

《计算与编程》第四次作业报告

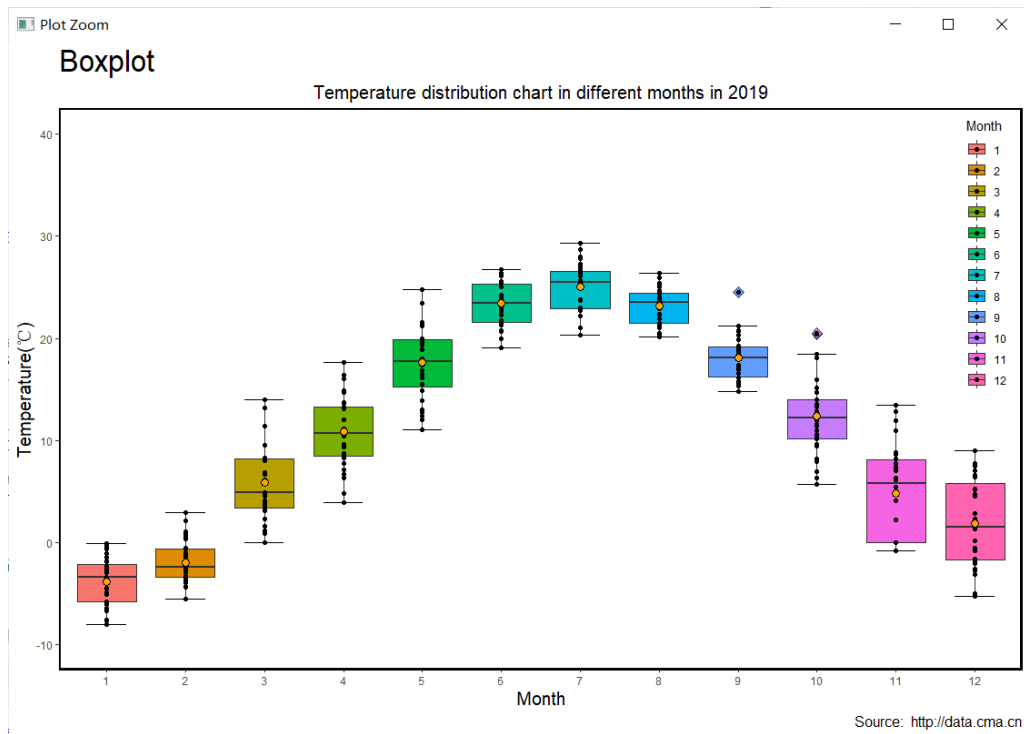
王超 12031012

第一题

思路 and 结果:

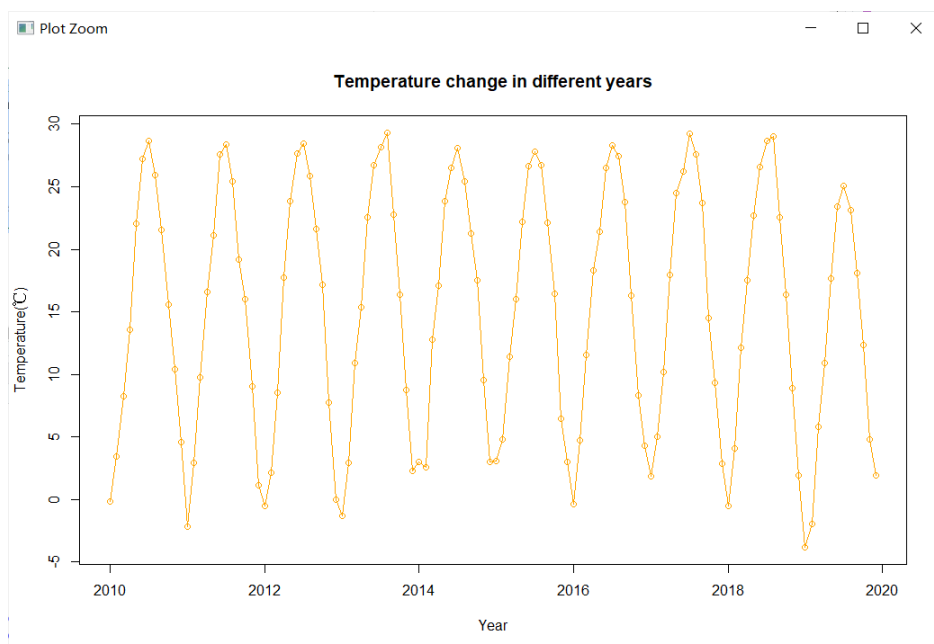
1.1

首先读取数据, 原始数据为新乡气象站近十年的气象数据, 目的是要得到该地 2019 年的各月份的气温分布情况, 最终所得的数据如下图所示, 源代码有具体的注释。



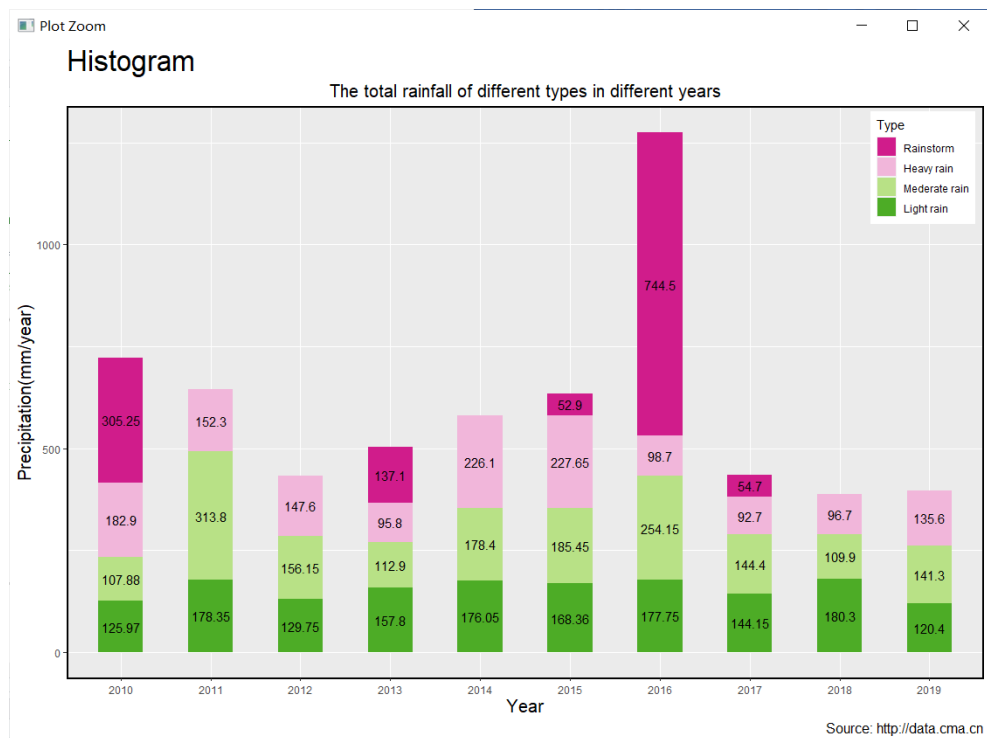
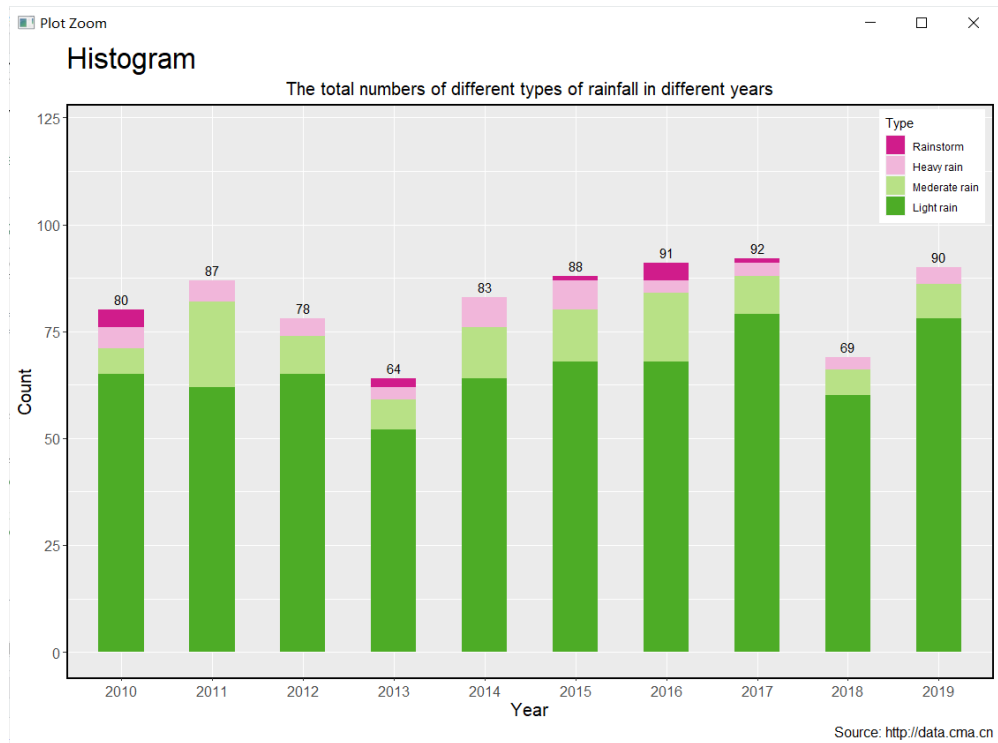
1.2

最终结果如下所示:



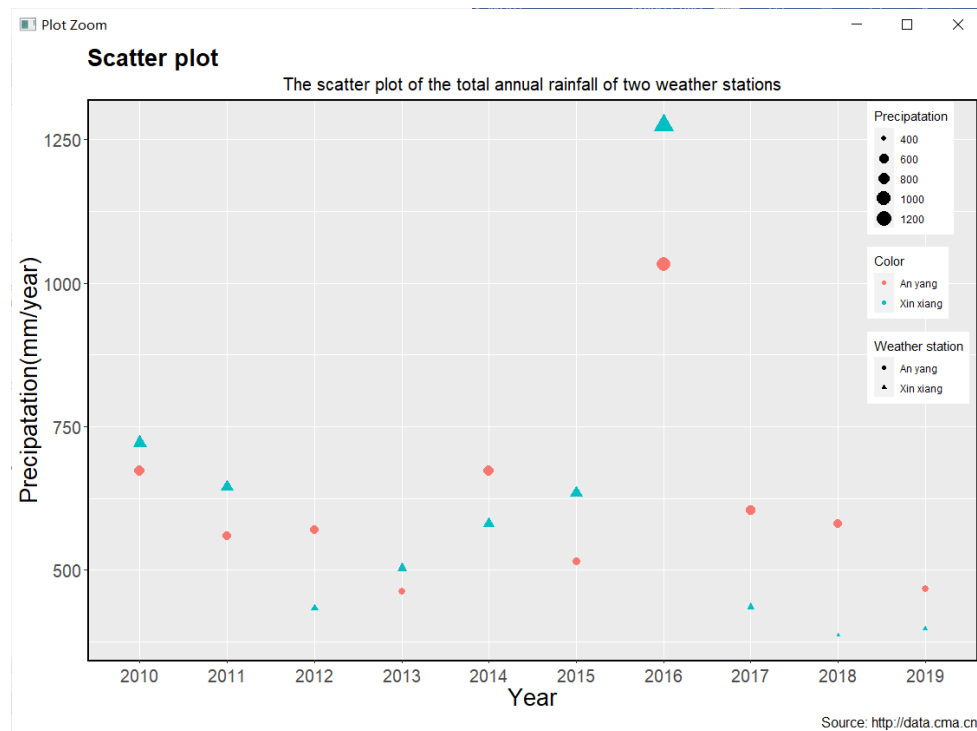
1.3

首先读取数据，原始数据为新乡气象站近十年的气象数据，目的是要得到该地各年的降雨类型及降雨量的分布情况，最终所得的数据如下图所示，源代码有具体的注释。



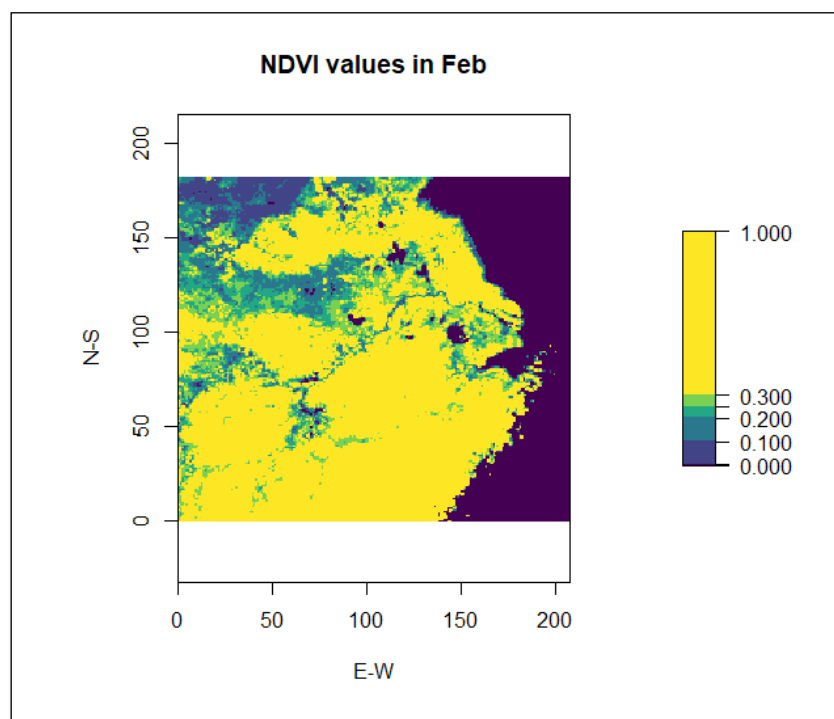
1.4

首先读取数据，原始数据为新乡气象站和安阳气象站近十年的降雨数据，目的是要比较这两个气象站近十年的年降雨总量，最终所得的数据如下图所示，源代码有具体的注释。：



1.5

下载长江入海口区域 2019 年 1 月份的 NDVI 数据（TIF 文件），总共有 208 * 183 个栅格点，根据 NDVI 值的大小对数据进行简单分类，不同类赋予不同颜色最终的结果如下图所示：

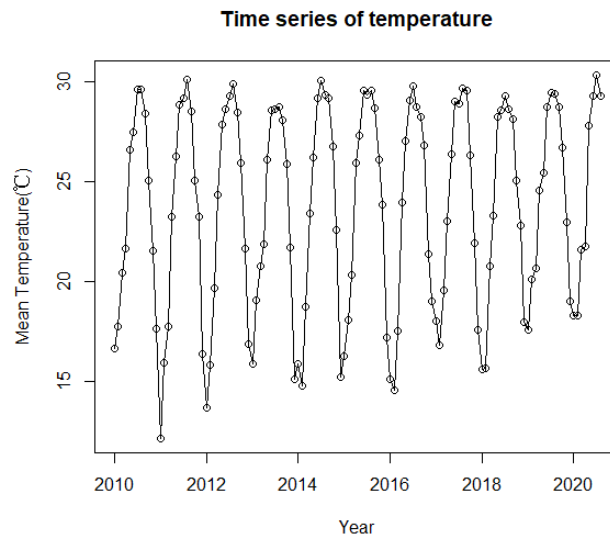


第二题

思路 and 结果:

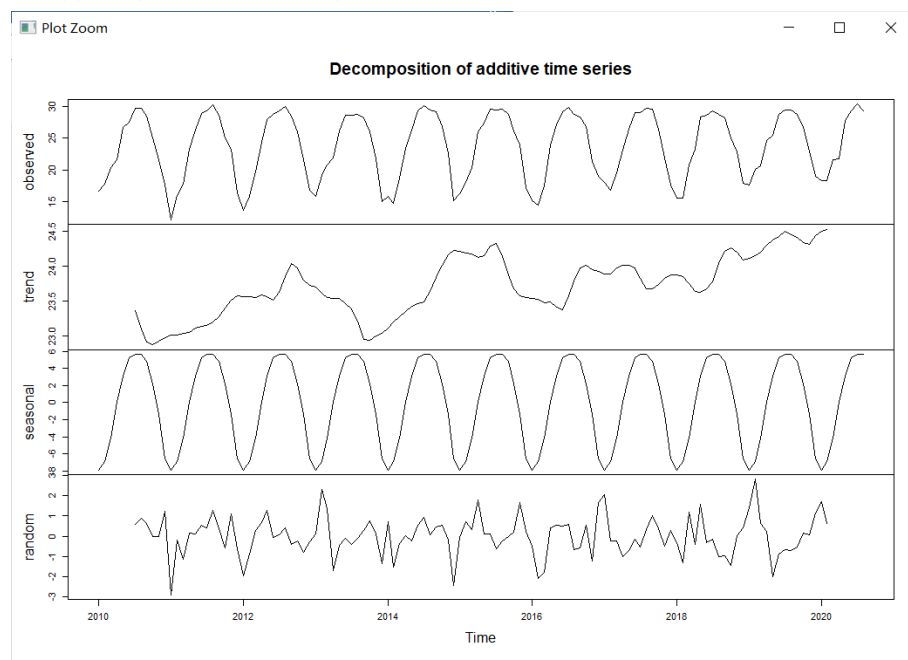
2.1

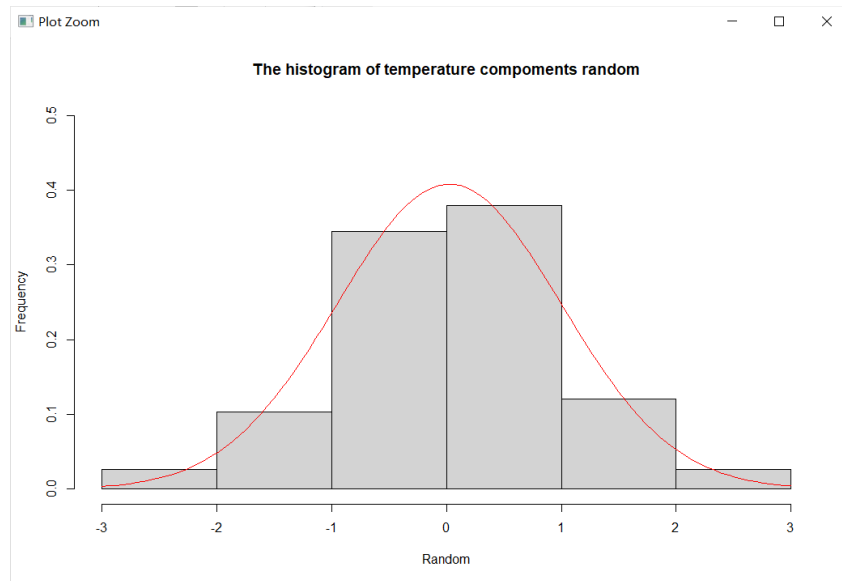
根据要求生成所需时间序列数据，再作图即可：



2.2

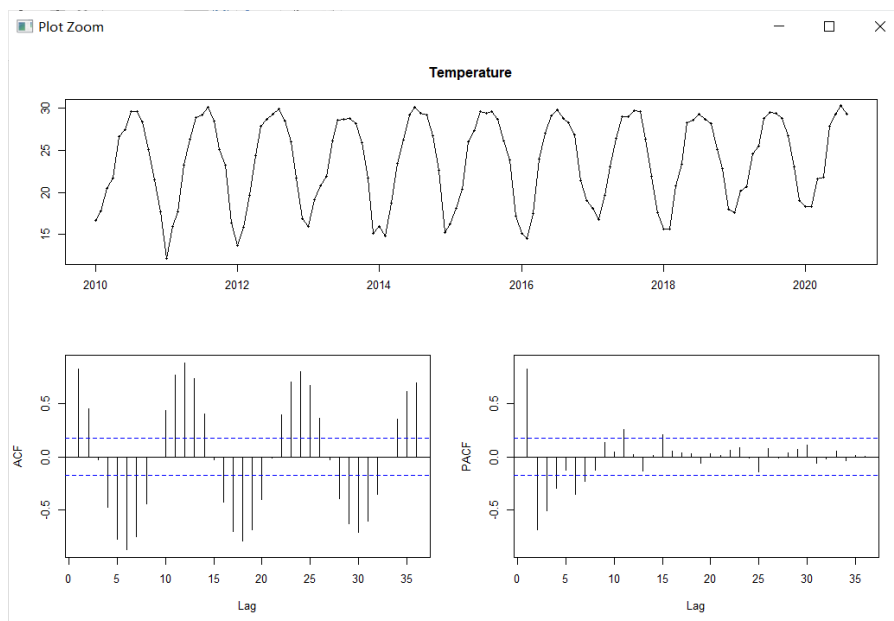
首先查看时间序列的组成部分，如下图所示，有明显的周期性，再检查误差分量的分布，发现符合高斯白噪声的分布。





2.3

首先查看原始数据的 ACF 和 PACF 分布情况，发现 ACF 具有周期性的趋势，而 PACF 渐近于平稳，因此首先考虑去除周期性的影响。

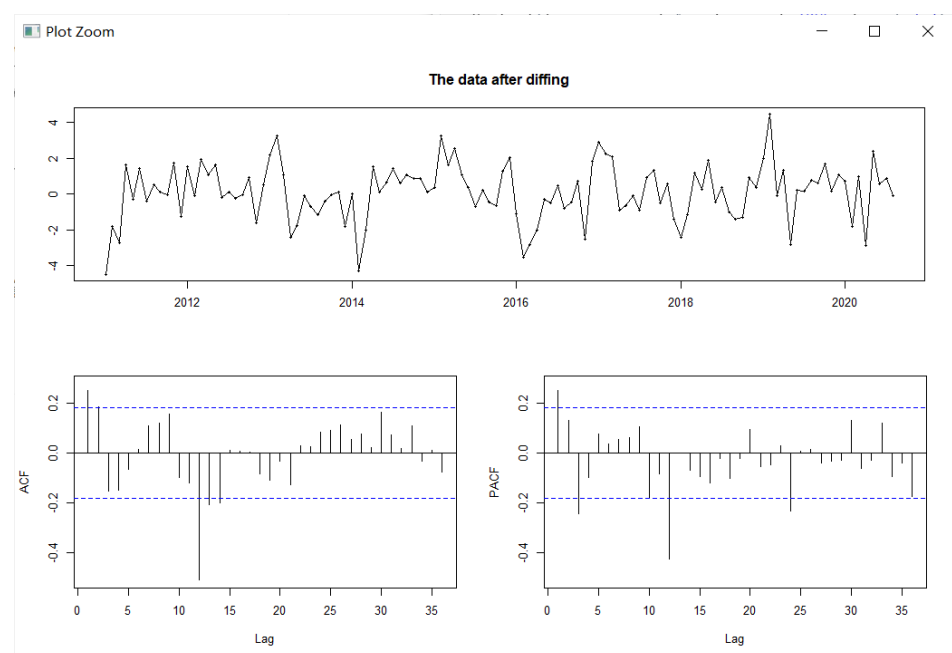
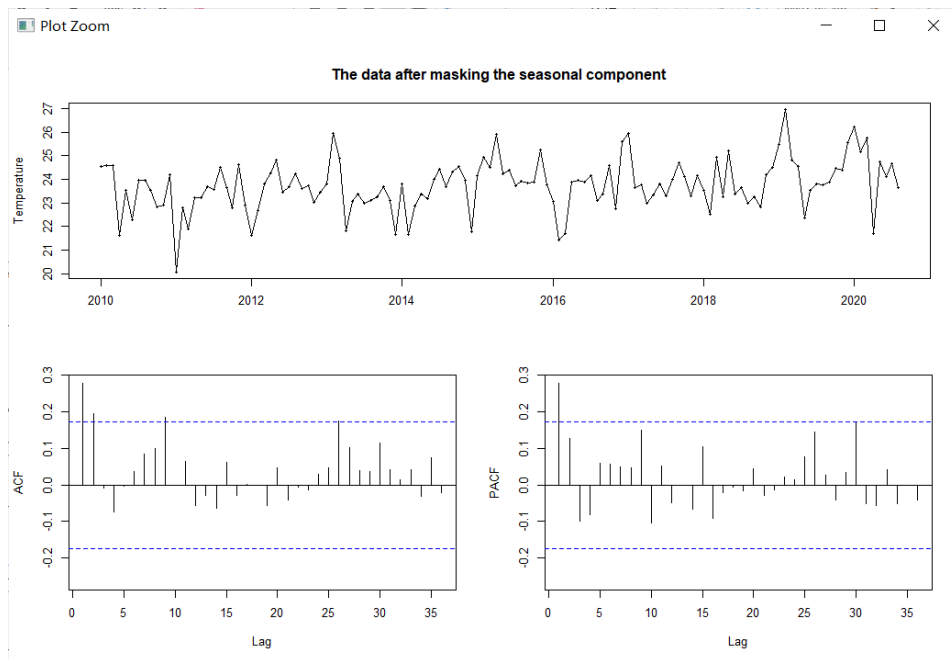


用两种进行了测试，如下图所示：

```
#去除周期性影响后的ACF和PACF（减去周期性的分量）
data_adjusted<-data_ts-temp_components$seasonal
plot(data_adjusted,main= "The data after masking the seasonal component")
tsdisplay(data_adjusted,main= "The data after masking the seasonal component",ylab="Temperature")

data_D1=diff(data_ts,12) #周期为12个月，进行差分（）差分法
tsdisplay(data_D1,main= "The data after differencing")
temp_components <- decompose(data_D1)
plot(temp_components) #展示差分后的时间序列组成部分，判断周期性影响是否被去除
```

各自的结果如下图所示：



以上的结果表面，再去除了周期性的影响后，数据具有平稳性，不需要进行非周期性的差分。同时由于具有周期性，考虑使用 $ARIMA(p,d,q)(P,D,Q)[Seasonal]$ 模型，由于只进行了周期性的差分，所以 $d=0$, $D=1$ ，而根据去除周期性因素后数据的 ACF 和 PACF 图可以看出，参数 P, Q 都为 1，为了进一步确定 p 和 q ，遍历 p, q 的取值，根据 AIC 最小来确定最优模型。手动获得的模型为：

```
> model

Call:
arima(x = data_ts, order = c(0, 0, 2), seasonal = list(order = c(1, 1, 1), period = 12))

Coefficients:
      ma1      ma2      sar1      sma1
  0.2333  0.2255 -0.0583 -0.7406
s.e.  0.0890  0.1091  0.1770  0.1783

sigma^2 estimated as 1.292: log likelihood = -184.81, aic = 379.61
```

再利用自动拟合函数得到自动拟合模型为：

```
> model_auto
Series: data_ts
ARIMA(0,0,2)(1,1,1)[12] with drift

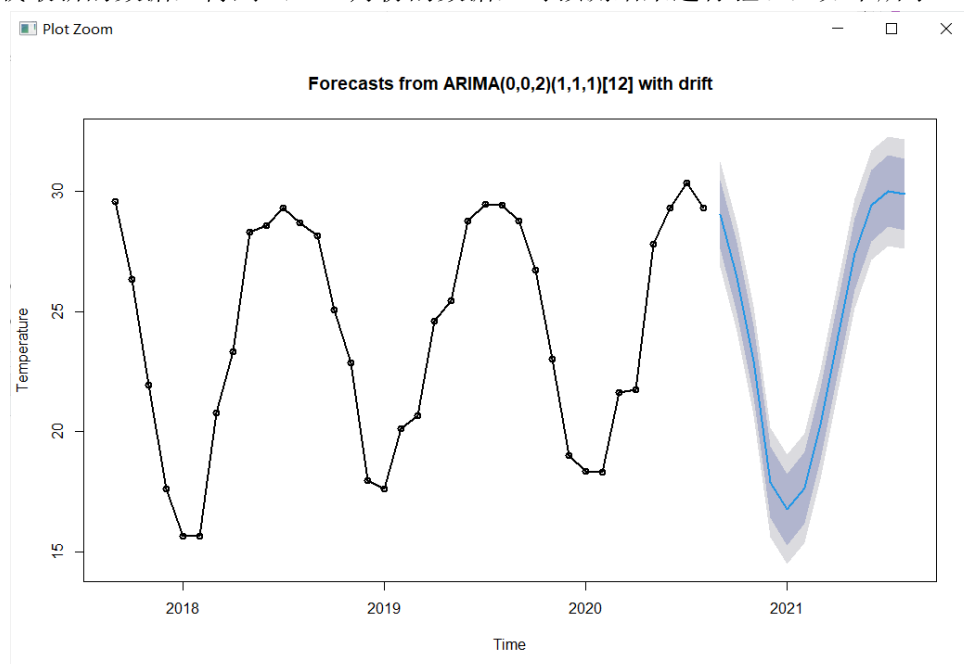
Coefficients:
      ma1      ma2      sar1      sma1      drift
      0.2090  0.1876 -0.0309 -0.8662  0.0085
s.e.  0.0902  0.1152  0.1684  0.2387  0.0037

sigma^2 estimated as 1.236: log likelihood=-182.72
AIC=377.43 AICc=378.2 BIC=393.95
```

再比较二者的 AIC 值，最终选取自动拟合的模型。

2.4

下载最新的数据，得到 9，10 月份的数据，对预测结果进行验证，如下所示：



```
> forecast(best_model, 2)
      Point Forecast   Lo 80   Hi 80   Lo 95   Hi 95
Sep 2020    29.04255  27.60715  30.47794  26.84730  31.23779
Oct 2020    26.40132  24.93491  27.86773  24.15864  28.64400
> dataobs
# A tibble: 11 x 3
# Groups:   year [1]
   year month mean_temp
<chr> <chr>    <dbl>
1 2020 01      18.3
2 2020 02      18.3
3 2020 03      21.6
4 2020 04      21.8
5 2020 05      27.8
6 2020 06      29.3
7 2020 07      30.3
8 2020 08      29.3
9 2020 09      28.8
10 2020 10      25.8
11 2020 11      24.3
```