

```
In [1]: import pandas as pd
```

```
In [2]: df=pd.read_csv("jelly.csv",engine='python')
```

## READING DATA

```
In [7]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>  
RangeIndex: 1704 entries, 0 to 1703  
Data columns (total 3 columns):  
#   Column      Non-Null Count  Dtype  
---  ---  
0   company    1704 non-null   object  
1   address    1698 non-null   object  
2   db          1702 non-null   object  
dtypes: object(3)  
memory usage: 40.1+ KB
```

## REMOVING UNWANTED COLUMN

```
In [4]: df.rename(columns={"Address":"address","Database Address":"db","Unnamed: 2":"company"},inplace=True)
```

```
In [6]: df.drop(["Unnamed: 0","Unnamed: 1","Unnamed: 4","Unnamed: 5"],axis=1,inplace=True)
```

```
In [14]: df.head(10)
```

```
Out[14]:
```

|   | company                               | address   | db  |
|---|---------------------------------------|---|---|
| 0 | KIC Financial Brokerage KSC           | Kuwait City (Qibla), Fahad Alsalem Street, Awt... | Fahad Alsalem Street Awtad Tower Floor#15 K...    |
| 1 | Oman Agriculture Development Co SAOG  | P.O Box No. 2398, Al Ghubra, Postal Code 130, ... | Sun Farms, Near Penguin RestaurantAl Suwaihrah... |
| 2 | National Hospitality Institute SAOC   | Building No.900, Way No.63, Al Fursan Street, ... | P O Box 854:131AlHamriyaAl Wadi Al Kabir , Nr ... |
| 3 | Qatar Real Estate Investment PJSC     | Alaqaria Tower P.O. Box 22311, Old Salata area... | P.O.Box : 22311 Doha Qatar                        |
| 4 | EFG Hermes IFA Financial Brokerage Co | Kuwait Stock Exchange Building, 4th floor, P.O... | 4TH FLOOR EAST ENTRANCEP.O. BOX 22695STOCK EX...  |
| 5 | Oman Air SAOC                         | PO Box 58 Postal Code 111 Muscat International... | P O Box 58:111CPOHead Quarters Muscat Airport ... |
| 6 | Dhofar Power Co                       | 1st Floor, A'Sahwa Complex Building, near Cham... | Near Chamber of Commerce Salalah A'Sahwa Compl... |
| 7 | Oman Hotels and Tourism SAOG          | PO Box 2031 ,Postal Code 112, Ruwi, Muscat, Su... | Way 133, Al Farahidi StreetRuwi, Muscat - Ruwi... |
| 8 | Sohar Poultry SAOG                    | Building No. 1678, Al Mawaleh, Railway 4720, S... | P O Box 2808:112RuwiRuwi , Next Al Jadeeda, Me... |
| 9 | Oman Medical Projects Co              | P.O. Box 8, Al Azaiba 130, Oman                   | P.O. BOX 8MUSCAT PRIVATE HOSPITAL BAUSHER STR...  |

```
In [9]: df = df.iloc[2:]
```

```
In [11]: df.reset_index(inplace=True)
```

```
In [13]: df.drop(["index"],axis=1,inplace=True)
```

## IMPORTING NLP LIBRARY

```
In [15]: from fuzzywuzzy import fuzz
```

```
C:\Users\Akash\anaconda3\lib\site-packages\fuzzywuzzy\fuzz.py:11: UserWarning: Using slow pure-python SequenceMatcher. Install python-Levenshtein to remove this warning
  warnings.warn('Using slow pure-python SequenceMatcher. Install python-Levenshtein to remove this warning')
```

```
In [28]: import itertools
```

```
In [32]: dist=[]
         for (i,j) in itertools.zip_longest(df["address"],df["db"]):
             dis=fuzz.token_set_ratio(str(i).lower(),str(j).lower())
             dist.append(dis)
```

```
In [33]: dist
```

```
Out[33]: [97,
          31,
          60,
          70,
          83,
          72,
          85,
          44,
          45,
          88,
          46,
          45,
          73,
          66,
          48,
          77,
          51,
          42,
          63,
          79,
          67,
          72,
          55,
```

74,  
62,  
50,  
34,  
59,  
50,  
34,  
67,  
44,  
65,  
30,  
58,  
45,  
40,  
91,  
54,  
54,  
45,  
100,  
65,  
71,  
100,  
41,  
49,  
37,  
87,  
35,  
88,  
73,  
35,  
47,  
57,  
53,  
95,  
72,  
80,  
38,  
77,  
100,

