The Factors of Choosing Best Areas for Families 2019 . A Case Study from Movinga: The New Way of Moving

MS4S08 – Group 8 – University of South Wales
Hossain KMI | Parvin M | Akter S | Nashirullah M | Siam M

Background Summary

This data comes from Movinga a global moving company that identifies the top cities around the world that offer the best conditions to raise a family in 2019. The dataset has been expanded to include geographic locations relative to latitude and longitude of cities. The case study is started by selecting 150 international cities known as attractive locations for raising a family. Then it's split into a number of factors into three categories that indicate whether the place is family friendly. These included key issues affecting quality of life in cities, such as housing, education, employment rates, and overall affordability, as well as family law, such as the amount of paid parental leave. The dataset is filtered manually with the help of python programming.

Aims and Outcomes

Aims

- To determine the factor of living cost by income in each country.
- To check the association between unemployment and education according to country.
- To evaluate the significant difference between the safety and health care of the country.

Expected Outcomes

- To explain the impact of the factors for choosing the best areas.
- To decide the best areas for families.

Distribution of Data

The data includes 18 columns and those are City, Country, Housing Affordability, Living Costs by Income, Unemployment (%), Education, Safety, Mobility, Air Quality (µg/m3), Healthcare, Kids' Activities, Paid Parental Leave (Days), Family Inclusivity, Neighborhood Safety, Family-Friendliness, Total, Latitude, Longitude.

Country	N Obs	Variable	Mean	Std Dev	N	Coeff of Variation	Skewness	Kurtosis
Canada	6	Living Costs by Income	82.9316667	6.6527541	6	8.0219708	-1.3768695	1.3217313
		Unemployment	5.4333333	1.4375906	6	26.4587223	0.6143792	-0.4421925
		Education	94.6066667	1.4351260	6	1.5169396	-0.7966897	-1.8339016
		Air Quality	13.6666667	2.8595920	6	20.9238442	0.5502301	-1.0497596
		Healthcare	77.2133333	1.1267416	6	1.4592578	-1.1929347	-0.0508040
France	10	Living Costs by Income	77.6780000	5.0754678	10	6.5339836	1.6231864	3.7122044
		Unemployment	9.2000000	2.0677417	10	22.4754528	0.7377772	-0.8866600
		Education	87.5190000	0.9281816	10	1.0605487	-1.4144705	2.9718481
		Air Quality	23.4750000	3.7459882	10	15.9573513	0.3907309	-1.5671536
		Healthcare	73.6420000	0.9715257	10	1.3192549	-0.4651536	-0.6604159
Germany	13	Living Costs by Income	88.2169231	5.2300898	13	5.9286695	0.7433263	0.8611762
		Unemployment	7.0776923	2.1063323	13	29.7601569	0.2327722	-0.5626451
		Education	88.8953846	0.9160205	13	1.0304478	0.3143868	0.3529408
		Air Quality	23.9584615	2.0454741	13	8.5375851	-0.4552912	-0.9047638
		Healthcare	72.5292308	1.1711423	13	1.6147177	0.7890144	0.0046245
UK	9	Living Costs by Income	70.5722222	3.4291026	9	4.8589976	-0.1111931	-0.6768997
		Unemployment	4.5833333	0.8739279	9	19.0675181	0.4709535	-1.1933705
		Education	87.6500000	1.5128698	9	1.7260351	-1.8787071	4.8532228
		Air Quality	17.8388889	3.1378832	9	17.5901271	0.2090611	-0.6056615
		Healthcare	64.5700000	4.5688812	9	7.0758574	1.4228689	0.6749489
US	48	Living Costs by Income	86.7214583	5.6981680	48	6.5706552	-0.9928988	1.5649614
		Unemployment	3.6981250	0.8287947	48	22.4112140	1.7714638	7.2917309
		Education	86.0995833	1.2308827	48	1.4296035	2.5264368	13.7221902
		Air Quality	15.7095833	3.8464822	48	24.4849410	1.0064952	2.2202715
		Healthcare	61.0902083	7.5848389	48	12.4158013	-0.1355720	-0.8887297

From the above table we can see:

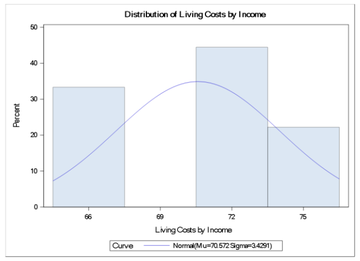
- Mean – highest average Education by country in Canada
- Standard Deviation - all measured on the same scale, US has the lowest in Unemployment sector so there is less variation and their employment rate is good. On the other side US has the highest variation in Healthcare sector.

