

Life expectancy of a person

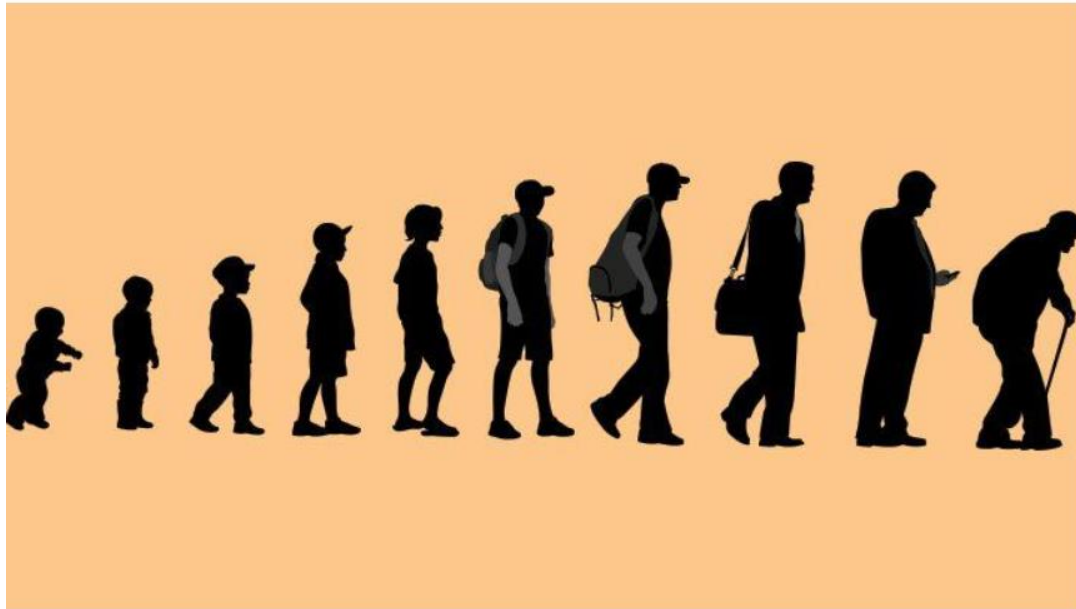
Anitha

Ananya

Saraswathi

Introduction

Life expectancy ?



Factors influencing

Country
Year
Status
Life expectancy
Adult Mortality
infant deaths
Alcohol
percentage expenditure
Hepatitis B
Measles
BMI
under-five deaths
Polio
Total expenditure
Diphtheria
HIV/AIDS
GDP
Population
thinness 1-19 years
thinness 5-9 years
Income composition of resources
Schooling

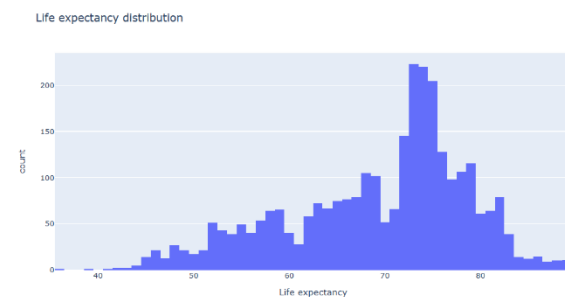
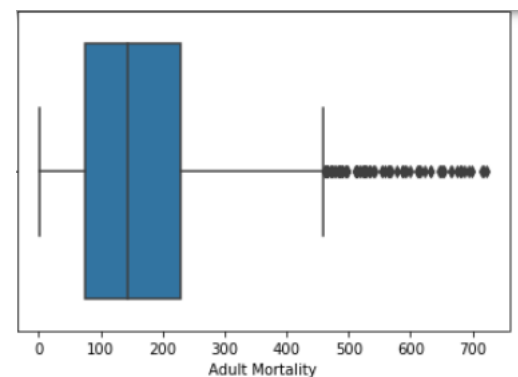
Data analysis/cleaning

