Insulin_Dosage_Visualizations_April.R

Sonal

Fri Feb 15 16:43:12 2019

```r
#Diabetes Analysis and Visualizations.
#VISUALIZATION 1:

#Reading the data-01.csv file in DiabetesDataFrame
diabetesDataFrame <- read.csv(file =
"/Users/Sonal/Desktop/Multi_Analysis/Project/data-01.csv")

#Form a subset of Regular Insulin dose given to the patient.
#The code of Regular Insulin Dose is 33.
#To make the date format of data set acceptable in R:
diabetesDataFrame$Date <- as.Date(diabetesDataFrame$Date,format="%m/%d/%y")

#Our aim is to analyze the dosage of different insulins that are given to the
patient over the months from April to September.
#We will first plot the data for April for all the 3 types of insulins and
further during the project analyze the dosage for months from May to
September.

#To make a subset that has all the regular insulin dosage given to the
patient in the month of APRIL we will write the following condition.
regularinsulindose <- subset(diabetesDataFrame,Code == c(33) & Date >= "1991-
04-01" & Date<="1991-04-30")
#To create a visualization with date and the dosage given in april:
plot(regularinsulindose$Date,regularinsulindose$Value, xlab="April
Data",ylab="Insulin Dosage")
```

April Data
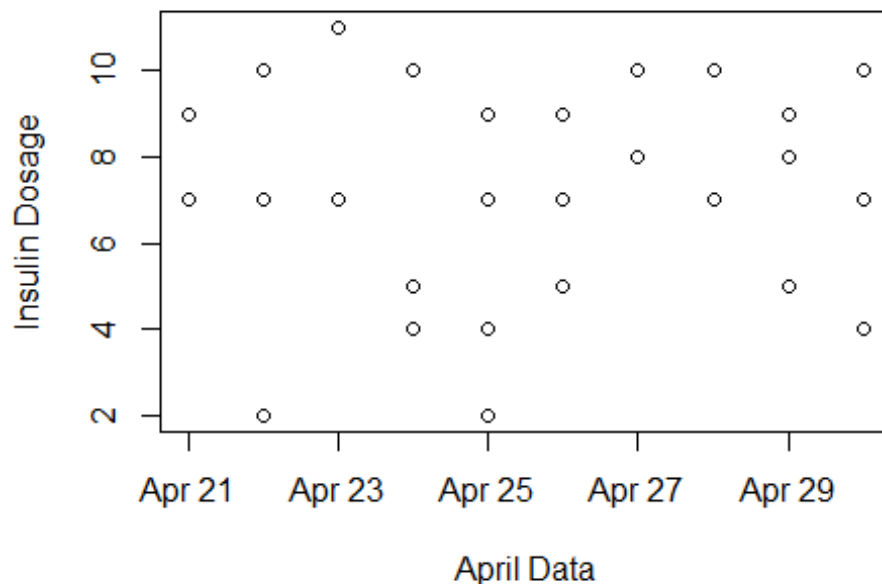
```
#From the visualization we can see that the insulin dosage provided to the
patient was the highest on April 23 with value over 10.


#VISUALIZATION 2:
#To make a subset that has all the NPH insulin dosage given to the patient in
the month of april we will write the following condition.
NPHinsulindose <- subset(diabetesDataFrame,Code == c(34) & Date >= "1991-04-
01" & Date<="1991-04-30")
#To create a visualization with date and the dosage given in april:
plot(NPHinsulindose$Date,NPHinsulindose$Value, xlab="April Data",ylab="NPH
Insulin Dosage")
