Task 1:

RMSE Scores (for 2014 and 2015 data):

	Linear Function	Quadratic Function	Cubic Function	Quartic Function
Training Data (2000 – 2013)	1.0672591184072877	0.9894912620578972	0.8718359094448799	0.8329835827343549
Testing Data (2014 – 2015)	4.269053041258142	4.150947401409626	3.8221538776086406	4.193473032433365

R2 Scores (for 2014 and 2015 data):

	Linear Function	Quadratic Function	Cubic Function	Quartic Function
Training Data (2000 – 2013)	0.4788209985493368	0.5520071442080722	0.6522103389963448	0.6825173143510774
Testing Data (2014 – 2015)	-1.8027395415725316	-1.6498061252240395	-1.2466528664542502	-1.7043777122254329

The best model to predict the dataset would be the Cubic function model because it has the least RMSE score and the closest R^2 value to 1. This means that it is the best model we have to predict the small dataset out of all the models we created.

Task 2:

For Developing Countries the Linear function and Quadratic functions would be the best to use as they have least RMSE score and R^2 value. For Developed countries the same would be the case as for Developed they also had lower RMSE and R^2 values.

Developed

Developing Countries	Linear Function	Quadratic Function	Cubic Function	Quartic Function
Training Data (2000 - 2013)	1.98	1.79	1.64	23.53
Testing Data (2014 and 2015)	2.404	2.43	2.67	28.55

Developing

Developed Countries	Linear Function	Quadratic Function	Cubic Function	Quartic Function
Training Data (2000 - 2013)	1.762	1.48	1.34	7.22
Testing Data (2014 and 2015)	2.58	2.60	4.90	16.62

R2 Scores:

Developed

~~Developing~~ Countries	Linear Function	Quadratic Function	Cubic Function	Quartic Function
Training Data (2000 - 2013)	0.13	0.279	0.37	-239.67
Testing Data (2014 and 2015)	-208.61	-294.31	-414.99	-11387

Developing	Developed Countries	Linear Function	Quadratic Function	Cubic Function	Quartic Function
	Training Data (2000 - 2013)	0.23	0.46	0.54	-234
	Testing Data (2014 and 2015)	-385	-351	-3597	-127450

Task 3:

For interesting results I found that Adult mortality didn't have much of an impact on Life expectancy as much as I thought. Lots of the values were negatively correlated which was also interesting.

Please fill this table (for testing with 2014 and 2015 data):

	RMSE	R2
Developing Country	3.08	-0.46
Developed Country	1.0	-100.93

Please fill the following table with the "regression coefficients" (for each variable):

	Adult Mortality	Alcohol	BMI	GDP	Schooling
Developing Country	-0.001	5.75	0.71	-0.0002	0.1311
Developed Country	-0.022	5.93	0.0066	-1.15	0.1898

Task 4:

There are many limitations to the model for example if the relationship between the variables is not linear then the predictions would not be as accurate. This would occur if the relationship between the independent and dependent variables was non-linear. Furthermore, the model was very susceptible to outliers and contained inconsistent data and this can significantly affect the regression line and make incorrect results. Life expectancy is something that is related to many variables, and we could improve this model by providing more variables to predict on. We also could try different models to model the data like decision tree or polynomial regression. Removing outliers in the dataset would also help make the model a better predicter of Life expectancy.