Acquiring data :Kaggle/WHO data set

	Country	Year	Status	Life expectancy	Adult Mortality	Infant deaths	Alcohol	percentage expenditure	Hepatitis B	Measles	...	Polio	Total expenditure	Diphtheria	HIV/AIDS	GC
0	Afghanistan	2015	Developing	65.0	263.0	62	0.01	71.279624	65.0	1154	...	6.0	8.16	65.0	0.1	584.2592
1	Afghanistan	2014	Developing	59.9	271.0	64	0.01	73.523582	62.0	492	...	58.0	8.18	62.0	0.1	612.6965
2	Afghanistan	2013	Developing	59.9	268.0	66	0.01	73.219243	64.0	430	...	62.0	8.13	64.0	0.1	631.7449
3	Afghanistan	2012	Developing	59.5	272.0	69	0.01	78.184215	67.0	2787	...	67.0	8.52	67.0	0.1	669.9590
4	Afghanistan	2011	Developing	59.2	275.0	71	0.01	7.097109	68.0	3013	...	68.0	7.87	68.0	0.1	63.5372

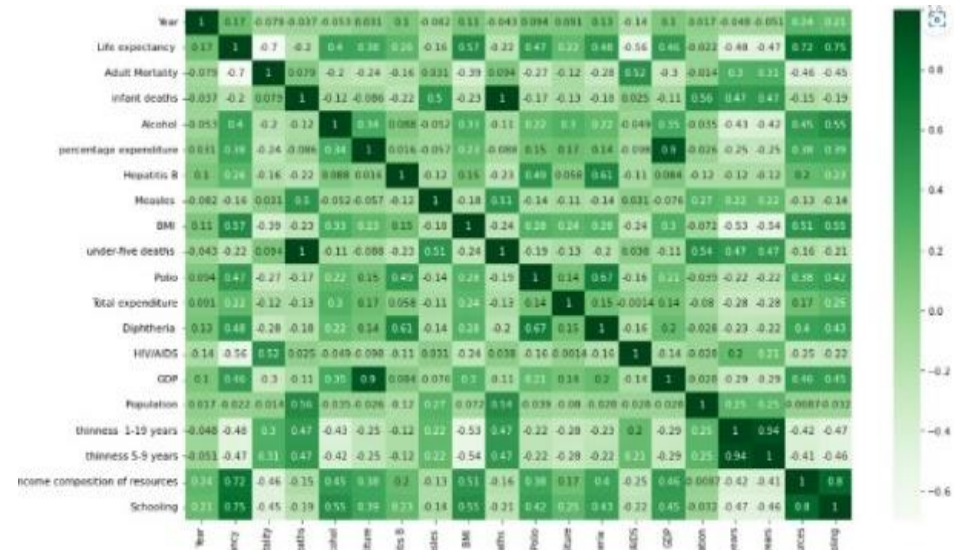
	Year	Life expectancy	Adult Mortality	Infant deaths	Alcohol	percentage expenditure	Hepatitis B	Measles	BMI	under-five deaths	Polio	expenditure
count	2938.000000	2928.000000	2928.000000	2938.000000	2744.000000	2938.000000	2385.000000	2938.000000	2904.000000	2938.000000	2919.000000	27
mean	2007.518720	69.224932	164.796448	30.303948	4.602861	738.251295	80.940461	2419.592240	38.321247	42.035739	82.550188	
std	4.613841	9.523867	124.292079	117.926501	4.052413	1987.914858	25.070016	11467.272489	20.044034	160.445548	23.428046	
min	2000.000000	36.300000	1.000000	0.000000	0.010000	0.000000	1.000000	0.000000	1.000000	0.000000	3.000000	
25%	2004.000000	63.100000	74.000000	0.000000	0.877500	4.685343	77.000000	0.000000	19.300000	0.000000	78.000000	
50%	2008.000000	72.100000	144.000000	3.000000	3.755000	64.912906	92.000000	17.000000	43.500000	4.000000	93.000000	
75%	2012.000000	75.700000	228.000000	22.000000	7.702500	441.534144	97.000000	360.250000	56.200000	28.000000	97.000000	
max	2015.000000	89.000000	723.000000	1800.000000	17.870000	19479.911610	99.000000	212183.000000	87.300000	2500.000000	99.000000	

