ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT OECD.Stat

# K Means Clustering

OECD Better Life Index

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT OECD.Stat

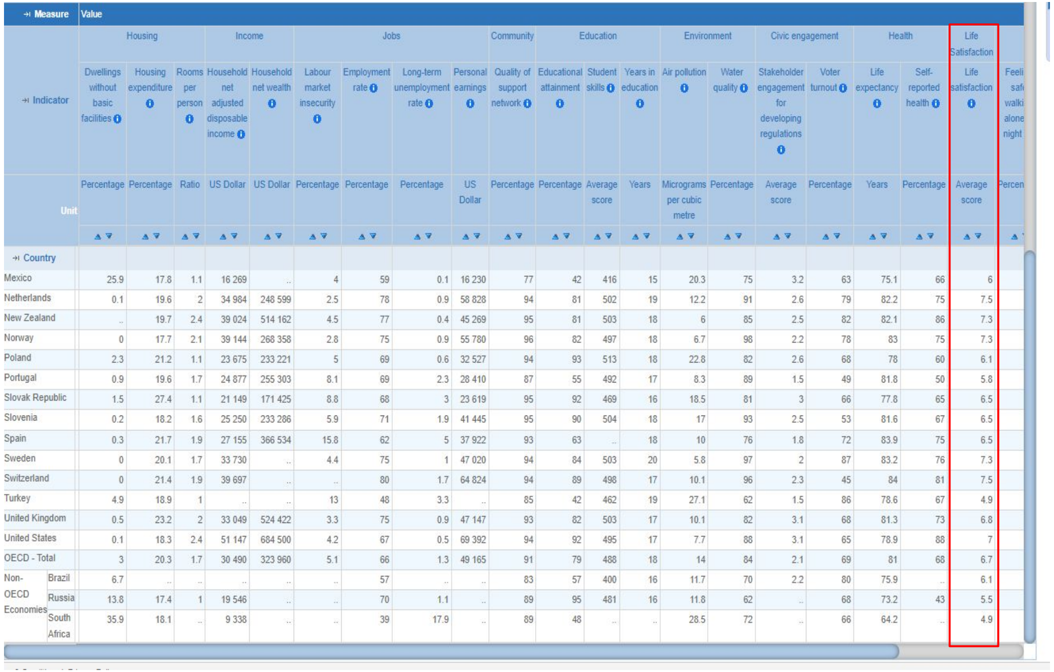
# Project Brief

### Problem statement

*“We want to maximise the Life Satisfaction across the OECD+ set of countries, reviewing what issues are potentially causing lags in Life Satisfaction worldwide.”*

### Defining ‘Life Satisfaction’

*“The indicator considers people's evaluation of their life as a whole. It is a weighted-sum of different response categories based on people's rates of their current life relative to the best and worst possible lives for them on a scale from 0 to 10, using the Cantril Ladder (known also as the "Self-Anchoring Striving Scale").”*



| Measure | Value | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Housing | | | Income | | Jobs | | | | Community | Education | | | Environment | | Civic engagement | | Health | | Life Satisfaction | |
| Indicator | Dwellings without basic facilities | Housing expenditure | Rooms per person | Household net adjusted disposable income | Household net wealth | Labour market insecurity | Employment rate | Long-term unemployment rate | Personal earnings | Quality of support network | Educational attainment | Student skills | Years in education | Air pollution | Water quality | Stakeholder engagement for developing regulations | Voter turnout | Life expectancy | Self-reported health | Life satisfaction | Feeli saf walki alone night |
| Unit | Percentage | Percentage | Ratio | US Dollar | US Dollar | Percentage | Percentage | Percentage | US Dollar | Percentage | Percentage | Average score | Years | Micrograms per cubic metre | Percentage | Average score | Percentage | Years | Percentage | Average score | Percen |
| Country | | | | | | | | | | | | | | | | | | | | | |
| Mexico | 25.9 | 17.8 | 1.1 | 16 269 | .. | 4 | 59 | 0.1 | 16 230 | 77 | 42 | 416 | 15 | 20.3 | 75 | 3.2 | 63 | 75.1 | 66 | 6 | |
| Netherlands | 0.1 | 19.6 | 2 | 34 984 | 248 599 | 2.5 | 78 | 0.9 | 58 828 | 94 | 81 | 502 | 19 | 12.2 | 91 | 2.6 | 79 | 82.2 | 75 | 7.5 | |
| New Zealand | .. | 19.7 | 2.4 | 39 024 | 514 162 | 4.5 | 77 | 0.4 | 45 269 | 95 | 81 | 503 | 18 | 6 | 85 | 2.5 | 82 | 82.1 | 86 | 7.3 | |
| Norway | 0 | 17.7 | 2.1 | 39 144 | 268 358 | 2.8 | 75 | 0.9 | 55 780 | 96 | 82 | 497 | 18 | 6.7 | 98 | 2.2 | 78 | 83 | 75 | 7.3 | |
| Poland | 2.3 | 21.2 | 1.1 | 23 675 | 233 221 | 5 | 69 | 0.6 | 32 527 | 94 | 93 | 513 | 18 | 22.8 | 82 | 2.6 | 68 | 78 | 60 | 6.1 | |
| Portugal | 0.9 | 19.6 | 1.7 | 24 877 | 255 303 | 8.1 | 69 | 2.3 | 28 410 | 87 | 55 | 492 | 17 | 8.3 | 89 | 1.5 | 49 | 81.8 | 50 | 5.8 | |
| Slovak Republic | 1.5 | 27.4 | 1.1 | 21 149 | 171 425 | 8.8 | 68 | 3 | 23 619 | 95 | 92 | 469 | 16 | 18.5 | 81 | 3 | 66 | 77.8 | 65 | 6.5 | |
| Slovenia | 0.2 | 18.2 | 1.6 | 25 250 | 233 286 | 5.9 | 71 | 1.9 | 41 445 | 95 | 90 | 504 | 18 | 17 | 93 | 2.5 | 53 | 81.6 | 67 | 6.5 | |
| Spain | 0.3 | 21.7 | 1.9 | 27 155 | 366 534 | 15.8 | 62 | 5 | 37 922 | 93 | 63 | .. | 18 | 10 | 76 | 1.8 | 72 | 83.9 | 75 | 6.5 | |
| Sweden | 0 | 20.1 | 1.7 | 33 730 | .. | 4.4 | 75 | 1 | 47 020 | 94 | 84 | 503 | 20 | 5.8 | 97 | 2 | 87 | 83.2 | 76 | 7.3 | |
| Switzerland | 0 | 21.4 | 1.9 | 39 697 | .. | .. | 80 | 1.7 | 64 824 | 94 | 89 | 498 | 17 | 10.1 | 96 | 2.3 | 45 | 84 | 81 | 7.5 | |
| Turkey | 4.9 | 18.9 | 1 | .. | .. | 13 | 48 | 3.3 | .. | 85 | 42 | 462 | 19 | 27.1 | 62 | 1.5 | 86 | 78.6 | 67 | 4.9 | |
| United Kingdom | 0.5 | 23.2 | 2 | 33 049 | 524 422 | 3.3 | 75 | 0.9 | 47 147 | 93 | 82 | 503 | 17 | 10.1 | 82 | 3.1 | 68 | 81.3 | 73 | 6.8 | |
| United States | 0.1 | 18.3 | 2.4 | 51 147 | 684 500 | 4.2 | 67 | 0.5 | 69 392 | 94 | 92 | 495 | 17 | 7.7 | 88 | 3.1 | 65 | 78.9 | 88 | 7 | |
| OECD - Total | 3 | 20.3 | 1.7 | 30 490 | 323 960 | 5.1 | 66 | 1.3 | 49 165 | 91 | 79 | 488 | 18 | 14 | 84 | 2.1 | 69 | 81 | 68 | 6.7 | |
| Non-OECD Economies Brazil | 6.7 | .. | .. | .. | .. | .. | 57 | .. | .. | 83 | 57 | 400 | 16 | 11.7 | 70 | 2.2 | 80 | 75.9 | .. | 6.1 | |
| Russia | 13.8 | 17.4 | 1 | 19 546 | .. | .. | 70 | 1.1 | .. | 89 | 95 | 481 | 16 | 11.8 | 62 | .. | 68 | 73.2 | 43 | 5.5 | |
| South Africa | 35.9 | 18.1 | .. | 9 338 | .. | .. | 39 | 17.9 | .. | 89 | 48 | .. | .. | 28.5 | 72 | .. | 66 | 64.2 | .. | 4.9 | |

