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

print(sns.get_dataset_names())
```

```
['anagrams', 'anscombe', 'attention', 'brain_networks', 'car_crashes',
'diamonds', 'dots', 'dowjones', 'exercise', 'flights', 'fmri',
'geyser', 'glue', 'healthexp', 'iris', 'mpg', 'penguins', 'planets',
'seaice', 'taxis', 'tips', 'titanic']
```

```
df=sns.load_dataset('car_crashes')
df.head(10)
```

```
   total  speeding  alcohol  not_distracted  no_previous  ins_premium
\
0   18.8     7.332    5.640          18.048       15.040       784.55

1   18.1     7.421    4.525          16.290       17.014      1053.48

2   18.6     6.510    5.208          15.624       17.856       899.47

3   22.4     4.032    5.824          21.056       21.280       827.34

4   12.0     4.200    3.360          10.920       10.680       878.41

5   13.6     5.032    3.808          10.744       12.920       835.50

6   10.8     4.968    3.888           9.396        8.856      1068.73

7   16.2     6.156    4.860          14.094       16.038      1137.87

8    5.9     2.006    1.593           5.900        5.900      1273.89

9   17.9     3.759    5.191          16.468       16.826      1160.13


   ins_losses abbrev
0      145.08     AL
1      133.93     AK
2      110.35     AZ
3      142.39     AR
4      165.63     CA
5      139.91     CO
6      167.02     CT
7      151.48     DE
8      136.05     DC
9      144.18     FL
```

```
df.shape
```

```
(51, 8)
```

```
df.isnull().sum()
```

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

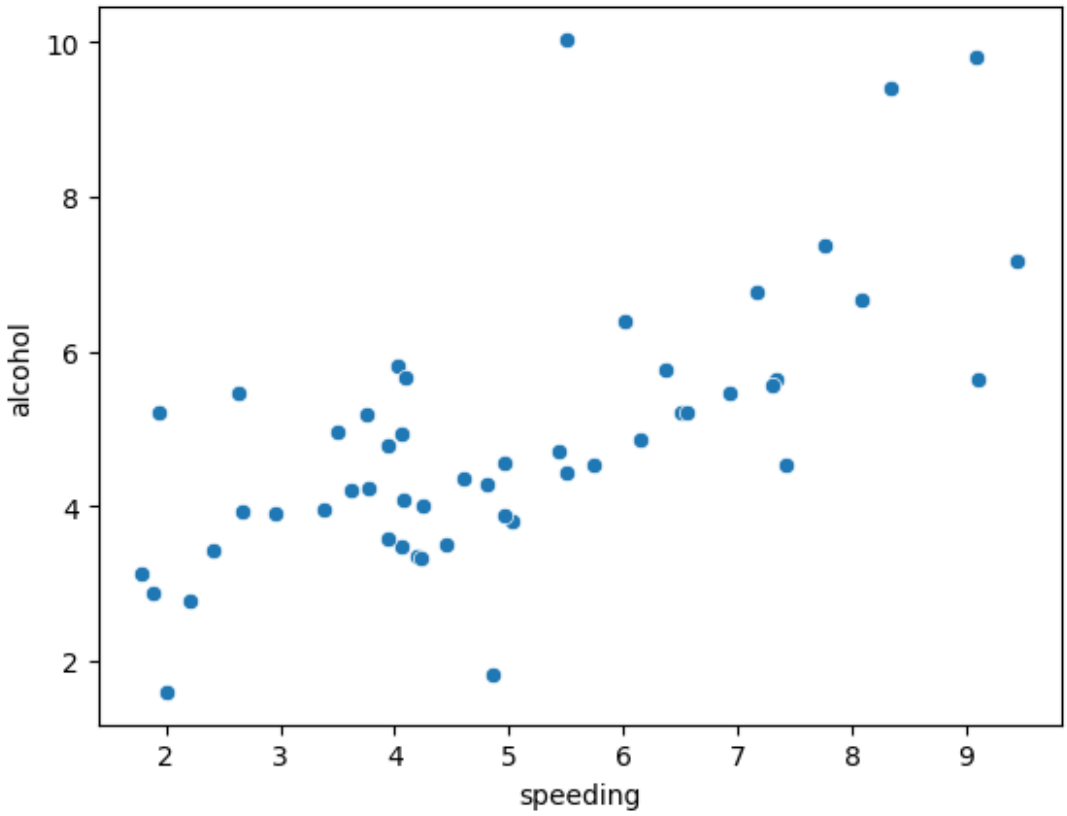
```
df.describe()
```

```
           total   speeding    alcohol  not_distracted  no_previous  \
count  51.000000  51.000000  51.000000       51.000000    51.000000
mean   15.790196   4.998196   4.886784       13.573176    14.004882
std     4.122002   2.017747   1.729133        4.508977     3.764672
min     5.900000   1.792000   1.593000        1.760000     5.900000
25%    12.750000   3.766500   3.894000       10.478000    11.348000
50%    15.600000   4.608000   4.554000       13.857000    13.775000
75%    18.500000   6.439000   5.604000       16.140000    16.755000
max    23.900000   9.450000  10.038000       23.661000    21.280000

       ins_premium  ins_losses
count    51.000000   51.000000
mean    886.957647  134.493137
std     178.296285   24.835922
min     641.960000   82.750000
25%     768.430000  114.645000
50%     858.970000  136.050000
75%    1007.945000  151.870000
max    1301.520000  194.780000
```

