BIOS 661 Project 1

Xinran Chu

10/07/2019

Background

The data was provided by UMD (Urban Ministry of Durham). Established in 1983, Urban Ministries of Durham works holistically to end homelessness and meet the emergency needs of very poor and hungry neighbors through its programs and in partnership with other nonprofits. UMD is a private, 501(c) 3 non-profit organization governed by a volunteer Board of Directors from the community. UMD offers food, shelter and a future to about 6,000 men, women and children annually.

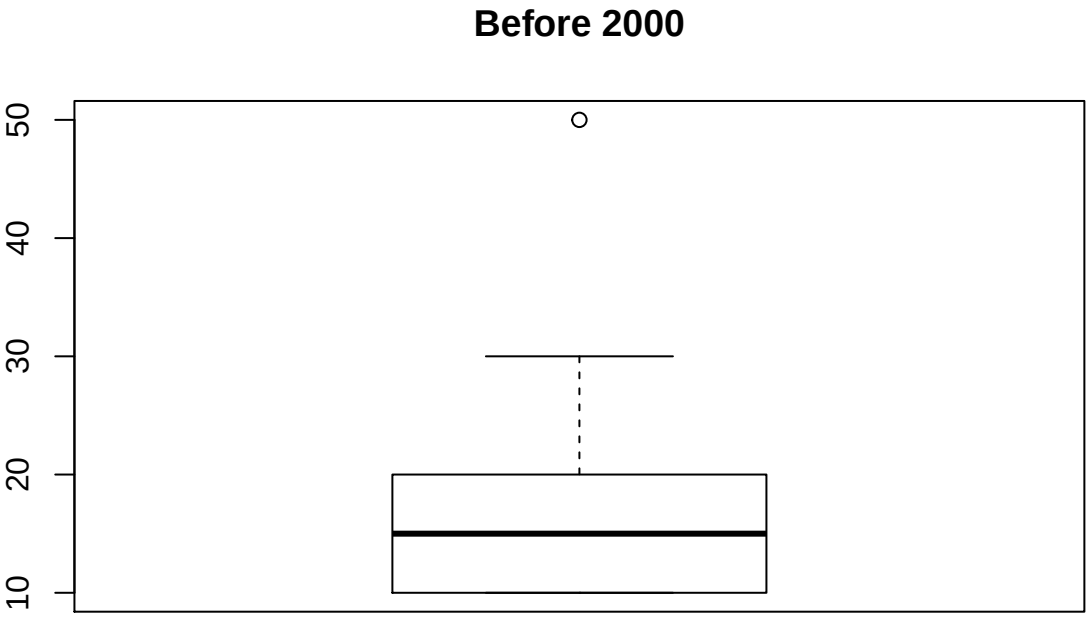
The data consists of 79838 observations and 13 variables (5 categorical variables and 8 numerical variables) from 1931 to 2019. However, since UMD was not established till 1983, I removed the observations with date prior to 1983. For this project, I used the variables 'Date' and 'Food Pounds'.