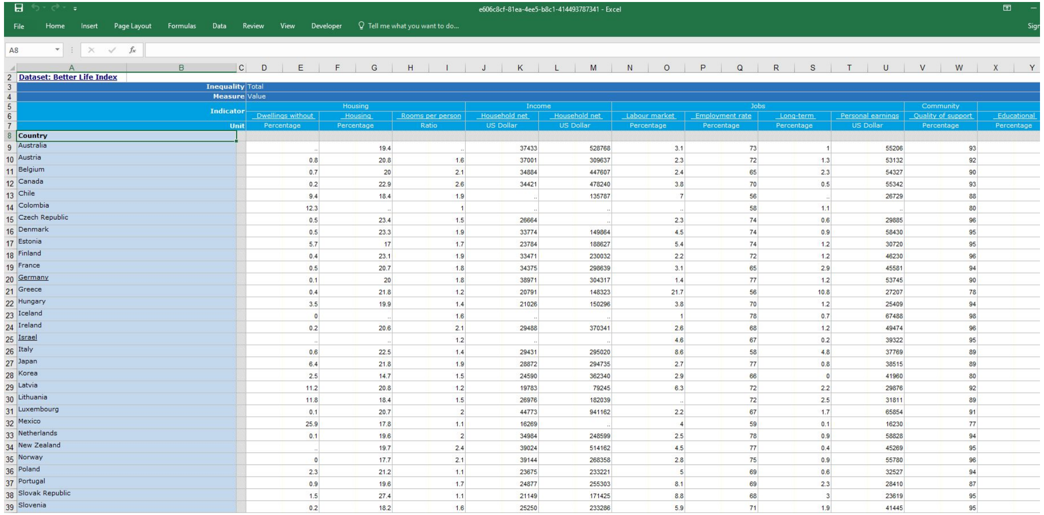
# Summary of methodology

- Using a K Means Clustering Algorithm to Group common countries together.
- Evaluating clusters relative to their ‘Life Satisfaction’ Score.
- Exploring why Countries have been grouped together, and what common issues may be lowering their Life Satisfaction Score.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT OECD.Stat

# Acquiring data and data cleaning

- Reformatted to machine friendly CSV format.
- Imported data came with a number of *null* values, meaning data loaded as ‘objects’
- Once strategy on handling *nulls* was decided, converted data types to Float64 and assigned countries to index.



Dataset: Better Life Index

Inequality Total

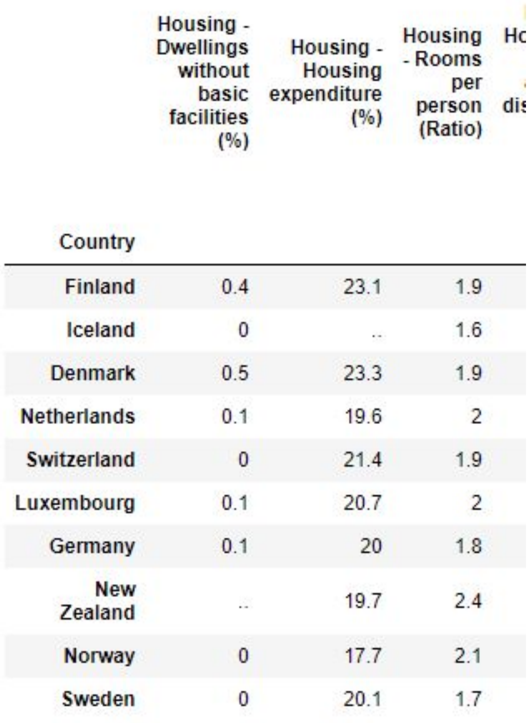
Measure Value

| Indicator | Housing | | | Income | | Jobs | | | | Community | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Dwellings without | Housing | Rooms per person | Household net | Household net | Labour market | Employment rate | Long-term | Personal earnings | Quality of support | Educational |
| Unit | Percentage | Percentage | Ratio | US Dollar | US Dollar | Percentage | Percentage | Percentage | US Dollar | Percentage | Percentage |
| Country | | | | | | | | | | | |
| Australia | .. | 19.4 | .. | 37433 | 528768 | 3.1 | 73 | 1 | 55206 | 93 | |
| Austria | 0.8 | 20.8 | 1.6 | 37001 | 309637 | 2.3 | 72 | 1.3 | 53132 | 92 | |
| Belgium | 0.7 | 20 | 2.1 | 34884 | 447607 | 2.4 | 65 | 2.3 | 54327 | 90 | |
| Canada | 0.2 | 22.9 | 2.6 | 34421 | 478240 | 3.8 | 70 | 0.5 | 55342 | 93 | |
| Chile | 9.4 | 18.4 | 1.9 | .. | 135787 | 7 | 56 | .. | 26729 | 88 | |
| Colombia | 12.3 | .. | 1 | .. | .. | .. | 58 | 1.1 | .. | 80 | |
| Czech Republic | 0.5 | 23.4 | 1.5 | 26664 | .. | 2.3 | 74 | 0.6 | 29885 | 96 | |
| Denmark | 0.5 | 23.3 | 1.9 | 33774 | 149864 | 4.5 | 74 | 0.9 | 58430 | 95 | |
| Estonia | 5.7 | 17 | 1.7 | 23784 | 188627 | 5.4 | 74 | 1.2 | 30720 | 95 | |
| Finland | 0.4 | 23.1 | 1.9 | 33471 | 230032 | 2.2 | 72 | 1.2 | 46230 | 96 | |
| France | 0.5 | 20.7 | 1.8 | 34375 | 298639 | 3.1 | 65 | 2.9 | 45581 | 94 | |
| Germany | 0.1 | 20 | 1.8 | 38971 | 304317 | 1.4 | 77 | 1.2 | 53745 | 90 | |
| Greece | 0.4 | 21.8 | 1.2 | 20791 | 148323 | 21.7 | 56 | 10.8 | 27207 | 78 | |
| Hungary | 3.5 | 19.9 | 1.4 | 21026 | 150296 | 3.8 | 70 | 1.2 | 25409 | 94 | |
| Iceland | 0 | .. | 1.6 | .. | .. | 1 | 78 | 0.7 | 67488 | 98 | |
| Ireland | 0.2 | 20.6 | 2.1 | 29488 | 370341 | 2.6 | 68 | 1.2 | 49474 | 96 | |
| Israel | .. | .. | 1.2 | .. | .. | 4.6 | 67 | 0.2 | 39322 | 95 | |
| Italy | 0.6 | 22.5 | 1.4 | 29431 | 295020 | 8.6 | 58 | 4.8 | 37769 | 89 | |
| Japan | 6.4 | 21.8 | 1.9 | 28872 | 294735 | 2.7 | 77 | 0.8 | 38515 | 89 | |
| Korea | 2.5 | 14.7 | 1.5 | 24590 | 362340 | 2.9 | 66 | 0 | 41960 | 80 | |
| Latvia | 11.2 | 20.8 | 1.2 | 19783 | 79245 | 6.3 | 72 | 2.2 | 29876 | 92 | |
| Lithuania | 11.8 | 18.4 | 1.5 | 26976 | 182039 | .. | 72 | 2.5 | 31811 | 89 | |
| Luxembourg | 0.1 | 20.7 | 2 | 44773 | 941162 | 2.2 | 67 | 1.7 | 65854 | 91 | |
| Mexico | 25.9 | 17.8 | 1.1 | 16269 | .. | 4 | 59 | 0.1 | 16230 | 77 | |
| Netherlands | 0.1 | 19.6 | 2 | 34984 | 248599 | 2.5 | 78 | 0.9 | 58828 | 94 | |
| New Zealand | .. | 19.7 | 2.4 | 39024 | 514162 | 4.5 | 77 | 0.4 | 45269 | 95 | |
| Norway | 0 | 17.7 | 2.1 | 39144 | 268358 | 2.8 | 75 | 0.9 | 55780 | 96 | |
| Poland | 2.3 | 21.2 | 1.1 | 23675 | 233221 | 5 | 69 | 0.6 | 32527 | 94 | |
| Portugal | 0.9 | 19.6 | 1.7 | 24877 | 255303 | 8.1 | 69 | 2.3 | 28410 | 87 | |
| Slovak Republic | 1.5 | 27.4 | 1.1 | 21149 | 171425 | 8.8 | 68 | 3 | 23619 | 95 | |
| Slovenia | 0.2 | 18.2 | 1.6 | 25250 | 233286 | 5.9 | 71 | 1.9 | 41445 | 95 | |

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT OECD.Stat

# Problem handling null values

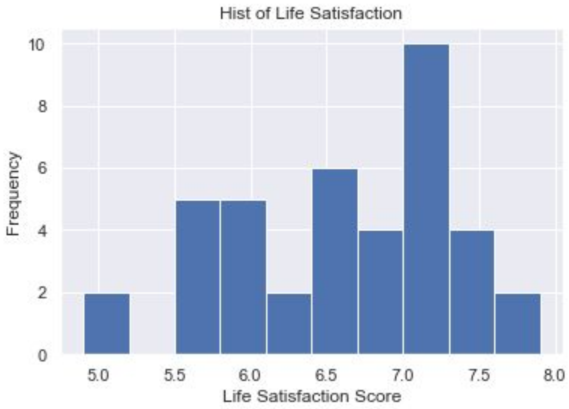
1) Ignore rows - Not viable due to the size of data set (40 Rows)
2) Impute values:
   a) Take Average of the dataset - outlier sensitive (countries like Mexico, South Africa and Russia have significant outlier values)
   b) KNN Classification - attempted a classification, but end result did not accurately model which countries are most like one another once tested.
   c) Replace Null with a ‘0’ value - *taken as least worst option*

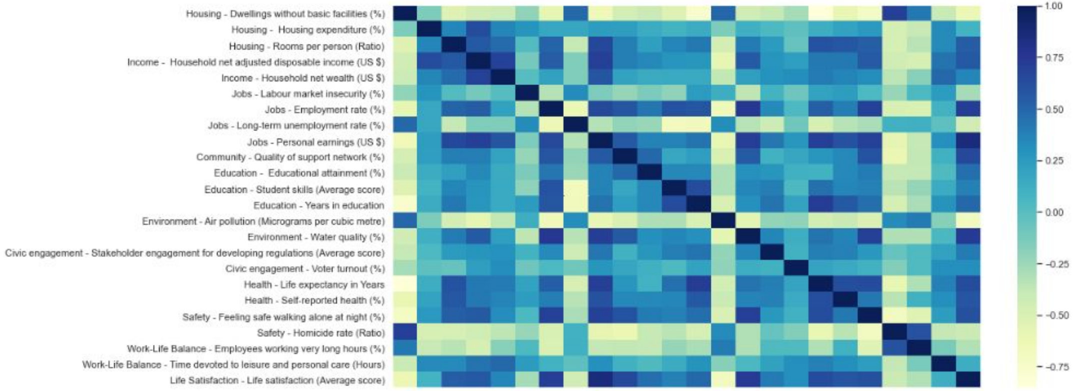


|  | Housing - Dwellings without basic facilities (%) | Housing - Housing expenditure (%) | Housing - Rooms per person (Ratio) | Hc dis |
|---|---|---|---|---|
| Country |  |  |  |  |
| Finland | 0.4 | 23.1 | 1.9 |  |
| Iceland | 0 | .. | 1.6 |  |
| Denmark | 0.5 | 23.3 | 1.9 |  |
| Netherlands | 0.1 | 19.6 | 2 |  |
| Switzerland | 0 | 21.4 | 1.9 |  |
| Luxembourg | 0.1 | 20.7 | 2 |  |
| Germany | 0.1 | 20 | 1.8 |  |
| New Zealand | .. | 19.7 | 2.4 |  |
| Norway | 0 | 17.7 | 2.1 |  |
| Sweden | 0 | 20.1 | 1.7 |  |

# Exploratory Analysis

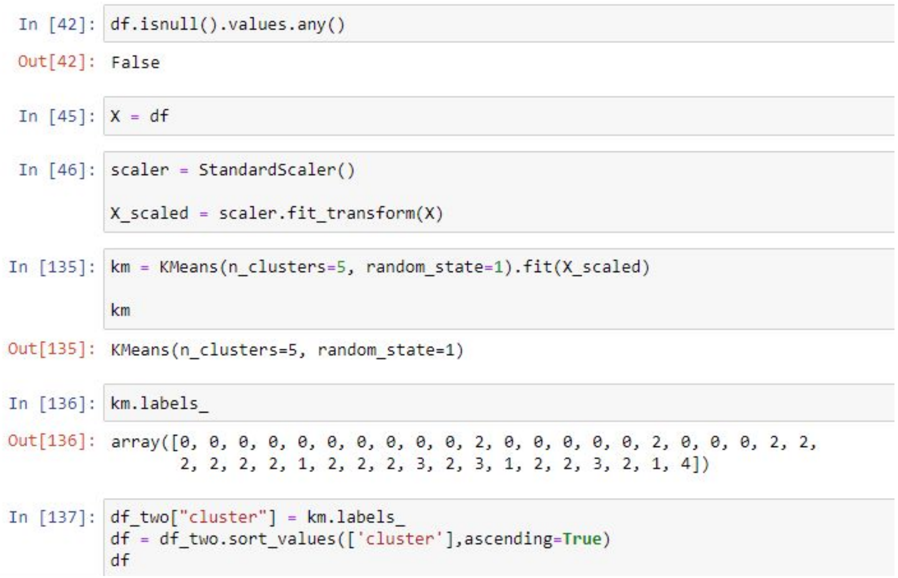
- Reviewed distribution of life satisfaction scores, general positive skew to the right.
- And potential correlations between Life Satisfaction to be reviewed after clustering.
- Economic features around Income, Wealth and Employment highly correlated to Satisfaction Score





# Creating the model: K Means cluster

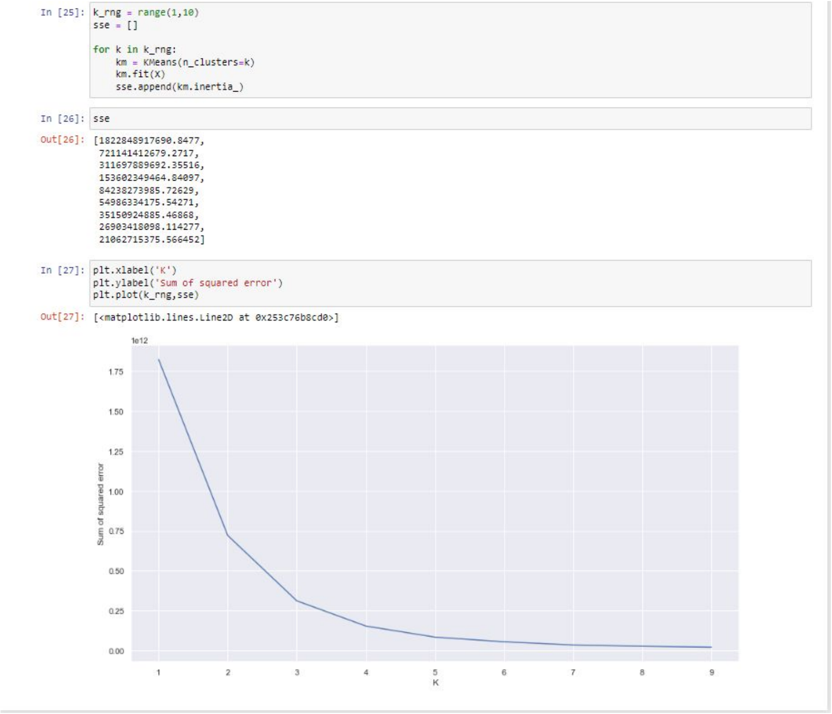
- Once dataset cleaned, assigned dataframe to X, ensured values were scaled and initiated the K Means model.
- Confirmed the K value I wanted (after some trial and error).
- Assigned the clustering labels to the new dataframe and exported as a final excel file.



```
In [42]:  df.isnull().values.any()

Out[42]:  False

In [45]:  X = df

In [46]:  scaler = StandardScaler()

          X_scaled = scaler.fit_transform(X)

In [135]: km = KMeans(n_clusters=5, random_state=1).fit(X_scaled)

          km

Out[135]: KMeans(n_clusters=5, random_state=1)

In [136]: km.labels_

Out[136]: array([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 2, 0, 0, 0, 2, 2,
                 2, 2, 2, 2, 1, 2, 2, 2, 3, 2, 3, 1, 2, 2, 3, 2, 1, 4])

In [137]: df_two["cluster"] = km.labels_
          df = df_two.sort_values(['cluster'],ascending=True)
          df
```

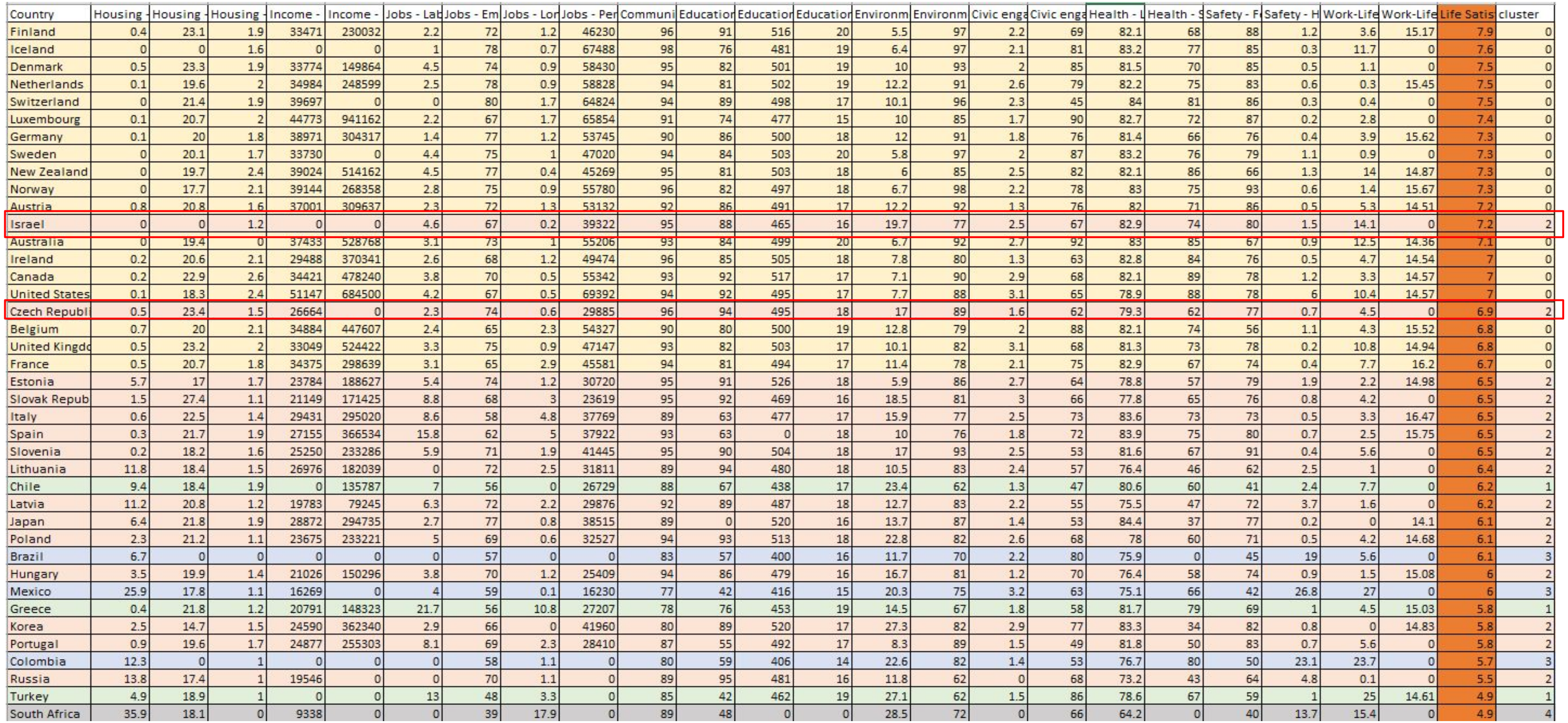
# Challenges in confirming the correct K value

- Originally used the ‘Elbow Technique’ as a guide, result recommended a K value of 3.
- However, clusters appeared to be too homogenous with Income/Wealth levels having too-high a weighting.
- After experimentation, chose a K-Value of 5.
- Higher number of clusters revealed better insights, despite size of dataset.



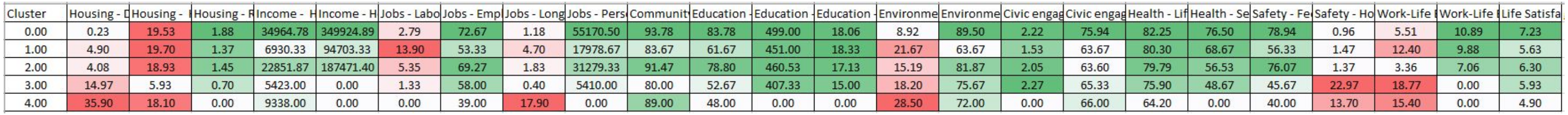
```
In [25]: k_rng = range(1,10)
         sse = []

         for k in k_rng:
             km = KMeans(n_clusters=k)
             km.fit(X)
             sse.append(km.inertia_)

In [26]: sse

Out[26]: [1822848917690.8477,
          721141412679.2717,
          311697889692.35516,
          153602349464.84097,
          84238273985.72629,
          54986334175.54271,
          35150924885.46868,
          26903418098.114277,
          21062715375.566452]

In [27]: plt.xlabel('K')
         plt.ylabel('Sum of squared error')
         plt.plot(k_rng,sse)

Out[27]: [<matplotlib.lines.Line2D at 0x253c76b8cd0>]
```

# Results Overview - Clusters Vs. Life Satisfaction Score



| Country | Housing - | Housing - | Housing - | Income - | Income - | Jobs - Lab | Jobs - Em | Jobs - Lor | Jobs - Per | Communi | Educatior | Educatior | Educatior | Environm | Environm | Civic enga | Civic enga | Health - L | Health - S | Safety - F | Safety - H | Work-Life | Work-Life | Life Satis | cluster |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Finland | 0.4 | 23.1 | 1.9 | 33471 | 230032 | 2.2 | 72 | 1.2 | 46230 | 96 | 91 | 516 | 20 | 5.5 | 97 | 2.2 | 69 | 82.1 | 68 | 88 | 1.2 | 3.6 | 15.17 | 7.9 | 0 |
| Iceland | 0 | 0 | 1.6 | 0 | 0 | 1 | 78 | 0.7 | 67488 | 98 | 76 | 481 | 19 | 6.4 | 97 | 2.1 | 81 | 83.2 | 77 | 85 | 0.3 | 11.7 | 0 | 7.6 | 0 |
| Denmark | 0.5 | 23.3 | 1.9 | 33774 | 149864 | 4.5 | 74 | 0.9 | 58430 | 95 | 82 | 501 | 19 | 10 | 93 | 2 | 85 | 81.5 | 70 | 85 | 0.5 | 1.1 | 0 | 7.5 | 0 |
| Netherlands | 0.1 | 19.6 | 2 | 34984 | 248599 | 2.5 | 78 | 0.9 | 58828 | 94 | 81 | 502 | 19 | 12.2 | 91 | 2.6 | 79 | 82.2 | 75 | 83 | 0.6 | 0.3 | 15.45 | 7.5 | 0 |
| Switzerland | 0 | 21.4 | 1.9 | 39697 | 0 | 0 | 80 | 1.7 | 64824 | 94 | 89 | 498 | 17 | 10.1 | 96 | 2.3 | 45 | 84 | 81 | 86 | 0.3 | 0.4 | 0 | 7.5 | 0 |
| Luxembourg | 0.1 | 20.7 | 2 | 44773 | 941162 | 2.2 | 67 | 1.7 | 65854 | 91 | 74 | 477 | 15 | 10 | 85 | 1.7 | 90 | 82.7 | 72 | 87 | 0.2 | 2.8 | 0 | 7.4 | 0 |
| Germany | 0.1 | 20 | 1.8 | 38971 | 304317 | 1.4 | 77 | 1.2 | 53745 | 90 | 86 | 500 | 18 | 12 | 91 | 1.8 | 76 | 81.4 | 66 | 76 | 0.4 | 3.9 | 15.62 | 7.3 | 0 |
| Sweden | 0 | 20.1 | 1.7 | 33730 | 0 | 4.4 | 75 | 1 | 47020 | 94 | 84 | 503 | 20 | 5.8 | 97 | 2 | 87 | 83.2 | 76 | 79 | 1.1 | 0.9 | 0 | 7.3 | 0 |
| New Zealand | 0 | 19.7 | 2.4 | 39024 | 514162 | 4.5 | 77 | 0.4 | 45269 | 95 | 81 | 503 | 18 | 6 | 85 | 2.5 | 82 | 82.1 | 86 | 66 | 1.3 | 14 | 14.87 | 7.3 | 0 |
| Norway | 0 | 17.7 | 2.1 | 39144 | 268358 | 2.8 | 75 | 0.9 | 55780 | 96 | 82 | 497 | 18 | 6.7 | 98 | 2.2 | 78 | 83 | 75 | 93 | 0.6 | 1.4 | 15.67 | 7.3 | 0 |
| Austria | 0.8 | 20.8 | 1.6 | 37001 | 309637 | 2.3 | 72 | 1.3 | 53132 | 92 | 86 | 491 | 17 | 12.2 | 92 | 1.3 | 76 | 82 | 71 | 86 | 0.5 | 5.3 | 14.51 | 7.2 | 0 |
| Israel | 0 | 0 | 1.2 | 0 | 0 | 4.6 | 67 | 0.2 | 39322 | 95 | 88 | 465 | 16 | 19.7 | 77 | 2.5 | 67 | 82.9 | 74 | 80 | 1.5 | 14.1 | 0 | 7.2 | 2 |
| Australia | 0 | 19.4 | 0 | 37433 | 528768 | 3.1 | 73 | 1 | 55206 | 93 | 84 | 499 | 20 | 6.7 | 92 | 2.7 | 92 | 83 | 85 | 67 | 0.9 | 12.5 | 14.36 | 7.1 | 0 |
| Ireland | 0.2 | 20.6 | 2.1 | 29488 | 370341 | 2.6 | 68 | 1.2 | 49474 | 96 | 85 | 505 | 18 | 7.8 | 80 | 1.3 | 63 | 82.8 | 84 | 76 | 0.5 | 4.7 | 14.54 | 7 | 0 |
| Canada | 0.2 | 22.9 | 2.6 | 34421 | 478240 | 3.8 | 70 | 0.5 | 55342 | 93 | 92 | 517 | 17 | 7.1 | 90 | 2.9 | 68 | 82.1 | 89 | 78 | 1.2 | 3.3 | 14.57 | 7 | 0 |
| United States | 0.1 | 18.3 | 2.4 | 51147 | 684500 | 4.2 | 67 | 0.5 | 69392 | 94 | 92 | 495 | 17 | 7.7 | 88 | 3.1 | 65 | 78.9 | 88 | 78 | 6 | 10.4 | 14.57 | 7 | 0 |
| Czech Republi | 0.5 | 23.4 | 1.5 | 26664 | 0 | 2.3 | 74 | 0.6 | 29885 | 96 | 94 | 495 | 18 | 17 | 89 | 1.6 | 62 | 79.3 | 62 | 77 | 0.7 | 4.5 | 0 | 6.9 | 2 |
| Belgium | 0.7 | 20 | 2.1 | 34884 | 447607 | 2.4 | 65 | 2.3 | 54327 | 90 | 80 | 500 | 19 | 12.8 | 79 | 2 | 88 | 82.1 | 74 | 56 | 1.1 | 4.3 | 15.52 | 6.8 | 0 |
| United Kingdo | 0.5 | 23.2 | 2 | 33049 | 524422 | 3.3 | 75 | 0.9 | 47147 | 93 | 82 | 503 | 17 | 10.1 | 82 | 3.1 | 68 | 81.3 | 73 | 78 | 0.2 | 10.8 | 14.94 | 6.8 | 0 |
| France | 0.5 | 20.7 | 1.8 | 34375 | 298639 | 3.1 | 65 | 2.9 | 45581 | 94 | 81 | 494 | 17 | 11.4 | 78 | 2.1 | 75 | 82.9 | 67 | 74 | 0.4 | 7.7 | 16.2 | 6.7 | 0 |
| Estonia | 5.7 | 17 | 1.7 | 23784 | 188627 | 5.4 | 74 | 1.2 | 30720 | 95 | 91 | 526 | 18 | 5.9 | 86 | 2.7 | 64 | 78.8 | 57 | 79 | 1.9 | 2.2 | 14.98 | 6.5 | 2 |
| Slovak Repub | 1.5 | 27.4 | 1.1 | 21149 | 171425 | 8.8 | 68 | 3 | 23619 | 95 | 92 | 469 | 16 | 18.5 | 81 | 3 | 66 | 77.8 | 65 | 76 | 0.8 | 4.2 | 0 | 6.5 | 2 |
| Italy | 0.6 | 22.5 | 1.4 | 29431 | 295020 | 8.6 | 58 | 4.8 | 37769 | 89 | 63 | 477 | 17 | 15.9 | 77 | 2.5 | 73 | 83.6 | 73 | 73 | 0.5 | 3.3 | 16.47 | 6.5 | 2 |
| Spain | 0.3 | 21.7 | 1.9 | 27155 | 366534 | 15.8 | 62 | 5 | 37922 | 93 | 63 | 0 | 18 | 10 | 76 | 1.8 | 72 | 83.9 | 75 | 80 | 0.7 | 2.5 | 15.75 | 6.5 | 2 |
| Slovenia | 0.2 | 18.2 | 1.6 | 25250 | 233286 | 5.9 | 71 | 1.9 | 41445 | 95 | 90 | 504 | 18 | 17 | 93 | 2.5 | 53 | 81.6 | 67 | 91 | 0.4 | 5.6 | 0 | 6.5 | 2 |
| Lithuania | 11.8 | 18.4 | 1.5 | 26976 | 182039 | 0 | 72 | 2.5 | 31811 | 89 | 94 | 480 | 18 | 10.5 | 83 | 2.4 | 57 | 76.4 | 46 | 62 | 2.5 | 1 | 0 | 6.4 | 2 |
| Chile | 9.4 | 18.4 | 1.9 | 0 | 135787 | 7 | 56 | 0 | 26729 | 88 | 67 | 438 | 17 | 23.4 | 62 | 1.3 | 47 | 80.6 | 60 | 41 | 2.4 | 7.7 | 0 | 6.2 | 1 |
| Latvia | 11.2 | 20.8 | 1.2 | 19783 | 79245 | 6.3 | 72 | 2.2 | 29876 | 92 | 89 | 487 | 18 | 12.7 | 83 | 2.2 | 55 | 75.5 | 47 | 72 | 3.7 | 1.6 | 0 | 6.2 | 2 |
| Japan | 6.4 | 21.8 | 1.9 | 28872 | 294735 | 2.7 | 77 | 0.8 | 38515 | 89 | 0 | 520 | 16 | 13.7 | 87 | 1.4 | 53 | 84.4 | 37 | 77 | 0.2 | 0 | 14.1 | 6.1 | 2 |
| Poland | 2.3 | 21.2 | 1.1 | 23675 | 233221 | 5 | 69 | 0.6 | 32527 | 94 | 93 | 513 | 18 | 22.8 | 82 | 2.6 | 68 | 78 | 60 | 71 | 0.5 | 4.2 | 14.68 | 6.1 | 2 |
| Brazil | 6.7 | 0 | 0 | 0 | 0 | 0 | 57 | 0 | 0 | 83 | 57 | 400 | 16 | 11.7 | 70 | 2.2 | 80 | 75.9 | 0 | 45 | 19 | 5.6 | 0 | 6.1 | 3 |
| Hungary | 3.5 | 19.9 | 1.4 | 21026 | 150296 | 3.8 | 70 | 1.2 | 25409 | 94 | 86 | 479 | 16 | 16.7 | 81 | 1.2 | 70 | 76.4 | 58 | 74 | 0.9 | 1.5 | 15.08 | 6 | 2 |
| Mexico | 25.9 | 17.8 | 1.1 | 16269 | 0 | 4 | 59 | 0.1 | 16230 | 77 | 42 | 416 | 15 | 20.3 | 75 | 3.2 | 63 | 75.1 | 66 | 42 | 26.8 | 27 | 0 | 6 | 3 |
| Greece | 0.4 | 21.8 | 1.2 | 20791 | 148323 | 21.7 | 56 | 10.8 | 27207 | 78 | 76 | 453 | 19 | 14.5 | 67 | 1.8 | 58 | 81.7 | 79 | 69 | 1 | 4.5 | 15.03 | 5.8 | 1 |
| Korea | 2.5 | 14.7 | 1.5 | 24590 | 362340 | 2.9 | 66 | 0 | 41960 | 80 | 89 | 520 | 17 | 27.3 | 82 | 2.9 | 77 | 83.3 | 34 | 82 | 0.8 | 0 | 14.83 | 5.8 | 2 |
| Portugal | 0.9 | 19.6 | 1.7 | 24877 | 255303 | 8.1 | 69 | 2.3 | 28410 | 87 | 55 | 492 | 17 | 8.3 | 89 | 1.5 | 49 | 81.8 | 50 | 83 | 0.7 | 5.6 | 0 | 5.8 | 2 |
| Colombia | 12.3 | 0 | 1 | 0 | 0 | 0 | 58 | 1.1 | 0 | 80 | 59 | 406 | 14 | 22.6 | 82 | 1.4 | 53 | 76.7 | 80 | 50 | 23.1 | 23.7 | 0 | 5.7 | 3 |
| Russia | 13.8 | 17.4 | 1 | 19546 | 0 | 0 | 70 | 1.1 | 0 | 89 | 95 | 481 | 16 | 11.8 | 62 | 0 | 68 | 73.2 | 43 | 64 | 4.8 | 0.1 | 0 | 5.5 | 2 |
| Turkey | 4.9 | 18.9 | 1 | 0 | 0 | 13 | 48 | 3.3 | 0 | 85 | 42 | 462 | 19 | 27.1 | 62 | 1.5 | 86 | 78.6 | 67 | 59 | 1 | 25 | 14.61 | 4.9 | 1 |
| South Africa | 35.9 | 18.1 | 0 | 9338 | 0 | 0 | 39 | 17.9 | 0 | 89 | 48 | 0 | 0 | 28.5 | 72 | 0 | 66 | 64.2 | 0 | 40 | 13.7 | 15.4 | 0 | 4.9 | 4 |

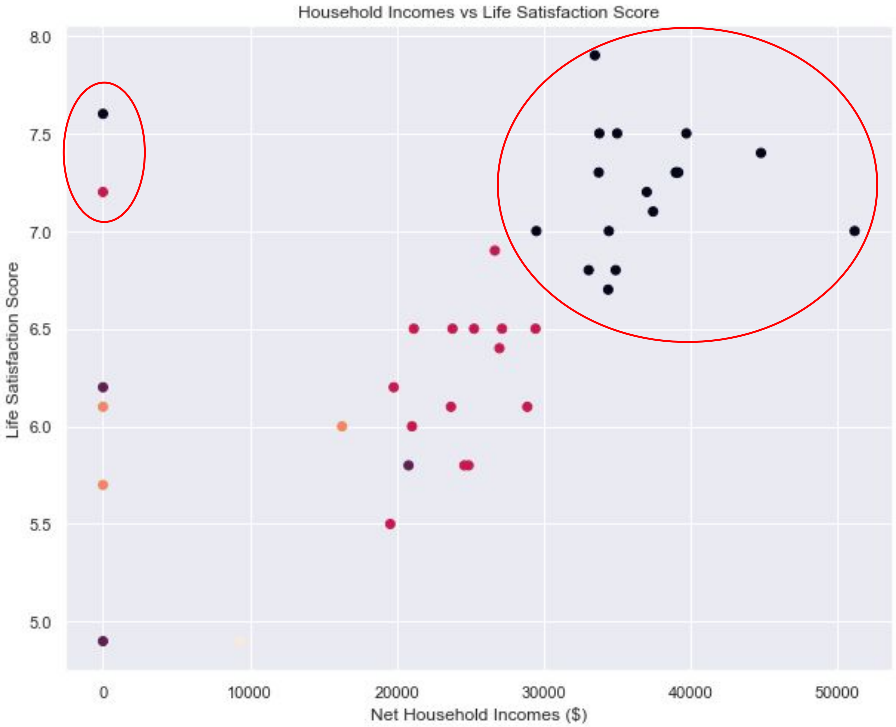
# Results Overview - Cluster Means



| Cluster | Housing - D | Housing - | Housing - R | Income - H | Income - H | Jobs - Labo | Jobs - Empl | Jobs - Long | Jobs - Pers | Communit | Education - | Education - | Education - | Environme | Environme | Civic engag | Civic engag | Health - Lif | Health - Se | Safety - Fe | Safety - Ho | Work-Life | Work-Life | Life Satisfa |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.00 | 0.23 | 19.53 | 1.88 | 34964.78 | 349924.89 | 2.79 | 72.67 | 1.18 | 55170.50 | 93.78 | 83.78 | 499.00 | 18.06 | 8.92 | 89.50 | 2.22 | 75.94 | 82.25 | 76.50 | 78.94 | 0.96 | 5.51 | 10.89 | 7.23 |
| 1.00 | 4.90 | 19.70 | 1.37 | 6930.33 | 94703.33 | 13.90 | 53.33 | 4.70 | 17978.67 | 83.67 | 61.67 | 451.00 | 18.33 | 21.67 | 63.67 | 1.53 | 63.67 | 80.30 | 68.67 | 56.33 | 1.47 | 12.40 | 9.88 | 5.63 |
| 2.00 | 4.08 | 18.93 | 1.45 | 22851.87 | 187471.40 | 5.35 | 69.27 | 1.83 | 31279.33 | 91.47 | 78.80 | 460.53 | 17.13 | 15.19 | 81.87 | 2.05 | 63.60 | 79.79 | 56.53 | 76.07 | 1.37 | 3.36 | 7.06 | 6.30 |
| 3.00 | 14.97 | 5.93 | 0.70 | 5423.00 | 0.00 | 1.33 | 58.00 | 0.40 | 5410.00 | 80.00 | 52.67 | 407.33 | 15.00 | 18.20 | 75.67 | 2.27 | 65.33 | 75.90 | 48.67 | 45.67 | 22.97 | 18.77 | 0.00 | 5.93 |
| 4.00 | 35.90 | 18.10 | 0.00 | 9338.00 | 0.00 | 0.00 | 39.00 | 17.90 | 0.00 | 89.00 | 48.00 | 0.00 | 0.00 | 28.50 | 72.00 | 0.00 | 66.00 | 64.20 | 0.00 | 40.00 | 13.70 | 15.40 | 0.00 | 4.90 |

# Cluster [0] Developed / Rich Economies with high income, wealth and employment

- The cluster we would ideally like everyone to be - *Highest Life Satisfaction*, with a clear distinction between them and the other countries.
- Cluster outperformed nearly all others across key index points (including: income, wealth and employment).
- Two outliers who has a *null* Income level, Iceland which was correctly classified into [0], and below this: Israel.
- Made up of western and northern European countries + North America.



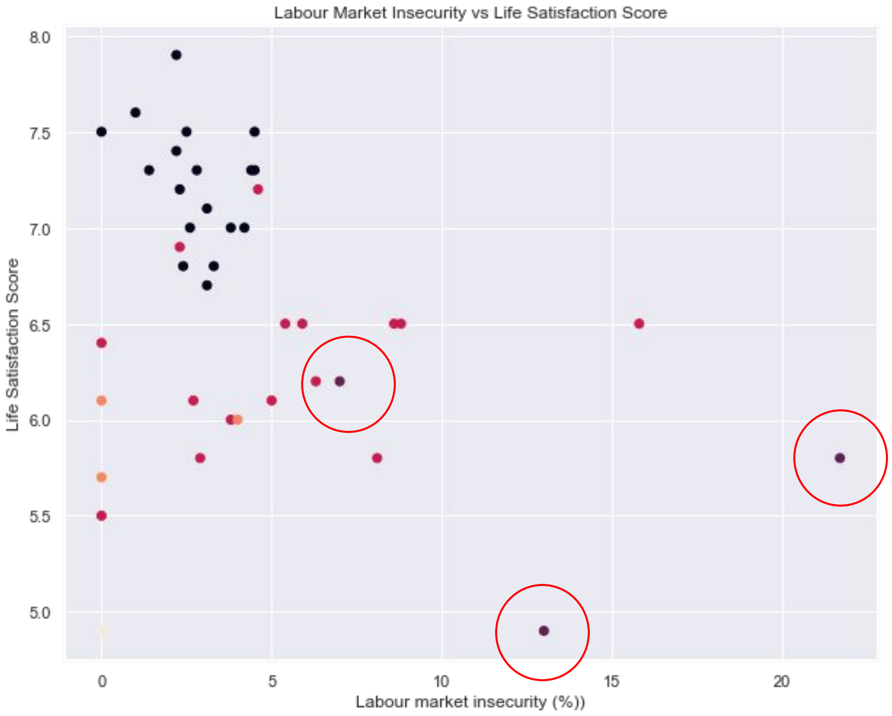
# Cluster [1] Chile, Turkey and Greece (Outlier States?)

- This cluster groups together Turkey, Greece and Chile.

- Cluster grouped together due to sharing a high-labour market insecurity rate, coupled with a high Housing expenditure (%), and second highest air pollution.

- Chile may have been misclassified here, as its labour market security is generally stronger. However its Housing Expenditure % and pollution level is concurrent with the others.

- Greece and Turkey seem more comparable across the majority of features.



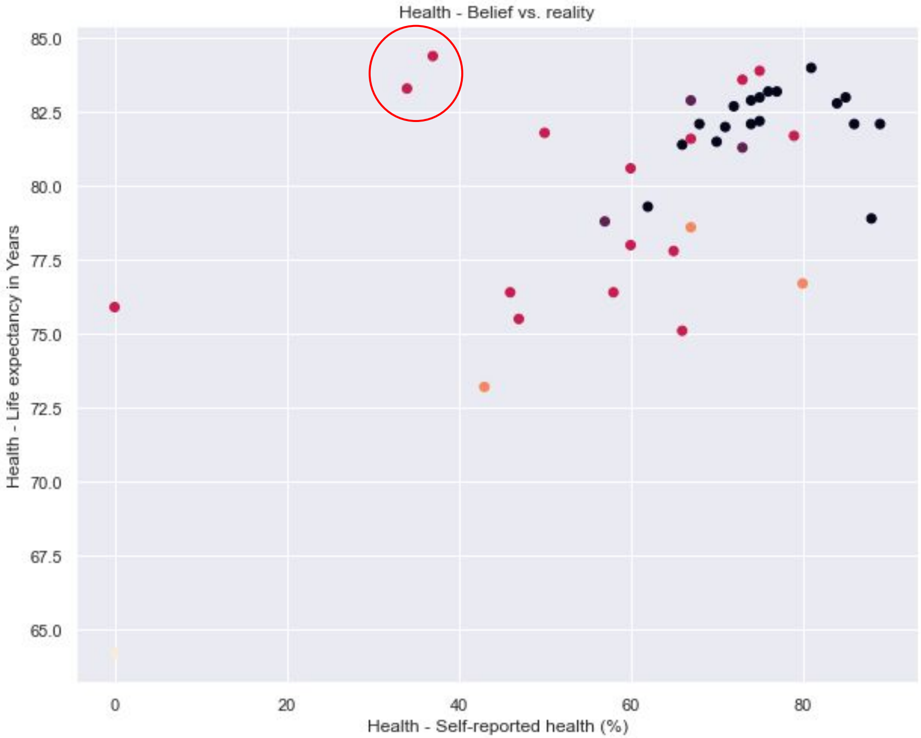
# Cluster [2] The Median Group with potential misclassification caused by cultural outlook

- Cluster includes many eastern and southern European states + Israel, Korea and Japan.

- Japan and Korea are interestings, as despite low unemployment, and higher income and wealth levels in their cluster - they self report some of the lowest Life Satisfaction scores.

- Additional observation is the disconnect between self-reported health levels, and actual life expectancy: Japan and Korea are the 1st and 5th longest lived, yet ranked themselves 38th and 39th in self-reported health respectively.

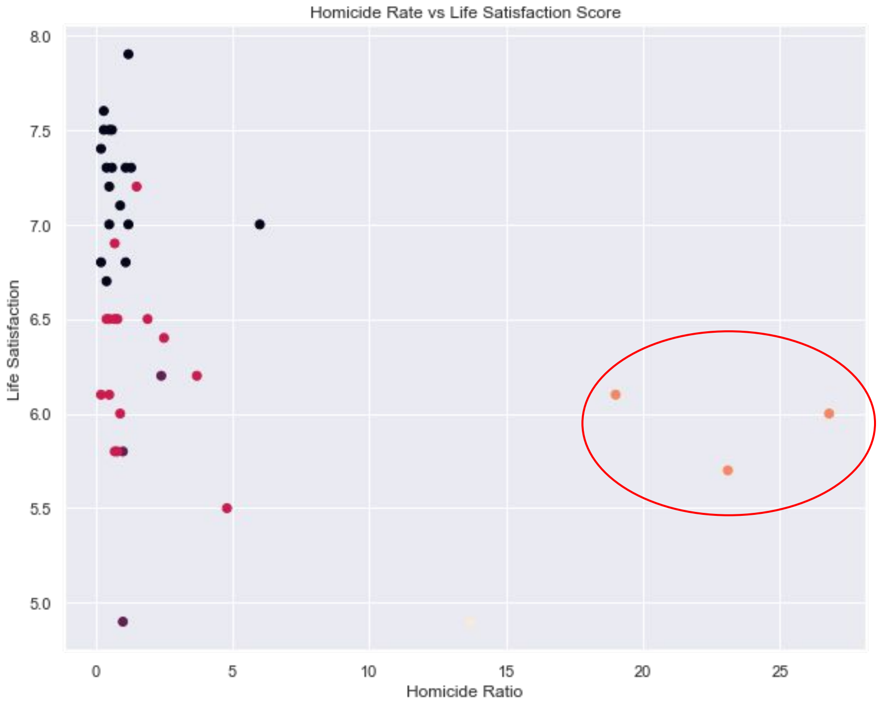
- Israel also had a number of null values in more model-sensitive features, such as: Income and Wealth levels. Based on it's position in the overview, it is more likely to be a constituent of cluster [0] (rich nations) if all information was known.



# Cluster [3] South America (minus Chile)

- Clear grouping of Columbia, Mexico and Brazil based largely on significantly higher Homicide rates relative to the rest of the world.

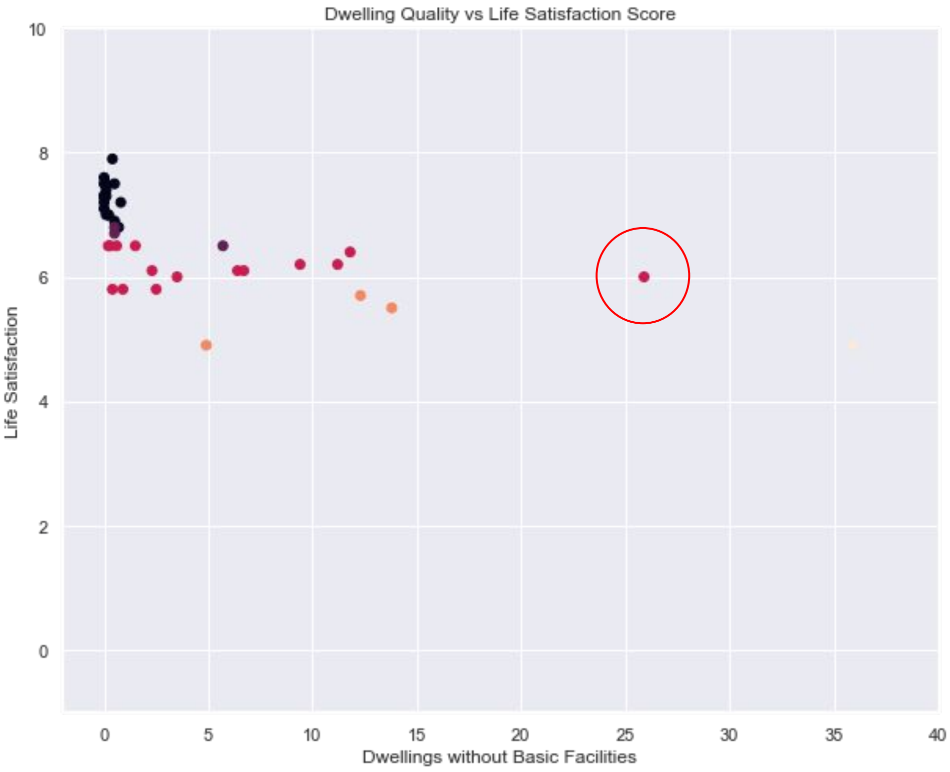
- However, this cluster also has the highest mean rate for *% Employees working very long hours.*



# Cluster [4] South Africa

- South Africa was separated out into its own cluster, largely due to the significantly higher rate of *Housing without Basic Facilities*

- Yet South Africa also has the highest Long-Term unemployment rate, at 17.9% (for context, Greece as the next closest sits at 10.8%)



# Conclusions

Based on broad analysis, key factors impacting Worldwide Life Satisfaction across our clusters:

1. *Higher homicide rate within Columbia, Brazil and Mexico.*
2. *High Labour market insecurity within Turkey and Greece.*
3. *Poor housing and infrastructure within South Africa.*
4. *Raising Income and Wealth levels within Southern and Eastern European State to the standards of Northern European states, and North America.*

*(Easier said than done on all of the above, however this is a broad based analysis!)*

# Q&A