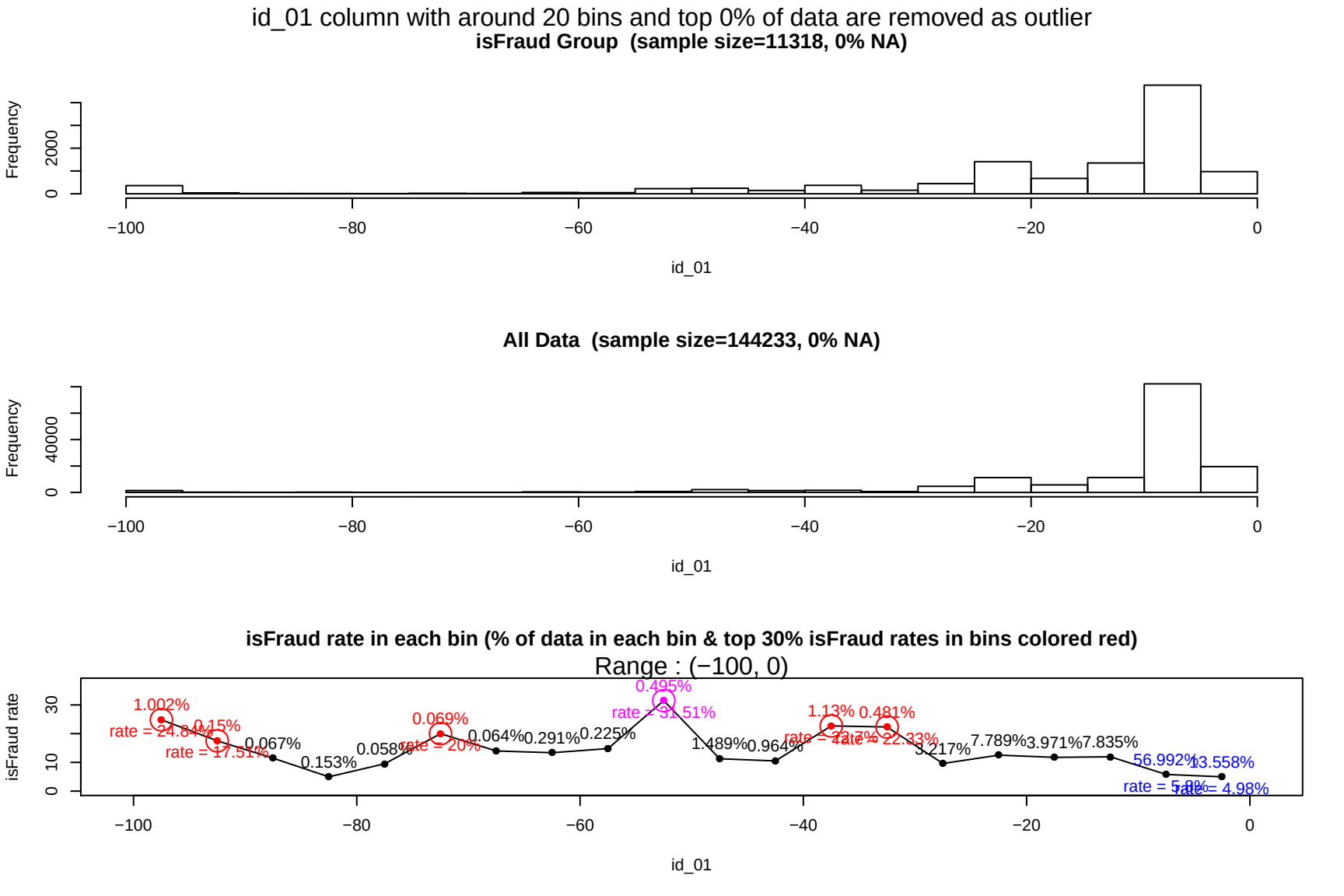
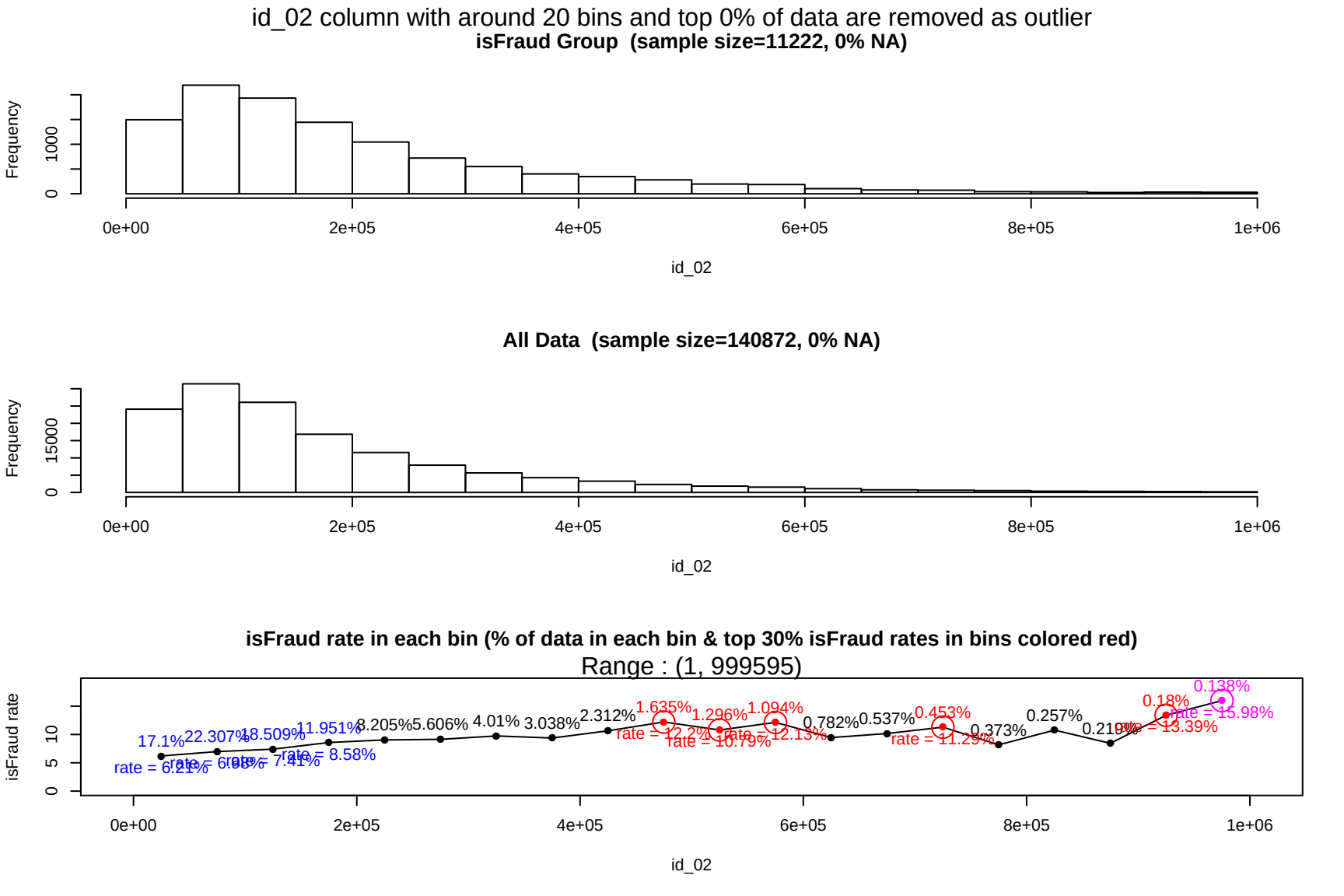
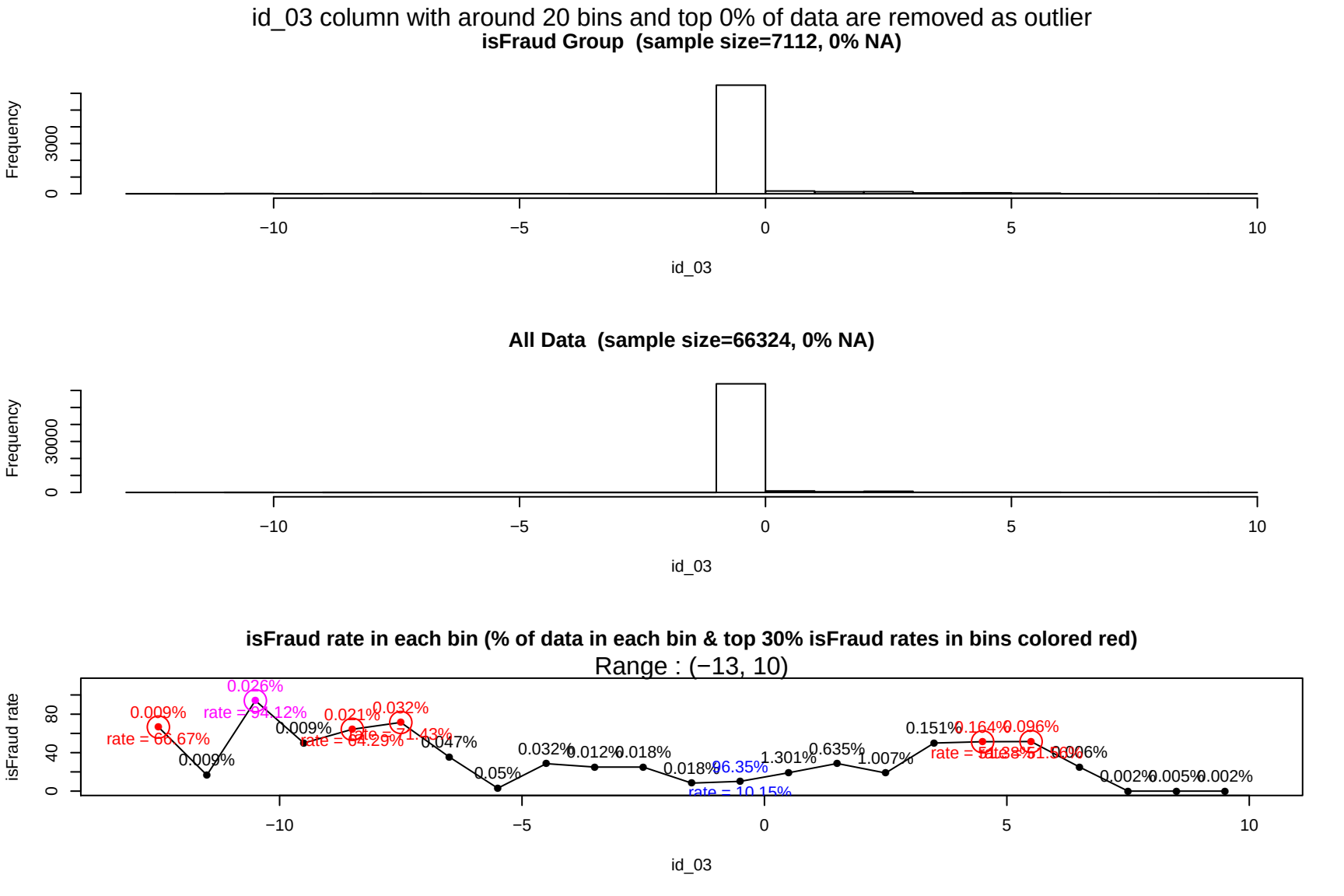
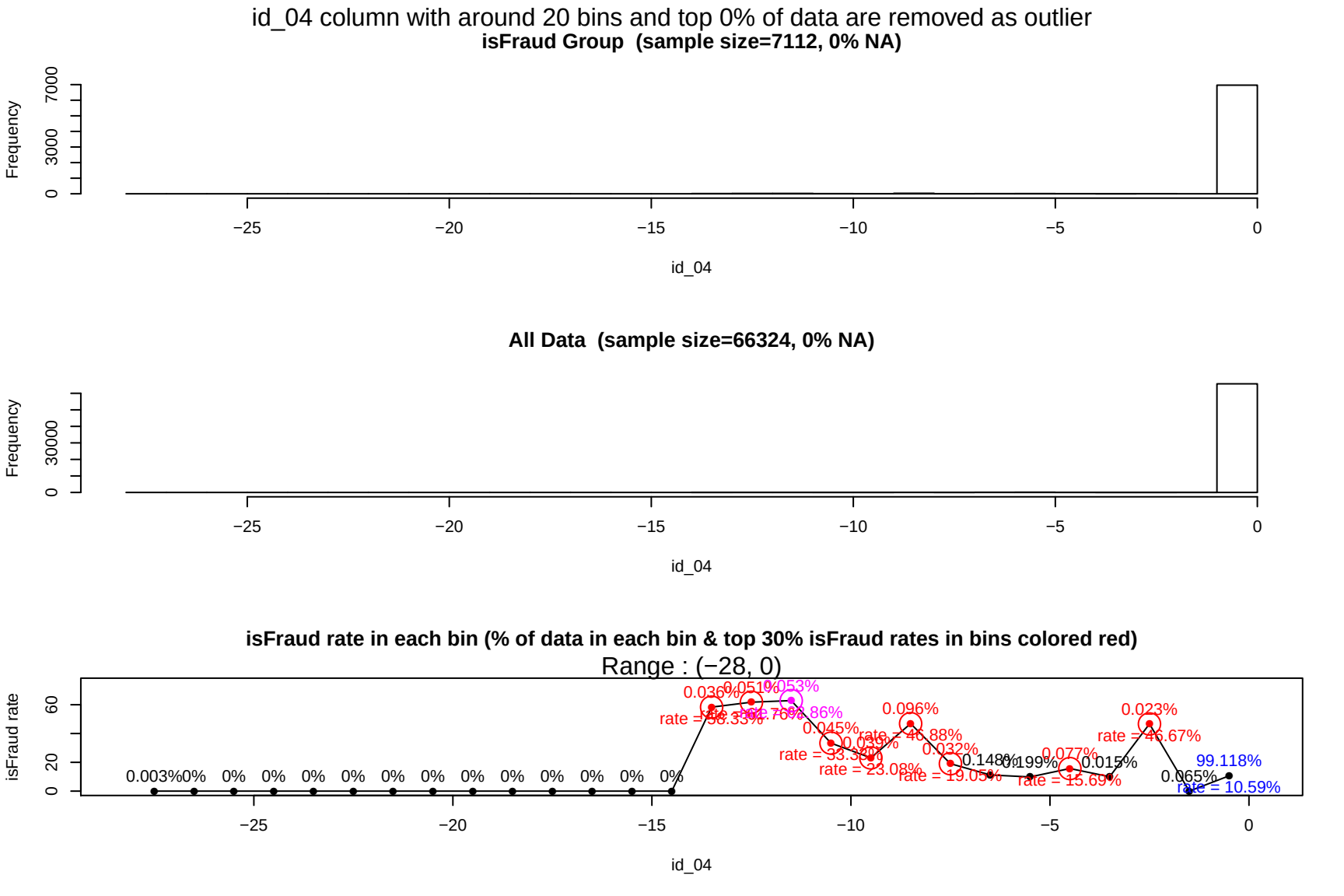
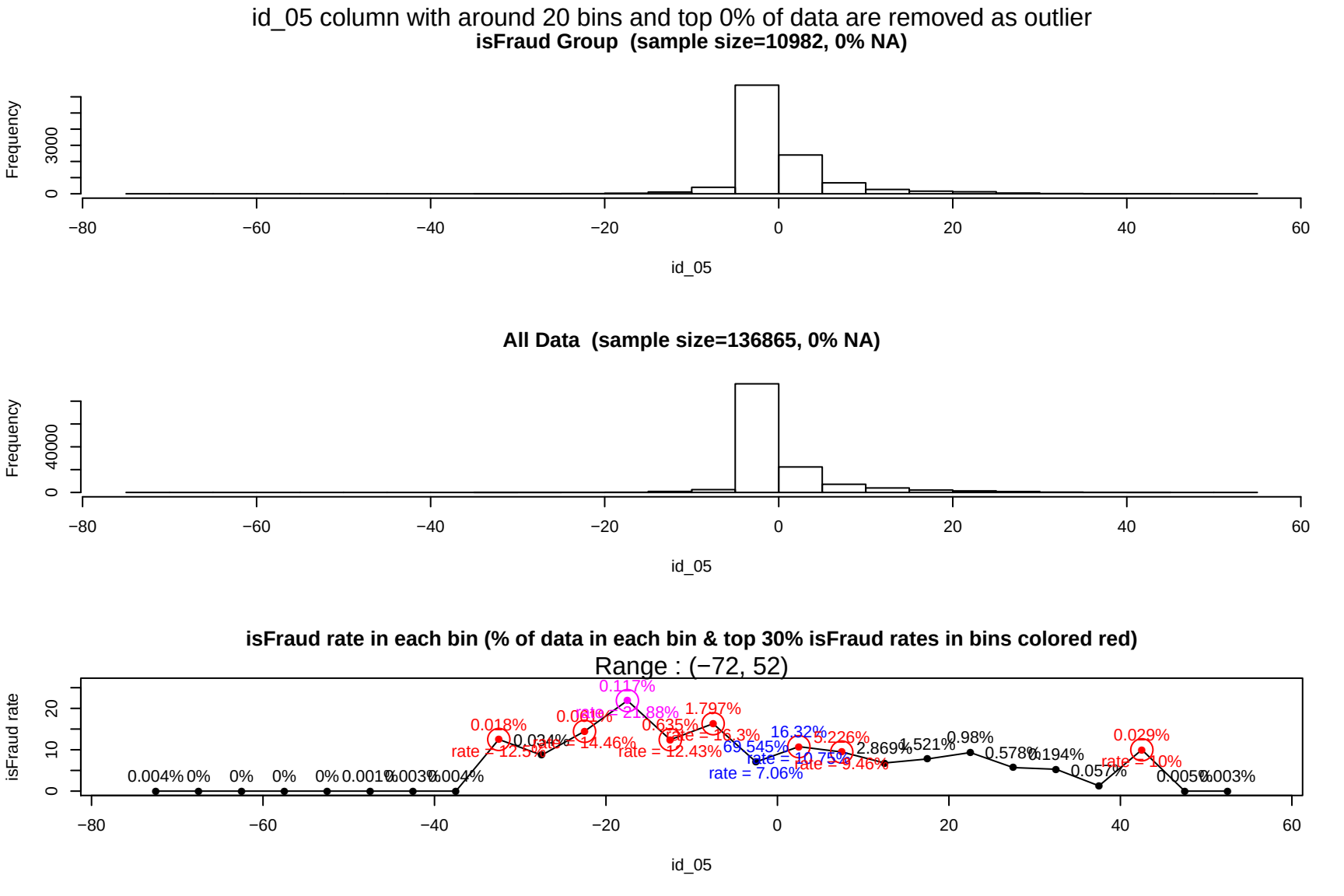
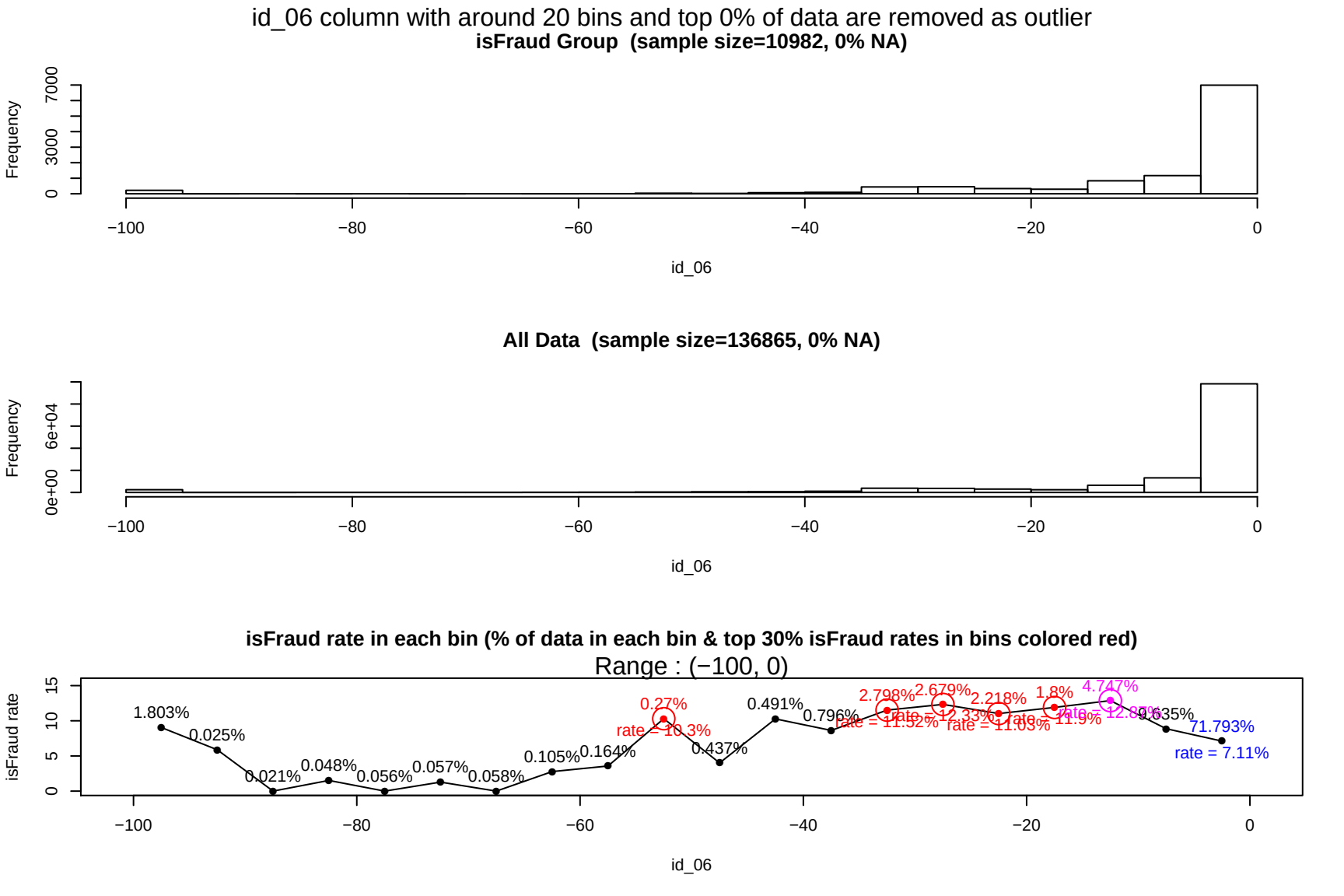

id_01 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=11318, 0% NA)
Frequency
0
2000
−100
−80
−60
−40
−20
0
id_01
All Data (sample size=144233, 0% NA)
Frequency
0
40000
−100
−80
−60
−40
−20
0
id_01
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (−100, 0)
isFraud rate
0
10
30
1.002%
rate = 24.84%
0.15%
rate = 17.51%
0.067%
0.153%
0.058%
0.069%
rate = 20%
0.064%
0.291%
0.225%
0.495%
rate = 31.51%
1.489%
0.964%
1.13%
0.481%
rate = 22.33%
3.217%
7.789%
3.971%
7.835%
56.992%
13.558%
rate = 4.98%
−100
−80
−60
−40
−20
0
id_01

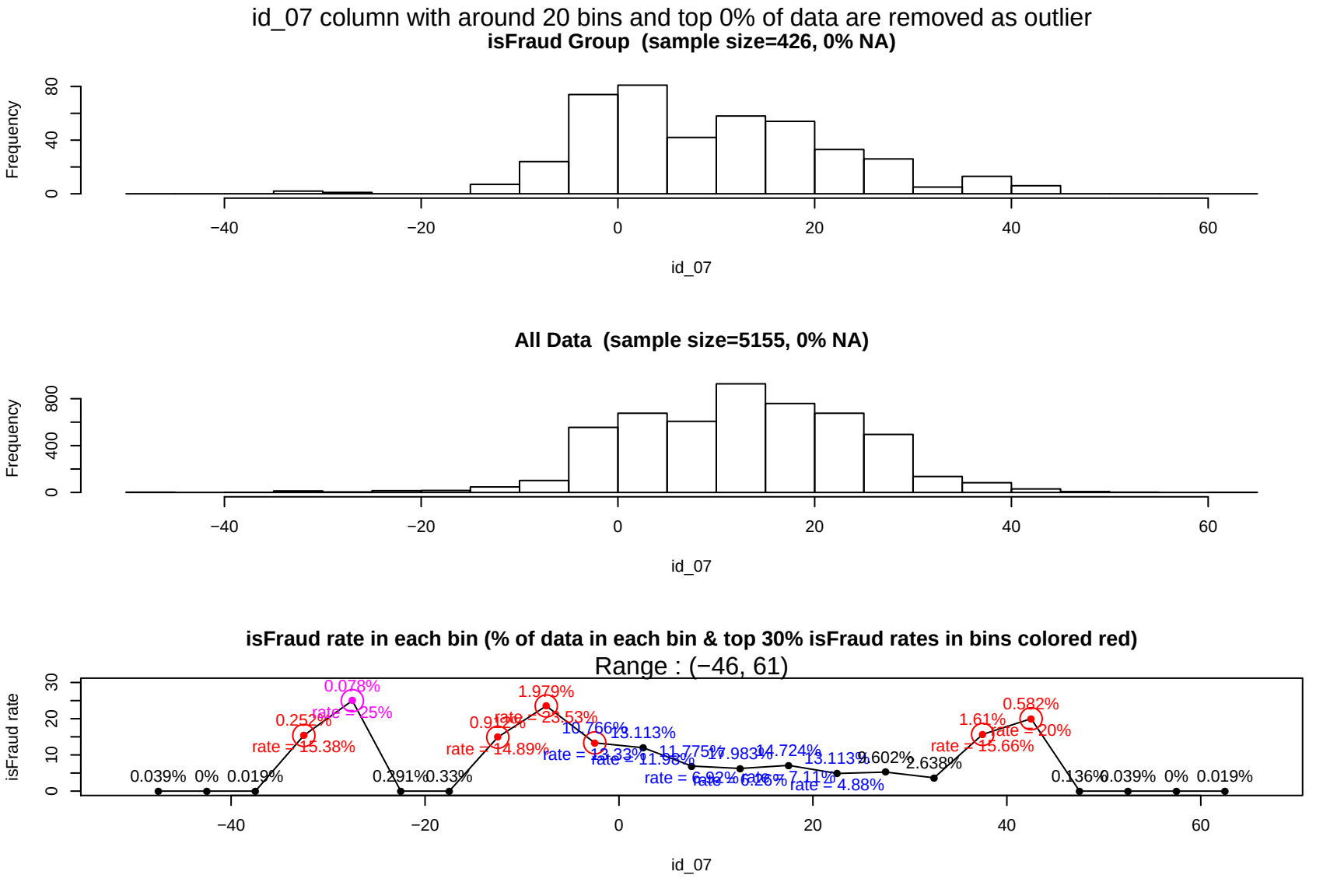

id_02 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=11222, 0% NA)
Frequency
0
1000
0e+00
2e+05
4e+05
6e+05
8e+05
1e+06
id_02
All Data (sample size=140872, 0% NA)
Frequency
0
15000
0e+00
2e+05
4e+05
6e+05
8e+05
1e+06
id_02
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (1, 999595)
isFraud rate
0
5
10
17.1%
22.307%
18.509%
11.951%
8.205%
5.606%
4.01%
3.038%
2.312%
1.635%
1.296%
1.094%
0.782%
0.537%
0.453%
0.373%
0.257%
0.21%
0.18%
0.138%
rate = 6.21%
rate = 8.58%
rate = 11.29%
rate = 13.39%
rate = 15.98%
0e+00
2e+05
4e+05
6e+05
8e+05
1e+06
id_02

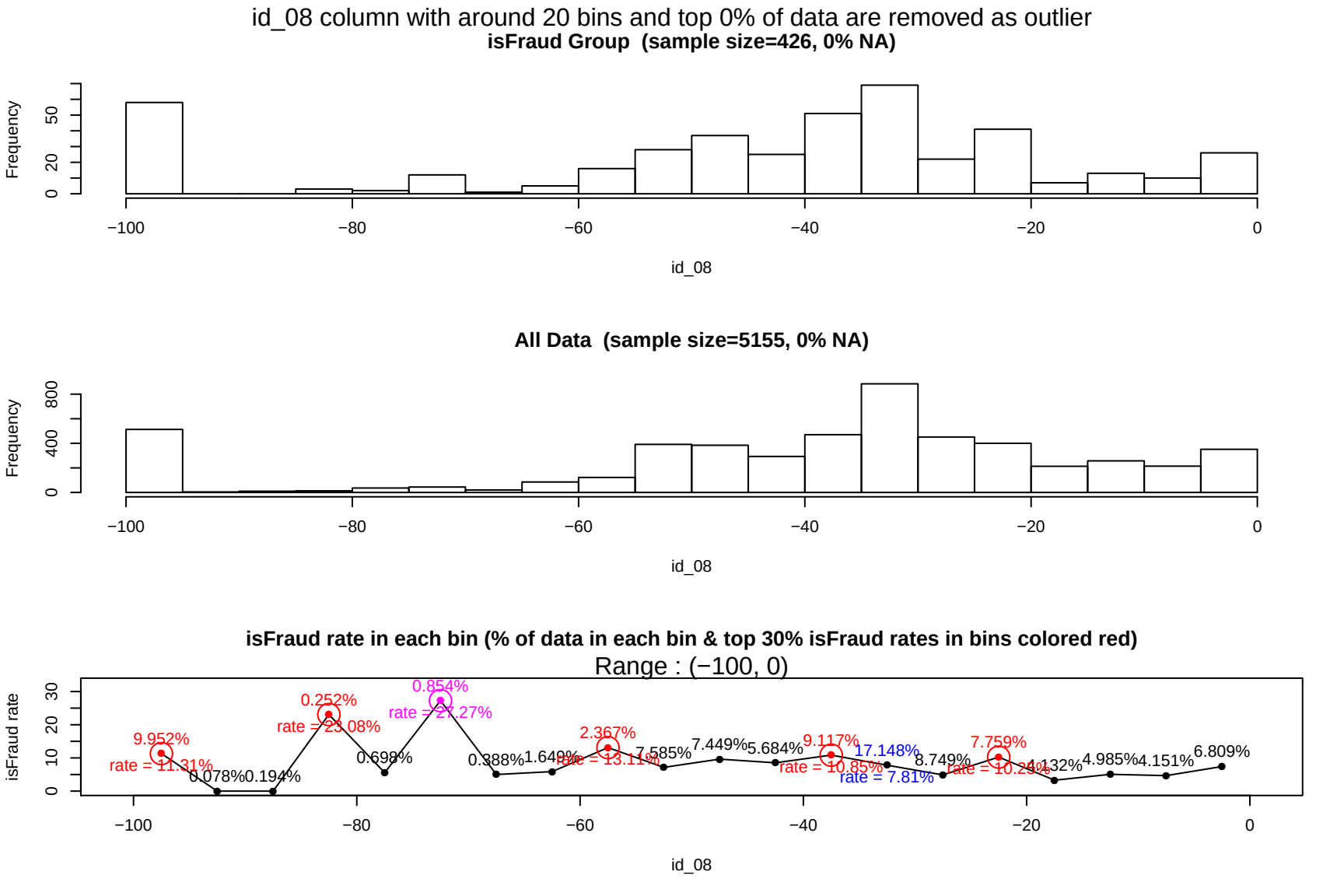

id_03 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=7112, 0% NA)
Frequency
0
3000
−10
−5
0
5
10
id_03
All Data (sample size=66324, 0% NA)
Frequency
0
30000
−10
−5
0
5
10
id_03
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (−13, 10)
isFraud rate
0
40
80
0.009%
rate = 66.67%
0.009%
0.026%
rate = 94.12%
0.009%
0.021%
rate = 64.29%
0.032%
rate = 71.43%
0.047%
0.05%
0.032%
0.012%
0.018%
0.018%
rate = 10.15%
96.35%
1.301%
0.635%
1.007%
0.151%
0.164%
0.096%
0.002%
0.005%
0.002%
−10
−5
0
5
10
id_03

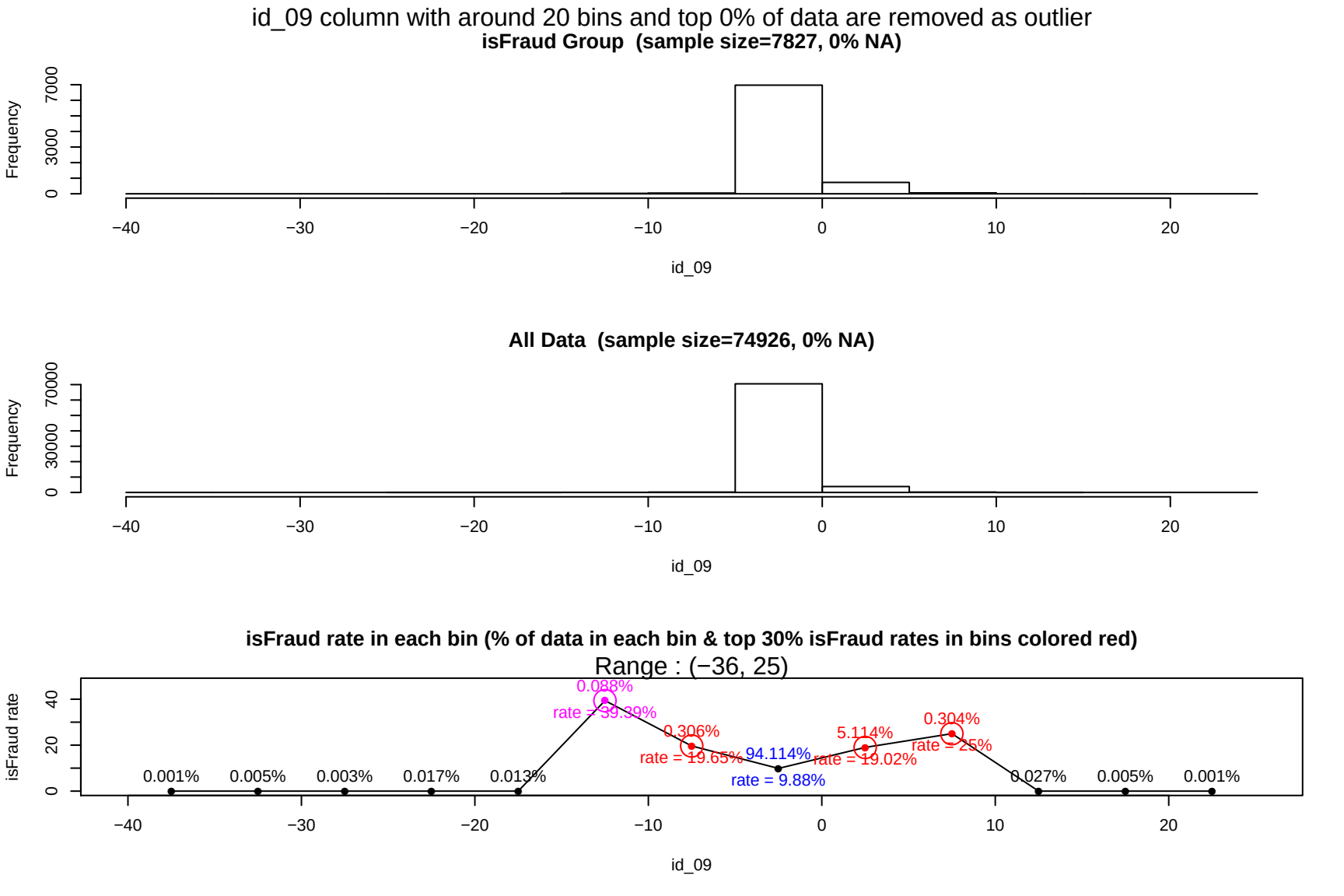

id_04 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=7112, 0% NA)
Frequency
0
3000
7000
−25
−20
−15
−10
−5
0
id_04
All Data (sample size=66324, 0% NA)
Frequency
0
30000
−25
−20
−15
−10
−5
0
id_04
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (−28, 0)
isFraud rate
0
20
60
0.003%
0%
0%
0%
0%
0%
0%
0%
0%
0%
0%
0%
0%
0%
0%
0.036%
0.051%
0.053%
rate = 56.33%
0.045%
rate = 33.33%
rate = 23.08%
0.096%
rate = 46.88%
0.032%
rate = 19.05%
0.148%
0.199%
0.077%
rate = 15.69%
0.015%
0.023%
rate = 46.67%
0.065%
99.118%
rate = 10.59%
−25
−20
−15
−10
−5
0
id_04

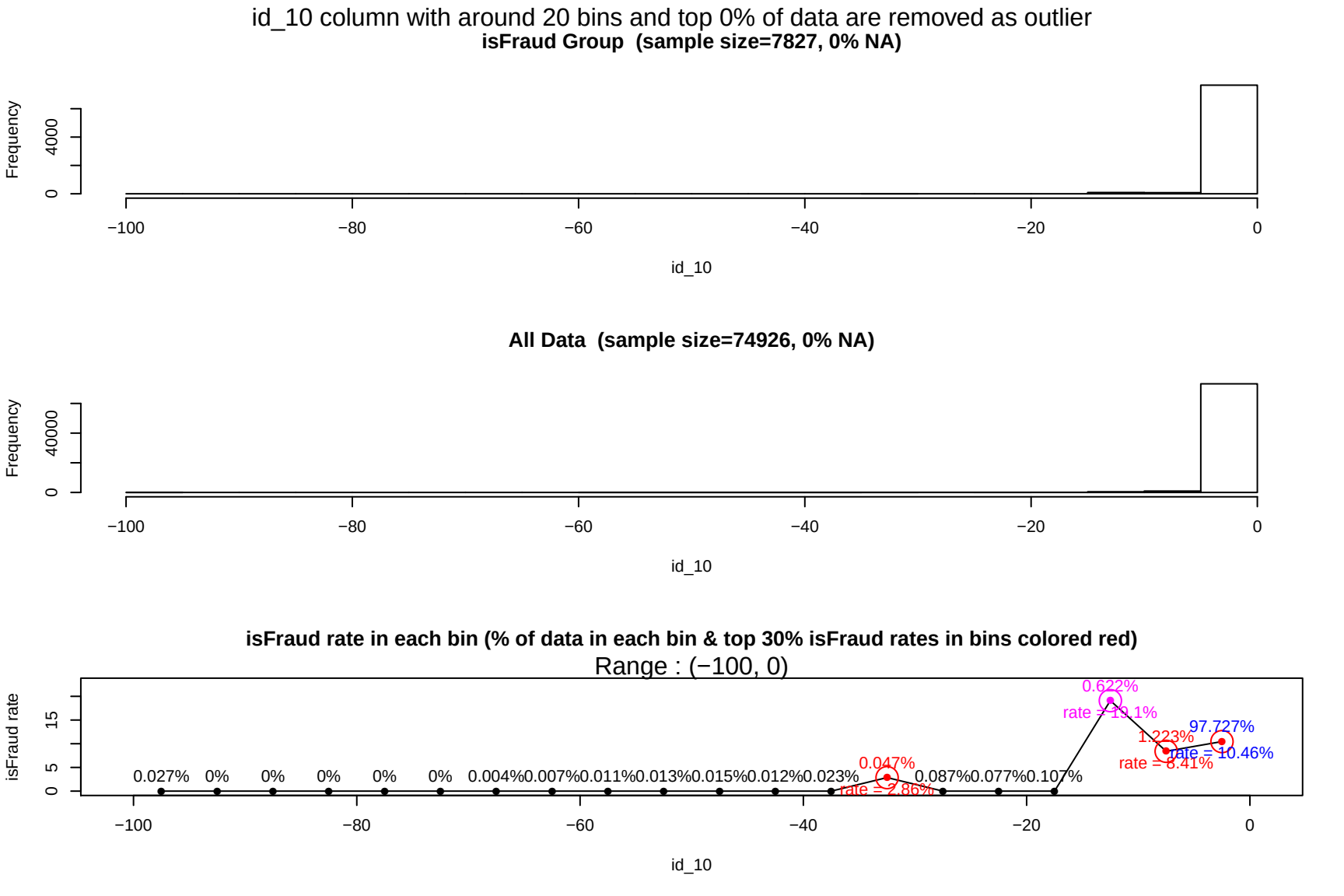

id_05 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=10982, 0% NA)
Frequency
0
3000
−80
−60
−40
−20
0
20
40
60
id_05
All Data (sample size=136865, 0% NA)
Frequency
0
40000
−80
−60
−40
−20
0
20
40
60
id_05
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (−72, 52)
isFraud rate
0
10
20
0.004%
0%
0%
0%
0%
0.001%
0.003%
0.004%
0.018%
rate = 12.5%
0.034%
0.061%
rate = 14.46%
0.117%
rate = 21.88%
rate = 12.43%
0.635%
1.797%
rate = 16.3%
69.545%
rate = 10.75%
rate = 7.06%
16.32%
5.226%
rate = 9.46%
2.869%
4.521%
0.98%
0.578%
0.194%
0.057%
0.029%
rate = 10%
0.005%
0.003%
−80
−60
−40
−20
0
20
40
60
id_05

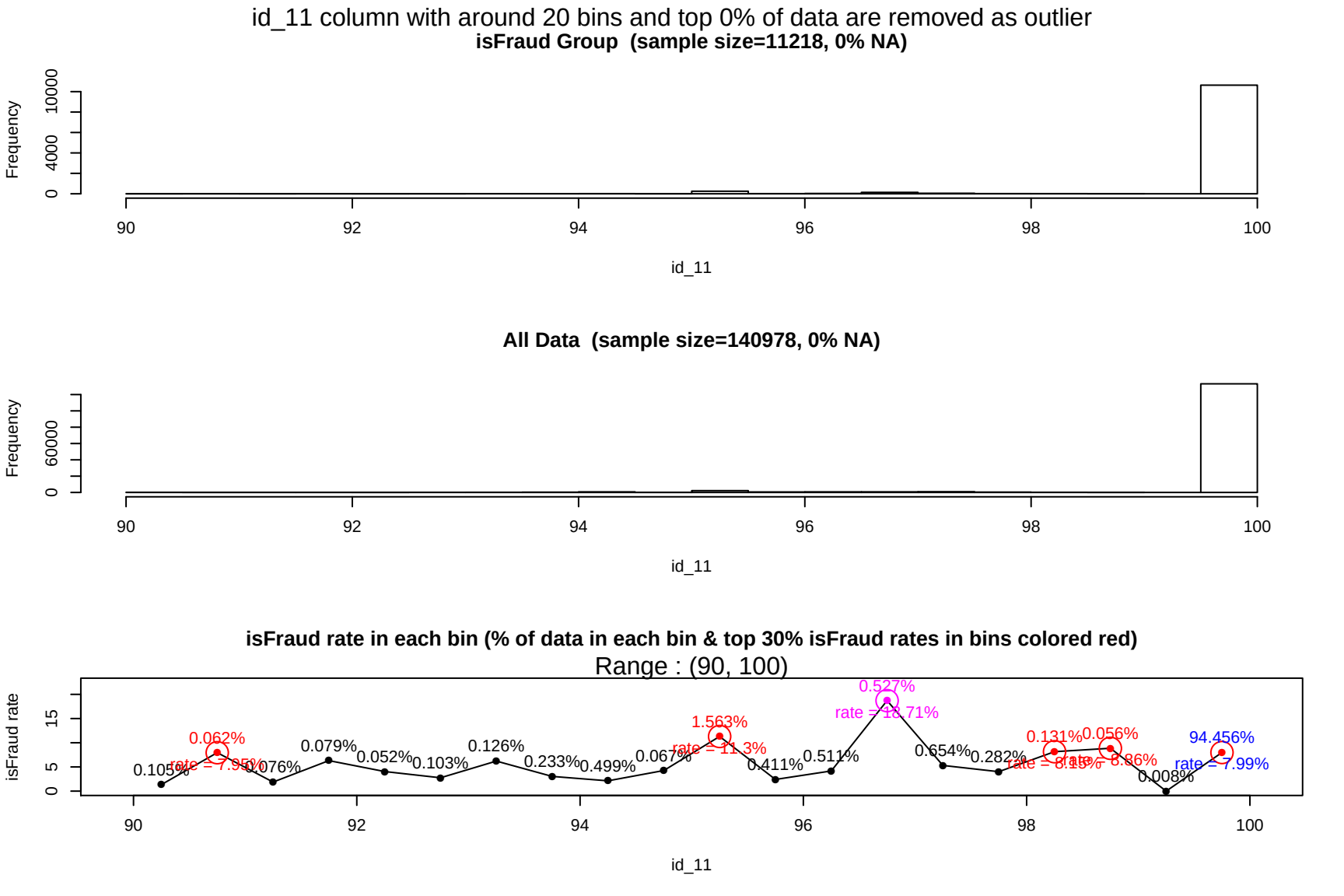

id_06 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=10982, 0% NA)
Frequency
0
3000
7000
−100
−80
−60
−40
−20
0
id_06
All Data (sample size=136865, 0% NA)
Frequency
0e+00
6e+04
−100
−80
−60
−40
−20
0
id_06
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (−100, 0)
isFraud rate
0
5
10
15
1.803%
0.025%
0.021%
0.048%
0.056%
0.057%
0.058%
0.105%
0.164%
0.27%
rate = 10.3%
0.437%
0.491%
0.796%
2.798%
2.679%
2.218%
1.8%
4.747%
rate = 11.52%
rate = 12.33%
rate = 11.03%
rate = 11.9%
rate = 12.89%
3.635%
71.793%
rate = 7.11%
−100
−80
−60
−40
−20
0
id_06

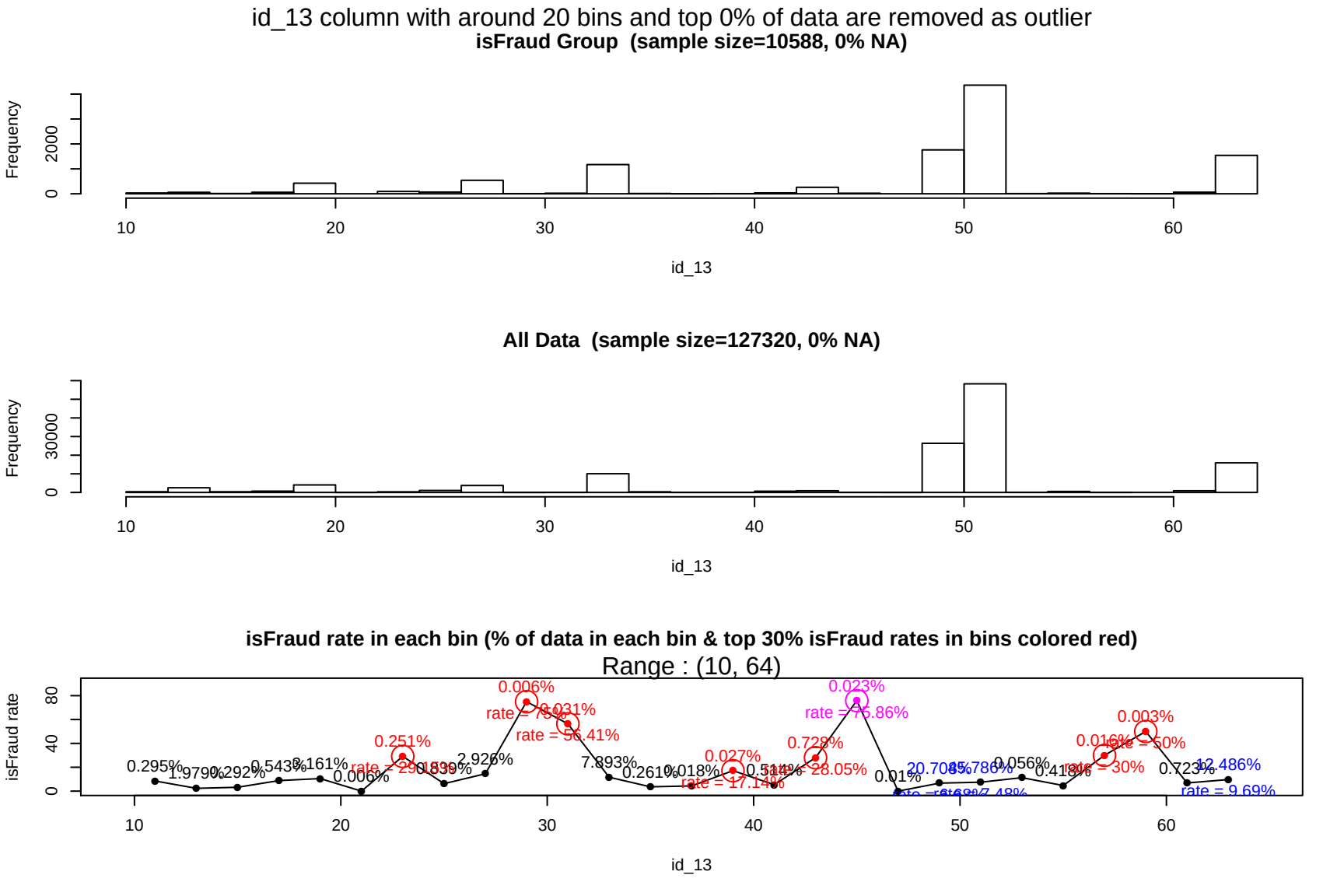

id_07 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=426, 0% NA)
Frequency
0
40
80
−40
−20
0
20
40
60
id_07
All Data (sample size=5155, 0% NA)
Frequency
0
400
800
−40
−20
0
20
40
60
id_07
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (−46, 61)
isFraud rate
0
10
20
30
0.078%
rate = 25%
0.252%
rate = 15.38%
1.979%
rate = 23.53%
0.913%
rate = 14.89%
10.766%
13.113%
0.039%
0%
0.019%
0.291%
0.33%
11.775%
17.983%
14.724%
13.113%
9.602%
2.638%
6.26%
7.11%
4.88%
1.61%
rate = 20%
0.582%
rate = 15.66%
0.136%
0.039%
0%
0.019%
−40
−20
0
20
40
60
id_07

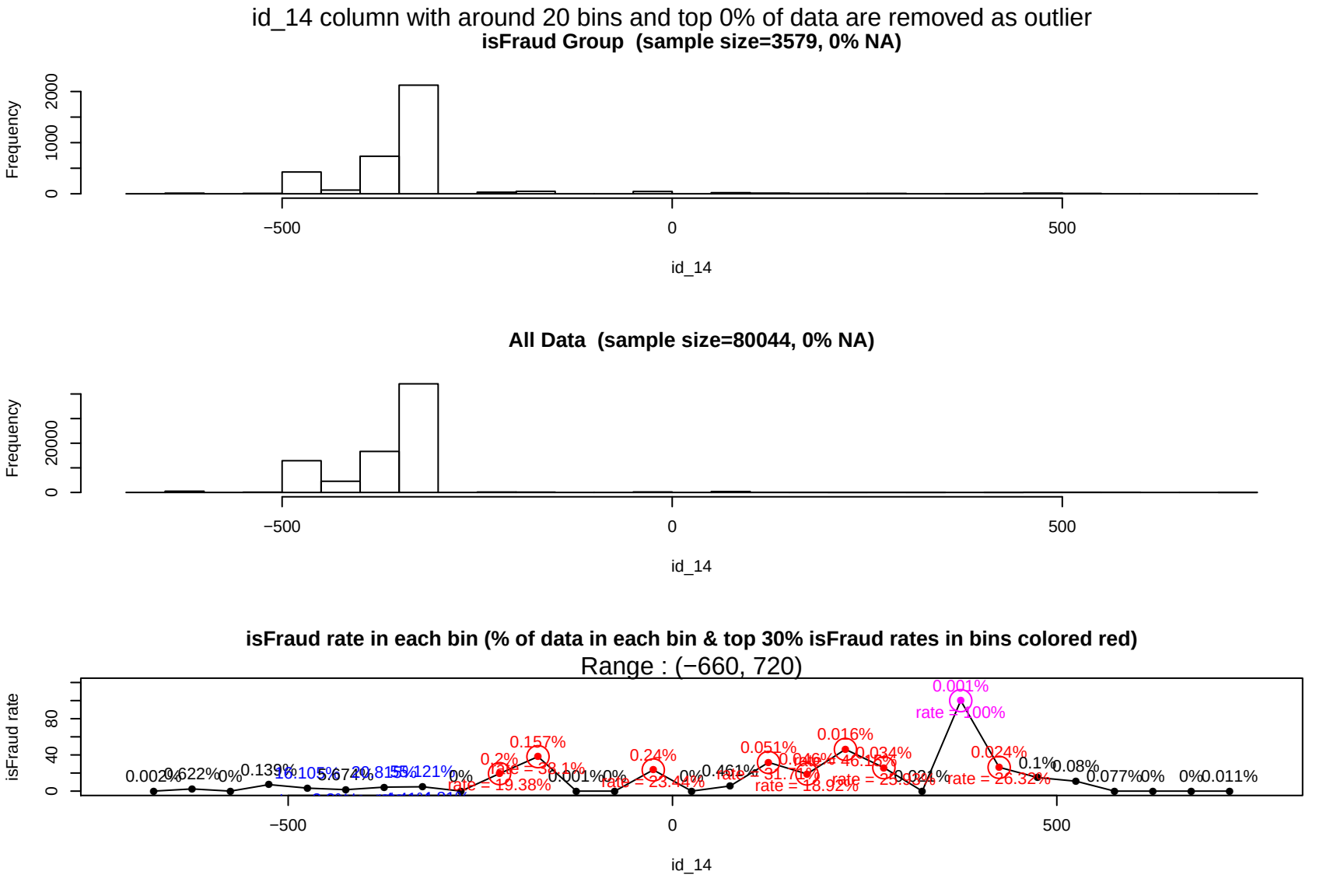

id_08 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=426, 0% NA)
Frequency
0
20
50
−100
−80
−60
−40
−20
0
id_08
All Data (sample size=5155, 0% NA)
Frequency
0
400
800
−100
−80
−60
−40
−20
0
id_08
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (−100, 0)
isFraud rate
0
10
20
30
9.952%
rate = 11.31%
0.078%
0.194%
0.252%
rate = 23.08%
0.698%
0.854%
rate = 27.27%
0.388%
1.649%
2.367%
rate = 13.11%
7.585%
7.449%
5.684%
9.117%
rate = 10.85%
17.148%
rate = 7.81%
8.749%
7.759%
rate = 10.28%
4.132%
4.985%
4.151%
6.809%
−100
−80
−60
−40
−20
0
id_08

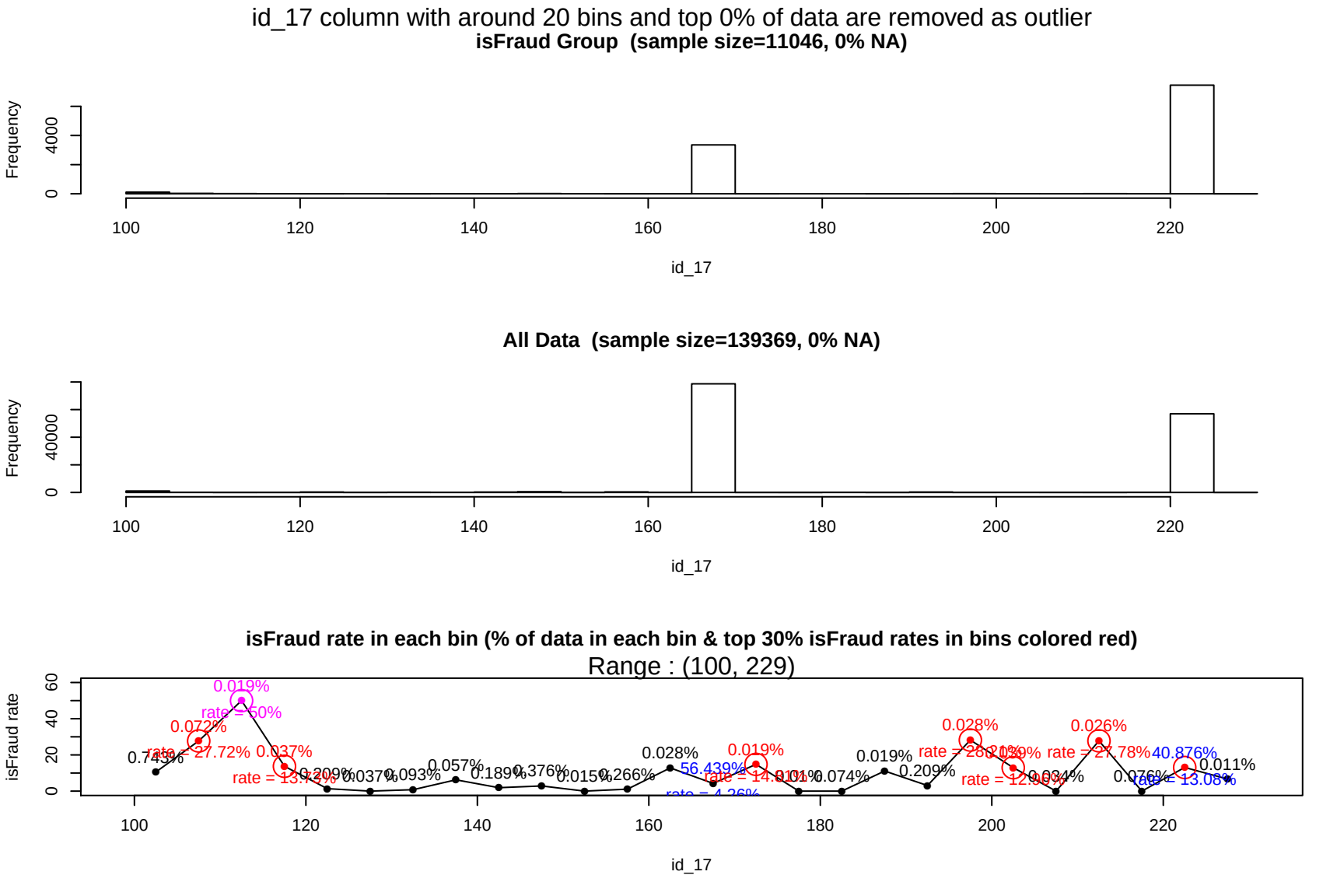

id_09 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=7827, 0% NA)
Frequency
0
3000
7000
−40
−30
−20
−10
0
10
20
id_09
All Data (sample size=74926, 0% NA)
Frequency
0
30000
70000
−40
−30
−20
−10
0
10
20
id_09
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (−36, 25)
isFraud rate
0
20
40
0.001%
0.005%
0.003%
0.017%
0.013%
0.088%
rate = 39.39%
0.306%
rate = 19.65%
94.114%
rate = 9.88%
5.114%
rate = 19.02%
0.304%
rate = 25%
0.027%
0.005%
0.001%
−40
−30
−20
−10
0
10
20
id_09

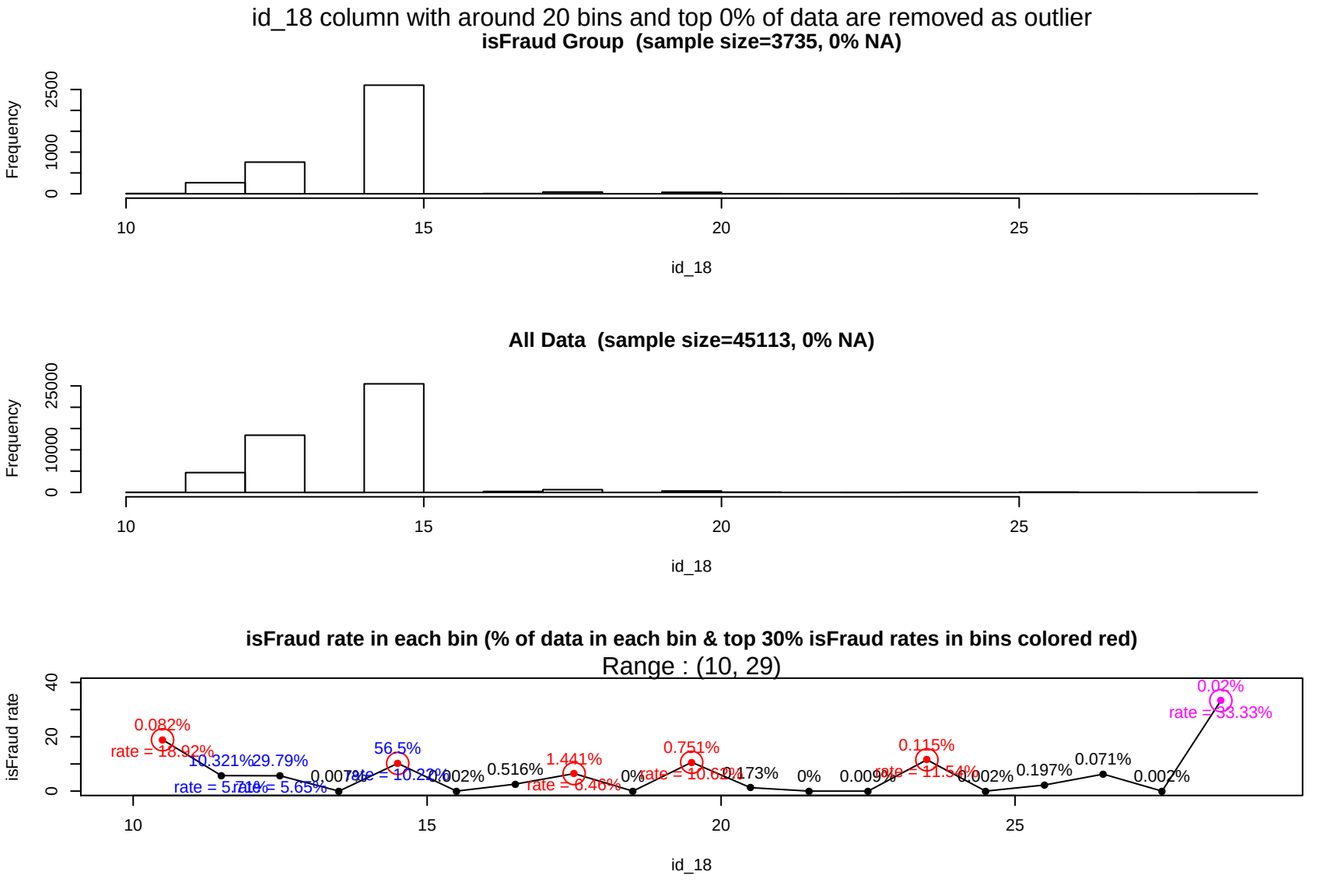

id_10 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=7827, 0% NA)
Frequency
0
4000
−100
−80
−60
−40
−20
0
id_10
All Data (sample size=74926, 0% NA)
Frequency
0
40000
−100
−80
−60
−40
−20
0
id_10
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (−100, 0)
isFraud rate
0
5
15
0.027%
0%
0%
0%
0%
0%
0.004%
0.007%
0.011%
0.013%
0.015%
0.012%
0.023%
0.047%
rate = 2.86%
0.087%
0.077%
0.107%
0.622%
rate = 19.1%
1.223%
rate = 8.41%
97.727%
rate = 10.46%
−100
−80
−60
−40
−20
0
id_10

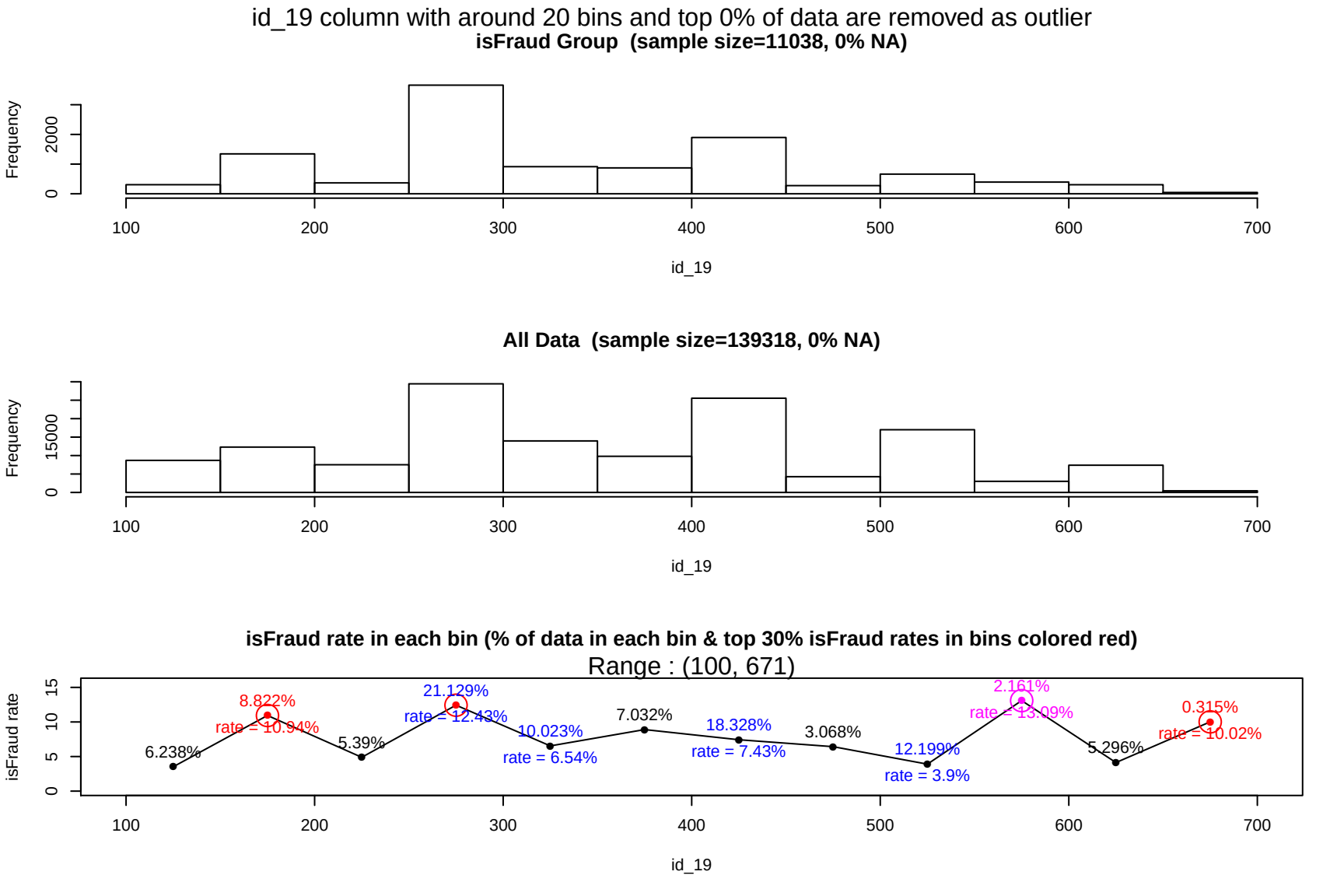

id_11 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=11218, 0% NA)
Frequency
0
4000
10000
90
92
94
96
98
100
id_11
All Data (sample size=140978, 0% NA)
Frequency
0
60000
90
92
94
96
98
100
id_11
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (90, 100)
isFraud rate
0
5
15
0.105%
0.062%
rate = 7.95%
0.076%
0.079%
0.052%
0.103%
0.126%
0.233%
0.499%
0.067%
1.563%
rate = 11.3%
0.411%
0.511%
0.527%
rate = 18.71%
0.654%
0.282%
0.131%
rate = 8.15%
0.056%
rate = 8.86%
0.008%
94.456%
rate = 7.99%
90
92
94
96
98
100
id_11

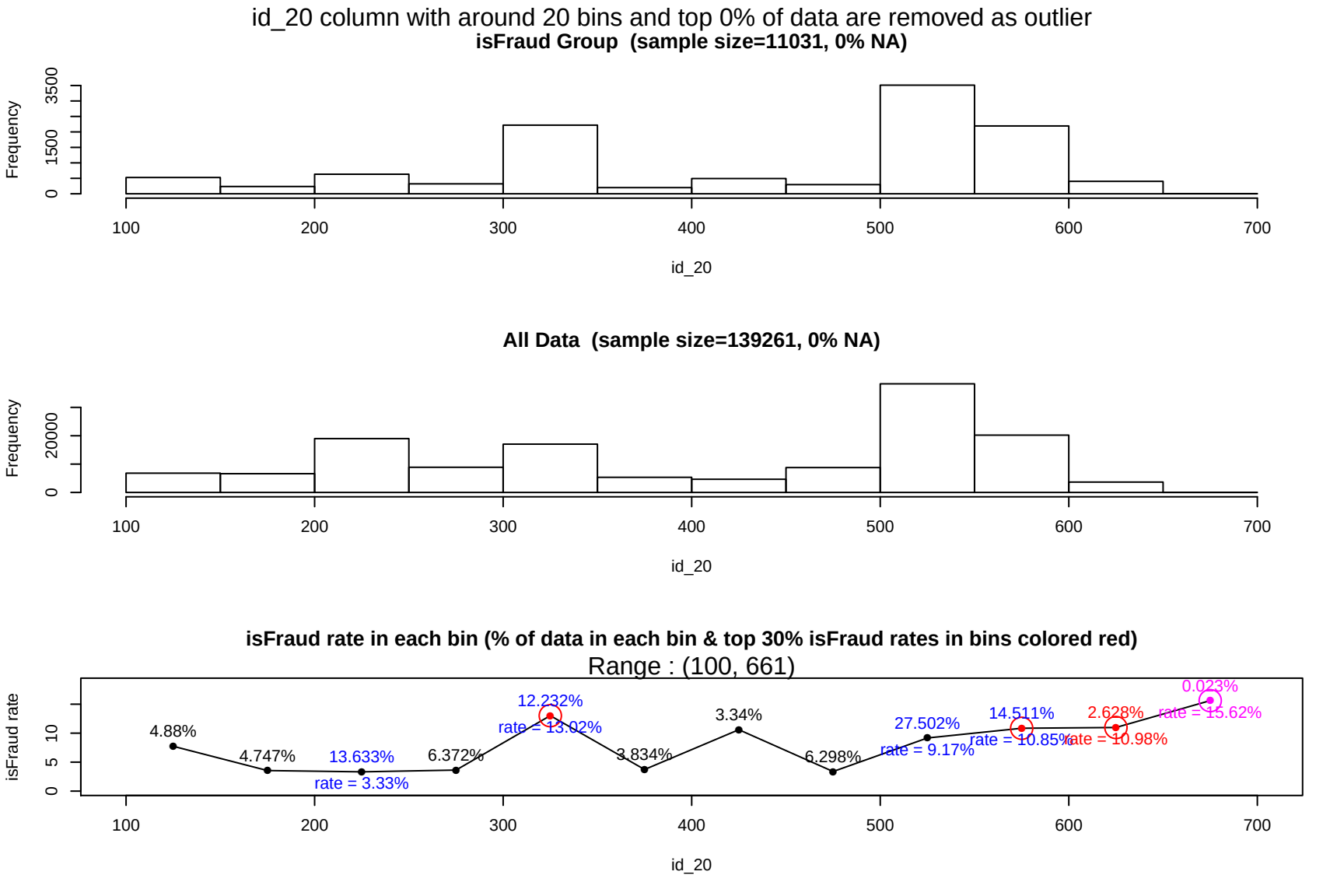

id_13 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=10588, 0% NA)
Frequency
0
2000
10
20
30
40
50
60
id_13
All Data (sample size=127320, 0% NA)
Frequency
0
30000
10
20
30
40
50
60
id_13
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (10, 64)
isFraud rate
0
40
80
0.295%
1.979%
0.292%
0.543%
0.161%
0.006%
0.251%
2.926%
0.006%
rate = 78.31%
rate = 56.41%
7.893%
0.261%
0.018%
0.027%
rate = 17.14%
0.729%
rate = 28.05%
0.023%
rate = 75.86%
0.01%
0.056%
0.418%
0.016%
0.003%
0.723%
12.486%
rate = 9.69%
10
20
30
40
50
60
id_13

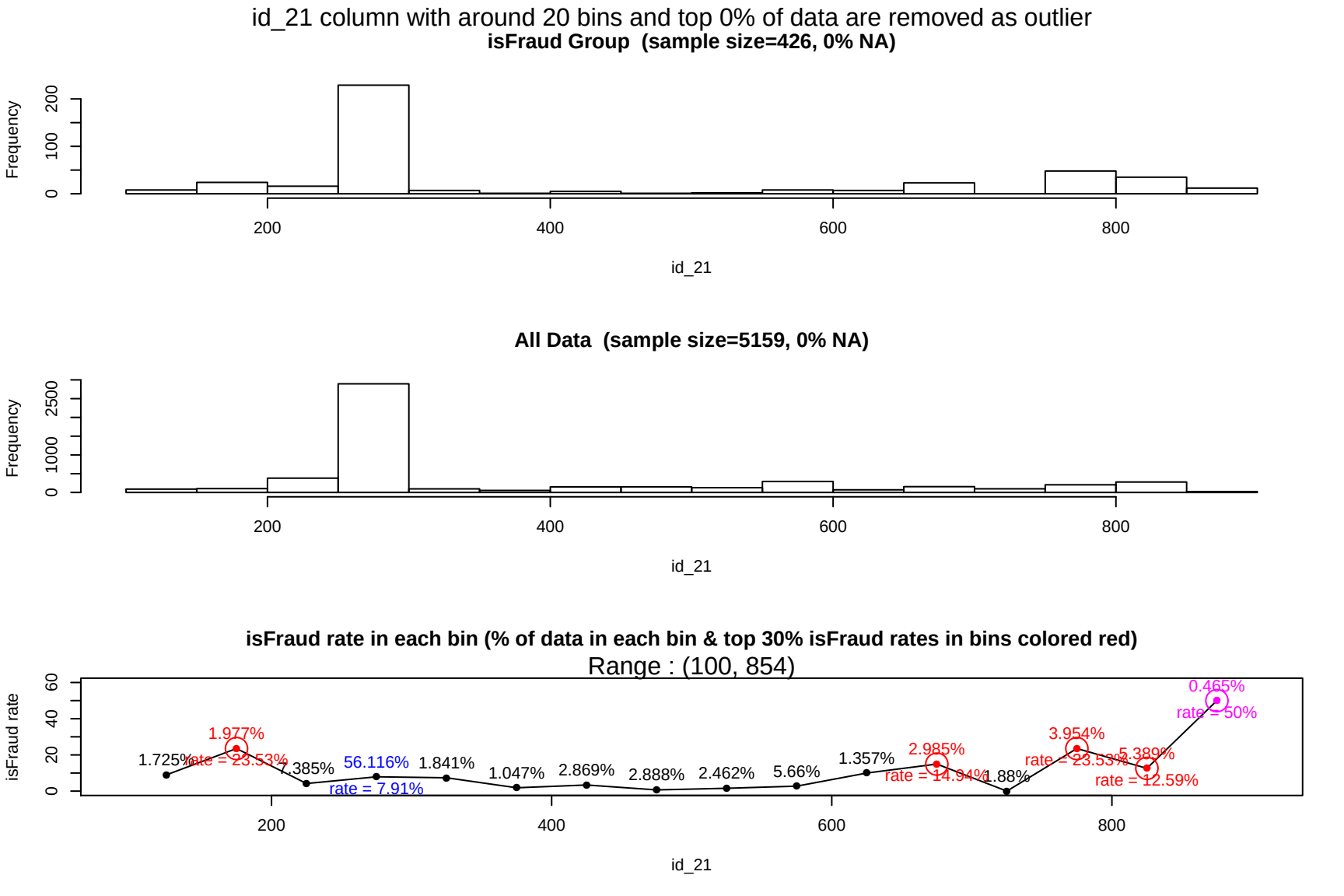

id_14 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=3579, 0% NA)
Frequency
0
1000
2000
−500
0
500
id_14
All Data (sample size=80044, 0% NA)
Frequency
0
20000
−500
0
500
id_14
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (−660, 720)
isFraud rate
0
40
80
0.002%
0.622%
0%
0.139%
0.105%
5.674%
0%
0.157%
0.2%
rate = 38.1%
rate = 19.38%
0.001%
0%
0.24%
rate = 23.44%
0%
0.461%
0.051%
rate = 31.76%
rate = 18.92%
0.016%
0.034%
rate = 46.15%
rate = 29.03%
0.001%
rate = 100%
0.024%
rate = 26.32%
0.1%
0.08%
0.077%
0%
0%
0.011%
−500
0
500
id_14

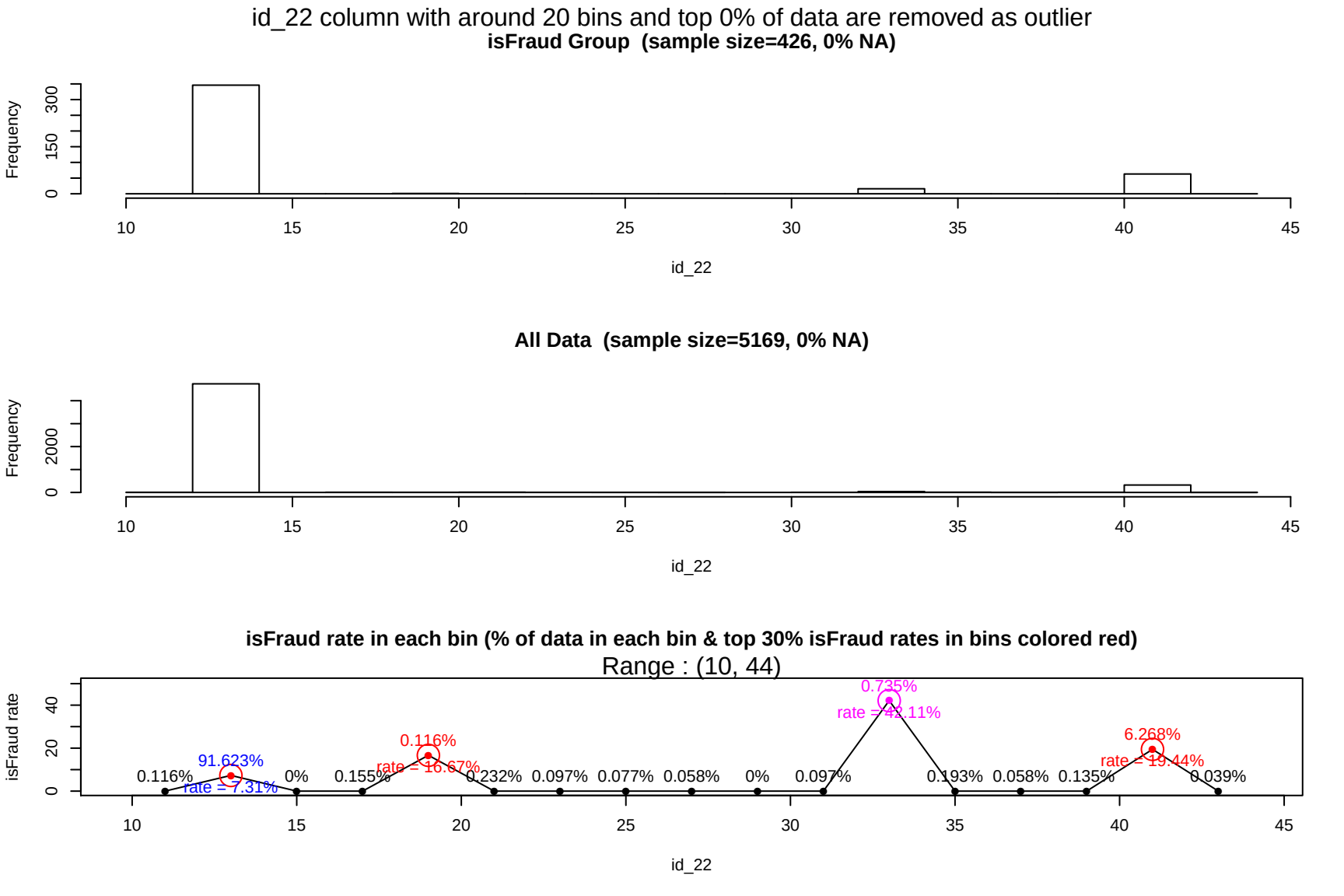

id_17 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=11046, 0% NA)
Frequency
0
4000
100
120
140
160
180
200
220
id_17
All Data (sample size=139369, 0% NA)
Frequency
0
40000
100
120
140
160
180
200
220
id_17
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (100, 229)
isFraud rate
0
20
40
60
0.743%
0.072%
rate = 27.72%
0.019%
rate = 50%
0.037%
0.209%
0.037%
0.093%
0.057%
0.189%
0.376%
0.015%
0.266%
0.028%
56.439%
0.019%
0.074%
0.019%
0.209%
0.028%
0.026%
rate = 27.78%
40.876%
0.011%
0.076%
100
120
140
160
180
200
220
id_17

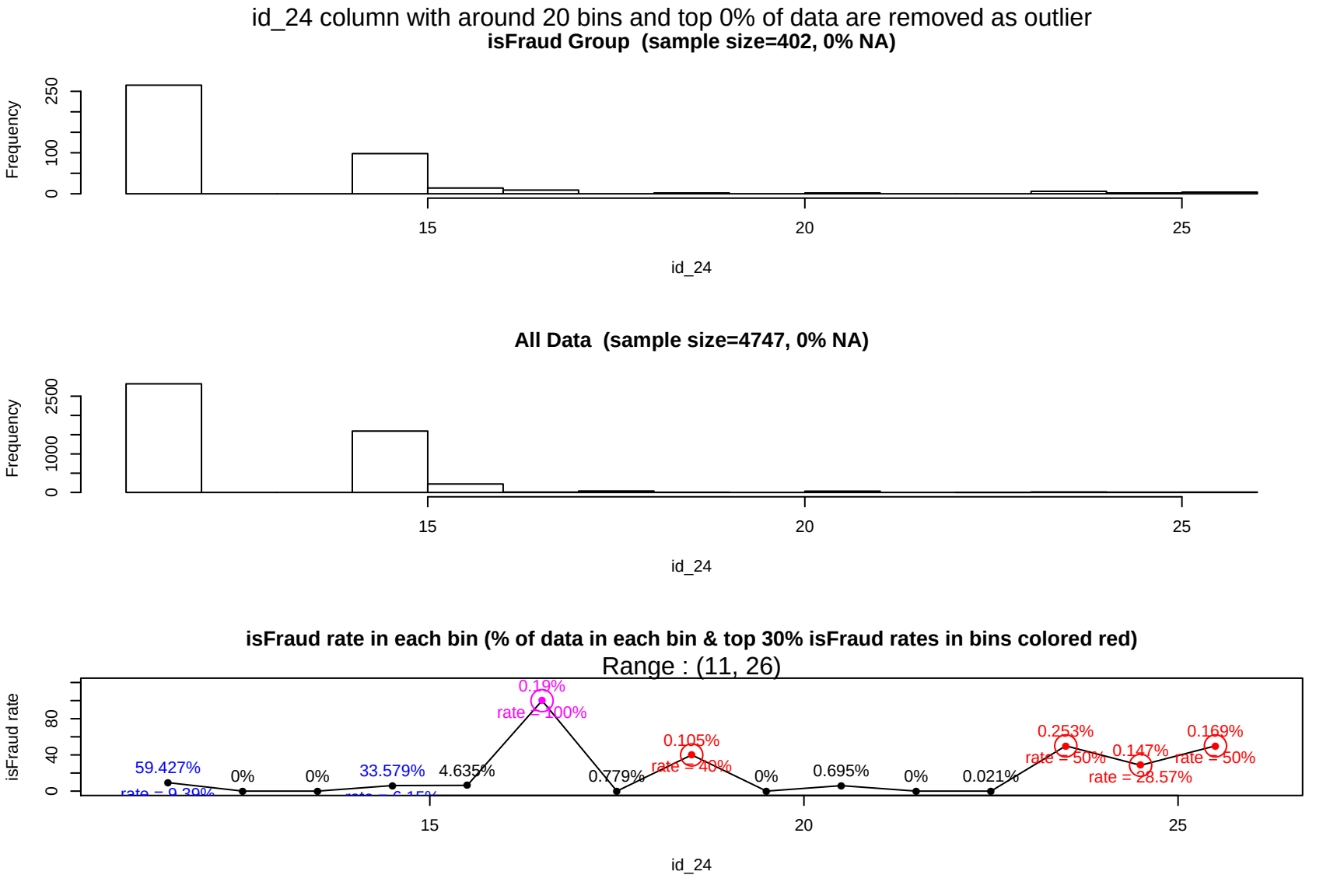

id_18 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=3735, 0% NA)
Frequency
0
1000
2500
10
15
20
25
id_18
All Data (sample size=45113, 0% NA)
Frequency
0
10000
25000
10
15
20
25
id_18
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (10, 29)
isFraud rate
0
20
40
0.082%
rate = 18.92%
10.321%
29.79%
rate = 5.65%
0.007%
56.5%
0.002%
0.516%
1.441%
rate = 6.46%
0%
0.751%
0.173%
0%
0.009%
0.115%
0.002%
0.197%
0.071%
0.002%
0.02%
rate = 33.33%
10
15
20
25
id_18

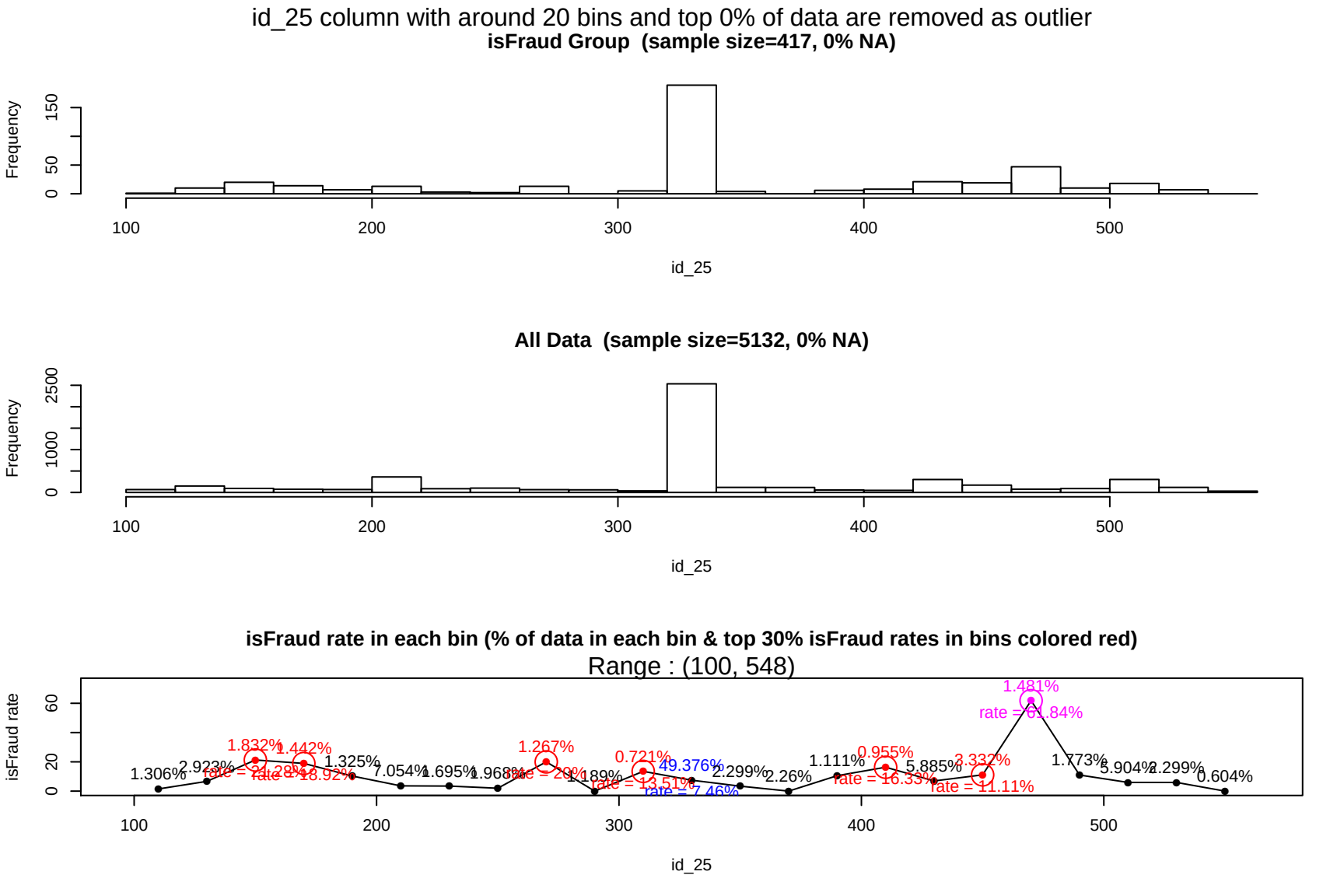

id_19 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=11038, 0% NA)
Frequency
0
2000
100
200
300
400
500
600
700
id_19
All Data (sample size=139318, 0% NA)
Frequency
0
15000
100
200
300
400
500
600
700
id_19
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (100, 671)
isFraud rate
0
5
10
15
6.238%
8.822%
rate = 10.94%
5.39%
21.129%
rate = 12.43%
10.023%
rate = 6.54%
7.032%
18.328%
rate = 7.43%
3.068%
12.199%
rate = 3.9%
2.161%
rate = 13.09%
5.296%
0.315%
rate = 10.02%
100
200
300
400
500
600
700
id_19

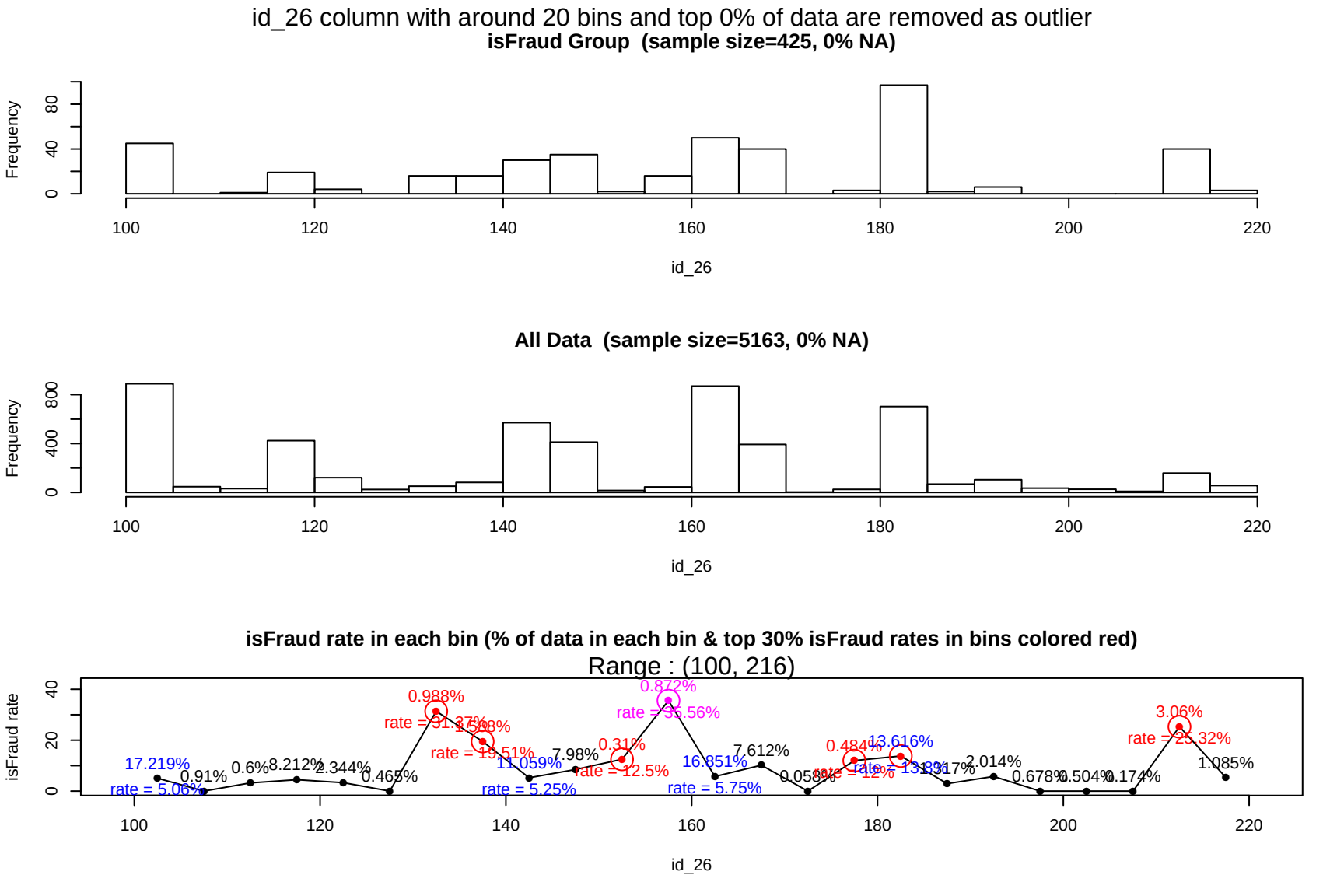

id_20 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=11031, 0% NA)
Frequency
0
1500
3500
100
200
300
400
500
600
700
id_20
All Data (sample size=139261, 0% NA)
Frequency
0
20000
100
200
300
400
500
600
700
id_20
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (100, 661)
isFraud rate
0
5
10
4.88%
4.747%
13.633%
rate = 3.33%
6.372%
12.232%
rate = 13.02%
3.834%
3.34%
6.298%
27.502%
rate = 9.17%
14.511%
rate = 10.85%
2.628%
rate = 10.98%
0.023%
rate = 15.62%
100
200
300
400
500
600
700
id_20

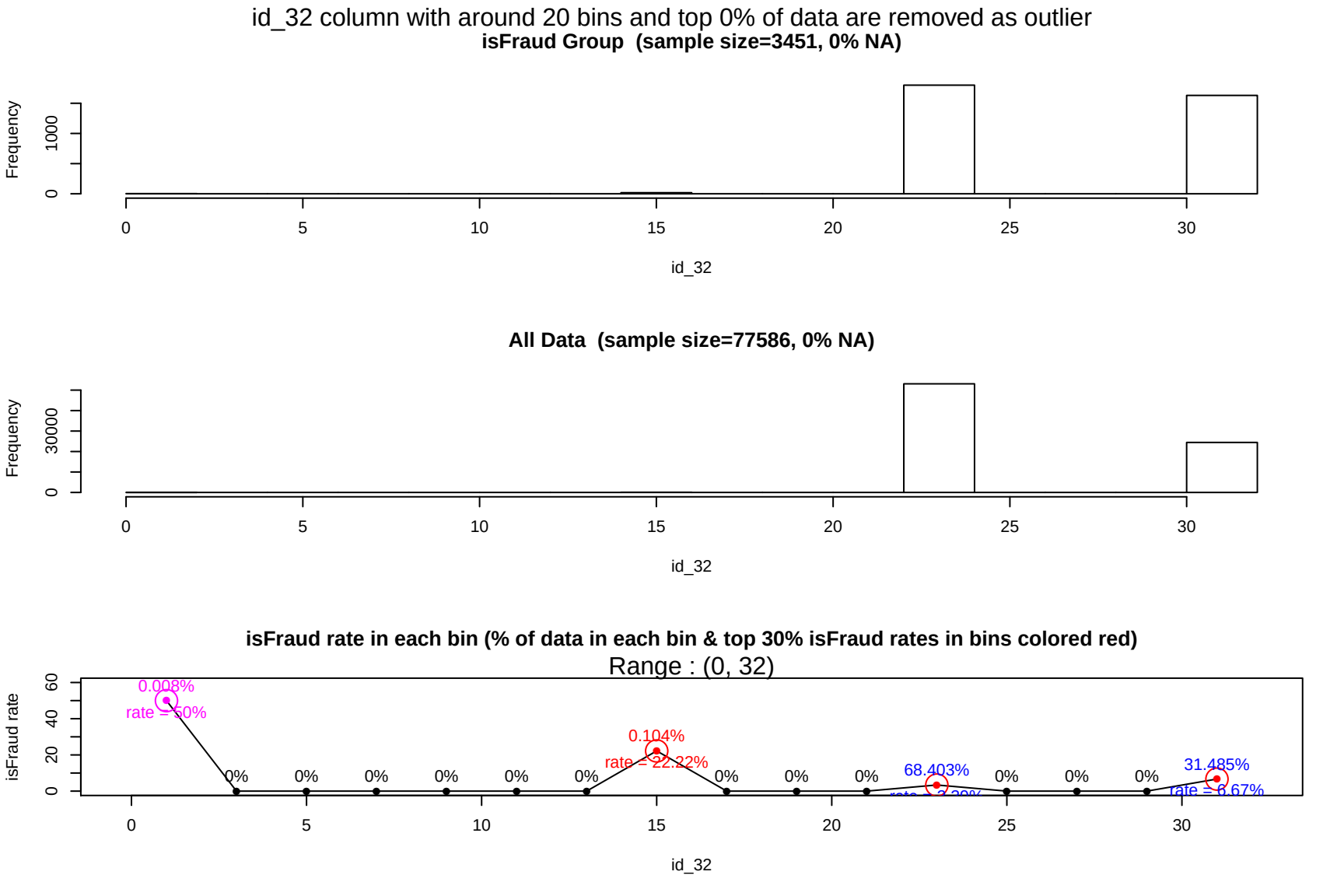

id_21 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=426, 0% NA)
Frequency
0
100
200
200
400
600
800
id_21
All Data (sample size=5159, 0% NA)
Frequency
0
1000
2500
200
400
600
800
id_21
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (100, 854)
isFraud rate
0
20
40
60
1.725%
1.977%
rate = 23.53%
7.385%
56.116%
rate = 7.91%
1.841%
1.047%
2.869%
2.888%
2.462%
5.66%
1.357%
2.985%
rate = 14.94%
1.88%
3.954%
rate = 23.53%
5.389%
rate = 12.59%
0.465%
rate = 50%
200
400
600
800
id_21


id_22 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=426, 0% NA)
Frequency
0
150
300
10
15
20
25
30
35
40
45
id_22
All Data (sample size=5169, 0% NA)
Frequency
0
2000
10
15
20
25
30
35
40
45
id_22
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (10, 44)
isFraud rate
0
20
40
0.116%
91.623%
rate = 7.31%
0%
0.155%
0.116%
rate = 16.67%
0.232%
0.097%
0.077%
0.058%
0%
0.097%
0.735%
rate = 42.11%
0.193%
0.058%
0.135%
6.268%
rate = 19.44%
0.039%
10
15
20
25
30
35
40
45
id_22


id_24 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=402, 0% NA)
Frequency
0
100
250
15
20
25
id_24
All Data (sample size=4747, 0% NA)
Frequency
0
1000
2500
15
20
25
id_24
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (11, 26)
isFraud rate
0
40
80
59.427%
0%
0%
33.579%
4.635%
0.19%
rate = 100%
0.779%
0.105%
rate = 40%
0%
0.695%
0%
0.021%
0.253%
rate = 50%
0.147%
rate = 28.57%
0.169%
rate = 50%
15
20
25
id_24


id_25 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=417, 0% NA)
Frequency
0
50
150
100
200
300
400
500
id_25
All Data (sample size=5132, 0% NA)
Frequency
0
1000
2500
100
200
300
400
500
id_25
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (100, 548)
isFraud rate
0
20
60
1.306%
2.923%
rate = 27.28%
1.832%
1.442%
rate = 18.92%
1.325%
7.054%
1.695%
1.968%
rate = 26%
1.267%
1.189%
0.721%
rate = 13.51%
49.376%
rate = 7.46%
2.299%
2.26%
1.111%
0.955%
rate = 16.33%
5.885%
2.332%
rate = 11.11%
1.481%
rate = 61.84%
1.773%
5.904%
2.299%
0.604%
100
200
300
400
500
id_25


id_26 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=425, 0% NA)
Frequency
0
40
80
100
120
140
160
180
200
220
id_26
All Data (sample size=5163, 0% NA)
Frequency
0
400
800
100
120
140
160
180
200
220
id_26
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (100, 216)
isFraud rate
0
20
40
17.219%
rate = 5.06%
0.91%
0.6%
8.212%
2.344%
0.465%
0.988%
rate = 19.51%
11.059%
rate = 5.25%
7.98%
0.31%
rate = 12.5%
0.872%
rate = 35.56%
16.851%
rate = 5.75%
7.612%
0.058%
0.484%
13.616%
2.014%
0.678%
0.504%
0.174%
3.06%
rate = 25.32%
1.085%
100
120
140
160
180
200
220
id_26


id_32 column with around 20 bins and top 0% of data are removed as outlier
isFraud Group (sample size=3451, 0% NA)
Frequency
0
1000
0
5
10
15
20
25
30
id_32
All Data (sample size=77586, 0% NA)
Frequency
0
30000
0
5
10
15
20
25
30
id_32
isFraud rate in each bin (% of data in each bin & top 30% isFraud rates in bins colored red)
Range : (0, 32)
isFraud rate
0
20
40
60
0.008%
rate = 50%
0%
0%
0%
0%
0%
0%
0.104%
rate = 22.22%
0%
0%
0%
68.403%
0%
0%
0%
31.485%
rate = 6.67%
0
5
10
15
20
25
30
id_32