For this assignment, we have used the data from world bank website. We have selected two categories of indicators which are climate change and aid effectiveness indicators. There are several things and operations that can be performed to understand the differences between different indicators of both categories. Different Python IDEs can be used to deal with the world bank data. For this assignment, we have used Jupyter Notebook, which is a very effective tool for dealing with this kind of data. We have selected 10 different countries to find the relationship between different indicators of two categories that we have selected. Along, with this we have used the data of recent 10 years which means the data from 2011 to 2020. The countries that we have selected are United States, Canada, India, Japan, China, Thailand, Costa Rica, Australia, Afghanistan, and New Zealand. It is also worth noting here, that we have created two different data frames for downloading the data from World Bank. For both of the data frames, the countries and time range are same, but the indicators are different in both data frames. Below is the screenshot that is telling us about the type of data that we have selected from the World Bank.

		YR2011	YR2012	YR2013	YR2014	\
economy	series					
AFG	AG.LND.FRST.ZS	1.850994	1.850994	1.850994	1.850994	
	AG.LND.IRIG.AG.ZS	5.391717	5.465576	5.518333	5.742548	
	EG.ELC.ACCS.ZS	43.222019	69.099998	68.290649	89.500000	
	ER.H2O.FWTL.ZS	NaN	43.015907	NaN	NaN	
AUS	AG.LND.FRST.ZS	16.955310	17.047689	17.140067	17.232446	
	AG.LND.IRIG.AG.ZS	0.495556	0.553121	0.639634	0.630650	
	EG.ELC.ACCS.ZS	100.000000	100.000000	100.000000	100.000000	
	ER.H2O.FWTL.ZS	NaN	3.133875	NaN	NaN	
CAN	AG.LND.FRST.ZS	38.734864	38.730258	38.725651	38.721045	
	AG.LND.IRIG.AG.ZS	NaN	NaN	NaN	NaN	
	EG.ELC.ACCS.ZS	100.000000	100.000000	100.000000	100.000000	
	ER.H2O.FWTL.ZS	NaN	1.278596	NaN	NaN	
CHN	AG.LND.FRST.ZS	21.491096	21.696596	21.902095	22.107595	
	AG.LND.IRIG.AG.ZS	10.206754	NaN	NaN	NaN	
	EG.ELC.ACCS.ZS	99.848724	99.961929	99.996445	100.000000	

Figure 1: Data Frame-1

economy	series	YR2011	YR2012	YR2013	YR2014	\
AFG	DT.ODA.ODAT.GI.ZS	NaN	NaN	NaN	NaN	
	SE.PRM.CMPT.ZS	NaN	NaN	NaN	NaN	
	SH.STA.ANVC.ZS	47.900000	51.200000	NaN	63.200000	
	SH.STA.STNT.ZS	NaN	NaN	40.400000	NaN	
AUS	DT.ODA.ODAT.GI.ZS	NaN	NaN	NaN	NaN	
	SE.PRM.CMPT.ZS	NaN	NaN	NaN	NaN	
	SH.STA.ANVC.ZS	NaN	NaN	NaN	NaN	
	SH.STA.STNT.ZS	NaN	NaN	NaN	NaN	
CAN	DT.ODA.ODAT.GI.ZS	NaN	NaN	NaN	NaN	
	SE.PRM.CMPT.ZS	NaN	NaN	NaN	NaN	
	SH.STA.ANVC.ZS	NaN	NaN	NaN	NaN	
	SH.STA.STNT.ZS	NaN	NaN	NaN	NaN	
CHN	DT.ODA.ODAT.GI.ZS	-0.017131	-0.004577	-0.014785	-0.019729	
	SE.PRM.CMPT.ZS	NaN	NaN	NaN	NaN	
	SH.STA.ANVC.ZS	93.700000	95.000000	95.600000	96.200000	
	SH.STA.STNT.ZS	NaN	NaN	8.100000	NaN	
CRI	DT.ODA.ODAT.GI.ZS	0.414668	0.303494	0.328816	0.545544	
	SE.PRM.CMPT.ZS	104.360451	102.410400	98.649651	98.426071	
	SH.STA.ANVC.ZS	98.100000	NaN	NaN	NaN	
	SH.STA.STNT.ZS	NaN	NaN	NaN	NaN	
IND	DT.ODA.ODAT.GI.ZS	0.453094	0.239972	0.388836	0.428159	
	SE.PRM.CMPT.ZS	92.763359	NaN	96.200890	97.571121	
	SH.STA.ANVC.ZS	NaN	NaN	NaN	NaN	
	SH.STA.STNT.ZS	NaN	NaN	NaN	38.700000	
JPN	DT.ODA.ODAT.GI.ZS	NaN	NaN	NaN	NaN	
	SE.PRM.CMPT.ZS	NaN	NaN	NaN	NaN	
	SH.STA.ANVC.ZS	NaN	NaN	NaN	NaN	
	SH.STA.STNT.ZS	NaN	NaN	NaN	NaN	
NZL	DT.ODA.ODAT.GI.ZS	NaN	NaN	NaN	NaN	
	SE.PRM.CMPT.ZS	NaN	NaN	NaN	NaN	
	SH.STA.ANVC.ZS	NaN	NaN	NaN	NaN	
	SH.STA.STNT.ZS	NaN	NaN	NaN	NaN	

Figure 2: Data Frame-2

If we look at the values in both figures, we can see that the values are not that meaningful. So, therefore we have transformed the values based on the mean. It means that for all of the indicators, now we will get the mean value.

economy	series	YR2011	YR2012	YR2013	YR2014	YR2015	YR2016	YR2017	YR2018	YR2019	YR2020
AFG	AG.LND.FRST.ZS	16.821577	29.858119	25.219992	32.364514	26.353963	35.344044	37.139350	34.533919	49.775496	49.775496
	AG.LND.IRIG.AG.ZS	16.821577	29.858119	25.219992	32.364514	26.353963	35.344044	37.139350	34.533919	49.775496	49.775496
	EG.ELC.ACCS.ZS	16.821577	29.858119	25.219992	32.364514	26.353963	35.344044	37.139350	34.533919	49.775496	49.775496
	ER.H2O.FWTL.ZS	16.821577	29.858119	25.219992	32.364514	26.353963	35.344044	37.139350	34.533919	49.775496	49.775496
AUS	AG.LND.FRST.ZS	39.150289	30.183671	39.259901	39.287699	39.314057	39.350721	30.316814	39.353871	58.710657	58.710657
	AG.LND.IRIG.AG.ZS	39.150289	30.183671	39.259901	39.287699	39.314057	39.350721	30.316814	39.353871	58.710657	58.710657
	EG.ELC.ACCS.ZS	39.150289	30.183671	39.259901	39.287699	39.314057	39.350721	30.316814	39.353871	58.710657	58.710657
	ER.H2O.FWTL.ZS	39.150289	30.183671	39.259901	39.287699	39.314057	39.350721	30.316814	39.353871	58.710657	58.710657
CAN	AG.LND.FRST.ZS	69.367432	46.669618	69.362826	69.360522	46.758775	69.356007	46.653857	69.351881	69.349819	69.347756
	AG.LND.IRIG.AG.ZS	69.367432	46.669618	69.362826	69.360522	46.758775	69.356007	46.653857	69.351881	69.349819	69.347756
	EG.ELC.ACCS.ZS	69.367432	46.669618	69.362826	69.360522	46.758775	69.356007	46.653857	69.351881	69.349819	69.347756
	ER.H2O.FWTL.ZS	69.367432	46.669618	69.362826	69.360522	46.758775	69.356007	46.653857	69.351881	69.349819	69.347756

Figure 3: Transformation Sample

Now, to check the correlation between year 2011 and 2019 of the first data frame which is climate change, we have visualized a scatter plot below.

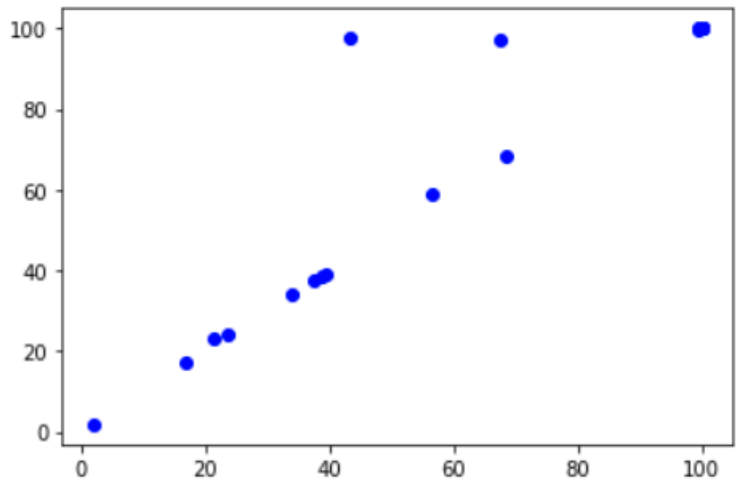

Figure 4: Scatter Plot

Here, we can clearly see that both 2011 and 2019 years have strong correlation which means that both the years are contributing equally to climate change factor.