library(tidyverse)

## -- Attaching packages ------------------------------------------------ tidyverse 1.2.1 --

## v ggplot2 3.1.0 v purrr 0.3.2   
## v tibble 2.1.1 v dplyr 0.8.0.1  
## v tidyr 0.8.3 v stringr 1.4.0   
## v readr 1.3.1 v forcats 0.4.0

## -- Conflicts --------------------------------------------------- tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

library(mice)

## Loading required package: lattice

##   
## Attaching package: 'mice'

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

## The following objects are masked from 'package:base':  
##   
## cbind, rbind

library(VIM)

## 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

## The following object is masked from 'package:purrr':  
##   
## transpose

## 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'

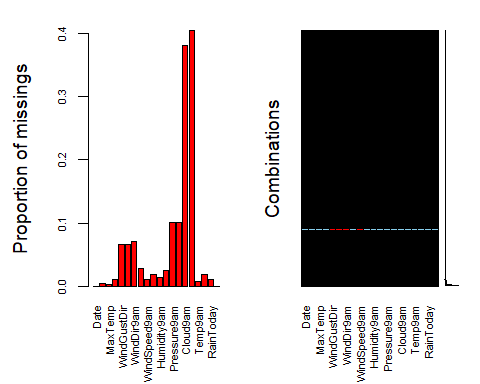
## The following object is masked from 'package:datasets':  
##   
## sleep

Rain=read.csv("Rain.csv")  
str(Rain)

## 'data.frame': 28003 obs. of 20 variables:  
## $ Date : Factor w/ 3223 levels "1/1/2009","1/1/2010",..: 1041 1049 885 893 950 992 103 271 280 47 ...  
## $ MinTemp : num 17.5 14.6 9.8 14.1 20.5 20.1 9.6 14 12.5 17.4 ...  
## $ MaxTemp : num 32.3 29.7 27.7 20.9 31.8 32.7 23.9 28.3 28.4 43 ...  
## $ Rainfall : num 1 0.2 NA 0 0 0 0 0 0 0 ...  
## $ WindGustDir : Factor w/ 16 levels "E","ENE","ESE",..: 14 15 15 2 15 15 14 14 5 8 ...  
## $ WindGustSpeed: int 41 56 50 22 41 48 41 48 37 39 ...  
## $ WindDir9am : Factor w/ 16 levels "E","ENE","ESE",..: 2 14 NA 12 14 4 16 14 11 11 ...  
## $ WindDir3pm : Factor w/ 16 levels "E","ENE","ESE",..: 8 14 15 1 14 15 12 16 9 12 ...  
## $ WindSpeed9am : int 7 19 NA 11 19 13 19 17 20 7 ...  
## $ WindSpeed3pm : int 20 24 22 9 20 30 11 24 9 17 ...  
## $ Humidity9am : int 82 55 50 69 54 56 44 43 38 40 ...  
## $ Humidity3pm : int 33 23 28 82 24 15 22 15 16 8 ...  
## $ Pressure9am : num 1011 1009 1013 1012 1008 ...  
## $ Pressure3pm : num 1006 1005 1010 1010 1006 ...  
## $ Cloud9am : int 7 NA 0 8 NA NA NA NA NA NA ...  
## $ Cloud3pm : int 8 NA NA 1 NA NA NA NA NA NA ...  
## $ Temp9am : num 17.8 20.6 17.3 17.2 23.8 24.6 14.9 17.9 17.2 25.6 ...  
## $ Temp3pm : num 29.7 28.9 26.2 18.1 30.8 32.1 22.1 27.6 26.6 41.5 ...  
## $ RainToday : Factor w/ 2 levels "No","Yes": 1 1 NA 1 1 1 1 1 1 1 ...  
## $ RainTomorrow : Factor w/ 2 levels "No","Yes": 1 1 1 2 1 1 1 1 1 1 ...

vim\_plot = aggr(Rain, numbers = TRUE, prop = c(TRUE, FALSE),cex.axis=.7)

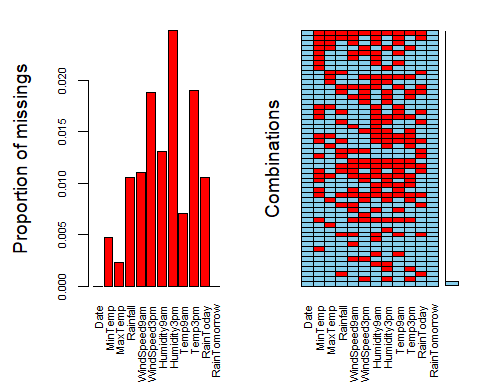
## Warning in plot.aggr(res, ...): not enough vertical space to display  
## frequencies (too many combinations)



Rain1 = Rain %>% select(-Pressure9am,-Pressure3pm,-Cloud3pm,-Cloud9am,-WindGustDir,-WindGustSpeed,-WindDir9am,-WindDir3pm)

vim\_plot = aggr(Rain1, numbers = TRUE, prop = c(TRUE, FALSE),cex.axis=.7)

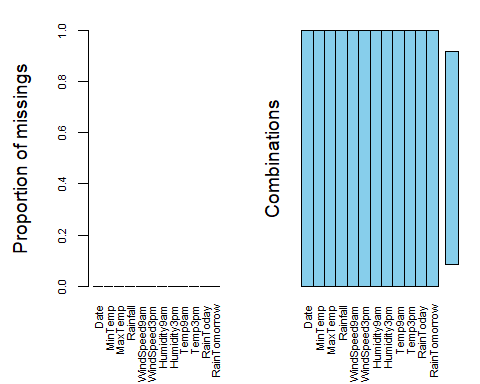
## Warning in plot.aggr(res, ...): not enough vertical space to display  
## frequencies (too many combinations)



Rain2= Rain1 %>% drop\_na()

vim\_plot = aggr(Rain2, numbers = TRUE, prop = c(TRUE, FALSE),cex.axis=.7)

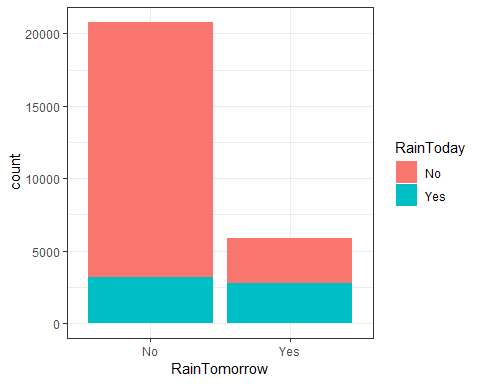
## Warning in plot.aggr(res, ...): not enough horizontal space to display  
## frequencies



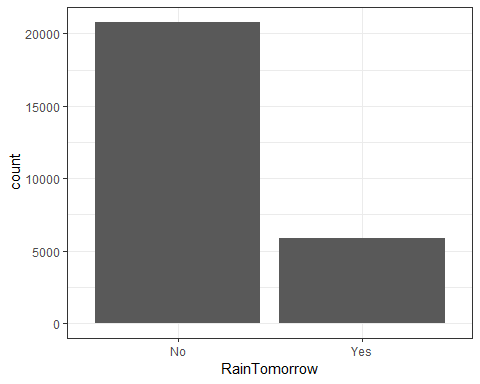
str(Rain2)

## 'data.frame': 26692 obs. of 12 variables:  
## $ Date : Factor w/ 3223 levels "1/1/2009","1/1/2010",..: 1041 1049 893 950 992 103 271 280 47 57 ...  
## $ MinTemp : num 17.5 14.6 14.1 20.5 20.1 9.6 14 12.5 17.4 19.8 ...  
## $ MaxTemp : num 32.3 29.7 20.9 31.8 32.7 23.9 28.3 28.4 43 32.7 ...  
## $ Rainfall : num 1 0.2 0 0 0 0 0 0 0 0 ...  
## $ WindSpeed9am: int 7 19 11 19 13 19 17 20 7 20 ...  
## $ WindSpeed3pm: int 20 24 9 20 30 11 24 9 17 28 ...  
## $ Humidity9am : int 82 55 69 54 56 44 43 38 40 34 ...  
## $ Humidity3pm : int 33 23 82 24 15 22 15 16 8 28 ...  
## $ Temp9am : num 17.8 20.6 17.2 23.8 24.6 14.9 17.9 17.2 25.6 27.6 ...  
## $ Temp3pm : num 29.7 28.9 18.1 30.8 32.1 22.1 27.6 26.6 41.5 27.1 ...  
## $ RainToday : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1 1 1 1 1 ...  
## $ RainTomorrow: Factor w/ 2 levels "No","Yes": 1 1 2 1 1 1 1 1 1 1 ...

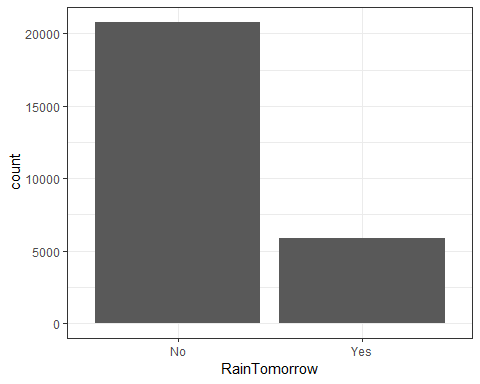
ggplot(Rain2, aes(x=RainTomorrow, fill = RainToday)) + geom\_bar() + theme\_bw()



