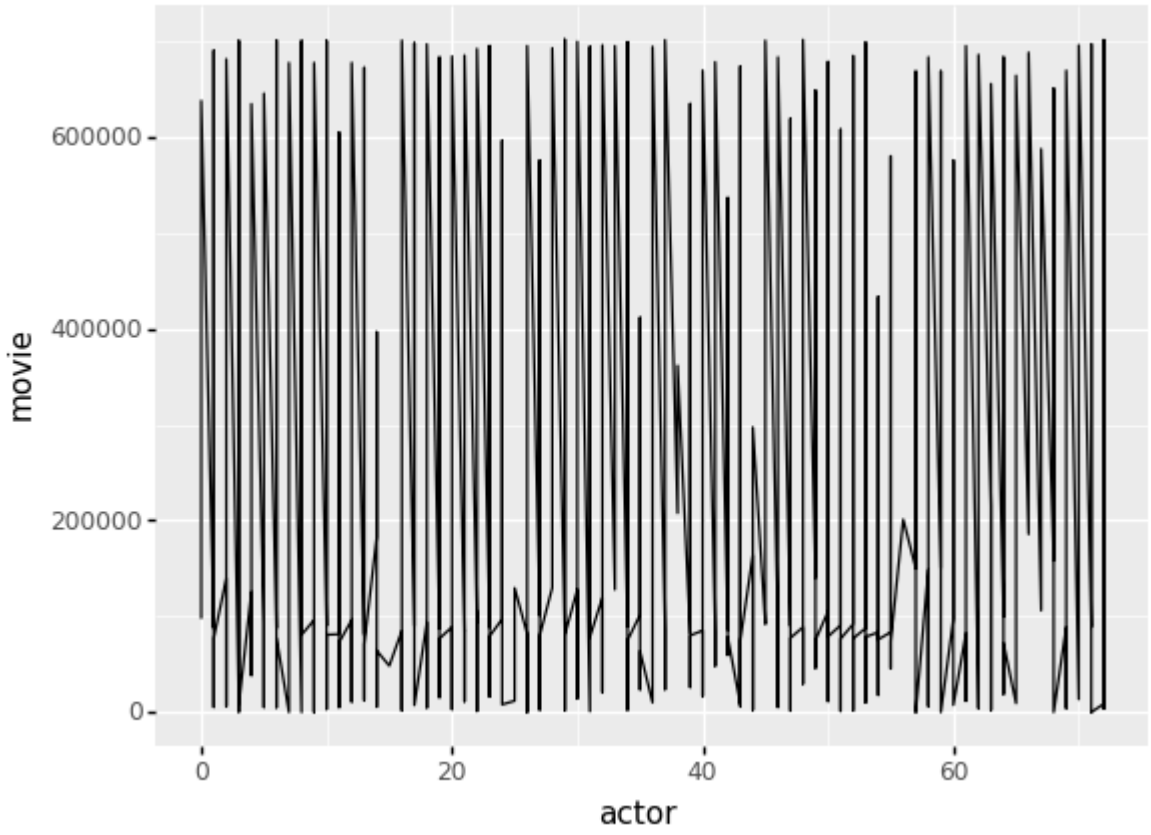
In [13]:
```
"Inicialmente, temos uma base de dados com duas colunas. A primeira delas com a
```

Out[13]:
```
'Inicialmente, temos uma base de dados com duas colunas. A primeira delas com a
identificação de um ator e a segunda com identificações de filmes que esse ator
trabalhou'
```

In [14]:
```
import pandas as pd

base = pd.read_csv('teste.csv')
base.head(3122)
print(base.head(3122))
```