```
sns.scatterplot(x="speeding",y="alcohol",data=df)
```

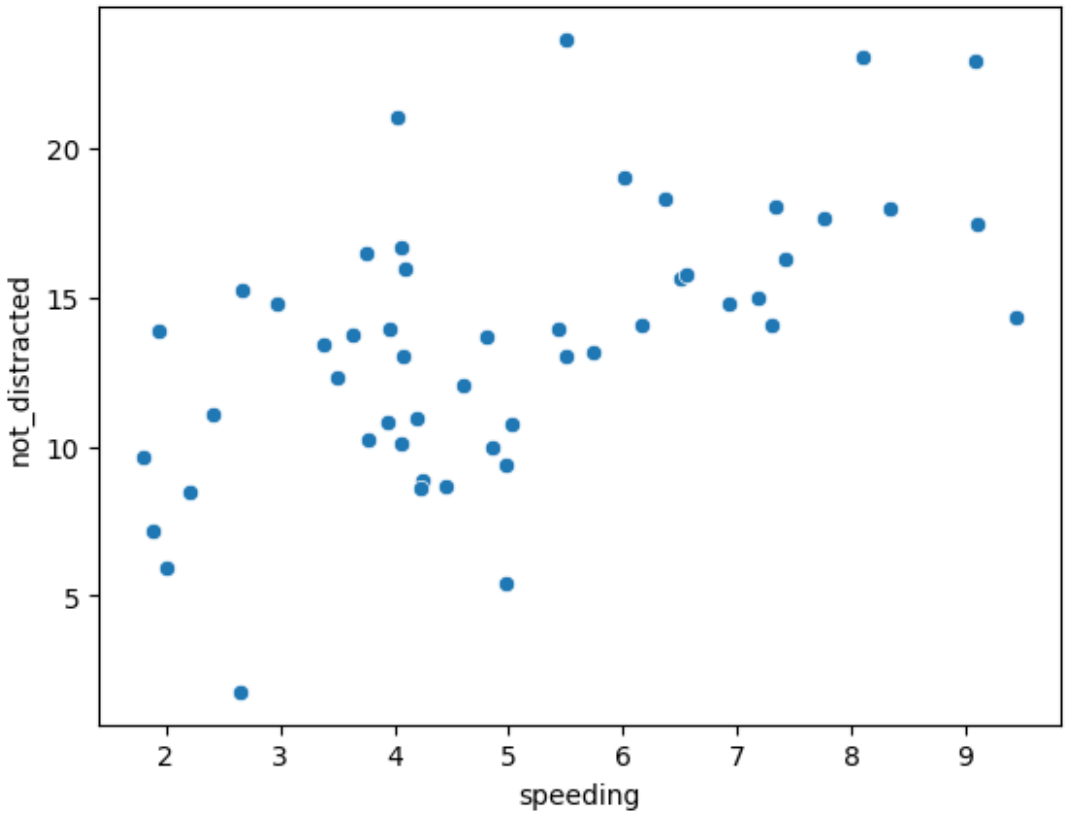
```
<Axes: xlabel='speeding', ylabel='alcohol'>
```



inference: more alcohol, more speed

```
sns.scatterplot(x="speeding",y="not_distracted",data=df)
```

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

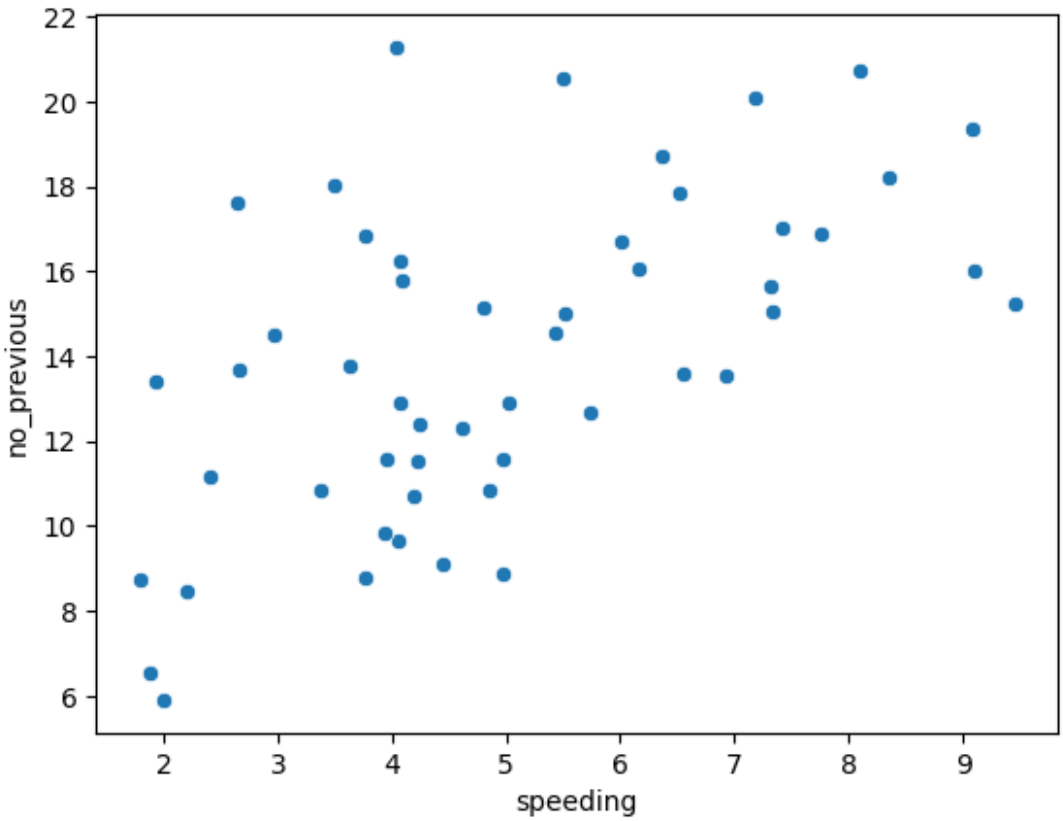


inference: more people are not distracted at lower speed.

sns.scatterplot(x="speeding",y="not_distracted",data=df)