Determine the significant differences of living costs by income in each country

H0: Samples came from a normally distributed population.
H1: At least one sample did not come from a normally distributed population.

Fitted Normal distribution for living costs by income for UK

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.16187496	Pr > D	>0.150
Cramer-von Mises	W-Sq	0.03456858	Pr > W-Sq	>0.250
Anderson-Darling	A-Sq	0.21013152	Pr > A-Sq	>0.250



Looking at the p-values for all tests for normality, no significant difference is found. Therefore we do not reject the null hypothesis, so there is no evidence to suggest that the distribution is not a normally distributed.

Is there any association between unemployment and education according to country?

H0: There is no association between unemployment and education.
H1: There is an association between unemployment and education.

As the samples are not normally distributed we have to apply Spearman correlation coefficient test.

Spearman Correlation Coefficients		Education
Unemployment	Canada	0.23540 0.6534
	France	-0.30064 0.3986
	Germany	-0.49728 0.0838
	UK	0.41667 0.2646
	US	-0.07852 0.5958

The correlation coefficients between Unemployment and Education are : $r_{(Canada)}$ = 0.23540, $r_{(France)}$ = -0.30064, $r_{(Germany)}$ = -0.49728, $r_{(UK)}$ = 0.41667, $r_{(US)}$ = -0.07852

Since p-value greater than 0.05, therefore we accept the null hypothesis and conclude that there is not significant relation between Education and Unemployment. But here, $r_{(US)}$ = -0.07852 (negatively corelated), so in US there is more employment opportunities according to education.

To evaluate the significant difference between the safety and health care of the country.

H0: There is no significant difference between safety and health care.
H1: There is significant difference between safety and health care.

Method - non-parametric one-way ANOVA (Kruskal-Wallis Test)

Kruskal-Wallis Test		
Chi-Square	DF	Pr > ChiSq
35.7509	35	0.4330

Test for US

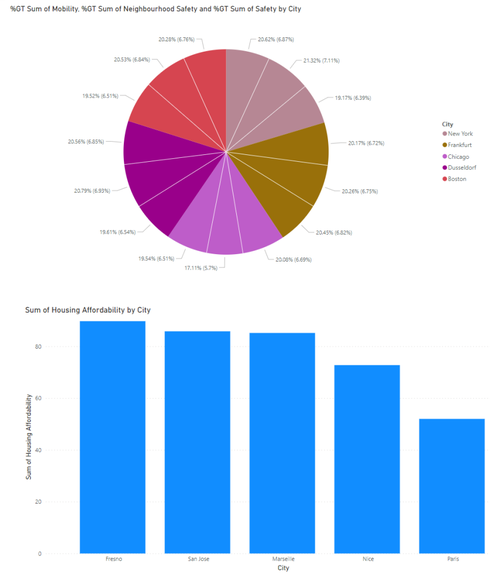
Kruskal-Wallis Test		
Chi-Square	DF	Pr > ChiSq
9.8901	11	0.5403

Test for Germany

Kruskal-Wallis Test		
Chi-Square	DF	Pr > ChiSq
5.2909	7	0.6245

Test for France

A Kruskal-Walls tests show that there is no significant difference between healthcare and safety. For US p>0.43 thus the null hypothesis and its not rejected.



With help of PowerBI visualization we can conclude that New York from US is in the top for Safety, Mobility and Neighborhood.

With help of PowerBI visualization we can conclude that Fresno from US is in the best area on Housing Affordability for families.

Conclusion

The study shows that Living cost by income, Education, Unemployment, Healthcare, Safety, Mobility, Air quality and Neighborhood have more impact on choosing the best area. From the analysis we can say that city from the US region is the best choice for families in times of 2019.

References

Movinga Best Cities for Families 2019 | Kaggle (www.Kaggle.com/datasets

Non Parametric Test - Definition, Types, Examples
(https://www.cuemath.com/data/non-parametric-test/)

Kruskal-Wallis Test: Definition, Formula, and Example
(https://www.statology.org/kruskal-wallis-test/)