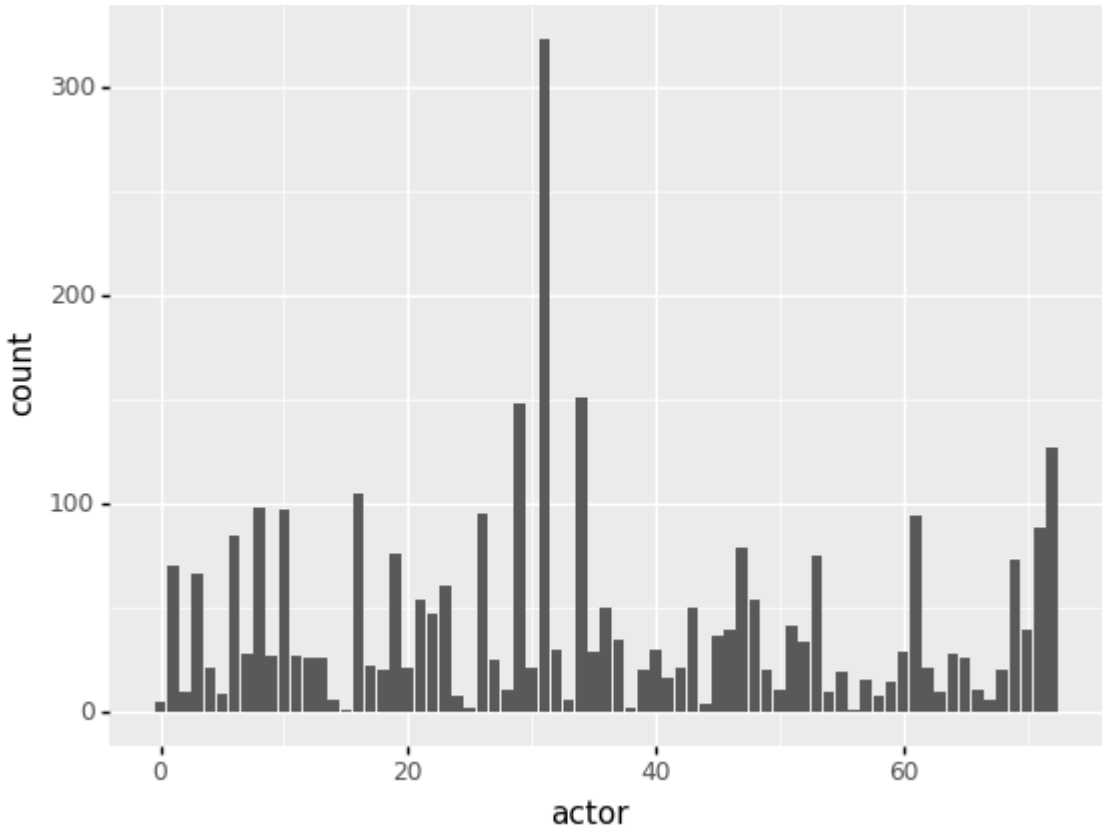
```
      actor   movie
0         0   97500
1         0  290039
2         0  326858
3         0  326859
4         0  637772
...     ...     ...
3117     72  667353
3118     72  673528
3119     72  677215
3120     72  680431
3121     72  701118

[3122 rows x 2 columns]
```

In [18]:
```
import pandas as pd
from plotnine import ggplot, aes, geom_line
base = pd.read_csv('teste.csv')
(
    ggplot(base)  # What data to use
    + aes(x="actor", y="movie")  # What variable to use
    + geom_line()  # Geometric object to use for drawing
)
```



