

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

ak = sns.load_dataset('car_crashes')
ak
```



	total	speeding	alcohol	not_distracted	no_previous	ins_premium	ins_losses
0	18.8	7.332	5.640	18.048	15.040	784.55	145.08
1	18.1	7.421	4.525	16.290	17.014	1053.48	133.93
2	18.6	6.510	5.208	15.624	17.856	899.47	110.35
3	22.4	4.032	5.824	21.056	21.280	827.34	142.39
4	12.0	4.200	3.360	10.920	10.680	878.41	165.63
5	13.6	5.032	3.808	10.744	12.920	835.50	139.91
6	10.8	4.968	3.888	9.396	8.856	1068.73	167.02
7	16.2	6.156	4.860	14.004	16.038	1137.87	151.48

```
ak.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 51 entries, 0 to 50
Data columns (total 8 columns):
#   Column                Non-Null Count  Dtype
---  -
0    total                  51 non-null    float64
1    speeding               51 non-null    float64
2    alcohol                51 non-null    float64
3    not_distracted         51 non-null    float64
4    no_previous            51 non-null    float64
5    ins_premium            51 non-null    float64
6    ins_losses             51 non-null    float64
7    abbrev                 51 non-null    object
dtypes: float64(7), object(1)
memory usage: 3.3+ KB
```

```
ak.head()
```

	total	speeding	alcohol	not_distracted	no_previous	ins_premium	ins_losses	abbrev
0	18.8	7.332	5.640	18.048	15.040	784.55	145.08	AL
1	18.1	7.421	4.525	16.290	17.014	1053.48	133.93	AK
2	18.6	6.510	5.208	15.624	17.856	899.47	110.35	AZ
3	22.4	4.032	5.824	21.056	21.280	827.34	142.39	AR
4	12.0	4.200	3.360	10.920	10.680	878.41	165.63	CA

```
ak.head(2)
```

	total	speeding	alcohol	not_distracted	no_previous	ins_premium	ins_losses	abbrev
0	18.8	7.332	5.640	18.048	15.040	784.55	145.08	AL
1	18.1	7.421	4.525	16.290	17.014	1053.48	133.93	AK

```
ak.tail(8)
```

	total	speeding	alcohol	not_distracted	no_previous	ins_premium	ins_losses	abbrev
43	19.4	7.760	7.372	17.654	16.878	1004.75	156.83	TX
44	11.3	4.859	1.808	9.944	10.848	809.38	109.48	UT
45	13.6	4.080	4.080	13.056	12.920	716.20	109.61	VT
46	12.7	2.413	3.429	11.049	11.176	768.95	153.72	VA
47	10.6	4.452	3.498	8.692	9.116	890.03	111.62	WA
48	23.8	8.092	6.664	23.086	20.706	992.61	152.56	WV
49	13.8	4.968	4.554	5.382	11.592	670.31	106.62	WI
50	17.4	7.308	5.568	14.094	15.660	791.14	122.04	WY

```
ak.tail()
```

```
total speeding alcohol not_distracted no_previous ins_premium ins_losses abbrev
46 12.7 2.413 3.429 11.049 11.176 768.95 153.72 VA

ak.shape

(51, 8)

49 13.8 4.968 4.554 5.382 11.592 670.31 106.62 WI

ak.describe()
```

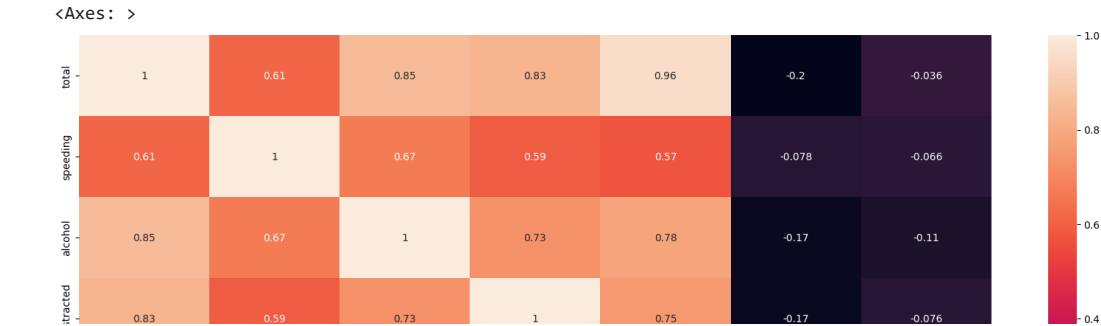
	total	speeding	alcohol	not_distracted	no_previous	ins_premium	ins_losses
count	51.000000	51.000000	51.000000	51.000000	51.000000	51.000000	51.000000
mean	15.790196	4.998196	4.886784	13.573176	14.004882	886.957647	134.493137
std	4.122002	2.017747	1.729133	4.508977	3.764672	178.296285	24.835922
min	5.900000	1.792000	1.593000	1.760000	5.900000	641.960000	82.750000
25%	12.750000	3.766500	3.894000	10.478000	11.348000	768.430000	114.645000
50%	15.600000	4.608000	4.554000	13.857000	13.775000	858.970000	136.050000
75%	18.500000	6.439000	5.604000	16.140000	16.755000	1007.945000	151.870000
max	23.900000	9.450000	10.038000	23.661000	21.280000	1301.520000	194.780000

