Distances for ML models

In [1]:
```
from scipy.spatial import distance
import pandas as pd
```

In [2]:
```
low_memory = False
```

In [3]:
```
df = pd.read_csv("ramen-ratings.csv")
df = df[df.Stars != 'Unrated']
df = df['Stars']
df = df.astype(float, errors = 'raise')
mean = float(df.mean())
mean
```

Out[3]:
```
3.6546759798214974
```

In [4]:
```
DataPoint1 = (mean, 5)
DataPoint2 = (df[0],5)
DataPoint1, DataPoint2
```

Out[4]:
```
((3.6546759798214974, 5), (3.75, 5))
```

In [5]:
```
StringPoint1 = "Zebra"
StringPoint2 = "Frogs"
```

Euclidean Distance

In [6]:
```
EuclideanDistance = distance.euclidean(DataPoint1, DataPoint2)
print("The Distance, Euclidean between Data Point 1 and 2 is : ", EuclideanDistance)
```

```
The Distance, Euclidean between Data Point 1 and 2 is :  0.09532402017850261
```

Manhattan Distance

In [7]:
```
ManhattanDistance = distance.cityblock(DataPoint1, DataPoint2)
print("The Distance, Manhattan, between Data Point 1 and 2 is : ", ManhattanDistance
```

```
The Distance, Manhattan, between Data Point 1 and 2 is :  0.09532402017850261
```

Minkowski Distance

In [8]:
```
MinkowskiDistance = distance.minkowski(DataPoint1, DataPoint2)
print("The Distance, Minkowski, between Data Point 1 and 2 is : ", MinkowskiDistance
```

```
The Distance, Minkowski, between Data Point 1 and 2 is :  0.09532402017850261
```

Hamming Distance

In [9]:

```
HammingDistance = distance.hamming(list(StringPoint1), list(StringPoint2))* len(Stri
print("The Distance, Hamming, between the String Points 1 and 2 is : ", HammingDista
```

```
The Distance, Hamming, between the String Points 1 and 2 is :  5.0
```