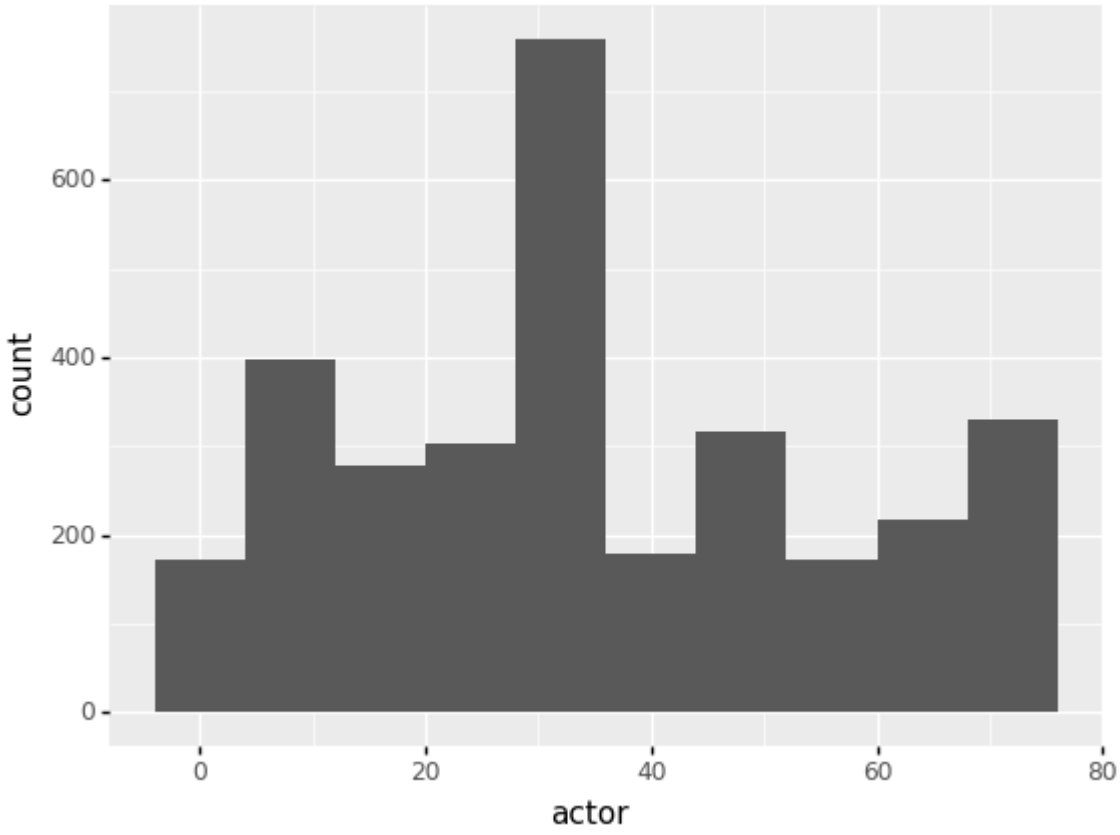
Out[18]: <ggplot: (8769511539417)>

In [2]:
```
import pandas as pd
from plotnine import ggplot, aes, geom_line, geom_point, geom_bar

base = pd.read_csv('teste.csv')

ggplot(base) + aes(x="actor") + geom_bar()
```



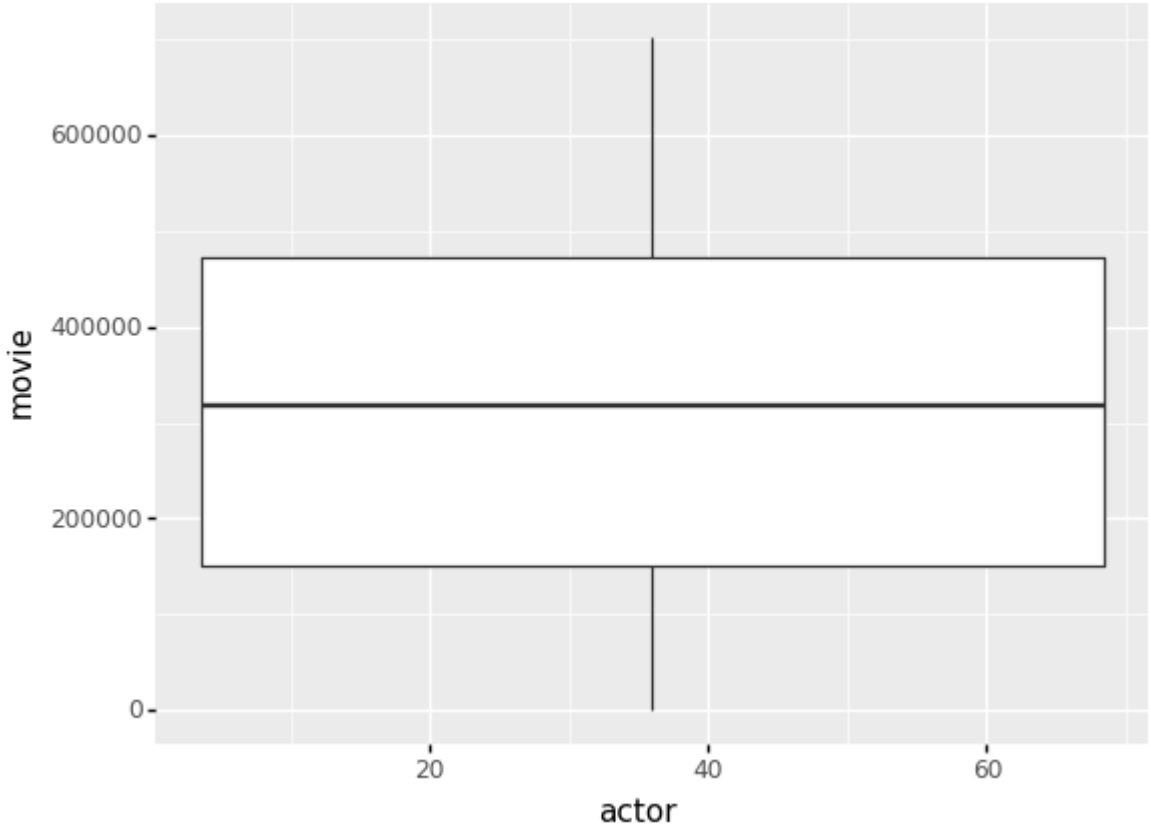
Out[2]: <ggplot: (8769589578427)>

In [3]:
```
import pandas as pd
from plotnine import ggplot, aes, geom_histogram

base = pd.read_csv('teste.csv')

ggplot(base) + aes(x="actor") + geom_histogram(bins=10)
```



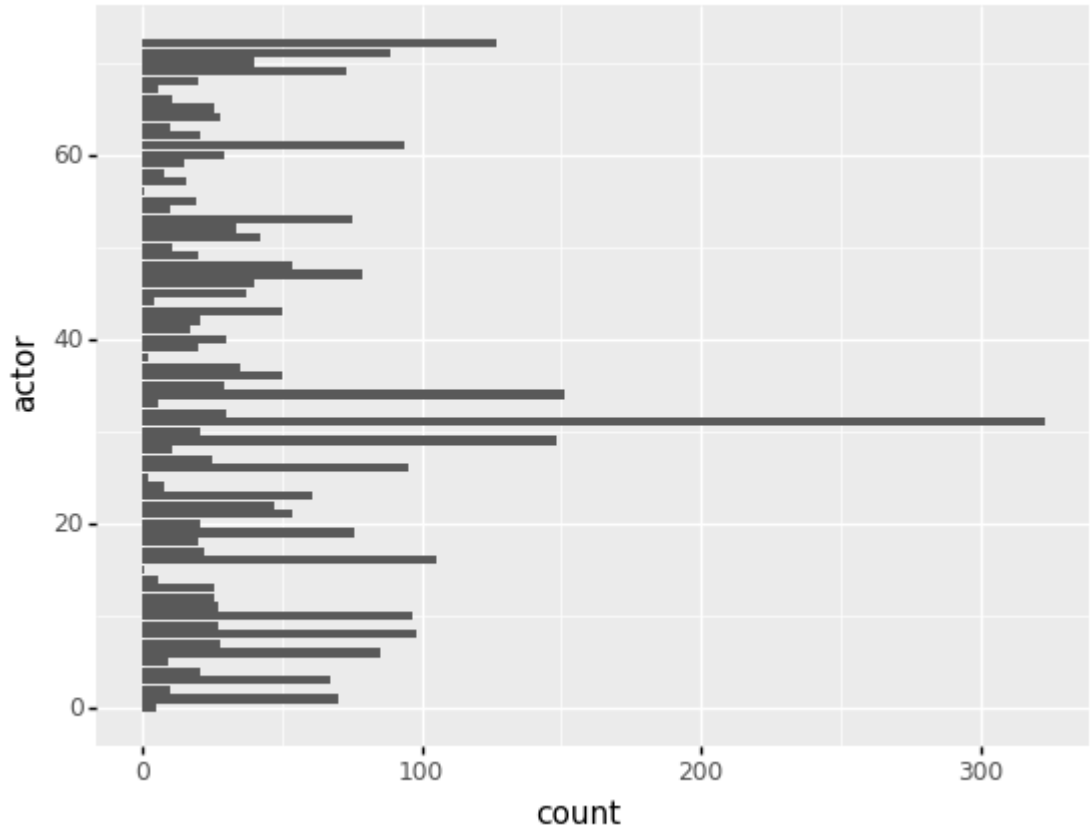
Out[3]: <ggplot: (8769514139450)>

In [5]:
```
import pandas as pd
from plotnine import ggplot, aes, geom_boxplot

base = pd.read_csv('teste.csv')

(
  ggplot(base)
  + aes(x="actor", y="movie")
  + geom_boxplot()
)
```



Out[5]: <ggplot: (8769511913054)>

In [8]:
```
import pandas as pd
from plotnine import ggplot, aes, geom_bar, coord_flip
base = pd.read_csv('teste.csv')

ggplot(base) + aes(x="actor") + geom_bar() + coord_flip()
```



Out[8]: <ggplot: (8769511679684)>

In [ ]:

In [ ]: