In [1]:
```
import pandas as pd
import numpy as np
from matplotlib import pyplot as plt
```

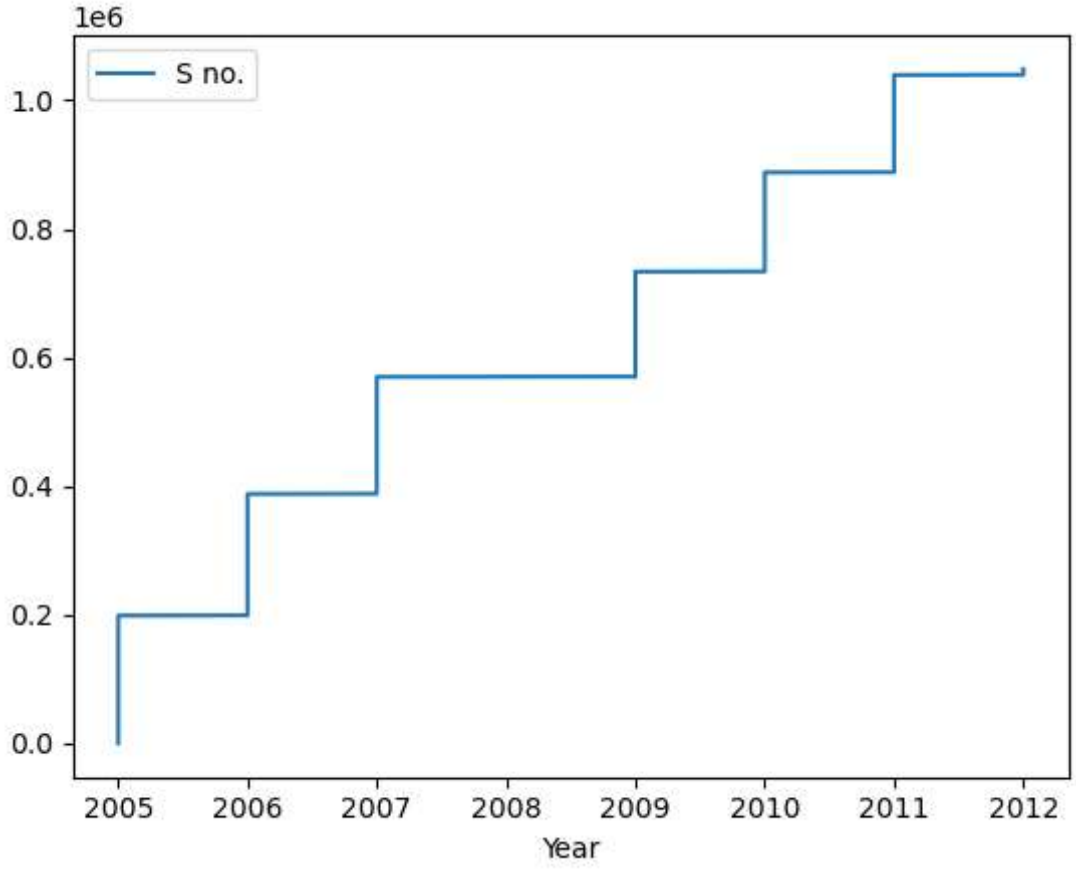
In [2]:
```
df=pd.read_csv('./Accident-filtered.csv')
df.head()
```

Out[2]:

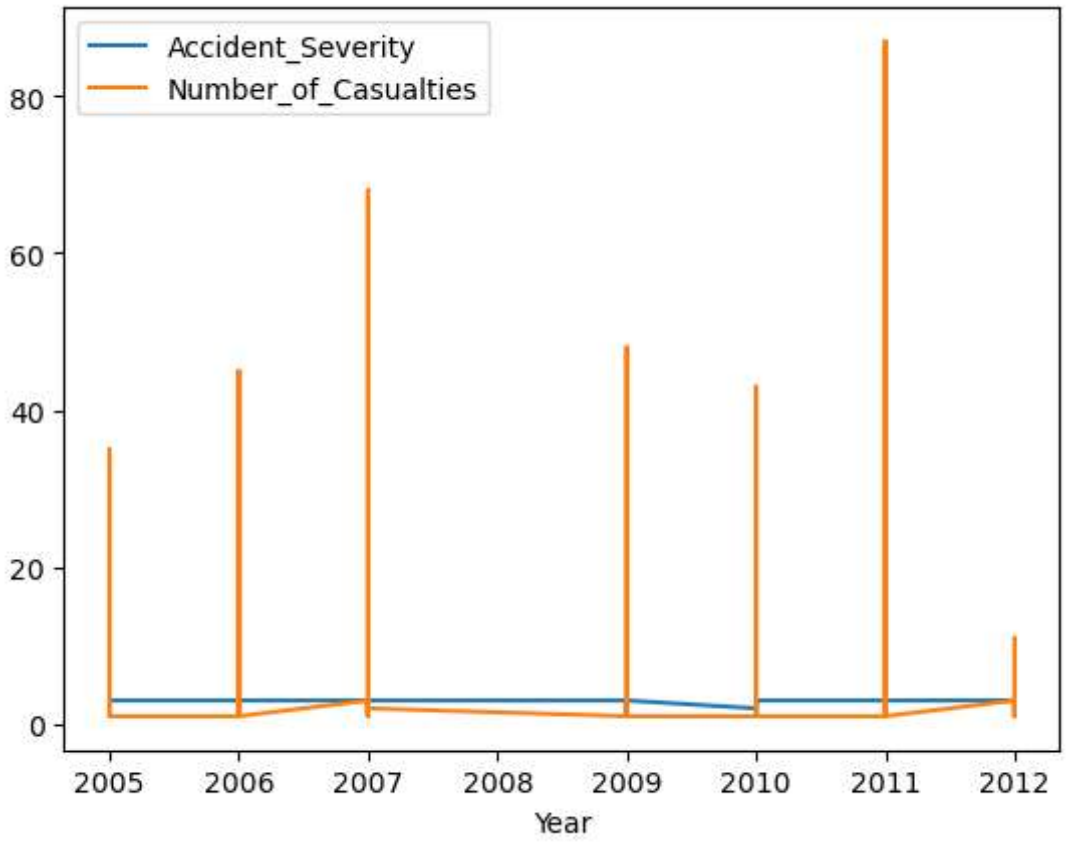
	S no.	Accident_Index	Location_Easting_OSGR	Location_Northing_OSGR	Longitude	Latitude	P
0	1	2.0000000E+12	342200.0	850290	-2.967122	57.539108	
1	2	2.0000000E+12	529710.0	277320	-0.095826	52.378530	
2	3	2.0000000E+12	565170.0	172960	0.374698	51.431513	
3	4	200501YE80011	535230.0	192430	-0.048247	51.614411	
4	5	200501RY10089	551510.0	177770	0.180406	51.478547	

