```python
In [1]:  # Data Imports
         import numpy as np
         import pandas as pd
         from pandas import Series,DataFrame

         # Math
         import math

         # Plot imports
         import matplotlib.pyplot as plt
         import seaborn as sns
         sns.set_style('whitegrid')
         %matplotlib inline
```

```python
In [2]:  clf = pd.read_excel("CLF Test.xlsx")
```

```python
In [3]:  clf.head()
```

Out[3]:

	Date	State	Sex	Age	Survey Result	Main Driver	Claim Type	Handling Type	Peril	Claim Number	Adju
0	2020-09-01	MA	M	46	1	Professional	Load Balance	Regular	Water	1.234568e+08	Les
1	2020-09-01	PA	F	70	1	Service	Load Balance	Regular	Lightning	1.234568e+08	Mar
2	2020-09-01	TX	M	36	0	Low Payment	Load Balance	Regular	Water	9.012346e+08	S
3	2020-09-01	IL	F	81	1	No Comment	Virtual	Virtual	Water	1.160494e+09	Ste
4	2020-09-02	VA	M	39	1	Service	Load Balance	Regular	Water	1.549383e+09	(

```python
In [4]:  clf['Survey Result'].count()
```

Out[4]: 76

```python
In [5]:  clf[['Survey Result','Adjuster']].mean()
```

Out[5]: Survey Result   0.697368
        dtype: float64

```python
In [6]:  clf['Survey Result'].value_counts()
```
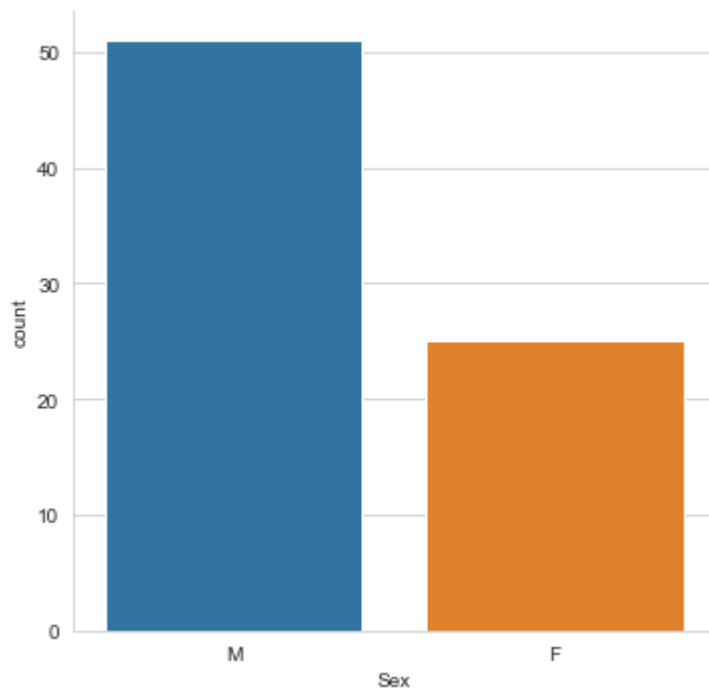
Out[6]: 1   53
        0   23
        Name: Survey Result, dtype: int64
```

```
In [7]: clf.columns
```

```
Out[7]: Index(['Date', 'State', 'Sex', 'Age', 'Survey Result', 'Main Driver',
               'Claim Type', 'Handling Type', 'Peril', 'Claim Number', 'Adjuster',
               'Survey Type', 'Payment Type', 'Days to Close', 'Coverage'],
              dtype='object')
```

```
In [8]: clf['Survey Result'].mean()
```

```
Out[8]: 0.6973684210526315
```

```
In [9]: clf[['Survey Result','Adjuster']]
```

Out[9]:

	Survey Result	Adjuster
0	1	Lessner
1	1	Martinez
2	0	Skeet
3	1	Stevens
4	1	Cook
...	...	...
71	1	Cook
72	0	Ortbal
73	0	Cook
74	1	Cook
75	1	Lessner

76 rows × 2 columns

```
In [10]: clf.describe()
```

Out[10]:

	Age	Survey Result	Claim Number	Days to Close
count	76.000000	76.000000	7.600000e+01	76.000000
mean	52.394737	0.697368	1.457716e+10	9.500000
std	16.169996	0.462450	8.587980e+09	14.566171
min	23.000000	0.000000	1.234568e+08	1.000000
25%	38.750000	0.000000	7.285494e+09	3.000000
50%	52.500000	1.000000	1.457716e+10	6.000000
75%	68.000000	1.000000	2.186883e+10	9.000000
max	81.000000	1.000000	2.916049e+10	114.000000

In [11]: `sns.catplot('Sex',data=clf,kind='count')`

Out[11]: `<seaborn.axisgrid.FacetGrid at 0x2813cab25c8>`

In [12]: `sns.catplot('Sex',data=clf,kind='count',hue='Survey Result')`

Out[12]: `<seaborn.axisgrid.FacetGrid at 0x2813cad8b08>`

In [13]: sns.catplot('Age',data=clf,kind='count',hue='Survey Result', aspect=2)

Out[13]: <seaborn.axisgrid.FacetGrid at 0x2813d2f9088>

In [14]: sns.catplot('Days to Close',data=clf,kind='count',hue='Survey Result')

Out[14]: <seaborn.axisgrid.FacetGrid at 0x2813d384748>