49,  
70,  
72,  
88,  
81,  
100,  
69,  
64,  
70,  
85,  
68,  
68,  
83,  
62,  
47,  
39,  
39,  
77,  
54,  
81,  
50,  
46,  
20,  
41,  
78,  
84,  
76,  
79,  
62,  
74,  
74,  
100,  
78,  
75,  
65,  
79,  
70,  
71,  
90,

64,  
62,  
56,  
58,  
100,  
78,  
100,  
27,  
42,  
73,  
47,  
93,  
71,  
70,  
59,  
68,  
74,  
92,  
100,  
68,  
90,  
62,  
45,  
61,  
41,  
100,  
50,  
32,  
100,  
88,  
80,  
44,  
72,  
38,  
32,  
87,  
95,  
77,  
35,

79,  
48,  
37,  
74,  
61,  
78,  
100,  
63,  
63,  
31,  
32,  
86,  
59,  
83,  
79,  
60,  
77,  
65,  
56,  
79,  
64,  
70,  
49,  
64,  
35,  
89,  
100,  
80,  
59,  
58,  
69,  
83,  
52,  
85,  
78,  
100,  
79,  
64,  
94,

66,  
67,  
88,  
53,  
89,  
59,  
68,  
65,  
86,  
100,  
53,  
80,  
51,  
59,  
94,  
83,  
100,  
52,  
45,  
34,  
100,  
86,  
100,  
85,  
100,  
58,  
100,  
72,  
46,  
86,  
43,  
74,  
100,  
39,  
80,  
50,  
35,  
76,  
50,



85,  
40,  
91,  
51,  
50,  
46,  
87,  
32,  
83,  
53,  
54,  
26,  
50,  
99,  
61,  
100,  
37,  
78,  
87,  
33,  
68,  
58,  
58,  
39,  
77,  
69,  
59,  
63,  
47,  
46,  
76,  
54,  
51,  
59,  
63,  
56,  
38,  
56,  
100,

67,  
100,  
70,  
26,  
58,  
69,  
63,  
78,  
89,  
64,  
65,  
91,  
78,  
77,  
73,  
43,  
84,  
80,  
89,  
30,  
68,  
80,  
85,  
70,  
62,  
83,  
57,  
62,  
93,  
74,  
61,  
39,  
57,  
100,  
82,  
100,  
71,  
81,  
90,

60,  
59,  
100,  
43,  
85,  
57,  
36,  
44,  
54,  
83,  
59,  
43,  
76,  
90,  
69,  
91,  
91,  
61,  
82,  
100,  
35,  
100,  
54,  
46,  
8,  
86,  
54,  
72,  
91,  
56,  
70,  
65,  
50,  
67,  
57,  
74,  
61,  
45,  
56,

42,  
34,  
65,  
59,  
37,  
71,  
65,  
78,  
77,  
68,  
77,  
54,  
91,  
59,  
29,  
51,  
81,  
38,  
85,  
84,  
100,  
83,  
75,  
89,  
31,  
67,  
71,  
100,  
48,  
29,  
69,  
45,  
32,  
65,  
64,  
40,  
80,  
36,  
63,

49,  
100,  
54,  
65,  
41,  
39,  
42,  
72,  
39,  
74,  
71,  
91,  
39,  
74,  
67,  
51,  
76,  
63,  
61,  
35,  
67,  
77,  
38,  
95,  
17,  
53,  
87,  
50,  
61,  
82,  
65,  
51,  
34,  
53,  
42,  
100,  
64,  
46,  
76,

64,  
59,  
32,  
100,  
63,  
67,  
33,  
81,  
83,  
36,  
76,  
74,  
31,  
72,  
52,  
64,  
40,  
31,  
55,  
78,  
63,  
55,  
91,  
100,  
93,  
32,  
89,  
36,  
100,  
62,  
91,  
36,  
58,  
57,  
38,  
57,  
56,  
74,  
59,

44,  
57,  
100,  
59,  
62,  
100,  
73,  
94,  
50,  
100,  
49,  
81,  
48,  
68,  
85,  
90,  
61,  
61,  
50,  
76,  
74,  
72,  
89,  
70,  
100,  
94,  
58,  
53,  
63,  
59,  
43,  
78,  
82,  
91,  
80,  
34,  
82,  
31,  
72,

60,  
53,  
43,  
74,  
38,  
71,  
57,  
58,  
42,  
100,  
42,  
78,  
68,  
42,  
32,  
78,  
100,  
100,  
78,  
80,  
47,  
70,  
82,  
64,  
53,  
89,  
64,  
83,  
80,  
89,  
58,  
87,  
53,  
74,  
60,  
66,  
86,  
38,  
82,