5 rows × 33 columns

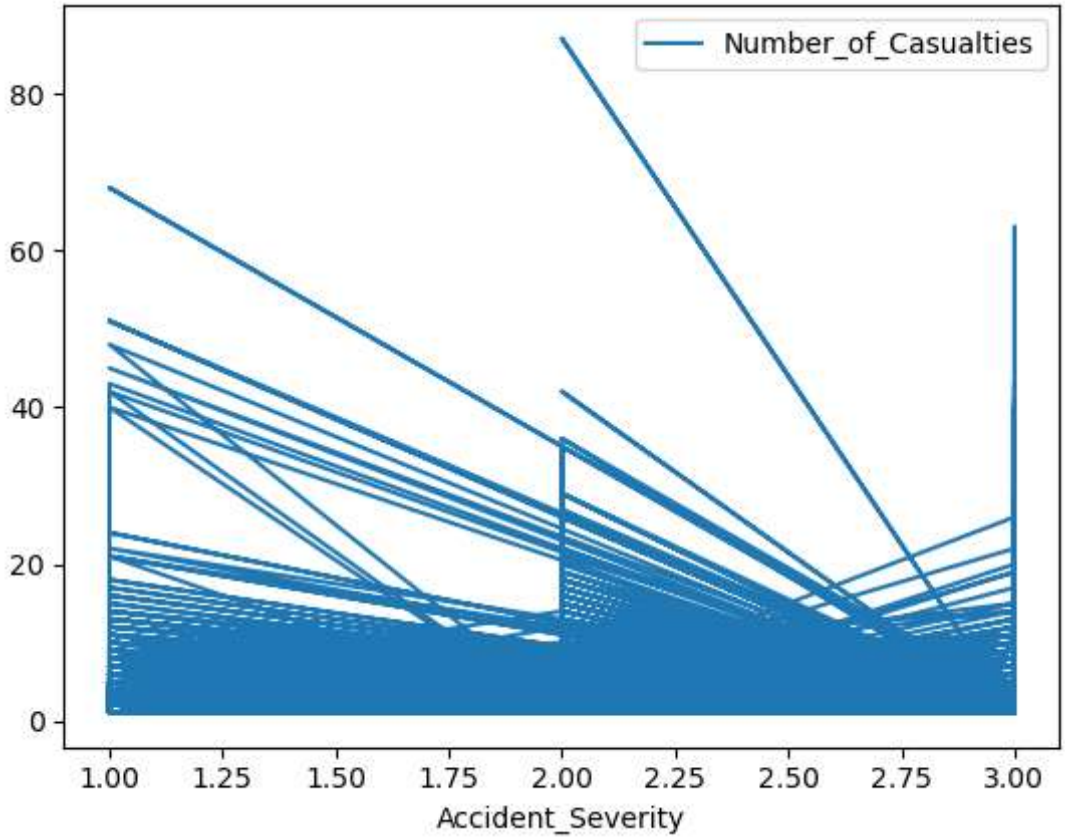
In [3]:
```
df.plot(x="Year", y="S no.")
plt.show()
```



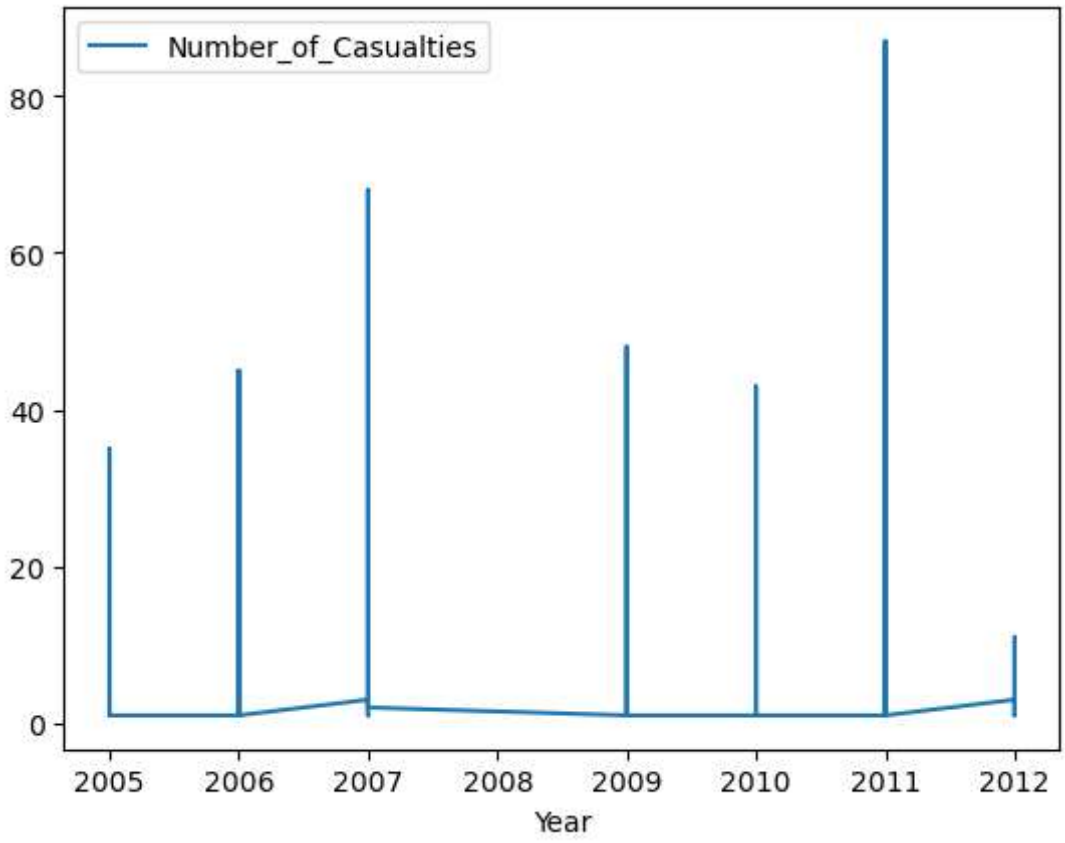
In [4]:
```
df.plot(x="Year", y=["Accident_Severity","Number_of_Casualties"])
plt.show()
```



