Reproducible Research - Assessment 1

José Miguel Reséndiz

Tuesday, February 10, 2015

INTRODUCCTION

It is now possible to collect a large amount of data about personal movement using activity monitoring devices such as a Fitbit, Nike Fuelband, or Jawbone Up. These type of devices are part of the "quantified self" movement - a group of enthusiasts who take measurements about themselves regularly to improve their health, to find patterns in their behavior, or because they are tech geeks. But these data remain under-utilized both because the raw data are hard to obtain and there is a lack of statistical methods and software for processing and interpreting the data.

Data

The data for this assignment can be downloaded from the course web site:

Dataset: Activity monitoring data [52K]

The variables included in this dataset are:

steps: Number of steps taking in a 5-minute interval (missing values are coded as NA)

date: The date on which the measurement was taken in YYYY-MM-DD format

· **interval:** Identifier for the 5-minute interval in which measurement was taken

The dataset is stored in a comma-separated-value (CSV) file and there are a total of 17,568 observations in this dataset.

GETING & CLEANING DATA

1.- LOAD DATA

Load Data

```
dataPerMov<-read.csv("activity.csv")
```

Raw Data

```
##   steps       date interval
## 1    NA 2012-10-01        0
## 2    NA 2012-10-01        5
## 3    NA 2012-10-01       10
## 4    NA 2012-10-01       15
## 5    NA 2012-10-01       20
## 6    NA 2012-10-01       25
```

Structure

```
## 'data.frame':    17568 obs. of  3 variables:
##  $ steps   : int  NA NA NA NA NA NA NA NA NA NA ...
##  $ date    : Factor w/ 61 levels "2012-10-01","2012-10-02",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ interval: int  0 5 10 15 20 25 30 35 40 45 ...
```

2.- Process/transform the data

We select only the rows with values different of NA, and tranform field 'date' to POSIXlt

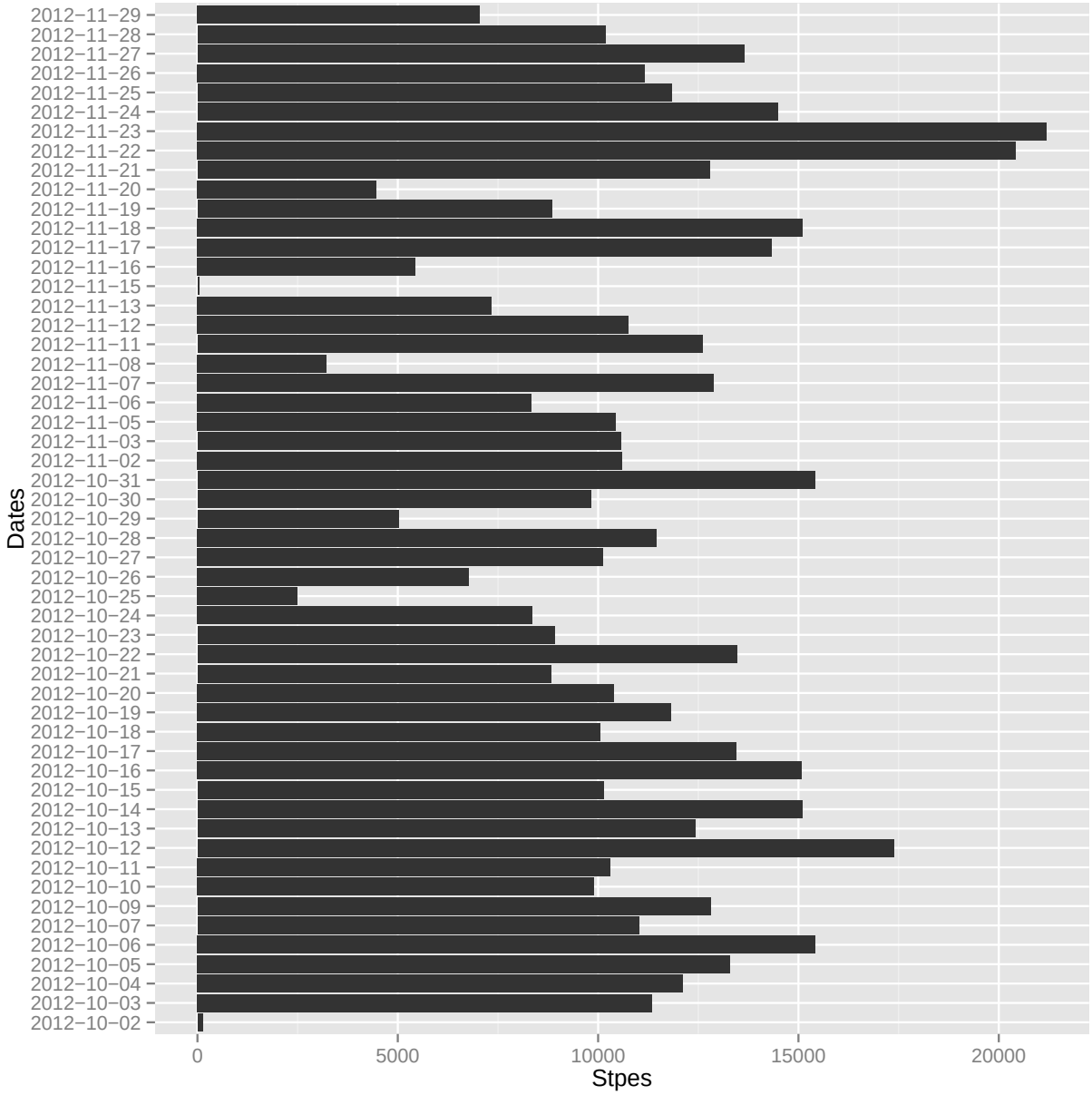
```
dataPerMov<-na.omit(dataPerMov)
dataPerMov$date <- as.POSIXlt(dataPerMov$date)
```

Output

```
##     steps       date interval
## 289     0 2012-10-02        0
## 290     0 2012-10-02        5
## 291     0 2012-10-02       10
## 292     0 2012-10-02       15
## 293     0 2012-10-02       20
## 294     0 2012-10-02       25
```

EXPLORATORY DATA ANALYSI

1.- What is mean total number of steps taken per day?



date

x

1

2012-10-02

0.44

2

2012-10-03

39.42

3

2012-10-04

42.07

4

2012-10-05

46.16

5

2012-10-06

53.54

6

2012-10-07

38.25

7

2012-10-09

44.48

8

2012-10-10

34.38

9

2012-10-11

35.78

10

2012-10-12

60.35

11

2012-10-13

43.15

12

2012-10-14

52.42

13

2012-10-15

35.20

14

2012-10-16

52.38

15

2012-10-17

46.71

16

2012-10-18

34.92

17

2012-10-19

41.07

18

2012-10-20

36.09

19

2012-10-21

30.63

20

2012-10-22

46.74

21

2012-10-23

30.97

22

2012-10-24

29.01

23

2012-10-25

8.65

24

2012-10-26

23.53

25

2012-10-27

35.14

26

2012-10-28

39.78

27

2012-10-29

17.42

28

2012-10-30

34.09

29

2012-10-31

53.52

30

2012-11-02

36.81

31

2012-11-03

36.70

32

2012-11-05

36.25

33

2012-11-06

28.94

34

2012-11-07

44.73

35

2012-11-08

11.18

36

2012-11-11

43.78

37

2012-11-12

37.38

38

2012-11-13

25.47

39

2012-11-15

0.14

40

2012-11-16

18.89

41

2012-11-17

49.79

42

2012-11-18

52.47

43

2012-11-19

30.70

44

2012-11-20

15.53

45

2012-11-21

44.40

46

2012-11-22

70.93

47

2012-11-23

73.59

48

2012-11-24

50.27

49

2012-11-25

41.09

50

2012-11-26

38.76

51

2012-11-27

47.38

52

2012-11-28

35.36

53

2012-11-29

24.47