AW_HealthDataPrep

Bo Simango

January 30, 2019

```r
# Created by Bo Simango
# M.Sc. candidate
# Supervisor: Dr. Daniel Fuller
# Memorial University of Newfoundland
# Date: 30 January 2019
# Function: This code appends the Health Data Collector app data by participant ID (107-111) and
 plots the HR, Steps and Calories vs. Time using two R packages ggplot2 and facet_wrap
# Data source: Health Data Collector App - AW data
# Output: Appended CSVs, ggplot facet wrap and mice imputation

# First install all libraries used throughout this exercise
```

```r
library(knitr)
```

```
## Warning: package 'knitr' was built under R version 3.5.2
```

```r
library(haven)
```

```
## Warning: package 'haven' was built under R version 3.5.2
```

```r
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.5.2
```

```r
library(plyr)
```

```
## Warning: package 'plyr' was built under R version 3.5.2
```

```r
library(readr)
```

```
## Warning: package 'readr' was built under R version 3.5.2
```

```r
# Set working directory
setwd("C:/Thesis")
```

```
# Read Health Data Collector App data
a_107 <- read.csv("applewatch_data107.csv")
a_108 <- read.csv("applewatch_data108.csv")
a_109 <- read.csv("applewatch_data109.csv")
a_110 <- read.csv("applewatch_data110.csv")
a_111 <- read.csv("applewatch_data111.csv")
```

```
# Summary of statistical descriptives of each participant ID
summary(a_107)
```

```
##             DateTime                        Username        DeviceName
##   13/12/2018 13:13: 1    acceltrial2019@walkabilly.ca:76    AppleWatch:76
##   13/12/2018 13:14: 1
##   13/12/2018 13:15: 1
##   13/12/2018 13:16: 1
##   13/12/2018 13:17: 1
##   13/12/2018 13:18: 1
##   (Other)        :70
##      Heart           Calories          Steps          Distance
##   Min.   : 55.00   Min.   : 0.5125   Min.   : 14.0   Min.   :0.00924
##   1st Qu.: 76.25   1st Qu.: 0.6100   1st Qu.: 89.5   1st Qu.:0.05337
##   Median : 94.00   Median : 0.6575   Median : 99.5   Median :0.06657
##   Mean   :104.51   Mean   : 1.7261   Mean   :163.9   Mean   :0.09256
##   3rd Qu.:146.88   3rd Qu.: 0.7560   3rd Qu.:153.5   3rd Qu.:0.08266
##   Max.   :156.00   Max.   :21.6190   Max.   :803.0   Max.   :0.60797
##   NA's   :50       NA's   :40        NA's   :29      NA's   :26
##      Part_ID
##   Min.   :107
##   1st Qu.:107
##   Median :107
##   Mean   :107
##   3rd Qu.:107
##   Max.   :107
##
```