In [5]:
```
df.plot(x="Accident_Severity",y="Number_of_Casualties")
plt.show()
```

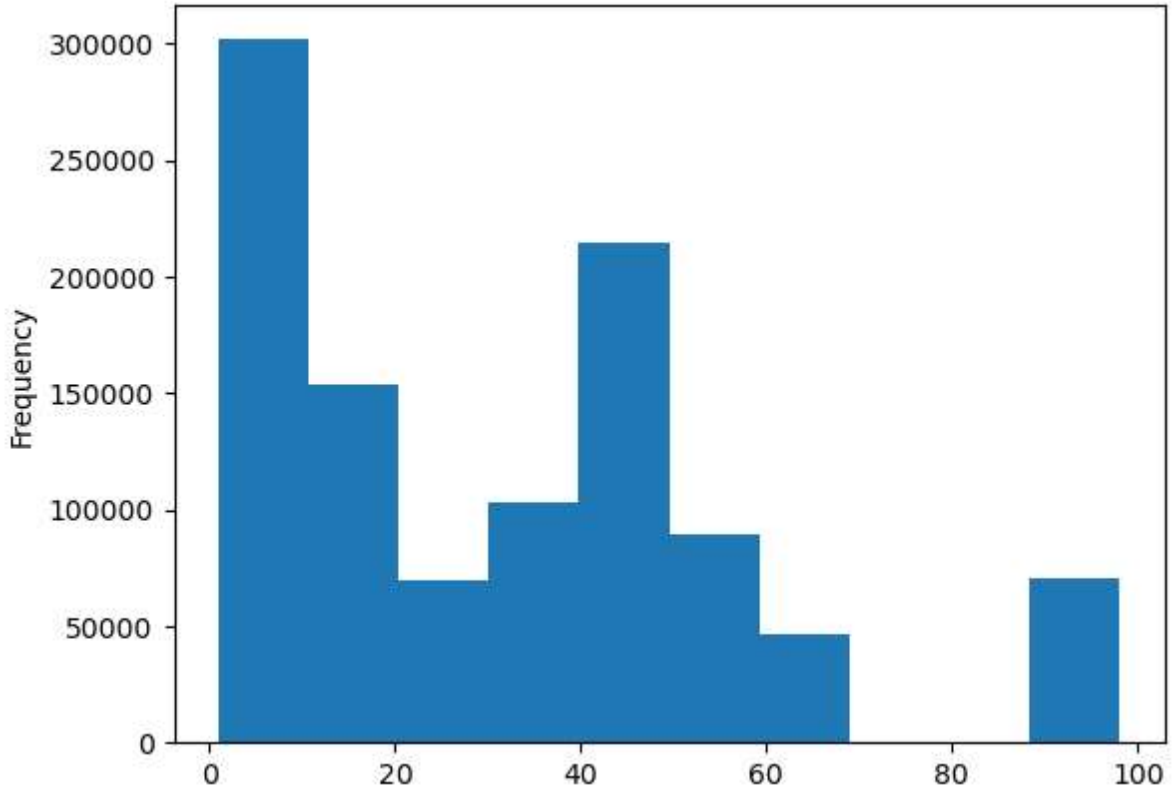


In [6]:
```
df.plot(x="Year", y="Number_of_Casualties")
plt.show()
```



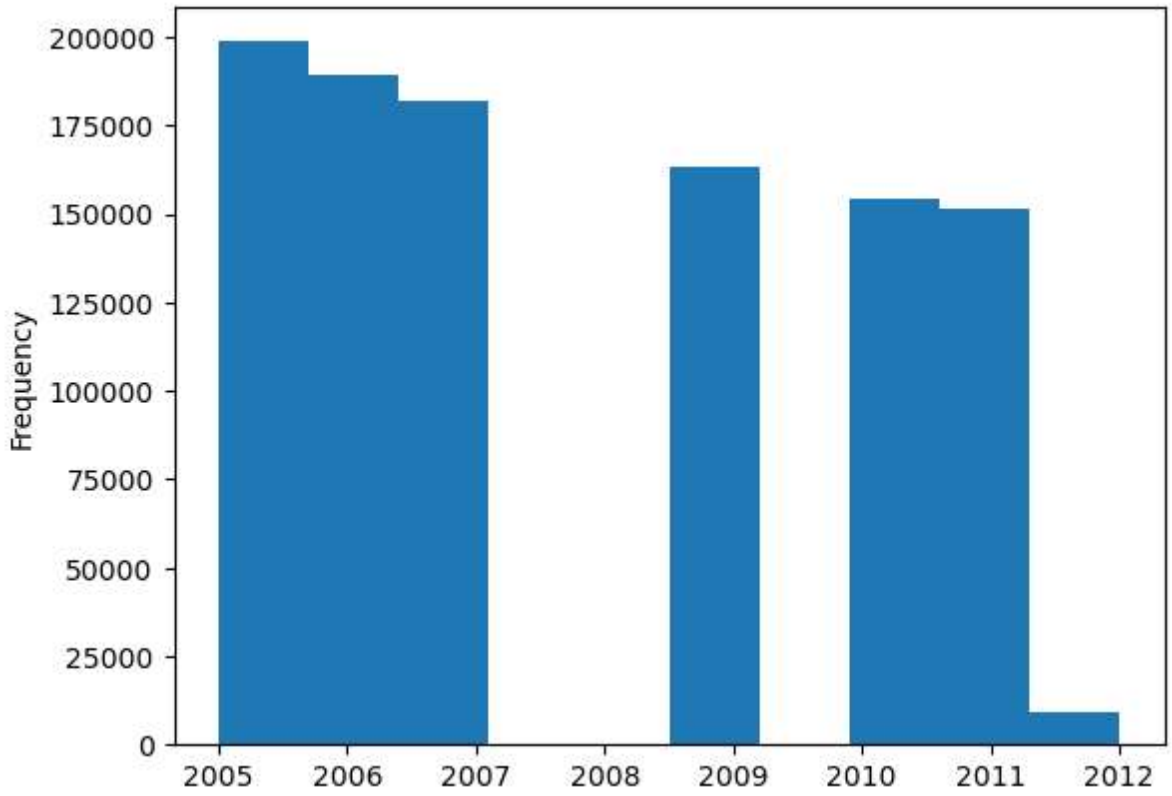
In [7]:
```
median_column = df["Police_Force"]
median_column.plot(kind="hist")
```

Out[7]:
```
<AxesSubplot:ylabel='Frequency'>
```



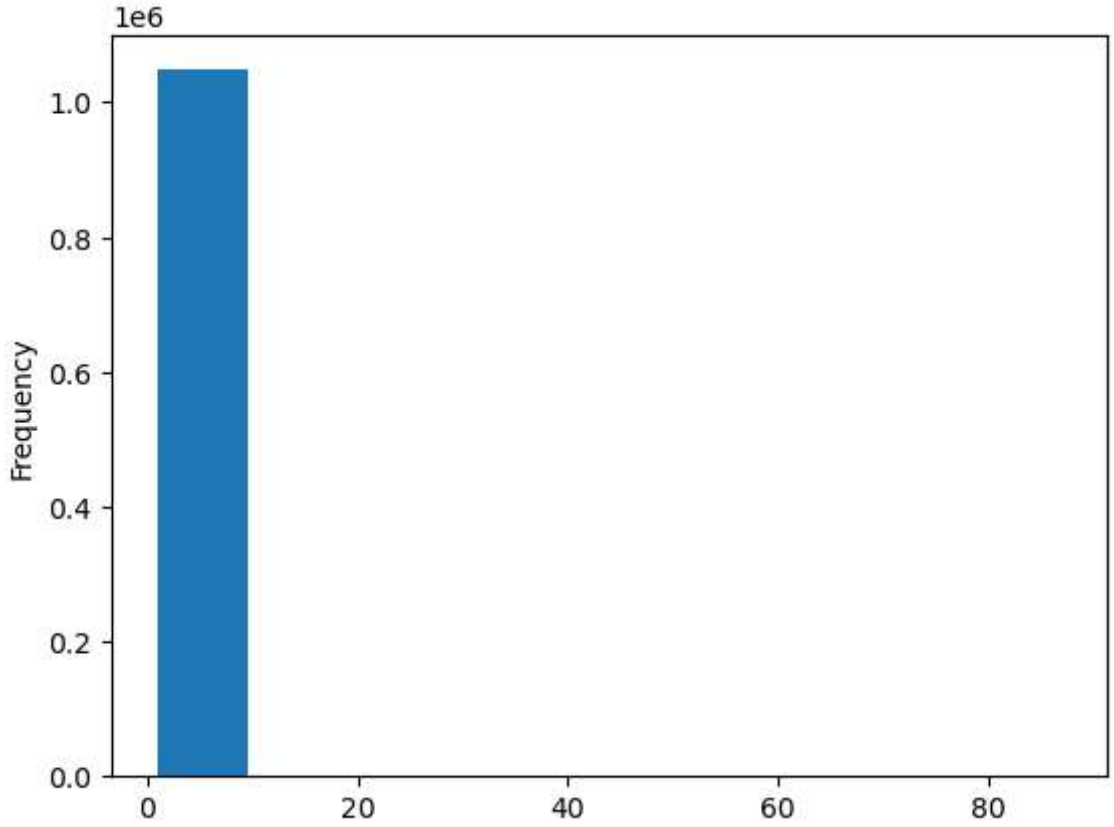
In [8]:
```
median_column = df["Year"]
median_column.plot(kind="hist")
```

Out[8]:
```
<AxesSubplot:ylabel='Frequency'>
```



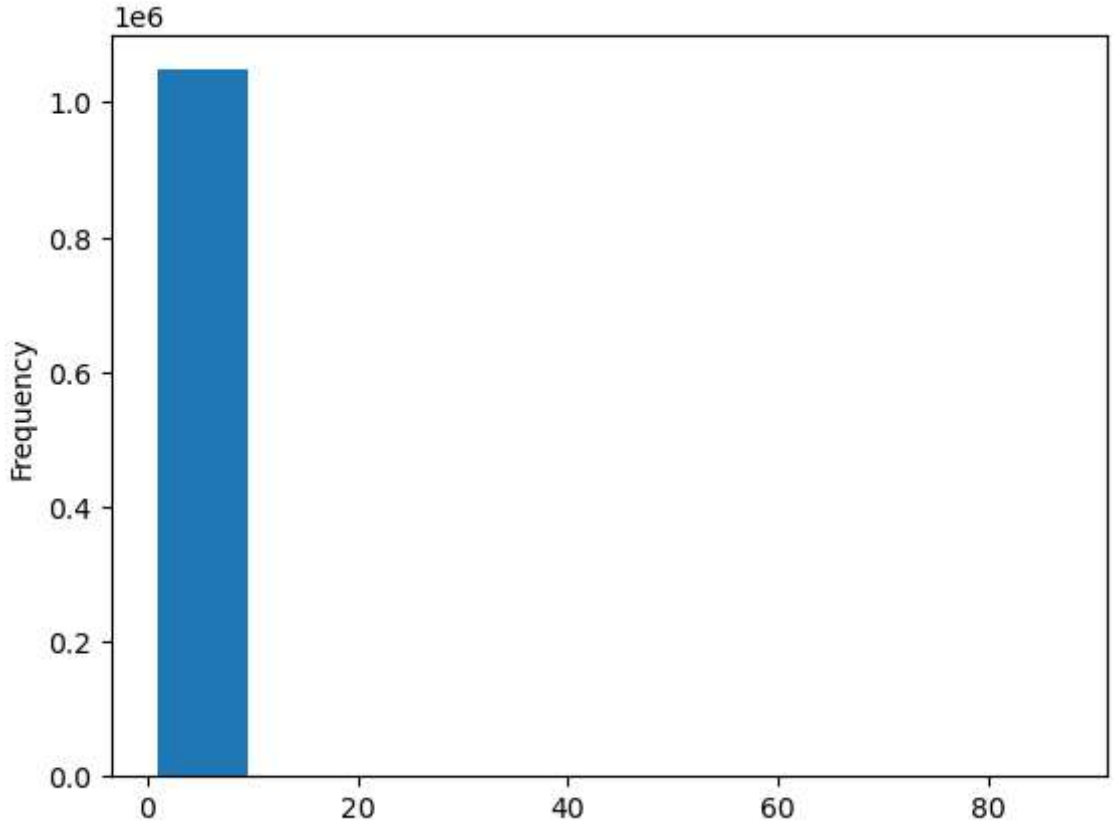
In [9]:
```
median_column = df["Number_of_Casualties"]
median_column.plot(kind="hist")
```

Out[9]:
```
<AxesSubplot:ylabel='Frequency'>
```



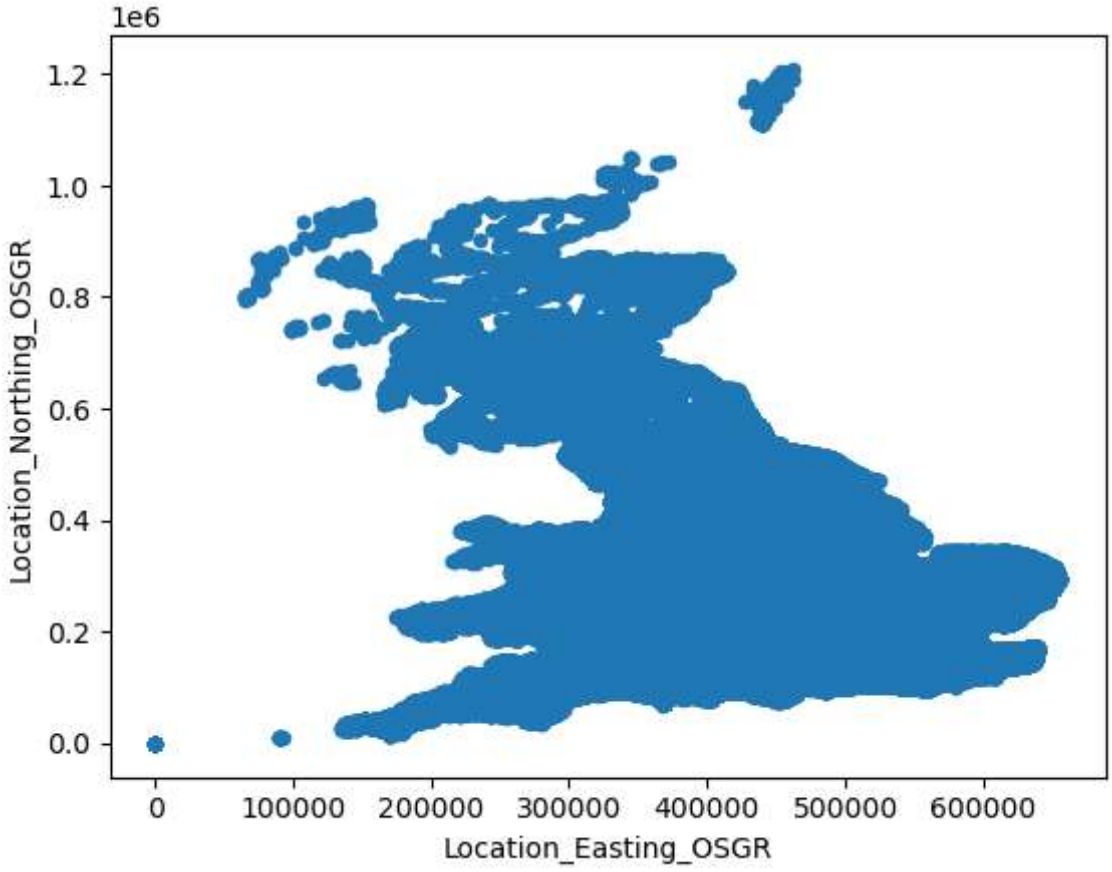
In [10]:
```
median_column = df["Number_of_Casualties"]
median_column.plot(kind="hist")
```

Out[10]:
```
<AxesSubplot:ylabel='Frequency'>
```



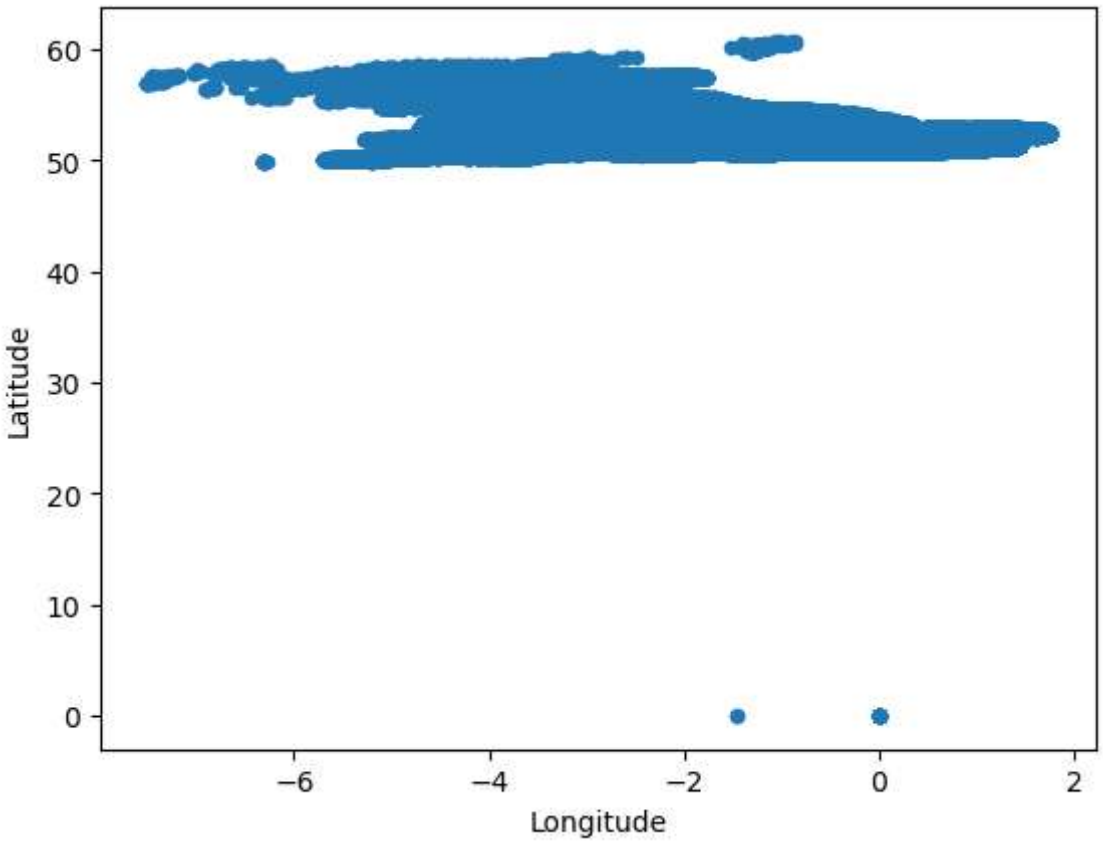
In [11]:
```
df.plot(x="Location_Easting_OSGR", y="Location_Northing_OSGR", kind="scatter")
```

Out[11]:
```
<AxesSubplot:xlabel='Location_Easting_OSGR', ylabel='Location_Northing_OSGR'>
```



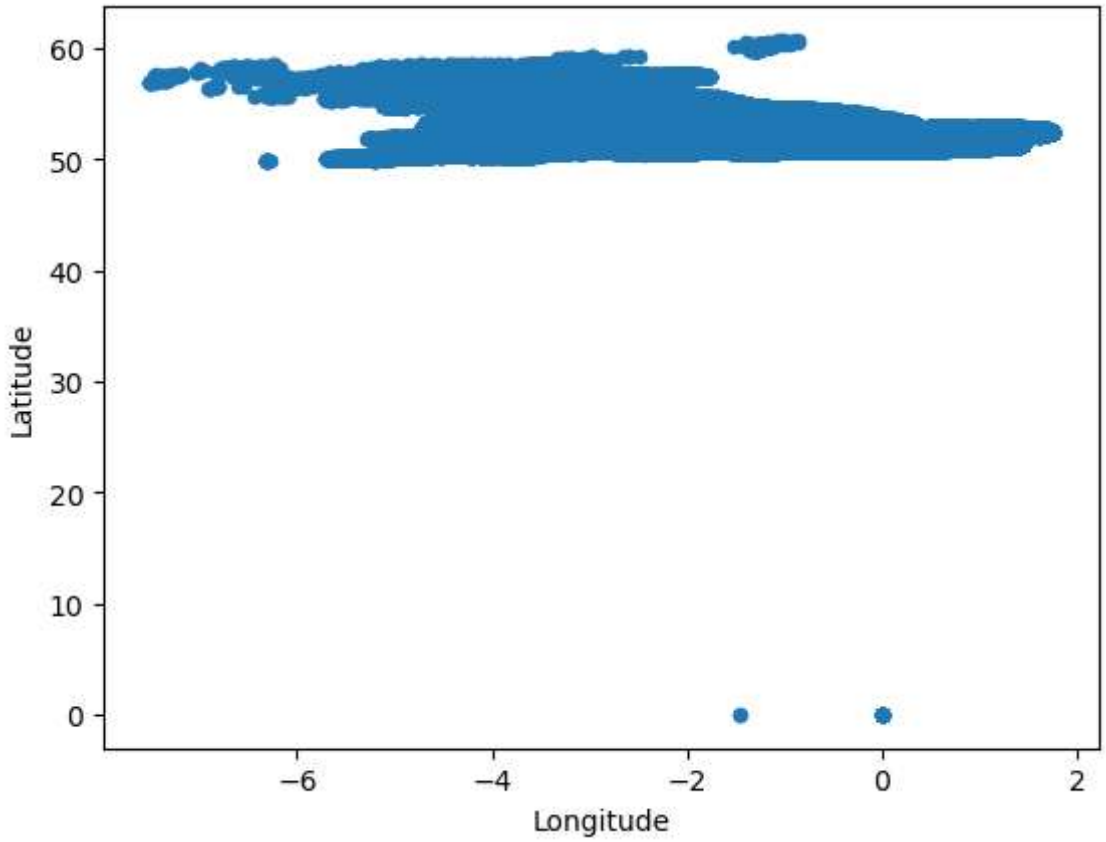
In [12]:
```
df.plot(x="Longitude", y="Latitude", kind="scatter")
```

Out[12]:
```
<AxesSubplot:xlabel='Longitude', ylabel='Latitude'>
```



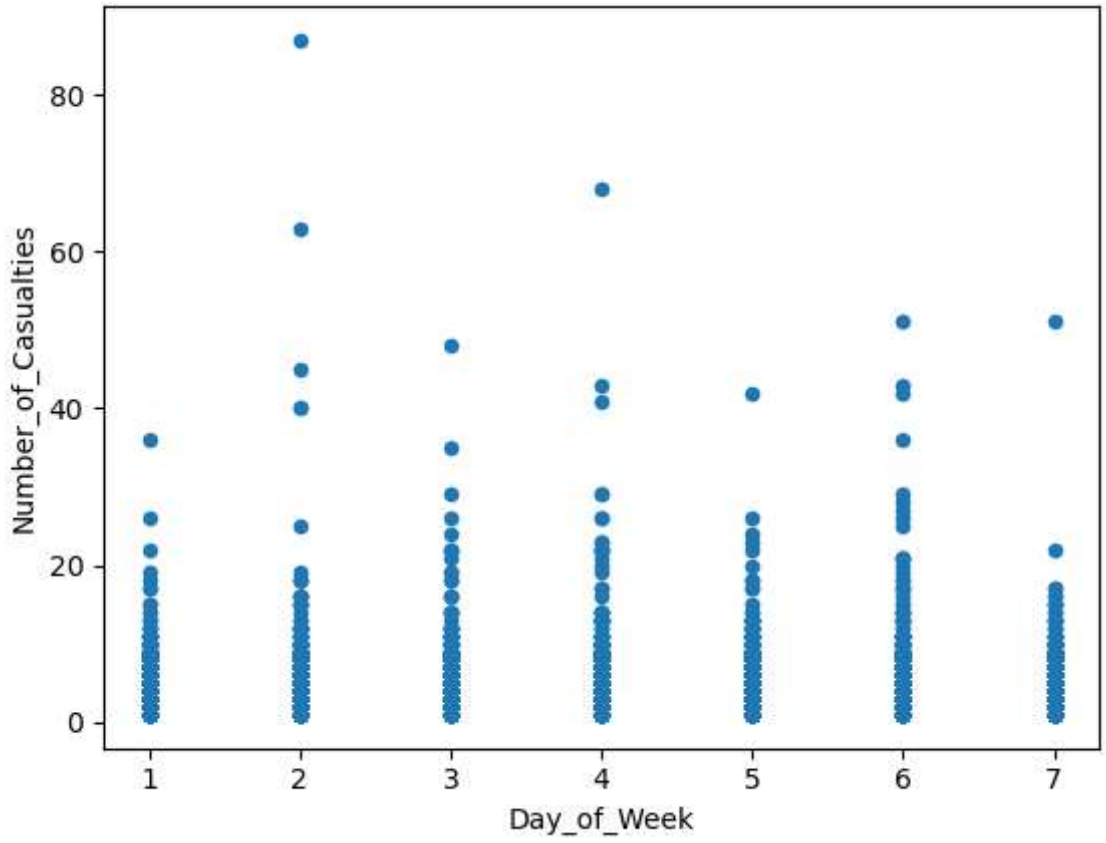
In [13]:
```
df.plot(x="Longitude", y="Latitude", kind="scatter")
```

Out[13]:
```
<AxesSubplot:xlabel='Longitude', ylabel='Latitude'>
```



In [14]:
```
df.plot(x="Day_of_Week", y="Number_of_Casualties", kind="scatter")
```

Out[14]:
```
<AxesSubplot:xlabel='Day_of_Week', ylabel='Number_of_Casualties'>
```



In [15]:
```
cat_totals = df.groupby("Road_Surface_Conditions")["S no."].sum().sort_values()
cat_totals
```

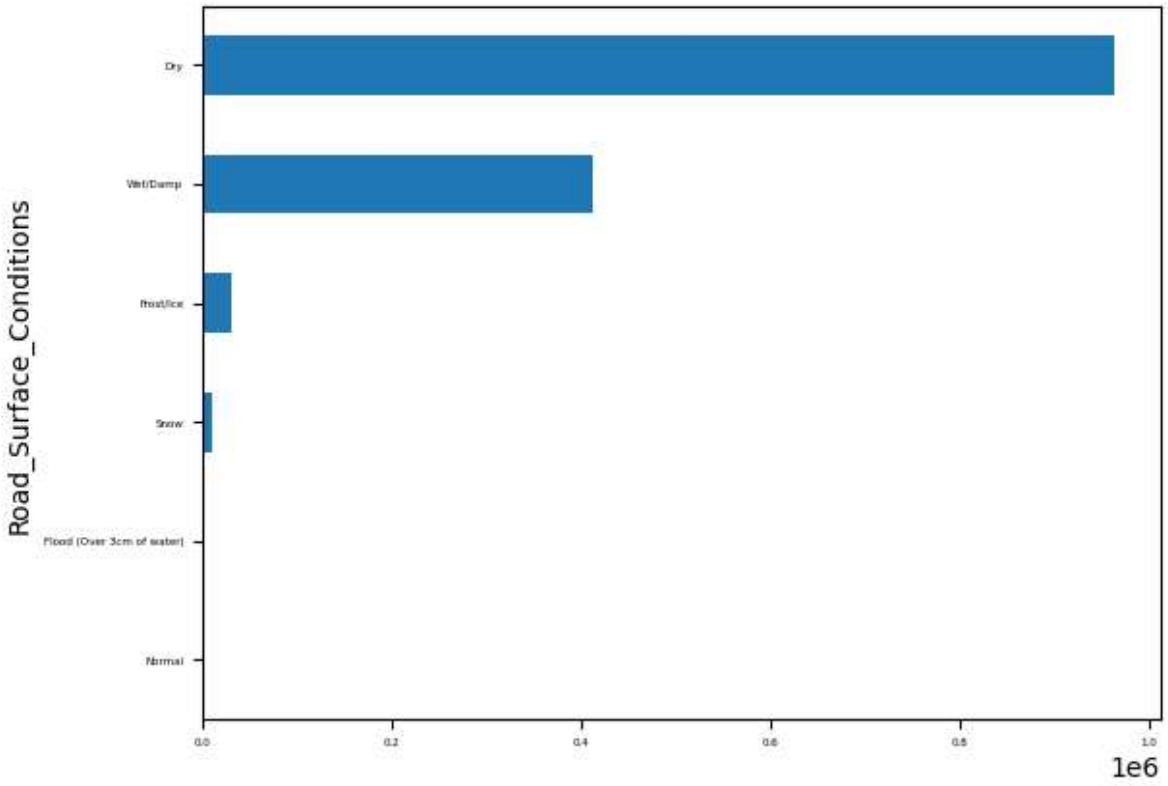
Out[15]:
```
Road_Surface_Conditions
Normal                          639833088
Flood (Over 3cm of water)       666207014
Snow                           4737404142
Frost/Ice                     14228062690
Wet/Damp                     149274362365
Dry                          380209420301
Name: S no., dtype: int64
```

In [16]:
```
cat_totals = df.groupby("Road_Surface_Conditions")["Number_of_Casualties"].sum().s
cat_totals
```

Out[16]:
```
Road_Surface_Conditions
Normal                          1480
Flood (Over 3cm of water)       1909
Snow                           10422
Frost/Ice                      31456
Wet/Damp                      413186
Dry                           964196
Name: Number_of_Casualties, dtype: int64
```