#We can observe from the visualization that the dosage of NPH Insulin was
always 13 or 14.
#Between 13 and 14, the dosage was mostly 14.


#VISUALIZATION 3:
#To make a subset that has all the Ultralente insulin dosage given to the
patient in the month of april we will write the following condition.
Ultralenteinsulindose <- subset(diabetesDataFrame,Code == c(35) & Date >=
"1991-04-01" & Date<="1991-04-30")
head(Ultralenteinsulindose)

## [1] Date  Time  Code  Value
## <0 rows> (or 0-length row.names)
```

```
#We can observe that there is no data for Ultralente Insulin.
#Hence we can conclude that, Ultralente Insulin dosage was never provided to
the patient.

#VISUALIZATION 4:
#Let us now visualize how many times between the span of April to September
has the patient shown Hypoglycemic symptoms.
#Hypoglycemia is a condition when the blood sugar (Glucose) level in the body
decreases drastically.
#For this visualiation we will use ggplot package and plot a histogram for
the same.
library(ggplot2)
```
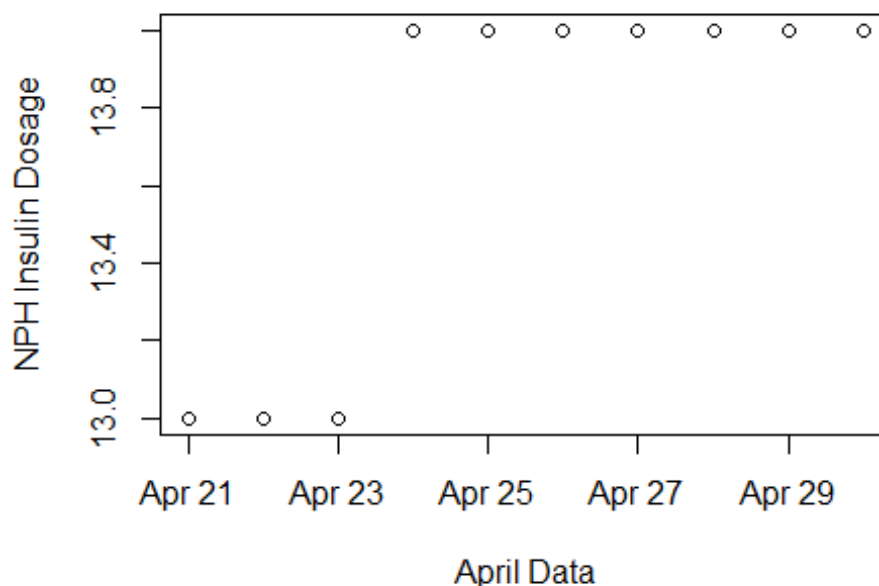


```
Hypoglycemia <- subset(diabetesDataFrame,Code == c(65))
head(Ultralenteinsulindose)

## [1] Date  Time  Code  Value
## <0 rows> (or 0-length row.names)
```

```
#We can observe that there is no data for Hypoglycemic Symptoms.
#Hence we can conclude that, the patient never showed any Hypoglycemic
Symptoms.

#VISUALIZATION 5:
#Let us now visualize
UnspecifiedGlucoseMeasurement <- subset(diabetesDataFrame,Code == c(48))
#The range of blood sugar level has mostly been between 70-170.
```
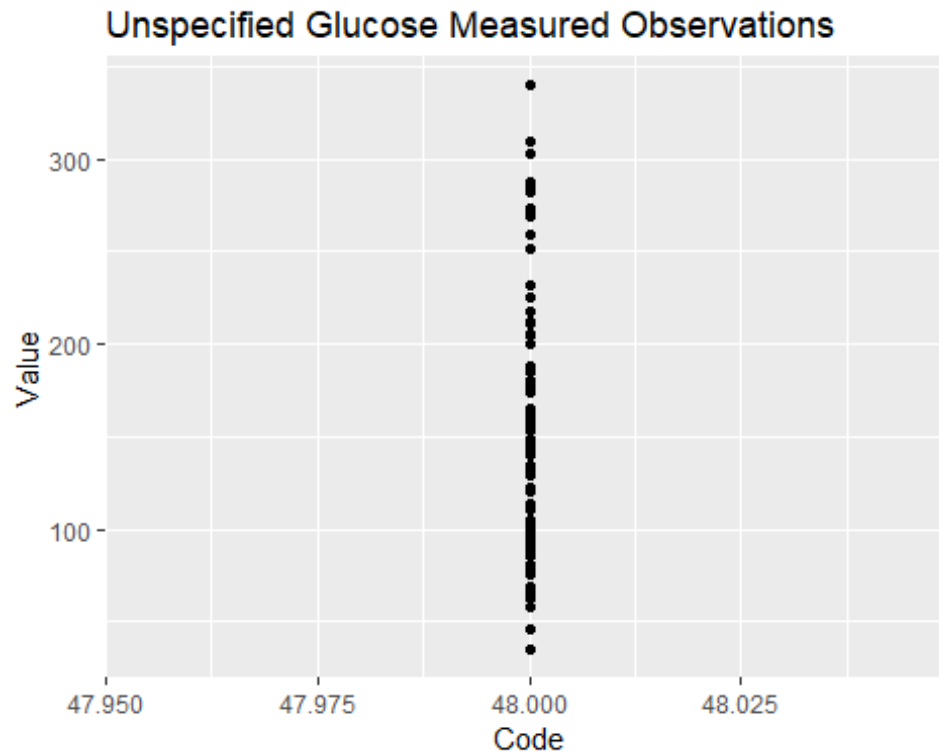
```
#The highest observed blood sugar level is close to 350.
ggplot(UnspecifiedGlucoseMeasurement,aes(x=Code,y=Value)) + geom_point() +
  labs(x = "Code",y = "Value", title = "Unspecified Glucose Measured
Observations")
```


Unspecified Glucose Measured Observations

```
#VISUALIZATION 6:
ggplot(UnspecifiedGlucoseMeasurement,aes(x=Code,y=Value)) + geom_boxplot() +
  labs(x = "Code",y = "Value", title = "Unspecified Glucose Measured
Observations")
```

Unspecified Glucose Measured Observations
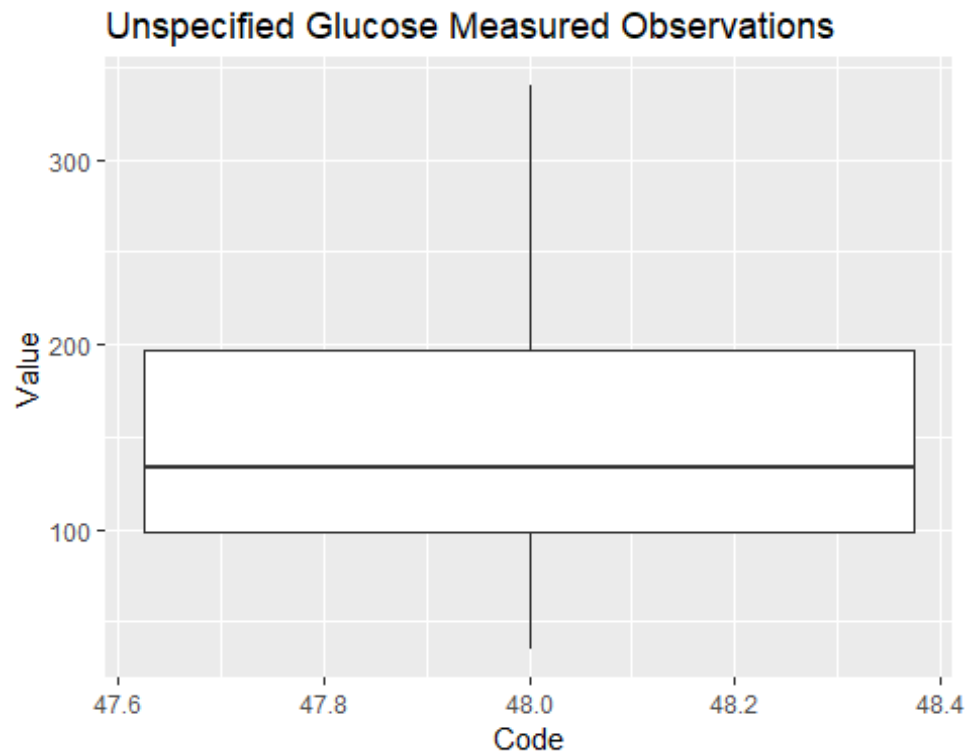
#By looking at the data, we can see that these random blood checks have been done on the patient at night.
#Mostly between the time of 10pm to 11pm.
#From this plot we can conclude that whenever these random checks have happened, the blood sugar level has been more the mean value.
#The highest number of values are observed between 130 to 200.