```
summary(a_108)
```

```
##                 DateTime                        Username          DeviceName
##   07/01/2019 10:30: 1    acceltrial2019@walkabilly.ca:71    AppleWatch:71
##   07/01/2019 10:31: 1
##   07/01/2019 10:32: 1
##   07/01/2019 10:33: 1
##   07/01/2019 10:34: 1
##   07/01/2019 10:35: 1
##   (Other)           :65
##       Heart            Calories            Steps            Distance
##   Min.   : 63.0   Min.   :0.4593   Min.   : 440.5   Min.   :0.02826
##   1st Qu.: 74.0   1st Qu.:0.6115   1st Qu.: 681.4   1st Qu.:0.05214
##   Median : 87.0   Median :0.7385   Median : 880.5   Median :0.06977
##   Mean   :121.4   Mean   :0.7040   Mean   : 892.2   Mean   :0.18025
##   3rd Qu.:183.8   3rd Qu.:0.7560   3rd Qu.:1184.2   3rd Qu.:0.32751
##   Max.   :187.0   Max.   :1.3805   Max.   :1251.5   Max.   :0.72492
##   NA's   :54      NA's   :5        NA's   :65       NA's   :45
##      Part_ID
##   Min.   :108
##   1st Qu.:108
##   Median :108
##   Mean   :108
##   3rd Qu.:108
##   Max.   :108
##
```

```
summary(a_109)
```

```
##                 DateTime                         Username          DeviceName
##   09/01/2019 10:21: 1    acceltrial2019@walkabilly.ca:104    AppleWatch:104
##   09/01/2019 10:22: 1
##   09/01/2019 10:23: 1
##   09/01/2019 10:24: 1
##   09/01/2019 10:25: 1
##   09/01/2019 10:26: 1
##   (Other)           :98
##       Heart            Calories            Steps            Distance
##   Min.   : 62.00   Min.   : 0.5800   Min.   :   10.0   Min.   :0.00721
##   1st Qu.: 68.45   1st Qu.: 0.6303   1st Qu.: 140.8   1st Qu.:0.04720
##   Median : 82.50   Median : 0.7465   Median : 476.5   Median :0.05634
##   Mean   : 86.09   Mean   : 3.8991   Mean   : 546.7   Mean   :0.14151
##   3rd Qu.: 88.00   3rd Qu.: 0.7560   3rd Qu.: 877.2   3rd Qu.:0.16813
##   Max.   :145.08   Max.   :21.0620   Max.   :1214.0   Max.   :0.76997
##   NA's   :70       NA's   :75        NA's   :90       NA's   :68
##      Part_ID
##   Min.   :109
##   1st Qu.:109
##   Median :109
##   Mean   :109
##   3rd Qu.:109
##   Max.   :109
##
```

```
summary(a_110)
```

```
##                DateTime                        Username        DeviceName
##  11/01/2019 11:37: 1    acceltrial2019@walkabilly.ca:91    AppleWatch:91
##  11/01/2019 11:38: 1
##  11/01/2019 11:39: 1
##  11/01/2019 11:40: 1
##  11/01/2019 11:41: 1
##  11/01/2019 11:42: 1
##  (Other)         :85
##      Heart           Calories           Steps           Distance
##  Min.   : 51.00   Min.   : 0.4240   Min.   : 415.5   Min.   :0.02541
##  1st Qu.: 56.75   1st Qu.: 0.6125   1st Qu.: 606.0   1st Qu.:0.04282
##  Median : 79.00   Median : 0.7292   Median : 764.0   Median :0.04672
##  Mean   : 87.30   Mean   : 0.9461   Mean   : 830.1   Mean   :0.09498
##  3rd Qu.:124.56   3rd Qu.: 0.7560   3rd Qu.: 985.2   3rd Qu.:0.07715
##  Max.   :128.91   Max.   :19.8390   Max.   :1444.0   Max.   :0.68262
##  NA's   :69       NA's   :19        NA's   :79       NA's   :39
##     Part_ID
##  Min.   :110
##  1st Qu.:110
##  Median :110
##  Mean   :110
##  3rd Qu.:110
##  Max.   :110
##
```

```
summary(a_111)
```

```
##                   DateTime                          Username          DeviceName
##   16/01/2019 10:56: 1    acceltrial2019@walkabilly.ca:72    AppleWatch:72
##   16/01/2019 10:57: 1
##   16/01/2019 10:58: 1
##   16/01/2019 10:59: 1
##   16/01/2019 11:00: 1
##   16/01/2019 11:01: 1
##   (Other)        :66
##      Heart            Calories           Steps             Distance
##   Min.   : 56.00   Min.   : 0.5100   Min.   :  16.0   Min.   :0.01232
##   1st Qu.: 77.75   1st Qu.: 0.6100   1st Qu.: 529.8   1st Qu.:0.05264
##   Median : 96.50   Median : 0.6523   Median : 721.0   Median :0.09616
##   Mean   :109.97   Mean   : 0.8685   Mean   : 842.5   Mean   :0.16137
##   3rd Qu.:159.57   3rd Qu.: 0.7362   3rd Qu.:1220.2   3rd Qu.:0.10104
##   Max.   :171.33   Max.   :10.4285   Max.   :1660.5   Max.   :0.95876
##   NA's   :41       NA's   :25        NA's   :65       NA's   :35
##      Part_ID
##   Min.   :111
##   1st Qu.:111
##   Median :111
##   Mean   :111
##   3rd Qu.:111
##   Max.   :111
##
```

