

Explore the Toyota Corolla Data Set Using R and R Studio

Didem Bulut Aykurt

MIS510-1 – Data Mining and Visualization

Colorado State University-Global Campus

Dr. Emmanuel Tsukerman

Now 20, 2022

## **Uploading and importing the Toyota Corolla CSV file into RStudio**

When do I start to learn new knowledge that comes up with why I need that tool? Where will I use it? What kind of problem would I solve? Why do most companies work with them for the R program? First, I would like to give some minor information about all my questions. The best reason to use R is that over 12K packages and libraries are available and handle large amounts of unstructured data, all for free or less than other programs. Also, safe and security are the reasons all companies need them. Let's look at my work; R has an analysis tool to create graphs and charts with big and heavy data. It has an excellent console to manage code, data, and results and display graphs and charts. I download the R program, then the RStudio software project. Open the RStudio; go File, New File, R Script, then new page pop up.

I worked with the ToyotaCorolla.xlsx dataset. I convert the xlsx file to a CSV file with excel. Load the file into R Studio with `<- read.csv("file path")`, `header=TRUE` (if the dataset has a True header)).

Hence, let's dive deep into the Toyota Corolla dataset with the R programming language with RStudio. The ToyotaCorolla dataset contains car detail information that helps to compare car quality-related prices. This dataset is available from CSUGlobal. The function of `"dim()"` shows dataset dimension as the Toyota corolla dataset includes 1436 observations with 39 variables such as Id, Model, Price, Age\_08\_04, Mfg\_Month, Mfg\_Year, KM, Fuel\_Type, HP, Met\_Color, Color, Automatic, CC, Doors, Cylinders, Gears, Quarterly\_Tax, Weight, Mfr\_Guarantee, BOVAG\_Guarantee, Guarantee\_Period, ABS, Airbag\_1, Airbag\_2, Airco, Automatic\_airco, Boardcomputer, CD\_Player, Central\_Lock, Power\_Windows, Power\_Steering, Radio, Mistlamps, Sport\_Model,

Backseat\_divider, Metallic\_Rim, Radio\_cassette, Parking\_Assistant, Tow\_Bar. The syntax of “view()” shows the whole table nicely on the other page. The “head()” shows all variables with six rows. The “mean()” and “summary()” display statistical results. When looking at the Price statistical result mean is the average number on the dataset of \$10,731, the median is the middle number on the data point of \$9,900, and the max is \$32,500. Using the sample data is an easy and fast way to compare results between samples. R has a syntax is “sample <- sample(row.names(dataset name), number of observation) to create a random sample. Each variable has its category with “class(variable name)” and levels with “unique(variable names), a new name with “colnames(dataset name) <-c(“put new column name”). Use the “model.matrix()” to create dummy table.

Concerns:

I am facing an issue with implementing complex code in R, like textbooks show levels() syntax that doesn't work with the Toyota corolla dataset; why the levels() function didn't work with the dataset?

**Figure 1:** *R code to import toyotacorolla.csv file into RStudio and explore the dataset.*

## Critical Thinking 1- Option 1

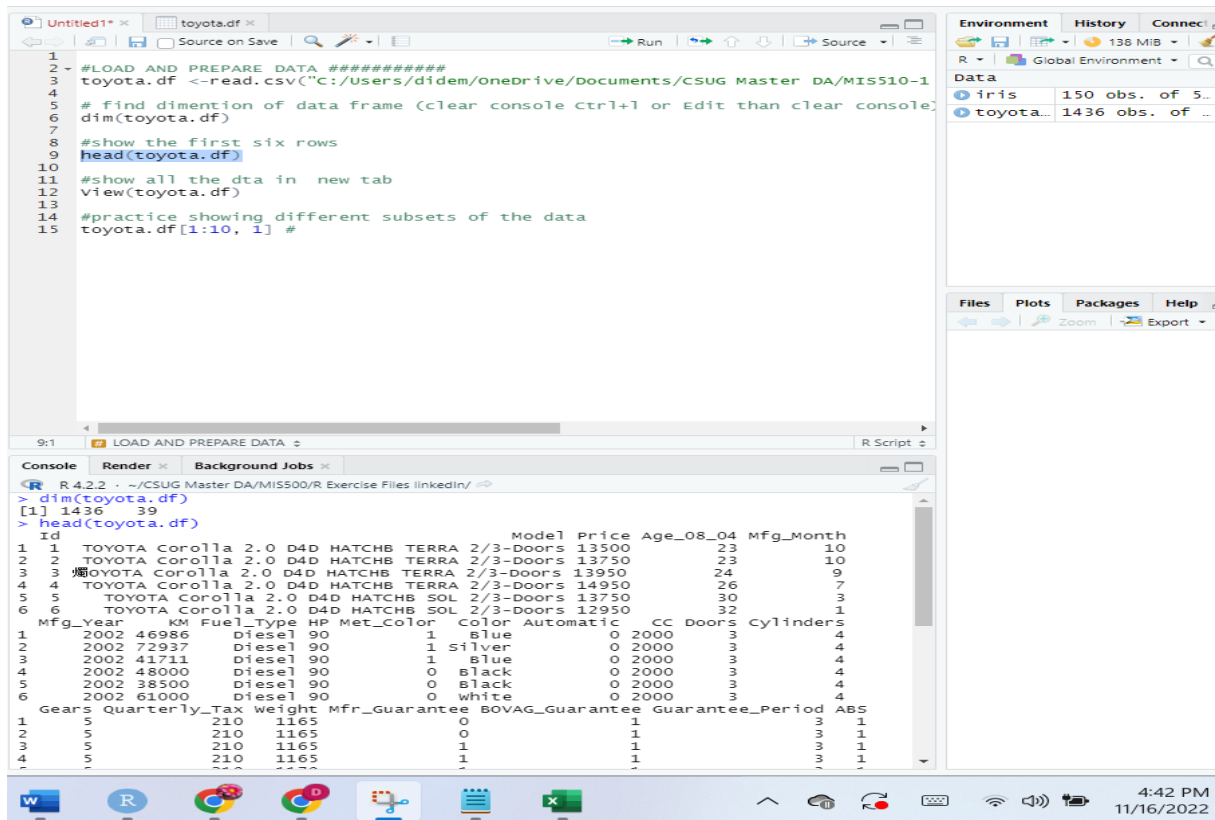


Figure 2: R code as a “view ()” syntax result table.

## Critical Thinking 1- Option 1

The screenshot displays the RStudio environment with a data frame named 'toyota.df' loaded. The data frame contains 17 rows and 7 columns: Id, Model, Price, Age\_08\_04, Mfg\_Month, and Mfg\_Year. The 'Model' column lists various Toyota Corolla models, including 'TOYOTA Corolla 2.0 D4D HATCHB TERRA 2/3-Doors' and 'TOYOTA Corolla 1.8 16V VVTU 3DR T SPORT 2/3-Doors'. The 'Price' column shows values ranging from 12,250 to 22,750. The 'Age\_08\_04' column shows ages from 23 to 32. The 'Mfg\_Month' column shows months from 1 to 11. The 'Mfg\_Year' column shows years from 2 to 2. The console shows the command 'view(toyota.df)' executed. The environment pane on the right shows the 'Global Environment' with the 'toyota...' object containing 1436 observations. The status bar at the bottom indicates the R version is 4.2.2 and the file path is '~/CSUG Master DA/MIS500/R Exercise Files/linkedIn/'.

Id	Model	Price	Age_08_04	Mfg_Month	Mfg_Year
1	TOYOTA Corolla 2.0 D4D HATCHB TERRA 2/3-Doors	13500	23	10	2
2	TOYOTA Corolla 2.0 D4D HATCHB TERRA 2/3-Doors	13750	23	10	2
3	TOYOTA Corolla 2.0 D4D HATCHB TERRA 2/3-Doors	13950	24	9	2
4	TOYOTA Corolla 2.0 D4D HATCHB TERRA 2/3-Doors	14950	26	7	2
5	TOYOTA Corolla 2.0 D4D HATCHB SOL 2/3-Doors	13750	30	3	2
6	TOYOTA Corolla 2.0 D4D HATCHB SOL 2/3-Doors	12950	32	1	2
7	TOYOTA Corolla 2.0 D4D 90 3DR TERRA 2/3-Doors	16900	27	6	2
8	TOYOTA Corolla 2.0 D4D 90 3DR TERRA 2/3-Doors	18600	30	3	2
9	TOYOTA Corolla 1800 T SPORT VVT I 2/3-Doors	21500	27	6	2
10	TOYOTA Corolla 1.9 D HATCHB TERRA 2/3-Doors	12950	23	10	2
11	TOYOTA Corolla 1.8 VVTU-i T-Sport 3-Doors 2/3-Doors	20950	25	8	2
12	TOYOTA Corolla 1.8 16V VVTU 3DR T SPORT BNS 2/3-Doors	19950	22	11	2
13	TOYOTA Corolla 1.8 16V VVTU 3DR T SPORT 2/3-Doors	19600	25	8	2
14	TOYOTA Corolla 1.8 16V VVTU 3DR T SPORT 2/3-Doors	21500	31	2	2
15	TOYOTA Corolla 1.8 16V VVTU 3DR T SPORT 2/3-Doors	22500	32	1	2
16	TOYOTA Corolla 1.8 16V VVTU 3DR T SPORT 2/3-Doors	22000	28	5	2
17	TOYOTA Corolla 1.8 16V VVTU 3DR T SPORT 2/3-Doors	22750	30	3	2

**Figur3:** The R code display different subsets of the Toyota corolla dataset.

### Critical Thinking 1- Option 1

```
1 #LOAD AND PREPARE DATA #####
2 toyota.df <-read.csv("C:/Users/didem/OneDrive/Documents/CSUG Master DA/MIS510-1 Data Mining_4 term/Module1/Toyota.csv")
3
4 # find dimation of data frame (clear console ctrl+l or Edit than clear console)
5 dim(toyota.df)
6
7 #show the first six rows
8 head(toyota.df)
9
10 #show all the dta in new tab
11 view(toyota.df)
12
13 #practice showing different subsets of the data
14
15 #display the first column with first 10 row
16 toyota.df[1:10, 1]
17
18 #display the first 10 rows of each columns
19 toyota.df[1:10, ]
20
21
22
23
```

22:1 LOAD AND PREPARE DATA R Script

Console Render Background Jobs

R 4.2.2 ~:/CSUG Master DA/MIS500/R Exercise Files linkedIn/

```
> toyota.df[1:10, 1]
[1] 1 2 3 4 5 6 7 8 9 10
> toyota.df[1:10, ]
  Id      Model Price Age_08_04 Mfg_Month
1 1 TOYOTA Corolla 2.0 D4D HATCHB TERRA 2/3-Doors 13500 23 10
2 2 TOYOTA Corolla 2.0 D4D HATCHB TERRA 2/3-Doors 13750 23 10
3 3 TOYOTA Corolla 2.0 D4D HATCHB TERRA 2/3-Doors 13950 24 9
4 4 TOYOTA Corolla 2.0 D4D HATCHB TERRA 2/3-Doors 14950 26 7
5 5 TOYOTA Corolla 2.0 D4D HATCHB SOL 2/3-Doors 13750 30 3
6 6 TOYOTA Corolla 2.0 D4D HATCHB SOL 2/3-Doors 12950 32 1
7 7 TOYOTA Corolla 2.0 D4D 90 3DR TERRA 2/3-Doors 16900 27 6
8 8 TOYOTA Corolla 2.0 D4D 90 3DR TERRA 2/3-Doors 18600 30 3
9 9 TOYOTA Corolla 1800 T SPORT VVT I 2/3-Doors 21500 27 6
10 10 TOYOTA Corolla 1.9 D HATCHB TERRA 2/3-Doors 12950 23 10
Mfg_Year KM Fuel_Type HP Met_Color Color Automatic CC Doors Cylinders
1 2002 46986 Diesel 90 1 Blue 0 2000 3 4
2 2002 72937 Diesel 90 1 Silver 0 2000 3 4
3 2002 41711 Diesel 90 1 Blue 0 2000 3 4
4 2002 48000 Diesel 90 0 Black 0 2000 3 4
5 2002 38500 Diesel 90 0 Black 0 2000 3 4
6 2002 61000 Diesel 90 0 white 0 2000 3 4
7 2002 94612 Diesel 90 1 Grey 0 2000 3 4
```

## Critical Thinking 1- Option 1

```
!1  
!2 #show the fifth row of first 10 columns  
!3 toyota.df[5, 1:10]  
!4  
!5 #show the fifth row of some columns  
!6 toyota.df[[5, c(1:2, 4, 8:10)]]  
!7
```

LOAD AND PREPARE DATA R Script

Render Background Jobs

R 4.2.2 · ~/CSUG Master DA/MIS500/R Exercise Files linkedIn/

```
toyota.df[5, 1:10]  
id Model Price Age_08_04 Mfg_Month Mfg_Year KM Fuel_Type HP  
5 TOYOTA Corolla 2.0 D4D HATCHB SOL 2/3-Doors 13750 30 3 2002 38500 Diesel 90  
Met_Color  
0  
toyota.df[5, c(1:2, 4, 8:10)]  
id Model Age_08_04 Fuel_Type HP Met_Color  
5 TOYOTA Corolla 2.0 D4D HATCHB SOL 2/3-Doors 30 Diesel 90 0
```

4:55 PM 11/16/2022

## Critical Thinking 1- Option 1

```

27
28 #show the whole first column
29 toyota.df[, 1]
30

```

9:1 LOAD AND PREPARE DATA R Script

nsolve Render Background Jobs

R 4.2.2 ~ /CSUG Master DA/MIS500/R Exercise Files linkedIn/

```

toyota.df[, 1]
[1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
[21] 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40
[41] 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
[61] 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80
[81] 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
[101] 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121
[121] 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141
[141] 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161
[161] 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181
[181] 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201
[201] 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221
[221] 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241
[241] 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261
[261] 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281
[281] 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301
[301] 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321
[321] 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342
[341] 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362
[361] 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382
[381] 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402
[401] 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422

```

4:57 PM 11/16/2022

```

31 #different way to show the whole first column
32 toyota.df$id
33
34 #show the first 10 rows of the first column
35 toyota.df$id[1:10]
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004

```

reached getoption("max.print") -- omitted 436 entries ]

```

#show the first 10 rows of the first column
toyota.df$id[1:10]
[1] 1 2 3 4 5 6 7 8 9 10

```

6:04 PM 11/16/2022



**Figure 4:** R code result of statistical information with Toyota corolla dataset.

The screenshot shows the R Studio environment. The script editor at the top contains the following R code:

```

30
31 #find the length of the first column
32 length(toyota.df[, 1])
33
34 #find the mean of the first column
35 mean(toyota.df[, 1])
36
37 #find summary statistics for each column
38 summary(toyota.df)

```

The console at the bottom shows the execution of these commands and their results:

```

> #find the length of the first column
> length(toyota.df[, 1])
[1] 1436
>
> #find the mean of the first column
> mean(toyota.df[, 1])
[1] 721.555
>
> #find summary statistics for each column
> summary(toyota.df)

```

Id		Model		Price	
Min. :	1.0	Length:	1436	Min. :	4350
1st Qu.:	361.8	Class :	character	1st Qu.:	8450
Median :	721.5	Mode :	character	Median :	9900
Mean :	721.6			Mean :	10731
3rd Qu.:	1081.2			3rd Qu.:	11950
Max. :	1442.0			Max. :	32500

Age_08_04		Mfg_Month		Mfg_Year		KM	
Min. :	1.00	Min. :	1.000	Min. :	1998	Min. :	1
1st Qu.:	44.00	1st Qu.:	3.000	1st Qu.:	1998	1st Qu.:	43000
Median :	61.00	Median :	5.000	Median :	1999	Median :	63390
Mean :	55.95	Mean :	5.549	Mean :	2000	Mean :	68533

**Figure 5:** R code creates a sample dataset from the database.

## Critical Thinking 1- Option 1

```

39
40 #random sample of 5 observations
41 samples <- sample(row.names(toyota.df), 5)
42 toyoata.df[samples, ]

```

40:1 LOAD AND PREPARE DATA R Script

Console Render Background Jobs

R 4.2.2 ~ /CSUG Master DA/MIS500/R Exercise Files linkedIn/

```

> #random sample of 5 observations
> samples <- sample(row.names(toyota.df), 5)
> toyoata.df[samples, ]

```

	Id	Model	Price	Age_08_04	Mfg_Month	Mfg_Year
1173	1178	TOYOTA Corolla 1.6 16V LIFTB LINEA TERRA 4/5-Doors	7750	78	3	1998
965	968	TOYOTA Corolla 1.6 16V HATCHB G6 2/3-Doors	10950	65	4	1999
1180	1185	TOYOTA Corolla 1.6 16V HATCHB GTSI 2/3-Doors	8950	79	2	1998
160	161	TOYOTA Corolla 1.6 16V VVT I 5DR SOL AUT4 4/5-Doors	19750	16	5	2003
919	922	TOYOTA Corolla 1.6 16V LIFTB LINEA LUNA 4/5-Doors	8950	65	4	1999

	KM	Fuel_Type	HP	Met_Color	Color	Automatic	CC	Doors	Cylinders	Gears	Quarterly_Tax	weight
1173	95651	Petrol	110	1	Blue	0	1600	5	4	5	85	1070
965	49942	Petrol	110	0	Red	0	1600	3	4	6	85	1065
1180	93039	Petrol	110	1	Blue	0	1600	3	4	5	69	1050
160	11754	Petrol	110	0	Grey	1	1600	5	4	5	100	1180
919	57374	Petrol	110	1	Blue	0	1600	5	4	5	85	1075

	Mfr_Guarantee	BOVAG_Guarantee	Guarantee_Period	ABS	Airbag_1	Airbag_2	Airco	Automatic_airco
1173	0	1	3	1	1	1	0	0
965	0	1	3	1	1	1	1	0
1180	0	1	3	0	1	0	0	0
160	0	1	3	1	1	1	1	1
919	0	1	3	1	1	1	1	0

	Boardcomputer	CD_Player	Central_Lock	Powered_windows	Power_Steering	Radio	Mistlamps
1173	1	1	1	1	1	1	1
965	1	1	1	1	1	1	1
1180	1	1	1	1	1	1	1
160	1	1	1	1	1	1	1
919	1	1	1	1	1	1	1

```

50 #oversample toyota with over 15 age
51 samples <- sample(row.names(toyota.df), 5, prob = ifelse(toyota.df$Age_08_04>15, 0.9, 0.01))
52 toyoata.df[samples, ]

```

50:36 LOAD AND PREPARE DATA R Script

Console Render Background Jobs

R 4.2.2 ~ /CSUG Master DA/MIS500/R Exercise Files linkedIn/

```

samples <- sample(row.names(toyota.df), 5, prob = ifelse(toyota.df$Age_08_04>15, 0.9, 0.01))
toyoata.df[samples, ]

```

	Id	Model	Price	Age_08_04	Mfg_Month	Mfg_Year
05	306	TOYOTA Corolla 1.4 16V VVT I LIFTB TERRA 4/5-Doors	13950	44	1	2001
139	1144	TOYOTA Corolla 1.6 3D Terra/Comf. 2/3-Doors	8750	76	5	1998
40	843	TOYOTA Corolla 1.6 16V SEDAN LINEA TERRA 4/5-Doors	10500	65	4	1999
26	327	TOYOTA Corolla 1.6 16V VVT I LIFTB TERRA 4/5-Doors	11950	41	4	2001
54	757	TOYOTA Corolla 1.6 16V HATCHB G6 2/3-Doors	9950	59	10	1999

	KM	Fuel_Type	HP	Met_Color	Color	Automatic	CC	Doors	Cylinders	Gears	Quarterly_Tax	weight
05	39516	Petrol	97	1	Grey	0	1400	5	4	5	85	1060
139	104344	Petrol	110	1	Green	0	1600	3	4	5	69	1050
40	67003	Petrol	110	1	Red	0	1600	4	4	5	69	1035
26	34000	Petrol	110	1	Red	0	1600	5	4	5	85	1075
54	80430	Petrol	110	1	Black	0	1600	3	4	6	85	1065

	Mfr_Guarantee	BOVAG_Guarantee	Guarantee_Period	ABS	Airbag_1	Airbag_2	Airco	Automatic_airco
05	1	1	3	1	1	1	0	0
139	0	1	6	0	1	0	1	0
40	1	1	3	1	1	1	1	0
26	0	1	3	1	1	1	1	0
54	0	1	3	1	1	1	1	0

	Boardcomputer	CD_Player	Central_Lock	Powered_windows	Power_Steering	Radio	Mistlamps
05	1	1	1	0	1	0	0
139	1	1	1	1	1	1	1
40	1	1	1	1	1	1	1
26	1	1	1	1	1	1	1
54	1	1	1	1	1	1	1

Figure 6: R code result of types of variables.

## Critical Thinking 1- Option 1

```
54 #print a list of variables to the screen
55 names(toyota.df)
```

55:1 LOAD AND PREPARE DATA R Script

Console Render Background Jobs

R 4.2.2 ~ /CSUG Master DA/MIS500/R Exercise Files linkedIn/

```
> names(toyota.df)
```

[1]	"Id"	"Model"	"Price"	"Age_08_04"
[5]	"Mfg_Month"	"Mfg_Year"	"KM"	"Fuel_Type"
[9]	"HP"	"Met_Color"	"Color"	"Automatic"
[13]	"CC"	"Doors"	"Cylinders"	"Gears"
[17]	"Quarterly_Tax"	"Weight"	"Mfr_Guarantee"	"BOVAG_Guarantee"
[21]	"Guarantee_Period"	"ABS"	"Airbag_1"	"Airbag_2"
[25]	"Airco"	"Automatic_airco"	"Boardcomputer"	"CD_Player"
[29]	"Central_Lock"	"Powered_windows"	"Power_Steering"	"Radio"
[33]	"Mistlamps"	"Sport_Model"	"Backseat_Divider"	"Metallic_Rim"
[37]	"Radio_cassette"	"Parking_Assistant"	"Tow_Bar"	

>

W R Chrome D VS Code Notepad X

6:23 PM 11/16/2022

```
i7 #print the list in a useful column format
i8 t(t(names(toyota.df)))
```

3:1 LOAD AND PREPARE DATA R Script

Console Render Background Jobs

R 4.2.2 ~ /CSUG Master DA/MIS500/R Exercise Files linkedIn/

```
:t(names(toyota.df)))
```

```
[,1]
1,] "Id"
2,] "Model"
3,] "Price"
4,] "Age_08_04"
5,] "Mfg_Month"
6,] "Mfg_Year"
7,] "KM"
8,] "Fuel_Type"
9,] "HP"
10,] "Met_Color"
11,] "Color"
12,] "Automatic"
13,] "CC"
14,] "Doors"
15,] "Cylinders"
16,] "Gears"
17,] "Quarterly_Tax"
18,] "Weight"
19,] "Mfr_Guarantee"
20,] "BOVAG_Guarantee"
```

W R Chrome D VS Code Notepad X

6:27 PM 11/16/2022

## Critical Thinking 1- Option 1

```

60 # change first column's name
61 colnames(toyota.df)[1] <-c("Row Number")
62 names(toyota.df)[1]

```

LOAD AND PREPARE DATA R Script

insole Render Background Jobs

R 4.2.2 · ~/CSUG Master DA/MIS500/R Exercise Files linkedIn/

```

3,] "Airbag_1"
4,] "Airbag_2"
5,] "Airco"
6,] "Automatic_airco"
7,] "Boardcomputer"
8,] "CD_Player"
9,] "Central_Lock"
0,] "Powered_windows"
1,] "Power_Steering"
2,] "Radio"
3,] "Mistlamps"
4,] "Sport_Model"
5,] "Backseat_Divider"
6,] "Metallic_Rim"
7,] "Radio_cassette"
8,] "Parking_Assistant"
9,] "Tow_Bar"
colnames(toyota.df)[1] <-c("Row Number")
colnames(toyota.df)[1] <-c("Row Number")
names(toyota.df)[1]
] "Row Number"

```

6:32 PM 11/16/2022

```

64 #Price is a integer variable
65 class(toyota.df$Price)
66
67 #show the different way variable category
68 class(toyota.df[,8])
69
70 #it can take one of levels
71
72 unique(toyota.df[,8])
73
74 #Model is a character variable
75 class(toyota.df$Model)
76
77

```

LOAD AND PREPARE DATA R Script

Console Render Background Jobs

R 4.2.2 · ~/CSUG Master DA/MIS500/R Exercise Files linkedIn/

```

mean :0.1455 mean :0.002786 mean :0.2779
3rd Qu.:0.0000 3rd Qu.:0.000000 3rd Qu.:1.0000
Max. :1.0000 Max. :1.000000 Max. :1.0000
> class(toyota.df$Model)
[1] "character"
> #Price is a integer variable
> class(toyota.df$Price)
[1] "integer"
>
> #show the different way variable category
> class(toyota.df[,8])
[1] "character"
>
> #it can take one of levels
>
> unique(toyota.df[,8])
[1] "Diesel" "Petrol" "CNG"
>
> #Model is a character variable
> class(toyota.df$Model)
[1] "character"
>

```

Files Plots Packages Help

Install Update

Na...	Description	...
	to MIME Types	
<input type="checkbox"/> ml...	Machine Learning Benchmark Problems	2.1-3
<input type="checkbox"/> mlr	Machine Learning in R	2.1-
<input type="checkbox"/> mlr3	Machine Learning in R - Next Generation	0.1-
<input type="checkbox"/> mlr...	Performance Measures for 'mlr3'	0.4-
<input type="checkbox"/> mlr...	Helper Functions for 'mlr3'	0.1-
<input type="checkbox"/> Mo...	Rapid Calculation of Model Metrics	1.2-
<input type="checkbox"/> mo...	Modelling Functions that Work with the Pipe	0.1-
<input type="checkbox"/> nlme	Linear and Nonlinear Mixed Effects Models	3.1-16C
<input type="checkbox"/> nnet	Feed-Forward Neural	7.3-18

1:53 PM 11/18/2022

**Figure 7:** *Create a dummy table with the Toyota corolla dataset in R code.*