```
sns.scatterplot(x="speeding",y="no_previous",data=df)
```

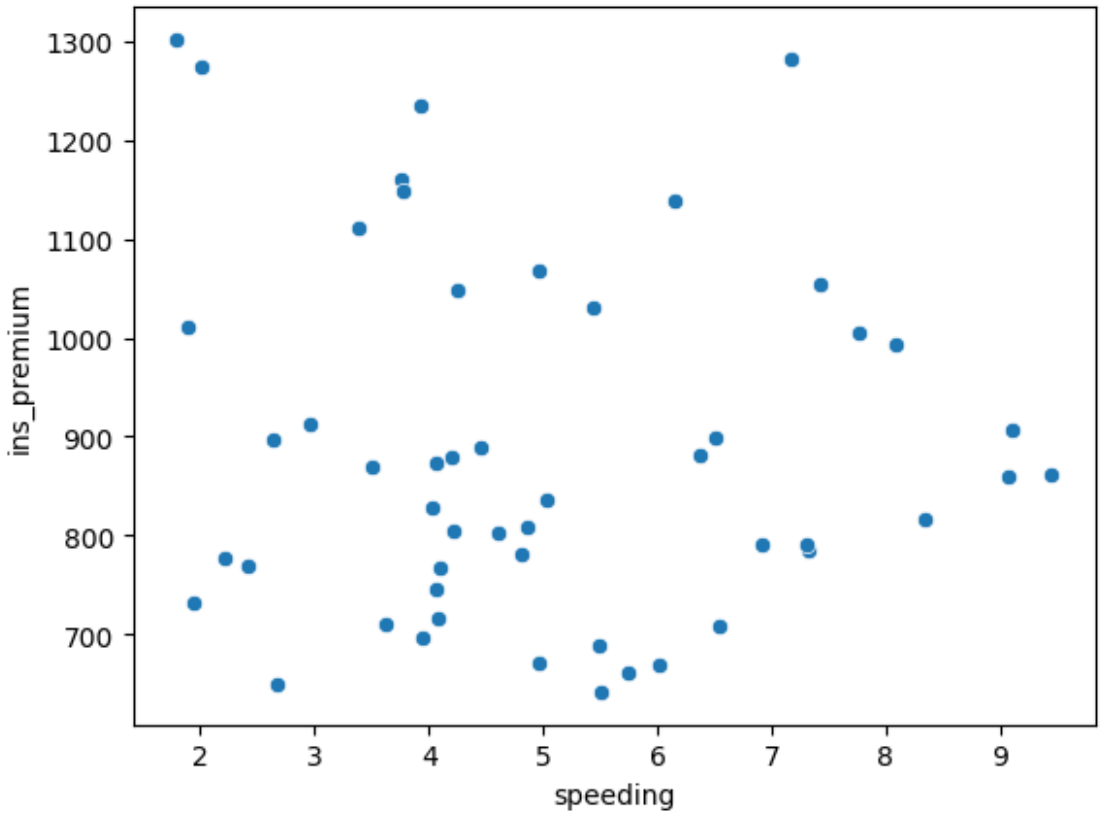
```
<Axes: xlabel='speeding', ylabel='no_previous'>
```



no particular conclusion

```
sns.scatterplot(x="speeding",y="ins_premium",data=df)
```

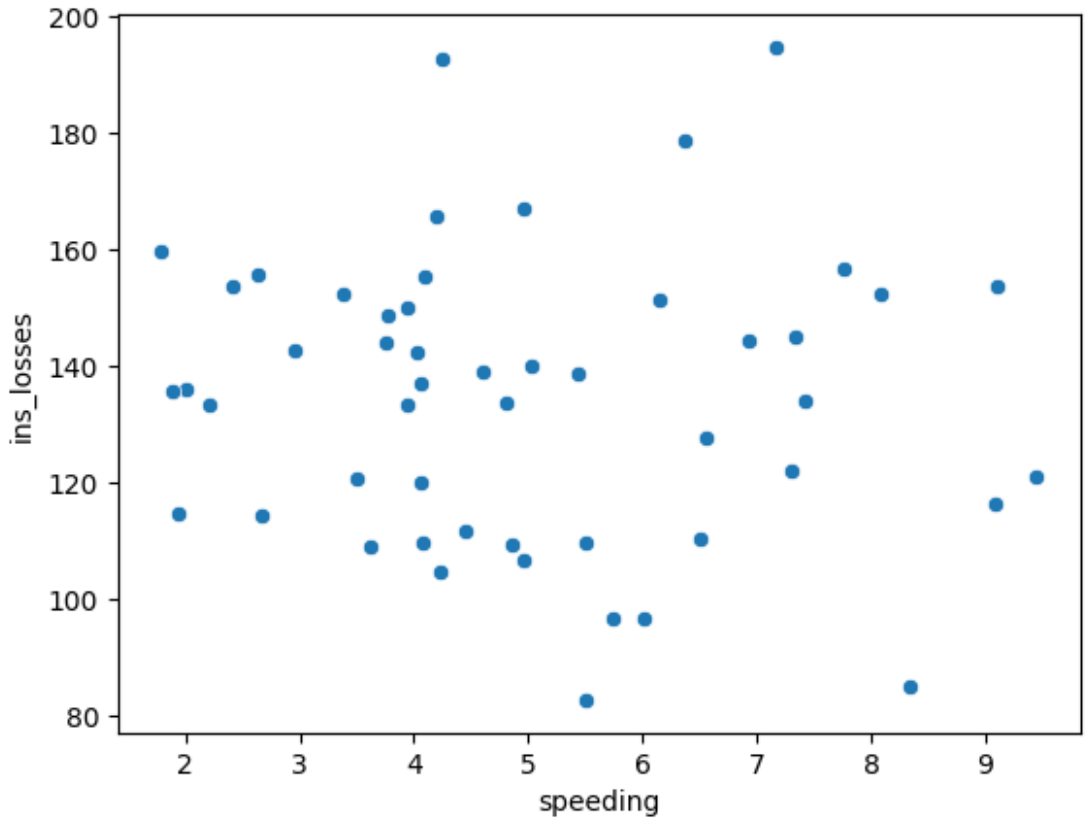
```
<Axes: xlabel='speeding', ylabel='ins_premium'>
```



no particular conclusion

```
sns.scatterplot(x="speeding",y="ins_losses",data=df)
```

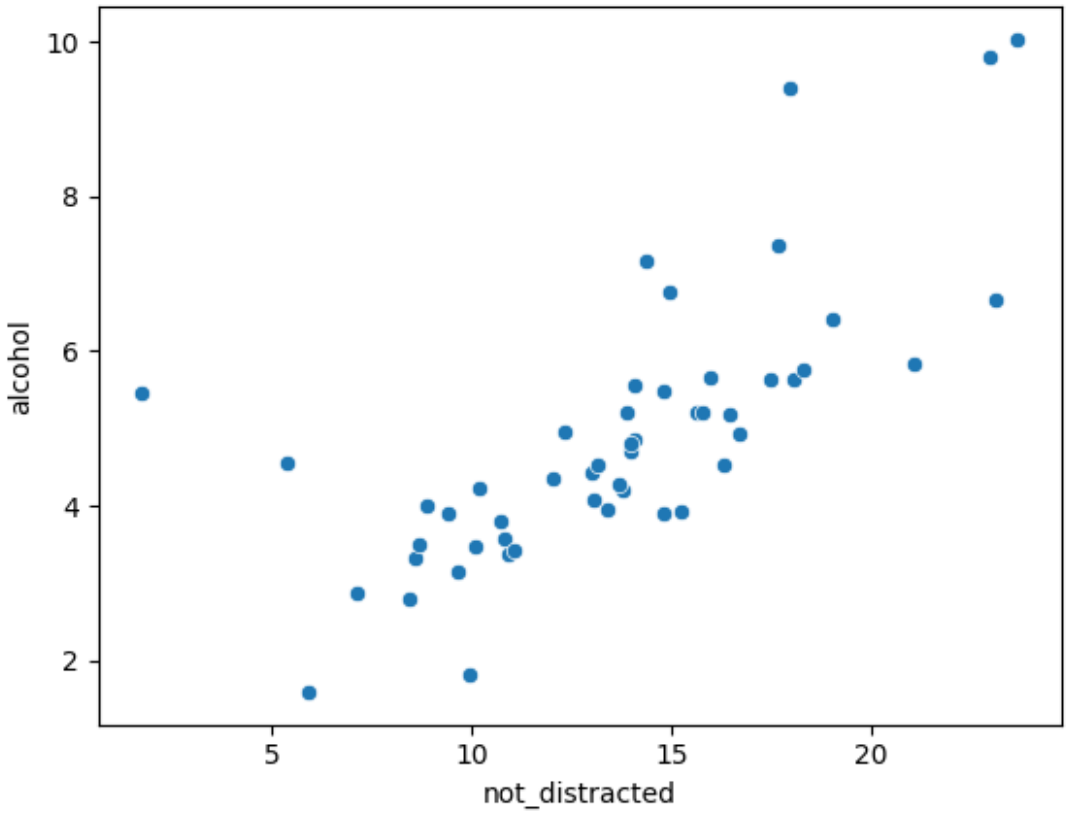
```
<Axes: xlabel='speeding', ylabel='ins_losses'>
```



no particular inference

```
sns.scatterplot(y="alcohol",x="not_distracted",data=df)
```

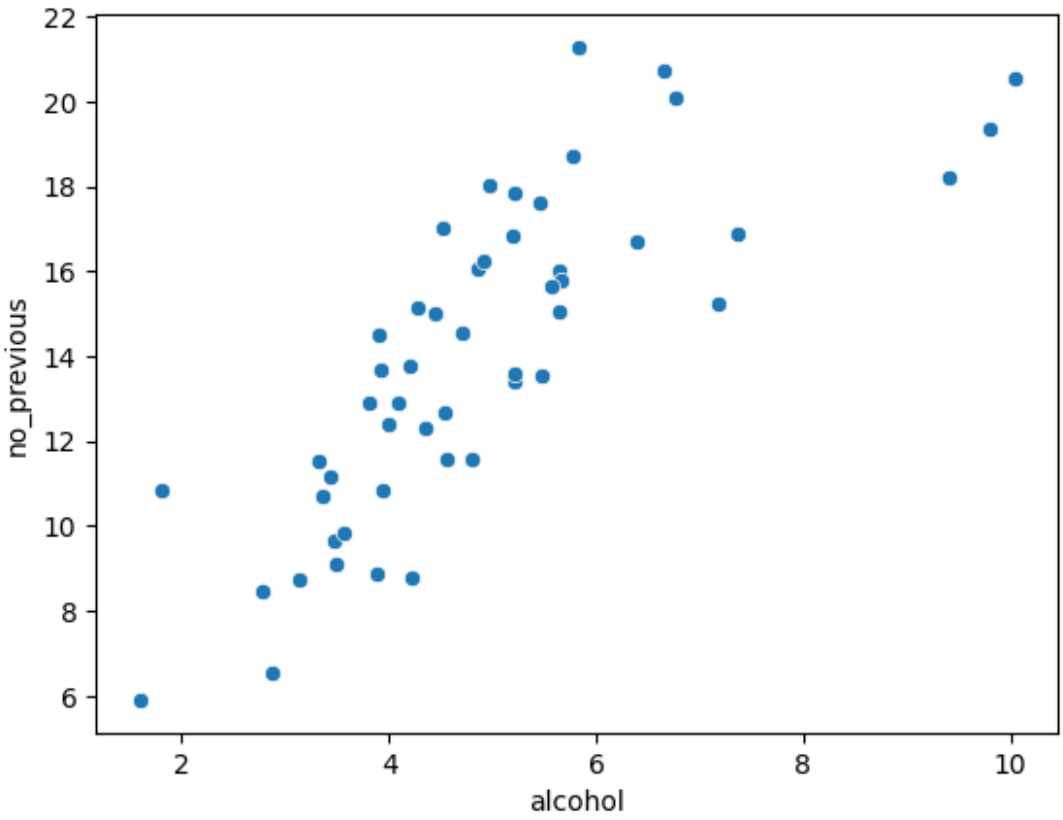
```
<Axes: xlabel='not_distracted', ylabel='alcohol'>
```



inference: less people are not distracted the more they drink alcohol. More alcohol you drink, the more distracted you get.

