#install.packages('gdata')  
#library("gdata")  
library("dplyr")

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library("tidyr")  
library("mice")

## Loading required package: lattice

##   
## Attaching package: 'mice'

## The following object is masked from 'package:tidyr':  
##   
## complete

library("VIM")

## Warning: package 'VIM' was built under R version 3.4.3

## Loading required package: colorspace

## Loading required package: grid

## Loading required package: data.table

##   
## Attaching package: 'data.table'

## The following objects are masked from 'package:dplyr':  
##   
## between, first, last

## VIM is ready to use.   
## Since version 4.0.0 the GUI is in its own package VIMGUI.  
##   
## Please use the package to use the new (and old) GUI.

## Suggestions and bug-reports can be submitted at: https://github.com/alexkowa/VIM/issues

##   
## Attaching package: 'VIM'

## The following object is masked from 'package:datasets':  
##   
## sleep

library("Hmisc")

## Loading required package: survival

## Loading required package: Formula

## Loading required package: ggplot2

##   
## Attaching package: 'Hmisc'

## The following objects are masked from 'package:dplyr':  
##   
## combine, src, summarize

## The following objects are masked from 'package:base':  
##   
## format.pval, round.POSIXt, trunc.POSIXt, units

library("lubridate")

##   
## Attaching package: 'lubridate'

## The following objects are masked from 'package:data.table':  
##   
## hour, isoweek, mday, minute, month, quarter, second, wday,  
## week, yday, year

## The following object is masked from 'package:base':  
##   
## date

library(corrplot)

## corrplot 0.84 loaded

library(reshape2)

##   
## Attaching package: 'reshape2'

## The following objects are masked from 'package:data.table':  
##   
## dcast, melt

## The following object is masked from 'package:tidyr':  
##   
## smiths

library(caret)

##   
## Attaching package: 'caret'

## The following object is masked from 'package:survival':  
##   
## cluster

library(AppliedPredictiveModeling)  
library(e1071)

##   
## Attaching package: 'e1071'

## The following object is masked from 'package:Hmisc':  
##   
## impute

library(cluster)  
library(lubridate)  
library(readxl)  
library(psych)

##   
## Attaching package: 'psych'

## The following object is masked from 'package:Hmisc':  
##   
## describe

## The following objects are masked from 'package:ggplot2':  
##   
## %+%, alpha

library(mice)  
library(MASS)

##   
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':  
##   
## select

library(reshape2)  
library(VIM)  
  
  
df = read.csv(file = "F:/HW 2/StudentData.csv",   
 na.strings = c("", " "),   
 header = TRUE)   
head(df)

## Brand.Code Carb.Volume Fill.Ounces PC.Volume Carb.Pressure Carb.Temp  
## 1 B 5.340000 23.96667 0.2633333 68.2 141.2  
## 2 A 5.426667 24.00667 0.2386667 68.4 139.6  
## 3 B 5.286667 24.06000 0.2633333 70.8 144.8  
## 4 A 5.440000 24.00667 0.2933333 63.0 132.6  
## 5 A 5.486667 24.31333 0.1113333 67.2 136.8  
## 6 A 5.380000 23.92667 0.2693333 66.6 138.4  
## PSC PSC.Fill PSC.CO2 Mnf.Flow Carb.Pressure1 Fill.Pressure  
## 1 0.104 0.26 0.04 -100 118.8 46.0  
## 2 0.124 0.22 0.04 -100 121.6 46.0  
## 3 0.090 0.34 0.16 -100 120.2 46.0  
## 4 NA 0.42 0.04 -100 115.2 46.4  
## 5 0.026 0.16 0.12 -100 118.4 45.8  
## 6 0.090 0.24 0.04 -100 119.6 45.6  
## Hyd.Pressure1 Hyd.Pressure2 Hyd.Pressure3 Hyd.Pressure4 Filler.Level  
## 1 0 NA NA 118 121.2  
## 2 0 NA NA 106 118.6  
## 3 0 NA NA 82 120.0  
## 4 0 0 0 92 117.8  
## 5 0 0 0 92 118.6  
## 6 0 0 0 116 120.2  
## Filler.Speed Temperature Usage.cont Carb.Flow Density MFR Balling  
## 1 4002 66.0 16.18 2932 0.88 725.0 1.398  
## 2 3986 67.6 19.90 3144 0.92 726.8 1.498  
## 3 4020 67.0 17.76 2914 1.58 735.0 3.142  
## 4 4012 65.6 17.42 3062 1.54 730.6 3.042  
## 5 4010 65.6 17.68 3054 1.54 722.8 3.042  
## 6 4014 66.2 23.82 2948 1.52 738.8 2.992  
## Pressure.Vacuum PH Oxygen.Filler Bowl.Setpoint Pressure.Setpoint  
## 1 -4.0 8.36 0.022 120 46.4  
## 2 -4.0 8.26 0.026 120 46.8  
## 3 -3.8 8.94 0.024 120 46.6  
## 4 -4.4 8.24 0.030 120 46.0  
## 5 -4.4 8.26 0.030 120 46.0  
## 6 -4.4 8.32 0.024 120 46.0  
## Air.Pressurer Alch.Rel Carb.Rel Balling.Lvl  
## 1 142.6 6.58 5.32 1.48  
## 2 143.0 6.56 5.30 1.56  
## 3 142.0 7.66 5.84 3.28  
## 4 146.2 7.14 5.42 3.04  
## 5 146.2 7.14 5.44 3.04  
## 6 146.6 7.16 5.44 3.02

tail(df)

## Brand.Code Carb.Volume Fill.Ounces PC.Volume Carb.Pressure Carb.Temp  
## 2566 C 5.320000 24.11333 0.2753333 62.8 136.0  
## 2567 <NA> 5.500000 24.04000 0.2486667 74.8 145.8  
## 2568 <NA> 5.506667 23.98000 0.2540000 74.4 146.0  
## 2569 B 5.346667 24.04000 0.2806667 68.0 141.8  
## 2570 B 5.366667 24.02667 0.2646667 68.2 141.6  
## 2571 A 5.253333 23.96000 0.1686667 67.2 142.4  
## PSC PSC.Fill PSC.CO2 Mnf.Flow Carb.Pressure1 Fill.Pressure  
## 2566 0.080 0.32 0.08 144.4 124.4 54.2  
## 2567 0.172 0.36 0.04 136.0 123.4 56.0  
## 2568 0.108 0.22 0.02 138.8 122.6 49.8  
## 2569 0.032 0.28 0.02 143.8 122.0 53.0  
## 2570 0.040 0.32 0.02 147.2 121.6 55.0  
## 2571 0.030 0.22 0.02 0.2 130.2 44.4  
## Hyd.Pressure1 Hyd.Pressure2 Hyd.Pressure3 Hyd.Pressure4 Filler.Level  
## 2566 6.8 28.4 35.8 88 111.8  
## 2567 4.8 25.4 33.6 86 101.0  
## 2568 7.4 25.8 34.2 84 109.6  
## 2569 9.0 26.2 34.6 102 109.8  
## 2570 11.2 26.6 34.8 100 110.4  
## 2571 -0.8 0.2 -1.2 NA 109.8  
## Filler.Speed Temperature Usage.cont Carb.Flow Density MFR Balling  
## 2566 4000 65.4 23.94 1190 0.62 708.6 1.966  
## 2567 4002 66.8 23.94 1220 0.78 716.8 2.366  
## 2568 3994 66.6 23.82 1192 0.64 730.8 2.016  
## 2569 3996 66.0 23.66 1182 0.44 733.2 1.516  
## 2570 3992 66.0 23.78 1182 0.46 720.8 1.566  
## 2571 NA 66.2 23.58 44 1.10 NA 3.166  
## Pressure.Vacuum PH Oxygen.Filler Bowl.Setpoint Pressure.Setpoint  
## 2566 -5.8 8.78 0.0252 110 50  
## 2567 -5.8 8.86 0.0026 110 50  
## 2568 -5.6 8.88 0.0026 110 50  
## 2569 -5.6 8.68 0.0026 110 50  
## 2570 -5.4 8.56 0.0026 110 50  
## 2571 -5.4 8.46 0.0248 110 50  
## Air.Pressurer Alch.Rel Carb.Rel Balling.Lvl  
## 2566 142.4 6.52 5.30 1.60  
## 2567 141.8 7.72 5.56 1.62  
## 2568 142.6 7.68 5.58 1.78  
## 2569 143.0 6.56 5.36 1.28  
## 2570 142.8 6.56 5.36 1.28  
## 2571 142.2 7.16 5.48 3.10

#Explore the data  
  
str(df)

## 'data.frame': 2571 obs. of 33 variables:  
## $ Brand.Code : Factor w/ 4 levels "A","B","C","D": 2 1 2 1 1 1 1 2 2 2 ...  
## $ Carb.Volume : num 5.34 5.43 5.29 5.44 5.49 ...  
## $ Fill.Ounces : num 24 24 24.1 24 24.3 ...  
## $ PC.Volume : num 0.263 0.239 0.263 0.293 0.111 ...  
## $ Carb.Pressure : num 68.2 68.4 70.8 63 67.2 66.6 64.2 67.6 64.2 72 ...  
## $ Carb.Temp : num 141 140 145 133 137 ...  
## $ PSC : num 0.104 0.124 0.09 NA 0.026 0.09 0.128 0.154 0.132 0.014 ...  
## $ PSC.Fill : num 0.26 0.22 0.34 0.42 0.16 0.24 0.4 0.34 0.12 0.24 ...  
## $ PSC.CO2 : num 0.04 0.04 0.16 0.04 0.12 0.04 0.04 0.04 0.14 0.06 ...  
## $ Mnf.Flow : num -100 -100 -100 -100 -100 -100 -100 -100 -100 -100 ...  
## $ Carb.Pressure1 : num 119 122 120 115 118 ...  
## $ Fill.Pressure : num 46 46 46 46.4 45.8 45.6 51.8 46.8 46 45.2 ...  
## $ Hyd.Pressure1 : num 0 0 0 0 0 0 0 0 0 0 ...  
## $ Hyd.Pressure2 : num NA NA NA 0 0 0 0 0 0 0 ...  
## $ Hyd.Pressure3 : num NA NA NA 0 0 0 0 0 0 0 ...  
## $ Hyd.Pressure4 : int 118 106 82 92 92 116 124 132 90 108 ...  
## $ Filler.Level : num 121 119 120 118 119 ...  
## $ Filler.Speed : int 4002 3986 4020 4012 4010 4014 NA 1004 4014 4028 ...  
## $ Temperature : num 66 67.6 67 65.6 65.6 66.2 65.8 65.2 65.4 66.6 ...  
## $ Usage.cont : num 16.2 19.9 17.8 17.4 17.7 ...  
## $ Carb.Flow : int 2932 3144 2914 3062 3054 2948 30 684 2902 3038 ...  
## $ Density : num 0.88 0.92 1.58 1.54 1.54 1.52 0.84 0.84 0.9 0.9 ...  
## $ MFR : num 725 727 735 731 723 ...  
## $ Balling : num 1.4 1.5 3.14 3.04 3.04 ...  
## $ Pressure.Vacuum : num -4 -4 -3.8 -4.4 -4.4 -4.4 -4.4 -4.4 -4.4 -4.4 ...  
## $ PH : num 8.36 8.26 8.94 8.24 8.26 8.32 8.4 8.38 8.38 8.5 ...  
## $ Oxygen.Filler : num 0.022 0.026 0.024 0.03 0.03 0.024 0.066 0.046 0.064 0.022 ...  
## $ Bowl.Setpoint : int 120 120 120 120 120 120 120 120 120 120 ...  
## $ Pressure.Setpoint: num 46.4 46.8 46.6 46 46 46 46 46 46 46 ...  
## $ Air.Pressurer : num 143 143 142 146 146 ...  
## $ Alch.Rel : num 6.58 6.56 7.66 7.14 7.14 7.16 6.54 6.52 6.52 6.54 ...  
## $ Carb.Rel : num 5.32 5.3 5.84 5.42 5.44 5.44 5.38 5.34 5.34 5.34 ...  
## $ Balling.Lvl : num 1.48 1.56 3.28 3.04 3.04 3.02 1.44 1.44 1.44 1.38 ...

summary(df)

## Brand.Code Carb.Volume Fill.Ounces PC.Volume   
## A : 293 Min. :5.040 Min. :23.63 Min. :0.07933   
## B :1239 1st Qu.:5.293 1st Qu.:23.92 1st Qu.:0.23917   
## C : 304 Median :5.347 Median :23.97 Median :0.27133   
## D : 615 Mean :5.370 Mean :23.97 Mean :0.27712   
## NA's: 120 3rd Qu.:5.453 3rd Qu.:24.03 3rd Qu.:0.31200   
## Max. :5.700 Max. :24.32 Max. :0.47800   
## NA's :10 NA's :38 NA's :39   
## Carb.Pressure Carb.Temp PSC PSC.Fill   
## Min. :57.00 Min. :128.6 Min. :0.00200 Min. :0.0000   
## 1st Qu.:65.60 1st Qu.:138.4 1st Qu.:0.04800 1st Qu.:0.1000   
## Median :68.20 Median :140.8 Median :0.07600 Median :0.1800   
## Mean :68.19 Mean :141.1 Mean :0.08457 Mean :0.1954   
## 3rd Qu.:70.60 3rd Qu.:143.8 3rd Qu.:0.11200 3rd Qu.:0.2600   
## Max. :79.40 Max. :154.0 Max. :0.27000 Max. :0.6200   
## NA's :27 NA's :26 NA's :33 NA's :23   
## PSC.CO2 Mnf.Flow Carb.Pressure1 Fill.Pressure   
## Min. :0.00000 Min. :-100.20 Min. :105.6 Min. :34.60   
## 1st Qu.:0.02000 1st Qu.:-100.00 1st Qu.:119.0 1st Qu.:46.00   
## Median :0.04000 Median : 65.20 Median :123.2 Median :46.40   
## Mean :0.05641 Mean : 24.57 Mean :122.6 Mean :47.92   
## 3rd Qu.:0.08000 3rd Qu.: 140.80 3rd Qu.:125.4 3rd Qu.:50.00   
## Max. :0.24000 Max. : 229.40 Max. :140.2 Max. :60.40   
## NA's :39 NA's :2 NA's :32 NA's :22   
## Hyd.Pressure1 Hyd.Pressure2 Hyd.Pressure3 Hyd.Pressure4   
## Min. :-0.80 Min. : 0.00 Min. :-1.20 Min. : 52.00   
## 1st Qu.: 0.00 1st Qu.: 0.00 1st Qu.: 0.00 1st Qu.: 86.00   
## Median :11.40 Median :28.60 Median :27.60 Median : 96.00   
## Mean :12.44 Mean :20.96 Mean :20.46 Mean : 96.29   
## 3rd Qu.:20.20 3rd Qu.:34.60 3rd Qu.:33.40 3rd Qu.:102.00   
## Max. :58.00 Max. :59.40 Max. :50.00 Max. :142.00   
## NA's :11 NA's :15 NA's :15 NA's :30   
## Filler.Level Filler.Speed Temperature Usage.cont   
## Min. : 55.8 Min. : 998 Min. :63.60 Min. :12.08   
## 1st Qu.: 98.3 1st Qu.:3888 1st Qu.:65.20 1st Qu.:18.36   
## Median :118.4 Median :3982 Median :65.60 Median :21.79   
## Mean :109.3 Mean :3687 Mean :65.97 Mean :20.99   
## 3rd Qu.:120.0 3rd Qu.:3998 3rd Qu.:66.40 3rd Qu.:23.75   
## Max. :161.2 Max. :4030 Max. :76.20 Max. :25.90   
## NA's :20 NA's :57 NA's :14 NA's :5   
## Carb.Flow Density MFR Balling   
## Min. : 26 Min. :0.240 Min. : 31.4 Min. :-0.170   
## 1st Qu.:1144 1st Qu.:0.900 1st Qu.:706.3 1st Qu.: 1.496   
## Median :3028 Median :0.980 Median :724.0 Median : 1.648   
## Mean :2468 Mean :1.174 Mean :704.0 Mean : 2.198   
## 3rd Qu.:3186 3rd Qu.:1.620 3rd Qu.:731.0 3rd Qu.: 3.292   
## Max. :5104 Max. :1.920 Max. :868.6 Max. : 4.012   
## NA's :2 NA's :1 NA's :212 NA's :1   
## Pressure.Vacuum PH Oxygen.Filler Bowl.Setpoint   
## Min. :-6.600 Min. :7.880 Min. :0.00240 Min. : 70.0   
## 1st Qu.:-5.600 1st Qu.:8.440 1st Qu.:0.02200 1st Qu.:100.0   
## Median :-5.400 Median :8.540 Median :0.03340 Median :120.0   
## Mean :-5.216 Mean :8.546 Mean :0.04684 Mean :109.3   
## 3rd Qu.:-5.000 3rd Qu.:8.680 3rd Qu.:0.06000 3rd Qu.:120.0   
## Max. :-3.600 Max. :9.360 Max. :0.40000 Max. :140.0   
## NA's :4 NA's :12 NA's :2   
## Pressure.Setpoint Air.Pressurer Alch.Rel Carb.Rel   
## Min. :44.00 Min. :140.8 Min. :5.280 Min. :4.960   
## 1st Qu.:46.00 1st Qu.:142.2 1st Qu.:6.540 1st Qu.:5.340   
## Median :46.00 Median :142.6 Median :6.560 Median :5.400   
## Mean :47.62 Mean :142.8 Mean :6.897 Mean :5.437   
## 3rd Qu.:50.00 3rd Qu.:143.0 3rd Qu.:7.240 3rd Qu.:5.540   
## Max. :52.00 Max. :148.2 Max. :8.620 Max. :6.060   
## NA's :12 NA's :9 NA's :10   
## Balling.Lvl   
## Min. :0.00   
## 1st Qu.:1.38   
## Median :1.48   
## Mean :2.05   
## 3rd Qu.:3.14   
## Max. :3.66   
## NA's :1

describe(df) #Need to consider what to do with zero values: Brand Code of zero for example

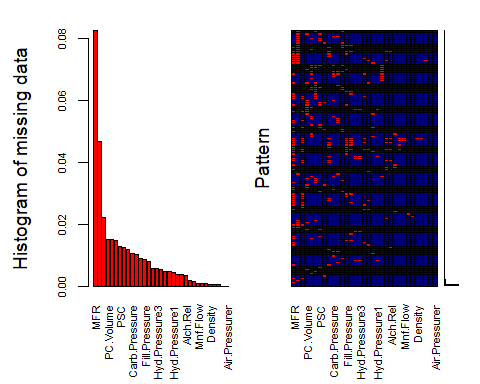
## vars n mean sd median trimmed mad min  
## Brand.Code\* 1 2451 2.51 1.00 2.00 2.51 0.00 1.00  
## Carb.Volume 2 2561 5.37 0.11 5.35 5.37 0.11 5.04  
## Fill.Ounces 3 2533 23.97 0.09 23.97 23.98 0.08 23.63  
## PC.Volume 4 2532 0.28 0.06 0.27 0.27 0.05 0.08  
## Carb.Pressure 5 2544 68.19 3.54 68.20 68.12 3.56 57.00  
## Carb.Temp 6 2545 141.09 4.04 140.80 140.99 3.85 128.60  
## PSC 7 2538 0.08 0.05 0.08 0.08 0.05 0.00  
## PSC.Fill 8 2548 0.20 0.12 0.18 0.18 0.12 0.00  
## PSC.CO2 9 2532 0.06 0.04 0.04 0.05 0.03 0.00  
## Mnf.Flow 10 2569 24.57 119.48 65.20 21.07 169.02 -100.20  
## Carb.Pressure1 11 2539 122.59 4.74 123.20 122.54 4.45 105.60  
## Fill.Pressure 12 2549 47.92 3.18 46.40 47.71 2.37 34.60  
## Hyd.Pressure1 13 2560 12.44 12.43 11.40 10.84 16.90 -0.80  
## Hyd.Pressure2 14 2556 20.96 16.39 28.60 21.05 13.34 0.00  
## Hyd.Pressure3 15 2556 20.46 15.98 27.60 20.51 13.94 -1.20  
## Hyd.Pressure4 16 2541 96.29 13.12 96.00 95.45 11.86 52.00  
## Filler.Level 17 2551 109.25 15.70 118.40 111.04 9.19 55.80  
## Filler.Speed 18 2514 3687.20 770.82 3982.00 3919.99 47.44 998.00  
## Temperature 19 2557 65.97 1.38 65.60 65.80 0.89 63.60  
## Usage.cont 20 2566 20.99 2.98 21.79 21.25 3.19 12.08  
## Carb.Flow 21 2569 2468.35 1073.70 3028.00 2601.14 326.17 26.00  
## Density 22 2570 1.17 0.38 0.98 1.15 0.15 0.24  
## MFR 23 2359 704.05 73.90 724.00 718.16 15.42 31.40  
## Balling 24 2570 2.20 0.93 1.65 2.13 0.37 -0.17  
## Pressure.Vacuum 25 2571 -5.22 0.57 -5.40 -5.25 0.59 -6.60  
## PH 26 2567 8.55 0.17 8.54 8.55 0.18 7.88  
## Oxygen.Filler 27 2559 0.05 0.05 0.03 0.04 0.02 0.00  
## Bowl.Setpoint 28 2569 109.33 15.30 120.00 111.35 0.00 70.00  
## Pressure.Setpoint 29 2559 47.62 2.04 46.00 47.60 0.00 44.00  
## Air.Pressurer 30 2571 142.83 1.21 142.60 142.58 0.59 140.80  
## Alch.Rel 31 2562 6.90 0.51 6.56 6.84 0.06 5.28  
## Carb.Rel 32 2561 5.44 0.13 5.40 5.43 0.12 4.96  
## Balling.Lvl 33 2570 2.05 0.87 1.48 1.98 0.21 0.00  
## max range skew kurtosis se  
## Brand.Code\* 4.00 3.00 0.38 -1.06 0.02  
## Carb.Volume 5.70 0.66 0.39 -0.47 0.00  
## Fill.Ounces 24.32 0.69 -0.02 0.86 0.00  
## PC.Volume 0.48 0.40 0.34 0.67 0.00  
## Carb.Pressure 79.40 22.40 0.18 -0.01 0.07  
## Carb.Temp 154.00 25.40 0.25 0.24 0.08  
## PSC 0.27 0.27 0.85 0.65 0.00  
## PSC.Fill 0.62 0.62 0.93 0.77 0.00  
## PSC.CO2 0.24 0.24 1.73 3.73 0.00  
## Mnf.Flow 229.40 329.60 0.00 -1.87 2.36  
## Carb.Pressure1 140.20 34.60 0.05 0.14 0.09  
## Fill.Pressure 60.40 25.80 0.55 1.41 0.06  
## Hyd.Pressure1 58.00 58.80 0.78 -0.14 0.25  
## Hyd.Pressure2 59.40 59.40 -0.30 -1.56 0.32  
## Hyd.Pressure3 50.00 51.20 -0.32 -1.57 0.32  
## Hyd.Pressure4 142.00 90.00 0.55 0.63 0.26  
## Filler.Level 161.20 105.40 -0.85 0.05 0.31  
## Filler.Speed 4030.00 3032.00 -2.87 6.71 15.37  
## Temperature 76.20 12.60 2.39 10.16 0.03  
## Usage.cont 25.90 13.82 -0.54 -1.02 0.06  
## Carb.Flow 5104.00 5078.00 -0.99 -0.58 21.18  
## Density 1.92 1.68 0.53 -1.20 0.01  
## MFR 868.60 837.20 -5.09 30.46 1.52  
## Balling 4.01 4.18 0.59 -1.39 0.02  
## Pressure.Vacuum -3.60 3.00 0.53 -0.03 0.01  
## PH 9.36 1.48 -0.29 0.06 0.00  
## Oxygen.Filler 0.40 0.40 2.66 11.09 0.00  
## Bowl.Setpoint 140.00 70.00 -0.97 -0.06 0.30  
## Pressure.Setpoint 52.00 8.00 0.20 -1.60 0.04  
## Air.Pressurer 148.20 7.40 2.25 4.73 0.02  
## Alch.Rel 8.62 3.34 0.88 -0.85 0.01  
## Carb.Rel 6.06 1.10 0.50 -0.29 0.00  
## Balling.Lvl 3.66 3.66 0.59 -1.49 0.02

md.pattern(df)

## Pressure.Vacuum Air.Pressurer Density Balling Balling.Lvl Mnf.Flow  
## 2038 1 1 1 1 1 1  
## 91 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 18 1 1 1 1 1 1  
## 17 1 1 1 1 1 1  
## 9 1 1 1 1 1 1  
## 10 1 1 1 1 1 1  
## 15 1 1 1 1 1 1  
## 9 1 1 1 1 1 1  
## 18 1 1 1 1 1 1  
## 27 1 1 1 1 1 1  
## 5 1 1 1 1 1 1  
## 3 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 100 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 7 1 1 1 1 1 1  
## 4 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 3 1 1 1 1 1 1  
## 3 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 7 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 3 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 3 1 1 1 1 1 1  
## 4 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 3 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 4 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 4 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 10 1 1 1 1 1 1  
## 30 1 1 1 1 1 1  
## 4 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 4 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 9 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 3 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 0 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 0  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 0 0 1 0  
## 0 0 1 1 1 2  
## Carb.Flow Bowl.Setpoint PH Usage.cont Alch.Rel Carb.Volume Carb.Rel  
## 2038 1 1 1 1 1 1 1  
## 91 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 0 1  
## 18 1 1 1 1 1 1 1  
## 17 1 1 1 1 1 1 1  
## 9 1 1 1 1 1 1 1  
## 10 1 1 1 1 1 1 1  
## 15 1 1 1 1 1 1 1  
## 9 1 1 1 1 1 1 1  
## 18 1 1 1 1 1 1 1  
## 27 1 1 1 1 1 1 1  
## 5 1 1 1 1 1 1 1  
## 3 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1 1  
## 2 1 1 1 0 1 1 1  
## 2 0 1 1 1 1 1 1  
## 100 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 2 1 0 1 1 1 1 1  
## 7 1 1 1 1 1 1 1  
## 4 1 1 1 1 0 1 1  
## 2 1 1 1 1 1 1 0  
## 3 1 1 1 1 1 1 1  
## 3 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 0 1  
## 1 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 7 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 3 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 0 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 3 1 1 1 1 1 1 1  
## 4 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 0 1 1 1  
## 3 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 4 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 4 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 10 1 1 1 1 1 1 1  
## 30 1 1 1 1 1 1 1  
## 4 1 1 1 1 1 1 1  
## 1 1 1 1 0 1 1 1  
## 4 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 0 1 1 0  
## 1 1 1 1 1 1 1 0  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 0 1  
## 2 1 1 1 1 1 0 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 0 1  
## 9 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 3 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 0 1  
## 1 1 1 0 1 1 1 1  
## 1 1 1 1 1 0 1 0  
## 1 1 1 1 1 1 1 0  
## 1 1 1 1 1 1 0 1  
## 1 1 1 1 1 1 0 1  
## 1 1 1 0 1 1 1 1  
## 1 1 1 1 1 0 1 0  
## 1 1 1 1 1 0 1 0  
## 1 1 1 0 1 0 1 0  
## 1 1 1 0 1 0 1 0  
## 2 2 4 5 9 10 10  
## Hyd.Pressure1 Oxygen.Filler Pressure.Setpoint Temperature  
## 2038 1 1 1 1  
## 91 1 1 1 1  
## 1 1 1 1 1  
## 18 1 1 1 1  
## 17 1 1 1 1  
## 9 1 1 1 1  
## 10 1 1 1 1  
## 15 1 1 1 1  
## 9 1 1 1 1  
## 18 1 1 1 1  
## 27 1 1 1 1  
## 5 1 1 1 1  
## 3 1 1 1 1  
## 2 1 1 1 1  
## 2 1 1 1 1  
## 2 1 1 1 1  
## 100 1 1 1 1  
## 1 1 0 1 1  
## 2 1 1 1 1  
## 7 1 1 0 1  
## 4 1 1 1 1  
## 2 1 1 1 1  
## 3 1 1 1 1  
## 3 1 1 1 1  
## 2 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 2 1 1 1 1  
## 1 1 1 1 1  
## 7 1 1 1 1  
## 1 1 1 1 1  
## 3 1 1 1 1  
## 2 1 1 1 1  
## 2 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 2 1 1 1 1  
## 1 1 1 1 1  
## 3 1 1 1 1  
## 4 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 3 1 1 1 1  
## 2 1 1 1 1  
## 2 1 1 1 1  
## 2 1 1 1 1  
## 1 1 1 1 1  
## 4 1 1 1 1  
## 1 1 1 1 1  
## 4 1 1 1 1  
## 1 1 1 1 1  
## 10 1 1 1 1  
## 30 1 1 1 1  
## 4 1 1 1 0  
## 1 1 1 1 1  
## 4 1 0 1 1  
## 1 1 1 0 1  
## 1 1 1 0 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 2 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 9 0 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 3 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 2 1 1 1 1  
## 2 1 1 1 1  
## 2 1 1 1 1  
## 1 1 1 1 0  
## 2 1 1 1 0  
## 2 1 0 1 1  
## 2 1 0 1 0  
## 1 1 1 0 1  
## 1 1 1 1 1  
## 1 0 1 1 1  
## 1 0 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 2 1 1 1 1  
## 1 1 1 1 0  
## 1 1 0 1 1  
## 1 1 1 0 1  
## 1 1 1 0 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 0 1 1  
## 1 1 1 1 1  
## 1 1 1 1 0  
## 1 1 1 1 1  
## 1 1 1 1 1  
## 1 1 1 1 0  
## 1 1 0 1 0  
## 1 1 1 1 0  
## 11 12 12 14  
## Hyd.Pressure2 Hyd.Pressure3 Filler.Level Fill.Pressure PSC.Fill  
## 2038 1 1 1 1 1  
## 91 1 1 1 1 1  
## 1 1 1 1 1 1  
## 18 1 1 1 1 1  
## 17 1 1 1 1 1  
## 9 1 1 1 1 1  
## 10 1 1 1 1 1  
## 15 1 1 1 1 1  
## 9 1 1 1 1 0  
## 18 1 1 1 1 1  
## 27 1 1 1 1 1  
## 5 1 1 1 1 1  
## 3 1 1 0 1 1  
## 2 1 1 1 1 1  
## 2 1 1 1 1 1  
## 2 1 1 1 1 1  
## 100 1 1 1 1 1  
## 1 1 1 1 1 1  
## 2 1 1 1 1 1  
## 7 1 1 1 1 1  
## 4 1 1 1 1 1  
## 2 1 1 1 1 1  
## 3 1 1 1 1 1  
## 3 1 1 1 1 1  
## 2 1 1 1 1 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 1  
## 2 1 1 1 1 1  
## 1 1 1 1 1 1  
## 7 1 1 1 1 1  
## 1 1 1 1 1 1  
## 3 1 1 1 1 1  
## 2 1 1 1 1 0  
## 2 1 1 1 1 0  
## 1 1 1 1 1 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 1  
## 2 1 1 1 1 1  
## 1 1 1 1 1 0  
## 3 0 0 1 1 1  
## 4 1 1 0 0 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 1  
## 3 1 1 1 1 1  
## 2 1 1 1 1 1  
## 2 1 1 1 1 1  
## 2 1 1 1 1 1  
## 1 1 1 1 1 1  
## 4 1 1 1 1 1  
## 1 1 1 1 1 0  
## 4 1 1 1 1 1  
## 1 1 1 1 0 1  
## 10 1 1 1 1 1  
## 30 1 1 1 1 1  
## 4 1 1 1 1 1  
## 1 1 1 1 1 1  
## 4 1 1 1 1 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 0  
## 1 1 1 1 1 1  
## 2 1 1 1 1 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 0  
## 1 1 1 1 1 1  
## 1 1 1 1 1 0  
## 1 1 1 1 0 1  
## 9 0 0 1 1 1  
## 1 1 1 0 0 1  
## 1 0 0 1 1 1  
## 1 1 1 1 1 1  
## 1 1 1 1 0 1  
## 1 1 1 0 1 1  
## 3 1 1 1 1 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 1  
## 2 1 1 1 0 1  
## 2 1 1 1 1 1  
## 2 1 1 0 1 1  
## 1 1 1 1 1 1  
## 2 1 1 1 1 1  
## 2 1 1 1 1 1  
## 2 1 1 1 1 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 0  
## 1 0 0 1 1 1  
## 1 0 0 1 1 1  
## 1 1 1 0 0 0  
## 1 1 1 1 0 1  
## 1 1 1 1 1 0  
## 1 1 1 1 0 1  
## 2 1 1 0 0 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 1  
## 1 1 1 1 0 1  
## 1 1 1 1 1 0  
## 1 1 1 0 0 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 1  
## 1 1 1 1 1 0  
## 1 1 1 1 1 1  
## 1 1 1 0 0 1  
## 1 1 1 0 0 1  
## 1 1 1 0 0 1  
## 1 1 1 0 0 1  
## 1 1 1 0 0 1  
## 15 15 20 22 23  
## Carb.Temp Carb.Pressure Hyd.Pressure4 Carb.Pressure1 PSC Fill.Ounces  
## 2038 1 1 1 1 1 1  
## 91 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 18 1 1 1 1 1 0  
## 17 1 1 1 1 1 1  
## 9 1 0 1 1 1 1  
## 10 0 1 1 1 1 1  
## 15 1 1 1 1 0 1  
## 9 1 1 1 1 1 1  
## 18 1 1 1 1 1 1  
## 27 1 1 1 0 1 1  
## 5 1 1 0 1 1 1  
## 3 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 100 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 7 1 1 1 1 1 1  
## 4 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 3 1 1 1 1 1 0  
## 3 1 1 1 1 1 1  
## 2 1 1 1 1 1 0  
## 1 1 0 1 1 1 1  
## 1 1 0 1 1 1 1  
## 2 0 1 1 1 1 1  
## 1 0 1 1 1 1 0  
## 7 0 0 1 1 1 1  
## 1 1 1 1 1 0 1  
## 3 1 1 1 1 0 1  
## 2 1 1 1 1 1 0  
## 2 1 1 1 1 0 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 0 1 1 1 1  
## 1 0 1 1 1 1 1  
## 1 1 1 1 1 0 1  
## 2 1 1 1 0 1 1  
## 1 1 1 1 0 1 1  
## 3 1 1 1 1 1 1  
## 4 1 1 1 1 1 1  
## 1 1 1 0 1 1 1  
## 1 1 1 1 1 1 1  
## 3 1 1 1 1 1 1  
## 2 1 1 1 1 1 0  
## 2 1 1 1 1 1 1  
## 2 1 0 1 1 1 1  
## 1 0 1 1 1 1 1  
## 4 1 1 1 1 0 1  
## 1 1 1 1 1 1 1  
## 4 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 10 1 1 0 1 1 1  
## 30 1 1 1 1 1 1  
## 4 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 4 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 0 0 1 1 1 1  
## 1 1 1 1 1 0 1  
## 1 1 1 1 1 1 0  
## 1 1 1 1 1 1 0  
## 2 1 0 1 1 1 1  
## 1 1 1 1 1 0 1  
## 1 1 1 1 1 1 0  
## 1 1 1 1 0 0 1  
## 1 1 1 1 0 1 1  
## 1 1 1 1 1 1 1  
## 9 1 1 1 1 1 1  
## 1 1 1 1 1 1 0  
## 1 1 1 1 1 1 1  
## 1 1 1 0 1 1 1  
## 1 1 1 0 1 1 1  
## 1 1 1 1 1 1 0  
## 3 1 1 1 1 1 1  
## 1 1 1 1 1 1 0  
## 1 1 1 1 1 1 1  
## 1 0 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 2 1 1 0 1 1 1  
## 2 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 2 1 1 0 1 1 1  
## 2 1 1 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 0 0  
## 1 1 1 1 1 1 1  
## 1 1 1 0 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 1 1 0 1  
## 1 1 1 0 1 1 1  
## 1 1 1 0 1 1 1  
## 2 1 1 1 1 1 1  
## 1 1 1 1 1 1 0  
## 1 1 1 1 1 1 0  
## 1 1 1 1 1 0 1  
## 1 1 1 1 1 1 1  
## 1 0 0 1 1 1 1  
## 1 1 1 0 1 1 1  
## 1 1 1 0 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 0 1 1 0 0  
## 1 0 0 1 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 0 1 1 1  
## 1 1 1 0 1 1 1  
## 1 1 1 1 1 1 1  
## 1 1 1 0 1 1 1  
## 26 27 30 32 33 38  
## PC.Volume PSC.CO2 Filler.Speed Brand.Code MFR   
## 2038 1 1 1 1 1 0  
## 91 1 1 1 0 1 1  
## 1 1 1 1 1 1 1  
## 18 1 1 1 1 1 1  
## 17 0 1 1 1 1 1  
## 9 1 1 1 1 1 1  
## 10 1 1 1 1 1 1  
## 15 1 1 1 1 1 1  
## 9 1 1 1 1 1 1  
## 18 1 0 1 1 1 1  
## 27 1 1 1 1 1 1  
## 5 1 1 1 1 1 1  
## 3 1 1 1 1 1 1  
## 2 1 1 0 1 1 1  
## 2 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 100 1 1 1 1 0 1  
## 1 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 7 1 1 1 1 1 1  
## 4 1 1 1 1 1 1  
## 2 1 1 1 1 1 1  
## 3 1 1 1 0 1 2  
## 3 0 1 1 0 1 2  
## 2 0 1 1 1 1 2  
## 1 1 1 1 1 1 2  
## 1 0 1 1 1 1 2  
## 2 1 1 1 0 1 2  
## 1 1 1 1 1 1 2  
## 7 1 1 1 1 1 2  
## 1 1 1 1 0 1 2  
## 3 0 1 1 1 1 2  
## 2 1 1 1 1 1 2  
## 2 1 1 1 1 1 2  
## 1 1 0 1 0 1 2  
## 1 1 0 1 1 1 2  
## 1 0 0 1 1 1 2  
## 1 1 0 1 1 1 2  
## 1 1 0 1 1 1 2  
## 1 1 0 1 1 1 2  
## 2 0 1 1 1 1 2  
## 1 1 1 1 1 1 2  
## 3 1 1 1 1 1 2  
## 4 1 1 1 1 1 2  
## 1 1 1 0 1 1 2  
## 1 1 0 1 1 1 2  
## 3 1 1 1 0 0 2  
## 2 1 1 1 1 0 2  
## 2 0 1 1 1 0 2  
## 2 1 1 1 1 0 2  
## 1 1 1 1 1 0 2  
## 4 1 1 1 1 0 2  
## 1 1 1 1 1 0 2  
## 4 1 0 1 1 0 2  
## 1 1 1 1 1 0 2  
## 10 1 1 1 1 0 2  
## 30 1 1 0 1 0 2  
## 4 1 1 1 1 0 2  
## 1 1 1 1 1 0 2  
## 4 1 1 1 1 0 2  
## 1 0 1 1 1 1 2  
## 1 1 0 1 1 1 2  
## 1 1 1 1 1 1 2  
## 1 1 1 1 1 0 2  
## 1 1 1 1 0 1 3  
## 1 0 1 1 0 1 3  
## 1 1 1 1 0 1 3  
## 1 1 0 1 1 1 3  
## 2 1 0 1 1 1 3  
## 1 1 0 1 0 1 3  
## 1 1 0 1 1 1 3  
## 1 0 1 1 1 1 3  
## 1 1 1 1 0 1 3  
## 1 0 1 1 1 1 3  
## 9 1 1 1 1 1 3  
## 1 1 1 1 1 1 3  
## 1 1 1 1 1 0 3  
## 1 1 1 1 0 0 3  
## 1 1 1 1 1 0 3  
## 1 1 1 1 1 0 3  
## 3 1 1 0 0 0 3  
## 1 1 1 0 1 0 3  
## 1 0 1 0 1 0 3  
## 1 1 1 0 1 0 3  
## 2 1 1 0 1 0 3  
## 2 1 1 0 1 0 3  
## 2 1 1 0 1 0 3  
## 1 1 1 1 0 0 3  
## 2 1 1 0 1 0 3  
## 2 1 1 1 1 0 3  
## 2 1 1 1 1 0 3  
## 1 1 1 1 0 0 3  
## 1 1 0 1 1 1 4  
## 1 0 1 1 1 1 4  
## 1 1 1 1 1 1 4  
## 1 1 1 1 0 1 4  
## 1 0 1 1 1 0 4  
## 1 1 1 1 0 0 4  
## 1 1 1 1 0 0 4  
## 2 1 1 0 1 0 4  
## 1 1 1 0 1 0 4  
## 1 0 1 1 1 0 4  
## 1 1 1 0 1 0 4  
## 1 1 1 0 1 0 4  
## 1 1 0 1 1 1 5  
## 1 1 1 1 1 0 5  
## 1 1 1 0 1 0 5  
## 1 1 1 1 0 0 5  
## 1 1 0 1 1 1 6  
## 1 1 0 1 1 0 6  
## 1 1 1 0 1 0 6  
## 1 1 1 1 0 0 7  
## 1 1 1 0 1 0 8  
## 1 1 1 0 1 0 9  
## 1 1 1 0 1 0 12  
## 39 39 57 120 212 844

aggr\_plot <- aggr(df, col=c('navyblue','red'), numbers=TRUE, sortVars=TRUE, labels=names(data), cex.axis=.7, gap=3, ylab=c("Histogram of missing data","Pattern"))

## Warning in plot.aggr(res, ...): not enough vertical space to display  
## frequencies (too many combinations)



##   
## Variables sorted by number of missings:   
## Variable Count  
## MFR 0.0824581875  
## Brand.Code 0.0466744457  
## Filler.Speed 0.0221703617  
## PC.Volume 0.0151691949  
## PSC.CO2 0.0151691949  
## Fill.Ounces 0.0147802412  
## PSC 0.0128354726  
## Carb.Pressure1 0.0124465189  
## Hyd.Pressure4 0.0116686114  
## Carb.Pressure 0.0105017503  
## Carb.Temp 0.0101127966  
## PSC.Fill 0.0089459354  
## Fill.Pressure 0.0085569817  
## Filler.Level 0.0077790743  
## Hyd.Pressure2 0.0058343057  
## Hyd.Pressure3 0.0058343057  
## Temperature 0.0054453520  
## Oxygen.Filler 0.0046674446  
## Pressure.Setpoint 0.0046674446  
## Hyd.Pressure1 0.0042784909  
## Carb.Volume 0.0038895371  
## Carb.Rel 0.0038895371  
## Alch.Rel 0.0035005834  
## Usage.cont 0.0019447686  
## PH 0.0015558149  
## Mnf.Flow 0.0007779074  
## Carb.Flow 0.0007779074  
## Bowl.Setpoint 0.0007779074  
## Density 0.0003889537  
## Balling 0.0003889537  
## Balling.Lvl 0.0003889537  
## Pressure.Vacuum 0.0000000000  
## Air.Pressurer 0.0000000000

#No missing elements  
#https://davetang.org/muse/2013/05/22/using-aggregate-and-apply-in-r/  
#https://www.r-bloggers.com/imputing-missing-data-with-r-mice-package/

#Review the correlations  
df\_cor <- df  
df\_cor$Brand.Code <- NULL #Need to remove because a categorical variable  
df\_cor$PH <- NULL #Need to remove because our dependent variable  
df\_cor <- df\_cor[complete.cases(df\_cor),]   
correlations <- cor(df\_cor)  
#corrplot(correlations, order = "hclust")  
  
class(correlations)

## [1] "matrix"

corr\_mat=cor(correlations,method="p")  
#corr\_mat[1:35,1:35]  
  
#corrplot(correlations, order="hclust")  
#corrplot(correlations, method="square")  
  
#Identify high correlations  
  
#threshold <- 0.8  
#tooHigh <- findCorrelation(correlations, cutoff = threshold, names = TRUE, verbose = TRUE)   
#tooHigh  
# a few correlations that we shall consider removing those columns  
  
#Take a look at distributions to identify outliers

# Explore the distributions of the remaining columns with fairly significant NAs to determine appropriate

# imputations

df = read.csv(file = "F:/HW 2/StudentData.csv",   
 na.strings = c("", " "),   
 header = TRUE)   
library(Hmisc)  
head(df)

## Brand.Code Carb.Volume Fill.Ounces PC.Volume Carb.Pressure Carb.Temp  
## 1 B 5.340000 23.96667 0.2633333 68.2 141.2  
## 2 A 5.426667 24.00667 0.2386667 68.4 139.6  
## 3 B 5.286667 24.06000 0.2633333 70.8 144.8  
## 4 A 5.440000 24.00667 0.2933333 63.0 132.6  
## 5 A 5.486667 24.31333 0.1113333 67.2 136.8  
## 6 A 5.380000 23.92667 0.2693333 66.6 138.4  
## PSC PSC.Fill PSC.CO2 Mnf.Flow Carb.Pressure1 Fill.Pressure  
## 1 0.104 0.26 0.04 -100 118.8 46.0  
## 2 0.124 0.22 0.04 -100 121.6 46.0  
## 3 0.090 0.34 0.16 -100 120.2 46.0  
## 4 NA 0.42 0.04 -100 115.2 46.4  
## 5 0.026 0.16 0.12 -100 118.4 45.8  
## 6 0.090 0.24 0.04 -100 119.6 45.6  
## Hyd.Pressure1 Hyd.Pressure2 Hyd.Pressure3 Hyd.Pressure4 Filler.Level  
## 1 0 NA NA 118 121.2  
## 2 0 NA NA 106 118.6  
## 3 0 NA NA 82 120.0  
## 4 0 0 0 92 117.8  
## 5 0 0 0 92 118.6  
## 6 0 0 0 116 120.2  
## Filler.Speed Temperature Usage.cont Carb.Flow Density MFR Balling  
## 1 4002 66.0 16.18 2932 0.88 725.0 1.398  
## 2 3986 67.6 19.90 3144 0.92 726.8 1.498  
## 3 4020 67.0 17.76 2914 1.58 735.0 3.142  
## 4 4012 65.6 17.42 3062 1.54 730.6 3.042  
## 5 4010 65.6 17.68 3054 1.54 722.8 3.042  
## 6 4014 66.2 23.82 2948 1.52 738.8 2.992  
## Pressure.Vacuum PH Oxygen.Filler Bowl.Setpoint Pressure.Setpoint  
## 1 -4.0 8.36 0.022 120 46.4  
## 2 -4.0 8.26 0.026 120 46.8  
## 3 -3.8 8.94 0.024 120 46.6  
## 4 -4.4 8.24 0.030 120 46.0  
## 5 -4.4 8.26 0.030 120 46.0  
## 6 -4.4 8.32 0.024 120 46.0  
## Air.Pressurer Alch.Rel Carb.Rel Balling.Lvl  
## 1 142.6 6.58 5.32 1.48  
## 2 143.0 6.56 5.30 1.56  
## 3 142.0 7.66 5.84 3.28  
## 4 146.2 7.14 5.42 3.04  
## 5 146.2 7.14 5.44 3.04  
## 6 146.6 7.16 5.44 3.02

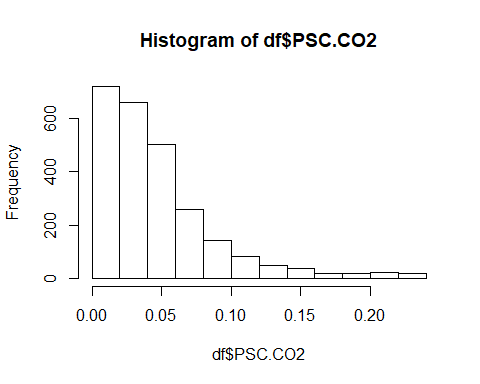
summary(df)

## Brand.Code Carb.Volume Fill.Ounces PC.Volume   
## A : 293 Min. :5.040 Min. :23.63 Min. :0.07933   
## B :1239 1st Qu.:5.293 1st Qu.:23.92 1st Qu.:0.23917   
## C : 304 Median :5.347 Median :23.97 Median :0.27133   
## D : 615 Mean :5.370 Mean :23.97 Mean :0.27712   
## NA's: 120 3rd Qu.:5.453 3rd Qu.:24.03 3rd Qu.:0.31200   
## Max. :5.700 Max. :24.32 Max. :0.47800   
## NA's :10 NA's :38 NA's :39   
## Carb.Pressure Carb.Temp PSC PSC.Fill   
## Min. :57.00 Min. :128.6 Min. :0.00200 Min. :0.0000   
## 1st Qu.:65.60 1st Qu.:138.4 1st Qu.:0.04800 1st Qu.:0.1000   
## Median :68.20 Median :140.8 Median :0.07600 Median :0.1800   
## Mean :68.19 Mean :141.1 Mean :0.08457 Mean :0.1954   
## 3rd Qu.:70.60 3rd Qu.:143.8 3rd Qu.:0.11200 3rd Qu.:0.2600   
## Max. :79.40 Max. :154.0 Max. :0.27000 Max. :0.6200   
## NA's :27 NA's :26 NA's :33 NA's :23   
## PSC.CO2 Mnf.Flow Carb.Pressure1 Fill.Pressure   
## Min. :0.00000 Min. :-100.20 Min. :105.6 Min. :34.60   
## 1st Qu.:0.02000 1st Qu.:-100.00 1st Qu.:119.0 1st Qu.:46.00   
## Median :0.04000 Median : 65.20 Median :123.2 Median :46.40   
## Mean :0.05641 Mean : 24.57 Mean :122.6 Mean :47.92   
## 3rd Qu.:0.08000 3rd Qu.: 140.80 3rd Qu.:125.4 3rd Qu.:50.00   
## Max. :0.24000 Max. : 229.40 Max. :140.2 Max. :60.40   
## NA's :39 NA's :2 NA's :32 NA's :22   
## Hyd.Pressure1 Hyd.Pressure2 Hyd.Pressure3 Hyd.Pressure4   
## Min. :-0.80 Min. : 0.00 Min. :-1.20 Min. : 52.00   
## 1st Qu.: 0.00 1st Qu.: 0.00 1st Qu.: 0.00 1st Qu.: 86.00   
## Median :11.40 Median :28.60 Median :27.60 Median : 96.00   
## Mean :12.44 Mean :20.96 Mean :20.46 Mean : 96.29   
## 3rd Qu.:20.20 3rd Qu.:34.60 3rd Qu.:33.40 3rd Qu.:102.00   
## Max. :58.00 Max. :59.40 Max. :50.00 Max. :142.00   
## NA's :11 NA's :15 NA's :15 NA's :30   
## Filler.Level Filler.Speed Temperature Usage.cont   
## Min. : 55.8 Min. : 998 Min. :63.60 Min. :12.08   
## 1st Qu.: 98.3 1st Qu.:3888 1st Qu.:65.20 1st Qu.:18.36   
## Median :118.4 Median :3982 Median :65.60 Median :21.79   
## Mean :109.3 Mean :3687 Mean :65.97 Mean :20.99   
## 3rd Qu.:120.0 3rd Qu.:3998 3rd Qu.:66.40 3rd Qu.:23.75   
## Max. :161.2 Max. :4030 Max. :76.20 Max. :25.90   
## NA's :20 NA's :57 NA's :14 NA's :5   
## Carb.Flow Density MFR Balling   
## Min. : 26 Min. :0.240 Min. : 31.4 Min. :-0.170   
## 1st Qu.:1144 1st Qu.:0.900 1st Qu.:706.3 1st Qu.: 1.496   
## Median :3028 Median :0.980 Median :724.0 Median : 1.648   
## Mean :2468 Mean :1.174 Mean :704.0 Mean : 2.198   
## 3rd Qu.:3186 3rd Qu.:1.620 3rd Qu.:731.0 3rd Qu.: 3.292   
## Max. :5104 Max. :1.920 Max. :868.6 Max. : 4.012   
## NA's :2 NA's :1 NA's :212 NA's :1   
## Pressure.Vacuum PH Oxygen.Filler Bowl.Setpoint   
## Min. :-6.600 Min. :7.880 Min. :0.00240 Min. : 70.0   
## 1st Qu.:-5.600 1st Qu.:8.440 1st Qu.:0.02200 1st Qu.:100.0   
## Median :-5.400 Median :8.540 Median :0.03340 Median :120.0   
## Mean :-5.216 Mean :8.546 Mean :0.04684 Mean :109.3   
## 3rd Qu.:-5.000 3rd Qu.:8.680 3rd Qu.:0.06000 3rd Qu.:120.0   
## Max. :-3.600 Max. :9.360 Max. :0.40000 Max. :140.0   
## NA's :4 NA's :12 NA's :2   
## Pressure.Setpoint Air.Pressurer Alch.Rel Carb.Rel   
## Min. :44.00 Min. :140.8 Min. :5.280 Min. :4.960   
## 1st Qu.:46.00 1st Qu.:142.2 1st Qu.:6.540 1st Qu.:5.340   
## Median :46.00 Median :142.6 Median :6.560 Median :5.400   
## Mean :47.62 Mean :142.8 Mean :6.897 Mean :5.437   
## 3rd Qu.:50.00 3rd Qu.:143.0 3rd Qu.:7.240 3rd Qu.:5.540   
## Max. :52.00 Max. :148.2 Max. :8.620 Max. :6.060   
## NA's :12 NA's :9 NA's :10   
## Balling.Lvl   
## Min. :0.00   
## 1st Qu.:1.38   
## Median :1.48   
## Mean :2.05   
## 3rd Qu.:3.14   
## Max. :3.66   
## NA's :1

summary(df$PSC.CO2)

## Min. 1st Qu. Median Mean 3rd Qu. Max. NA's   
## 0.00000 0.02000 0.04000 0.05641 0.08000 0.24000 39

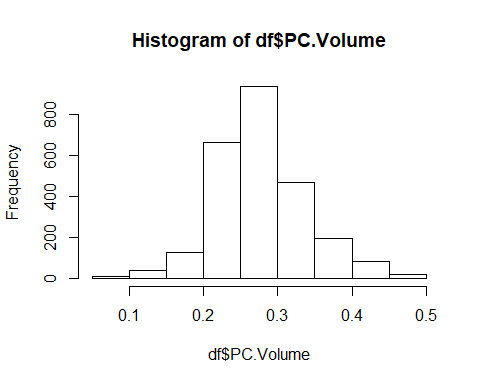
hist(df$PSC.CO2) #left skewed



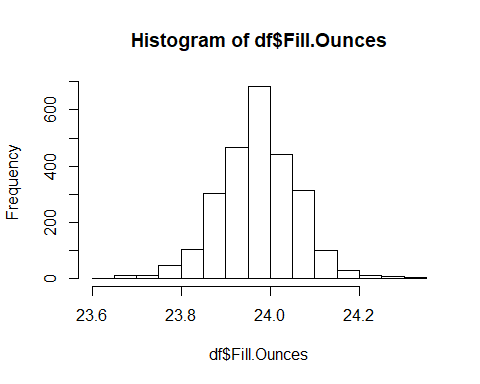
df$PSC.CO2 <- impute(as.matrix(df$PSC.CO2), "median")  
table(df$PSC.CO2)

##   
## 0 0.02 0.04 0.06 0.08 0.1 0.12 0.14 0.16 0.18 0.2 0.22 0.24   
## 108 613 700 502 257 144 82 49 39 20 18 22 17

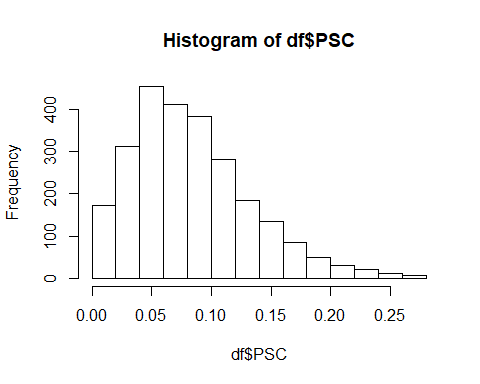
hist(df$PC.Volume) #Fairly normal distribution so mean should suffice



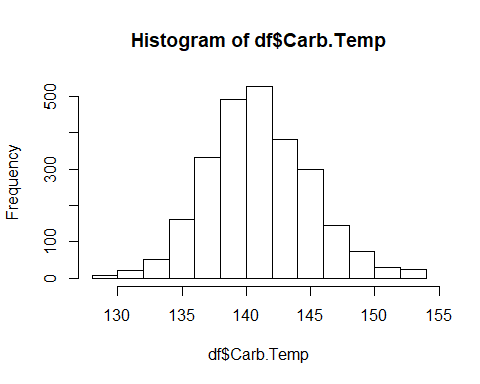
df$PC.Volume <- impute(as.matrix(df$PC.Volume), "mean")  
  
hist(df$Fill.Ounces) #Normal, leverage the mean



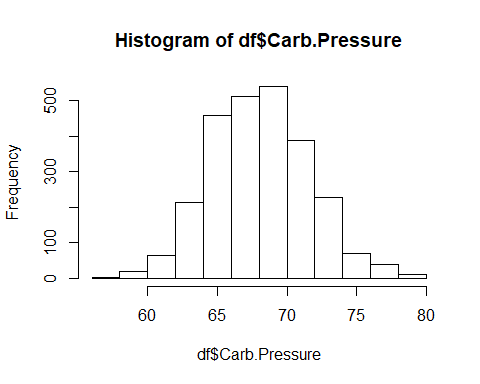
df$Fill.Ounces <- impute(as.matrix(df$Fill.Ounces), "mean")  
  
hist(df$PSC) #left skewed, so opt for the median



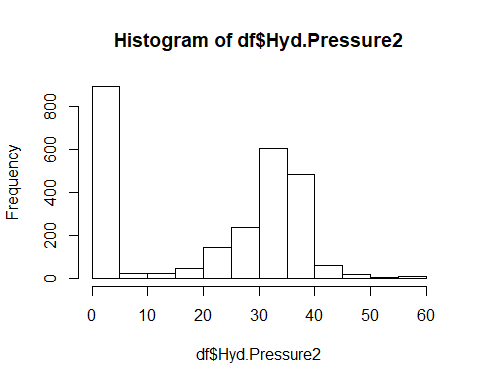
df$PSC <- impute(as.matrix(df$PSC), "median")  
  
hist(df$Carb.Temp) #Normal, leverage the mean



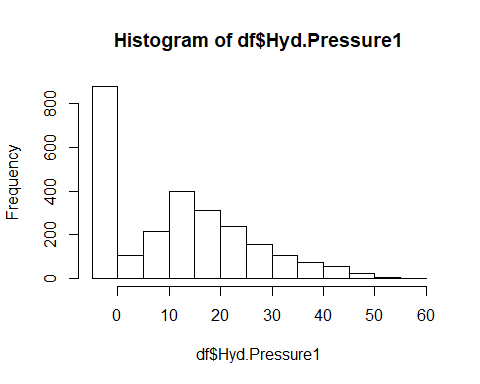
df$Carb.Temp <- impute(as.matrix(df$Carb.Temp), "median")  
  
hist(df$Carb.Pressure) #normal



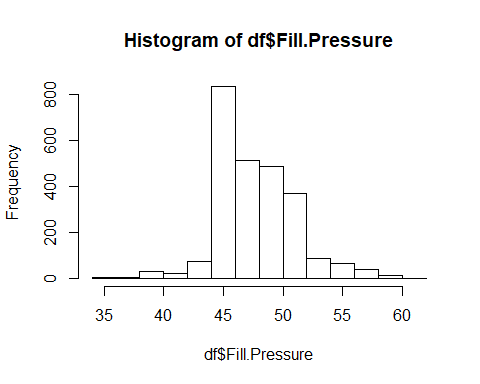
df$Carb.Pressure <- impute(as.matrix(df$Carb.Pressure), "mean")  
  
hist(df$Hyd.Pressure2) #a lot of zero values which seem problematic - use median, may be smarter to remove the 14 rows



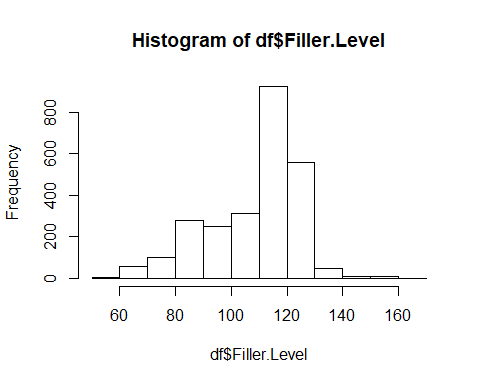
df$Hyd.Pressure2 <- impute(as.matrix(df$Hyd.Pressure2), "median")  
  
hist(df$Hyd.Pressure1) #also a lot of zero values



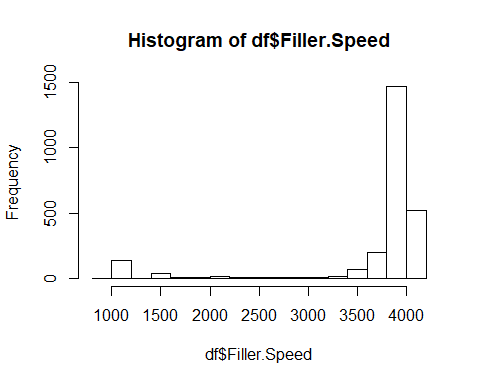
df$Hyd.Pressure1 <- impute(as.matrix(df$Hyd.Pressure1), "median")  
  
hist(df$Fill.Pressure) #right skewed



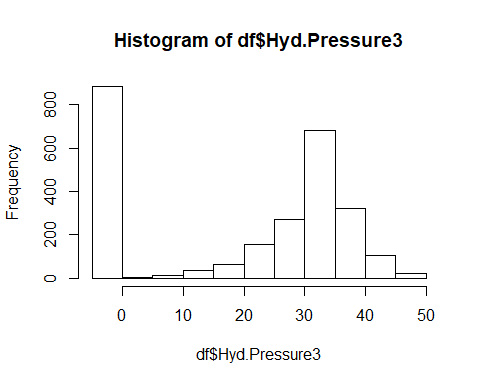
df$Fill.Pressure <- impute(as.matrix(df$Fill.Pressure), "median")  
  
hist(df$Filler.Level) #right skewed



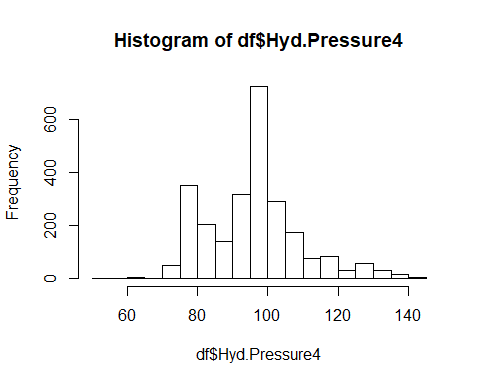
df$Filler.Level <- impute(as.matrix(df$Filler.Level), "median")  
  
hist(df$Filler.Speed) #far right skewed, may be wiser to simply remove rows, 3



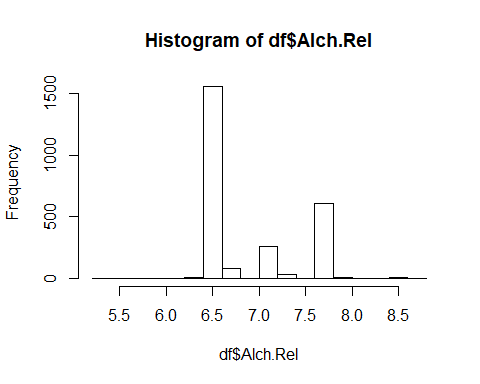
df$Filler.Speed <- impute(as.matrix(df$Filler.Speed), "median")  
  
#hist(df$Carb.Pressure) #normal  
#df$Carb.Pressure <- impute(as.matrix(df$Carb.Pressure), "mean")  
  
hist(df$Hyd.Pressure3) #a low of zeros and then right skewed



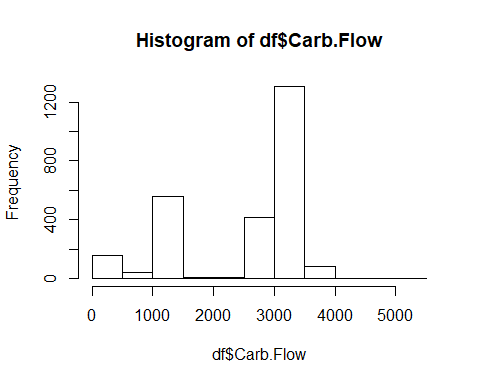
df$Hyd.Pressure3 <- impute(as.matrix(df$Hyd.Pressure3), "median")  
  
hist(df$Hyd.Pressure4) #normal



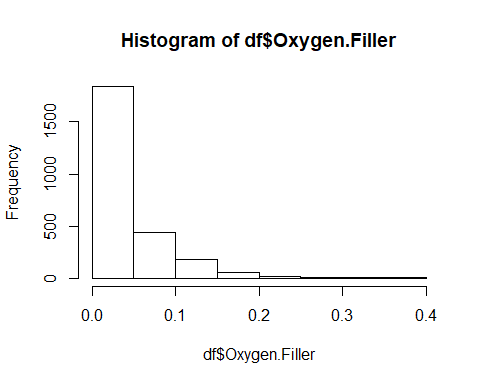
df$Hyd.Pressure4 <- impute(as.matrix(df$Hyd.Pressure4), "mean")  
  
hist(df$Alch.Rel) #odd distribution - median best, but may prefer to remove rows



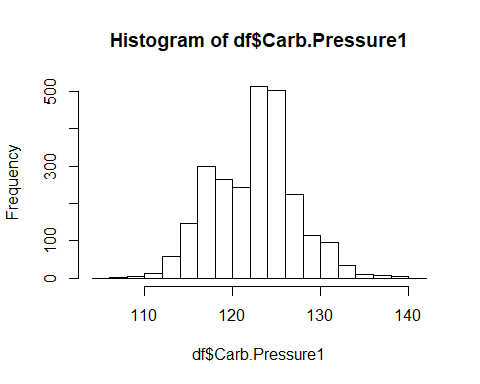
df$Alch.Rel <- impute(as.matrix(df$Alch.Rel), "median")  
  
hist(df$Carb.Flow) #odd distribution - median best, but may prefer to remove rows



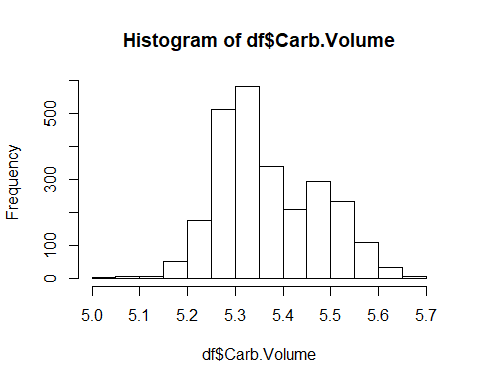
df$Carb.Flow <- impute(as.matrix(df$Carb.Flow), "median")  
  
hist(df$Oxygen.Filler) #left skewed



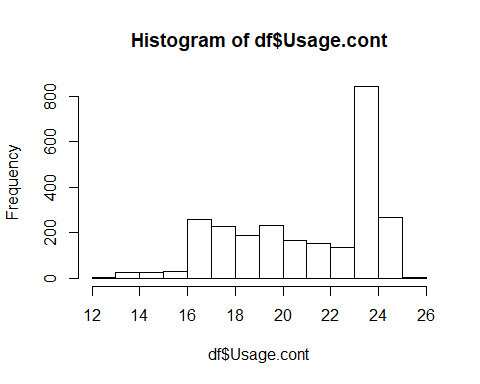
df$Oxygen.Filler <- impute(as.matrix(df$Oxygen.Filler), "median")  
  
hist(df$Carb.Pressure1) #right skewed



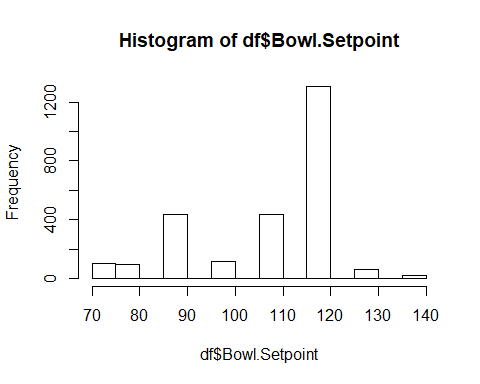
df$Carb.Pressure1 <- impute(as.matrix(df$Carb.Pressure1), "median")  
  
hist(df$Carb.Volume) #left skewed



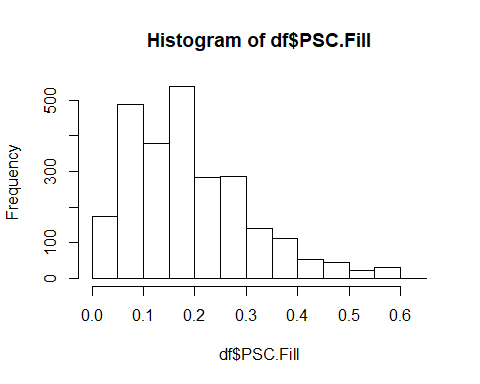
df$Carb.Volume <- impute(as.matrix(df$Carb.Volume), "median")  
  
hist(df$Usage.cont) #use median



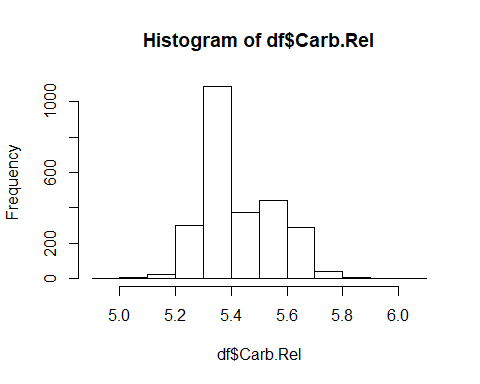
df$Usage.cont <- impute(as.matrix(df$Usage.cont), "median")  
  
hist(df$Bowl.Setpoint) #use median



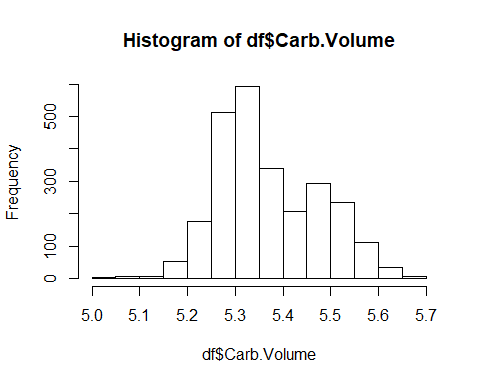
df$Bowl.Setpoint <- impute(as.matrix(df$Bowl.Setpoint), "median")  
  
hist(df$PSC.Fill)



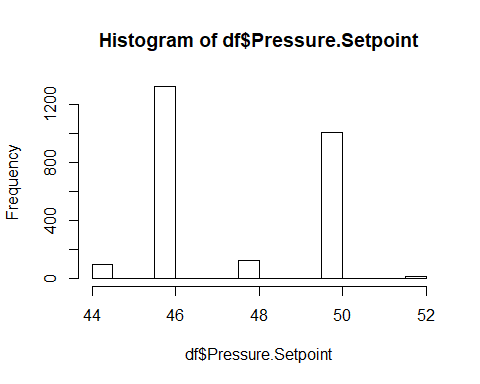
df$PSC.Fill <- impute(as.matrix(df$PSC.Fill), "median")  
  
hist(df$Carb.Rel)



df$Carb.Rel <- impute(as.matrix(df$Carb.Rel), "median")  
  
hist(df$Carb.Volume)



df$Carb.Volume <- impute(as.matrix(df$Carb.Volume), "median")  
  
hist(df$Pressure.Setpoint)

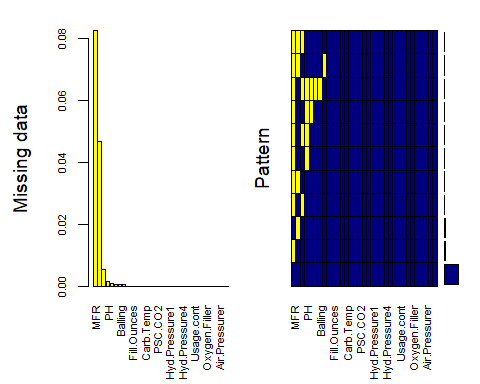


df$Pressure.Setpoint <- impute(as.matrix(df$Pressure.Setpoint), "median")

# Let’s check again for NAs and see the severity

library(VIM)  
mice\_plot <- aggr(df, col=c('navyblue','yellow'),  
 numbers=TRUE, sortVars=TRUE,  
 labels=names(df), cex.axis=.7,  
 gap=3, ylab=c("Missing data","Pattern"))

## Warning in plot.aggr(res, ...): not enough horizontal space to display  
## frequencies



##   
## Variables sorted by number of missings:   
## Variable Count  
## MFR 0.0824581875  
## Brand.Code 0.0466744457  
## Temperature 0.0054453520  
## PH 0.0015558149  
## Mnf.Flow 0.0007779074  
## Density 0.0003889537  
## Balling 0.0003889537  
## Balling.Lvl 0.0003889537  
## Carb.Volume 0.0000000000  
## Fill.Ounces 0.0000000000  
## PC.Volume 0.0000000000  
## Carb.Pressure 0.0000000000  
## Carb.Temp 0.0000000000  
## PSC 0.0000000000  
## PSC.Fill 0.0000000000  
## PSC.CO2 0.0000000000  
## Carb.Pressure1 0.0000000000  
## Fill.Pressure 0.0000000000  
## Hyd.Pressure1 0.0000000000  
## Hyd.Pressure2 0.0000000000  
## Hyd.Pressure3 0.0000000000  
## Hyd.Pressure4 0.0000000000  
## Filler.Level 0.0000000000  
## Filler.Speed 0.0000000000  
## Usage.cont 0.0000000000  
## Carb.Flow 0.0000000000  
## Pressure.Vacuum 0.0000000000  
## Oxygen.Filler 0.0000000000  
## Bowl.Setpoint 0.0000000000  
## Pressure.Setpoint 0.0000000000  
## Air.Pressurer 0.0000000000  
## Alch.Rel 0.0000000000  
## Carb.Rel 0.0000000000

# Review the correlations

library(corrplot)  
df\_cor <- df  
df\_cor$Brand.Code <- NULL #Need to remove because a categorical variable  
df\_cor$PH <- NULL #Need to remove because our dependent variable  
#correlations <- cor(df\_cor)  
#corrplot(correlations)  
#corrplot(correlations, order = "hclust")

# Identfiy high correlations

#threshold <- 0.8  
#tooHigh <- findCorrelation(correlations, cutoff = threshold, names = TRUE, verbose = TRUE)   
#tooHigh #eight columns - we should remove

# REMEMBER THIS FOR EDITING THE TEST DATA!!

df\_cor$Balling <- NULL  
df\_cor$Hyd.Pressure3 <- NULL  
df\_cor$Alch.Rel <- NULL  
df\_cor$Balling.Lvl <- NULL  
df\_cor$Density <- NULL  
df\_cor$Density <- NULL  
df\_cor$Carb.Volume <- NULL  
df\_cor$Bowl.Setpoint <- NULL  
df\_cor$Filler.Speed <- NULL

# a few correlations that we shall consider removing those columns

# boxplots and histograms to check distributions

library(ggplot2)  
d <- melt(df\_cor)

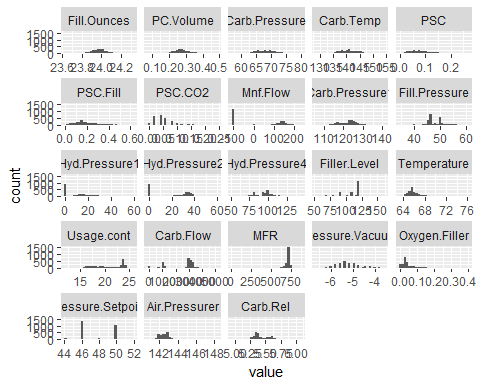
## No id variables; using all as measure variables

## Warning: attributes are not identical across measure variables; they will  
## be dropped

ggplot(d,aes(x = value)) +   
 facet\_wrap(~variable,scales = "free\_x") +   
 geom\_histogram()

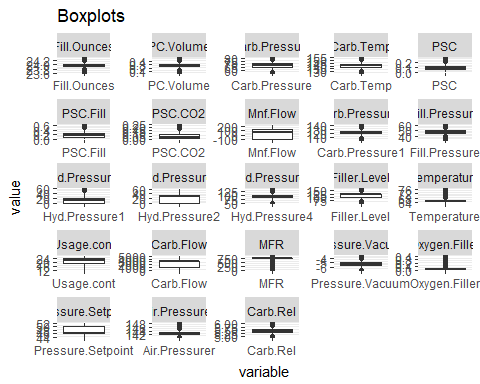
## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

## Warning: Removed 228 rows containing non-finite values (stat\_bin).

 # Zero values for some variables still need to be considered: Mnf.Flow, Hyd.Pressure1 &2

ggplot(d, aes(x=variable, y=value)) +  
 facet\_wrap(~variable,scales = "free") +  
 geom\_boxplot() +  
 ggtitle("Boxplots")

## Warning: Removed 228 rows containing non-finite values (stat\_boxplot).

 #All the more reason to consider doing something with the zeros…

# Check for skewness

#skewValues <- apply(df\_cor, 2, skewness)   
#View(skewValues) #quite a lot of high, negative values likely due to the zeros

# backwards stepwise regression

df = read.csv(file = "F:/HW 2/StudentData.csv",   
 na.strings = c("", " "),   
 header = TRUE)   
head(df)

## Brand.Code Carb.Volume Fill.Ounces PC.Volume Carb.Pressure Carb.Temp  
## 1 B 5.340000 23.96667 0.2633333 68.2 141.2  
## 2 A 5.426667 24.00667 0.2386667 68.4 139.6  
## 3 B 5.286667 24.06000 0.2633333 70.8 144.8  
## 4 A 5.440000 24.00667 0.2933333 63.0 132.6  
## 5 A 5.486667 24.31333 0.1113333 67.2 136.8  
## 6 A 5.380000 23.92667 0.2693333 66.6 138.4  
## PSC PSC.Fill PSC.CO2 Mnf.Flow Carb.Pressure1 Fill.Pressure  
## 1 0.104 0.26 0.04 -100 118.8 46.0  
## 2 0.124 0.22 0.04 -100 121.6 46.0  
## 3 0.090 0.34 0.16 -100 120.2 46.0  
## 4 NA 0.42 0.04 -100 115.2 46.4  
## 5 0.026 0.16 0.12 -100 118.4 45.8  
## 6 0.090 0.24 0.04 -100 119.6 45.6  
## Hyd.Pressure1 Hyd.Pressure2 Hyd.Pressure3 Hyd.Pressure4 Filler.Level  
## 1 0 NA NA 118 121.2  
## 2 0 NA NA 106 118.6  
## 3 0 NA NA 82 120.0  
## 4 0 0 0 92 117.8  
## 5 0 0 0 92 118.6  
## 6 0 0 0 116 120.2  
## Filler.Speed Temperature Usage.cont Carb.Flow Density MFR Balling  
## 1 4002 66.0 16.18 2932 0.88 725.0 1.398  
## 2 3986 67.6 19.90 3144 0.92 726.8 1.498  
## 3 4020 67.0 17.76 2914 1.58 735.0 3.142  
## 4 4012 65.6 17.42 3062 1.54 730.6 3.042  
## 5 4010 65.6 17.68 3054 1.54 722.8 3.042  
## 6 4014 66.2 23.82 2948 1.52 738.8 2.992  
## Pressure.Vacuum PH Oxygen.Filler Bowl.Setpoint Pressure.Setpoint  
## 1 -4.0 8.36 0.022 120 46.4  
## 2 -4.0 8.26 0.026 120 46.8  
## 3 -3.8 8.94 0.024 120 46.6  
## 4 -4.4 8.24 0.030 120 46.0  
## 5 -4.4 8.26 0.030 120 46.0  
## 6 -4.4 8.32 0.024 120 46.0  
## Air.Pressurer Alch.Rel Carb.Rel Balling.Lvl  
## 1 142.6 6.58 5.32 1.48  
## 2 143.0 6.56 5.30 1.56  
## 3 142.0 7.66 5.84 3.28  
## 4 146.2 7.14 5.42 3.04  
## 5 146.2 7.14 5.44 3.04  
## 6 146.6 7.16 5.44 3.02

names(df)

## [1] "Brand.Code" "Carb.Volume" "Fill.Ounces"   
## [4] "PC.Volume" "Carb.Pressure" "Carb.Temp"   
## [7] "PSC" "PSC.Fill" "PSC.CO2"   
## [10] "Mnf.Flow" "Carb.Pressure1" "Fill.Pressure"   
## [13] "Hyd.Pressure1" "Hyd.Pressure2" "Hyd.Pressure3"   
## [16] "Hyd.Pressure4" "Filler.Level" "Filler.Speed"   
## [19] "Temperature" "Usage.cont" "Carb.Flow"   
## [22] "Density" "MFR" "Balling"   
## [25] "Pressure.Vacuum" "PH" "Oxygen.Filler"   
## [28] "Bowl.Setpoint" "Pressure.Setpoint" "Air.Pressurer"   
## [31] "Alch.Rel" "Carb.Rel" "Balling.Lvl"

lreg<- glm(df$Brand.Code~.,data=df,family="binomial")  
summary(lreg)

##   
## Call:  
## glm(formula = df$Brand.Code ~ ., family = "binomial", data = df)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -4.2294 0.0008 0.0040 0.0209 4.0795   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -4.330e+02 1.778e+02 -2.436 0.014866 \*   
## Carb.Volume 2.815e+01 1.406e+01 2.002 0.045307 \*   
## Fill.Ounces 3.170e+00 4.126e+00 0.768 0.442272   
## PC.Volume 7.793e+00 7.170e+00 1.087 0.277078   
## Carb.Pressure -9.386e-01 7.388e-01 -1.271 0.203888   
## Carb.Temp 7.740e-01 5.880e-01 1.316 0.188045   
## PSC 5.000e+00 7.588e+00 0.659 0.509955   
## PSC.Fill -7.073e-01 2.625e+00 -0.269 0.787623   
## PSC.CO2 -6.544e+00 8.628e+00 -0.758 0.448211   
## Mnf.Flow -2.423e-02 6.980e-03 -3.471 0.000518 \*\*\*  
## Carb.Pressure1 1.451e-01 1.089e-01 1.333 0.182664   
## Fill.Pressure 4.719e-01 2.194e-01 2.151 0.031498 \*   
## Hyd.Pressure1 -2.815e-02 4.637e-02 -0.607 0.543792   
## Hyd.Pressure2 2.097e-02 7.397e-02 0.283 0.776850   
## Hyd.Pressure3 -6.958e-02 7.082e-02 -0.983 0.325840   
## Hyd.Pressure4 -1.167e-01 4.676e-02 -2.495 0.012590 \*   
## Filler.Level 2.175e-02 9.933e-02 0.219 0.826649   
## Filler.Speed -2.609e-03 1.669e-03 -1.563 0.118158   
## Temperature 2.127e-01 3.642e-01 0.584 0.559152   
## Usage.cont 2.506e-01 1.460e-01 1.717 0.086038 .   
## Carb.Flow -3.467e-03 7.956e-04 -4.357 1.32e-05 \*\*\*  
## Density 9.334e-01 3.106e+00 0.300 0.763810   
## MFR 1.621e-02 9.413e-03 1.722 0.085015 .   
## Balling 3.899e+00 4.610e+00 0.846 0.397777   
## Pressure.Vacuum -3.582e-01 1.437e+00 -0.249 0.803197   
## PH 1.478e+00 3.399e+00 0.435 0.663629   
## Oxygen.Filler 1.210e+01 1.025e+01 1.181 0.237745   
## Bowl.Setpoint -4.728e-02 1.002e-01 -0.472 0.637021   
## Pressure.Setpoint -2.221e-01 2.615e-01 -0.849 0.395726   
## Air.Pressurer 1.127e-01 3.516e-01 0.321 0.748542   
## Alch.Rel 2.608e+01 4.188e+00 6.228 4.72e-10 \*\*\*  
## Carb.Rel -5.049e+00 6.651e+00 -0.759 0.447846   
## Balling.Lvl -2.492e+01 5.346e+00 -4.662 3.13e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 1411.644 on 2037 degrees of freedom  
## Residual deviance: 93.228 on 2005 degrees of freedom  
## (533 observations deleted due to missingness)  
## AIC: 159.23  
##   
## Number of Fisher Scoring iterations: 11

lreg$coefficients

## (Intercept) Carb.Volume Fill.Ounces PC.Volume   
## -4.329978e+02 2.815441e+01 3.169907e+00 7.792588e+00   
## Carb.Pressure Carb.Temp PSC PSC.Fill   
## -9.386309e-01 7.739700e-01 4.999992e+00 -7.072880e-01   
## PSC.CO2 Mnf.Flow Carb.Pressure1 Fill.Pressure   
## -6.543532e+00 -2.422767e-02 1.451407e-01 4.719329e-01   
## Hyd.Pressure1 Hyd.Pressure2 Hyd.Pressure3 Hyd.Pressure4   
## -2.814773e-02 2.096573e-02 -6.958351e-02 -1.166683e-01   
## Filler.Level Filler.Speed Temperature Usage.cont   
## 2.175244e-02 -2.608699e-03 2.127322e-01 2.506088e-01   
## Carb.Flow Density MFR Balling   
## -3.466593e-03 9.334085e-01 1.621173e-02 3.898585e+00   
## Pressure.Vacuum PH Oxygen.Filler Bowl.Setpoint   
## -3.581516e-01 1.478199e+00 1.209600e+01 -4.728159e-02   
## Pressure.Setpoint Air.Pressurer Alch.Rel Carb.Rel   
## -2.220743e-01 1.127068e-01 2.608173e+01 -5.048555e+00   
## Balling.Lvl   
## -2.492427e+01

#measure the 95% confidence intervals for the 40 set data  
confint(lreg, level=0.95)

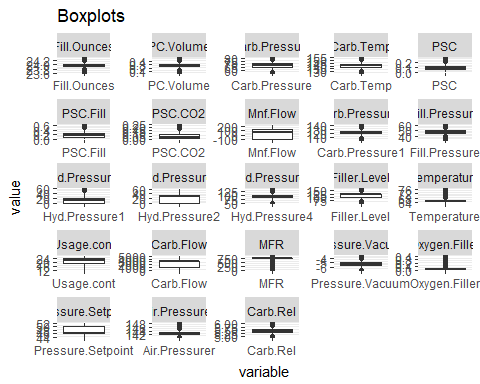
## Waiting for profiling to be done...

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred  
  
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## 2.5 % 97.5 %  
## (Intercept) -8.111584e+02 -1.043755e+02  
## Carb.Volume 2.028564e+00 5.756427e+01  
## Fill.Ounces -4.893746e+00 1.150597e+01  
## PC.Volume -6.155980e+00 2.215355e+01  
## Carb.Pressure -2.462482e+00 4.404040e-01  
## Carb.Temp -3.272268e-01 1.987317e+00  
## PSC -9.600937e+00 2.045816e+01  
## PSC.Fill -5.909472e+00 4.482754e+00  
## PSC.CO2 -2.304007e+01 1.106857e+01  
## Mnf.Flow -3.926049e-02 -1.151470e-02  
## Carb.Pressure1 -6.574295e-02 3.670532e-01  
## Fill.Pressure 6.949850e-02 9.301001e-01  
## Hyd.Pressure1 -1.210219e-01 6.320082e-02  
## Hyd.Pressure2 -1.305002e-01 1.635989e-01  
## Hyd.Pressure3 -2.094582e-01 7.291056e-02  
## Hyd.Pressure4 -2.162475e-01 -3.226052e-02  
## Filler.Level -1.900333e-01 1.515944e-01  
## Filler.Speed -5.643476e-03 1.828542e-03  
## Temperature -4.862743e-01 9.490101e-01  
## Usage.cont -3.064311e-02 5.517861e-01  
## Carb.Flow -5.272380e-03 -2.097482e-03  
## Density -5.358083e+00 7.059994e+00  
## MFR -6.956209e-03 3.422138e-02  
## Balling -5.144302e+00 1.296268e+01  
## Pressure.Vacuum -3.243509e+00 2.459081e+00  
## PH -5.032941e+00 8.428569e+00  
## Oxygen.Filler -1.120450e+01 3.192591e+01  
## Bowl.Setpoint -1.862666e-01 1.628123e-01  
## Pressure.Setpoint -7.502647e-01 2.795246e-01  
## Air.Pressurer -5.633142e-01 8.401317e-01  
## Alch.Rel 1.910948e+01 3.587090e+01  
## Carb.Rel -1.861096e+01 7.398637e+00  
## Balling.Lvl -3.670085e+01 -1.547905e+01

#when run separetly no errors  
#step(lreg,data=df,direction="backward")  
  
#plotting results  
ggplot(d, aes(x=variable, y=value)) +  
 facet\_wrap(~variable,scales = "free") +  
 geom\_boxplot() +  
 ggtitle("Boxplots")

## Warning: Removed 228 rows containing non-finite values (stat\_boxplot).

 # from the above MFflow, Hyd.Pressure4, Carb.Flow,Balling.Lvl has significant relation with Brand.Code

df = read.csv(file = “F:/HW 2/StudentData.csv”, na.strings = c(“”, " “), header = TRUE) save\_plots = T set.seed(0) library(caret) library(”tidyr“) library(AppliedPredictiveModeling) data(df)

pred <- df%>%select(-Yield) yield <- df%>%select(Yield)

require(pls) set.seed (1000)

trainingmydata = createDataPartition( yield, p=0.8 )

Predictorstraining = Predictors[trainingmydata$Resample1,] yieldtraining = yield[trainingmydata$Resample1]

Predictorstesting = Predictors[-trainingmydata$Resample1,] yieldtesting = yield[-trainingmydata$Resample1] #**PLS MODEL** #Build some linear models and predict the performance on the testing data set: set.seed(0) plsmodel = train(Predictorstraining, yieldtraining, method=“pls”, tuneLength=40, preProcess=c(“center”,“scale”), trControl=trainControl(method=“repeatedcv”,repeats=5)) yes = predict( plsmodel, newdata=Predictorstesting) r2\_pls = cor(yes,yieldtesting,method=“pearson”)^2 rmse\_pls = sqrt( mean( (yes-yieldtesting)^2 ) ) print( sprintf( “%-10s: Testing R^2= %10.6f; RMSE= %10.6f”, “PLS”, r2\_pls, rmse\_pls ) )

# variable importance

varimp <- varImp(plsmodel) ggplot(varimp, top = 10) + ggtitle(“Importance of top 10 predictors for PLS model”)