```
corr = ak.corr()
corr

<ipython-input-12-cd014e0cc39d>:1: FutureWarning: The default value of numeric_only in DataFrame.corr is
corr = ak.corr()
```

	total	speeding	alcohol	not_distracted	no_previous	ins_premium	ins_losses
total	1.000000	0.611548	0.852613	0.827560	0.956179	-0.199702	-0.036011
speeding	0.611548	1.000000	0.669719	0.588010	0.571976	-0.077675	-0.065928
alcohol	0.852613	0.669719	1.000000	0.732816	0.783520	-0.170612	-0.112547
not_distracted	0.827560	0.588010	0.732816	1.000000	0.747307	-0.174856	-0.075970
no_previous	0.956179	0.571976	0.783520	0.747307	1.000000	-0.156895	-0.006359
ins_premium	-0.199702	-0.077675	-0.170612	-0.174856	-0.156895	1.000000	0.623116
ins_losses	-0.036011	-0.065928	-0.112547	-0.075970	-0.006359	0.623116	1.000000

```
plt.subplots(figsize = (20,10))
sns.heatmap(corr,annot=True)
```



ak["total"].value_counts()

```
14.1    2
12.8    2
13.6    2
21.4    2
19.4    2
23.9    2
14.9    1
14.7    1
11.6    1
11.2    1
18.4    1
12.3    1
16.8    1
19.9    1
17.6    1
18.2    1
11.1    1
19.5    1
11.3    1
12.7    1
10.6    1
23.8    1
13.8    1
16.1    1
18.8    1
9.6     1
18.1    1
18.6    1
22.4    1
12.0    1
10.8    1
16.2    1
5.9     1
17.9    1
15.6    1
17.5    1
15.3    1
14.5    1
15.7    1
17.8    1
20.5    1
15.1    1
12.5    1
8.2     1
17.4    1
Name: total, dtype: int64
```

ak.alcohol.value_counts()

```
5.208    2
5.640    1
4.218    1
4.704    1
3.480    1
3.136    1
4.968    1
3.567    1
10.038   1
4.794    1
5.771    1
3.328    1
5.642    1
9.799    1
9.416    1
6.402    1
5.655    1
7.372    1
1.808    1
4.080    1
3.429    1
3.498    1
6.664    1
```

```

4.554      1
5.215      1
5.474      1
4.525      1
5.456      1
5.824      1
3.360      1
3.808      1
3.888      1
4.860      1
1.593      1
5.191      1
3.900      1
7.175      1
4.437      1
4.352      1
4.205      1
3.925      1
4.272      1
4.922      1
6.765      1
4.530      1
4.000      1
2.870      1
3.948      1
2.784      1
5.568      1
Name: alcohol, dtype: int64

```

```
ak.isnull().any()
```

```

total      False
speeding    False
alcohol     False
not_distracted  False
no_previous  False
ins_premium  False
ins_losses  False
abbrev      False
dtype: bool

```

