REPORT SERIES WITH DLOOKR

Exploratory Data Analysis Report

Author:
dlookr package

Version:
0.3.12

March 6, 2020

Contents

Chapter 1

Introduction

The EDA Report provides exploratory data analysis information on objects that inherit data.frame and data.frame.

1.1 Information of Dataset

The dataset that generated the EDA Report is an 'data.frame' object. It consists of 4,898 observations and 12 variables.

1.2 Information of Variables

Table 1.1: Information of Variables

variables	types	missing_count	missing_percent	unique_count	unique_rate
V1	numeric	0	0	68	0.0138832
V2	numeric	0	0	125	0.0255206
V3	numeric	0	0	87	0.0177624
V4	numeric	0	0	310	0.0632911
V5	numeric	0	0	160	0.0326664
V6	numeric	0	0	132	0.0269498
V7	numeric	0	0	251	0.0512454
V8	numeric	0	0	890	0.1817068
V9	numeric	0	0	103	0.0210290
V10	numeric	0	0	79	0.0161290
V11	numeric	0	0	103	0.0210290
Class	factor	0	0	7	0.0014292

The target variable of the data is 'Class', and the data type of the variable is factor.

1.3 About EDA Report

EDA reports provide information and visualization results that support the EDA process. In particular, it provides a variety of information to understand the relationship between the target variable and the rest of the variables of interest.

Chapter 2

Univariate Analysis

2.1 Descriptive Statistics

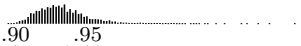
edaData
12 Variables 4898 Observations

V1

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
4898	0	68	0.998	6.855	0.925	5.6	5.9	6.3	6.8	7.3	7.9	8.3

```
lowest :  3.8  3.9  4.2  4.4  4.5, highest: 10.2 10.3 10.7 11.8 14.2
```

V2



n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
4898	0	125	0.999	0.2782	0.1055	0.15	0.17	0.21	0.26	0.32	0.40	0.46

```
lowest : 0.080 0.085 0.090 0.100 0.105, highest: 0.910 0.930 0.965 1.005 1.100
```

V3

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
4898	0	87	0.999	0.3342	0.1258	0.17	0.22	0.27	0.32	0.39	0.49	0.54

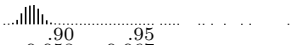
```
lowest : 0.00 0.01 0.02 0.03 0.04, highest: 0.91 0.99 1.00 1.23 1.66
```

V4

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
4898	0	310	1	6.391	5.548	1.1	1.2	1.7	5.2	9.9	14.0	15.7

```
lowest :  0.60  0.70  0.80  0.90  0.95, highest: 22.60 23.50 26.05 31.60 65.80
```

V5



n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
4898	0	160	0.999	0.04577	0.01708	0.027	0.030	0.036	0.043	0.050	0.058	0.067

```
lowest : 0.009 0.012 0.013 0.014 0.015, highest: 0.255 0.271 0.290 0.301 0.346
```

V6

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
4898	0	132	1	35.31	18.5	11	15	23	34	46	57	63

```
lowest :   2.0   3.0   4.0   5.0   6.0, highest: 128.0 131.0 138.5 146.5 289.0
```

V7

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
4898	0	251	1	138.4	47.79	75	87	108	134	167	195	212

```
lowest :   9.0  10.0  18.0  19.0  21.0, highest: 307.5 313.0 344.0 366.5 440.0
```

V8

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
4898	0	890	1	0.994	0.003351	0.9896	0.9903	0.9917	0.9937	0.9961	0.9981	0.9990

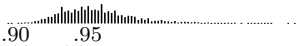
```
lowest : 0.98711 0.98713 0.98722 0.98740 0.98742, highest: 1.00240 1.00241 1.00295 1.01030 1.03898
```

V9

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
4898	0	103	1	3.188	0.1684	2.96	3.00	3.09	3.18	3.28	3.38	3.46

```
lowest : 2.72 2.74 2.77 2.79 2.80, highest: 3.77 3.79 3.80 3.81 3.82
```

V10



n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
4898	0	79	0.999	0.4898	0.1243	0.34	0.36	0.41	0.47	0.55	0.64	0.71

```
lowest : 0.22 0.23 0.25 0.26 0.27, highest: 0.99 1.00 1.01 1.06 1.08
```

V11

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
4898	0	103	0.999	10.51	1.398	8.9	9.0	9.5	10.4	11.4	12.4	12.7

```
lowest :  8.00  8.40  8.50  8.60  8.70, highest: 13.80 13.90 14.00 14.05 14.20
```

Class

n	missing	distinct
4898	0	7

```
lowest : 1 2 3 4 5, highest: 3 4 5 6 7
```

Value	1	2	3	4	5	6	7
Frequency	20	163	1457	2198	880	175	5
Proportion	0.004	0.033	0.297	0.449	0.180	0.036	0.001

2.2 Normality Test of Numerical Variables

2.2.1 Statistics and Visualization of (Sample) Data

V1

normality test : Shapiro-Wilk normality test
statistic : 0.97656, p-value : 1.15015E-27

type	skewness	kurtosis
original	0.6476	5.1687
log transformation	0.0768	3.9575
sqrt transformation	0.3512	4.2870

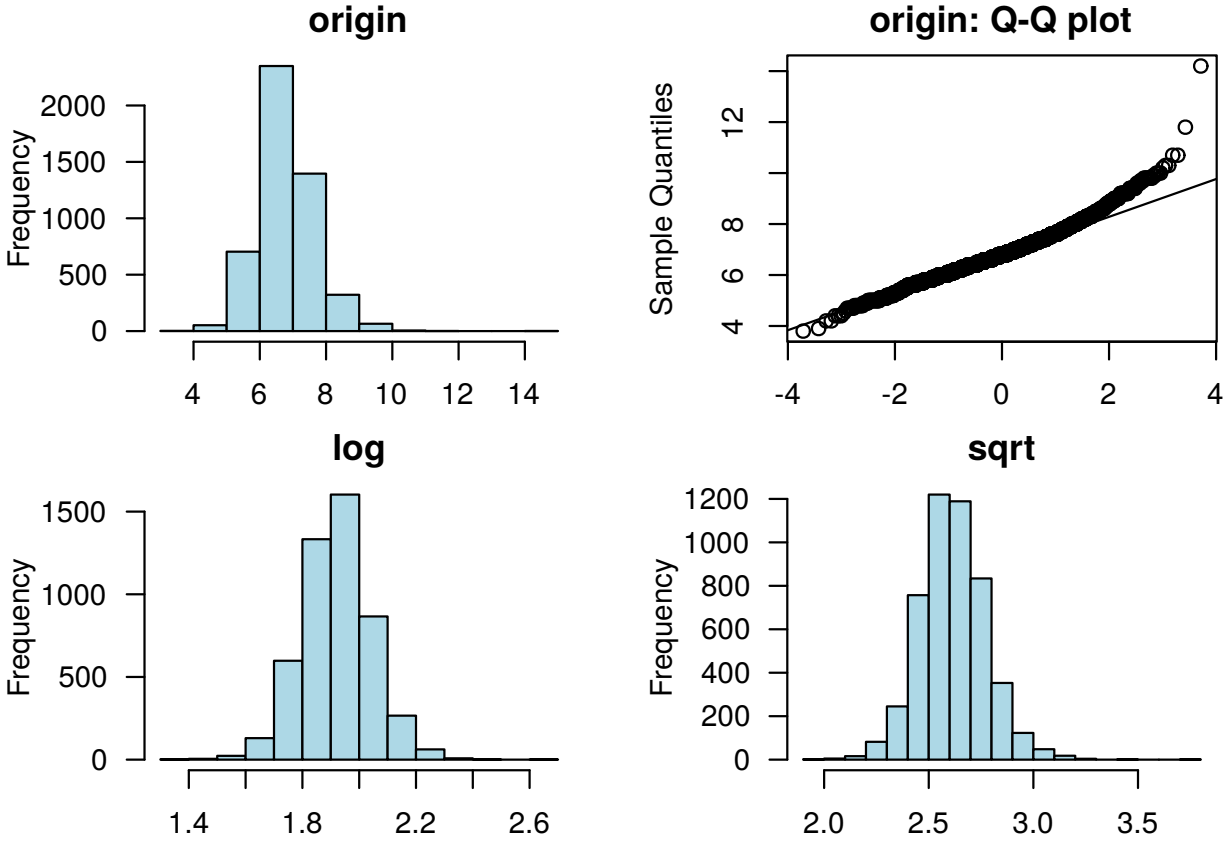


Figure 2.1: V1

V2

normality test : Shapiro-Wilk normality test
statistic : 0.90455, p-value : 4.5868E-48

type	skewness	kurtosis
original	1.5765	8.0852
log transformation	0.1393	3.4267
sqrt transformation	0.7881	4.6038

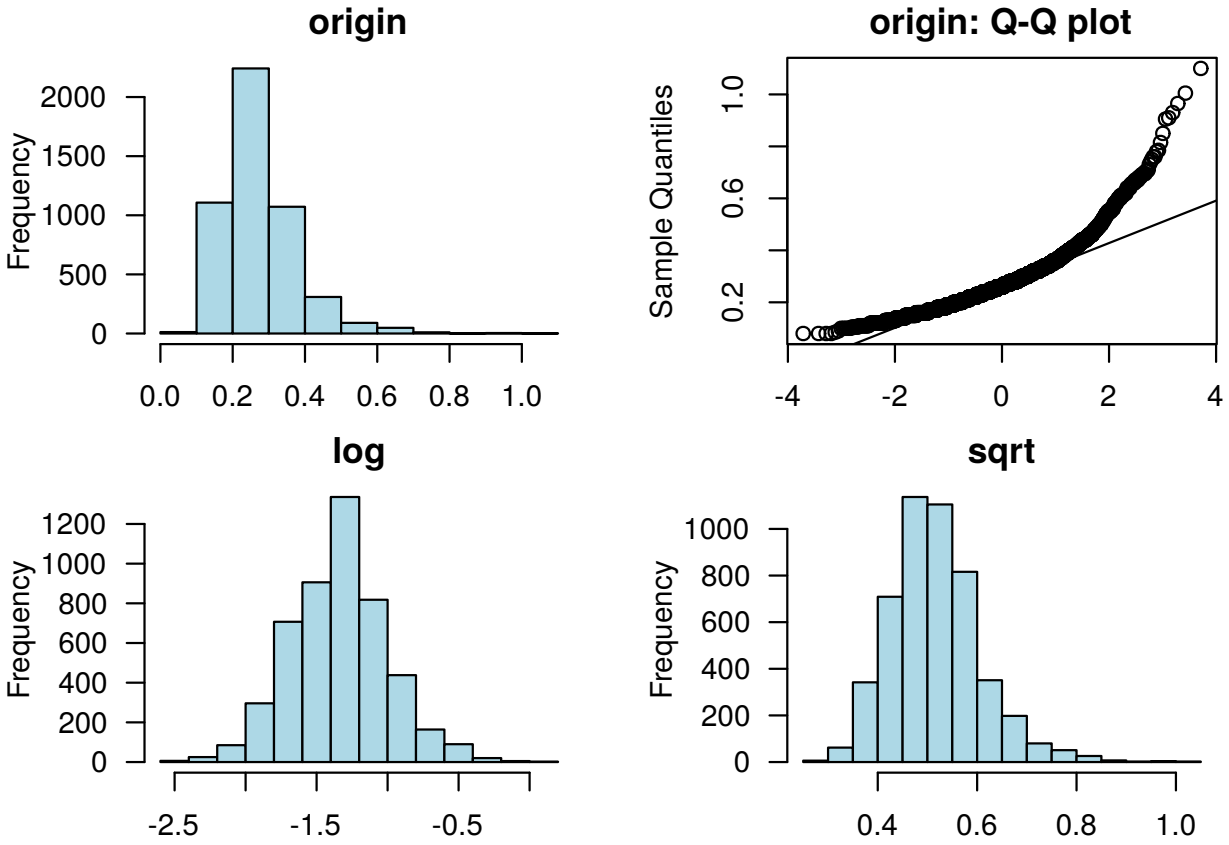


Figure 2.2: V2

V3

normality test : Shapiro-Wilk normality test
statistic : 0.92225, p-value : 1.01318E-44

type	skewness	kurtosis
original	1.2815	9.1674
log transformation		
sqrt transformation	-0.4267	7.7804

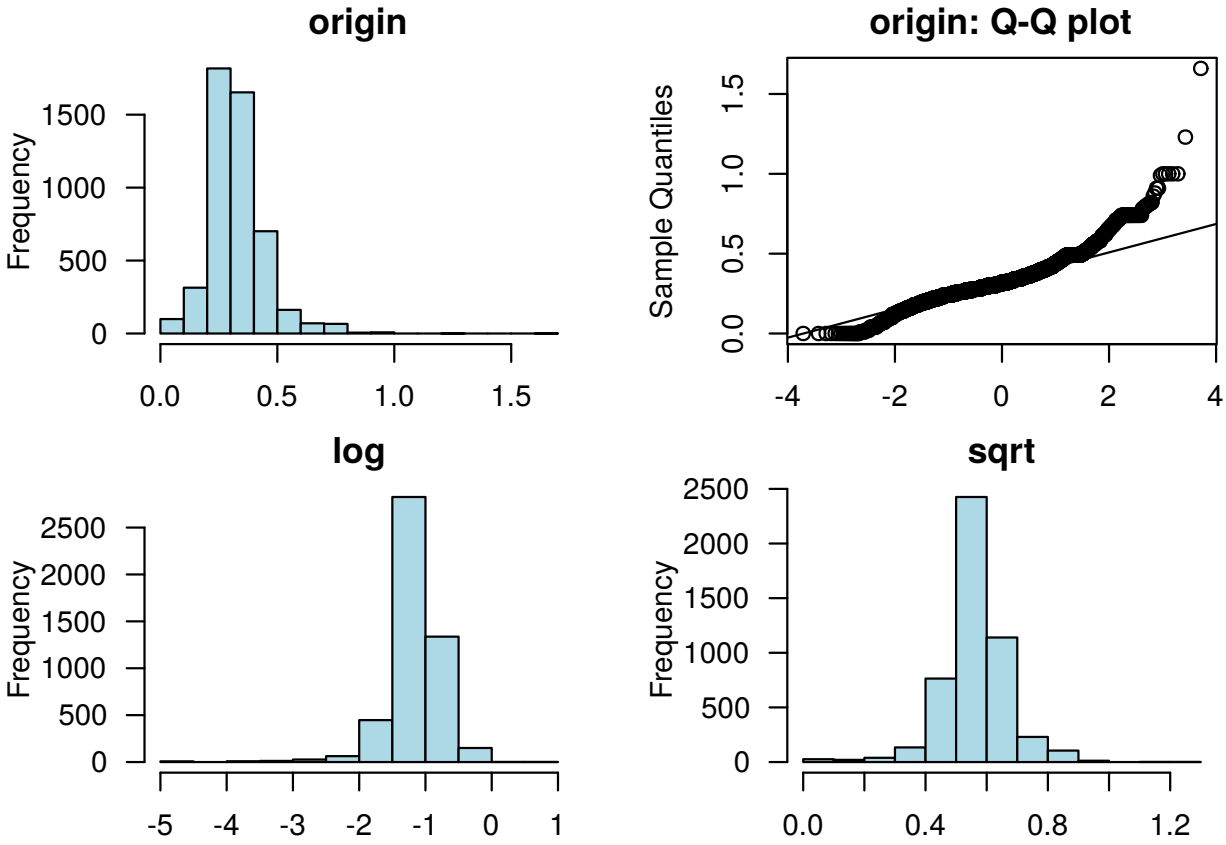


Figure 2.3: V3

V4

normality test : Shapiro-Wilk normality test
statistic : 0.88457, p-value : 2.82071E-51

type	skewness	kurtosis
original	1.0768	6.4651
log transformation	-0.1611	1.6478
sqrt transformation	0.3161	2.0561

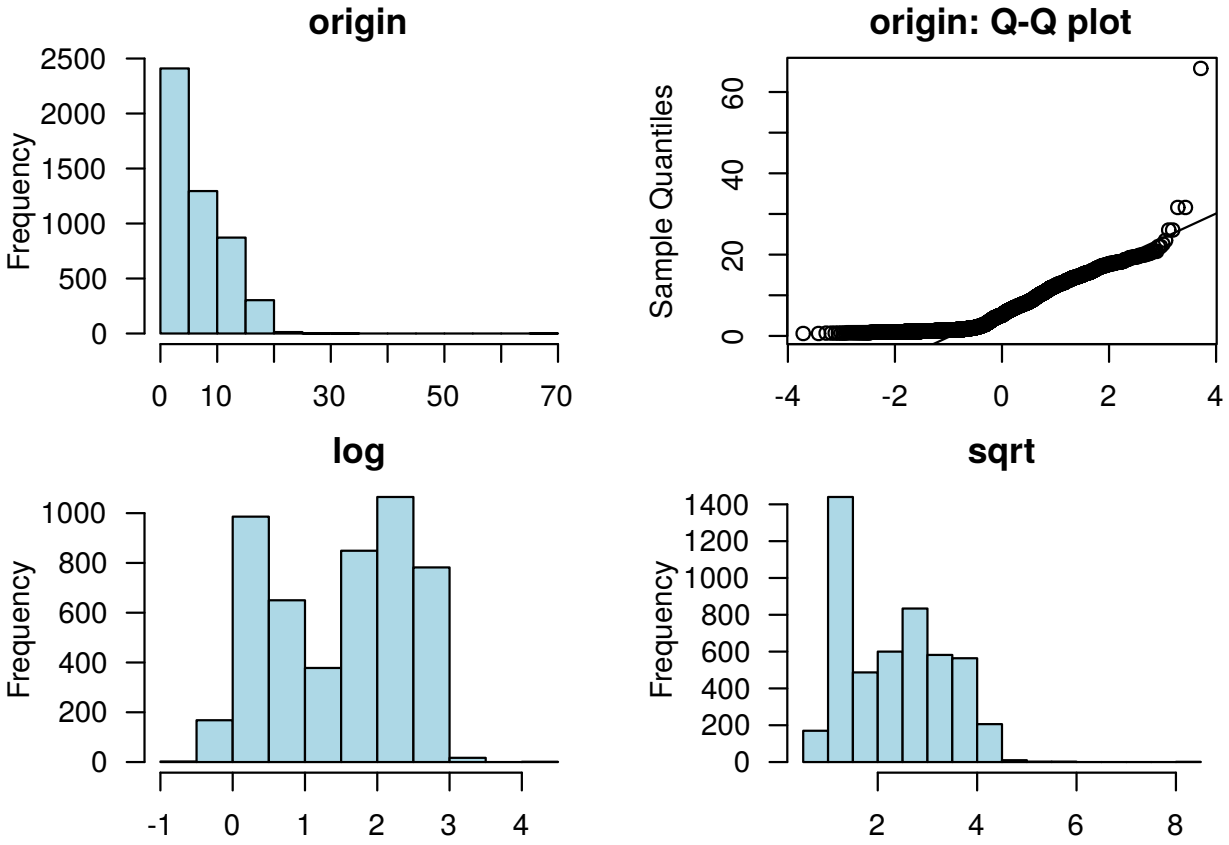


Figure 2.4: V4

V5

normality test : Shapiro-Wilk normality test
statistic : 0.59081, p-value : 2.14058E-75

type	skewness	kurtosis
original	5.0218	40.5250
log transformation	1.1338	8.2934
sqrt transformation	2.8499	18.0404

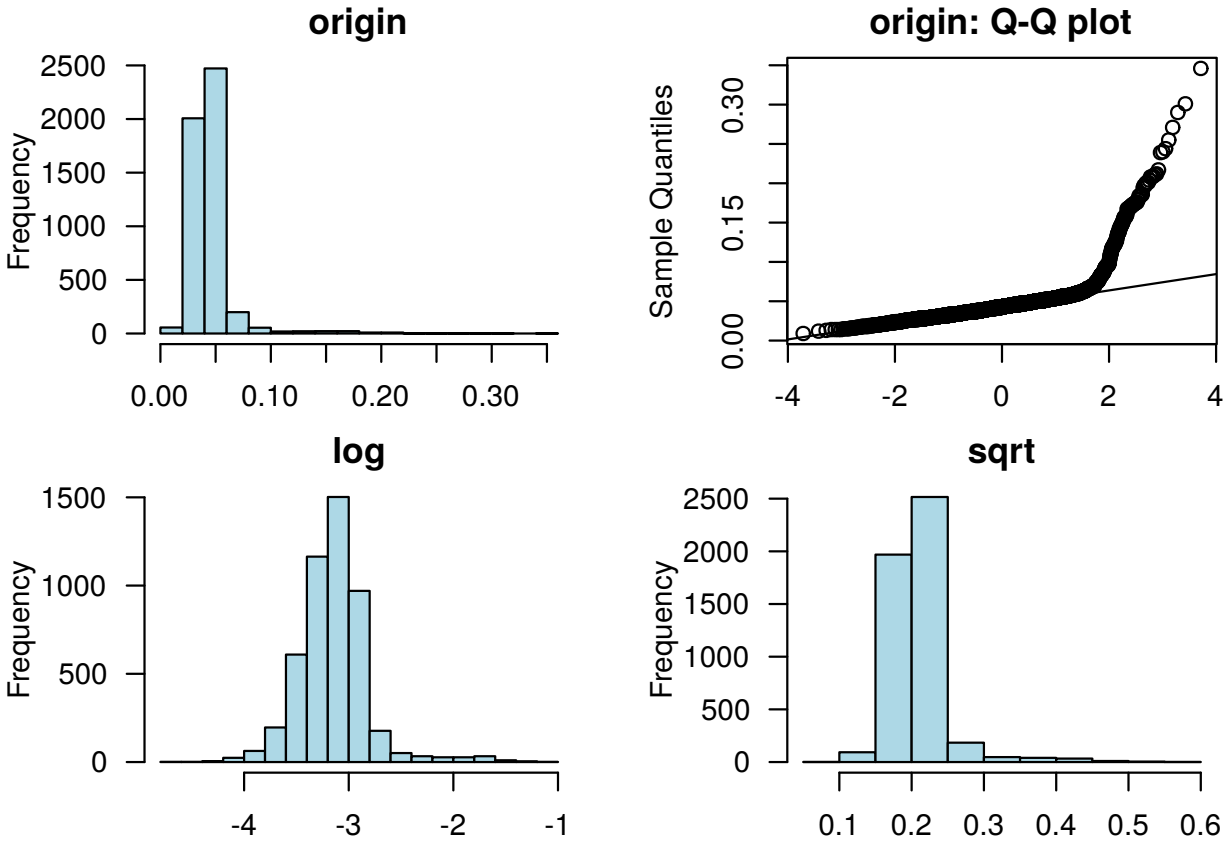


Figure 2.5: V5

V6

normality test : Shapiro-Wilk normality test
statistic : 0.94207, p-value : 3.85785E-40

type	skewness	kurtosis
original	1.4063	14.4534
log transformation	-0.9360	4.4854
sqrt transformation	0.0497	3.8396

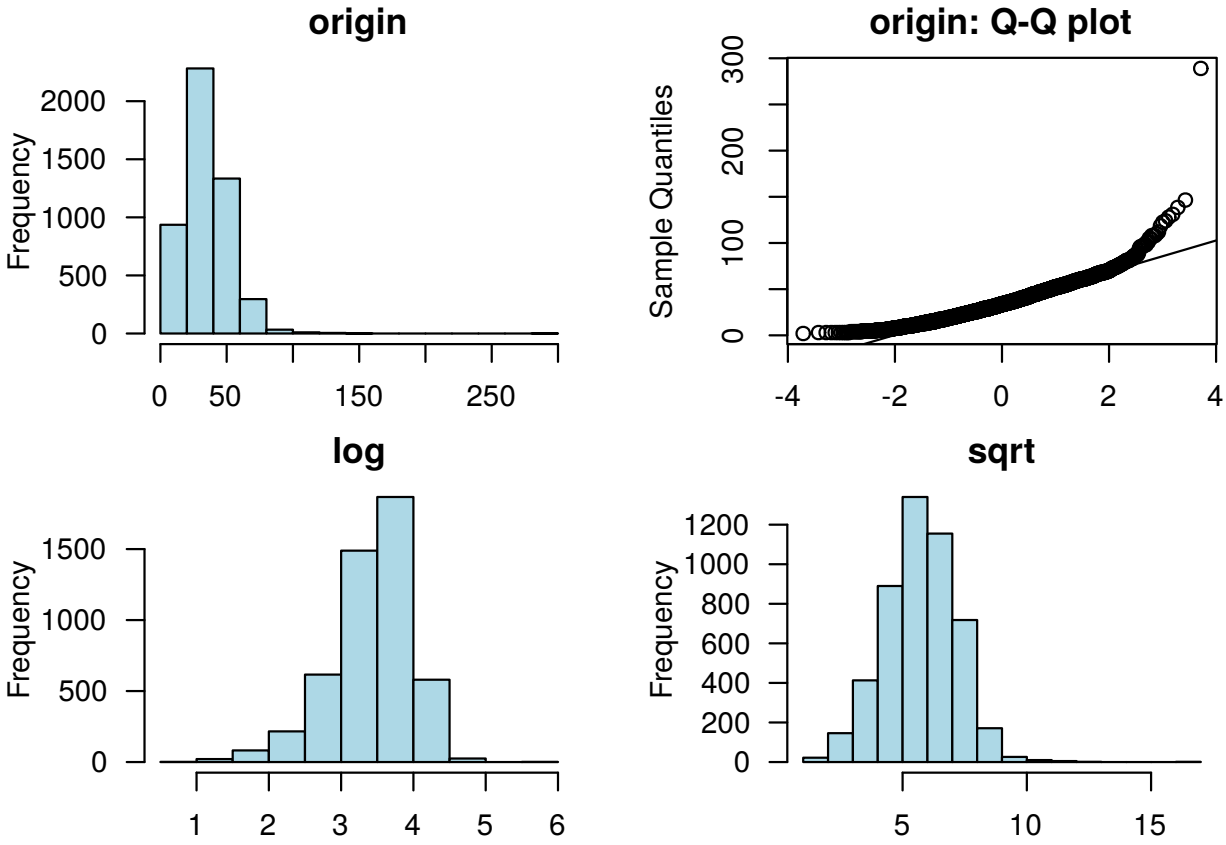


Figure 2.6: V6

V7

normality test : Shapiro-Wilk normality test
statistic : 0.98901, p-value : 4.38345E-19

type	skewness	kurtosis
original	0.3906	3.5700
log transformation	-0.9839	6.4623
sqrt transformation	-0.1639	3.4353

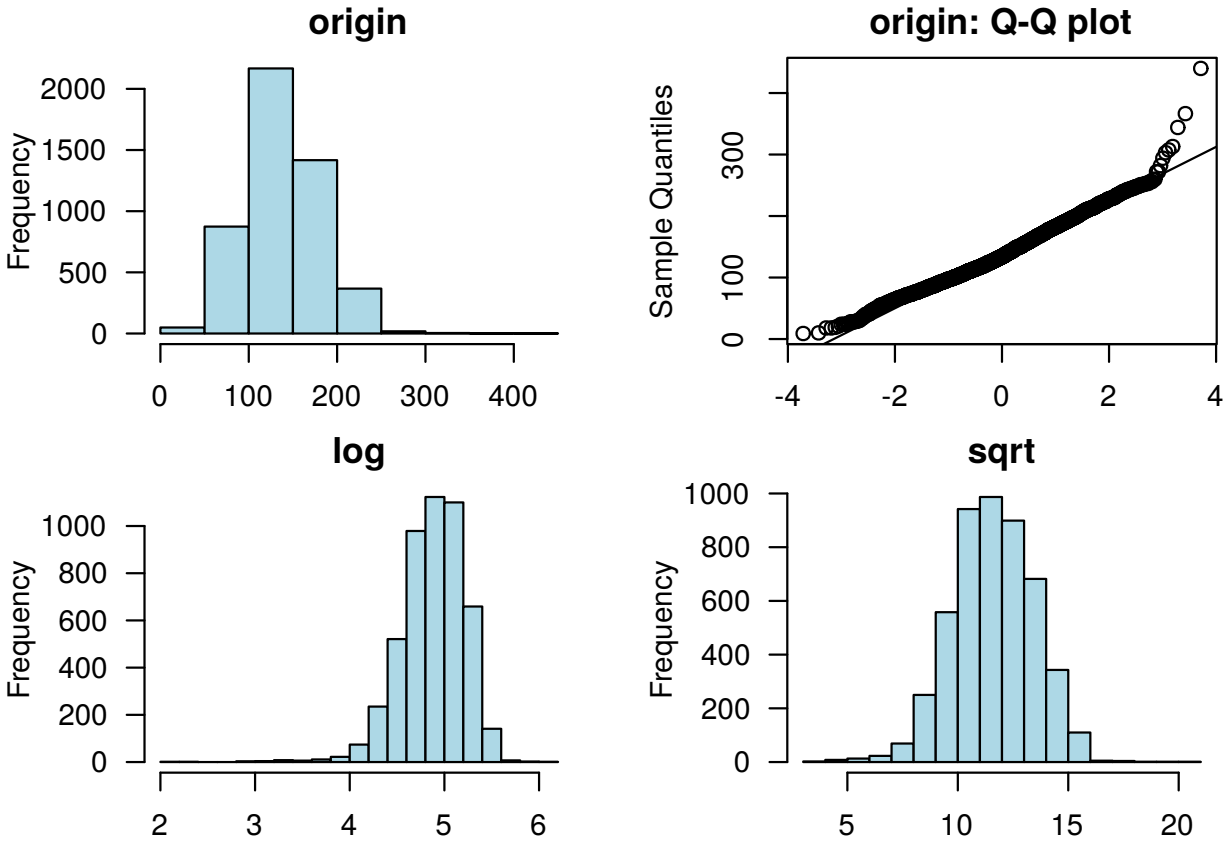


Figure 2.7: V7

V8

normality test : Shapiro-Wilk normality test
statistic : 0.9548, p-value : 1.78089E-36

type	skewness	kurtosis
original	0.9775	12.7826
log transformation	0.9307	11.9521
sqrt transformation	0.9537	12.3583

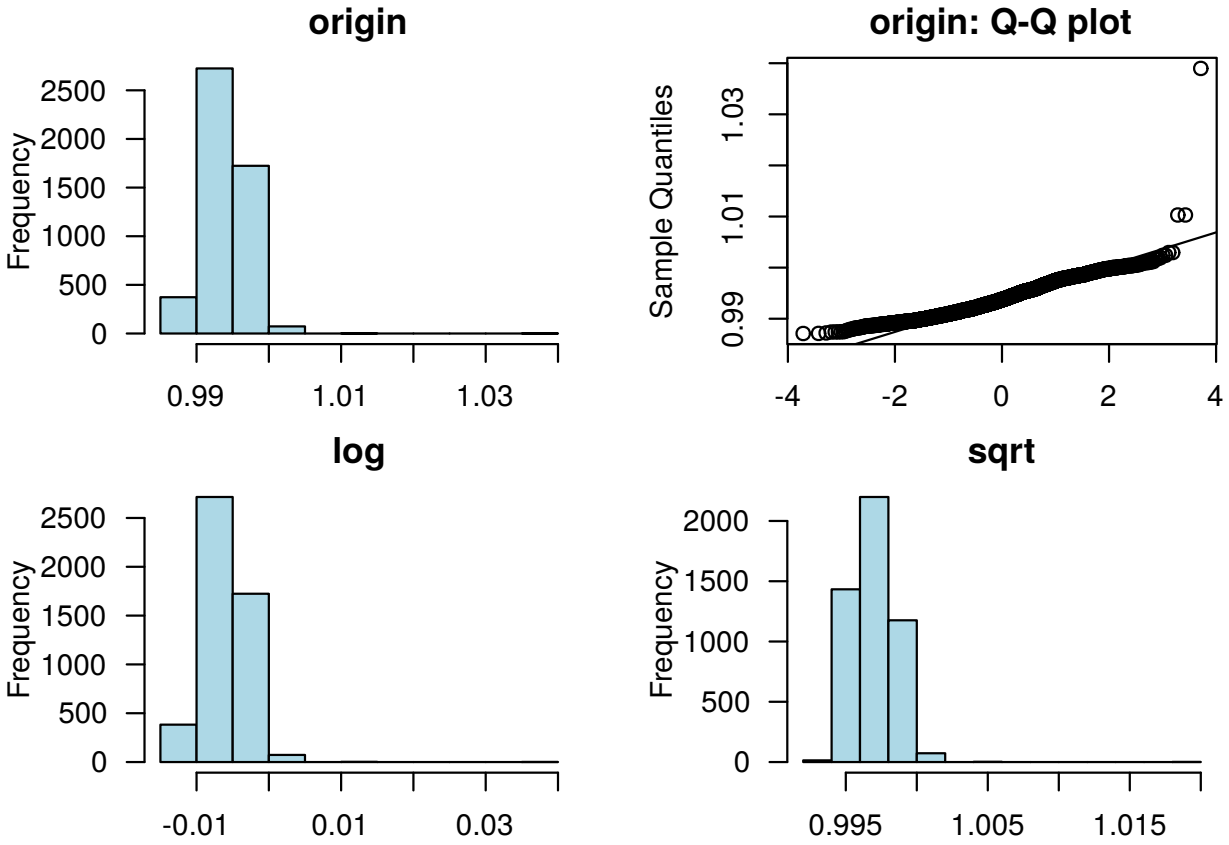


Figure 2.8: V8

V9

normality test : Shapiro-Wilk normality test
statistic : 0.9881, p-value : 6.50552E-20

type	skewness	kurtosis
original	0.4576	3.5290
log transformation	0.2987	3.3057
sqrt transformation	0.3775	3.4055

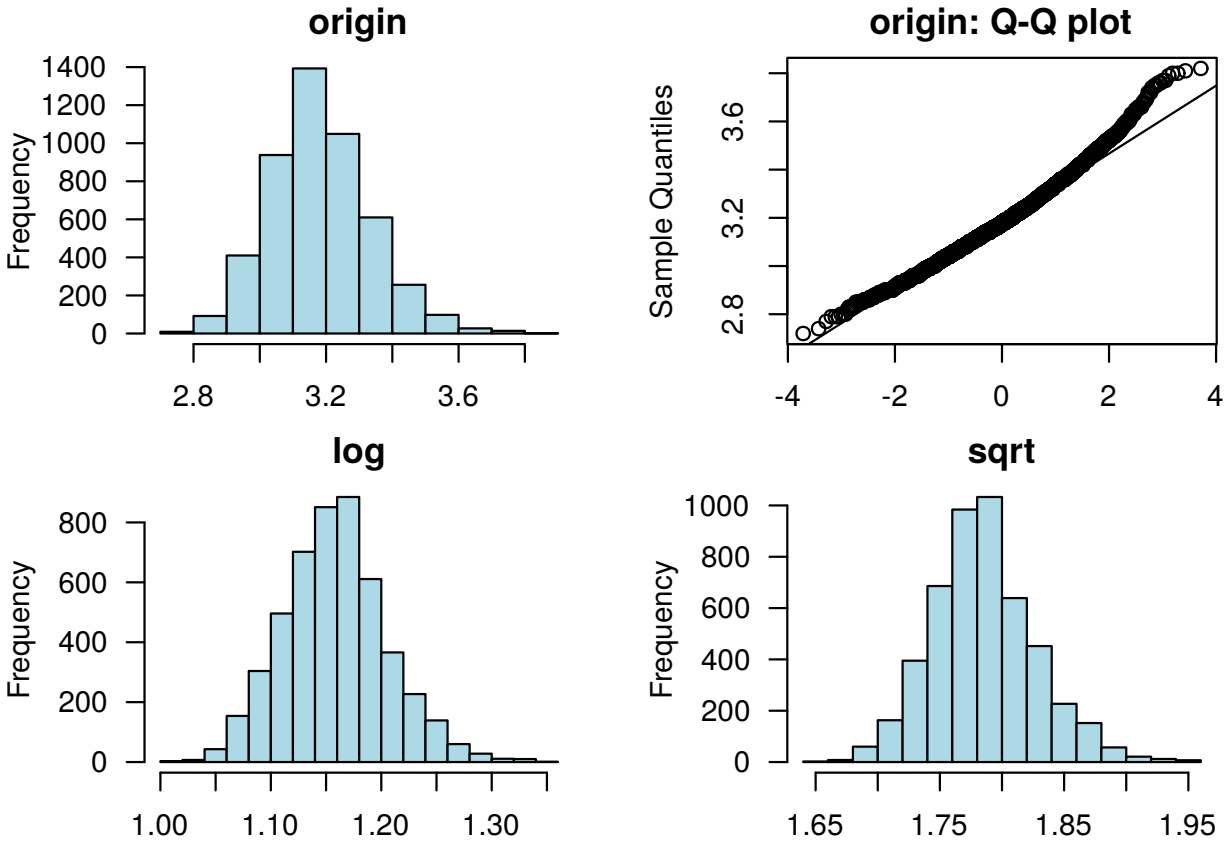


Figure 2.9: V9

V10

normality test : Shapiro-Wilk normality test
statistic : 0.95161, p-value : 1.82198E-37

type	skewness	kurtosis
original	0.9769	4.5881
log transformation	0.2337	3.1372
sqrt transformation	0.5921	3.6139

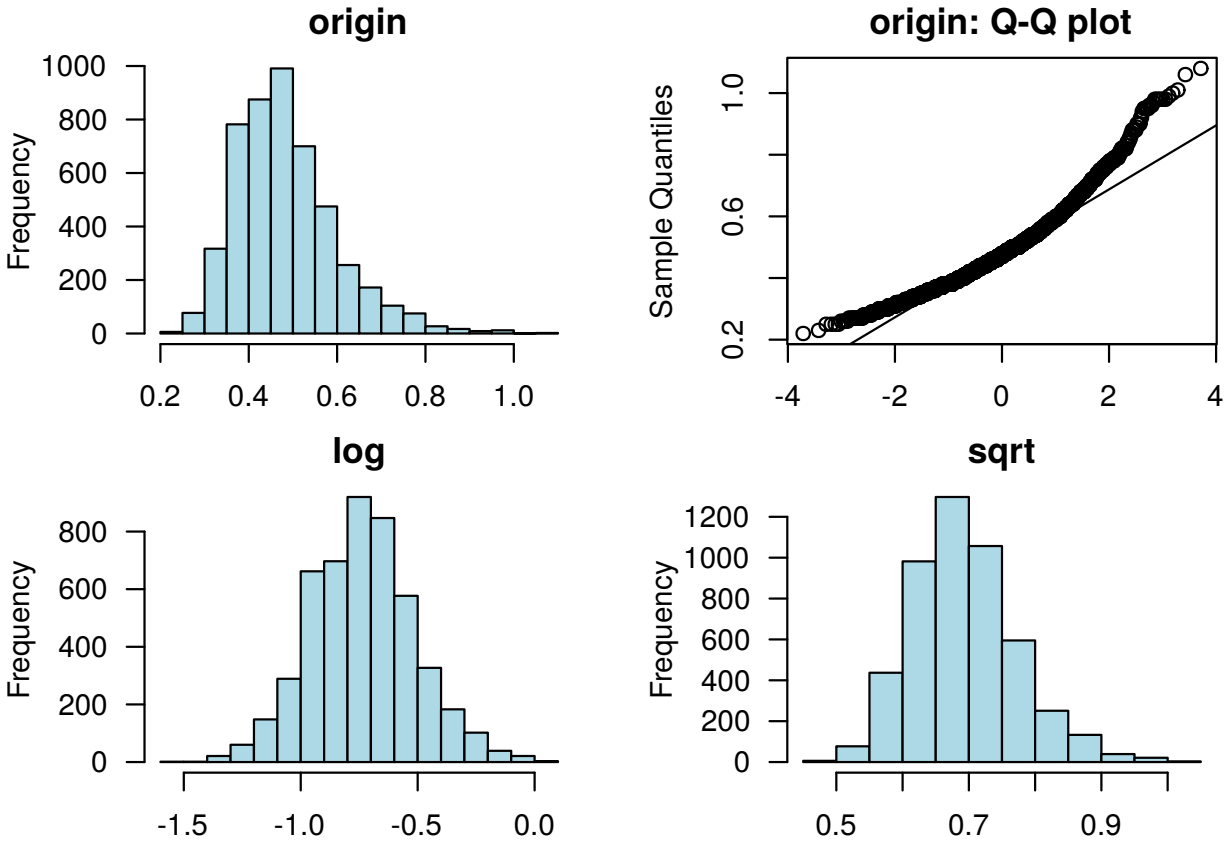


Figure 2.10: V10

V11

normality test : Shapiro-Wilk normality test
statistic : 0.9553, p-value : 2.56901E-36

type	skewness	kurtosis
original	0.4872	2.3011
log transformation	0.3100	2.1036
sqrt transformation	0.3978	2.1909

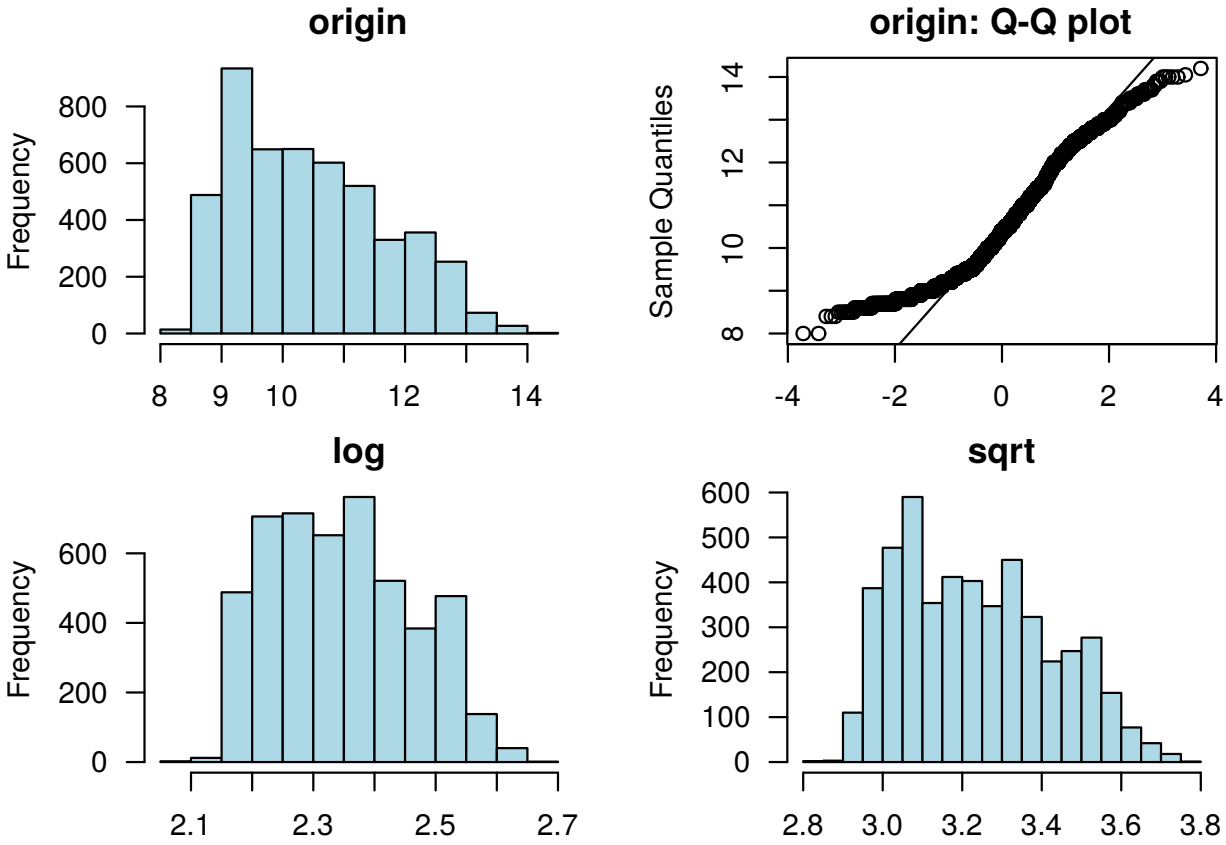


Figure 2.11: V11

Chapter 3

Relationship Between Variables

3.1 Correlation Coefficient

3.1.1 Correlation Coefficient by Variable Combination

Table 3.1: The correlation coefficients (0.5 or more)

Variable1	Variable2	Correlation Coefficient
V8	V4	0.839
V11	V8	-0.780
V7	V6	0.616
V8	V7	0.530

3.1.2 Correlation Plot of Numerical Variables

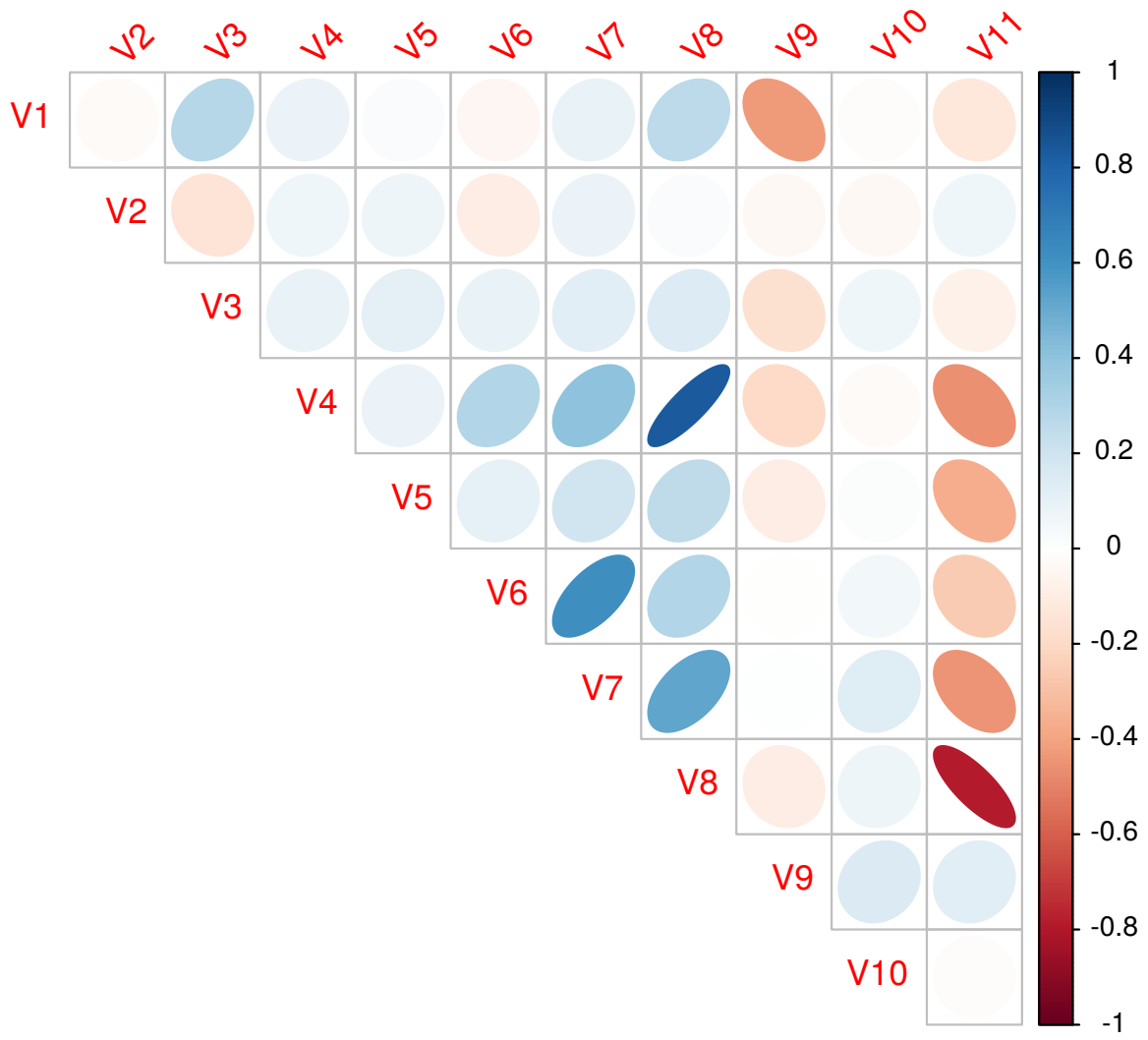


Figure 3.1: The correlation coefficient of numerical variables

Chapter 4

Target based Analysis

4.1 Grouped Descriptive Statistics

4.1.1 Grouped Numerical Variables

V1

Table 4.1: V1

	7	6	5	4	3	2	1
n	5.00	175.00	880.00	2,198.00	1,457.00	163.00	20.00
NA	0.00	0.00	0.00	0.00	0.00	0.00	0.00
mean	7.42	6.66	6.73	6.84	6.93	7.13	7.60
sd	0.98	0.82	0.76	0.84	0.84	1.08	1.72
se(mean)	0.44	0.06	0.03	0.02	0.02	0.08	0.39
IQR	0.50	1.10	1.00	1.00	1.00	1.20	1.95
skewness	1.76	-0.51	0.14	0.75	0.60	0.78	0.55
kurtosis	3.36	0.14	0.43	3.28	0.82	0.30	0.88
0%	6.60	3.90	4.20	3.80	4.50	4.80	4.20
1%	6.61	4.77	5.00	5.00	5.16	5.26	4.50
5%	6.66	5.20	5.60	5.60	5.70	5.80	5.72
10%	6.72	5.40	5.80	5.90	6.00	6.00	6.07
20%	6.84	6.00	6.10	6.20	6.30	6.30	6.18
25%	6.90	6.20	6.20	6.30	6.40	6.40	6.58
30%	6.94	6.40	6.40	6.40	6.50	6.50	6.77
40%	7.02	6.60	6.50	6.60	6.60	6.80	7.02
50%	7.10	6.80	6.70	6.80	6.80	6.90	7.30
60%	7.22	6.94	6.90	7.00	7.10	7.12	7.72
70%	7.34	7.10	7.10	7.20	7.30	7.50	8.36
75%	7.40	7.30	7.20	7.30	7.40	7.60	8.53
80%	7.74	7.30	7.30	7.50	7.60	8.00	8.70
90%	8.42	7.60	7.70	7.90	8.00	8.68	9.49
95%	8.76	8.00	8.00	8.30	8.40	9.20	10.38
99%	9.03	8.20	8.70	9.01	9.40	9.94	11.51
100%	9.10	8.20	9.20	14.20	10.30	10.20	11.80

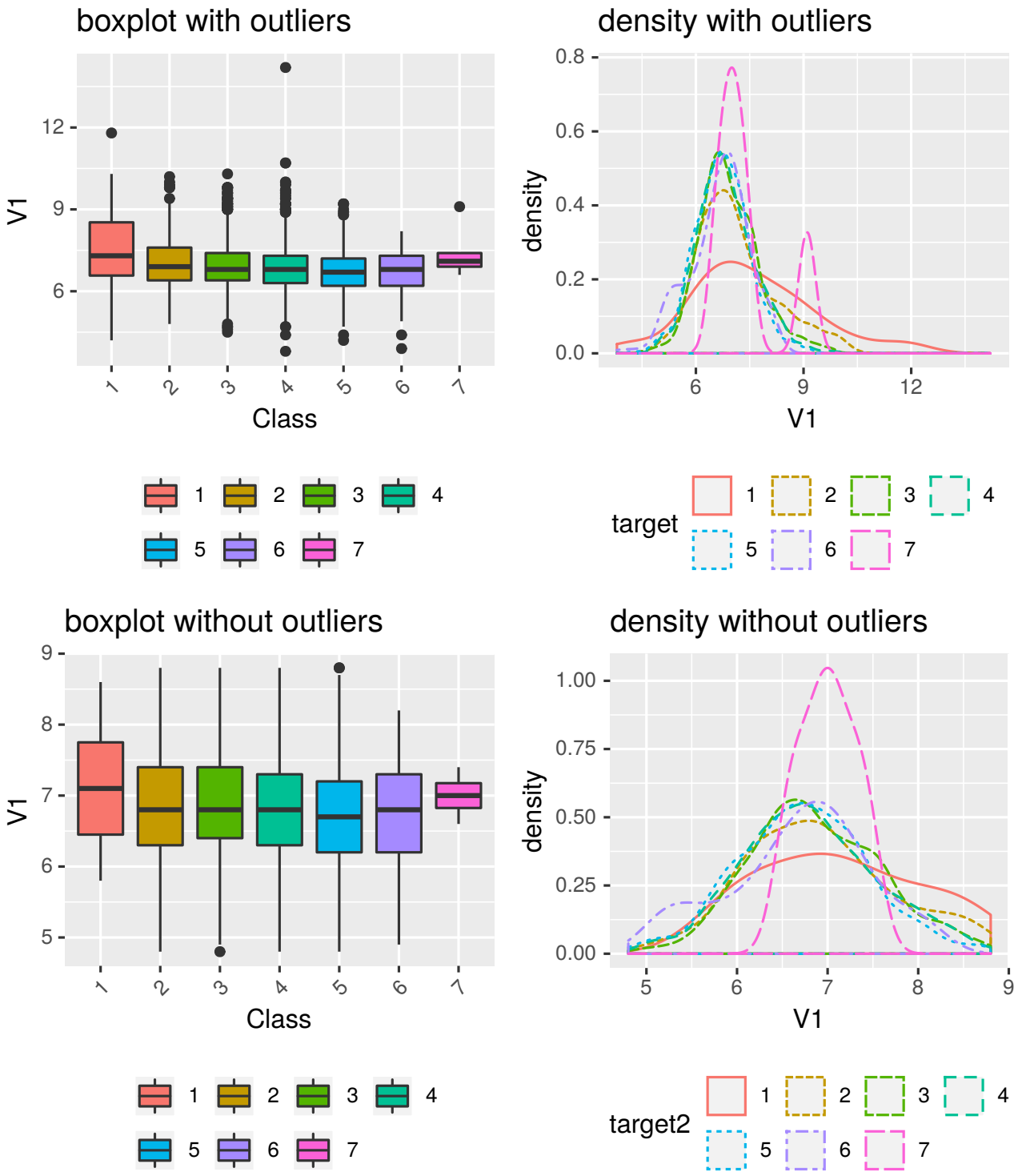

boxplot with outliers
V1
Class
density with outliers
density
V1
target
boxplot without outliers
V1
Class
density without outliers
density
V1
target2
1
2
3
4
5
6
7

Figure 4.1: V1

V2

Table 4.2: V2

	7	6	5	4	3	2	1
n	5.00	175.00	880.00	2,198.00	1,457.00	163.00	20.00
NA	0.00	0.00	0.00	0.00	0.00	0.00	0.00
mean	0.30	0.28	0.26	0.26	0.30	0.38	0.33
sd	0.06	0.11	0.09	0.09	0.10	0.17	0.14
se(mean)	0.03	0.01	0.00	0.00	0.00	0.01	0.03
IQR	0.10	0.13	0.13	0.10	0.10	0.19	0.18
skewness	0.45	0.99	0.81	1.53	1.43	1.40	1.03
kurtosis	-3.06	0.97	1.01	5.05	3.78	2.32	-0.19
0%	0.24	0.12	0.08	0.08	0.10	0.11	0.17
1%	0.24	0.12	0.11	0.12	0.14	0.16	0.18
5%	0.24	0.15	0.14	0.15	0.17	0.19	0.20
10%	0.25	0.16	0.16	0.17	0.20	0.21	0.21
20%	0.26	0.19	0.18	0.19	0.23	0.26	0.23
25%	0.26	0.20	0.19	0.20	0.24	0.27	0.24
30%	0.26	0.21	0.20	0.21	0.25	0.28	0.24
40%	0.27	0.24	0.23	0.23	0.27	0.30	0.26
50%	0.27	0.26	0.25	0.25	0.28	0.32	0.26
60%	0.31	0.28	0.28	0.27	0.31	0.36	0.32
70%	0.34	0.32	0.30	0.29	0.33	0.42	0.35
75%	0.36	0.33	0.32	0.30	0.34	0.46	0.41
80%	0.36	0.36	0.33	0.32	0.36	0.53	0.48
90%	0.36	0.44	0.39	0.37	0.42	0.63	0.55
95%	0.36	0.47	0.44	0.42	0.49	0.67	0.59
99%	0.36	0.58	0.50	0.58	0.62	0.96	0.63
100%	0.36	0.66	0.76	0.96	0.90	1.10	0.64

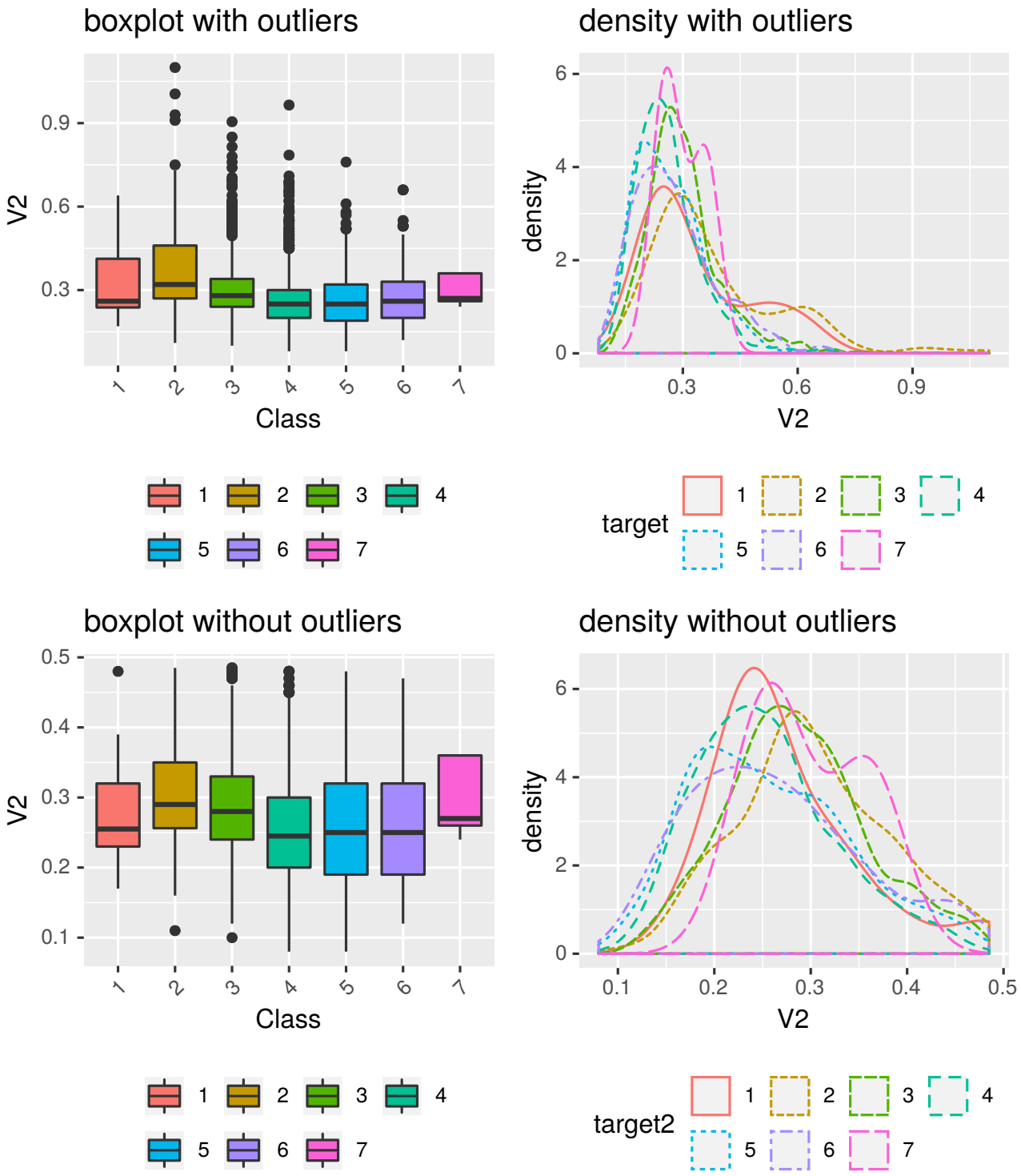

boxplot with outliers
V2
0.9
0.6
0.3
1
2
3
4
5
6
7
Class
1
2
3
4
5
6
7
density with outliers
density
6
4
2
0
0.3
0.6
0.9
V2
target
1
2
3
4
5
6
7
boxplot without outliers
V2
0.5
0.4
0.3
0.2
0.1
1
2
3
4
5
6
7
Class
1
2
3
4
5
6
7
density without outliers
density
6
4
2
0
0.1
0.2
0.3
0.4
0.5
V2
target2
1
2
3
4
5
6
7

Figure 4.2: V2

V3

Table 4.3: V3

	7	6	5	4	3	2	1
n	5.00	175.00	880.00	2,198.00	1,457.00	163.00	20.00
NA	0.00	0.00	0.00	0.00	0.00	0.00	0.00
mean	0.39	0.33	0.33	0.34	0.34	0.30	0.34
sd	0.08	0.09	0.08	0.12	0.14	0.16	0.08
se(mean)	0.04	0.01	0.00	0.00	0.00	0.01	0.02
IQR	0.11	0.08	0.08	0.11	0.17	0.21	0.13
skewness	0.29	0.55	0.75	2.00	0.70	0.47	-0.03
kurtosis	-1.83	4.00	2.87	11.73	1.00	0.47	-1.10
0%	0.29	0.04	0.01	0.00	0.00	0.00	0.21
1%	0.29	0.11	0.14	0.09	0.04	0.00	0.21
5%	0.30	0.22	0.23	0.20	0.14	0.04	0.22
10%	0.31	0.26	0.25	0.23	0.18	0.10	0.23
20%	0.33	0.27	0.27	0.26	0.23	0.17	0.25
25%	0.34	0.28	0.28	0.27	0.24	0.19	0.26
30%	0.34	0.29	0.28	0.28	0.26	0.22	0.28
40%	0.35	0.30	0.30	0.30	0.29	0.26	0.33
50%	0.36	0.32	0.31	0.32	0.32	0.29	0.34
60%	0.40	0.34	0.33	0.34	0.35	0.34	0.36
70%	0.43	0.36	0.35	0.37	0.39	0.37	0.38
75%	0.45	0.36	0.36	0.38	0.41	0.40	0.38
80%	0.46	0.37	0.38	0.40	0.45	0.45	0.40
90%	0.47	0.44	0.42	0.49	0.51	0.49	0.44
95%	0.48	0.48	0.49	0.52	0.60	0.59	0.46
99%	0.49	0.55	0.57	0.74	0.74	0.73	0.47
100%	0.49	0.74	0.74	1.66	1.00	0.88	0.47

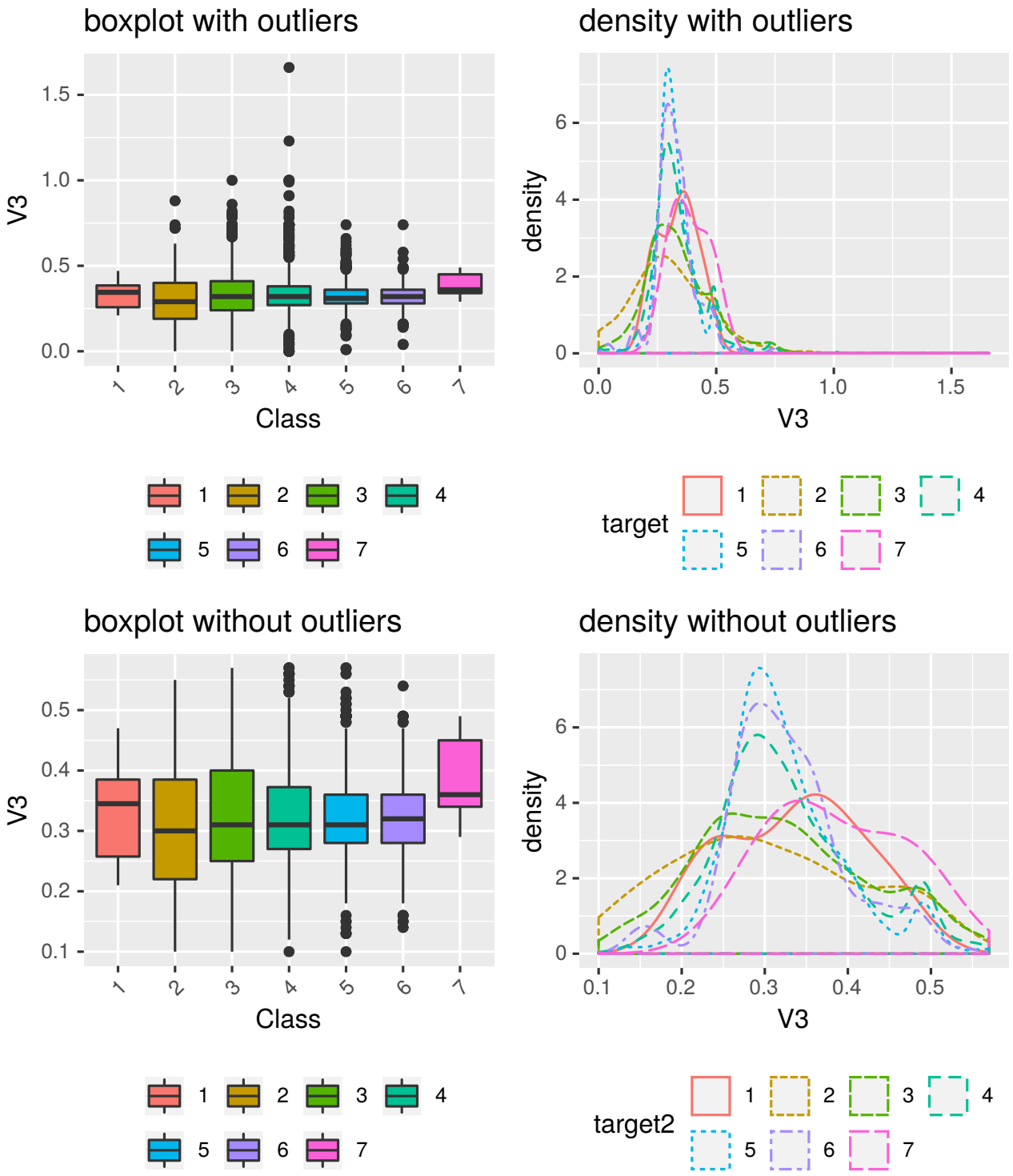

boxplot with outliers
V3
Class
density with outliers
density
V3
target
boxplot without outliers
V3
Class
density without outliers
density
V3
target2
1
2
3
4
5
6
7

Figure 4.3: V3

V4

Table 4.4: V4

	7	6	5	4	3	2	1
n	5.00	175.00	880.00	2,198.00	1,457.00	163.00	20.00
NA	0.00	0.00	0.00	0.00	0.00	0.00	0.00
mean	4.12	5.67	5.19	6.44	7.33	4.63	6.39
sd	3.76	4.26	4.30	5.17	5.33	4.16	5.32
se(mean)	1.68	0.32	0.14	0.11	0.14	0.33	1.19
IQR	2.20	6.10	5.62	8.20	9.70	5.80	9.11
skewness	1.88	0.85	1.13	1.43	0.49	1.11	0.66
kurtosis	3.50	-0.50	0.25	7.66	-0.79	0.20	-0.99
0%	1.60	0.80	0.90	0.70	0.60	0.70	0.70
1%	1.62	1.00	1.00	0.90	0.80	0.80	0.79
5%	1.68	1.17	1.20	1.10	1.10	0.90	1.13
10%	1.76	1.40	1.30	1.20	1.20	1.10	1.19
20%	1.92	1.80	1.60	1.50	1.60	1.22	1.52
25%	2.00	2.10	1.70	1.70	1.80	1.30	1.59
30%	2.04	2.20	1.90	2.00	2.29	1.40	1.88
40%	2.12	3.56	2.50	3.50	5.10	1.78	3.26
50%	2.20	4.30	3.65	5.30	7.00	2.50	4.60
60%	3.00	5.82	4.84	7.10	8.20	4.40	6.46
70%	3.80	7.18	6.30	8.60	10.30	6.02	9.83
75%	4.20	8.20	7.32	9.90	11.50	7.10	10.70
80%	5.48	9.12	8.80	11.16	12.30	8.42	11.02
90%	8.04	13.90	12.80	13.80	15.20	10.96	15.15
95%	9.32	14.20	14.40	16.00	17.20	12.99	15.58
99%	10.34	14.58	15.55	18.95	19.30	15.34	16.08
100%	10.60	14.80	19.25	65.80	23.50	17.55	16.20

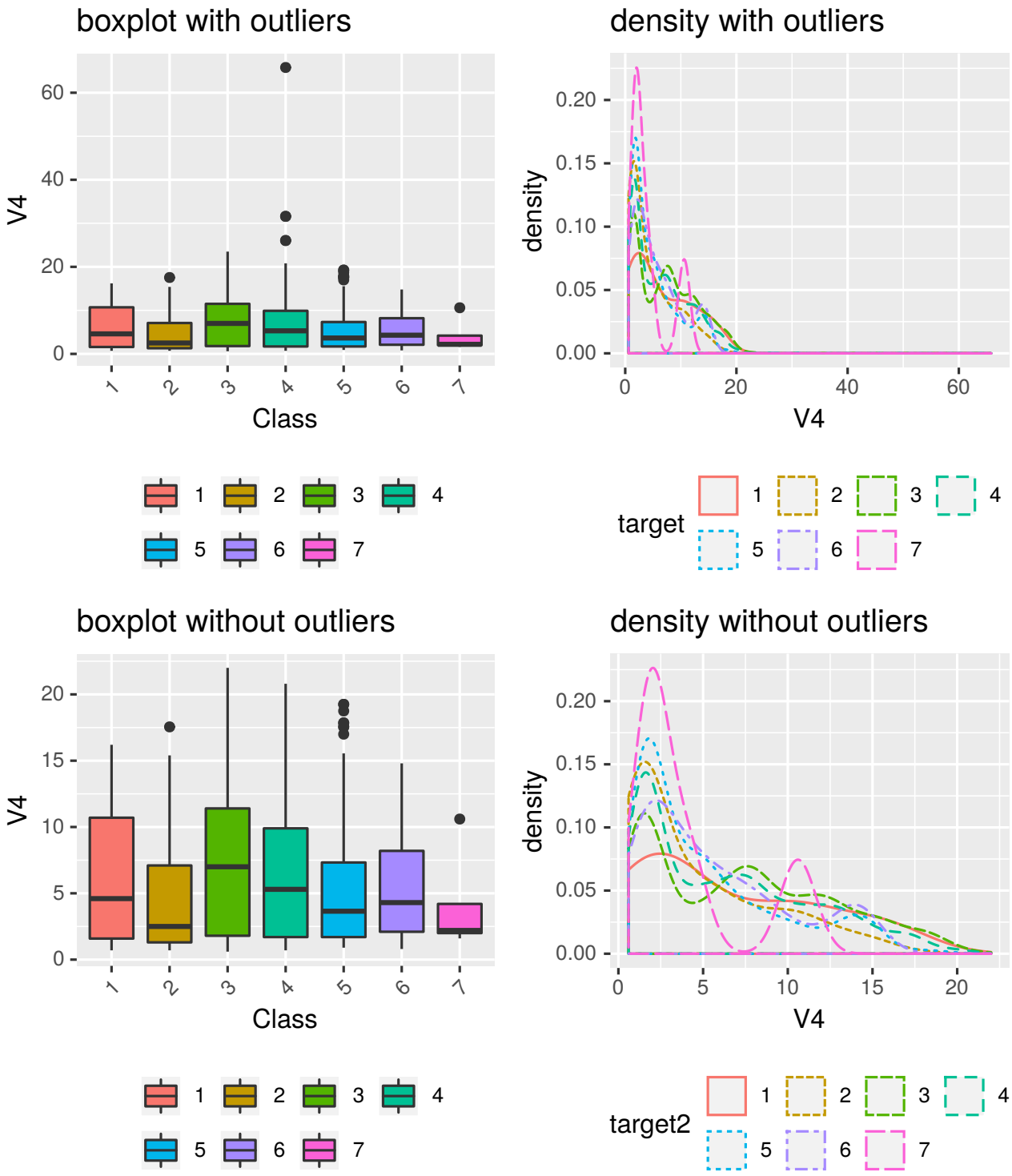

boxplot with outliers
V4
Class
density with outliers
density
V4
target
1
2
3
4
5
6
7
boxplot without outliers
V4
Class
density without outliers
density
V4
target2

Figure 4.4: V4

V5

Table 4.5: V5

	7	6	5	4	3	2	1
n	5.00	175.00	880.00	2,198.00	1,457.00	163.00	20.00
NA	0.00	0.00	0.00	0.00	0.00	0.00	0.00
mean	0.03	0.04	0.04	0.05	0.05	0.05	0.05
sd	0.01	0.01	0.01	0.02	0.03	0.03	0.05
se(mean)	0.00	0.00	0.00	0.00	0.00	0.00	0.01
IQR	0.01	0.01	0.01	0.01	0.01	0.02	0.02
skewness	-0.52	2.96	2.00	4.61	4.76	5.64	3.93
kurtosis	-2.51	16.26	13.73	28.85	31.21	46.73	16.56
0%	0.02	0.01	0.01	0.02	0.01	0.01	0.02
1%	0.02	0.02	0.02	0.02	0.02	0.02	0.02
5%	0.02	0.03	0.02	0.03	0.03	0.03	0.03
10%	0.02	0.03	0.03	0.03	0.03	0.03	0.03
20%	0.02	0.03	0.03	0.03	0.04	0.04	0.03
25%	0.02	0.03	0.03	0.04	0.04	0.04	0.04
30%	0.02	0.03	0.03	0.04	0.04	0.04	0.04
40%	0.03	0.03	0.04	0.04	0.04	0.04	0.04
50%	0.03	0.04	0.04	0.04	0.05	0.05	0.04
60%	0.03	0.04	0.04	0.04	0.05	0.05	0.05
70%	0.03	0.04	0.04	0.05	0.05	0.05	0.05
75%	0.03	0.04	0.04	0.05	0.05	0.05	0.05
80%	0.03	0.04	0.05	0.05	0.06	0.06	0.06
90%	0.03	0.06	0.05	0.06	0.06	0.06	0.07
95%	0.03	0.06	0.06	0.06	0.09	0.08	0.08
99%	0.03	0.08	0.06	0.16	0.18	0.14	0.21
100%	0.04	0.12	0.14	0.26	0.35	0.29	0.24

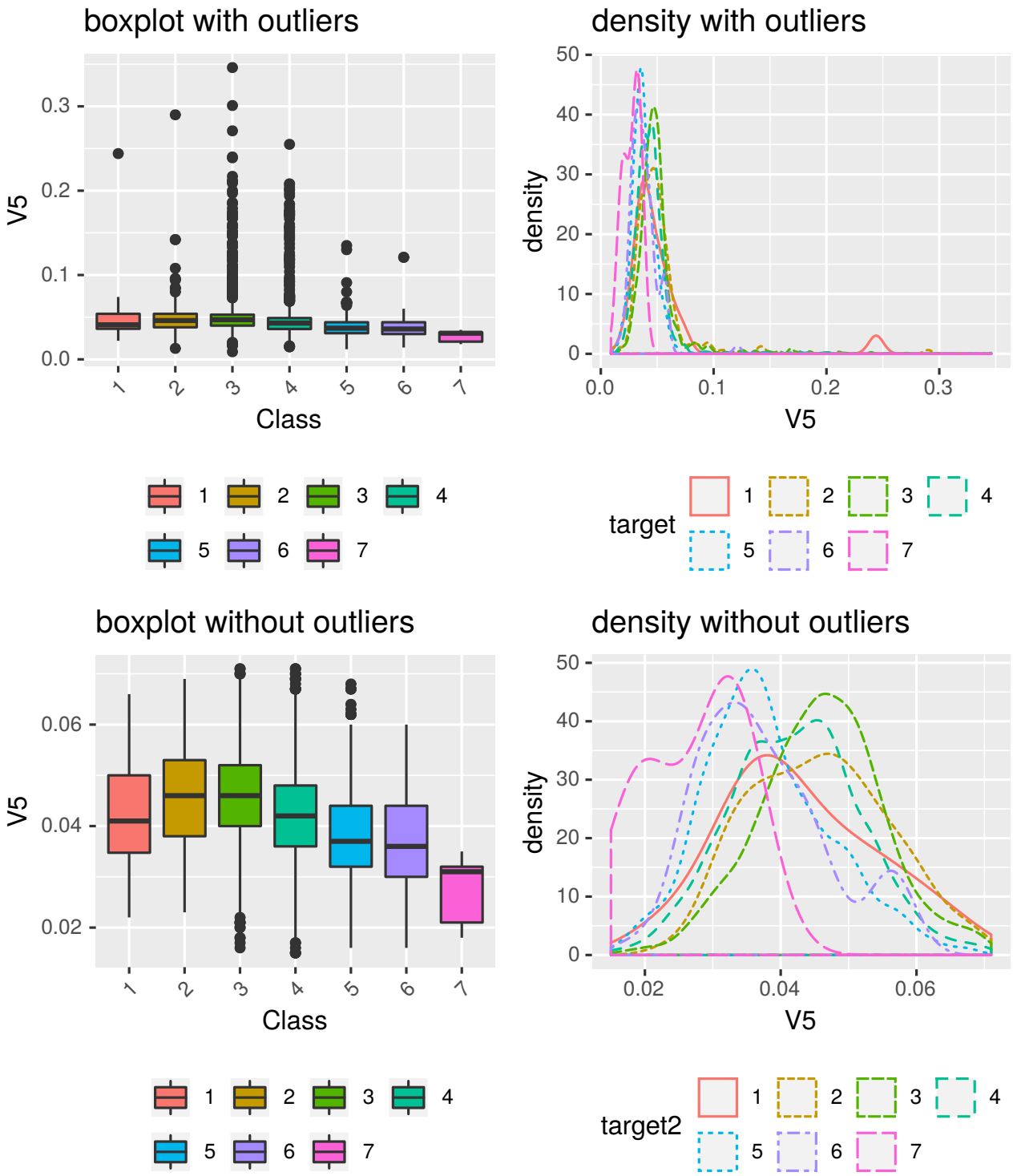

boxplot with outliers
V5
0.0
0.1
0.2
0.3
1
2
3
4
5
6
7
Class
1
2
3
4
5
6
7
density with outliers
density
0
10
20
30
40
50
0.0
0.1
0.2
0.3
V5
target
1
2
3
4
5
6
7
boxplot without outliers
V5
0.02
0.04
0.06
1
2
3
4
5
6
7
Class
1
2
3
4
5
6
7
density without outliers
density
0
10
20
30
40
50
0.02
0.04
0.06
V5
target2
1
2
3
4
5
6
7

Figure 4.5: V5

V6

Table 4.6: V6

	7	6	5	4	3	2	1
n	5.00	175.00	880.00	2,198.00	1,457.00	163.00	20.00
NA	0.00	0.00	0.00	0.00	0.00	0.00	0.00
mean	33.40	36.72	34.13	35.65	36.43	23.36	53.33
sd	13.43	16.20	13.24	15.74	18.15	20.39	69.42
se(mean)	6.00	1.22	0.45	0.34	0.48	1.60	15.52
IQR	4.00	16.50	16.00	22.00	28.00	21.50	34.25
skewness	2.04	1.46	0.81	0.68	0.48	2.36	2.41
kurtosis	4.33	4.22	1.74	0.79	0.40	8.64	6.53
0%	24.00	6.00	5.00	3.00	2.00	3.00	5.00
1%	24.12	10.70	9.79	7.00	5.56	3.00	5.00
5%	24.60	15.00	14.00	13.85	10.00	5.00	5.00
10%	25.20	18.40	18.00	17.00	13.00	6.00	5.00
20%	26.40	24.80	23.00	22.00	19.60	8.00	7.40
25%	27.00	28.00	25.00	24.00	22.00	9.00	13.25
30%	27.20	29.00	27.00	26.00	24.00	10.00	15.70
40%	27.60	31.00	29.00	30.00	30.00	14.00	21.60
50%	28.00	35.00	33.00	34.00	35.00	18.00	33.50
60%	29.20	38.00	36.00	38.00	41.00	23.00	36.50
70%	30.40	43.00	40.00	42.90	47.00	28.00	41.30
75%	31.00	44.50	41.00	46.00	50.00	30.50	47.50
80%	36.20	45.20	45.00	49.00	53.00	34.00	74.90
90%	46.60	53.00	50.00	56.00	60.00	49.60	126.25
95%	51.80	62.60	58.00	63.00	66.00	61.90	153.63
99%	55.96	98.34	73.11	78.05	81.00	93.05	261.92
100%	57.00	105.00	108.00	112.00	131.00	138.50	289.00

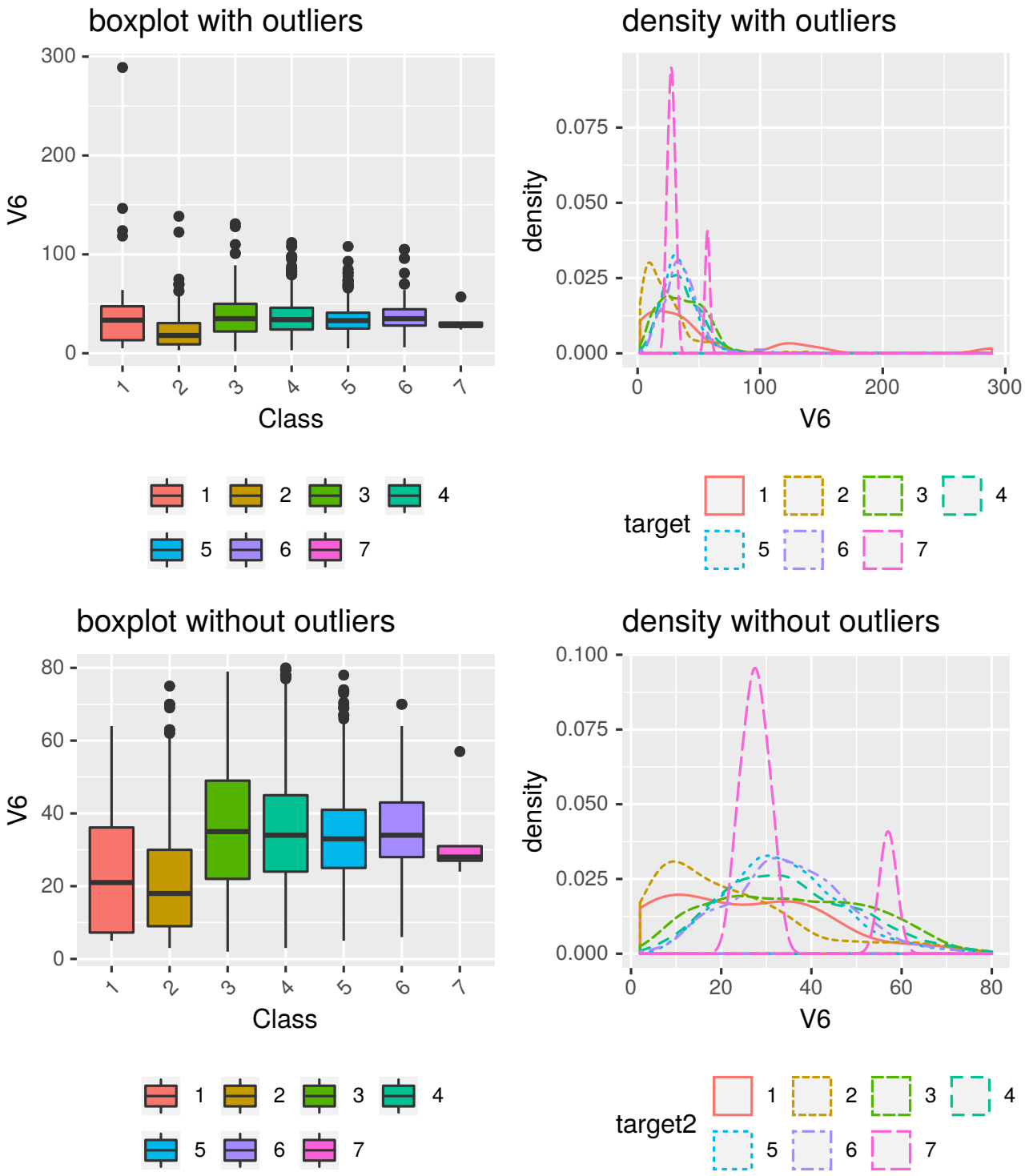

boxplot with outliers
V6
Class
density with outliers
density
V6
target
boxplot without outliers
V6
Class
density without outliers
density
V6
target2
1
2
3
4
5
6
7

Figure 4.6: V6

V7

Table 4.7: V7

	7	6	5	4	3	2	1
n	5.00	175.00	880.00	2,198.00	1,457.00	163.00	20.00
NA	0.00	0.00	0.00	0.00	0.00	0.00	0.00
mean	116.00	126.17	125.11	137.05	150.90	125.28	170.60
sd	19.82	33.01	32.74	41.29	44.09	52.75	107.76
se(mean)	8.87	2.50	1.10	0.88	1.15	4.13	24.10
IQR	11.00	47.50	43.25	56.75	61.00	86.50	104.25
skewness	-0.92	0.55	0.50	0.36	-0.03	0.21	0.95
kurtosis	1.77	0.07	0.29	-0.14	0.08	-0.65	0.93
0%	85.00	59.00	34.00	18.00	9.00	10.00	19.00
1%	86.12	67.14	64.79	54.97	53.56	25.62	21.66
5%	90.60	75.70	78.00	76.00	76.00	50.00	32.30
10%	96.20	86.40	87.00	88.00	93.00	61.00	54.60
20%	107.40	97.00	97.00	101.00	114.00	76.00	90.00
25%	113.00	102.50	101.00	107.25	121.00	85.00	105.75
30%	114.20	107.00	106.00	112.00	128.00	93.80	110.40
40%	116.60	115.00	114.00	121.00	140.00	106.80	129.60
50%	119.00	122.00	122.00	132.00	151.00	117.00	159.50
60%	121.00	128.00	130.40	145.00	164.00	136.20	188.00
70%	123.00	142.80	140.00	158.00	176.10	161.40	203.10
75%	124.00	150.00	144.25	164.00	182.00	171.50	210.00
80%	127.00	155.00	150.00	172.00	187.00	178.00	218.20
90%	133.00	170.00	168.00	193.00	206.40	194.60	313.40
95%	136.00	181.80	189.00	212.00	220.00	202.80	370.18
99%	138.40	212.50	212.42	240.06	248.44	238.49	426.03
100%	139.00	212.50	229.00	294.00	344.00	272.00	440.00

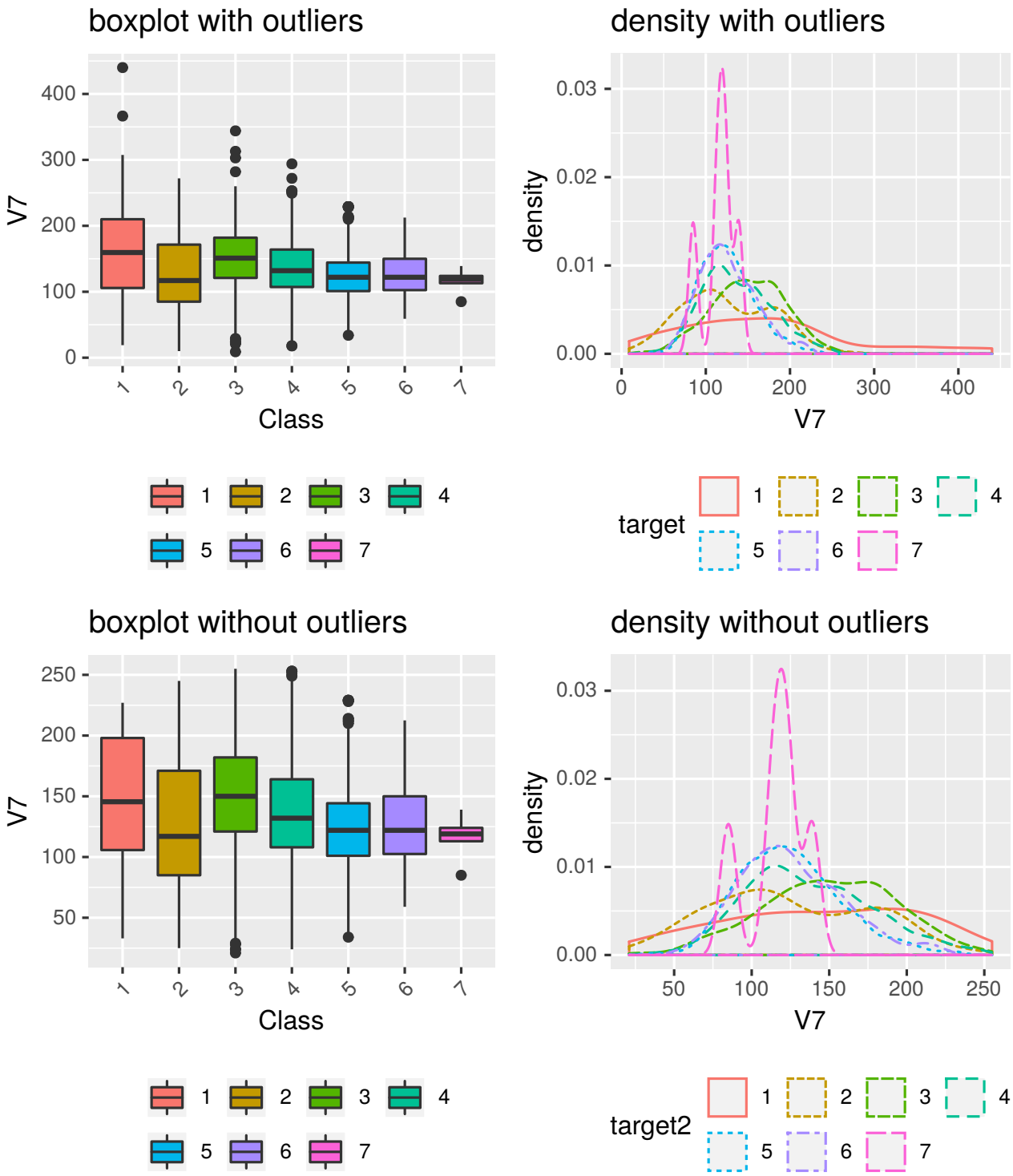

boxplot with outliers
V7
Class
density with outliers
density
V7
target
boxplot without outliers
V7
Class
density without outliers
density
V7
target2
1
2
3
4
5
6
7

Figure 4.7: V7

V8

Table 4.8: V8

	7	6	5	4	3	2	1
n	5.00	175.00	880.00	2,198.00	1,457.00	163.00	20.00
NA	0.00	0.00	0.00	0.00	0.00	0.00	0.00
mean	0.99	0.99	0.99	0.99	1.00	0.99	0.99
sd	0.00	0.00	0.00	0.00	0.00	0.00	0.00
se(mean)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
IQR	0.00	0.00	0.00	0.00	0.00	0.00	0.00
skewness	2.16	1.14	0.98	1.85	0.06	0.19	0.36
kurtosis	4.73	0.90	0.35	21.97	-0.57	-0.58	-1.00
0%	0.99	0.99	0.99	0.99	0.99	0.99	0.99
1%	0.99	0.99	0.99	0.99	0.99	0.99	0.99
5%	0.99	0.99	0.99	0.99	0.99	0.99	0.99
10%	0.99	0.99	0.99	0.99	0.99	0.99	0.99
20%	0.99	0.99	0.99	0.99	0.99	0.99	0.99
25%	0.99	0.99	0.99	0.99	0.99	0.99	0.99
30%	0.99	0.99	0.99	0.99	0.99	0.99	0.99
40%	0.99	0.99	0.99	0.99	0.99	0.99	0.99
50%	0.99	0.99	0.99	0.99	1.00	0.99	0.99
60%	0.99	0.99	0.99	0.99	1.00	0.99	1.00
70%	0.99	0.99	0.99	1.00	1.00	1.00	1.00
75%	0.99	0.99	0.99	1.00	1.00	1.00	1.00
80%	0.99	0.99	0.99	1.00	1.00	1.00	1.00
90%	0.99	1.00	1.00	1.00	1.00	1.00	1.00
95%	1.00	1.00	1.00	1.00	1.00	1.00	1.00
99%	1.00	1.00	1.00	1.00	1.00	1.00	1.00
100%	1.00	1.00	1.00	1.04	1.00	1.00	1.00

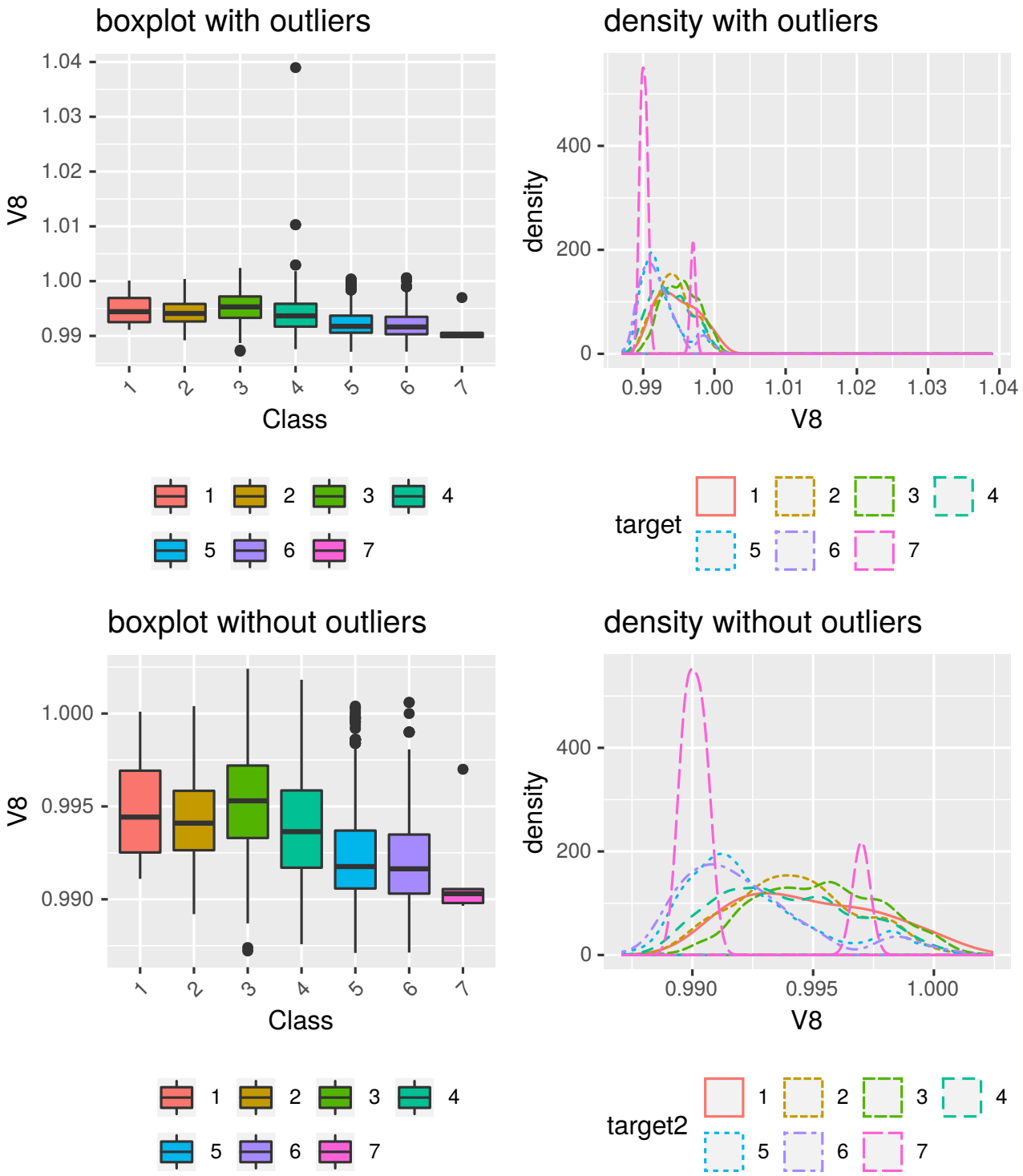

boxplot with outliers
V8
1.04
1.03
1.02
1.01
1.00
0.99
1
2
3
4
5
6
7
Class
1
2
3
4
5
6
7
density with outliers
density
400
200
0
0.99
1.00
1.01
1.02
1.03
1.04
V8
target
1
2
3
4
5
6
7
boxplot without outliers
V8
1.000
0.995
0.990
1
2
3
4
5
6
7
Class
1
2
3
4
5
6
7
density without outliers
density
400
200
0
0.990
0.995
1.000
V8
target2
1
2
3
4
5
6
7

Figure 4.8: V8

V9

Table 4.9: V9

	7	6	5	4	3	2	1
n	5.00	175.00	880.00	2,198.00	1,457.00	163.00	20.00
NA	0.00	0.00	0.00	0.00	0.00	0.00	0.00
mean	3.31	3.22	3.21	3.19	3.17	3.18	3.19
sd	0.08	0.15	0.16	0.15	0.14	0.16	0.21
se(mean)	0.04	0.01	0.01	0.00	0.00	0.01	0.05
IQR	0.09	0.21	0.22	0.20	0.16	0.21	0.29
skewness	0.00	0.09	0.27	0.42	0.65	0.66	0.10
kurtosis	-1.11	-0.47	-0.12	0.59	1.44	0.39	-0.92
0%	3.20	2.94	2.84	2.72	2.79	2.83	2.87
1%	3.20	2.94	2.90	2.88	2.87	2.89	2.87
5%	3.22	2.95	2.96	2.96	2.96	2.96	2.89
10%	3.23	3.01	3.00	3.00	3.00	3.00	2.90
20%	3.26	3.08	3.08	3.06	3.06	3.05	3.00
25%	3.28	3.12	3.10	3.08	3.08	3.07	3.04
30%	3.28	3.14	3.12	3.11	3.10	3.08	3.05
40%	3.28	3.18	3.16	3.14	3.13	3.12	3.14
50%	3.28	3.23	3.20	3.18	3.16	3.16	3.21
60%	3.32	3.26	3.25	3.22	3.19	3.19	3.24
70%	3.35	3.30	3.30	3.25	3.22	3.24	3.26
75%	3.37	3.33	3.32	3.28	3.24	3.28	3.33
80%	3.38	3.35	3.36	3.31	3.27	3.32	3.38
90%	3.39	3.41	3.42	3.38	3.34	3.40	3.45
95%	3.40	3.45	3.48	3.46	3.42	3.49	3.53
99%	3.41	3.56	3.59	3.60	3.58	3.64	3.55
100%	3.41	3.59	3.82	3.81	3.79	3.72	3.55

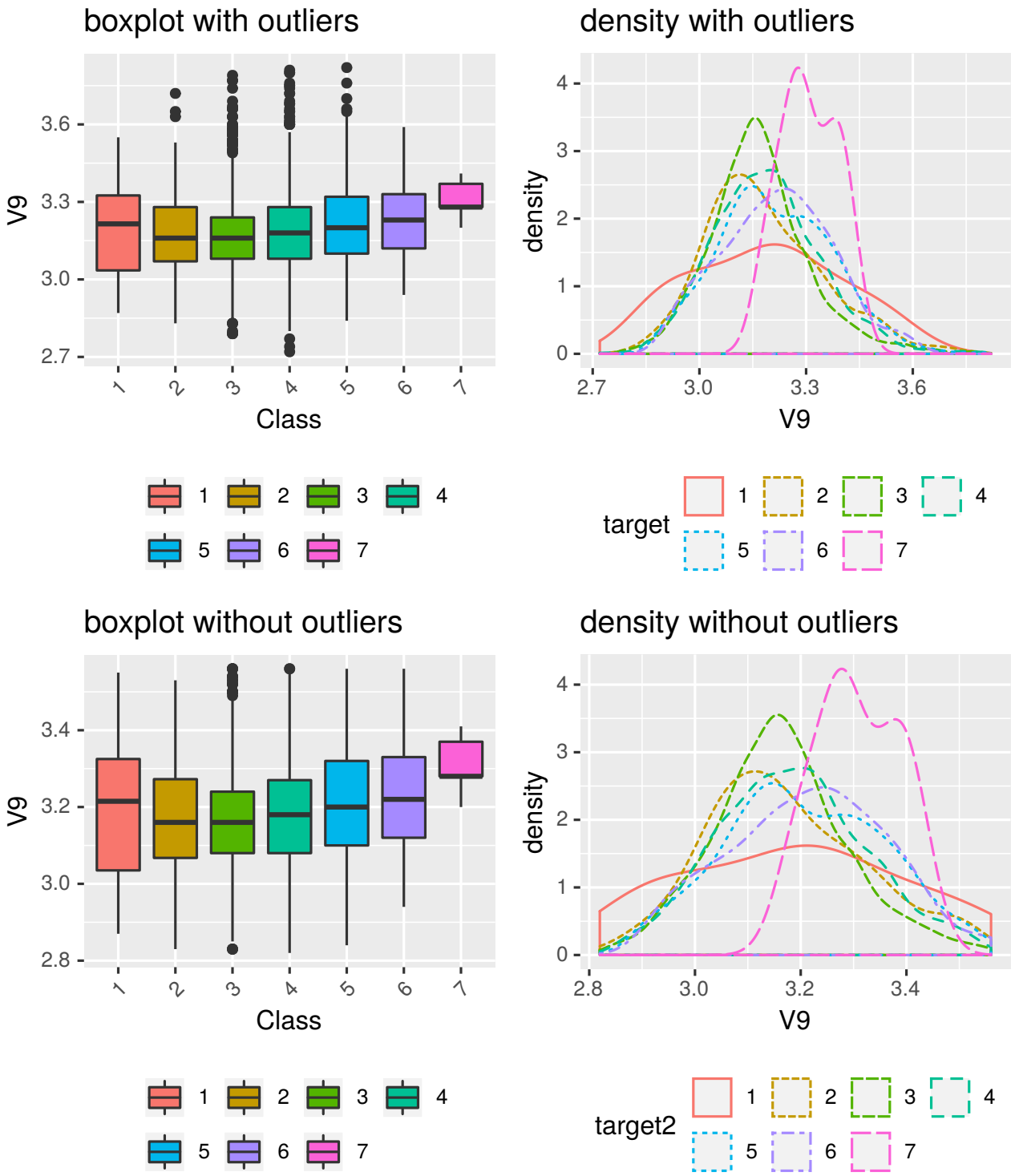

boxplot with outliers
V9
3.6
3.3
3.0
2.7
1
2
3
4
5
6
7
Class
1
2
3
4
5
6
7
density with outliers
density
4
3
2
1
0
2.7
3.0
3.3
3.6
V9
target
1
2
3
4
5
6
7
boxplot without outliers
V9
3.4
3.2
3.0
2.8
1
2
3
4
5
6
7
Class
1
2
3
4
5
6
7
density without outliers
density
4
3
2
1
0
2.8
3.0
3.2
3.4
V9
target2
1
2
3
4
5
6
7

Figure 4.9: V9

V10

Table 4.10: V10

	7	6	5	4	3	2	1
n	5.00	175.00	880.00	2,198.00	1,457.00	163.00	20.00
NA	0.00	0.00	0.00	0.00	0.00	0.00	0.00
mean	0.47	0.49	0.50	0.49	0.48	0.48	0.47
sd	0.09	0.15	0.13	0.11	0.10	0.12	0.12
se(mean)	0.04	0.01	0.00	0.00	0.00	0.01	0.03
IQR	0.06	0.20	0.17	0.14	0.11	0.16	0.16
skewness	0.89	0.97	0.95	0.98	0.85	0.70	0.54
kurtosis	1.52	0.64	1.22	1.70	1.11	0.54	-0.25
0%	0.36	0.25	0.22	0.23	0.27	0.25	0.28
1%	0.36	0.26	0.29	0.30	0.30	0.28	0.29
5%	0.37	0.32	0.33	0.34	0.34	0.31	0.32
10%	0.38	0.33	0.37	0.36	0.37	0.34	0.36
20%	0.41	0.37	0.39	0.39	0.40	0.37	0.38
25%	0.42	0.38	0.41	0.41	0.42	0.38	0.38
30%	0.43	0.39	0.43	0.42	0.43	0.40	0.39
40%	0.44	0.41	0.44	0.45	0.45	0.45	0.43
50%	0.46	0.46	0.48	0.48	0.47	0.47	0.44
60%	0.47	0.49	0.52	0.50	0.49	0.49	0.50
70%	0.48	0.55	0.55	0.54	0.51	0.53	0.53
75%	0.48	0.58	0.58	0.55	0.53	0.54	0.54
80%	0.51	0.60	0.60	0.57	0.55	0.57	0.57
90%	0.56	0.69	0.68	0.64	0.62	0.64	0.63
95%	0.58	0.77	0.74	0.70	0.68	0.68	0.64
99%	0.60	0.91	0.92	0.82	0.76	0.82	0.72
100%	0.61	0.95	1.08	1.06	0.88	0.87	0.74

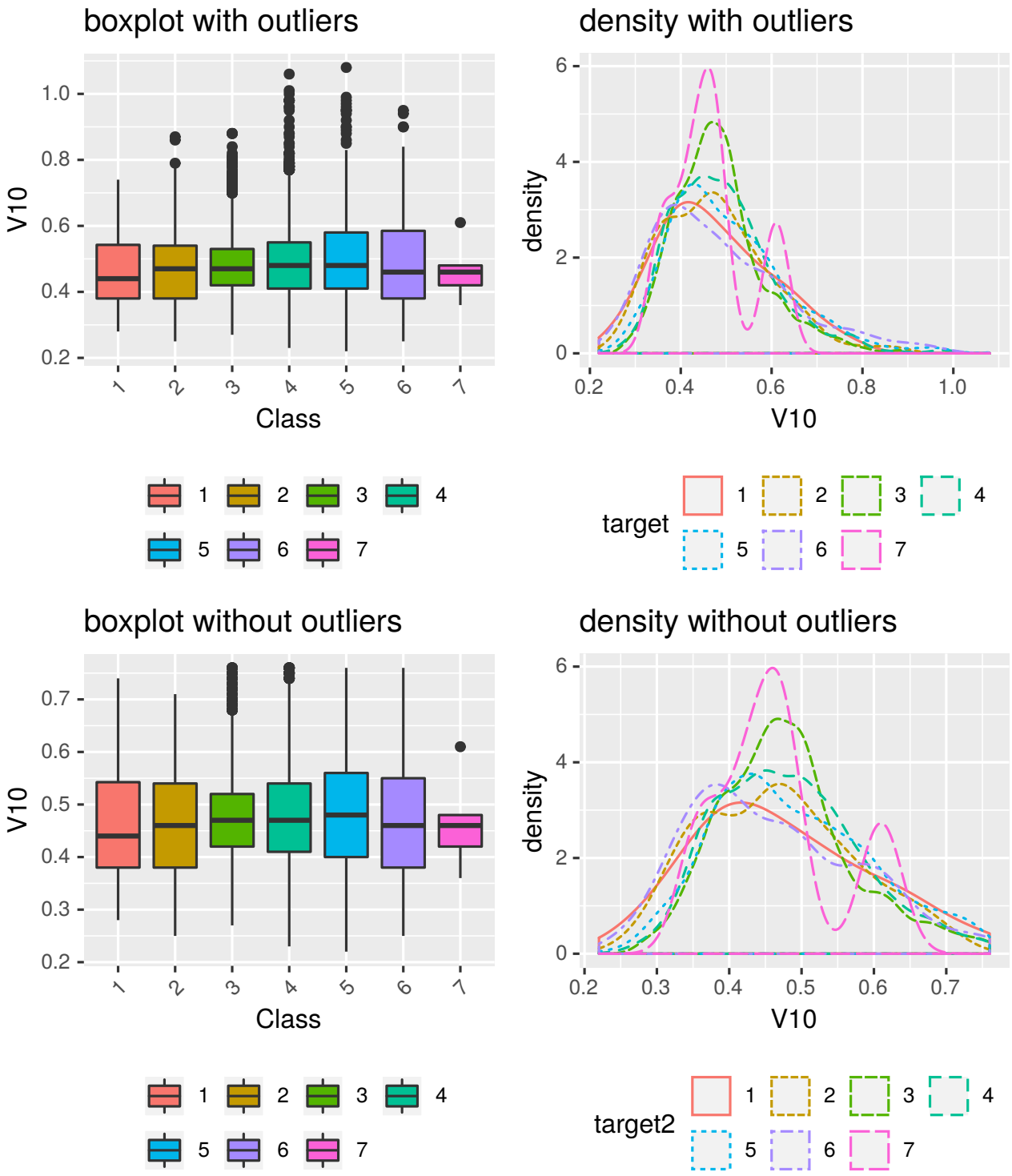


Figure 4.10: V10

V11

Table 4.11: V11

	7	6	5	4	3	2	1
n	5.00	175.00	880.00	2,198.00	1,457.00	163.00	20.00
NA	0.00	0.00	0.00	0.00	0.00	0.00	0.00
mean	12.18	11.64	11.37	10.58	9.81	10.15	10.35
sd	1.01	1.28	1.25	1.15	0.85	1.00	1.22
se(mean)	0.45	0.10	0.04	0.02	0.02	0.08	0.27
IQR	0.30	1.60	1.70	1.80	1.10	1.35	1.45
skewness	-2.03	-0.91	-0.31	0.40	1.08	0.71	0.02
kurtosis	4.32	0.08	-0.55	-0.71	1.09	0.24	-0.40
0%	10.40	8.50	8.60	8.50	8.00	8.40	8.00
1%	10.48	8.80	8.70	8.80	8.60	8.60	8.10
5%	10.80	8.80	9.00	9.00	8.80	8.81	8.47
10%	11.20	9.18	9.50	9.20	8.90	9.00	9.04
20%	12.00	10.70	10.40	9.50	9.10	9.20	9.34
25%	12.40	11.00	10.60	9.60	9.20	9.40	9.55
30%	12.42	11.20	10.80	9.80	9.30	9.50	9.67
40%	12.46	11.60	11.10	10.10	9.40	9.80	9.98
50%	12.50	12.00	11.40	10.50	9.50	10.10	10.45
60%	12.58	12.30	11.80	10.80	9.80	10.20	10.70
70%	12.66	12.50	12.20	11.20	10.10	10.50	11.00
75%	12.70	12.60	12.30	11.40	10.30	10.75	11.00
80%	12.74	12.70	12.50	11.60	10.50	11.06	11.10
90%	12.82	12.96	12.90	12.30	11.00	11.50	11.77
95%	12.86	13.20	13.20	12.60	11.40	11.98	12.41
99%	12.89	13.40	13.70	13.14	12.36	12.78	12.56
100%	12.90	14.00	14.20	14.00	13.60	13.50	12.60

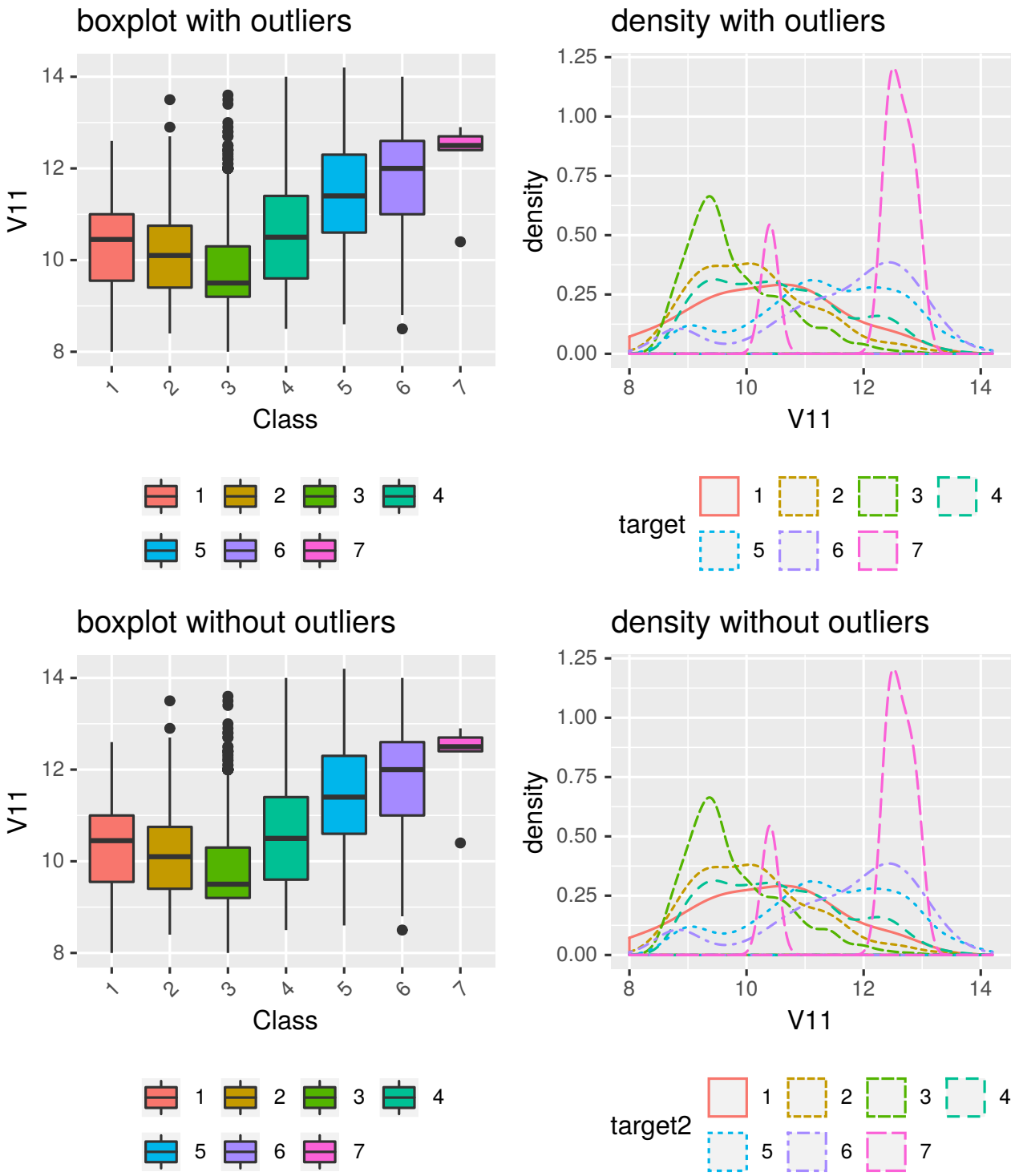


Figure 4.11: V11

4.1.2 Grouped Categorical Variables

There is no categorical variable.

4.2 Grouped Relationship Between Variables

4.2.1 Grouped Correlation Coefficient

Table 4.12: The correlation coefficients (0.5 or more)

Class	Variable1	Variable2	Correlation Coefficient
1	V7	V6	0.776
1	V9	V1	-0.755
1	V8	V4	0.709
1	V11	V8	-0.614
1	V10	V5	0.568
1	V5	V2	0.551
1	V6	V3	-0.549
1	V10	V7	0.543
1	V10	V8	0.537
2	V8	V4	0.745
2	V11	V8	-0.703
2	V7	V6	0.658
2	V8	V7	0.625
2	V3	V1	0.515
2	V11	V7	-0.514
3	V8	V4	0.893
3	V11	V8	-0.676
3	V7	V6	0.664
4	V8	V4	0.856
4	V11	V8	-0.744
4	V7	V6	0.596
5	V11	V8	-0.838
5	V8	V4	0.823
5	V8	V7	0.577
5	V11	V5	-0.555
5	V7	V6	0.533
5	V11	V2	0.503
6	V11	V8	-0.876
6	V8	V4	0.818
6	V7	V6	0.621
6	V8	V1	0.560
6	V11	V2	0.523
6	V11	V4	-0.522
6	V11	V5	-0.512
7	V8	V1	0.982
7	V11	V8	-0.969
7	V8	V4	0.949
7	V11	V4	-0.926
7	V11	V1	-0.923
7	V5	V2	-0.920

Table 4.12: The correlation coefficients (0.5 or more) *(continued)*

Class	Variable1	Variable2	Correlation Coefficient
7	V4	V1	0.918
7	V9	V1	-0.828
7	V9	V4	-0.817
7	V5	V3	0.786
7	V10	V6	-0.764
7	V9	V8	-0.757
7	V9	V7	-0.752
7	V5	V1	0.730
7	V11	V9	0.702
7	V7	V2	-0.696
7	V3	V2	-0.690
7	V8	V5	0.654
7	V10	V7	-0.595
7	V6	V5	-0.593
7	V3	V1	0.550
7	V2	V1	-0.515
7	V11	V5	-0.510
7	V7	V5	0.509

4.2.2 Grouped Correlation Plot of Numerical Variables

- Grouped Correlation Case of (Class == 1)
- Grouped Correlation Case of (Class == 2)
- Grouped Correlation Case of (Class == 3)
- Grouped Correlation Case of (Class == 4)
- Grouped Correlation Case of (Class == 5)
- Grouped Correlation Case of (Class == 6)
- Grouped Correlation Case of (Class == 7)

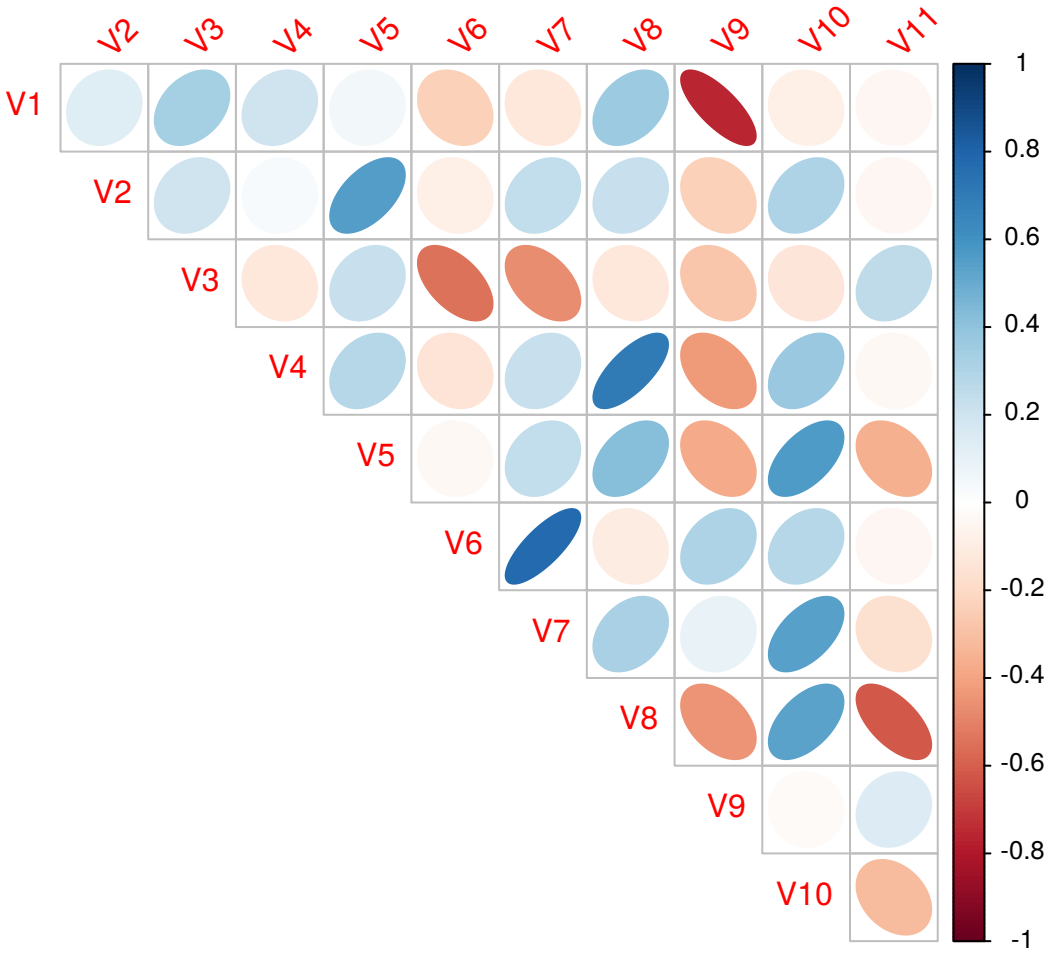


Figure 4.12: Correlation Matrix Plot (Class == 1)

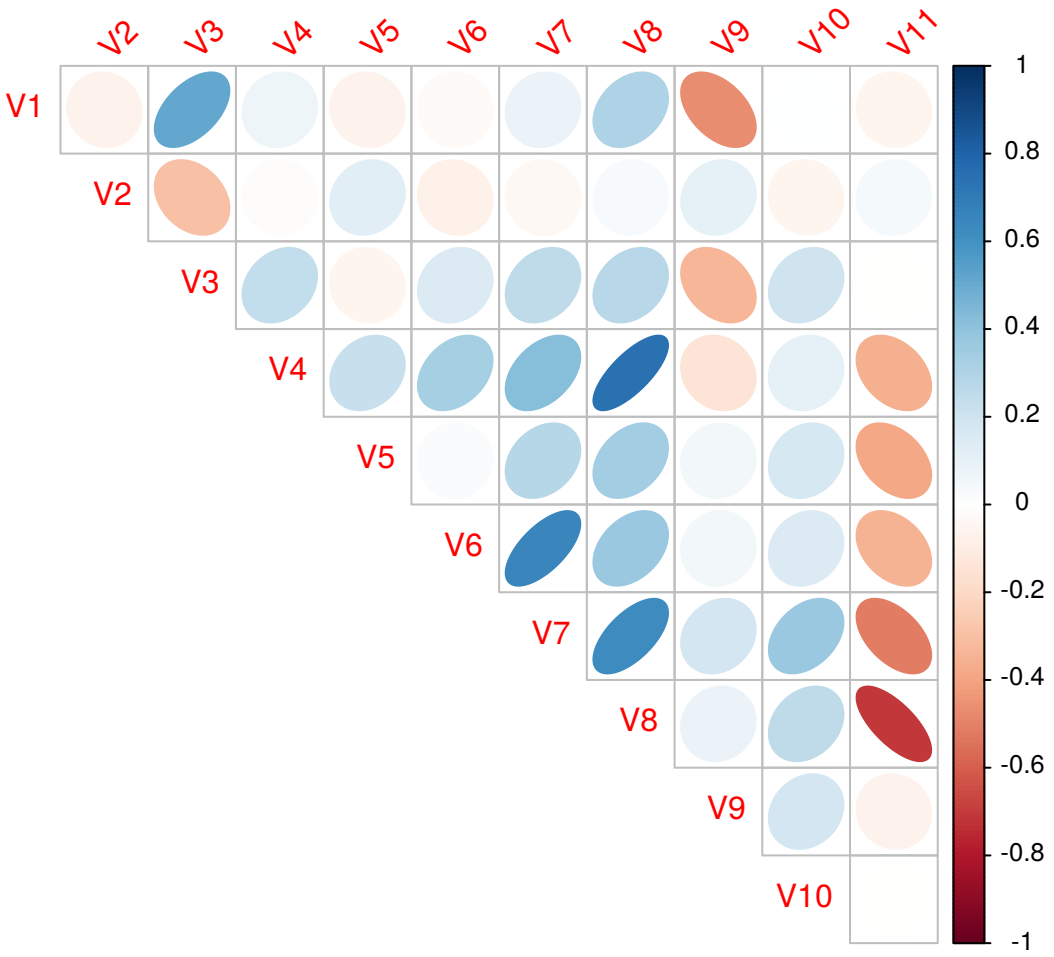


Figure 4.13: Correlation Matrix Plot (Class == 2)

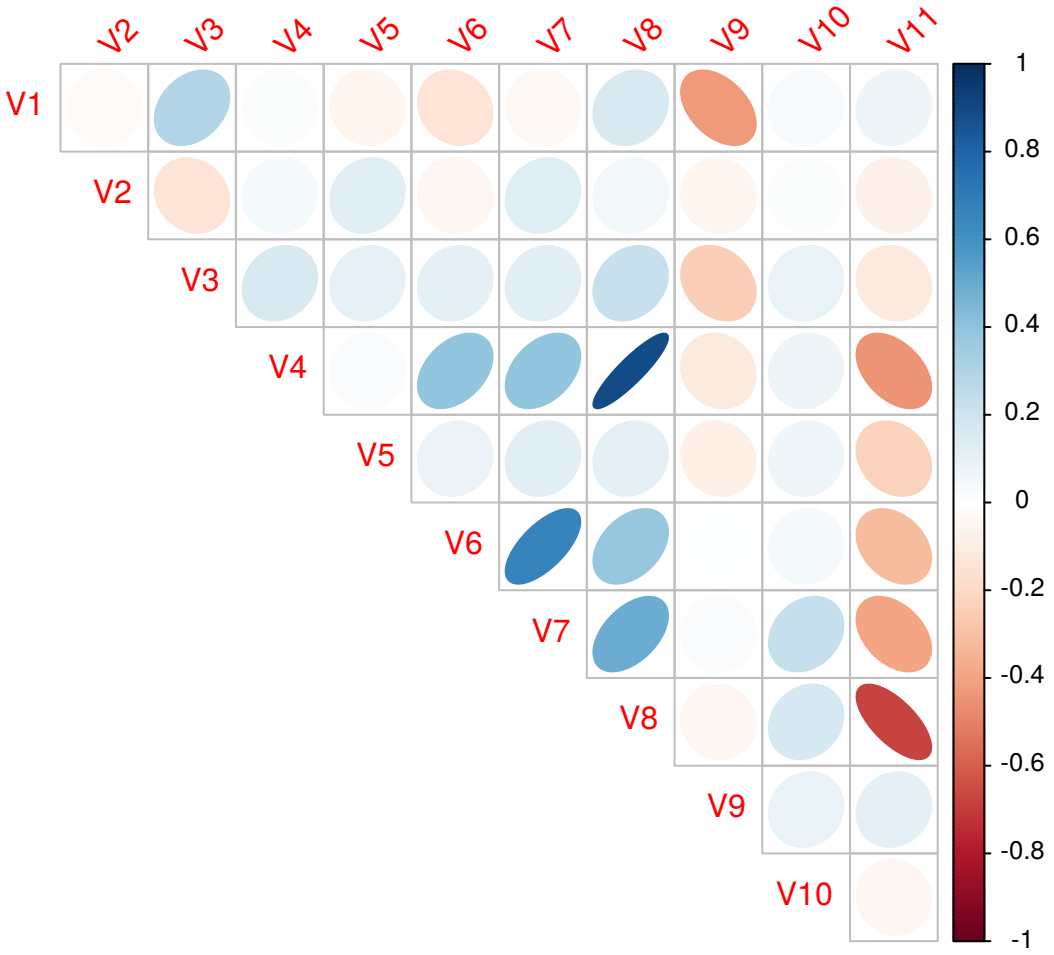


Figure 4.14: Correlation Matrix Plot (Class == 3)

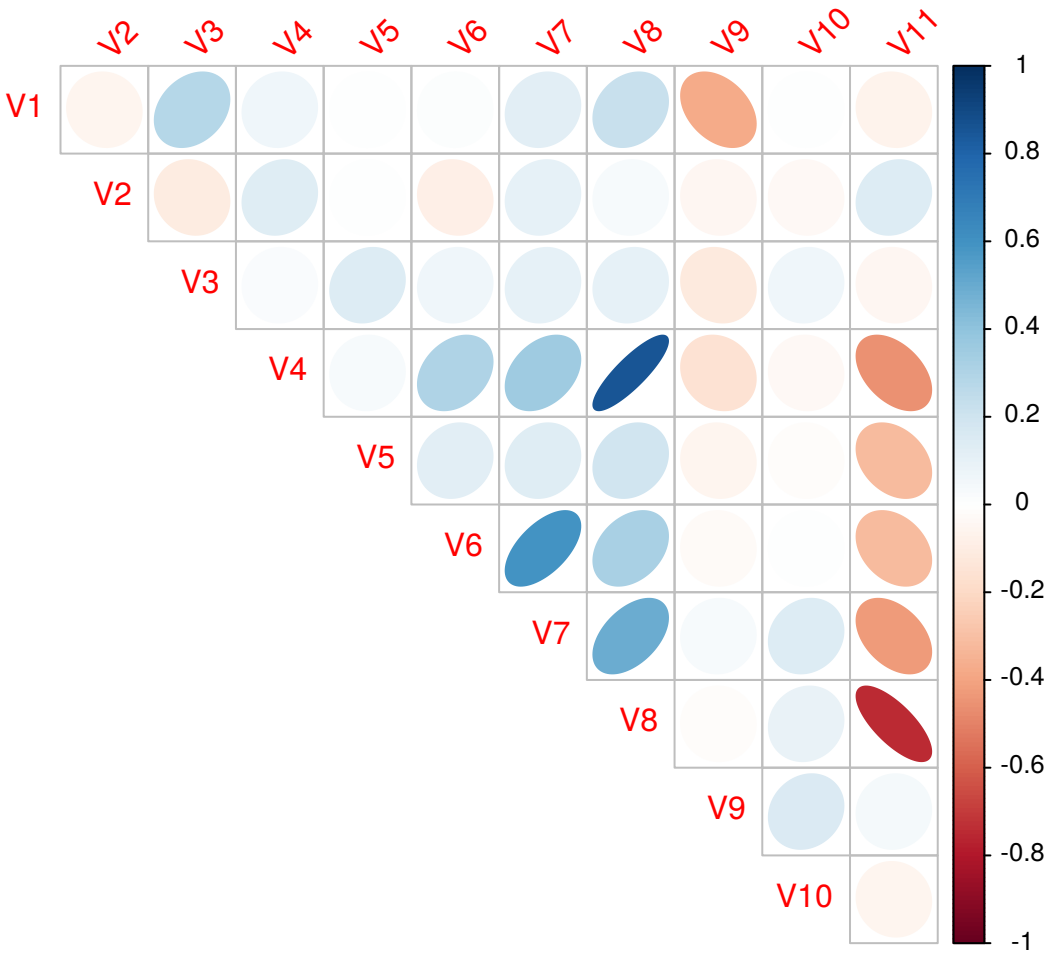


Figure 4.15: Correlation Matrix Plot (Class == 4)

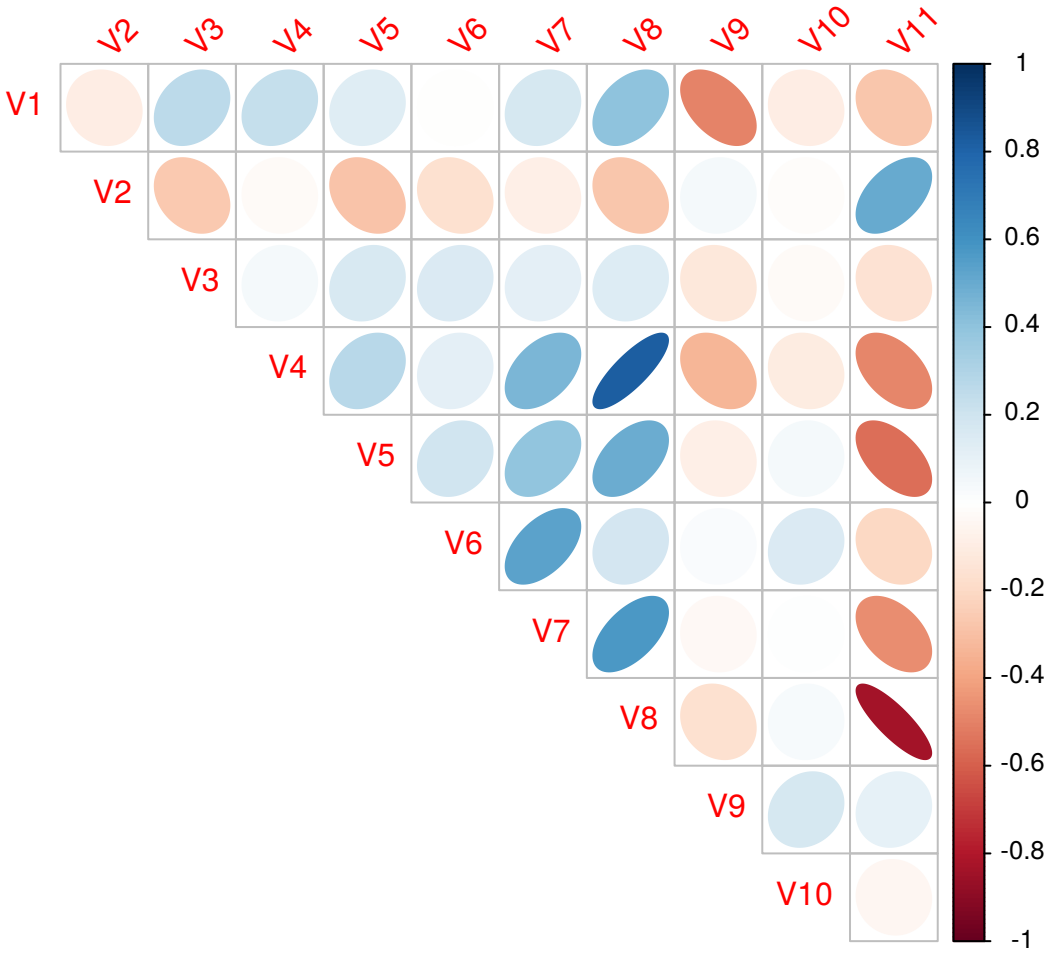


Figure 4.16: Correlation Matrix Plot (Class == 5)

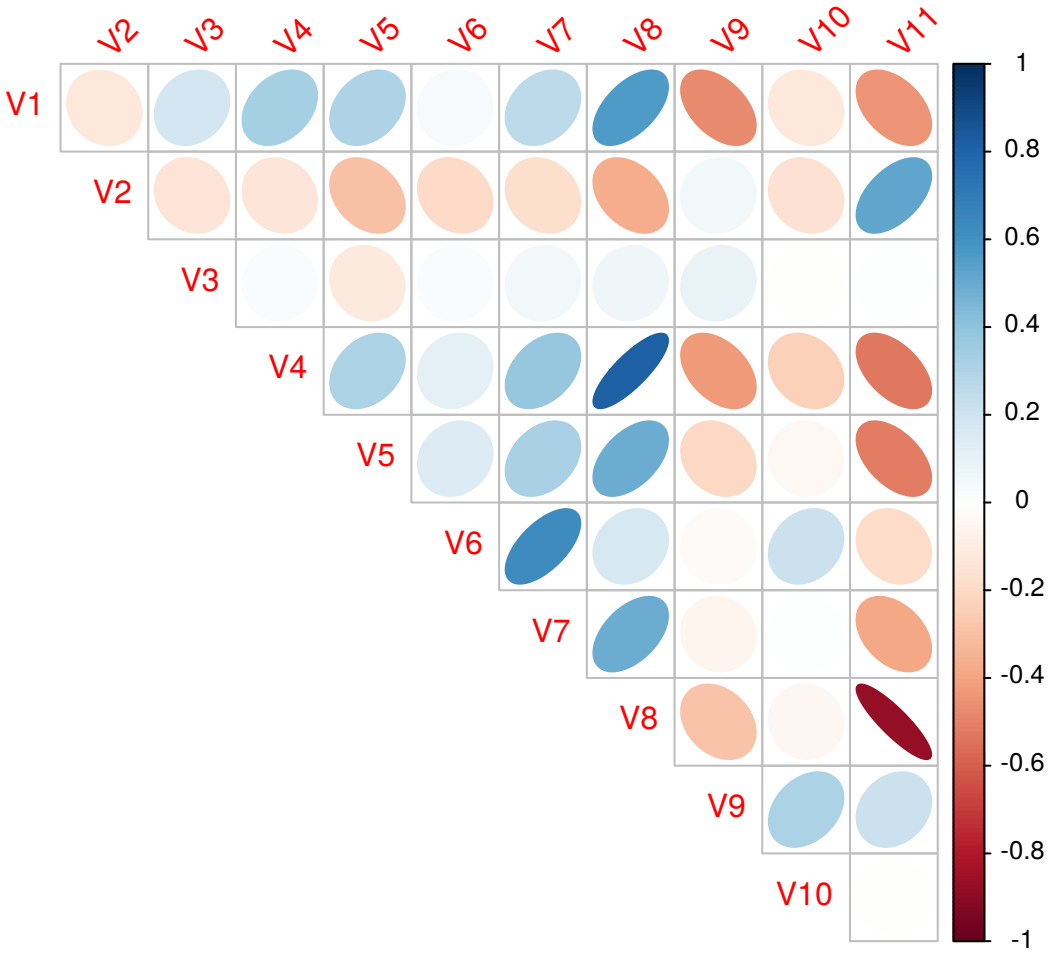


Figure 4.17: Correlation Matrix Plot (Class == 6)

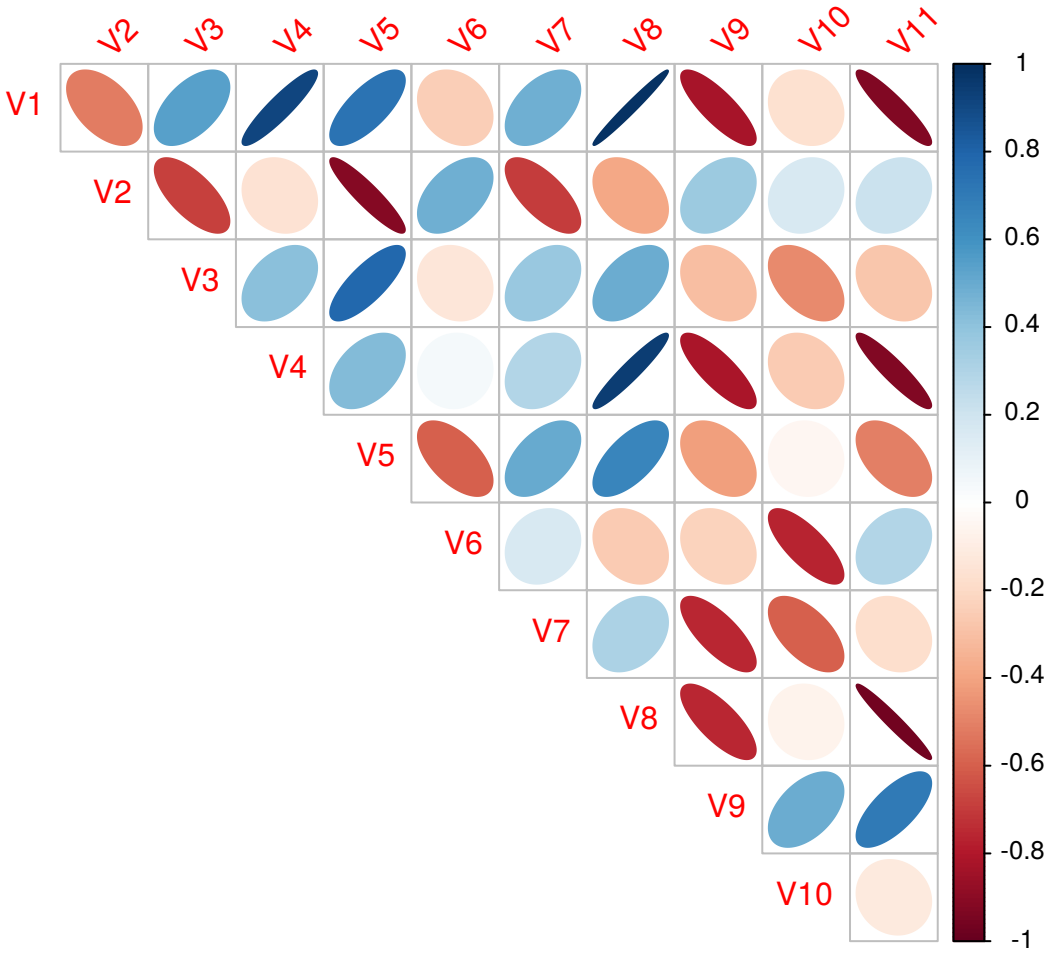


Figure 4.18: Correlation Matrix Plot (Class == 7)