In [38]:
```
import pandas as pd
```

In [39]:
```
car_sales = pd.read_csv("data/car-sales.csv")
car_sales
```

Out[39]:

	Make	Colour	Odometer (KM)	Doors	Price
0	Toyota	White	150043	4	$4,000.00
1	Honda	Red	87899	4	$5,000.00
2	Toyota	Blue	32549	3	$7,000.00
3	BMW	Black	11179	5	$22,000.00
4	Nissan	White	213095	4	$3,500.00
5	Toyota	Green	99213	4	$4,500.00
6	Honda	Blue	45698	4	$7,500.00
7	Honda	Blue	54738	4	$7,000.00
8	Toyota	White	60000	4	$6,250.00
9	Nissan	White	31600	4	$9,700.00

In [40]:
```
# Show the count of each pair of Make and Doors values
pd.crosstab(car_sales["Make"], car_sales["Doors"])
```

Out[40]:

Doors	3	4	5
Make			
BMW	0	0	1
Honda	0	3	0
Nissan	0	2	0
Toyota	1	3	0

In [41]:
```
# Group the rows according to Make column
# Show mean for each numeric columns
# Instead of mean some other functions also can used. (Ex: sum, std, . . .)
car_sales.groupby(["Make"]).mean()
```

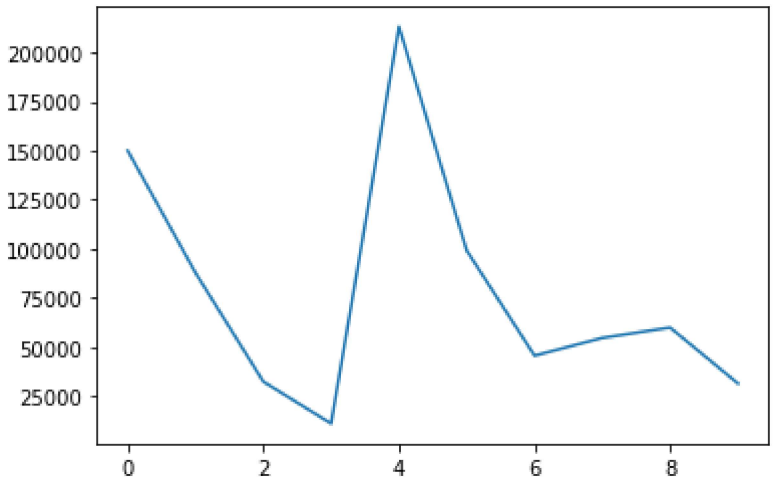
Out[41]:

	Odometer (KM)	Doors
Make		
BMW	11179.000000	5.00
Honda	62778.333333	4.00
Nissan	122347.500000	4.00

Make	Odometer (KM)	Doors
Toyota	85451.250000	3.75

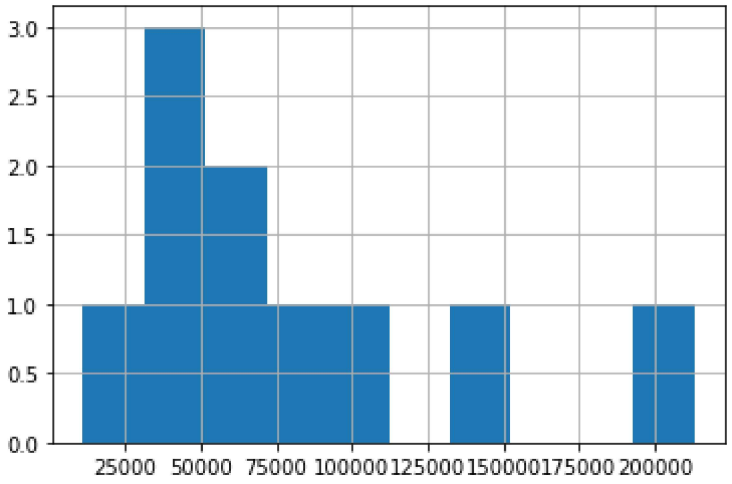
In [42]:
```
# Plot Odometer (KM) column with index
car_sales["Odometer (KM)"].plot()
```

Out[42]: <AxesSubplot:>



In [43]:
```
# Plot histogram for Odometer (KM) column with index
car_sales["Odometer (KM)"].hist()
```

Out[43]: <AxesSubplot:>



In [47]:
```
# Convert String column(Price) to an integer column

# replace , . $ symbols with ''
# convert resulting string to an integer
# devide that value by 100 (Because we remove decimal symbol also)
car_sales["Price"] = car_sales["Price"].replace('[\$\,\.]', '', regex=True).astype(int)
```

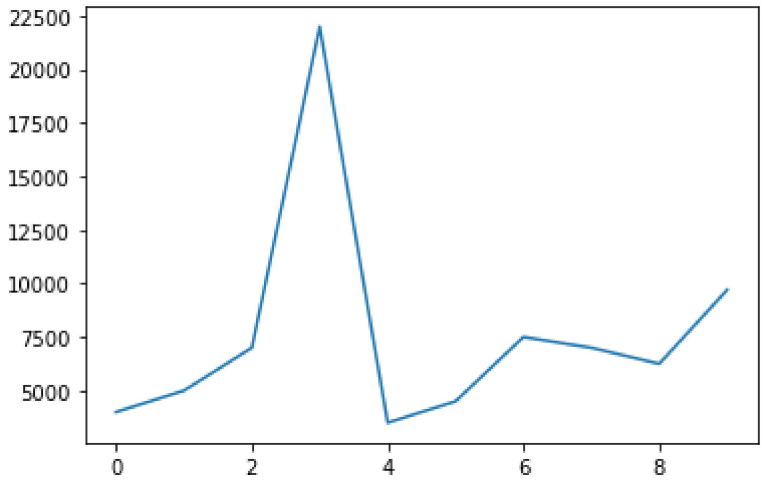
In [48]:

```
car_sales["Price"]
```

Out[48]:
```
0     4000.0
1     5000.0
2     7000.0
3    22000.0
4     3500.0
5     4500.0
6     7500.0
7     7000.0
8     6250.0
9     9700.0
Name: Price, dtype: float64
```

In [49]:
```
car_sales["Price"].plot()
```

Out[49]:
```
<AxesSubplot:>
```



In [ ]:
```

```