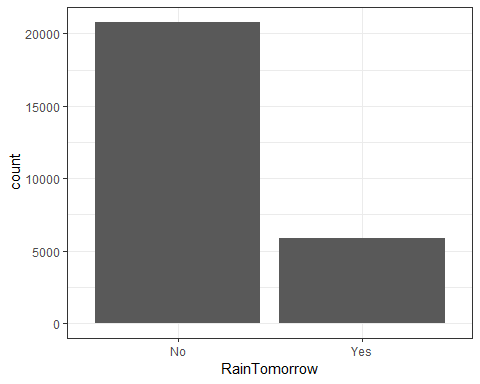
ggplot(Rain2, aes(x=RainTomorrow, fill = Temp3pm)) + geom\_bar() + theme\_bw()



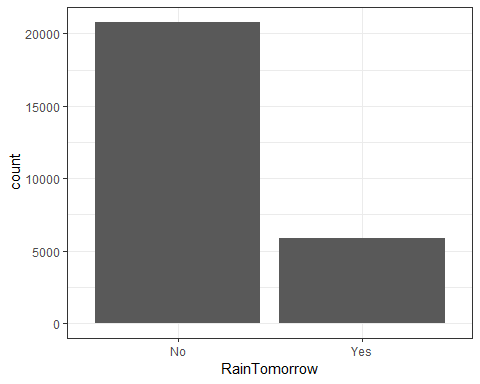
ggplot(Rain2, aes(x=RainTomorrow, fill = Temp9am)) + geom\_bar() + theme\_bw()



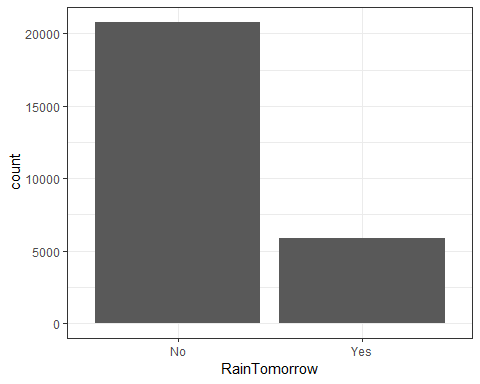
ggplot(Rain2, aes(x=RainTomorrow, fill = Humidity3pm)) + geom\_bar() + theme\_bw()



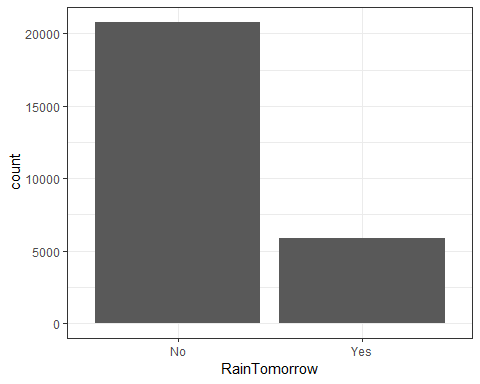
ggplot(Rain2, aes(x=RainTomorrow, fill = Humidity9am)) + geom\_bar() + theme\_bw()



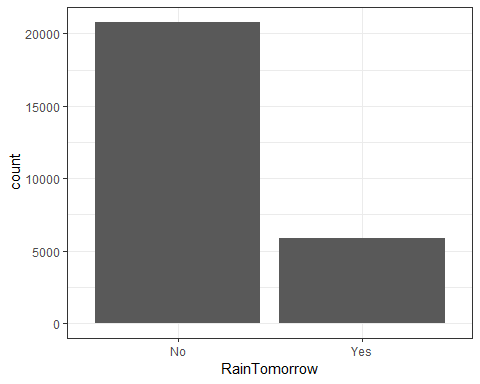
ggplot(Rain2, aes(x=RainTomorrow, fill = WindSpeed3pm)) + geom\_bar() + theme\_bw()



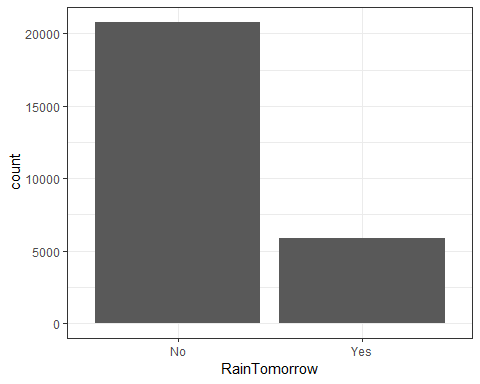
ggplot(Rain2, aes(x=RainTomorrow, fill = WindSpeed9am)) + geom\_bar() + theme\_bw()



ggplot(Rain2, aes(x=RainTomorrow, fill = Rainfall)) + geom\_bar() + theme\_bw()



ggplot(Rain2, aes(x=RainTomorrow, fill = MaxTemp)) + geom\_bar() + theme\_bw()



ggplot(Rain2, aes(x=RainTomorrow, fill = MinTemp)) + geom\_bar() + theme\_bw()