```
# Append all participant ID CSVs
applewatch_all <- rbind (a_107, a_108, a_109, a_110, a_111)
```

```
# Summary of statistical descriptives of each participant ID
summary(applewatch_all)
```
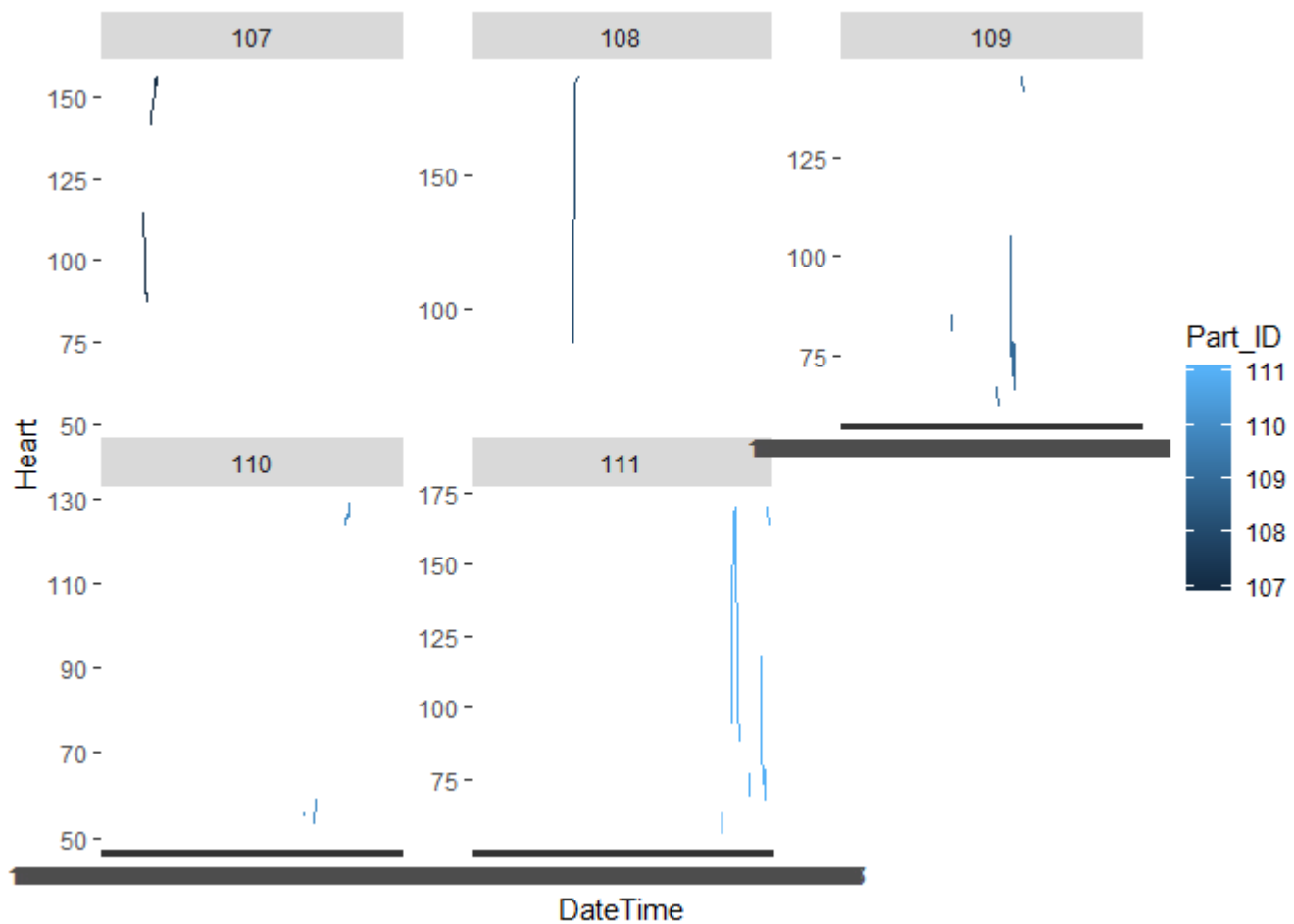
```
##                DateTime                              Username
##   13/12/2018 13:13:  1     acceltrial2019@walkabilly.ca:414
##   13/12/2018 13:14:  1
##   13/12/2018 13:15:  1
##   13/12/2018 13:16:  1
##   13/12/2018 13:17:  1
##   13/12/2018 13:18:  1
##   (Other)         :408
##        DeviceName        Heart             Calories            Steps
##   AppleWatch:414   Min.   : 51.00   Min.   : 0.424   Min.   :   10.00
##                    1st Qu.: 70.00   1st Qu.: 0.610   1st Qu.:   92.75
##                    Median : 86.85   Median : 0.732   Median :  156.00
##                    Mean   :100.29   Mean   : 1.322   Mean   :  425.23
##                    3rd Qu.:125.91   3rd Qu.: 0.756   3rd Qu.:  701.38
##                    Max.   :187.00   Max.   :21.619   Max.   : 1660.50
##                    NA's   :284      NA's   :164      NA's   :  328
##       Distance          Part_ID
##   Min.   :0.00721   Min.   :107
##   1st Qu.:0.04749   1st Qu.:108
##   Median :0.06412   Median :109
##   Mean   :0.12596   Mean   :109
##   3rd Qu.:0.09616   3rd Qu.:110
##   Max.   :0.95876   Max.   :111
##   NA's   :213
```

```
#Scatterplot of Heart rate vs. Time via Facet Wrap
  ggplot(applewatch_all, aes(x = DateTime, y = Heart, colour=Part_ID, group=Part_ID)) +
    geom_line() +
  facet_wrap(~Part_ID, scales ="free_y")
```
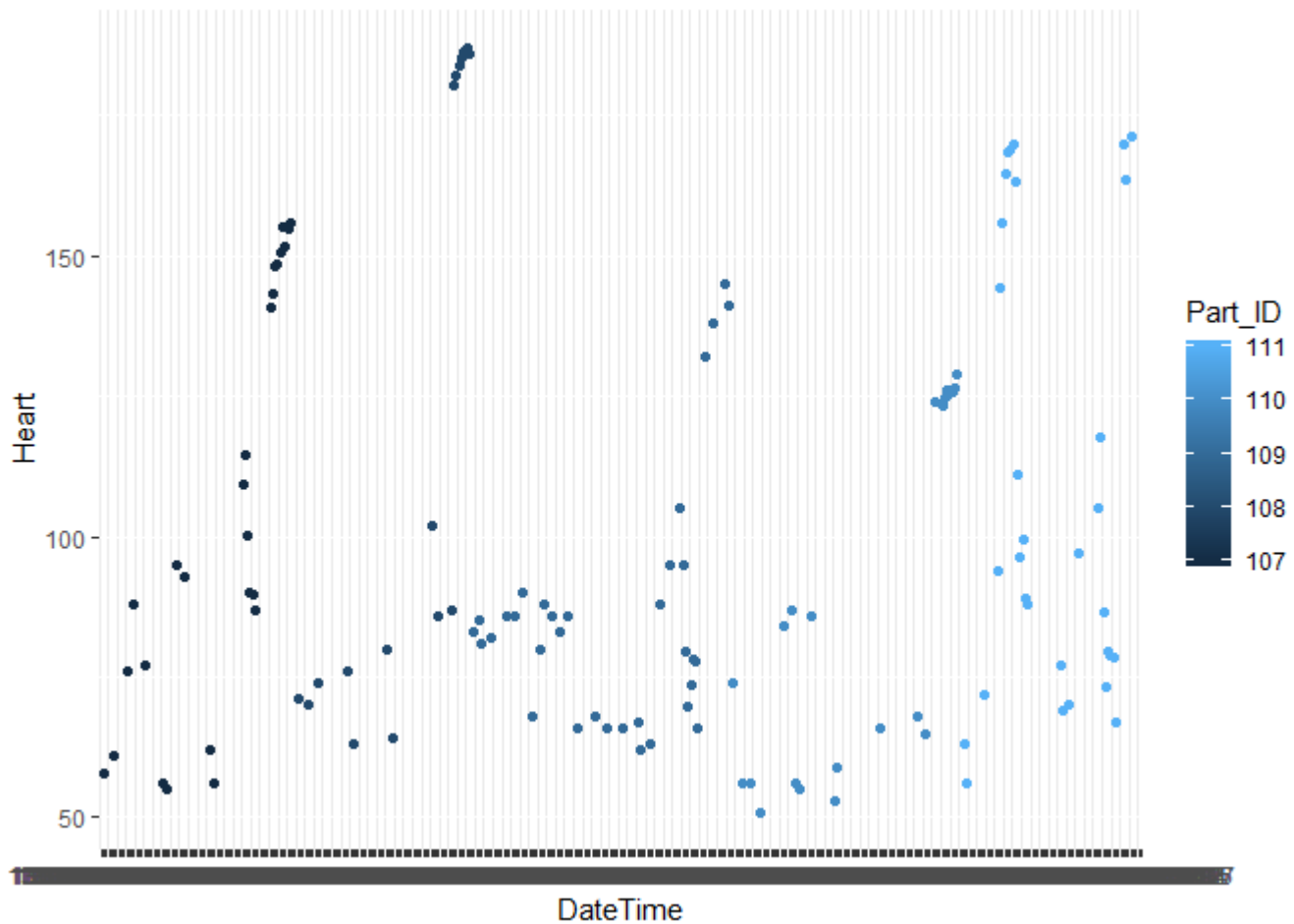
```
## Warning: Removed 11 rows containing missing values (geom_path).
```

```
#Scatterplot of Heart rate vs. Time, Condensed Scatter Plot
ggplot(applewatch_all, aes(x = DateTime, y = Heart, colour=Part_ID, group=Part_ID)) +
  geom_point()
```

```
## Warning: Removed 284 rows containing missing values (geom_point).
```
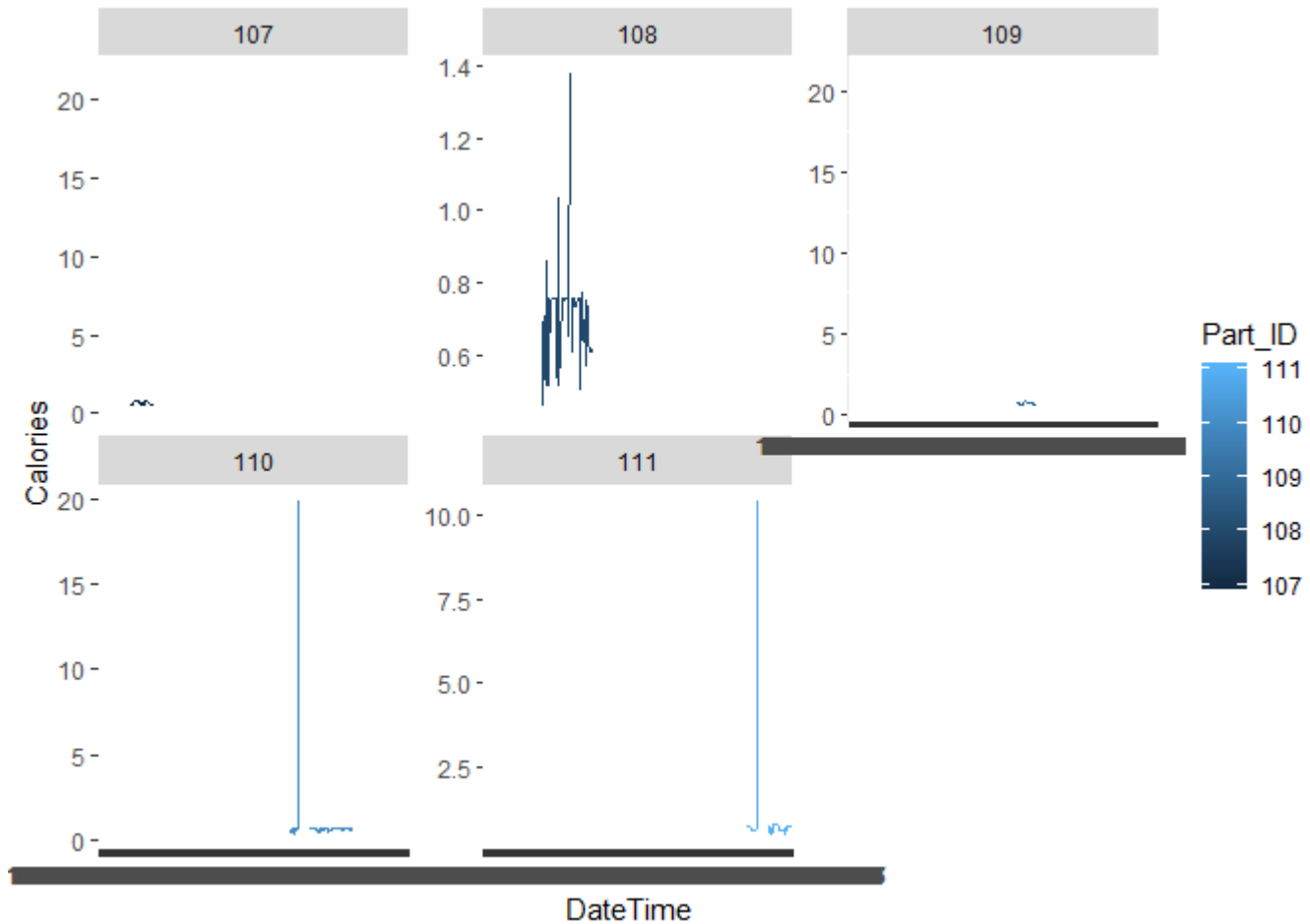
```
facet_wrap(~Part_ID, scales ="free_y")
```

```
## <ggproto object: Class FacetWrap, Facet, gg>
##     compute_layout: function
##     draw_back: function
##     draw_front: function
##     draw_labels: function
##     draw_panels: function
##     finish_data: function
##     init_scales: function
##     map_data: function
##     params: list
##     setup_data: function
##     setup_params: function
##     shrink: TRUE
##     train_scales: function
##     vars: function
##     super:  <ggproto object: Class FacetWrap, Facet, gg>
```

```
#Scatterplot of Calories vs. Time via Facet Wrap
  ggplot(applewatch_all, aes(x = DateTime, y = Calories, colour=Part_ID, group=Part_ID)) +
    geom_line() +
  facet_wrap(~Part_ID, scales ="free_y")
```
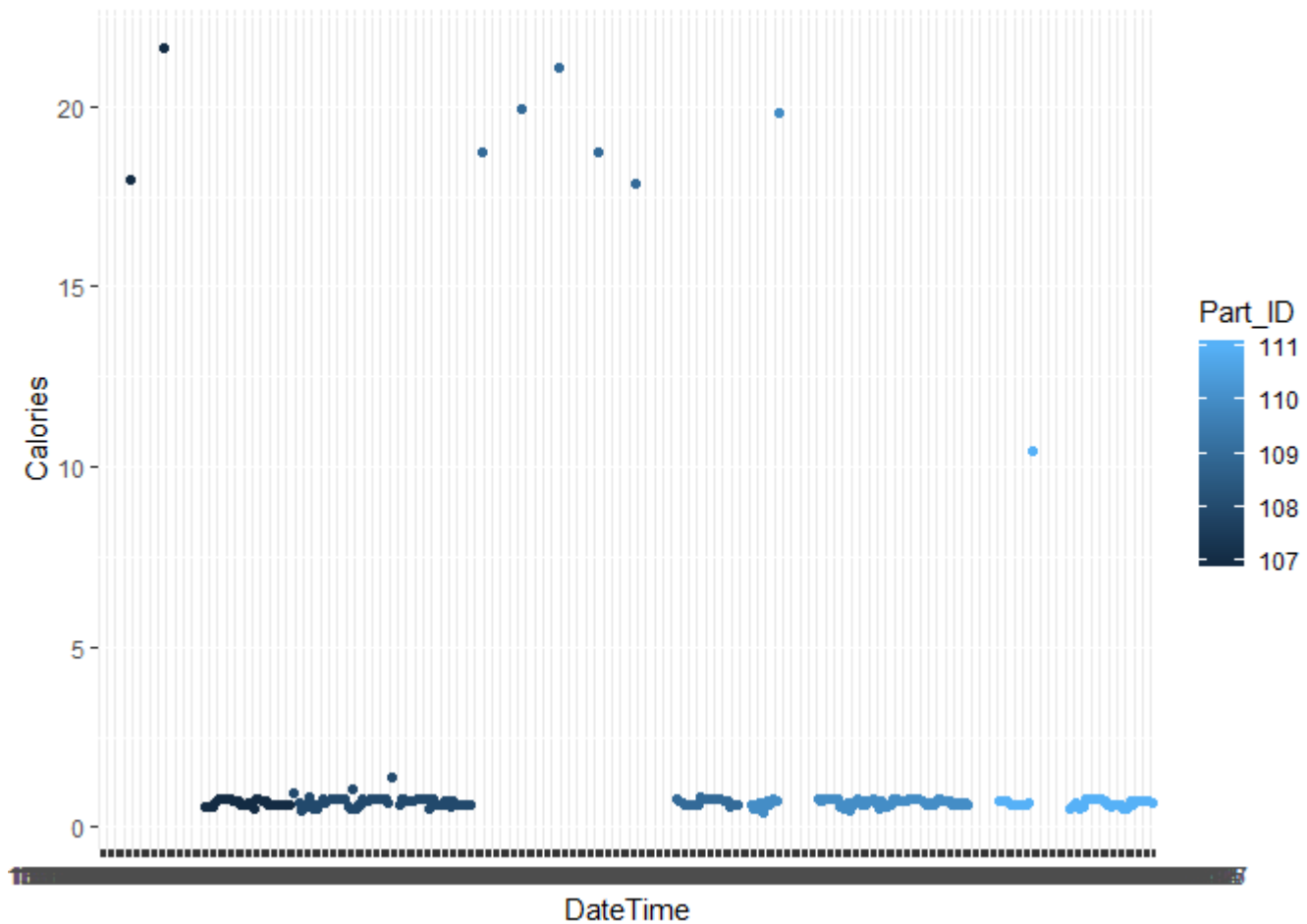
```
## Warning: Removed 33 rows containing missing values (geom_path).
```



```
#Scatterplot of Calories vs. Time, Condensed Scatter Plot
  ggplot(applewatch_all, aes(x = DateTime, y = Calories, colour=Part_ID, group=Part_ID)) +
    geom_point()
```

```
## Warning: Removed 164 rows containing missing values (geom_point).
```
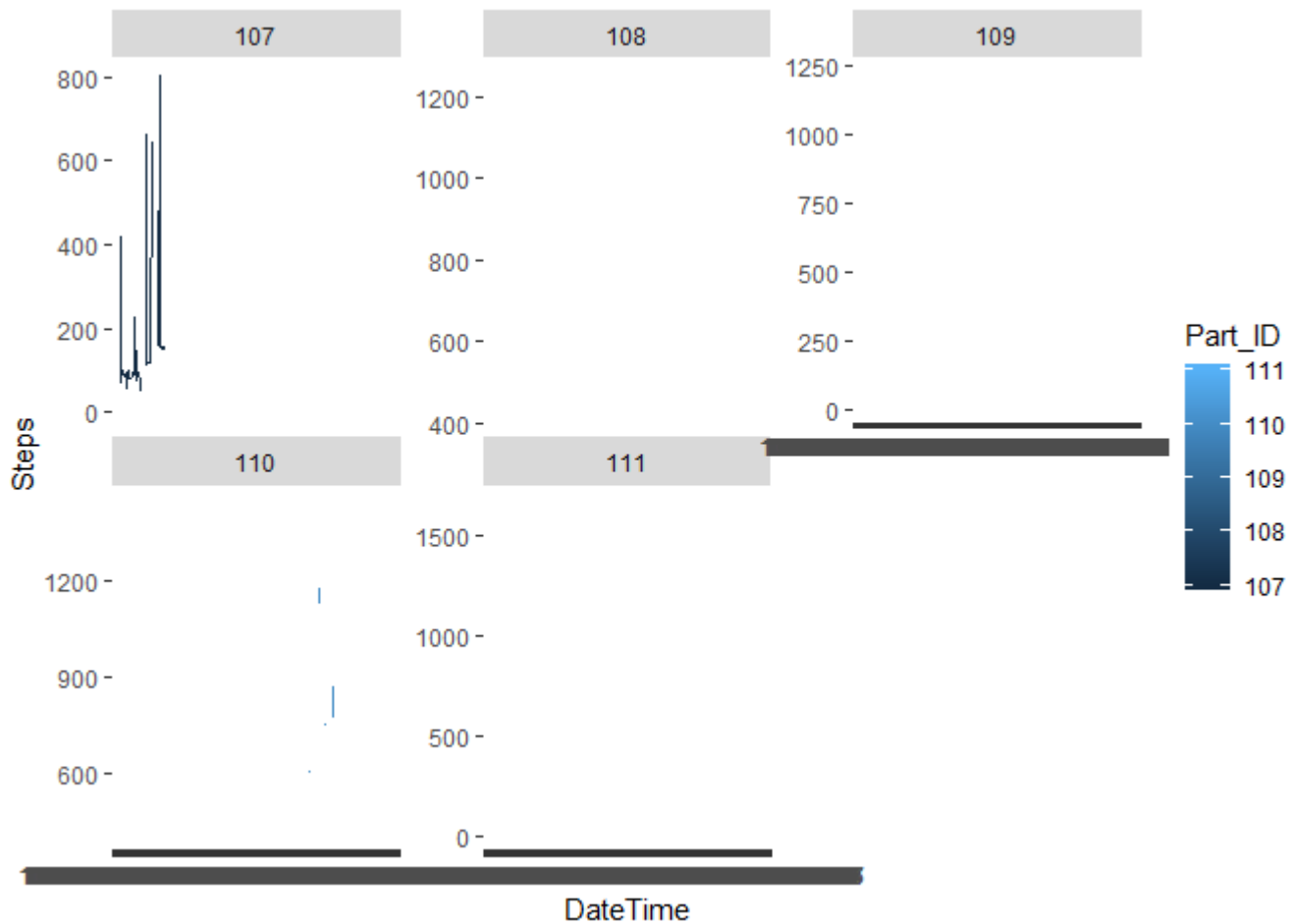
```
facet_wrap(~Part_ID, scales ="free_y")
```

```
## <ggproto object: Class FacetWrap, Facet, gg>
##     compute_layout: function
##     draw_back: function
##     draw_front: function
##     draw_labels: function
##     draw_panels: function
##     finish_data: function
##     init_scales: function
##     map_data: function
##     params: list
##     setup_data: function
##     setup_params: function
##     shrink: TRUE
##     train_scales: function
##     vars: function
##     super:  <ggproto object: Class FacetWrap, Facet, gg>
```

```
#Scatterplot of Steps vs. Time via Facet Wrap
  ggplot(applewatch_all, aes(x = DateTime, y = Steps, colour=Part_ID, group=Part_ID)) +
    geom_line() +
    facet_wrap(~Part_ID, scales ="free_y")
```
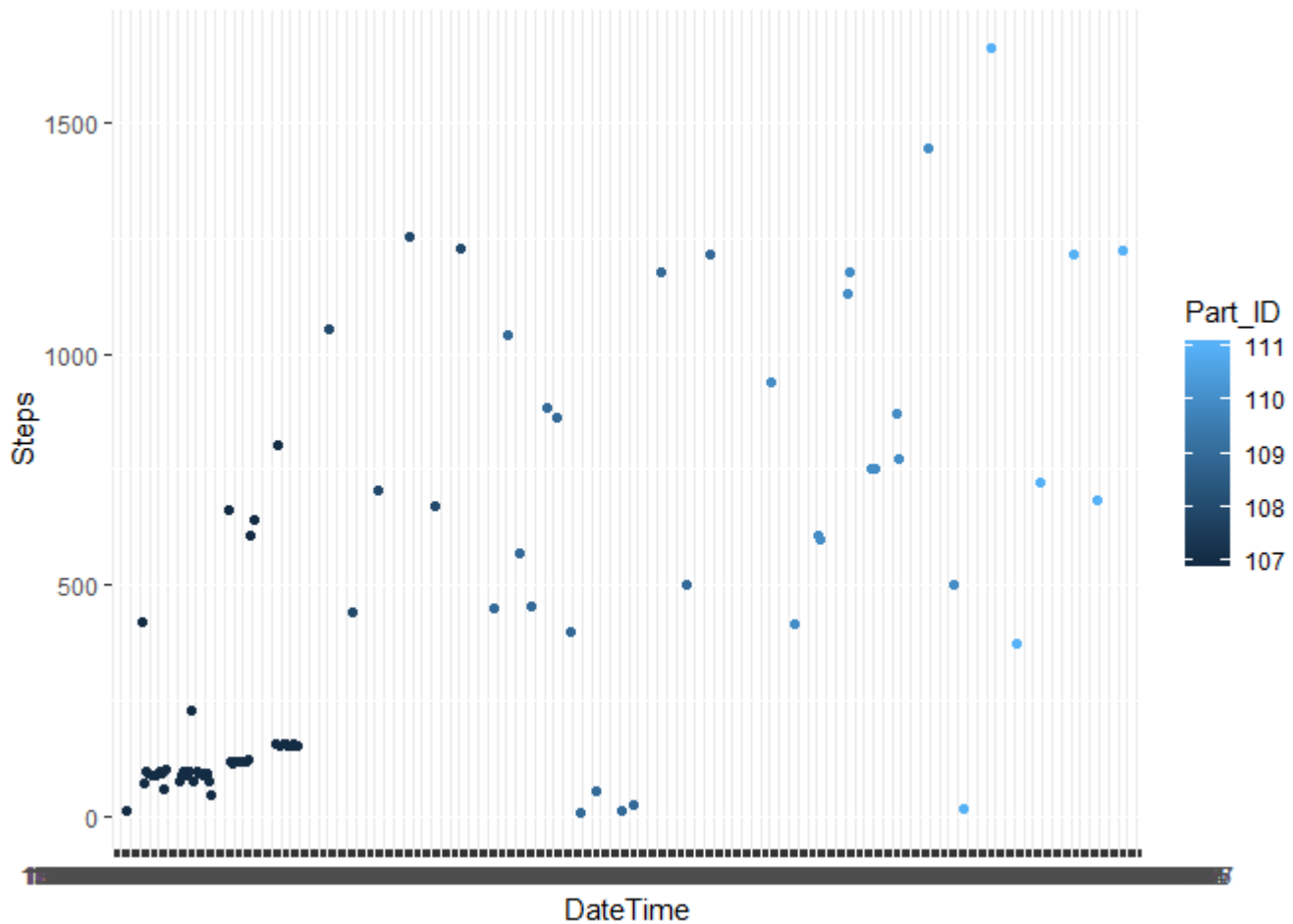
```
## Warning: Removed 62 rows containing missing values (geom_path).
```



```
#Scatterplot of Steps vs. Time, Condensed Scatter Plot
  ggplot(applewatch_all, aes(x = DateTime, y = Steps, colour=Part_ID, group=Part_ID)) +
    geom_point()
```

```
## Warning: Removed 328 rows containing missing values (geom_point).
```

```
facet_wrap(~Part_ID, scales ="free_y")
```

```
## <ggproto object: Class FacetWrap, Facet, gg>
##     compute_layout: function
##     draw_back: function
##     draw_front: function
##     draw_labels: function
##     draw_panels: function
##     finish_data: function
##     init_scales: function
##     map_data: function
##     params: list
##     setup_data: function
##     setup_params: function
##     shrink: TRUE
##     train_scales: function
##     vars: function
##     super:  <ggproto object: Class FacetWrap, Facet, gg>
```

```
setwd("C:/Thesis")
write.csv (applewatch_all, file='applewatch_all.csv')
```