

# Final

Semih Barutcu

3/6/2020

```
library(pacman)

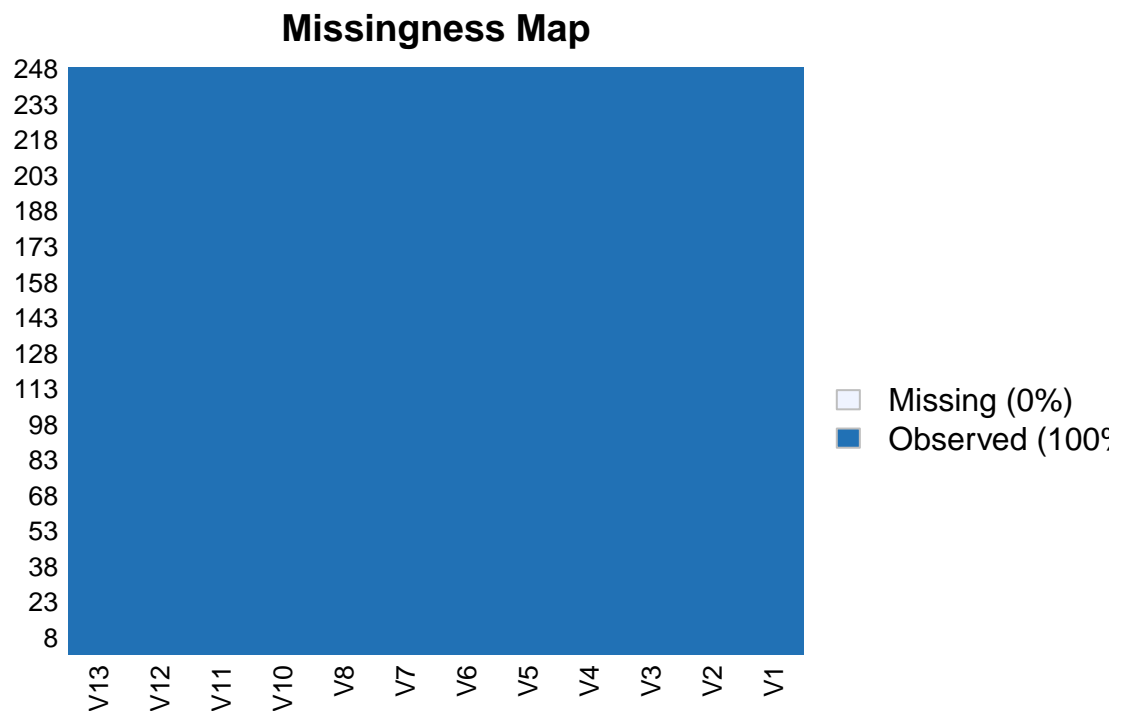
p_load(mlbench, tidyverse, ggplot2, Amelia, ROCR, caret, randomForest, tictoc, Boruta)

data("Ozone")

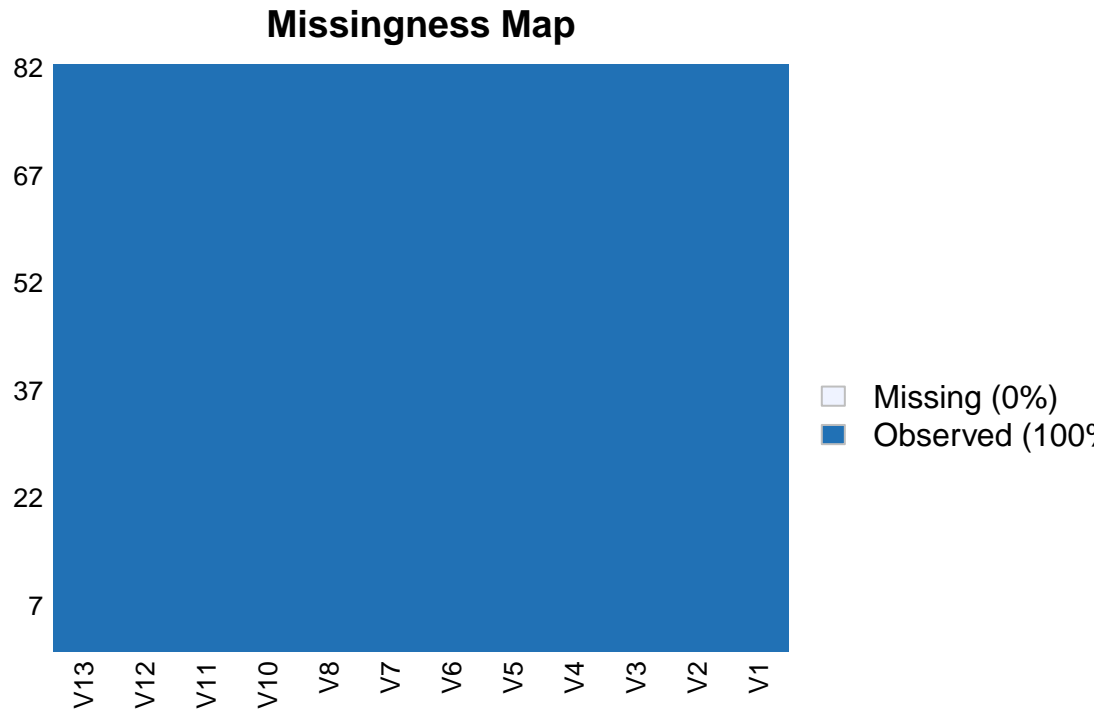
df <- Ozone %>% select(-V9) %>% na.omit()

idx <- sample(nrow(df), round(0.75*nrow(df)))
train <- df[idx,]
test <- df[-idx,]

missmap(train)
```



```
missmap(test)
```



```
tic()
set.seed(300)
m_rf_Boruta <- Boruta(V4 ~ ., data = train, doTrace = 2, ntree = 500)
```

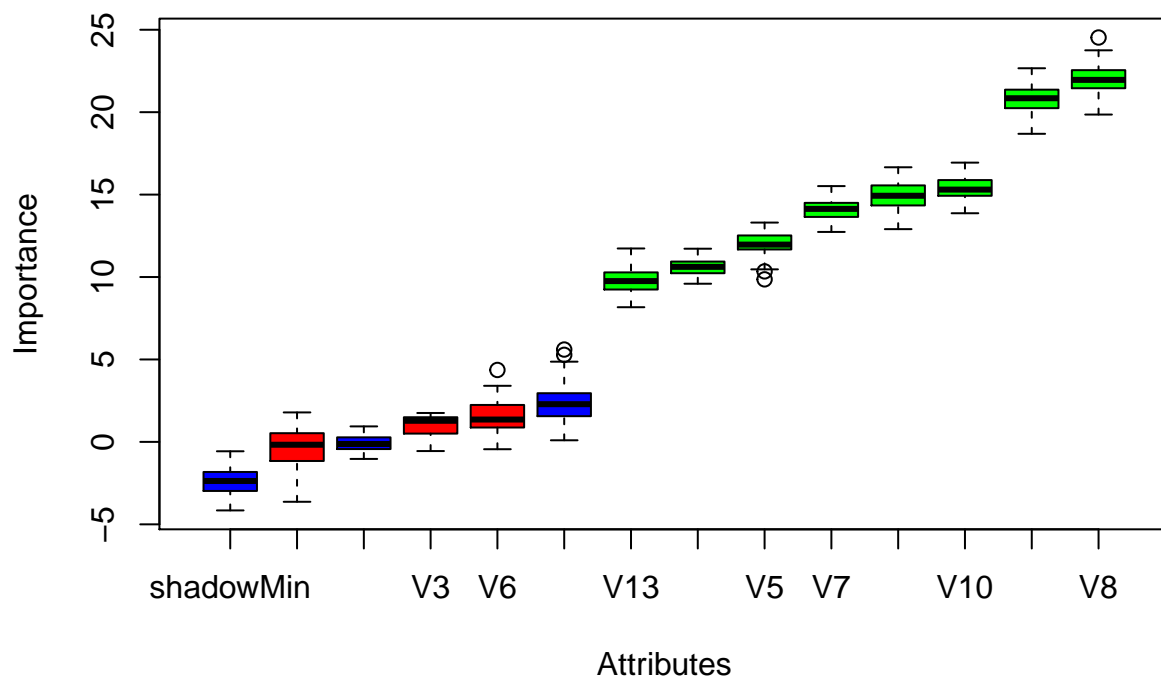
```
## 1. run of importance source...
## 2. run of importance source...
## 3. run of importance source...
## 4. run of importance source...
## 5. run of importance source...
## 6. run of importance source...
## 7. run of importance source...
## 8. run of importance source...
## 9. run of importance source...
## 10. run of importance source...
## 11. run of importance source...
## After 11 iterations, +0.46 secs:
## confirmed 8 attributes: V1, V10, V11, V12, V13 and 3 more;
## still have 3 attributes left.
```

```
## 12. run of importance source...
## 13. run of importance source...
## 14. run of importance source...
## 15. run of importance source...
## After 15 iterations, +0.65 secs:
## rejected 2 attributes: V2, V3;
## still have 1 attribute left.
## 16. run of importance source...
## 17. run of importance source...
## 18. run of importance source...
## 19. run of importance source...
## 20. run of importance source...
## 21. run of importance source...
## 22. run of importance source...
## 23. run of importance source...
## 24. run of importance source...
## 25. run of importance source...
## 26. run of importance source...
## 27. run of importance source...
## 28. run of importance source...
## 29. run of importance source...
## 30. run of importance source...
## 31. run of importance source...
## 32. run of importance source...
## 33. run of importance source...
## 34. run of importance source...
## 35. run of importance source...
## 36. run of importance source...
## 37. run of importance source...
## 38. run of importance source...
## 39. run of importance source...
## 40. run of importance source...
## 41. run of importance source...
## 42. run of importance source...
## 43. run of importance source...
## 44. run of importance source...
```

```
## 45. run of importance source...
## 46. run of importance source...
## 47. run of importance source...
## 48. run of importance source...
## 49. run of importance source...
## After 49 iterations, +2.1 secs:
## rejected 1 attribute: V6;
## no more attributes left.
toc()

## 2.14 sec elapsed
m_rf_Boruta

## Boruta performed 49 iterations in 2.138168 secs.
## 8 attributes confirmed important: V1, V10, V11, V12, V13 and 3 more;
## 3 attributes confirmed unimportant: V2, V3, V6;
plot(m_rf_Boruta)
```



```
getConfirmedFormula(m_rf_Boruta)

## V4 ~ V1 + V5 + V7 + V8 + V10 + V11 + V12 + V13
## <environment: 0x00000000292668c8>
```

```
attStats(m_rf_Boruta)
```

##		meanImp	medianImp	minImp	maxImp	normHits	decision
##	V1	10.5948050	10.6177257	9.5937484	11.718461	1.00000000	Confirmed
##	V2	-0.3362956	-0.1716562	-3.6319212	1.787859	0.02040816	Rejected
##	V3	0.8872903	1.2806926	-0.5553607	1.753566	0.02040816	Rejected
##	V5	11.9636157	11.9671960	9.8601622	13.302449	1.00000000	Confirmed
##	V6	1.5680120	1.3554150	-0.4442294	4.362534	0.26530612	Rejected
##	V7	14.1211329	14.1332628	12.7356914	15.515389	1.00000000	Confirmed
##	V8	21.9625295	21.9582947	19.8564084	24.528168	1.00000000	Confirmed
##	V10	15.3606706	15.3042561	13.8700326	16.940108	1.00000000	Confirmed
##	V11	14.9258768	14.9304138	12.9053498	16.658032	1.00000000	Confirmed
##	V12	20.8465705	20.8451462	18.6882141	22.663190	1.00000000	Confirmed
##	V13	9.7438767	9.7545129	8.1691584	11.733186	1.00000000	Confirmed

```
plotImpHistory(m_rf_Boruta)
```