```
ak.isnull().sum()
```

```

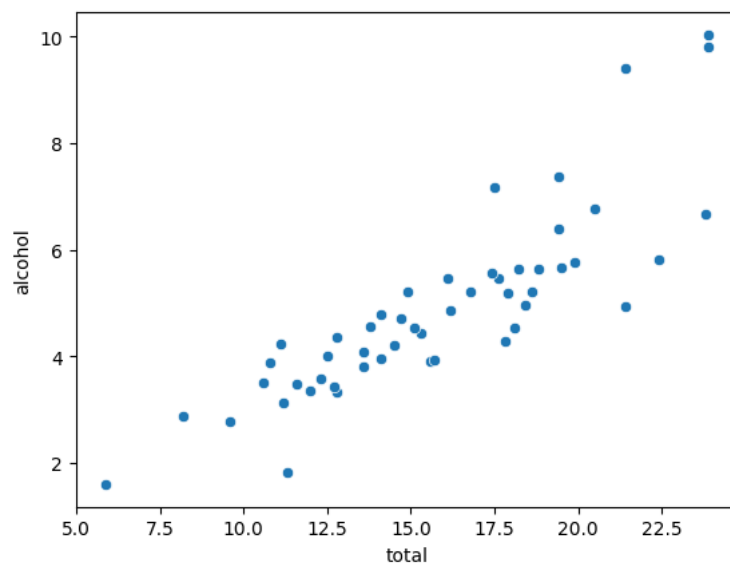
total      0
speeding    0
alcohol     0
not_distracted  0
no_previous  0
ins_premium  0
ins_losses  0
abbrev      0
dtype: int64

```

DATA VISUALIZAION

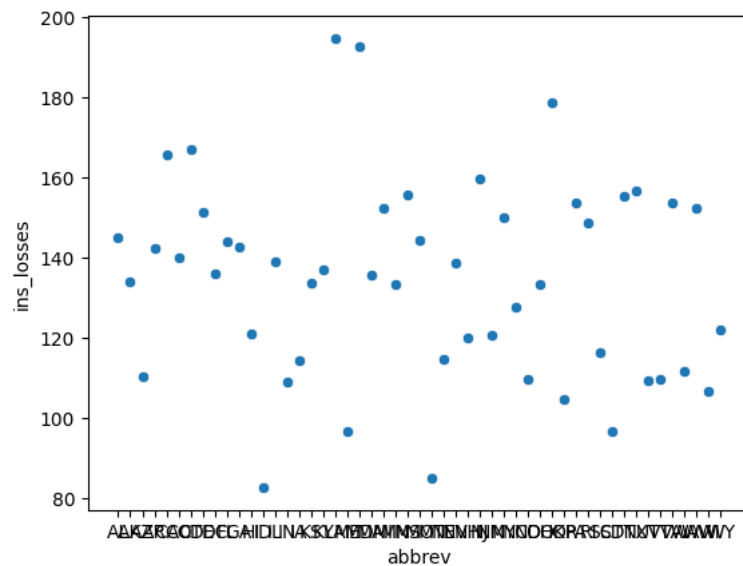
```
sns.scatterplot(x="total",y="alcohol", data=ak)
```

```
<Axes: xlabel='total', ylabel='alcohol'>
```



```
sns.scatterplot(x="abbrev",y="ins_losses",data=ak)
```

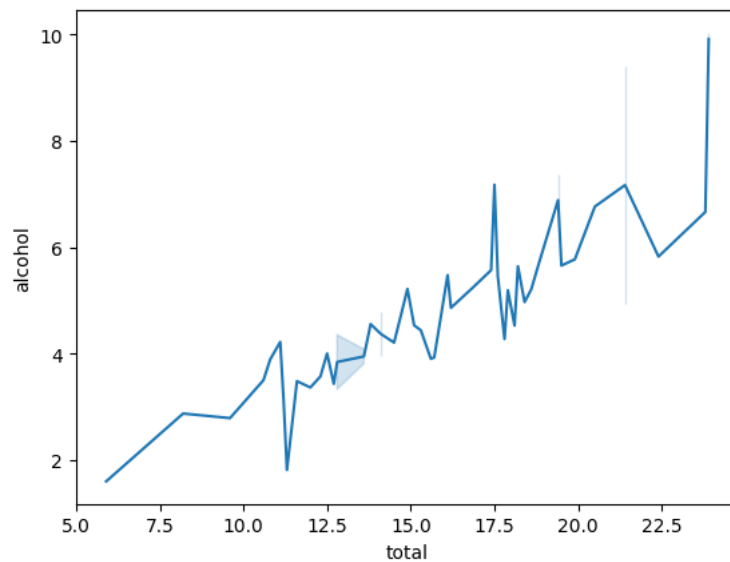
<Axes: xlabel='abbrev', ylabel='ins_losses'>



```
#Lineplot
```

```
sns.lineplot(y="alcohol",x="total",data=ak)
```

<Axes: xlabel='total', ylabel='alcohol'>



```
#Lineplot
```

```
sns.lineplot(y="alcohol",x="total",data=ak,ci=None)
```

```
<ipython-input-21-d491e3384d9c>:2: FutureWarning:
```

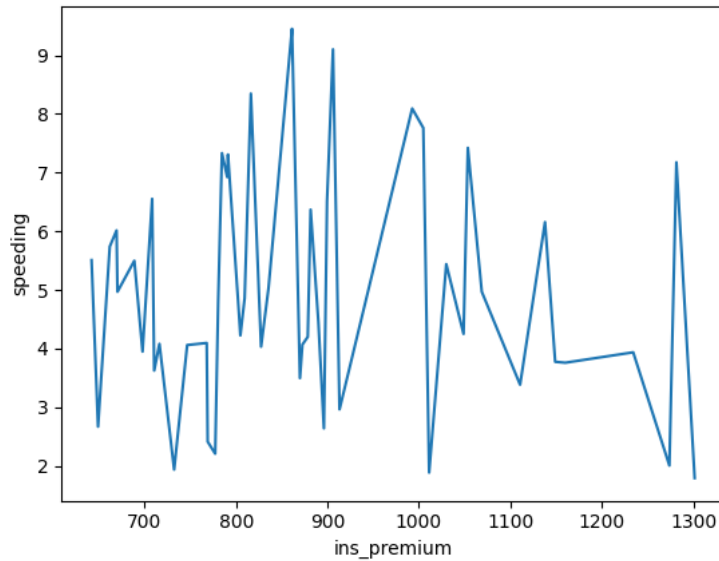
The `ci` parameter is deprecated. Use `errorbar=None` for the same effect.

```
sns.lineplot(y="alcohol",x="total",data=ak,ci=None)
```

```
<Axes: xlabel='total', ylabel='alcohol'>
```

```
sns.lineplot(x="ins_premium",y="speeding",data=ak)
```

```
<Axes: xlabel='ins_premium', ylabel='speeding'>
```



```
#Displot
```

```
sns.distplot(ak["total"])
```

```
<ipython-input-23-18a78abad740>:2: UserWarning:
```

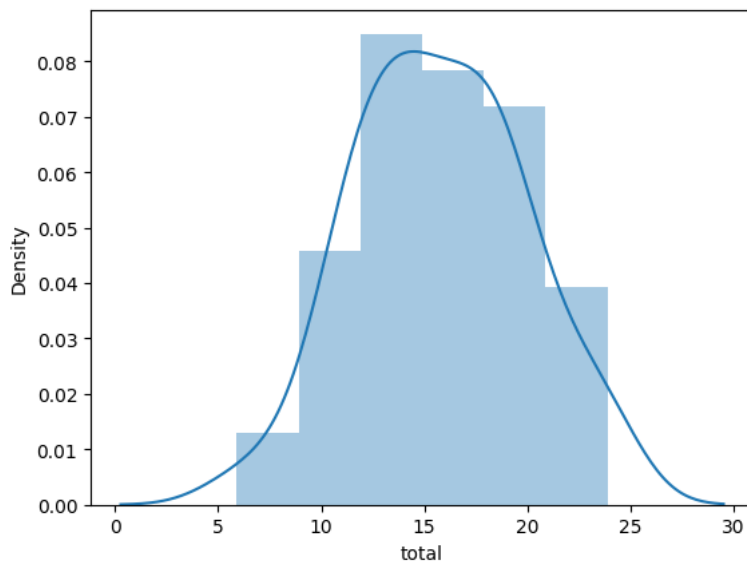
`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see <https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751>

```
sns.distplot(ak["total"])
```

```
<Axes: xlabel='total', ylabel='Density'>
```



```
sns.distplot(ak["not_distracted"])
```

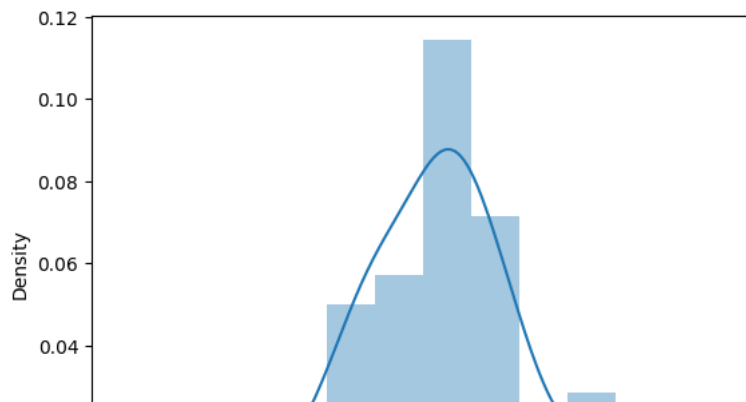
```
<ipython-input-24-bb0c8d7ed882>:1: UserWarning:
```

```
`distplot` is a deprecated function and will be removed in seaborn v0.14.0.
```

Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see <https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751>

```
sns.distplot(ak["not_distracted"])  
<Axes: xlabel='not_distracted', ylabel='Density'>
```



```
#Relationplot  
sns.relplot(x="total", y="ins_losses", data=ak, hue="abbrev")
```



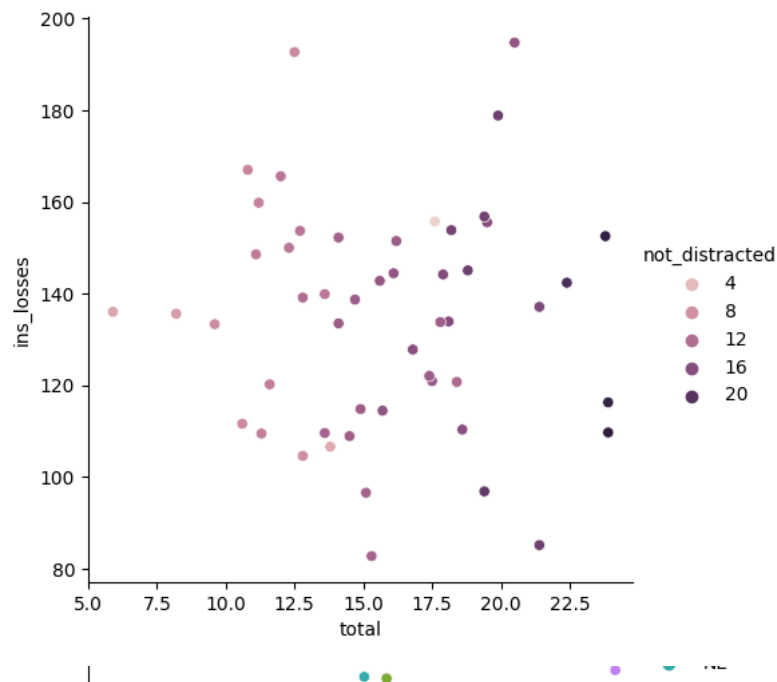
```
<seaborn.axisgrid.FacetGrid at 0x7e3d2e3066b0>
```

```
abbrev
```

```
#Relationplot
```

```
sns.relplot(x="total",y="ins_losses",data=ak,hue="not_distracted")
```

```
<seaborn.axisgrid.FacetGrid at 0x7e3d2e496fe0>
```

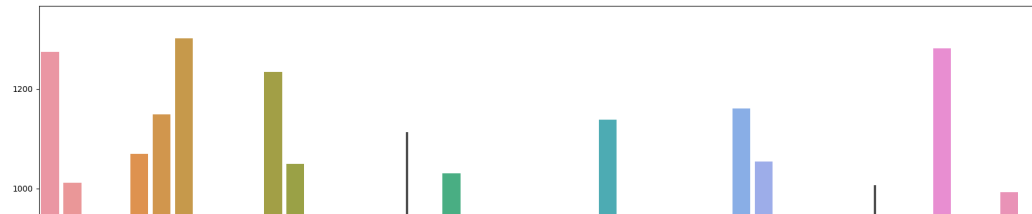