84,  
86,  
46,  
46,  
72,  
75,  
93,  
39,  
96,  
83,  
66,  
38,  
91,  
55,  
94,  
32,  
33,  
81,  
67,  
46,  
76,  
67,  
75,  
45,  
90,  
73,  
70,  
91,  
79,  
49,  
50,  
77,  
57,  
100,  
89,  
100,  
78,  
100,  
86,

42,  
56,  
54,  
47,  
70,  
73,  
68,  
38,  
40,  
82,  
71,  
87,  
77,  
82,  
74,  
59,  
79,  
100,  
84,  
96,  
95,  
57,  
78,  
73,  
52,  
45,  
45,  
67,  
67,  
58,  
67,  
31,  
92,  
38,  
54,  
51,  
40,  
60,  
80,

55,  
51,  
57,  
57,  
82,  
82,  
79,  
100,  
80,  
74,  
100,  
44,  
62,  
67,  
91,  
48,  
76,  
31,  
45,  
78,  
41,  
96,  
93,  
80,  
68,  
59,  
68,  
82,  
64,  
38,  
52,  
89,  
32,  
100,  
54,  
87,  
80,  
54,  
61,

81,  
56,  
98,  
79,  
62,  
96,  
54,  
43,  
78,  
87,  
100,  
81,  
91,  
57,  
80,  
45,  
32,  
64,  
63,  
48,  
38,  
88,  
100,  
42,  
65,  
90,  
91,  
65,  
80,  
68,  
55,  
86,  
64,  
46,  
68,  
100,  
100,  
60,  
61,

94,  
91,  
65,  
93,  
59,  
32,  
80,  
55,  
86,  
94,  
30,  
68,  
44,  
63,  
67,  
56,  
71,  
55,  
62,  
44,  
62,  
63,  
71,  
58,  
42,  
91,  
55,  
50,  
67,  
57,  
66,  
60,  
78,  
34,  
58,  
59,  
71,  
46,  
68,

64,  
83,  
58,  
73,  
65,  
36,  
76,  
55,  
63,  
45,  
41,  
41,  
78,  
100,  
21,  
48,  
52,  
36,  
88,  
76,  
53,  
60,  
95,  
46,  
64,  
80,  
51,  
47,  
55,  
49,  
100,  
61,  
45,  
48,  
25,  
75,  
51,  
80,  
31,

90,  
47,  
33,  
46,  
100,  
70,  
75,  
61,  
50,  
40,  
30,  
91,  
48,  
68,  
64,  
93,  
28,  
85,  
92,  
62,  
45,  
74,  
44,  
70,  
49,  
94,  
38,  
77,  
69,  
37,  
100,  
50,  
65,  
35,  
90,  
43,  
74,  
77,  
100,

46,  
76,  
100,  
44,  
90,  
78,  
64,  
52,  
49,  
55,  
60,  
70,  
32,  
43,  
40,  
92,  
45,  
100,  
81,  
97,  
100,  
100,  
33,  
44,  
39,  
65,  
32,  
62,  
67,  
71,  
95,  
64,  
11,  
28,  
41,  
100,  
74,  
53,  
77,



64,  
59,  
63,  
53,  
74,  
79,  
100,  
66,  
56,  
90,  
44,  
38,  
93,  
54,  
79,  
74,  
75,  
57,  
62,  
77,  
56,  
55,  
91,  
61,  
86,  
83,  
89,  
42,  
60,  
39,  
94,  
68,  
78,  
43,  
100,  
45,  
46,  
82,  
70,

79,  
25,  
72,  
50,  
81,  
65,  
63,  
67,  
73,  
71,  
55,  
62,  
70,  
95,  
79,  
79,  
77,  
100,  
29,  
85,  
82,  
61,  
46,  
59,  
88,  
57,  
40,  
46,  
65,  
65,  
71,  
45,  
37,  
90,  
89,  
65,  
48,  
82,  
40,

56,  
40,  
84,  
78,  
69,  
70,  
40,  
84,  
44,  
59,  
100,  
34,  
81,  
69,  
37,  
100,  
100,  
75,  
52,  
82,  
36,  
78,  
100,  
72,  
85,  
83,  
68,  
80,  
81,  
59,  
78,  
100,  
76,  
59,  
91,  
77,  
93,  
34,  
24,