In [17]:
```
cat_totals.plot(kind="barh", fontsize=4)
```

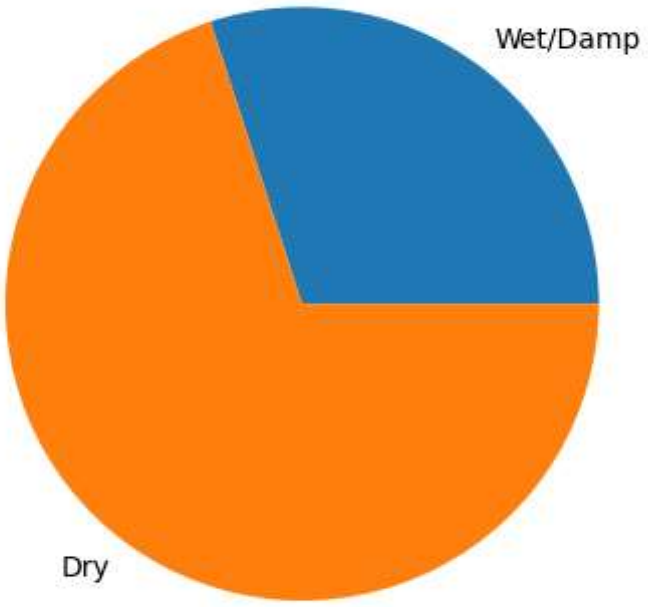
Out[17]:
```
<AxesSubplot:ylabel='Road_Surface_Conditions'>
```



In [18]:
```
small_cat_totals = cat_totals[cat_totals < 100_000]
big_cat_totals = cat_totals[cat_totals > 100_000]
small_sums = pd.Series([small_cat_totals.sum()])
```

In [19]:
```
big_cat_totals.plot(kind="pie", label="")
```

Out[19]:
```
<AxesSubplot:>
```



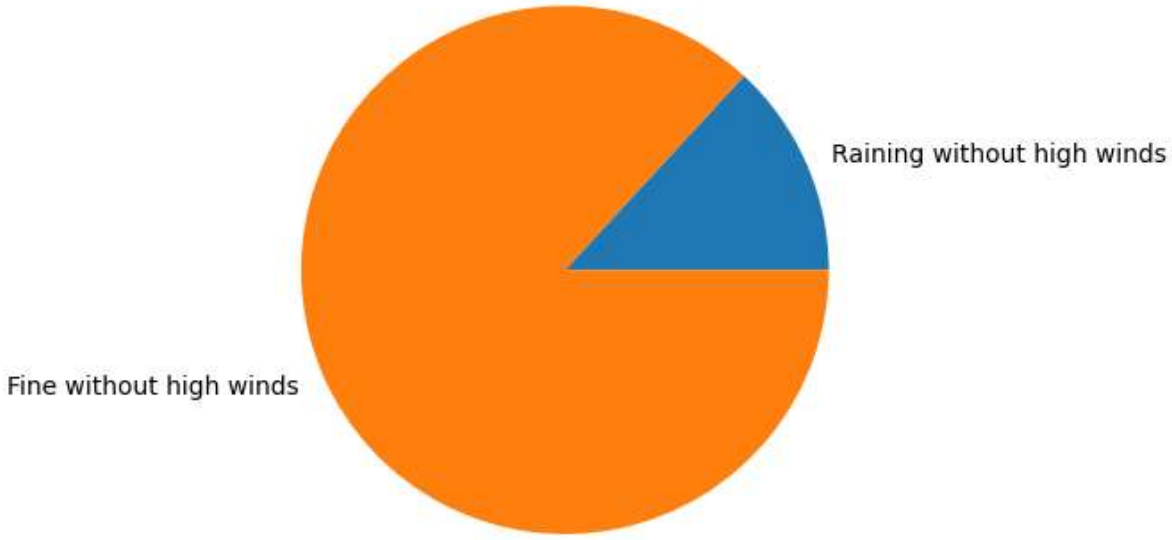
In [20]:
```
cat2_totals = df.groupby("Weather_Conditions")["Number_of_Casualties"].sum().sort_
```

In [21]:
```
cat2_totals
```

Out[21]:
```
Weather_Conditions
Snowing with high winds          1769
Fog or mist                      8273
Snowing without high winds      11489
Fine with high winds            18472
Raining with high winds         19395
Unknown                         25062
Other                           34522
Raining without high winds     171253
Fine without high winds       1132414
Name: Number_of_Casualties, dtype: int64
```

In [22]:
```
small_cat2_totals = cat2_totals[cat2_totals < 100_000]
big_cat2_totals = cat2_totals[cat2_totals > 100_000]
small_sums = pd.Series([small_cat2_totals.sum()])
big_cat2_totals.plot(kind="pie", label="")
```

Out[22]:
```
<AxesSubplot:>
```



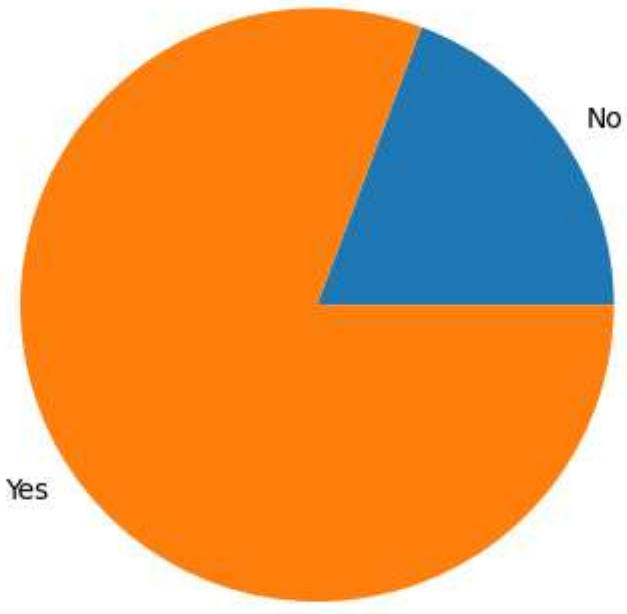
In [23]:
```
cat3_totals = df.groupby("Did_Police_Officer_Attend_Scene_of_Accident")["S no."].s
```

In [24]:
```
cat3_totals
```

Out[24]:
```
Did_Police_Officer_Attend_Scene_of_Accident
No     105968649954
Yes    443786639646
Name: S no., dtype: int64
```

In [25]:
```
small_cat3_totals = cat3_totals[cat3_totals < 100_000]
big_cat3_totals = cat3_totals[cat3_totals > 100_000]
small_sums = pd.Series([small_cat3_totals.sum()])
big_cat3_totals.plot(kind="pie", label="")
```

Out[25]:
```
<AxesSubplot:>
```



In [ ]:
```
top_medians = df[df["S no."] > 60000].sort_values("S no.")
top_medians.plot(x="S no.", y=["Number_of_Casualties", "Accident_Severity", "Numbe
```

In [ ]:
```

```