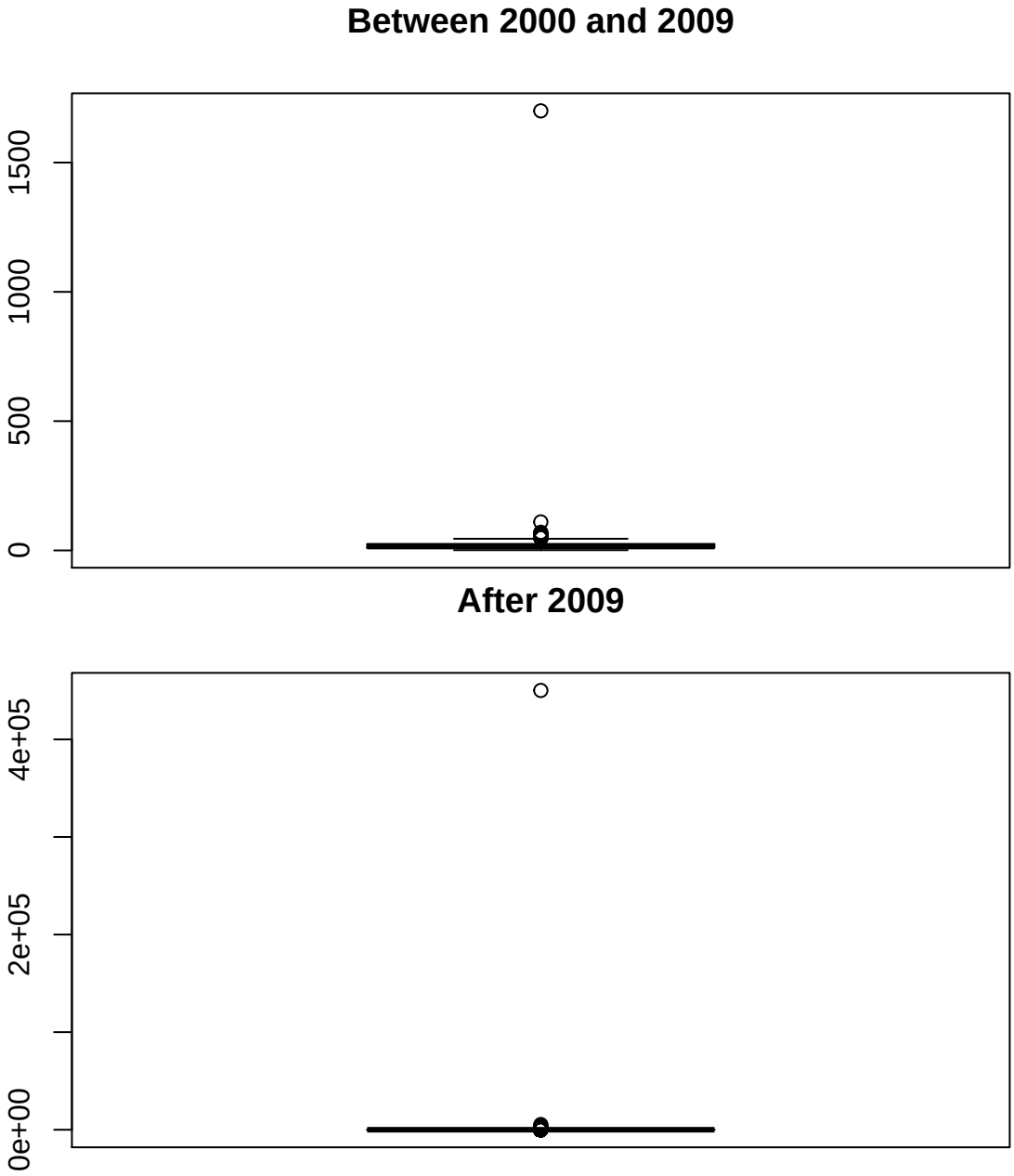
Research Question

My focus of this project is to look at the amount of food provided by UMD along with its trend in three seperate time periods: before 2000, between 2000 and 2010, and on and after 2010.

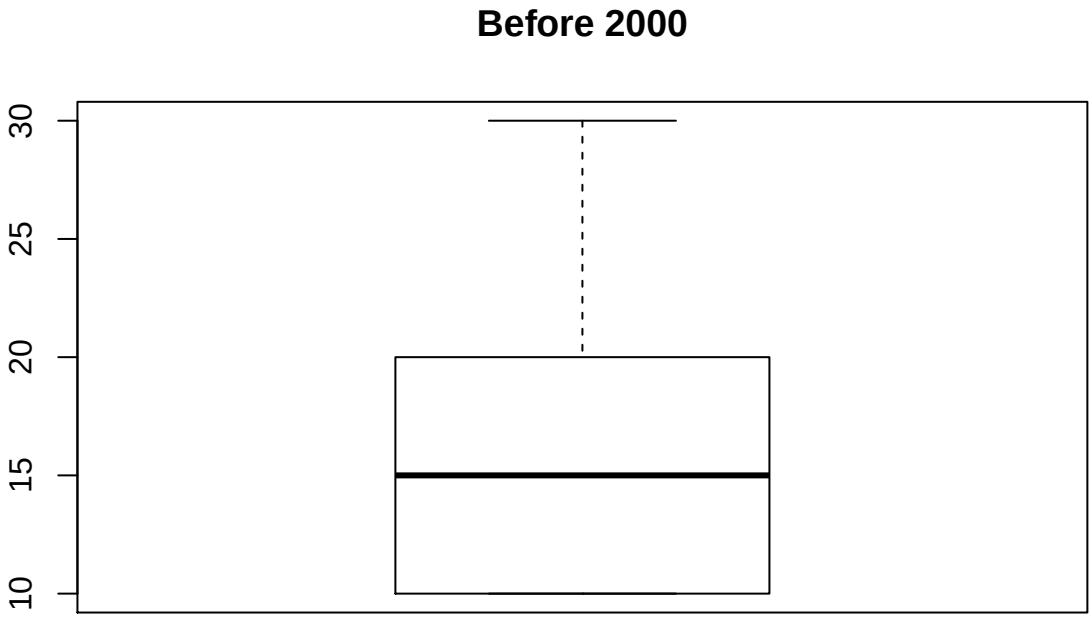
Remove outliers

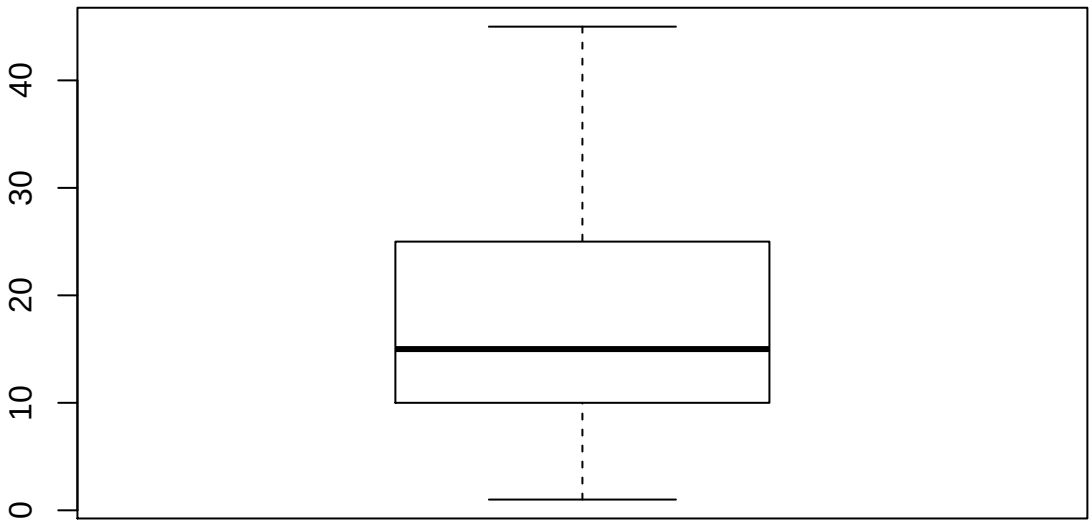
First check the boxplots to identify outliers

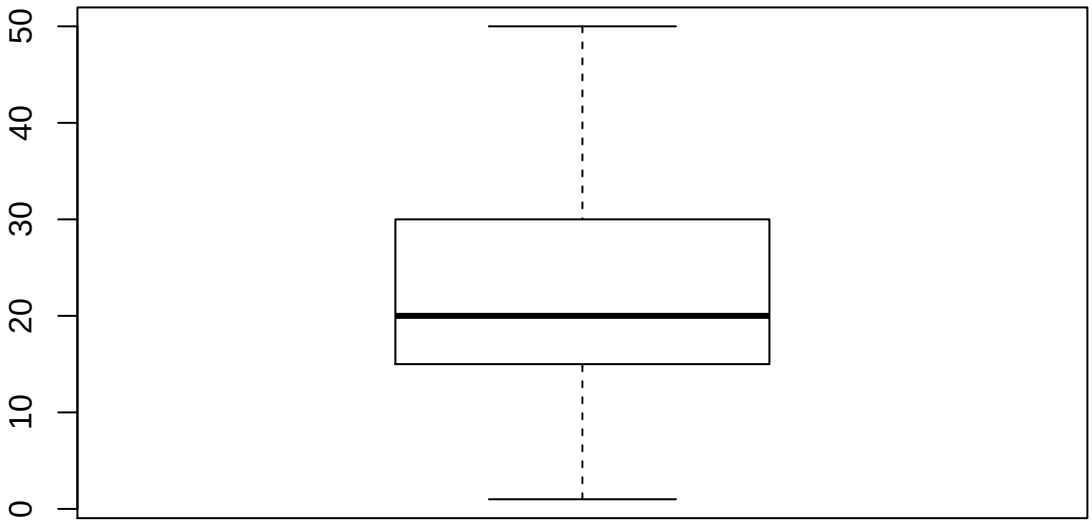



Between 2000 and 2009
0
500
1000
1500
After 2009
0e+00
2e+05
4e+05

After removing outliers, if we check now with boxplots, we notice that those pesky outliers are gone







Summary of the amount of food in pounds distributed in each period (after removing outliers).

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   10.00   10.00   15.00   16.11   20.00   30.00
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.00   10.00   15.00   17.66   25.00   45.00
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.00   15.00   20.00   24.51   30.00   50.00
```

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

```
## [1] 154428
```

```
## [1] 930893
```

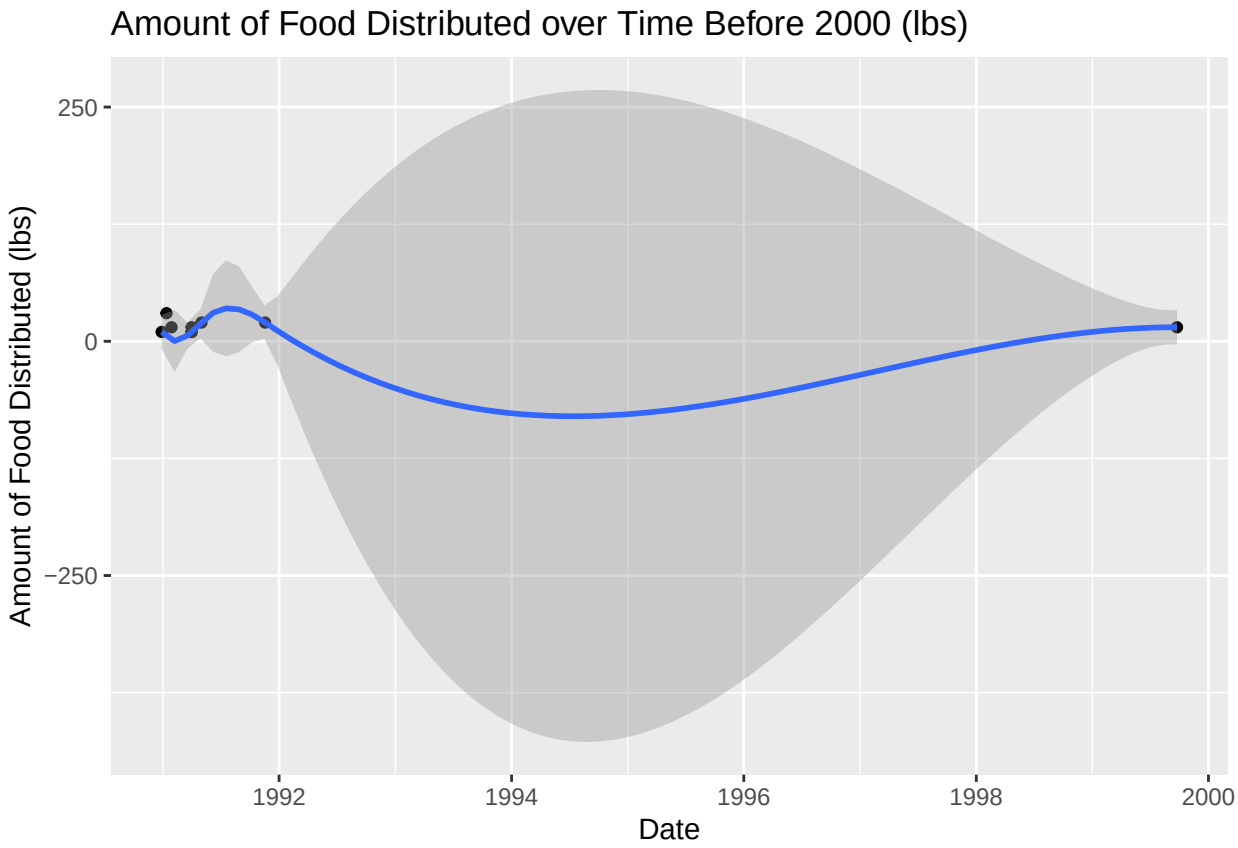
A total of 145 pounds of food were provided before 2000

A total of 154,428 pounds of food were provided between 2000 and 2009

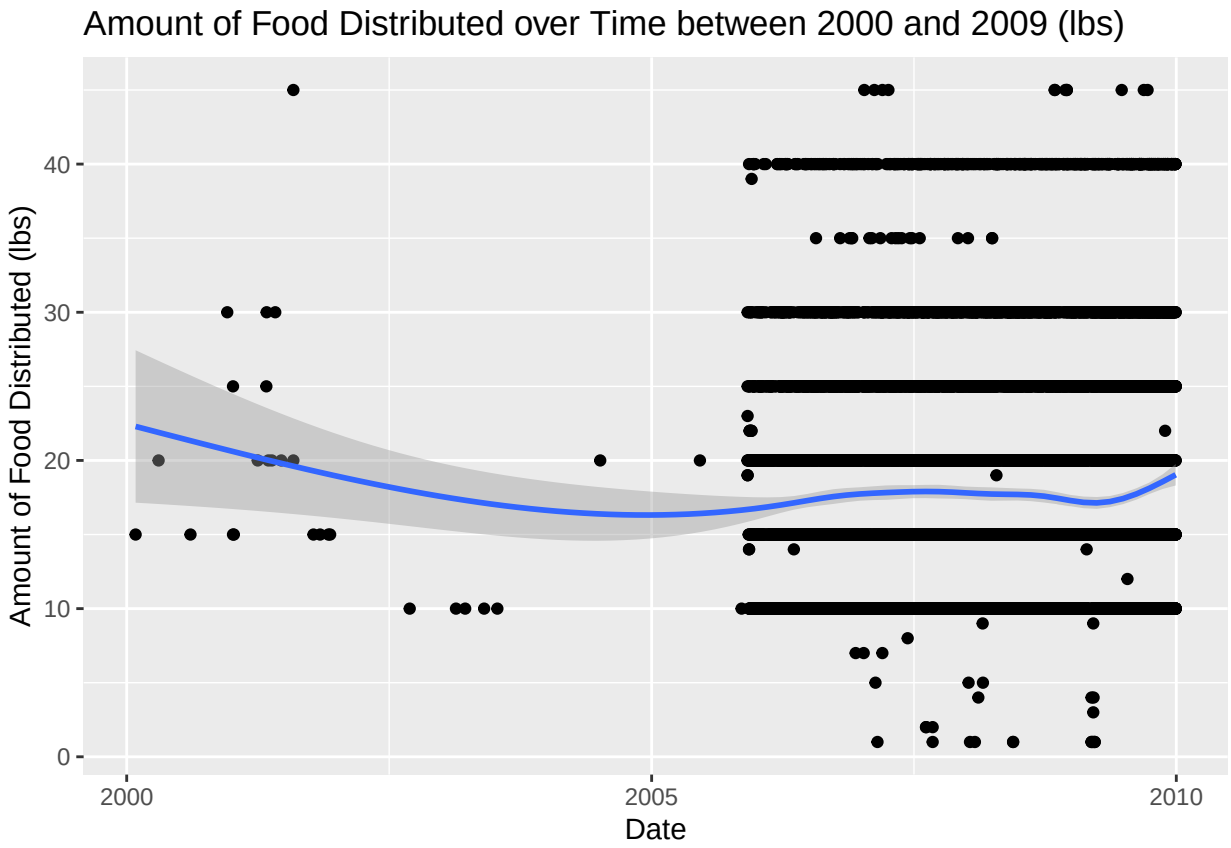
A total of 930,893 pounds of food were provided after 2009

Plot the trend for each period

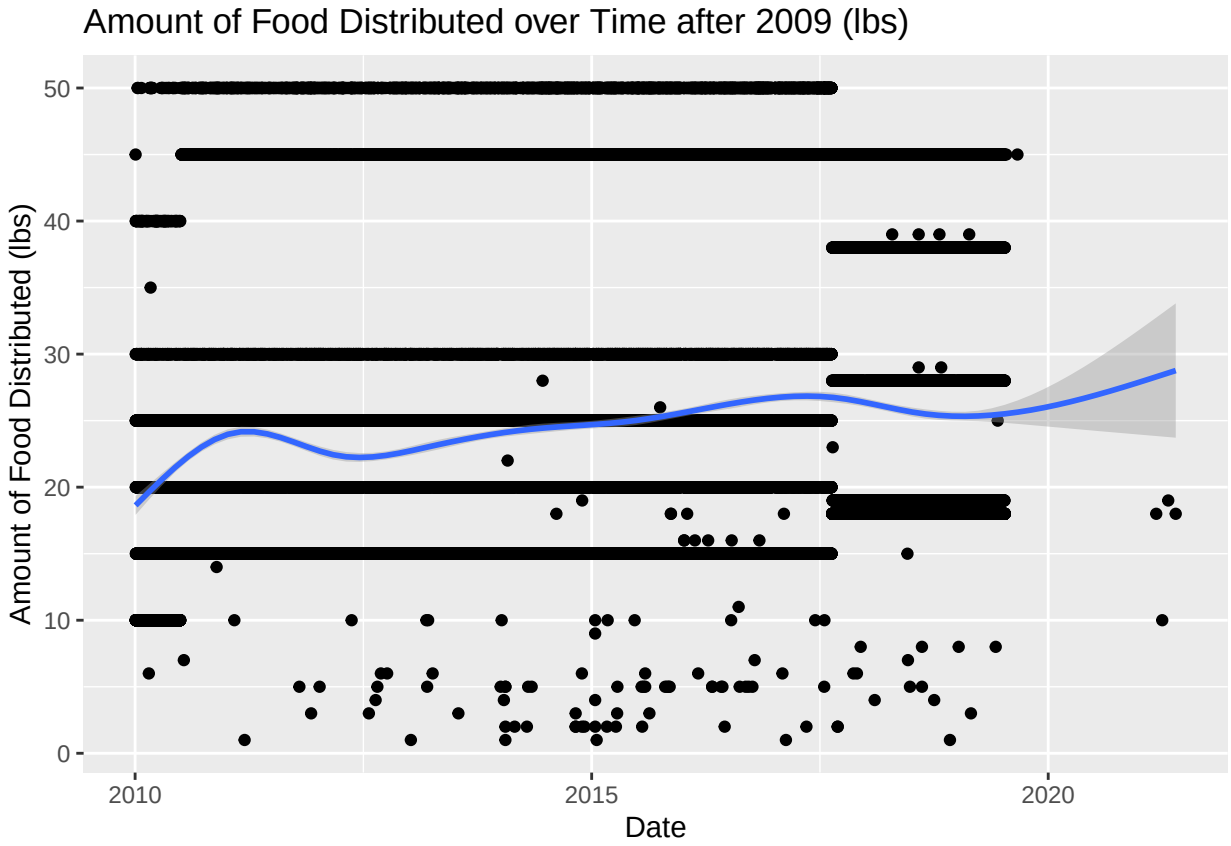
```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```




```
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```




```
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```



Before 2000, the amount of food distributed in pounds by UMD first went up till 1991, then went down and reached its minimum in between 1994 and 1995, and then started going up again.

Between 2000 and 2009, the amount of food distributed first went down, reaching its minimum in about 2005, and was faily steady afterwards.

After 2009, the amount of food distributed showed an overall increasing trend, with slight fluctuations.

From the numerical summaries and the plots, we observe that the total amount of food provided by UMD has been increasing in large magnitude. However, within each time period, there were fluctuations in the amount of food distributed by UMD. Perhaps in the future, we can conduct further analysis to identify the causes of these fluctuations.