```
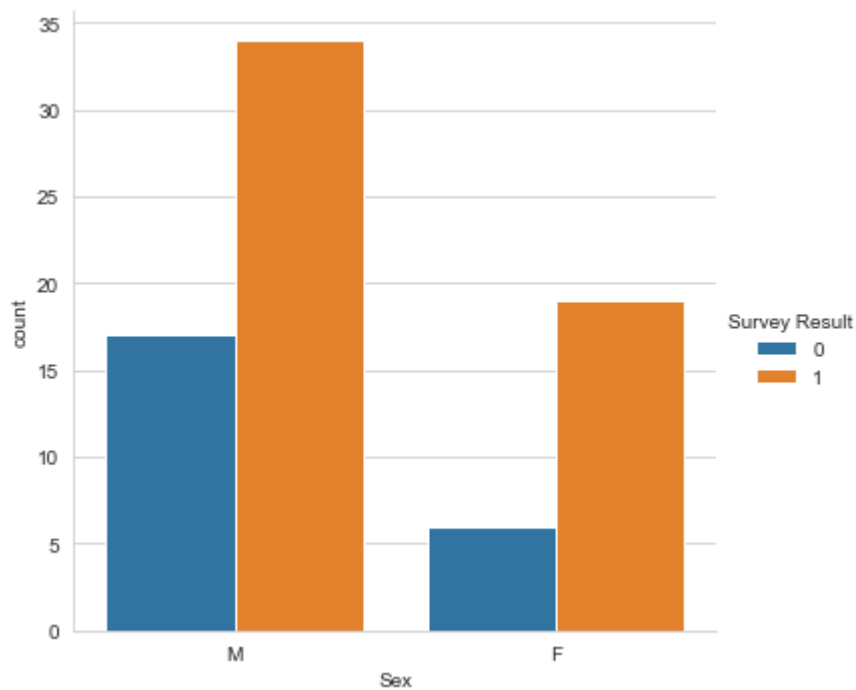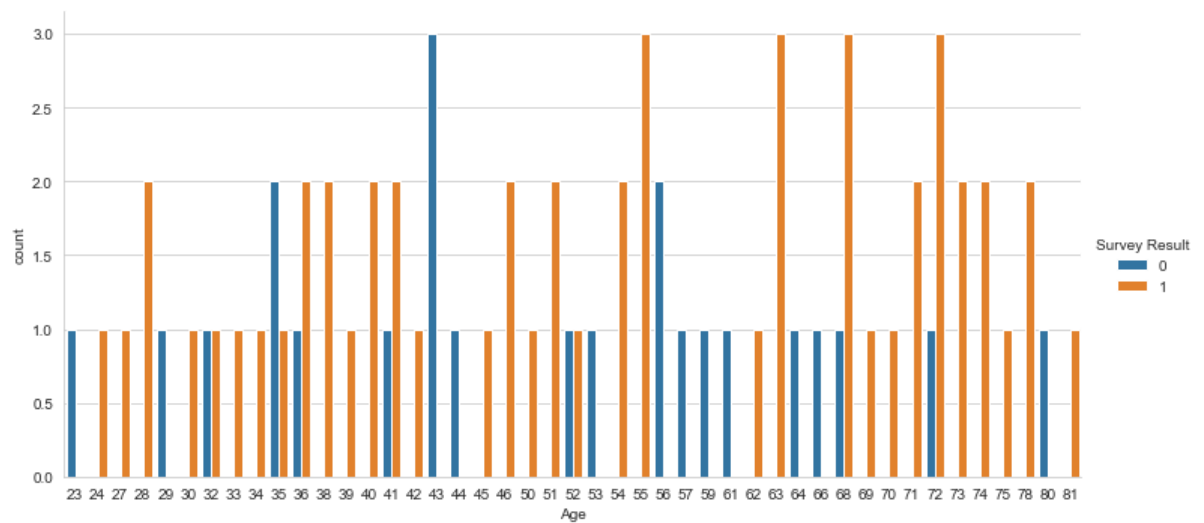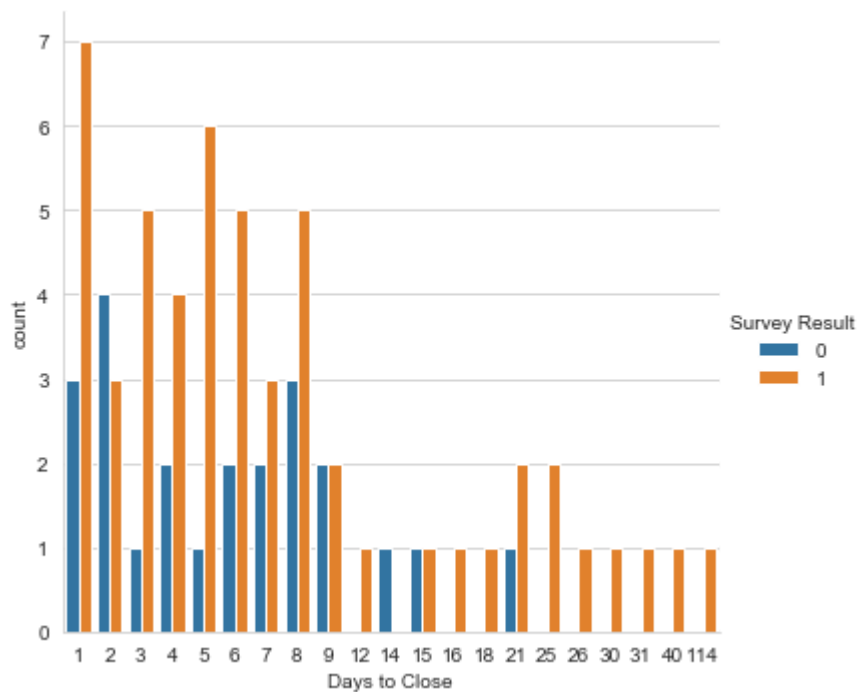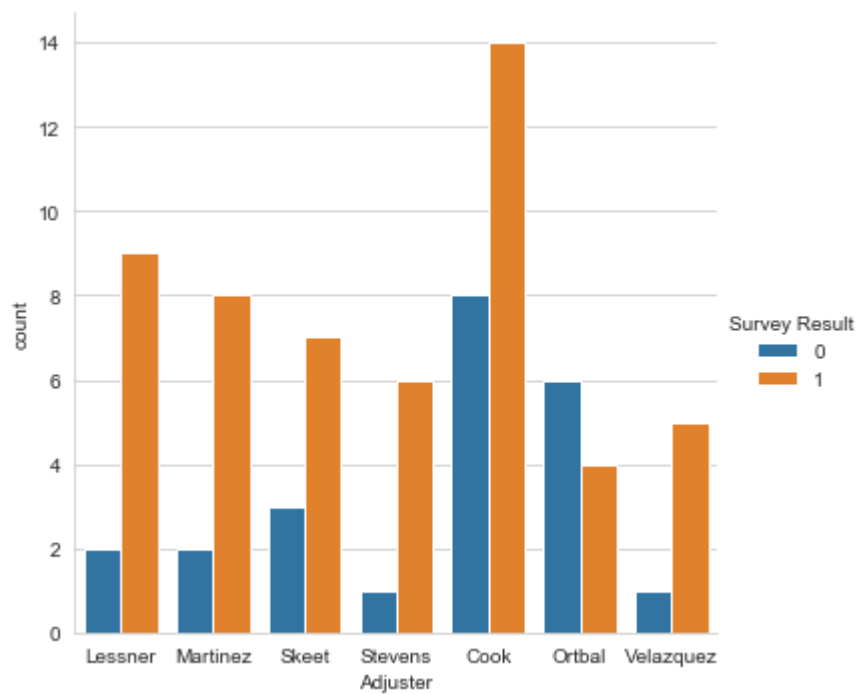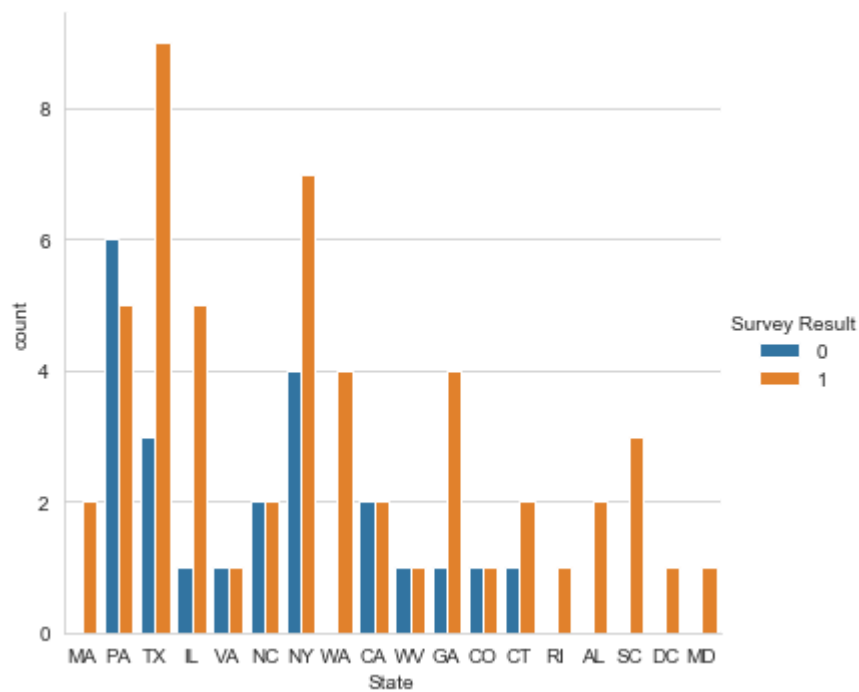In [15]: sns.catplot('Adjuster',data=clf,kind='count',hue='Survey Result')
```

Out[15]: <seaborn.axisgrid.FacetGrid at 0x2813d2f09c8>


```
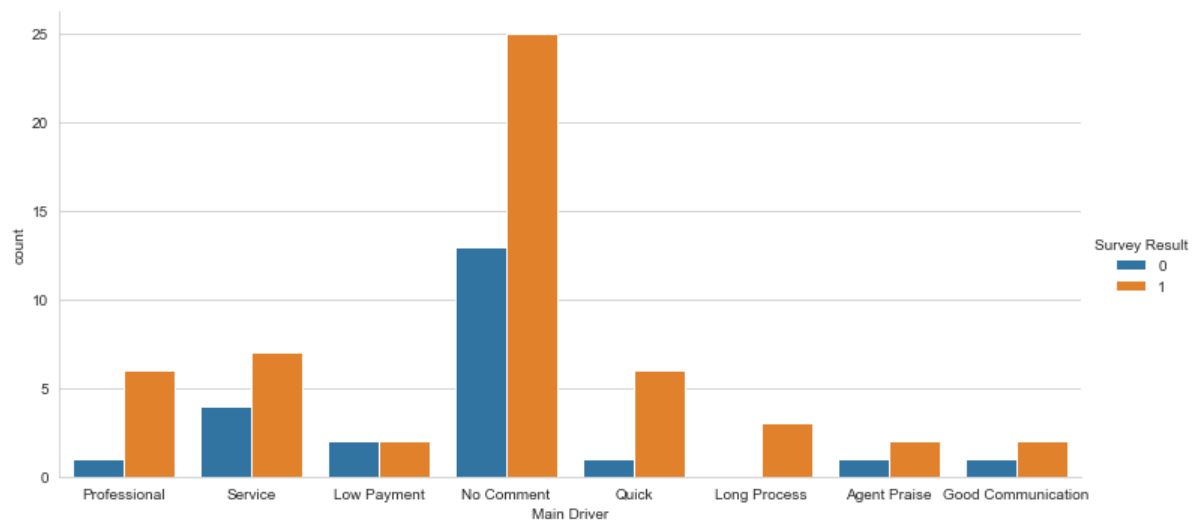In [16]: sns.catplot('State',data=clf,kind='count',hue='Survey Result')
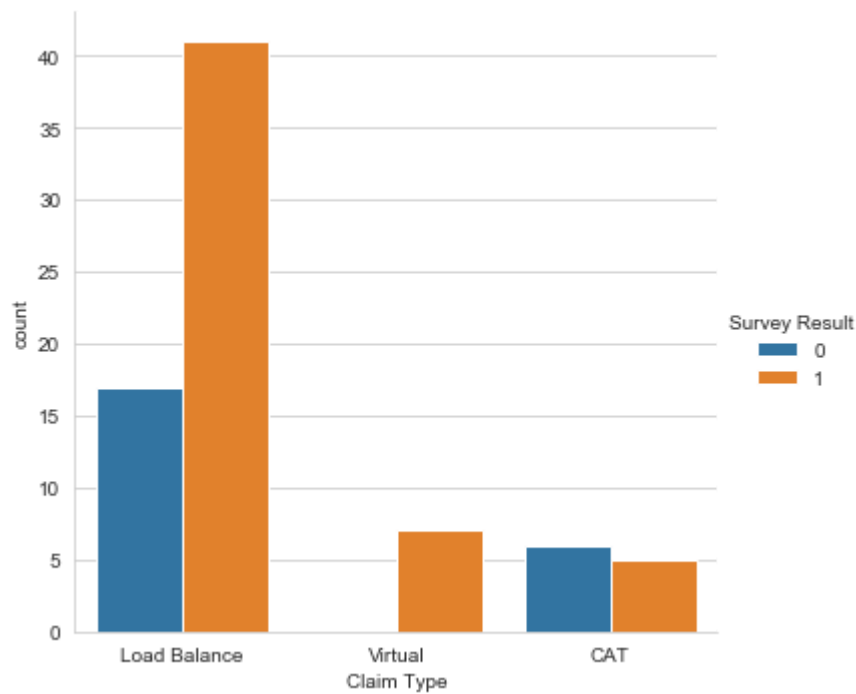```

Out[16]: <seaborn.axisgrid.FacetGrid at 0x2813d559cc8>

In [17]: sns.catplot('Main Driver',data=clf,kind='count',hue='Survey Result', aspect=2)

Out[17]: <seaborn.axisgrid.FacetGrid at 0x2813d84ea48>



In [18]: sns.catplot('Claim Type',data=clf,kind='count',hue='Survey Result')

Out[18]: <seaborn.axisgrid.FacetGrid at 0x2813d824788>

```
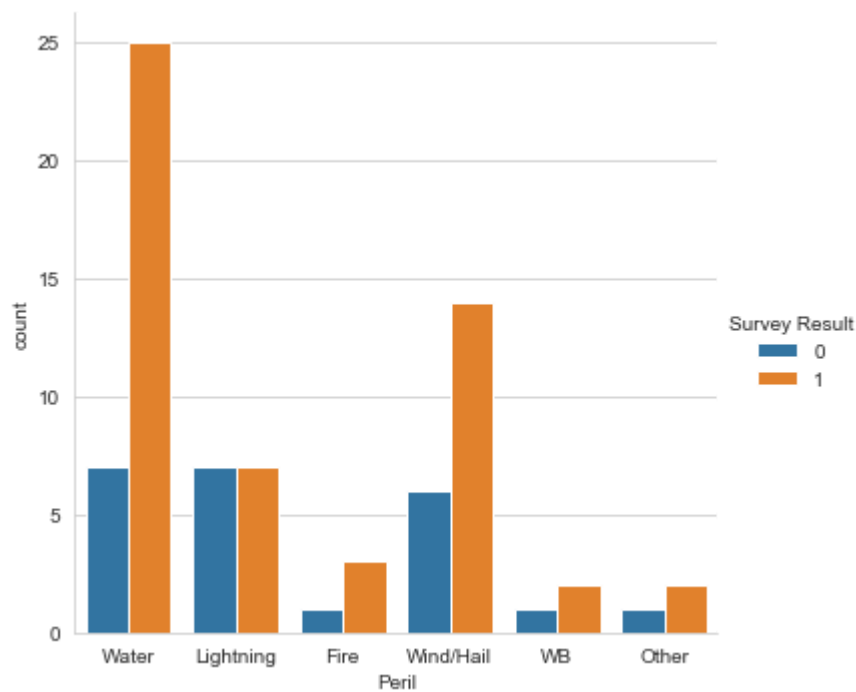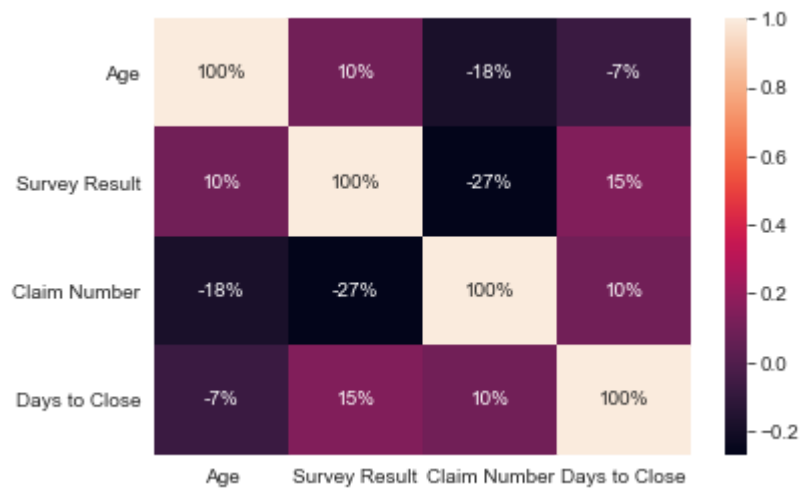In [19]: sns.catplot('Peril',data=clf,kind='count',hue='Survey Result')
```

Out[19]: <seaborn.axisgrid.FacetGrid at 0x2813d96cc08>


```
In [20]: sns.heatmap(clf.corr(),annot=True, fmt='.0%')
```

Out[20]: <matplotlib.axes._subplots.AxesSubplot at 0x2813da07748>


```
In [21]: clf.corr()
```

Out[21]:

	Age	Survey Result	Claim Number	Days to Close
Age	1.000000	0.101775	-0.180733	-0.070534
Survey Result	0.101775	1.000000	-0.268735	0.145484
Claim Number	-0.180733	-0.268735	1.000000	0.098845
Days to Close	-0.070534	0.145484	0.098845	1.000000

```
In [22]: pd.get_dummies(clf['Sex']).head()
```

Out[22]:

	F	M
0	0	1
1	1	0
2	0	1
3	1	0
4	0	1

```
In [ ]:
```