Univariate analysis, Histogram analysis for data distribution analysis

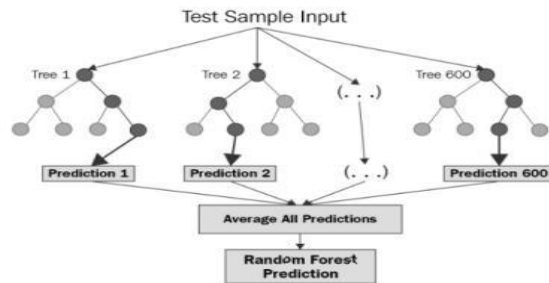


Heat matrix , Correlation matrix for parameter dependency analysis

	Year	Life expectancy	Adult Mortality	infant deaths	Alcohol	percentage expenditure	Hepatitis B	Measles	BMI	under-five deaths	Polio	Total expenditure	Diphtheria
Year	1.000000	0.170033	-0.079052	-0.037415	-0.052990	0.031400	0.104333	-0.082493	0.108974	-0.042937	0.094158	0.090740	0.134337
Life expectancy	0.170033	1.000000	-0.696359	-0.196557	0.404877	0.381864	0.256762	-0.157586	0.567694	-0.222529	0.465556	0.218086	0.479495
Adult Mortality	-0.079052	-0.696359	1.000000	0.078756	-0.195848	-0.242860	-0.162476	0.031176	-0.387017	0.094146	-0.274823	-0.115281	-0.275131
infant deaths	-0.037415	-0.196557	0.078756	1.000000	-0.115638	-0.085612	-0.223566	0.501128	-0.227279	0.996629	-0.170689	-0.128616	-0.175171
Alcohol	-0.052990	0.404877	-0.195848	-0.115638	1.000000	0.341285	0.087549	-0.051827	0.330408	-0.112370	0.221734	0.296942	0.222020



Algorithm analysis, error calculation



	mean_absolute_error	mean_squared_error	train_accuracy	test_accuracy
Linear Reg	3.070265	4.158004	0.819106	0.811760
Ridge	3.065506	4.159855	0.818929	0.811592
Lasso	3.430167	4.639142	0.775612	0.765675
SVR	2.507029	3.670730	0.859027	0.853294
Decision Tree	1.660998	2.739242	1.000000	0.918303
RF	1.225884	1.960517	0.994215	0.958151
Ada Boost	2.271010	2.923890	0.912369	0.906918
Grad Boost	1.638806	2.281133	0.962593	0.943344

```
] results.T
```

```
]:
```

	Linear Reg	Ridge	Lasso	SVR	Decision Tree	RF	Ada Boost	Grad Boost
mean_absolute_error	3.070265	3.065506	3.430167	2.507029	1.660998	1.225884	2.271010	1.638806
mean_squared_error	4.158004	4.159855	4.639142	3.670730	2.739242	1.960517	2.923890	2.281133
train_accuracy	0.819106	0.818929	0.775612	0.859027	1.000000	0.994215	0.912369	0.962593
test_accuracy	0.811760	0.811592	0.765675	0.853294	0.918303	0.958151	0.906918	0.943344

GUI implementation

tk

Predicted values: font = "50" fg="Newblue")

Afghanistan
Albania
Algeria
Angola
Antigua and Barbuda
Argentina
Armenia
Australia
Austria
Azerbaijan
Bahamas
Bahrain
Bangladesh
Barbados
Belarus
Belgium
Belize
Benin
Bhutan

BMI

☐ ALcoholic

☐ HIV

Estimate

Year *
2021

Status *
1

Adult Mortality *
145

Infant Deaths *
90

Alcohol *
0.23

Percentage Expenditure *
79

Hepatitis B *
234

Measles *
90

BMI *
30

Under-Five Deaths *
80

Polio *
32

Total Expenditure *
1235

Results, what we have done?

tk

Honduras
Hungary
Iceland
India
Indonesia
Iran (Islamic Republic of)
Iraq
Ireland
Israel
Italy
Jamaica
Japan
Jordan
Kazakhstan
Kenya
Kiribati
Kuwait
Kyrgyzstan

BMI

21

☒ ALcoholic

☐ HIV

Predicted value=

[74.93424
267]

Estimate

tk

Haiti
Honduras
Hungary
Iceland
India
Indonesia
Iran (Islamic Republic of)
Iraq
Ireland
Israel
Italy
Jamaica
Japan
Jordan
Kazakhstan
Kenya
Kiribati
Kuwait

BMI

21

☒ ALcoholic

☒ HIV

Predicted value=

[72.40167
961]

Estimate