Simple document

Problem 1

Load the moderndive library, and use the following code to load the early_january_weather dataset:

```
library(moderndive)
data("early_january_weather")
```

The following code loads the early_january_weather dataset and includes a description

```
data_set <- data("early_january_weather")
```

Describing the Data Set

```
mean(early_january_weather$temp)
```

```
## [1] 39.58212
```

```
range(early_january_weather$temp)
```

```
## [1] 24.08 57.92
```

```
median(early_january_weather$temp)
```

```
## [1] 39.02
```

```
range(early_january_weather$humid)
```

```
## [1]  32.86 100.00
```

```
mean(early_january_weather$humid)
```

```
## [1] 65.4767
```

```
range(early_january_weather$wind_speed)
```

```
## [1]  0.00000 24.16638
```

This data set contains 358 observations of 15 variables, including: origin, year, month, day, hour, temperature, dew, humidity, wind direction, wind speed, wind gust,precipitation, pressure, visibility. Values for temperature range from 24.08 to 57.92, the mean of temperature is 39.5821, median is 39.02. Values for humidity range from 32.86 to 100, mean humidity is 65.4767. Values for wind speee range from 0 to 24.14.

Describe the size of the dataset (using nrow and ncol)

```r
nrow(early_january_weather)
```

```
## [1] 358
```

```r
ncol(early_january_weather)
```

```
## [1] 15
```

Describe the mean temperature

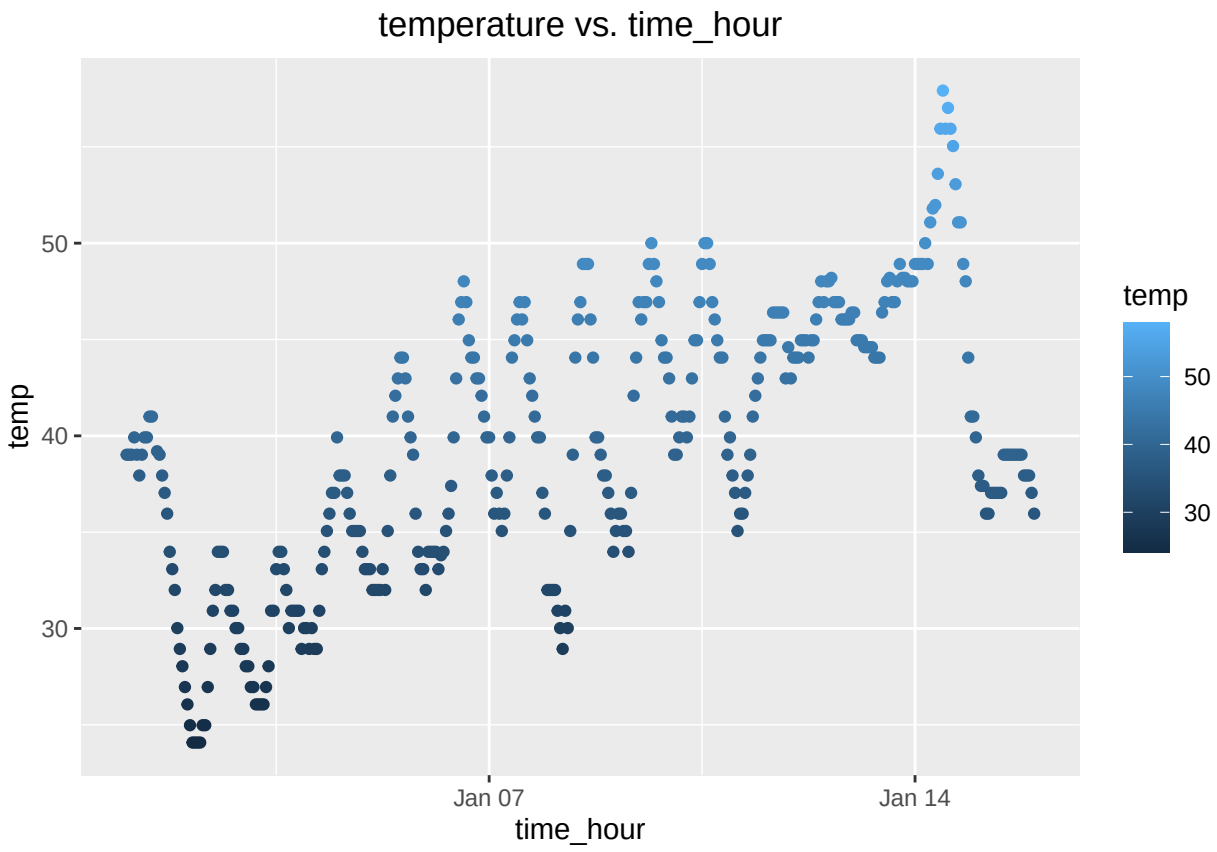
```r
mean(early_january_weather$temp)
```

```
## [1] 39.58212
```

Make a scatterplot of temp (y) vs time_hour (x)

```r
library(ggplot2)

ggplot(early_january_weather,aes(x=time_hour,y=temp, color=temp))+geom_point()+ggtitle("temperature vs.
```



```
ggsave("scatterplot_temp vs. time_hour.pdf")
```

```
## Saving 6.5 x 4.5 in image
```

Describe this plot: Temperatures are changing all the time, and in general, early January is showing an upward trend. However, we can also clearly see that the temperature is falling and then rising on a daily basis, constantly repeating itself. In addition, the day at the beginning of the month and the day in the middle of the month the temperature drops significantly.

#Problem 2

```
typeof(early_january_weather)
```

```
## [1] "list"
```

```
data_set_vector <- as.matrix(data_set)
nrow(early_january_weather)
```

```
## [1] 358
```

```
ncol(early_january_weather)
```

```
## [1] 15
```

```
mean(early_january_weather[["temp"]]) # or
```

```
## [1] 39.58212
```

```
mean(early_january_weather$temp)
```

```
## [1] 39.58212
```