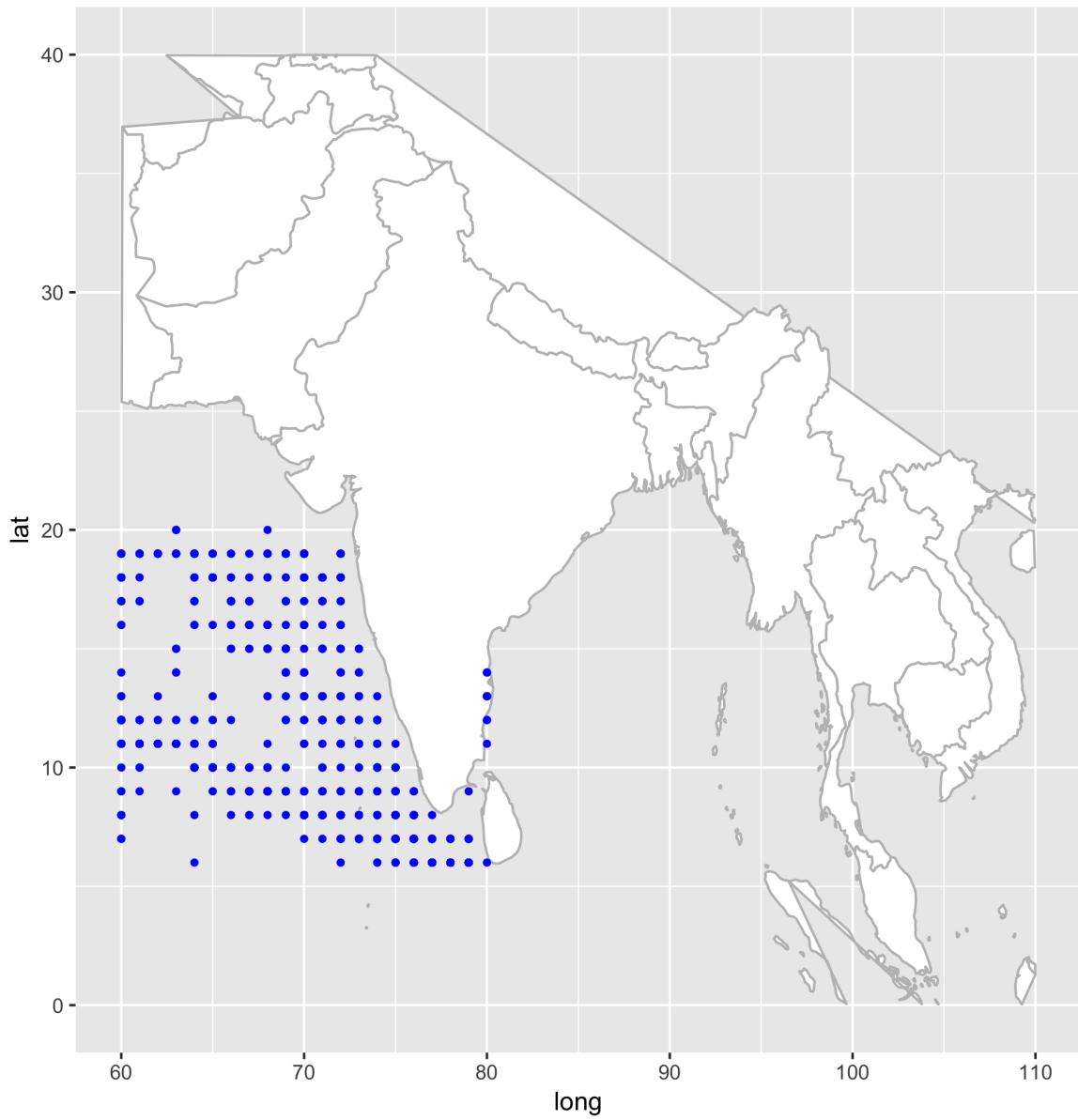
Subcontinent West 2012



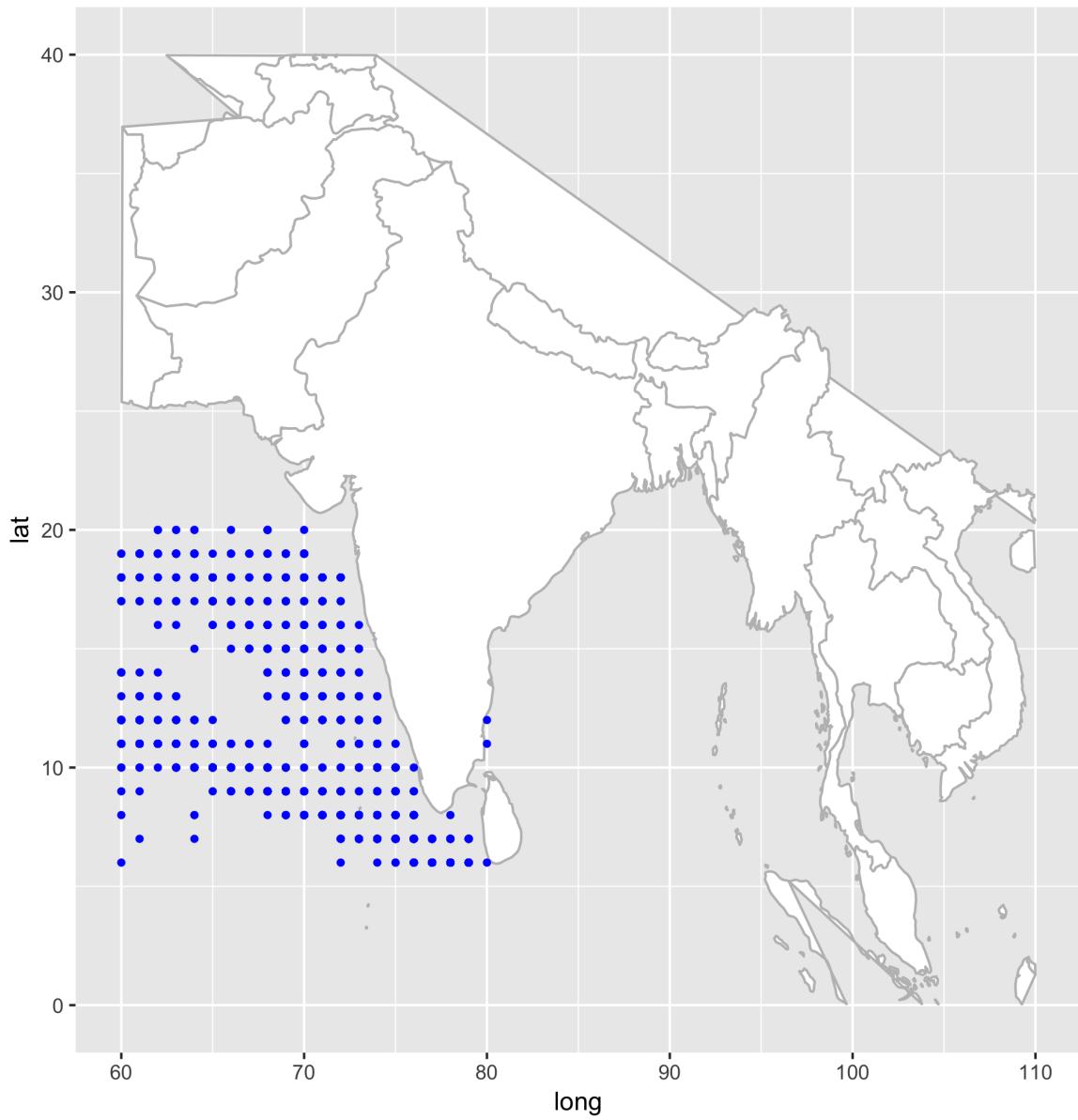
Subcontinent West 2013



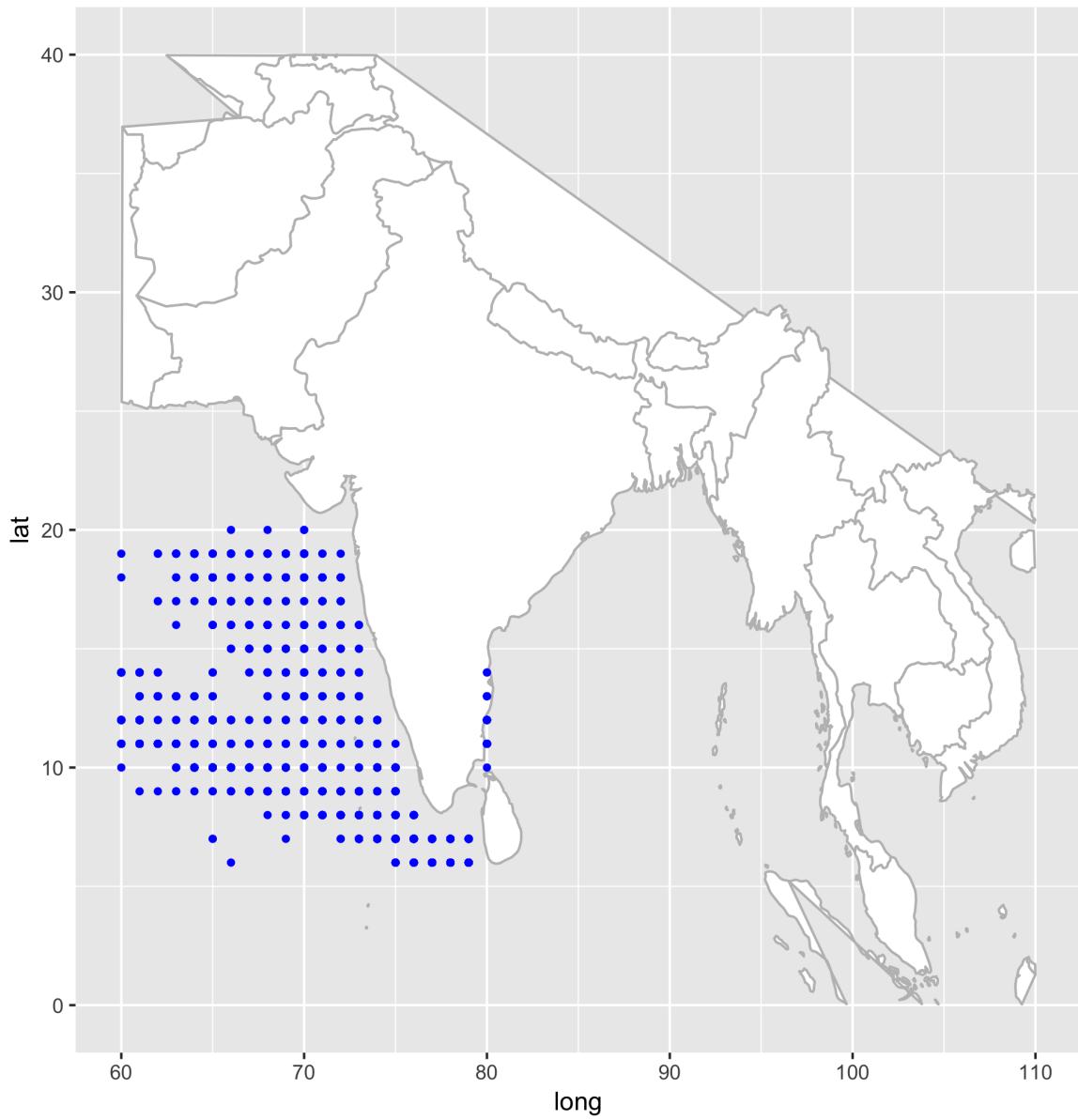
Subcontinent West 2014



Subcontinent West 2015



Subcontinent West 2016



Collaboration

Sibo Zhu and Li Liu are in charge of all coding parts and data cleaning. While Xiaoqian Xue and Danni Fu are in charge of all analysis of data and EDA.