78,  
62,  
74,  
85,  
32,  
75,  
42,  
61,  
100,  
59,  
89,  
43,  
70,  
60,  
68,  
55,  
74,  
81,  
75,  
91,  
100,  
78,  
82,  
100,  
100,  
70,  
100,  
40,  
100,  
93,  
46,  
38,  
89,  
70,  
84,  
40,  
38,  
75,  
56,

```
62,  
73,  
...]
```

```
In [34]: df["dist"]=dist
```

```
In [35]: df.head()
```

```
Out[35]:
```

|   | company                               | address   | db  | dist |
|---|---------------------------------------|---|---|------|
| 0 | KIC Financial Brokerage KSC           | Kuwait City (Qibla), Fahad Alsalem Street, Awt... | Fahad Alsalem Street Awtad Tower Floor#15 K...    | 97   |
| 1 | Oman Agriculture Development Co SAOG  | P.O Box No. 2398, Al Ghubra, Postal Code 130, ... | Sun Farms, Near Penguin RestaurantAl Suwaihrah... | 31   |
| 2 | National Hospitality Institute SAOC   | Building No.900, Way No.63, Al Fursan Street, ... | P O Box 854:131AlHamriyaAl Wadi Al Kabir , Nr ... | 60   |
| 3 | Qatar Real Estate Investment PJSC     | Alaqaria Tower P.O. Box 22311, Old Salata area... | P.O.Box : 22311 Doha Qatar                        | 70   |
| 4 | EFG Hermes IFA Financial Brokerage Co | Kuwait Stock Exchange Building, 4th floor, P.O... | 4TH FLOOR EAST ENTRANCEP.O. BOX 22695STOCK EX...  | 83   |

## TAKING VALUES HAVING SIMILARITY MORE THAN 90 PERCENT

```
In [50]: similarity=df[df["dist"]>90]  
similarity
```

```
Out[50]:
```

|    | company                        | address   | db   | dist |
|----|--------------------------------|---|--|------|
| 0  | KIC Financial Brokerage KSC    | Kuwait City (Qibla), Fahad Alsalem Street, Awt... | Fahad Alsalem Street Awtad Tower Floor#15 K... | 97   |
| 37 | Al Maha Financial Services LLC | Office Number 21 – 22 Second Floor Royal Tower... | P.O. BOX 1065 MUSCAT Oman                      | 91   |

|      | company                                | address   | db  | dist |
|------|--|---|---|------|
| 41   | First Finance Co                       | P.O. Box 7258 First Finance Company Building 3... | P.O.Box 7258 Doha Qatar                             | 100  |
| 44   | Watheeqa Group Holding Co              | P.O : Box 28009, Safat 13141, Kuwait              | P.O. Box 28009 Safat Kuwait                         | 100  |
| 56   | United Securities - Oman               | Building no.: 1/299, Way no.: 5007, Compound n... | P O Box 2566 Ruwi PC 112 Oman                       | 95   |
| ...  | ...                                    | ...   | ...   | ...  |
| 1666 | Kuwait Banking Association             | 24th Floor, Al-Tijaria Tower, Al-Soor street, ... | Al-Soor Street Al-Tijaria Tower 24th FloorSh...     | 99   |
| 1678 | Trelco International WLL               | P.O.Box: 3988, Doha, Qatar.                       | P.O.Box 3988 Doha Qatar                             | 100  |
| 1679 | Alijarah Holding                       | Alijarah Holding, Alhilal - D Ring Road, P.O. ... | D Ring RoadP.O. Box: 24141 Doha Qatar               | 91   |
| 1683 | Qatar Science and Technology Park      | PO Box 210000, Qatar Science & Technology Park... | AL GHARRAF AH AREA P.O. BOX 5825QATAR SCIENCE AN... | 91   |
| 1687 | Al Waha Contracting and Trading Co WLL | P.O.Box:1486, Doha, Qatar                         | P.O.Box : 1486 Doha Qatar                           | 100  |

201 rows × 4 columns

## Values having lowest similarity

In [56]: `df[df["dist"]<70]`

Out[56]:

|   | company                              | address   | db  | dist |
|---|--------------------------------------|---|---|------|
| 1 | Oman Agriculture Development Co SAOG | P.O Box No. 2398, Al Ghubra, Postal Code 130, ... | Sun Farms, Near Penguin RestaurantAl Suwaihrah... | 31   |
| 2 | National Hospitality Institute SAOC  | Building No.900, Way No.63, Al Fursan Street, ... | P O Box 854:131AlHamriyaAl Wadi Al Kabir , Nr ... | 60   |

|      | company   | address   | db  | dist |
|------|---|---|---|------|
| 7    | Oman Hotels and Tourism SAOG                      | PO Box 2031 ,Postal Code 112, Ruwi, Muscat, Su... | Way 133, Al Farahidi StreetRuwi, Muscat - Ruwi... | 44   |
| 8    | Sohar Poultry SAOG                                | Building No. 1678, Al Mawaleh, Railway 4720, S... | P O Box 2808:112RuwiRuwi , Next Al Jadeeda, Me... | 45   |
| 10   | Al Ahlia Detergent Co ( Under Liquidation )       | P.O.Box: 18, Postal Code: 322, Sohar, Sultanat... | Road Number 5 Sohar Industrial Estate 145-14...   | 46   |
| ...  | ...   | ...   | ...   | ...  |
| 1696 | Hepworth PME Qatar WLL                            | P.O. Box: 50207, Mesaieed, Doha, Qatar.           | Industrial AreaBuilding no. 17 Mesaieed ...       | 60   |
| 1697 | Global Direct Television General Trading and C... | Shop 6 & 7 & 8, Building 47, P.O Box 23982, Sa... | 6 Ring Road,Near Tala CentreAl Farwaniya, Deje... | 39   |
| 1698 | Ta'allum Training and Education Group             | PO.Box: 4219, Doha, Qatar                         | F/1, Al Mutanna ComplexSalwa Road, Al AsiriGSS... | 37   |
| 1699 | Manar Al Omran Scaffolding Rent and Trading       | Entrance C, 1st Flr, Al Kuwari Bldg, Al Sadd S... | Barwa Commercial AvenueIndustrial Area Sou...     | 49   |
| 1701 | Morad Yousuf Behbehani Group                      | P.O. Box 146 Safat,13002 Safat, Kuwait.           | 4th Ring Road, Mohammed Bin Al Qasim StreetAl ... | 34   |

1000 rows × 4 columns

## VALUES WITH MODERATE SIMILARITIES

In [59]: `df[(df["dist"]>=70) & (df["dist"]<=90)]`

Out[59]:

|   | company                               | address   | db   | dist |
|---|---------------------------------------|---|--|------|
| 3 | Qatar Real Estate Investment PJSC     | Alaqaria Tower P.O. Box 22311, Old Salata area... | P.O.Box : 22311 Doha Qatar                       | 70   |
| 4 | EFG Hermes IFA Financial Brokerage Co | Kuwait Stock Exchange Building, 4th floor, P.O... | 4TH FLOOR EAST ENTRANCEP.O. BOX 22695STOCK EX... | 83   |

|      | company  | address  |   | db  | dist |
|------|--|--|---|-----|------|
| 5    | Oman Air SAOC                                    | PO Box 58 Postal Code 111<br>Muscat International... | P O Box 58:111CPOHead Quarters<br>Muscat Airport ...  |     | 72   |
| 6    | Dhofar Power Co                                  | 1st Floor, A'Sahwa Complex<br>Building, near Cham... | Near Chamber of Commerce<br>Salalah A'Sahwa Compl...  |     | 85   |
| 9    | Oman Medical Projects<br>Co                      | P.O. Box 8, Al Azaiba 130,<br>Oman                   | P.O. BOX 8MUSCAT PRIVATE<br>HOSPITAL BAUSHER STR...   |     | 88   |
| ...  | ...  | ...  | ...   | ... | ...  |
| 1688 | Addhia Trading and<br>Contracting CO LLC         | PO Box 1912, P C 112, Way<br>4006, Bldg 439, Flat... | Way 4006Al Ghubra NorthBuilding<br>439 Flat 102...    |     | 74   |
| 1690 | Burgan National<br>Information Systems Co<br>WLL | 2nd Floor, Building 4, Al<br>Murqab Area, Block 3... | Murqab, block 3, omar bin al-<br>khattab streetBui... |     | 86   |
| 1691 | Al Sayer Soft Drinks<br>Factory Co WLL           | Al Sayer Soft Drinks Factory<br>Company Building,... | Sabhan Industrial Area - 1East of<br>Internationa...  |     | 70   |
| 1695 | Scrap Real Estate Co LLC                         | P O Box: 25418 Safat 13115<br>Kuwait                 | P.O. Box 25418SAFAT Safat Kuwait                      |     | 83   |
| 1700 | International Inspection<br>Centre Co. W.L.L.    | East Ahmadi, Block 6, Plot<br>142, P.O Box 296, S... | Block- 6Plot -142, East Ahmadi<br>Kuwait City...      |     | 84   |

501 rows × 4 columns

In [ ]:

In [ ]: