assignment-2

September 13, 2023

Importing Libraries

```
[1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

Loading Dataset

```
[18]: dataset=pd.read_csv("Car_crashes.csv")
```

[19]: dataset

```
[19]:
                           State \
      0
                         Alabama
      1
                         Alaska
      2
                        Arizona
      3
                       Arkansas
      4
                     California
      5
                       Colorado
      6
                    Connecticut
      7
                       Delaware
      8
          District of Columbia
      9
                        Florida
      10
                         Georgia
      11
                         Hawaii
      12
                           Idaho
      13
                       Illinois
      14
                         Indiana
      15
                            Iowa
      16
                         Kansas
      17
                       Kentucky
      18
                      Louisiana
      19
                          Maine
      20
                       Maryland
      21
                  Massachusetts
      22
                       Michigan
      23
                      Minnesota
```

24	Mississip	pi						
25	Missou	-						
26	Montar	na						
27	Nebrasl	ka						
28	Neva	da						
29	New Hampshin	re						
30	New Jerse	ey						
31	New Mexic	-						
32	New Yor							
33	North Carolin	na						
34	North Dakot	ta						
35	Oh:							
36	Oklahor	na						
37	Orego	on						
38	Pennsylvan							
39	Rhode Islam							
40	South Carolin	na						
41	South Dakot	ta						
42	Tennesse							
43	Texa	as						
44	Uta	ah						
45	Vermon	nt						
46	Virgin							
47								
47	Washingto	on						
48	West Virgin	ia						
48 49	West Virgin Wiscons	ia in						
48	West Virgin	ia in						
48 49	West Virgin Wiscons	ia in						
48 49	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	collisions per	billion	miles '	\
48 49 50	West Virgin Wiscons	ia in ng	in	fatal	_	billion	miles	\
48 49 50	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8	billion	miles `	\
48 49 50	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	_	billion	miles `	\
48 49 50	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1	billion	miles `	\
48 49 50 0 1 2	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1 18.6	billion	miles `	\
48 49 50 0 1 2 3	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1 18.6 22.4	billion	miles `	\
48 49 50 0 1 2 3 4	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1 18.6 22.4 12.0	billion	miles	`
48 49 50 0 1 2 3	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1 18.6 22.4	billion	miles	`
48 49 50 0 1 2 3 4 5	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1 18.6 22.4 12.0 13.6	billion	miles	`
48 49 50 0 1 2 3 4 5 6	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1 18.6 22.4 12.0 13.6 10.8	billion	miles	\
48 49 50 0 1 2 3 4 5 6 7	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1 18.6 22.4 12.0 13.6 10.8 16.2	billion	miles	\
48 49 50 0 1 2 3 4 5 6 7 8	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1 18.6 22.4 12.0 13.6 10.8	billion	miles	\
48 49 50 0 1 2 3 4 5 6 7	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1 18.6 22.4 12.0 13.6 10.8 16.2 5.9	billion	miles	\
48 49 50 0 1 2 3 4 5 6 7 8 9	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1 18.6 22.4 12.0 13.6 10.8 16.2 5.9	billion	miles	\
48 49 50 0 1 2 3 4 5 6 7 8 9 10	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1 18.6 22.4 12.0 13.6 10.8 16.2 5.9 17.9 15.6	billion	miles	\
48 49 50 0 1 2 3 4 5 6 7 8 9 10 11	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1 18.6 22.4 12.0 13.6 10.8 16.2 5.9 17.9 15.6 17.5	billion	miles	\
48 49 50 0 1 2 3 4 5 6 7 8 9 10	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1 18.6 22.4 12.0 13.6 10.8 16.2 5.9 17.9 15.6	billion	miles	\
48 49 50 0 1 2 3 4 5 6 7 8 9 10 11 12	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1 18.6 22.4 12.0 13.6 10.8 16.2 5.9 17.9 15.6 17.5 15.3	billion	miles	\
48 49 50 0 1 2 3 4 5 6 7 8 9 10 11 12 13	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1 18.6 22.4 12.0 13.6 10.8 16.2 5.9 17.9 15.6 17.5 15.3 12.8	billion	miles	\
48 49 50 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1 18.6 22.4 12.0 13.6 10.8 16.2 5.9 17.9 15.6 17.5 15.3 12.8 14.5	billion	miles	\
48 49 50 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1 18.6 22.4 12.0 13.6 10.8 16.2 5.9 17.9 15.6 17.5 15.3 12.8 14.5 15.7	billion	miles	\
48 49 50 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1 18.6 22.4 12.0 13.6 10.8 16.2 5.9 17.9 15.6 17.5 15.3 12.8 14.5	billion	miles	\
48 49 50 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15	West Virgin: Wiscons: Wyomin	ia in ng	in	fatal	18.8 18.1 18.6 22.4 12.0 13.6 10.8 16.2 5.9 17.9 15.6 17.5 15.3 12.8 14.5 15.7	billion	miles	\

18	20.5
19	15.1
20	12.5
21	8.2
22	14.1
23	9.6
24	17.6
25	16.1
26	21.4
27	14.9
28	14.7
29	11.6
30	11.2
31	18.4
32	12.3
33	16.8
34	23.9
35	14.1
36	19.9
37	12.8
38	18.2
39	11.1
40	23.9
41	19.4
42	19.5
43	19.4
44	11.3
45	13.6
46	12.7
47	10.6
48	23.8
49	13.8
50	17.4
0	Percentage Of Drivers Involved In Fatal Collisions Who Were Speeding \
0	39 44
1	41
2	35
3	18
4	35
5	37
6	46
7	38
8	34
9	21
10	19
11	54

13	12	36
14		
15		
17		
18 35 38 20 34 21 23 22 24 23 22 24 24 23 25 26 43 26 26 26 26 26 26 26 2	16	27
19	17	19
20	18	35
21		
22 24 23 23 24 25 26 26 26 26 26 26 26		
23 23 24 15 25 26 26 26 26 26 26 2		
24		
25		
26 39 27 13 28 37 29 35 36 36 36 36 36 36 32 32		
27		
28		
29		
30		
31		
32		
33 39 34 23 35 28 32 32 37 32 37 33 34 40 40 41 41 41 41 41 4		
34		
35		
36		
37		
\$\frac{50}{34}\$ \$\frac{40}{40}\$ \$\frac{38}{38}\$ \$\frac{41}{41}\$ \$\frac{31}{31}\$ \$\frac{42}{42}\$ \$\frac{40}{43}\$ \$\frac{40}{44}\$ \$\frac{43}{43}\$ \$\frac{43}{45}\$ \$\frac{30}{46}\$ \$\frac{19}{47}\$ \$\frac{42}{48}\$ \$\frac{34}{49}\$ \$\frac{36}{50}\$ \$\frac{50}{42}\$ Percentage Of Drivers Involved In Fatal Collisions Who Were Alcohol-Impaired \tag{0}{1} \$\frac{30}{1} \$\frac{30}{1} \$\frac{30}{1} \$\frac{30}{1} \$\frac{30}{25} \$\frac{2}{28} \$\frac{3}{3} \$\frac{30}{1} \$\frac{1}{25} \$\frac{2}{28} \$\frac{3}{3} \$\frac{3}{26}\$ \$\frac{1}{26}\$ \$\		
39		
40		
41		
42		
43 44 45 46 47 48 48 49 49 49 49 40 40 41 43 44 49 40 43 44 40 41 41 41 42 48 49 40 41 41 42 48 49 40 41 41 42 48 49 40 41 41 41 42 48 48 49 40 49 49 40 40 41 42 48 48 49 49 40 49 49 40 40 40 40 41 42 48 49 49 40 40 40 41 42 48 49 49 40 40 40 40 40 40 40 40 40 40 40 40 40		
44 45 30 46 19 42 42 48 34 49 36 50 42		
45		
19 47 48 48 49 50 Percentage Of Drivers Involved In Fatal Collisions Who Were Alcohol-Impaired \ 30 1 25 2 38 3 30 1 25 28 3		
47 48 49 49 40 40 40 40 40 40 40 40 40 40 40 40 40		
48		
36 42 Percentage Of Drivers Involved In Fatal Collisions Who Were Alcohol-Impaired 36 37 38 38 38 38 38 38 38 38 38 38 38 38 38		
Percentage Of Drivers Involved In Fatal Collisions Who Were Alcohol-Impaired 30 1 25 2 28 3 26		
Percentage Of Drivers Involved In Fatal Collisions Who Were Alcohol-Impaired 30 1 25 2 28 3 26		
\ 0 30 1 25 28 3 26		
0 30 1 25 2 28 3 26		Percentage Of Drivers Involved In Fatal Collisions Who Were Alcohol-Impaired
1 25 2 28 3 26	\	
2 28 3 26	0	
3 26		
4 28		
	4	28

5	28
6	36
7	30
8	27
9	29
10	25
11	41
12	29
13	34
14	29
15	25
16	24
17	23
18	33
19	30
20	32
21	35
22	28
23	29
24	31
25	34
26	44
27	35
28	32
29	30
30	28
31	27
32	29
33	31
34	42
35	34
36	29
37	26
38	31
39	38
40	41
41	33
42	29
43	38
44	16
45	30
46	27
47	33
48	28
49	33
50	
50	32

$\hbox{ Percentage Of Drivers Involved In Fatal Collisions Who Were Not Distracted } \\$

\	
0	96
1	90
2	84
3	94
4	91
5	79
6	87
7	87
	100
9	92
10	95
11	82
12	85
13	94
14	95
15	97
16	77
17	78
18	73
19	87
20	71
21	87
22	95
23	88
24	10
25	92
26	84
27	93
28	95
29	87
30	86
31	67
32	88
33	94
34	99
35	99
36	92
37	67
38	96
39	92
40	96
41	98
42	82
43	91
44	88

45				96						
46				87						
47				82						
48				97						
49				39						
50				81						
								_		_
_	Percentage Of Drivers		In Fatal	Collisions	Who	Had	Not	Been	Involv	ed
	Any Previous Accidents	\								
0				80						
1				94						
2				96						
3				95						
4				89						
5				95						
6				82						
7				99						
8				100						
9				94						
10				93						
11				87						
12				98						
13				96						
14				95						
15				87						
16				85						
17				76						
18				98						
19				84						
20				99						
21				80						
22				77						
23				88						
24				100						
25				84						
26				85						
27				90						
28				99						
29				83						
30				78						
31				98						
32				80						
33				81						
34				86						
35				82						
36				94						
37				90						

38 39 40 41 42 43 44 45 46 47 48 49 50		88 79 81 86 81 87 96 95 88 86 87 84
	Car Insurance Premiums (\$)	\
0	784.55	
1 2	1053.48 899.47	
3	827.34	
4	878.41	
5 6	835.50 1068.73	
7	1137.87	
8	1273.89	
9	1160.13	
10 11	913.15 861.18	
12	641.96	
13	803.11	
14	710.46	
15 16	649.06 780.45	
17	872.51	
18	1281.55	
19 20	661.88 1048.78	
21	1046.76	
22	1110.61	
23	777.18	
24	896.07	
25 26	790.32 816.21	
27	732.28	
28	1029.87	
29	746.54	
30 31	1301.52 869.85	

32									
J_		1234.31							
33		708.24							
34		688.75							
35		697.73							
36		881.51							
37		804.71							
38		905.99							
39		1148.99							
40		858.97							
41		669.31							
42		767.91							
43		1004.75							
44		809.38							
45		716.20							
46		768.95							
47		890.03							
48		992.61							
49		670.31							
50		791.14							
	Losses incurred by	insurance	companies	for	collisions	per	insured	driver	(\$)
0					145.08				
1					133.93				
2					110.35				
3					142.39				
4					165.63				
5									
5 6					139.91				
6					139.91 167.02				
6 7					139.91 167.02 151.48				
6 7 8					139.91 167.02 151.48 136.05				
6 7					139.91 167.02 151.48				
6 7 8					139.91 167.02 151.48 136.05				
6 7 8 9					139.91 167.02 151.48 136.05 144.18				
6 7 8 9 10 11					139.91 167.02 151.48 136.05 144.18 142.80 120.92				
6 7 8 9 10 11 12					139.91 167.02 151.48 136.05 144.18 142.80 120.92 82.75				
6 7 8 9 10 11 12					139.91 167.02 151.48 136.05 144.18 142.80 120.92 82.75 139.15				
6 7 8 9 10 11 12					139.91 167.02 151.48 136.05 144.18 142.80 120.92 82.75				
6 7 8 9 10 11 12					139.91 167.02 151.48 136.05 144.18 142.80 120.92 82.75 139.15				
6 7 8 9 10 11 12 13 14					139.91 167.02 151.48 136.05 144.18 142.80 120.92 82.75 139.15 108.92 114.47				
6 7 8 9 10 11 12 13 14 15 16					139.91 167.02 151.48 136.05 144.18 142.80 120.92 82.75 139.15 108.92 114.47 133.80				
6 7 8 9 10 11 12 13 14 15 16 17					139.91 167.02 151.48 136.05 144.18 142.80 120.92 82.75 139.15 108.92 114.47 133.80 137.13				
6 7 8 9 10 11 12 13 14 15 16 17 18					139.91 167.02 151.48 136.05 144.18 142.80 120.92 82.75 139.15 108.92 114.47 133.80 137.13 194.78				
6 7 8 9 10 11 12 13 14 15 16 17 18 19					139.91 167.02 151.48 136.05 144.18 142.80 120.92 82.75 139.15 108.92 114.47 133.80 137.13 194.78 96.57				
6 7 8 9 10 11 12 13 14 15 16 17 18					139.91 167.02 151.48 136.05 144.18 142.80 120.92 82.75 139.15 108.92 114.47 133.80 137.13 194.78				
6 7 8 9 10 11 12 13 14 15 16 17 18 19 20					139.91 167.02 151.48 136.05 144.18 142.80 120.92 82.75 139.15 108.92 114.47 133.80 137.13 194.78 96.57 192.70				
6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21					139.91 167.02 151.48 136.05 144.18 142.80 120.92 82.75 139.15 108.92 114.47 133.80 137.13 194.78 96.57 192.70 135.63				
6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22					139.91 167.02 151.48 136.05 144.18 142.80 120.92 82.75 139.15 108.92 114.47 133.80 137.13 194.78 96.57 192.70 135.63 152.26				
6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23					139.91 167.02 151.48 136.05 144.18 142.80 120.92 82.75 139.15 108.92 114.47 133.80 137.13 194.78 96.57 192.70 135.63 152.26 133.35				
6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22					139.91 167.02 151.48 136.05 144.18 142.80 120.92 82.75 139.15 108.92 114.47 133.80 137.13 194.78 96.57 192.70 135.63 152.26				
6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23					139.91 167.02 151.48 136.05 144.18 142.80 120.92 82.75 139.15 108.92 114.47 133.80 137.13 194.78 96.57 192.70 135.63 152.26 133.35				

```
26
                                                          85.15
      27
                                                         114.82
      28
                                                         138.71
      29
                                                         120.21
      30
                                                         159.85
      31
                                                         120.75
      32
                                                         150.01
      33
                                                         127.82
      34
                                                         109.72
      35
                                                         133.52
      36
                                                         178.86
      37
                                                         104.61
      38
                                                         153.86
      39
                                                         148.58
      40
                                                         116.29
      41
                                                          96.87
      42
                                                         155.57
      43
                                                         156.83
      44
                                                         109.48
      45
                                                         109.61
      46
                                                         153.72
      47
                                                         111.62
      48
                                                         152.56
      49
                                                         106.62
      50
                                                         122.04
[20]: dataset.head()
[20]:
              State \
      0
            Alabama
      1
             Alaska
      2
            Arizona
      3
           Arkansas
      4 California
         Number of drivers involved in fatal collisions per billion miles \
                                                          18.8
      0
      1
                                                          18.1
      2
                                                          18.6
      3
                                                          22.4
      4
                                                          12.0
         Percentage Of Drivers Involved In Fatal Collisions Who Were Speeding \
      0
      1
                                                            41
      2
                                                            35
      3
                                                            18
```

```
Percentage Of Drivers Involved In Fatal Collisions Who Were Alcohol-Impaired
      0
                                                          30
                                                          25
      1
      2
                                                          28
      3
                                                          26
      4
                                                          28
         Percentage Of Drivers Involved In Fatal Collisions Who Were Not Distracted \
      0
                                                          96
                                                          90
      1
      2
                                                          84
      3
                                                          94
      4
                                                          91
         Percentage Of Drivers Involved In Fatal Collisions Who Had Not Been Involved
      In Any Previous Accidents \
                                                          80
      0
      1
                                                          94
      2
                                                          96
      3
                                                          95
      4
                                                          89
         Car Insurance Premiums ($)
      0
                              784.55
      1
                             1053.48
      2
                              899.47
      3
                              827.34
      4
                              878.41
         Losses incurred by insurance companies for collisions per insured driver ($)
      0
                                                      145.08
                                                      133.93
      1
      2
                                                      110.35
                                                      142.39
      3
      4
                                                      165.63
[21]: dataset.tail()
[21]:
                  State \
      46
               Virginia
             Washington
      47
      48
          West Virginia
      49
              Wisconsin
                Wyoming
      50
```

```
Number of drivers involved in fatal collisions per billion miles \
46
                                                  12.7
47
                                                  10.6
48
                                                  23.8
49
                                                  13.8
50
                                                  17.4
    Percentage Of Drivers Involved In Fatal Collisions Who Were Speeding \
46
47
                                                     42
48
                                                     34
49
                                                     36
50
                                                     42
    Percentage Of Drivers Involved In Fatal Collisions Who Were Alcohol-Impaired
\
                                                     27
46
                                                     33
47
                                                     28
48
49
                                                     33
50
                                                     32
    Percentage Of Drivers Involved In Fatal Collisions Who Were Not Distracted
\
46
                                                     87
47
                                                     82
48
                                                     97
49
                                                     39
50
                                                     81
    Percentage Of Drivers Involved In Fatal Collisions Who Had Not Been Involved
In Any Previous Accidents \
46
                                                     88
47
                                                     86
48
                                                     87
49
                                                     84
50
                                                     90
    Car Insurance Premiums ($) \
46
                        768.95
47
                        890.03
48
                         992.61
49
                        670.31
                        791.14
50
```

Losses incurred by insurance companies for collisions per insured driver (\$)

```
      46
      153.72

      47
      111.62

      48
      152.56

      49
      106.62

      50
      122.04
```

Data Visualisation

1) Univariate

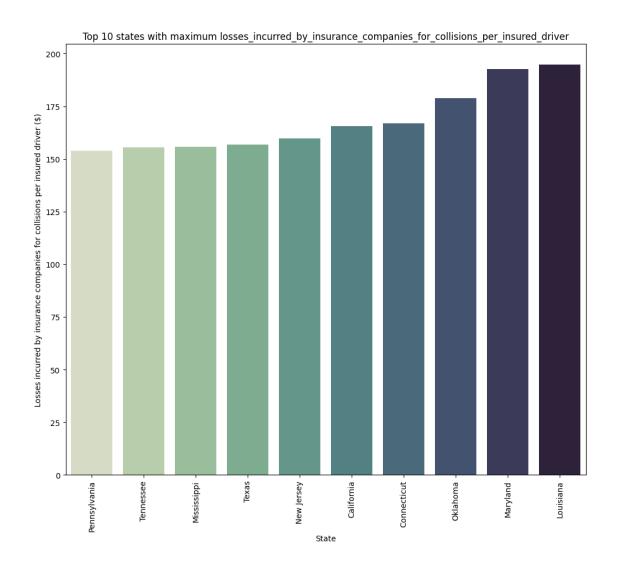
```
[34]: plt.figure(figsize=(12,10))
      d2 = dataset[["State",'Number of drivers involved in fatal collisions per_
       ⇒billion miles', 'Percentage Of Drivers Involved In Fatal Collisions Who Were⊔
       →Speeding', 'Percentage Of Drivers Involved In Fatal Collisions Who Were
       Alcohol-Impaired', 'Percentage Of Drivers Involved In Fatal Collisions Who
       \hookrightarrowWere Not Distracted', 'Percentage Of Drivers Involved In Fatal Collisions Who_\sqcup
       →Had Not Been Involved In Any Previous Accidents', 'Car Insurance Premiums,
       _{\hookrightarrow}($)','Losses incurred by insurance companies for collisions per insured_{\sqcup}
       odriver ($)']].sort_values(by='Losses incurred by insurance companies for⊔

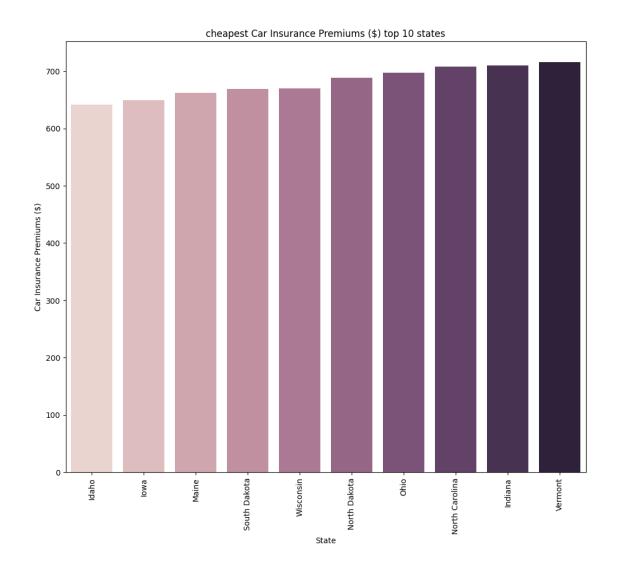
→collisions per insured driver ($)').tail(10)
      sns.barplot(x=d2['State'], y=d2['Losses incurred by insurance companies for_
       ⇔collisions per insured driver ($)'],palette=sns.

cubehelix_palette(len(d2['State']),start=0.5,rot=-0.75))

      plt.xticks(rotation= 90)
      plt.xlabel('State')
      plt.ylabel('Losses incurred by insurance companies for collisions per insured∪

driver ($)')
      plt.title('Top 10 states with maximum<sub>□</sub>
       -losses incurred by insurance companies for collisions per insured driver')
      plt.show()
```





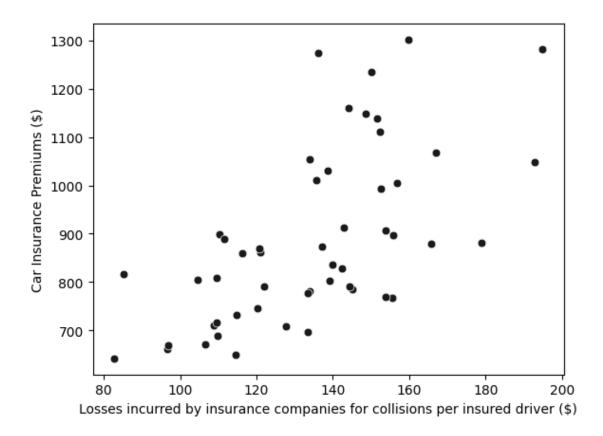
2) Bivariate

```
[39]: sns.scatterplot(x="Losses incurred by insurance companies for collisions per

insured driver ($)", y="Car Insurance Premiums ($)", data=dataset, s=40,

color=".1")
```

[39]: <Axes: xlabel='Losses incurred by insurance companies for collisions per insured driver (\$)', ylabel='Car Insurance Premiums (\$)'>



The scatterplot shows a linear relationship between car insurance premium and the losses incurred

3) Multivariate

```
d5 = dataset[["State",'Number of drivers involved in fatal collisions per_

⇒billion miles','Percentage Of Drivers Involved In Fatal Collisions Who Were

⇒Speeding','Percentage Of Drivers Involved In Fatal Collisions Who Were

⇒Alcohol-Impaired','Percentage Of Drivers Involved In Fatal Collisions Who

⇒Were Not Distracted','Percentage Of Drivers Involved In Fatal Collisions Who

⇒Had Not Been Involved In Any Previous Accidents','Car Insurance Premiums

⇒($)','Losses incurred by insurance companies for collisions per insured

⇒driver ($)']].sort_values(by='Losses incurred by insurance companies for

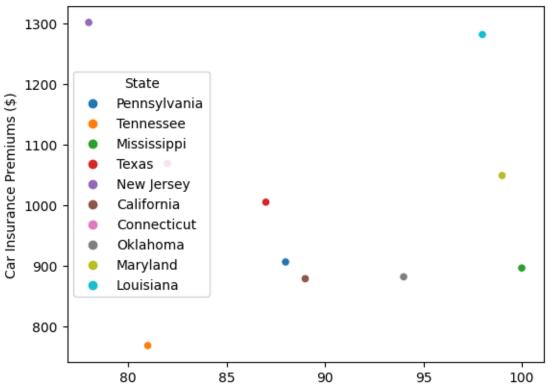
⇒collisions per insured driver ($)').tail(10)

ax = sns.scatterplot(x=d5["Percentage Of Drivers Involved In Fatal Collisions

⇒Who Had Not Been Involved In Any Previous Accidents"], y=d5["Car Insurance

⇒Premiums ($)"],

hue=d5["State"])
```



ntage Of Drivers Involved In Fatal Collisions Who Had Not Been Involved In Any Prev

Car insurance premiums are really high when % of drivers not involved in previous accident is below 85%. This also means that there are higher chances that the insurance company will have to bear the cost for the same insured driver again and if the insurance premium is low in those states then more likely that the company will incur huge losses. Thus to reduce the loss margins the insurance premiums are high in those states which again signifies that the states have a lot of high risk drivers