```
#barplot
```

```
plt.subplots(figsize=(22,15))
```

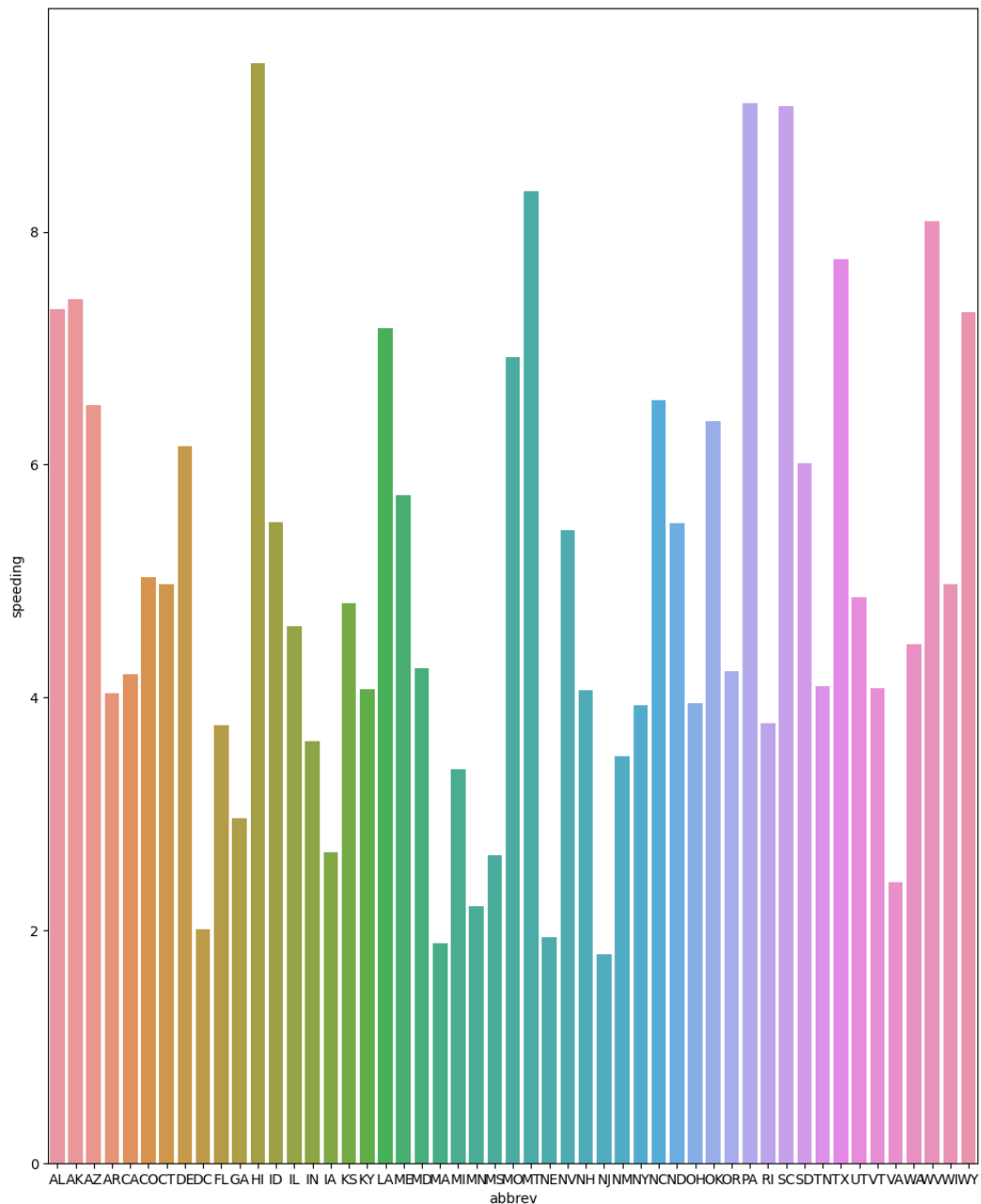
```
sns.barplot(x="total",y="ins_premium",data=ak)
```

<Axes: xlabel='total', ylabel='ins_premium'>



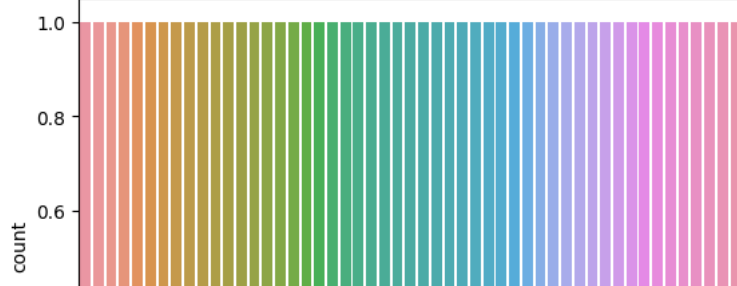
```
#barplot
plt.subplots(figsize=(12,15))
sns.barplot(y="speeding",x="abbrev",data=ak)
```

<Axes: xlabel='abbrev', ylabel='speeding'>



```
#Count Plot
sns.countplot(x="abbrev",data=ak)
```

<Axes: xlabel='abbrev', ylabel='count'>



```
#boxplot
plt.subplots(figsize=(22,15))
sns.boxplot(x="total",y="ins_losses",data=ak)
```

<Axes: xlabel='total', ylabel='ins_losses'>