```
sns.scatterplot(x="alcohol",y="no_previous",data=df)
```

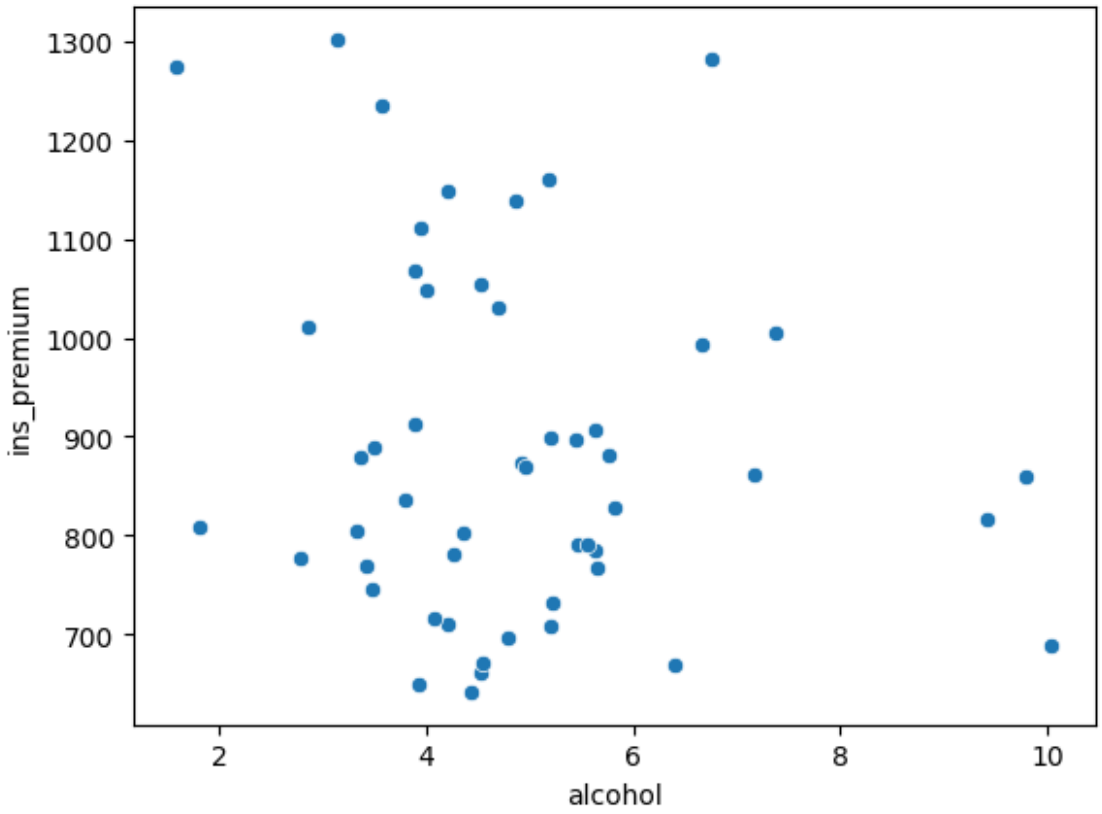
```
<Axes: xlabel='alcohol', ylabel='no_previous'>
```



more alcohol more previous crashes.

```
sns.scatterplot(x="alcohol",y="ins_premium",data=df)
```

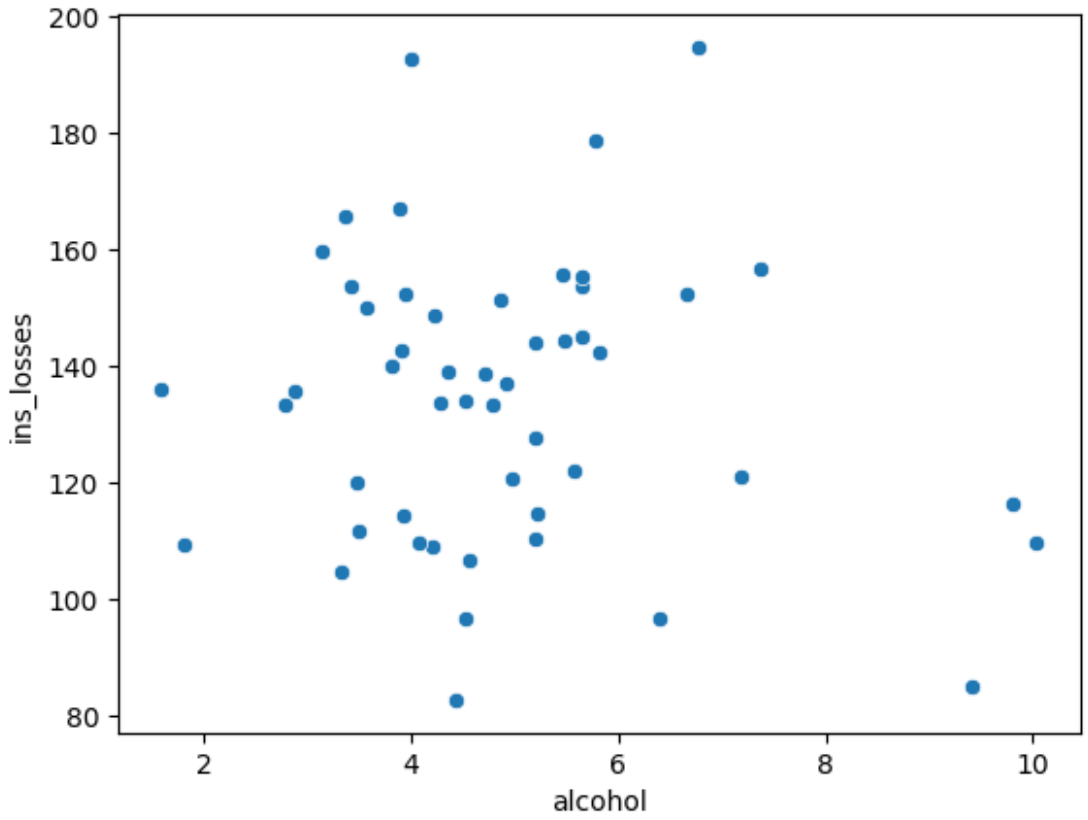
```
<Axes: xlabel='alcohol', ylabel='ins_premium'>
```



no particular conclusion

```
sns.scatterplot(x="alcohol",y="ins_losses",data=df)
```

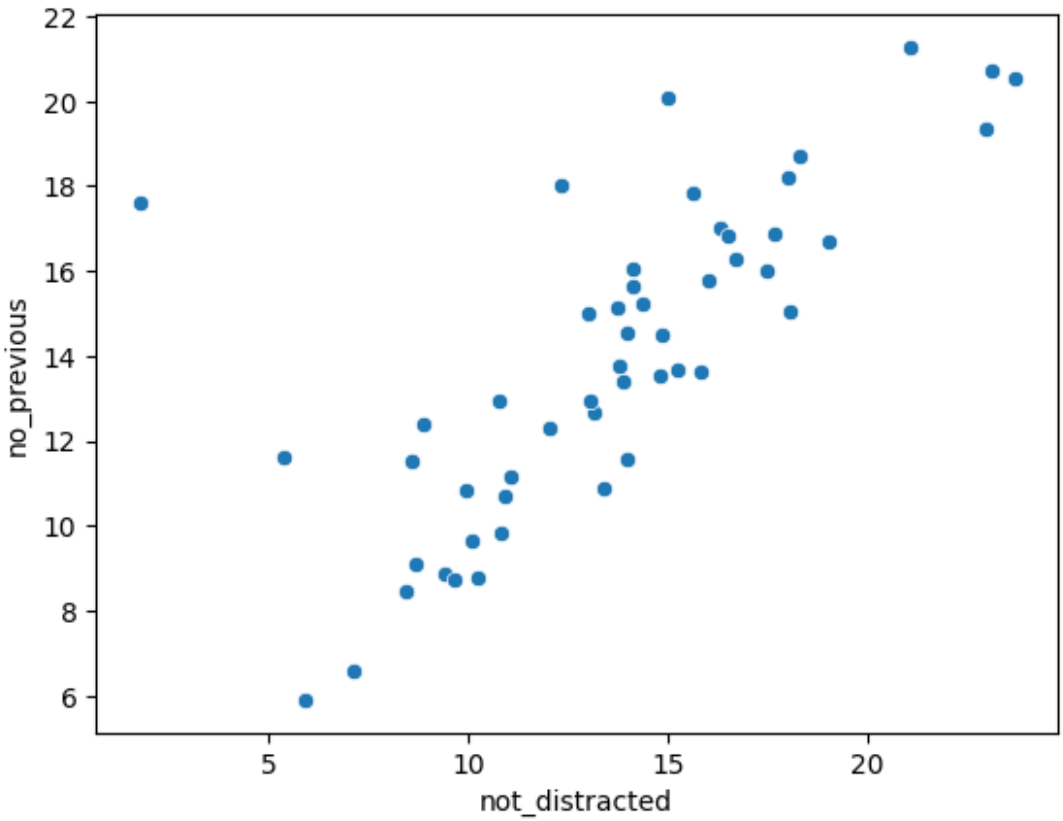
```
<Axes: xlabel='alcohol', ylabel='ins_losses'>
```



no particular conclusion

```
sns.scatterplot(x="not_distracted",y="no_previous",data=df)
```

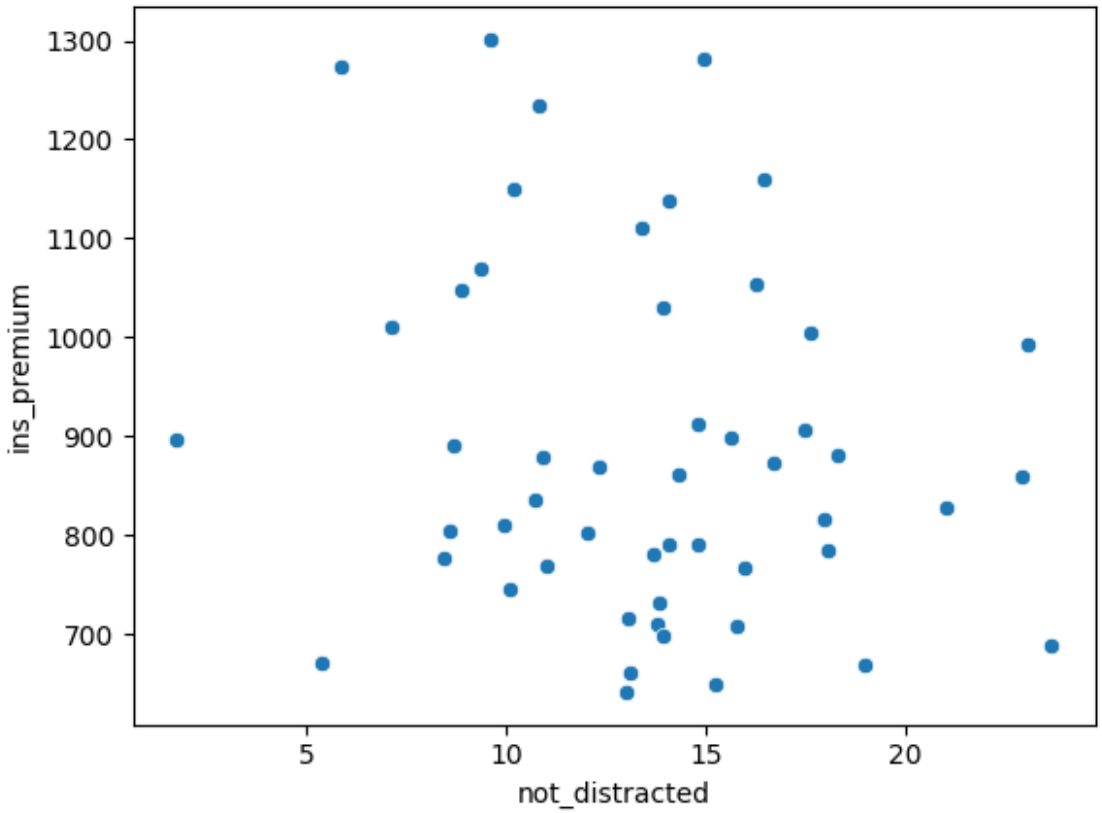
```
<Axes: xlabel='not_distracted', ylabel='no_previous'>
```



inference: not distracted is proportional to no previous accidents. Less distracted you are, the less are the chances of you having previous accidents

```
sns.scatterplot(x="not_distracted",y="ins_premium",data=df)
```

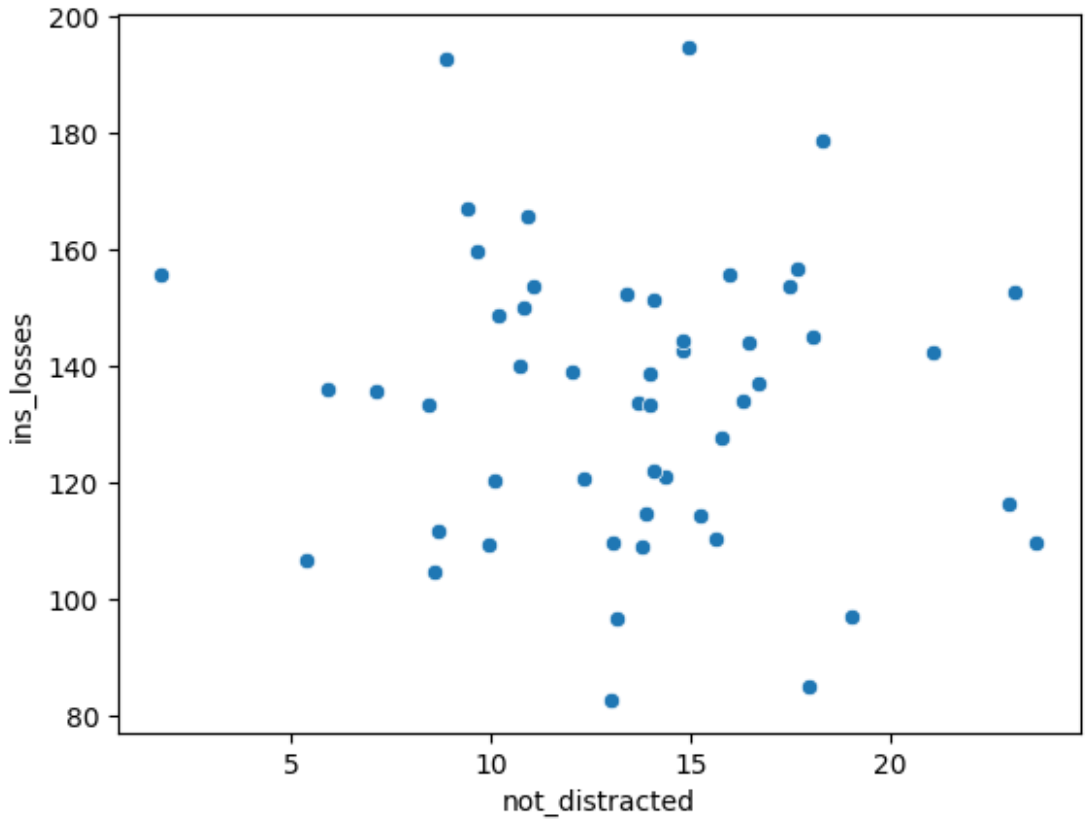
```
<Axes: xlabel='not_distracted', ylabel='ins_premium'>
```



no particular conclusion

```
sns.scatterplot(x="not_distracted",y="ins_losses",data=df)
```

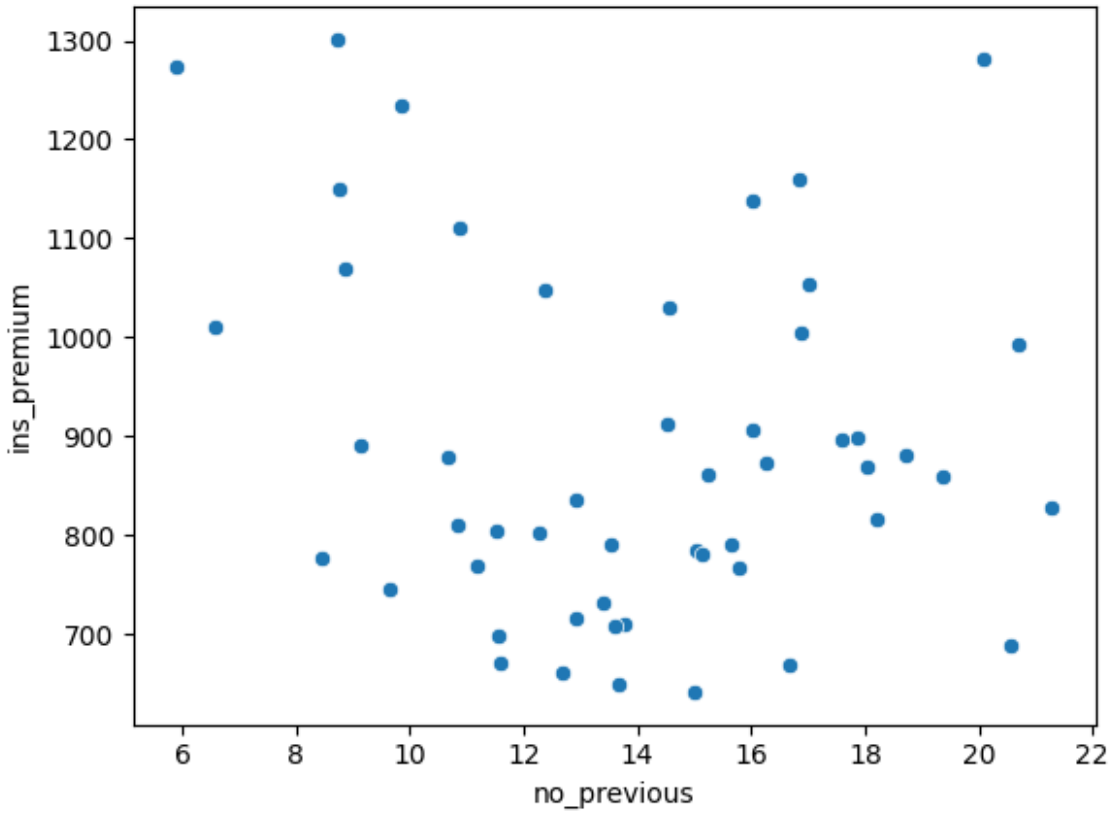
```
<Axes: xlabel='not_distracted', ylabel='ins_losses'>
```



no particular conclusion

```
sns.scatterplot(x="no_previous",y="ins_premium",data=df)
```

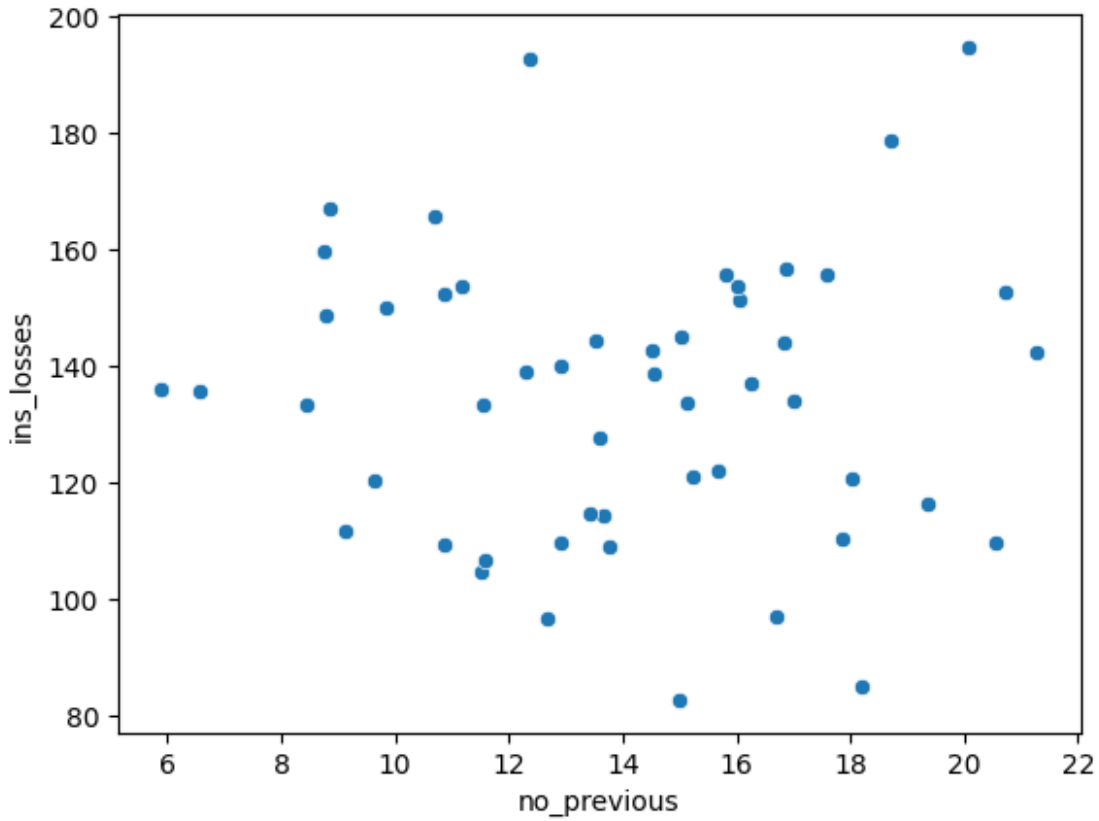
```
<Axes: xlabel='no_previous', ylabel='ins_premium'>
```



no particular conclusion

```
sns.scatterplot(x="no_previous",y="ins_losses",data=df)
```

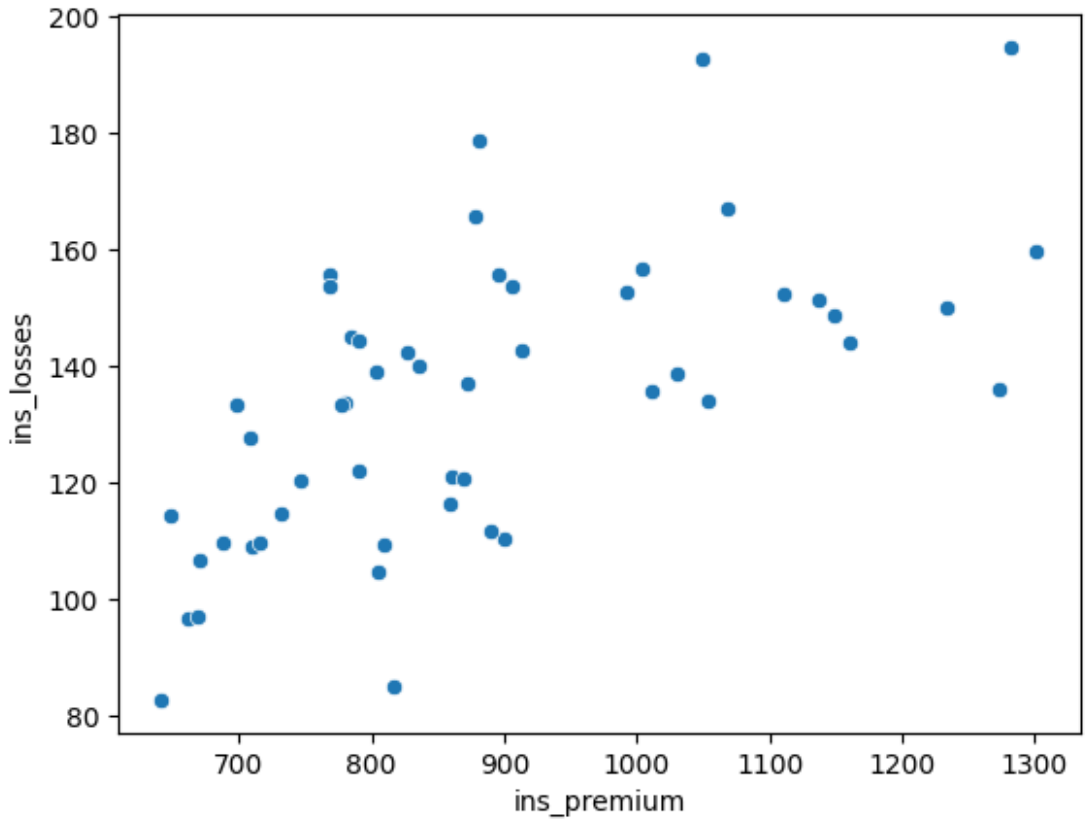
```
<Axes: xlabel='no_previous', ylabel='ins_losses'>
```



no particular conclusion

```
sns.scatterplot(x="ins_premium",y="ins_losses",data=df)
```

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



no particular conclusion