George Washington University
School of Business

ANTI-CORRUPTION GEO-ANALYTICS ENABLEMENT

DELOITTE

Xiang Fan, Vibhu Verma, Jennifer Nguyen, Lingxi Yang, Yue Zheng

TABLE OF CONTENT

Section 1 – Abstract

This project is to conduct an analysis to provide expanded capabilities to shape future anti-corruption investment coupled with global factors in shaping USAID strategic action in deploying its resources. Our final deliverables including tailored dashboards, clustering models, and predictive models will answer two key business questions.
They are i) which makes a successful anti-corruption project? and ii) where should USAID invest in future anti-corruption projects? By answering these two key business questions, we will achieve the goal to help USAID gauge the effectiveness of these Anti-Corruption projects to fight corruption and promote economic growth.

Section 2 – Introduction

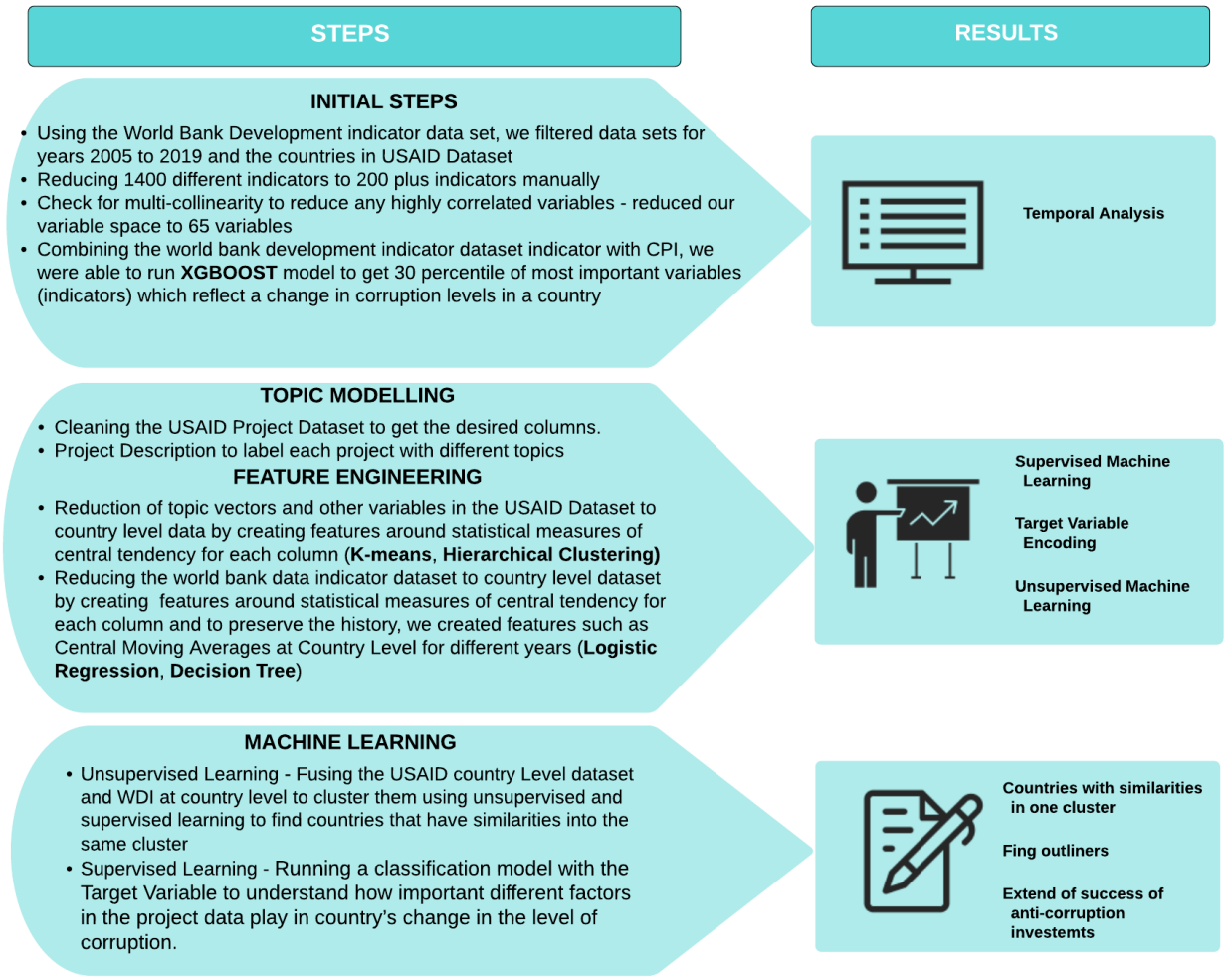
The project's core data set is the USAID database of 300 anti-corruption projects that was conducted between the year 2007 and 2013. The dataset provides details about each project such as the country and region where it was operated, the amount of money spent and the corruption assessment.

To support our analyses, we have also expanded the data source to also include development indicators to reflect the corruption level from reputable sources World Bank. We believe that there can be a relationship between corruption and development indicators. Discovering that relationship will shed a light on how to fight corruption.

Our analyses aim to identify the trends and patterns among countries and regions when it comes to corruption. Based on the patterns (pointed out by historical corruption and development data), we would like to categorize those countries and regions into several buckets so that the USAID would have a clearer guide on how to alleviate corruptions in different categories.

Section 3 – Methodology

Key Methodology Overview:



General Methodology Step:

- Step 1: Gather and Melt data
  - After gathering data from USAID and world bank, we decide 3 dataset: USAID Anticorruption Projects Database, Corruption Perception Index (CPI) and Development Indicators from World Bank.First,we merge CPI data and indicator dataset together to analysis and build models ,and then melt selected indicators and CPI with USAID database(after wrangling and analysis).
- Step 2: Wrangling data using python
  - After gathering the data from USAID and world bank ,we find that these database are not clean and really need to do data wrangling (like clean N/A by using random forest to predict and calculate the correlation heatmap and report to delete highly related indicators)
- Step 3: Explore the relationship and select indicators using machine learning techniques (python)
  - We have more than 1000 indicators in our world bank dataset, but most of them are

not significant to CPI.So we must use machine learning methods to select most significant indicators and use them for future analysis. So we use xgboost to select, which has a higher accuracy.

- Step 4: Do text mining using Natural Language Processing

  - In order to provide a more objective result, the up to date impact score, Natural Language Processing(NLP) analysis is applied. NLP is the automatic manipulation of natural language by software[2]. It studies how humans communicate with each other in speech and text. Our USAID database has a lot of text description ,so NLP will be helpful. We use this method to do sentiment analysis and LDA for the USAID database.

- Step 5: Analysis using python

  - We use several python methods (such as unsupervised machine learning and calculate method .We have done feature engineering for clustering models (unsupervised) on melted database ,and calculates Aggregated Indicator for all World Bank Development Indicators**,** as well as calculate growth rate for the aggregated indicator.

- Step 6: Visualize Data trend and pattern Using Tableau

  - Tableau software is a data visualization tool which is currently being used in the BI industry. It is one of the best ways to change or transform the raw set of data into an easily understandable format with zero technical skills and coding knowledge. In this project, we use tableau to visualize our world bank dataset(after wrangling and analysis) based on country-level and region-level to show the pattern of different countries and regions for each significant indicator(after selected by python) and compare them.

- Step 7: Create Tableau Dashboard

  - Tableau dashboard is a good way to clearly show our visualization analysis and our conclusion. So this method will be used for our final result display and visualization display.

Section 4 - Data Sources

Three datasets are used for our analysis: **USAID Anticorruption Projects Database, Corruption Perception Index (CPI) and Development Indicators from the World Bank.**

The USAID Anticorruption Projects Database includes information about USAID projects with anticorruption interventions implemented worldwide between 2007 and 2013. The Database includes about 300 long-term country projects and regional or short-term projects. The information for different projects in this dataset includes but is not limited to project description and results, implementation timeframe, Control of Corruption Index, project value, implementer and amount spent. This dataset is significant and meaningful to help us analyze which countries or regions are worthy of anti-corruption investment and how much money should be invested.

The Corruption Perception Index (CPI)dataset gives us the information about the annual perceived level of corruption for each country. It contains the corruption perception index value for different countries through different years. However, this index does not really help us to determine the effectiveness of USAID'S relative level of investment to fight corruption and promote economic growth, therefore, we chose some development indicators which are most related to corruption from the World Development Indicators dataset to help us answer our business questions.

The World Development Indicators dataset presents the most current and accurate global development

data, which includes national and regional estimates. This dataset includes indicators about various development categories for different countries such as agriculture and food security, economic growth, education, energy and extractives, health, trade and so on. We will choose the indicators that are mostly correlated to corruption for our analysis.

Section 5 – Data processing

5.1 Initial Steps

Step1: Merge datasets
a.Add two columns, 'Area' and 'Income Group' into table 'World Development Indicators' by separating the country, area, and income group information and merging the Metadata Country API information.
b.Unmelt the data table from Step a based on indicators. So, in the way, the indicators moved from rows to columns.
c.Right outer join the CPI table with the table from Step b on country and year
In this way, we successfully merged into two main dataset:WDI dataset(include indicators and CPI) ,USAID project dataset(containing the detailed information for all the anti-corruption project portfolios)

Step2: Filter Years
The World Development Index data has too many missing values originally, which was very unfavorable for our analysis and modeling. And we found the data before 2005 has a large number of null values in the most indicators and countries. So in order to reduce null values, we decided to only keep the year from **2005 to 2019**.

Step 3: Manual reduction of variables to lesser variables
There are more than 1400 indicators in the WDI database originally and most of them seem to have little relevance with corruption. To reduce the computational complexity and enhance the analysis accuracy, we decided to manually reduce variables to lesser variables. We divided WDI into **15 different categories** (based on the description and categories in the world bank dataset), which are:

- Agriculture and Food Security
- Climate Change
- Economic Growth
- Education
- Energy and Extractives
- Environment and Natural Resources
- Financial Sector Development
- Health, Nutrition and Population
- Macroeconomic Vulnerability and Debt

- Poverty
- Private Sector Development
- Public Sector Management
- Social Development
- Social Protection and Labor
- Trade

After manually selecting the relevant indicator variables from the 15 categories, we have 185 indicators left for future processing and analysis.

Step 4:Multicollinearity to reduce the number of variables

The indicators with high proportions of missing values could affect the analysis accuracy. So we deleted the indicators with more than 40% missing values by using Python. After that, we used Python to delete the indicators with high correlation (over 95%) to avoid multicollinearity in our models. After completing this step, we had 65 development indicators left.

Step 5: Imputing data

By applying the Random Forest algorithm, we imputed and filled all the missing values for the development indicators.

Step 6: Reducing the 65 selected variables into 20 most important variables by using CPI as the regressor.

By running regression and other ML models, we selected the 20 most related development indicators from the World Bank. We tried multiple regression models including stepwise regression model, lasso regression model and xgboost model. After comparing the test error, we decided to use the result of xgboost model, which has the lowest test error, to select the most related development indicators from the World Bank. Based on the XGBoost's "Feature Importance" chart, we selected **20 most important variables** shown as below:

- Gni Growth (Annual %),
- Intentional Homicides (Per 100,000 People)
- Food Production Index (2004-2006 = 100)
- Claims On Central Government (Annual Growth As % Of Broad Money)
- Foreign Direct Investment, Net Inflows (% Of Gdp)
- Inflation, Consumer Prices (Annual %)
- Claims On Private Sector (Annual Growth As % Of Broad Money)
- Inflation, Gdp Deflator: Linked Series (Annual %)
- Broad Money Growth (Annual %)
- Foreign Direct Investment, Net Outflows (% Of Gdp)
- Bank Liquid Reserves To Bank Assets Ratio (%)
- Total Tax And Contribution Rate (% Of Profit)
- Exports Of Goods And Services (% Of Gdp)
- Net Barter Terms Of Trade Index (2000 = 100)
- Total Reserves (% Of Total External Debt)
- Price Level Ratio Of Ppp Conversion Factor (Gdp) To Market Exchange Rate

- Merchandise Exports To Low- And Middle-Income Economies Within Region (% Of Total Merchandise Exports)
- Merchandise Exports To Low- And Middle-Income Economies Outside Region (% Of Total Merchandise Exports)
- Gross Savings (% Of Gdp)
- Domestic Private Health Expenditure (% Of Current Health Expenditure)

Step 7: Filtering the USAID for the countries which have data related to awd_amount
After exploring the dataset, we found there are four countries that have no anti-corruption projects with valid awd_amount value, which are Laos, Libya, Russian, and Senegal. In this case, these countries will not contribute to our further analysis on investment amount, so we filtered them out.

USAID DATASET

Step1: Topic Modelling

Data Preprocessing

- Using gensim built in module for simple preprocessing to remove punctuations
- Using lemmatization to further limit to only noun adjective word and adverbs thereby removing pronounce from the tokens, which will map each word to it's base state based on a built in dictionary
- We used TF-IDF vectorizer to get the input vectors from our corpus for each document, which accounts for over usage of few words in a document belonging to a corpus by penalising them
- Building a TF-IDF vectorizer using minimum required occurrences of a word in 10 documents stopwords from English, number of characters for each word greater than three and ngram range to trigram.

Using the USAID data set at project level topic we did topic modeling against project description columns to Assign topics to different projects based on topic modeling using Latent Dirichlet Allocation. This will help us assign a particular topic number for each project in our USAID dataset.

Selecting the appropriate number of topics was based on cross validation using different hyperparameters which aimed on getting the highest coherence score, and also making sure we have good intertropical distance(separation in topics) in our TSNE plot, we got 5 topics as the best model.

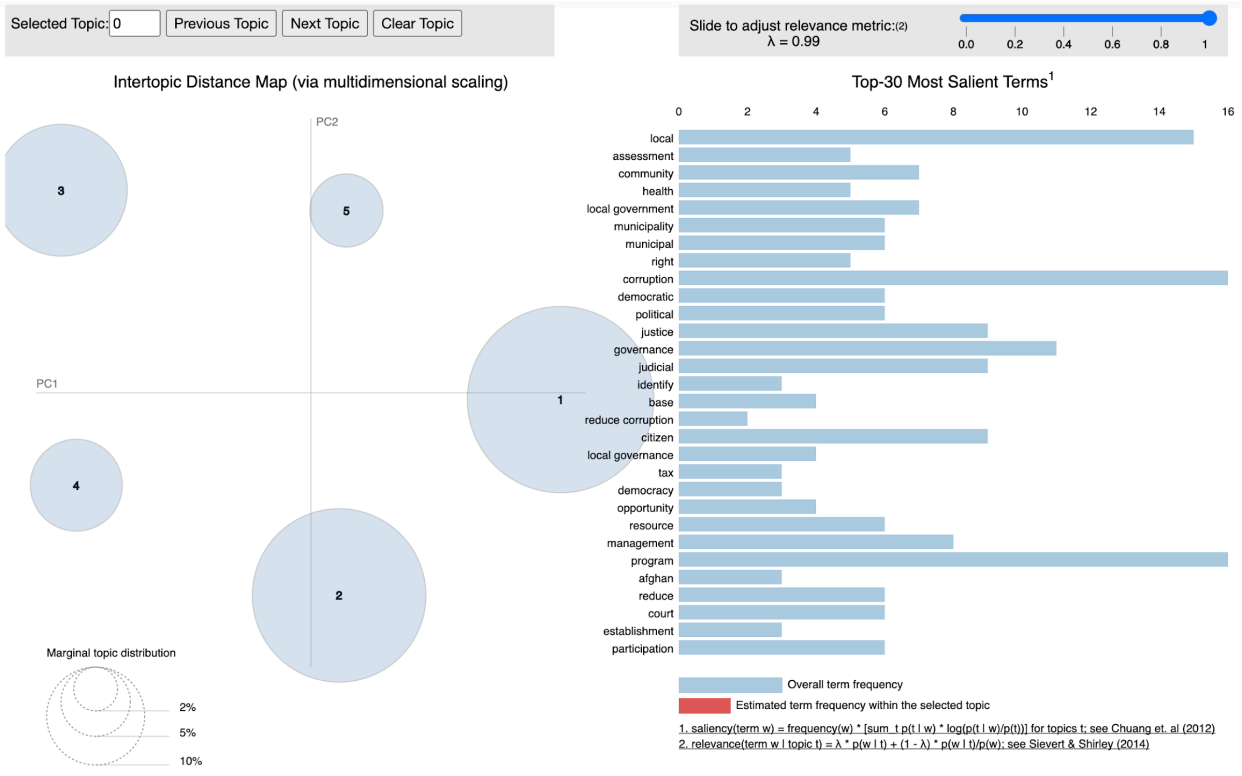
We gathered topic scores for all the five topics for different project descriptions in our data set as new features in the USAID data set, which will be used when we are reducing this data set into country level data set from project level data set

From our LDA plot in two dimensions we can use different values of Lambda to get most salient terms smaller value of Lambda will give us terms which particularly belong to one document and rear in the

corpus and larger values of Lambda will give us most common terms in that particular document based on frequency.

Using the top features in each topic we can say that following about the each topic:
Topic 0: Governance development
Topic 1: Technical support for public and private investment
Topic 2: Corruption reduction in the health sector
Topic 3: Projects that aim to strengthen governance of programs in democratic, political and health sector
Topic 4: Rule of law strengthening for civil society

(**Coherence** measures the degree of semantic similarity between high scoring words in a particular topic. These measurements help distinguish between topics that are semantically interpretable topics and topics that are artifacts of statistical inference)



Step2: Sentiment Analysis

We used sentiment analysis on both project description and project result columns to create more features on our project level USAID data set, also we included subjectivity score for both the columns as new features in the data set.

5.2 Reduction of data to the country level

Step1: Feature Engineering

USAID

We used different measures around the central tendency such as **mean, median, sum, count, min, max** to get multiple features at country level which helps us preserve information and reduce the project level USAID data set into country level dataset.

We also used Inter categorical counts, mean topic scores for all the projects for a country as features in our data set.

WDI DATASET

Not only using the measures of central tendency like the USAID data set we used central moving averages over different years as new features.
This helps us preserve more features from our temporal data set and have a richer dataset for Both supervised and unsupervised machine learning.
We created 5 central moving averages for all the 20 indicators in our data set for all the countries as follows:

2005-2007	MA(3)
2008-2010	MA(3)
2011-2013	MA(3)
2014-2016	MA(3)
2017-2019	MA(3)

Step2: Target Encoding for Supervised Learning

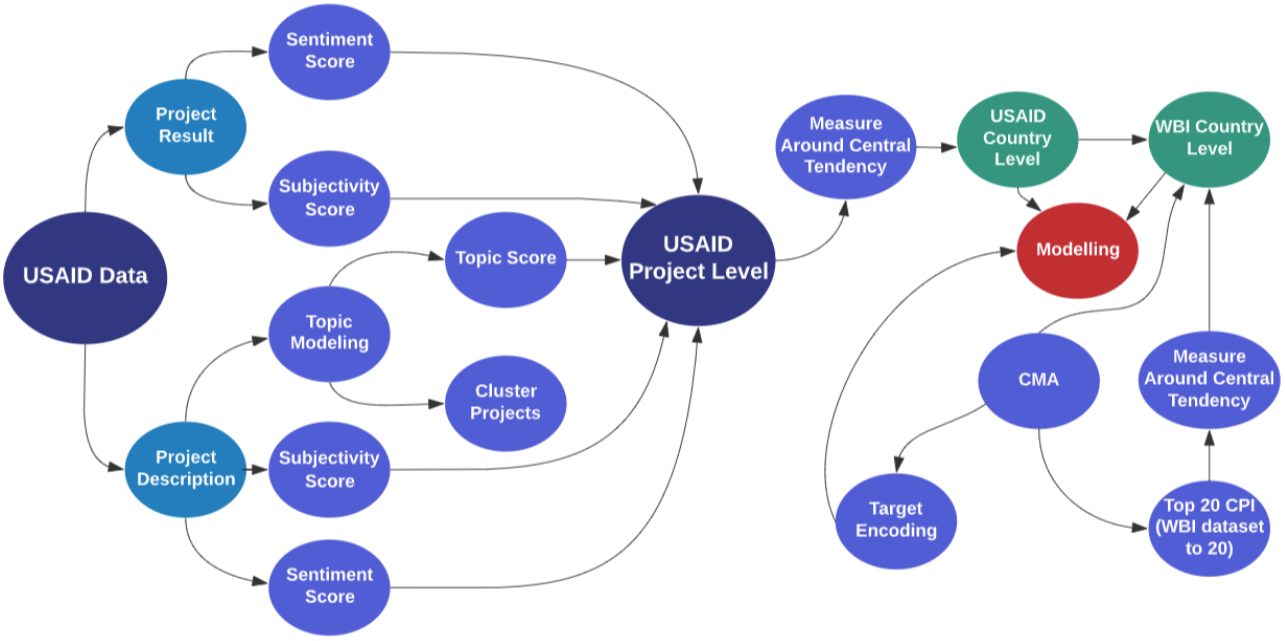
Steps

- Using the central moving averages for five different periods for each country for each indicator we calculate it differences between two consecutive periods to get 4 different values of '**positive change'** or '**no positive change'** in a particular indicator, also including with it the direction of indicator and rounding of negative values to minus one and positive values to one gave us +1 and -1 for a particular indicator between different time periods
- Using Mode of '**positive change(+1)'** or '**no positive change (-1)'** for each indicator for each country we can get one value for each particular indicator at country level
- Using awaited submission of these values with respect to F score from our initial **XGBOOST** model

gives us the target variable as '**positive change'** or '**no positive change'** in a particular country, this can be used as a target variable in a supervised classification model using USAID data set as independent variables.

The target variable looks as below:

	Country Name	Target	Result_Inv
0	Afghanistan	-0.073611	No Positive Change
1	Albania	0.010923	Positive Change
2	Angola	-0.114136	No Positive Change
3	Armenia	0.015039	Positive Change
4	Azerbaijan	-0.250277	No Positive Change
...	...	...	...
66	Ukraine	0.306950	Positive Change
67	Vietnam	-0.066962	No Positive Change
68	West Bank and Gaza	0.011398	Positive Change
69	Yemen	-0.403356	No Positive Change
70	Zambia	-0.139465	No Positive Change



Section 6 - Machine Learning and Temporal analysis

We fused the USAID country Level data set and WDI at country level to cluster them using unsupervised to find countries that have similar and dissimilar results.

6.1 Unsupervised Learning

DATA

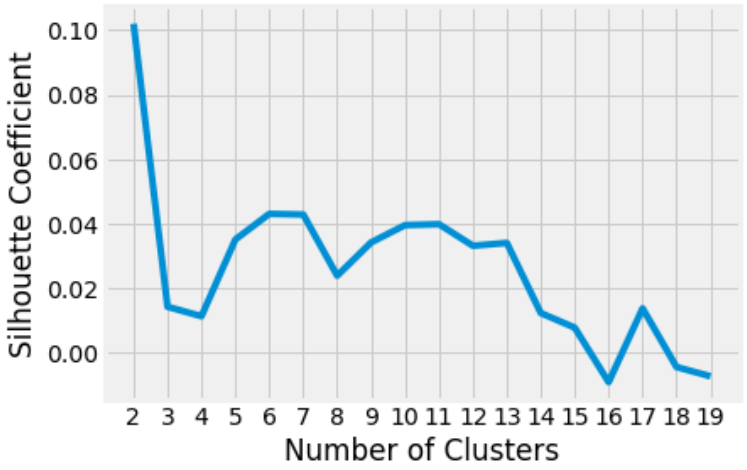
We use the country level USAID dataset and country level World Bank development indicator dataset merged together on the country column.
The data has been normalised to avoid giving equal weightage to all the columns while clustering in N-dimensional space to increase the robustness of the clusters.

K-means clustering

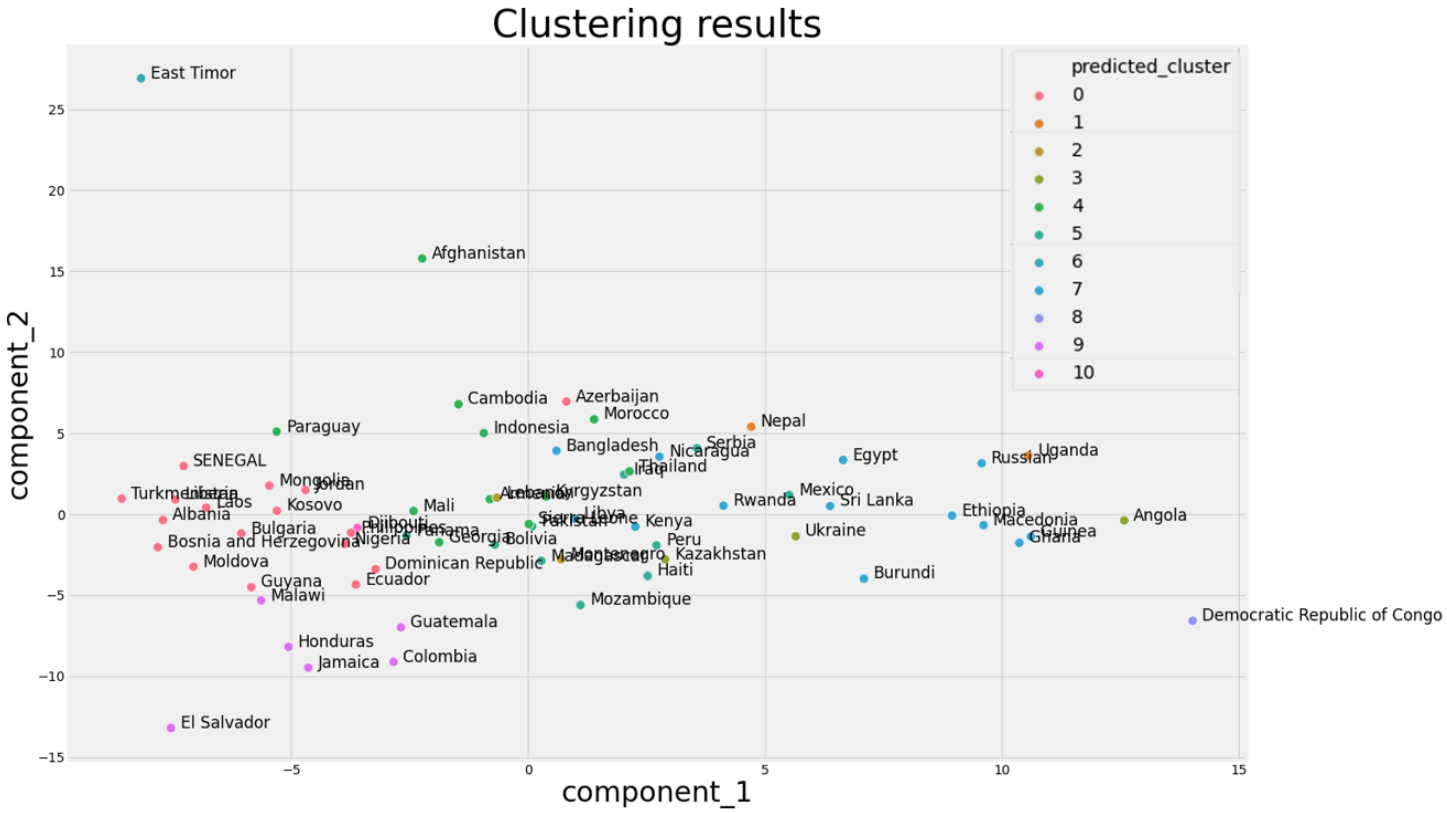
To find the similarities among the countries based on development indicators and the USAID country level data set information, we built a k-means clustering model.

To determine the optimal number of clusters, we applied the elbow method, which is to run k-means clustering on the dataset for a range of values of k (0 to 20), and for each value of k calculate the sum of squared errors (SSE). As we can see from the plot below, the line chart looks like an arm, then the "elbow" on the arm is the value of k that is the best, which should be 11. So we determine to have 11 clusters in our K-means clustering model.

The silhouette plot displays a measure of how close each point in one cluster is to points in the neighboring clusters and thus provides a way to assess parameters like number of clusters visually. From the silhouette plot below shows that when the number of clusters is 11, it has the silhouette coefficient of 0.04.



After determining the optimal number of clusters, we built a k-means clustering model for all the countries based on development indicators and the USAID country level data set information. The plot below shows the clustering results for all the countries. The x axis represents the PCA-1 value, and the y axis represents the PCA-2 value. There are 11 clusters in total which are denoted by different colors in the plot.

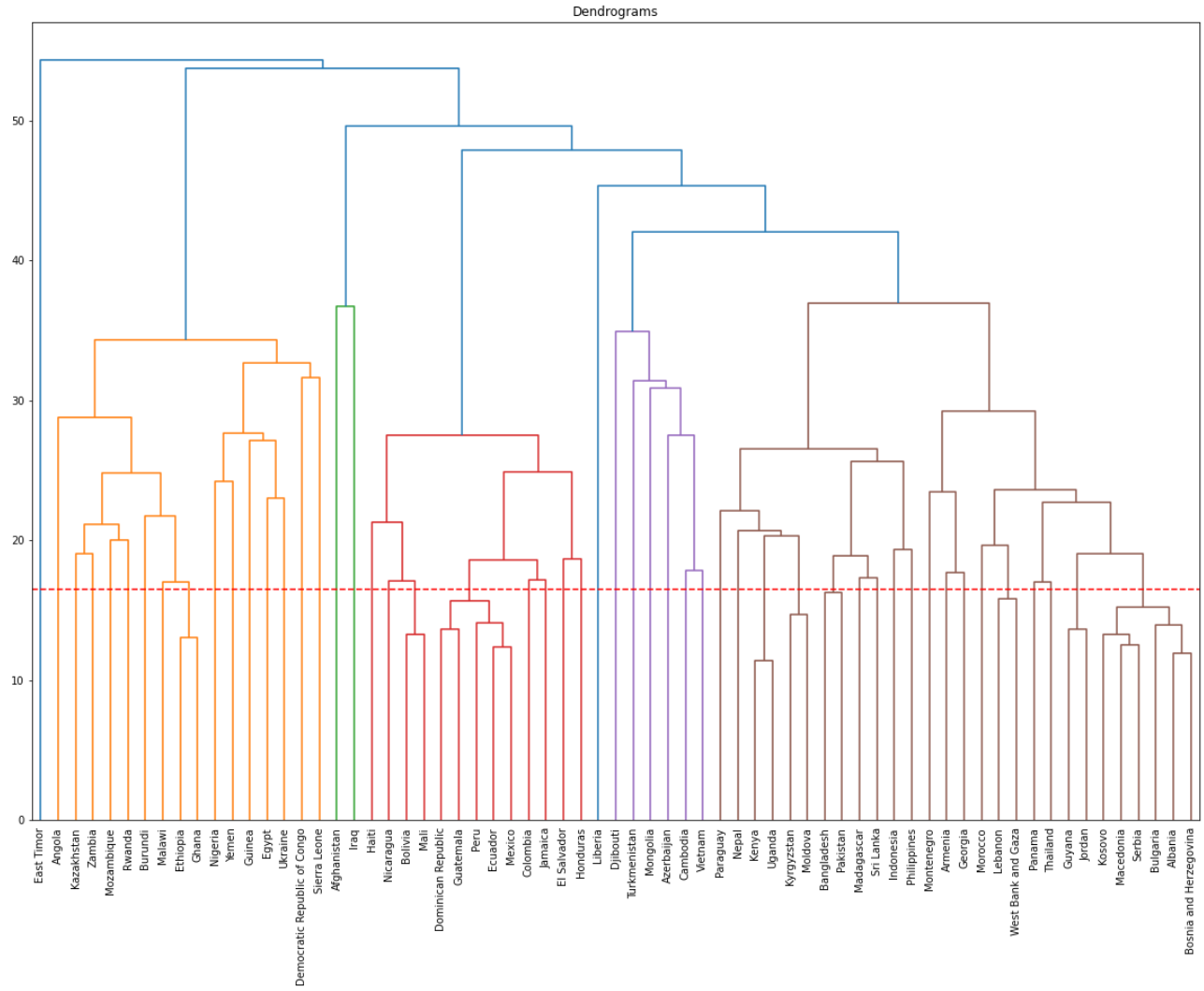


From the plot, we can see that though most clusters divide the countries nicely, there're still some overlaps among the clusters, such as cluster 3 and cluster 7. Both East Timor and Democratic Republic of Congo both belong to the clusters which have only one single country, and they are both at the edge of the plots, which means East Timor and Democratic Republic of Congo have little similarities with the other countries. The clustering for these countries could help USAID shape strategic action based on the countries with similar properties. For example, by using the clustering results as reference, USAID could consider applying the projects which resulted in good impacts to the countries within the same cluster of that country. The detailed PCA data and clustering results can be found in Appendix.

Hierarchical Clustering

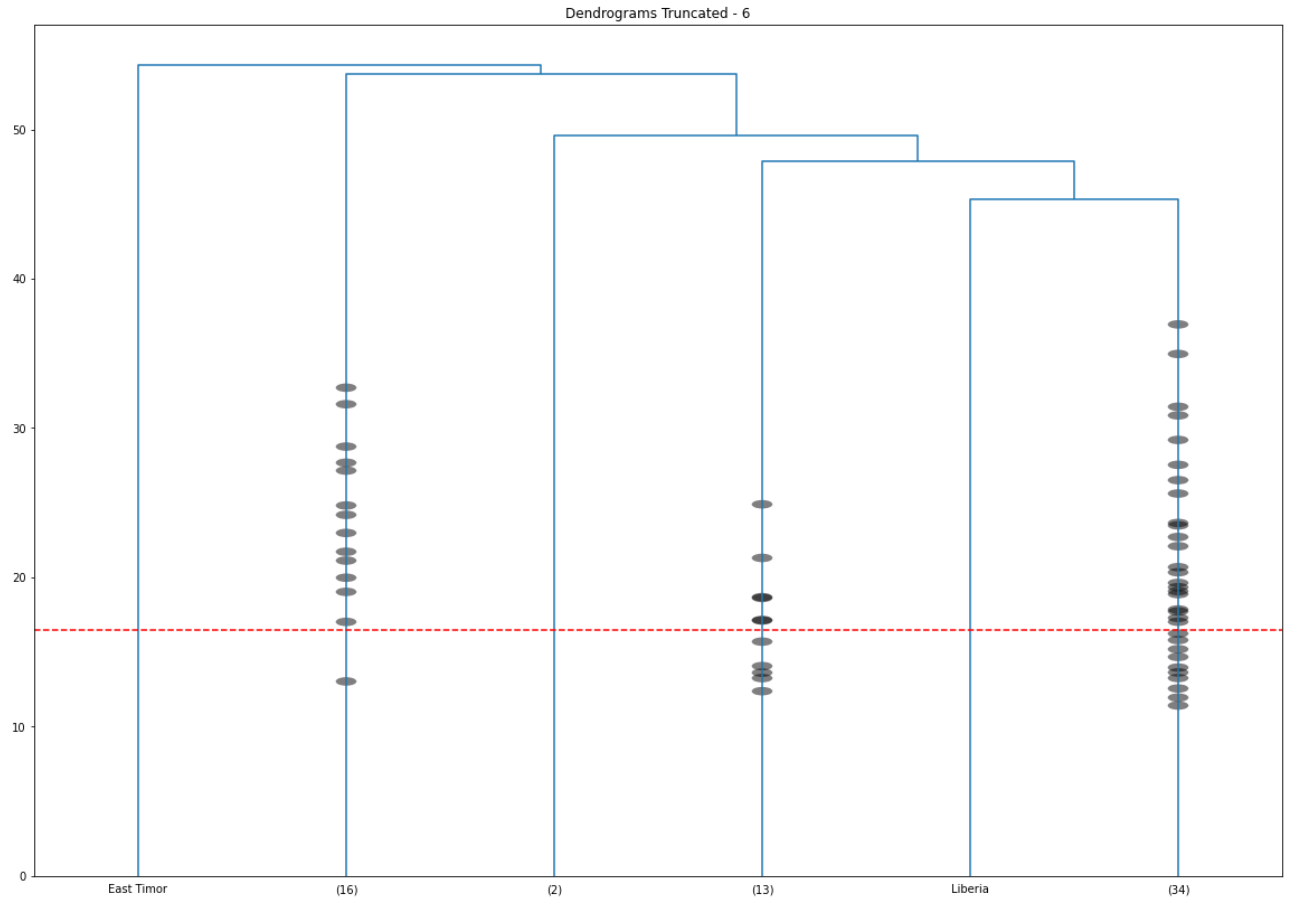
Using hierarchical clustering we get the following dendrogram
We have used absolute linkage between 2 observations to calculate the distance and Ward method. The logic behind choosing a value of distance to draw I got off height is based on a sudden increase in the absolute linkage difference.
With 67 countries in a data set we have 66 merges in total, and we can see that by eyeballing the cut-off line should lie somewhere between 15 and 25, which would give us outliers we need in each cluster.

The following clusters contain the last p joints, which will help us visualize outliers going top to bottom.



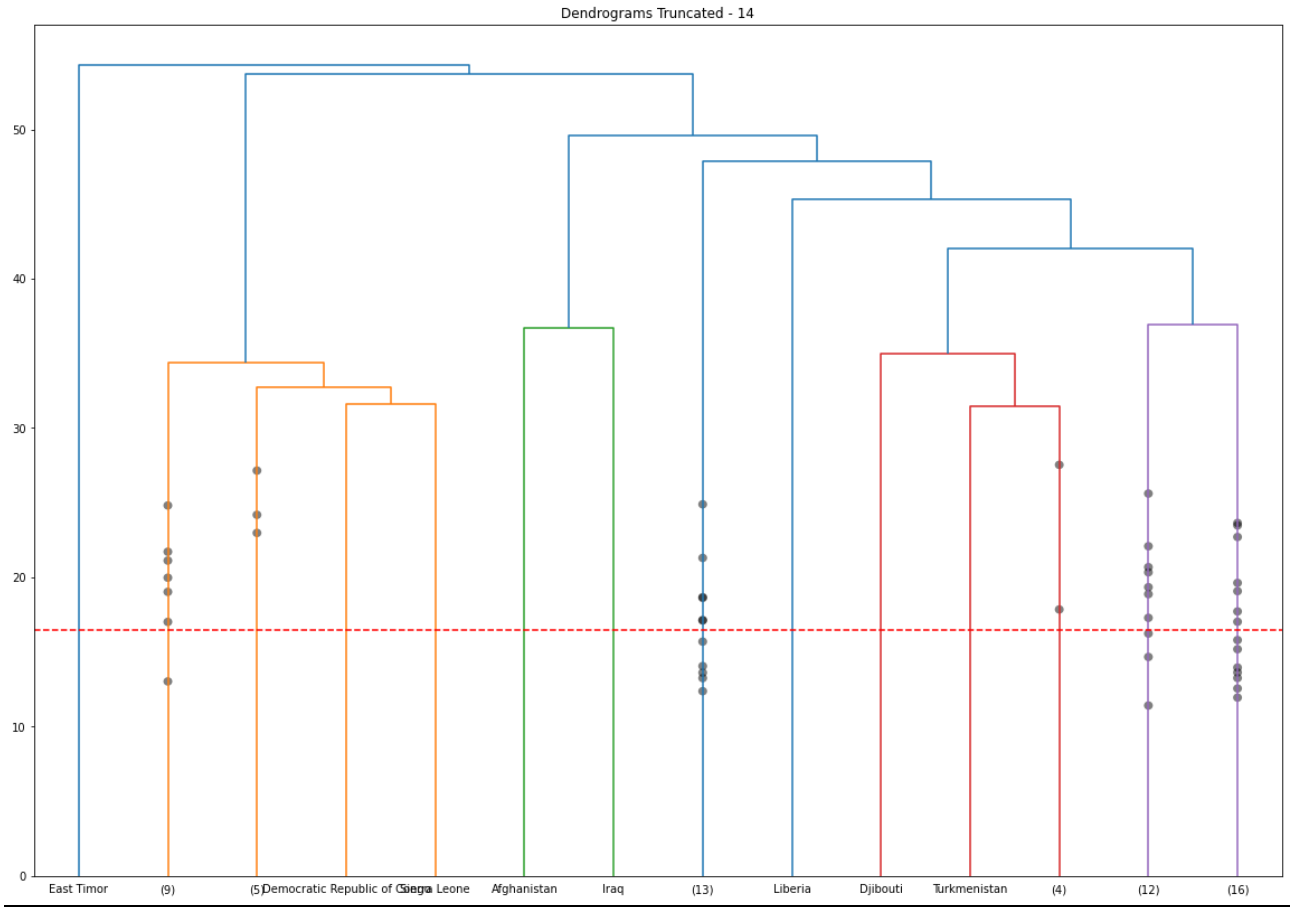
P = 6

We can see that East Timor and Liberia are two outliers in our data set because they don't belong to any definite cluster before they were fused into a cluster in the last 6 linkages.



P = 14

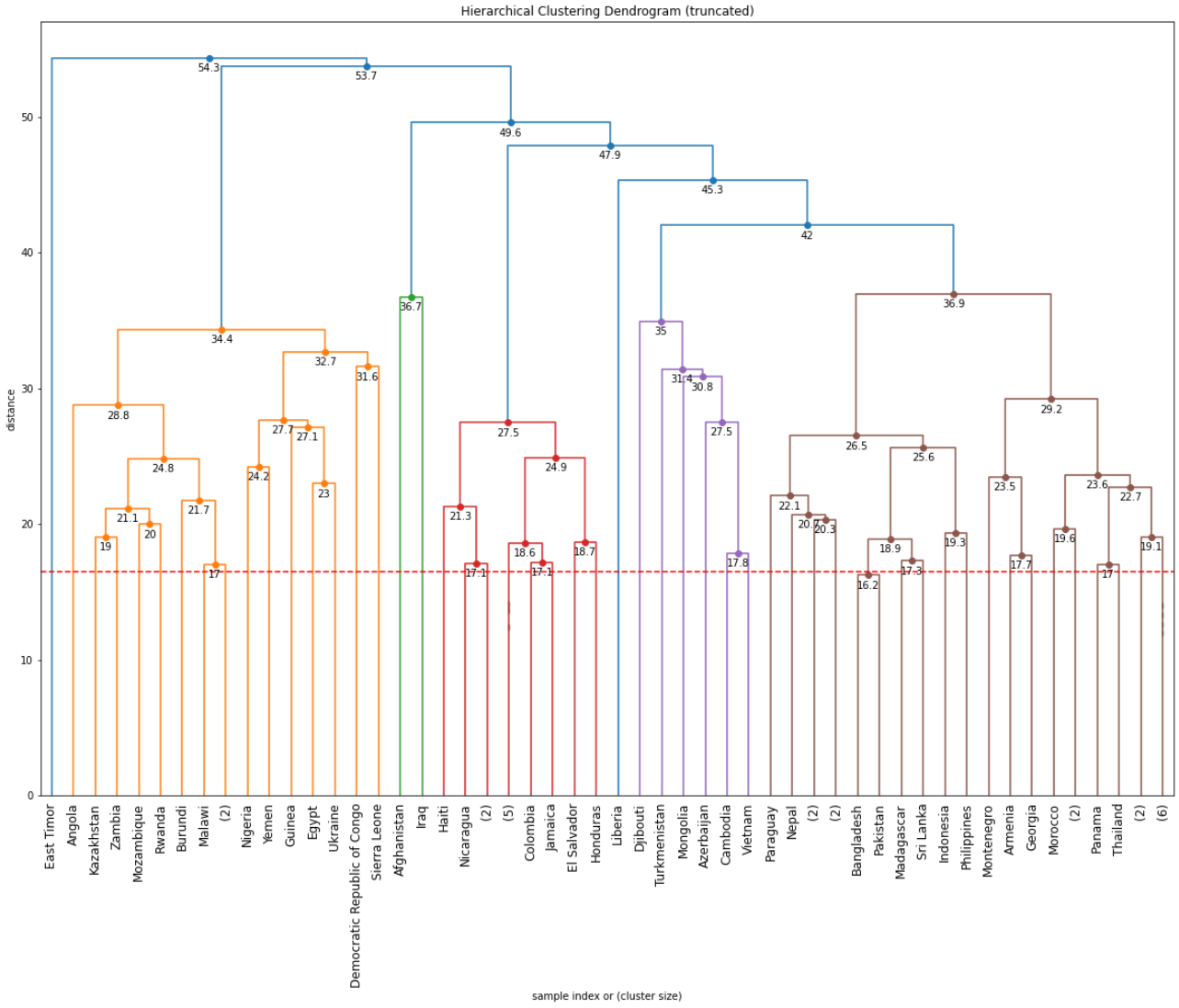
In addition to the previous two, Democratic Republic of Congo, Sierra Leone, Afghanistan, Iraq, Djibouti, Turkmenistan gets added into our linkages at the very end which are potential outliers.



Choosing the right cut off point:

Choosing the right cutoff point Was based on inconsistency method which uses different values of depth to calculate the standard deviation of difference between linkages and we can choose a cut of point where we see a sudden increase in standard deviation of the linkage, using this for different values of depth we chose the cut off height as **16.5** to get her potential outliers in the data set.

Using the above distance we can say that the following countries are outliers which later got added into the clusters:



6.2 Supervised Learning

Data

We used the country-level USAID dataset against the target variable to get the importance of different features within the dataset and correlate that to the success of projects in a particular country.
The data has been normalised to avoid crazy coefficients, and increase accuracy.

Data Preprocessing

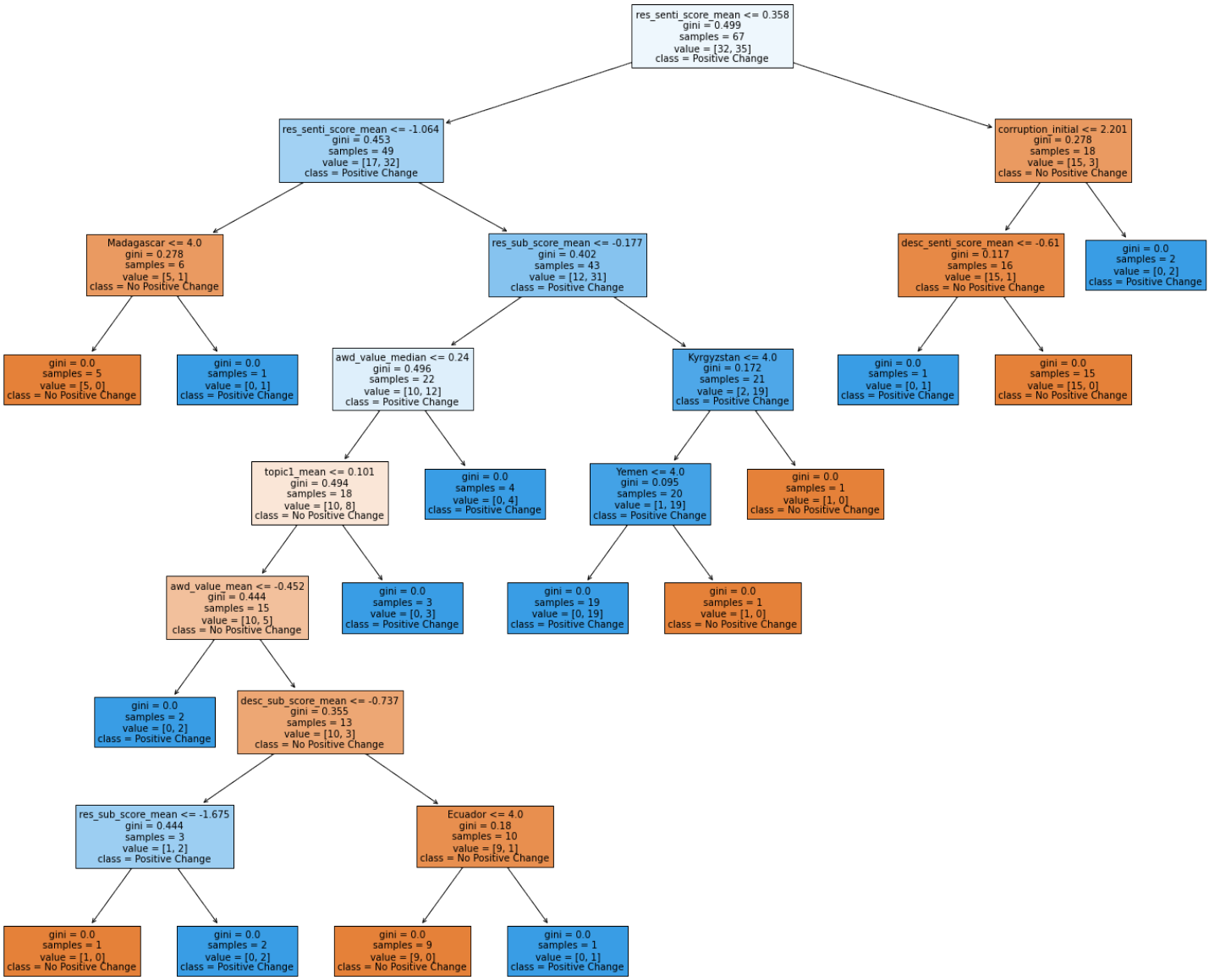
In our dataset, there are four ordinal categorical attributes(having natural ordering values). They are (1)corruption_initial, (2) stability_initial, (3)change_stability, (4)change_corruption.

We retained the order values by encoding them:
(1)corruption_initial_order= [High Corruption, Very High Corruption]
(2) stability_initia_order = [Unstable, Very Unstable]
(3)change_stability_order = [Worse, Same, Better]
(4)change_corruption_order = [Worse, Same, Better]

Moreover, we used countries as dummy variables so that we could compare their potential of success if invested by USAID in the end.

Classification and Regression Tree (CART)

Decision Tree classification results are presented in the image below.



Based on the decision tree graph above, we started with 67 samples at root and then split them into two child nodes with 49 and 18 samples. The cut-off point for that initial split is res_senti_score_mean<= 0.358. The graph also demonstrates that the most important variables for classification are res_senti_score_mean (the average of project result's sentiment scores) and corruption_initial (the initial corruption level of a country).
Looking at the final split, the decision tree with a depth of 8 performs well by separating class of

positive change or no positive change with gini=0 (gini measures how often the chosen elements are labeled incorrectly.)

Limits on Classification and Regression Tree:
Since our dataset is small, poor results can be easily drawn. Also, it is inadequate to predict continuous variables by applying decision tree algorithms. Moreover, it is hard to tell projects effectiveness in country-level, we decide to further our analysis of result we got from logistic regression and lasso regression

Logistics Regression with all countries

We applied Ridge and Lasso logistic regression. In both models, We achieved 100% accuracy, with 32/32 correct predictions for "Positive Change" and 35/35 correct predictions for "No Positive Change". Our dataset is quite balanced between the two classes. Thus, accuracy is a good measurement for model performance.

Recall (=TP/(TP+FN)) is the model's ability to find all positive data points in a dataset. Precision (=TP/(TP+FP)) tells us the probability of the model being correct when it predicts an observation to be positive. Both of the measurements are 100% as well.

	Precision	Recall
Positive Change	1	1
No Positive Change	1	1
Accuracy		
Macro Average	1	1
Weighted Average	1	1

Since our Lasso logistic regression models have excellent performance and their results are the easiest to interpret compared to those of CART's and Ridge, we have decided to choose Lasso Logistic Regression as our final models for supervised machine learning section. **Our recommendations that are presented in Section 7 will be based entirely on Lasso logistic regression models' results.**

Ridge Regression with all countries - Result Interpretation:

Top ten countries where USAID should invest in future anti-corruption projects:
From the results of Ridge logistic regression model, we selected the top 10 countries that have the highest coefficient, meaning the highest probability of getting a Positive outcome if invested.

	Variable	Coefficient	Odds
1	Ghana	0.475596	1.60897286

2	Bulgaria	0.428475	1.53491499
3	Ecuador	0.4221	1.52516103
4	Turkmenistan	0.380037	1.46233869
5	Mexico	0.346703	1.41439659
6	Jordan	0.336844	1.40052057
7	Thailand	0.310531	1.36414929
8	Kosovo	0.278918	1.32169896
9	Armenia	0.277331	1.31960309
10	Georgia	0.27214	1.31277078

Explanations for Top Three Countries:
Each country serves as a dummy variable. The countries whose coefficients are equal to 0 are the reference point. Since we don't have any countries whose coefficients are exactly 0, we picked Indonesia as the reference point as Indonesia's coefficient is 0.05, the closest to 0 out of all countries'. The odds of Indonesia will be 1.06.

Ghana:
If we invest in Ghana instead of Indonesia, the probability of having a Positive outcome will increase by 52% (1.609/1.06 - 1.06/1.06 =0.52). We divided both the coefficients of Ghana and Indonesia by 1.06 for scaling purposes.

Bulgaria:
If we invested in Bulgaria instead of Indonesia, the probability of having a Positive outcome will increase by 44.8% (1.5349/1.06-1.06/1.06=0.448).

Ecuador:
If we invested in Ecuador instead of Indonesia, the probability of having a Positive outcome will increase by 43.89% (1.5252/1.06 - 1.06/1.06 = 0.4389).

Top ten important features:

	Variable	Coefficient	Odds
1	awd_amount_median	0.426212	1.53144541
2	corruption_initial	0.412166	1.51008509
3	intervention_type_Enforcement_count	0.303623	1.35475822

4	Rule_of_Law_countLocal_Government_countGovern a...	0.261632	1.2990484
5	Rule_of_Law_countLocal_Government_countCivil_S.. .	0.251954	1.28653686
6	Civil_Society_count	0.23865	1.26953412
7	Media_count	0.226258	1.25389913
8	topic2_mean	0.216121	1.24125256
9	dominant_topic_2_count	0.214914	1.23975527
10	Rule_of_Law_countEconomic_Growth_count	0.196926	1.21765393

Explanation for the most influential variables:

1. **awd_amount_median:**

If we increase the number of projects that fall into the $10M - 25M bucket by 1 project, the probability of having a Positive outcome will increase by 53.14%(1.5314-1=0.5314), all else equal.

2. **corruption_initial:**

If USAID invests in a country labeled as "Very High Corruption" instead of "High Corruption", the probability of having a Positive outcome will increase by 51%, all else equal.

3. **intervention_type_Enforcement_count:**

If we increase the number of projects whose intervention type is Enforcement by 1 project, the probability of having a Positive outcome will increase by 35.48%, all else equal.

Lasso Regression with all countries - Result Interpretation:

Top countries where USAID should invest in future anti-corruption projects:

We further conducted a lasso logistic regression which performs feature selections (it can shrink multiple coefficients to 0). In the table below, we listed all countries that have a higher probability of getting a Positive outcome than the reference countries if we invested. Reference countries are countries whose coefficients are 0. The list of reference countries are Afghanistan, Albania, Angola, Bangladesh, Bosnia, Herzegovina, Burundi, Colombia, Democratic Republic of Congo, Ethiopia, Honduras, Indonesia, Iraq, Jordan, Kenya, Kosovo, Lebanon, Liberia, Mali, Mongolia, Mozambique, Nicaragua, Nigeria, Philippines, Rwanda, Serbia, Sierra Leone, Uganda, Ukraine, West Bank and Gaza.

	Country	**Coefficient**	**Odds**
1	Ghana	0.513189	1.67061029
2	Mexico	0.453834	1.57433664

3	Bulgaria	0.342294	1.40817424
4	Ecuador	0.321615	1.37935362
5	Guatemala	0.301838	1.35234213
6	Dominican Republic	0.26336	1.3012951
7	Jamaica	0.201935	1.22376846
8	Thailand	0.107474	1.11346191
9	Guinea	0.102459	1.10789188
10	Georgia	0.097731	1.10266613
11	Nepal	0.077016	1.08005936
12	Madagascar	0.055745	1.05732803
13	Turkmenistan	0.041618	1.04249617
14	Panama	0.040032	1.04084408
15	Sri Lanka	0.019561	1.01975357
16	Cambodia	0.015383	1.01550193
17	Moldova	0.014343	1.01444635
18	Armenia	0.004681	1.00469197

Explanation of Top three countries:

Ghana:

If we invested in Ghana instead of any of the reference countries, the probability of having a Positive outcome will increase by 67.01% (1.6701-1=0.6701), all else equal.

Mexico:

If we invested in Mexico instead of any of the reference countries, the probability of having a Positive outcome will increase by 57.43% (1.5743-1=0.5743), all else equal.

Bulgaria:

If we invested in Bulgaria instead of any reference countries, the probability of having a Positive outcome will increase by 40.82% (1.4082-1=0.4082), all else equal.

All features that have positive effects on anti-corruption projects' outcome:

	Variable	Coefficient	Odds
1	awd_amount_median	1.303142	3.68084373
2	corruption_initial	0.829138	2.29134275
3	topic2_mean	0.647296	1.9103682
4	intervention_type_Enforcement_count	0.623627	1.86568262
5	Rule_of_Law_countLocal_Governmen t_countGoverna...	0.329737	1.39060235
6	Civil_Society_count	0.228148	1.25627124
7	change_stability	0.196529	1.21717062
8	Rule_of_Law_countLocal_Governmen t_countCivil_S...	0.185308	1.20358909
9	Media_count	0.175152	1.1914273
10	Local_Government_countMedia_count Governance_count	0.156955	1.16994297
11	Rule_of_Law_countEconomic_Growt h_count	0.104892	1.11059066
12	dominant_topic_2_count	0.028854	1.02927431

Explanation of these variables:

1. **awd_amount_median:**

If we increase the number of projects that fall into the $10M - 25M bucket by 1 project, the probability of having a Positive outcome will increase by 268% (3.68 - 1= 2.68), all else equal.

2. **corruption_initial**

If USAID invests in a country labeled as "Very High Corruption" instead of "High Corruption", the probability of having a Positive outcome will increase by 229%, all else equal.

3. **topic2_mean**

If USAID executes more projects that fall into topic 2 - Corruption reduction in the health sector, then the probability of having a Positive outcome will increase, all else equal.

4. **intervention_type_Enforcement_count**

If we increase the number of projects whose intervention type is Enforcement by 1 project, the probability of having a Positive outcome will increase by 86.57%, all else equal.

5. **Rule_of_Law_countLocal_Government_countGovernance_count**

If the most common sector of anti-corruption projects in a country is "Rule of Law, Local Government, and Governance", the probability of having a Positive outcome will increase by 39.06% compared to countries where the most common sector is a reference sector (including Independent_Agencies, Rule_of_Law, Elections_and_Political_Processes, Local_Government,

Legislature, Economic_Governance, Disaster_Recovery, Economic_Growth, Civil_Society, Governance, Healthcare, Public_Finance, Private_Sector, Democracy_and_Governance, Gender_Equality, Natural_Resource, Education), all else equal.

6. **Civil_Society_count**

If the most common sector of anti-corruption projects in a country is Civil Society, the probability of having a Positive outcome will increase by 25.62% compared to countries where the most common sector is a reference sector, all else equal.

7. **change_stability**

If USAID invests in a country whose stability is "worse" after the 2007-2013 anti-corruption project, the probability of having a Positive outcome will increase by 21.72% compared to investing in countries whose stability did not change, all else equal. If USAID invests in a country whose stability is "same" after the 2007-2013 anti-corruption project, the probability of having a Positive outcome will increase by 21.72% compared to investing in countries whose stability is better, all else equal. If USAID invests in a country whose stability is "worse" after the 2007-2013 anti-corruption project, the probability of having a Positive outcome will increase by 48.15% compared to investing in countries whose stability is better, all else equal.

8. **Rule_of_Law_countLocal_Government_countCivil_Society_count**

If the most common sector of anti-corruption projects in a country is "Rule of Law, Local Government, and Civil Society", the probability of having a Positive outcome will increase by 20.36% compared to countries where the most common sector is a reference sector, all else equal.

9. **Media_count**

If we increase the number of projects whose sector is Media by 1 project, the probability of having a Positive outcome will increase by 19.14%, all else equal.

10. **Local_Government_countMedia_countGovernance_count**

If the most common sector of anti-corruption projects in a country is "Local Government, Media, and Governance", the probability of having a Positive outcome will increase by 17% compared to countries where the most common sector is a reference sector, all else equal.

11. **Rule_of_Law_countEconomic_Growth_count**

If the most common sector of anti-corruption projects in a country is "Rule of Law, Economic Growth", the probability of having a Positive outcome will increase by 11.06% compared to countries where the most common sector is a reference sector, all else equal.

12. **dominant_topic_2_count**

If we increase the number of projects whose dominant topic is topic 2 - Corruption reduction in the health sector by 1 project, the probability of having a Positive outcome will increase by 2.93%, all else equal.

Features that most NEGATIVELY affects a project's success:

	Variable	Coefficient	Odds
2	Local_Government_count	-0.4985	0.607444
3	Agriculture_and_Food_Security_count	-0.34321	0.709491
4	Economic_Growth_count	-0.32662	0.721361

5	Environment_count	-0.27889	0.756621
6	Public_Finance_count	-0.2522	0.777091
7	intervention_type_Citizen_Media_Business_Engagement_Participation_count	-0.23019	0.794381
8	Rule_of_Law_countLegislature_countEconomic_Growth_countCivil_Society_countPublic_Finance_count	-0.1592	0.852822
9	dominant_topic_1_count	-0.07378	0.928879
10	Rule_of_Law_countGovernance_count	-0.0571	0.944504
11	Civil_Society_countGovernance_countHealthcare_count	-0.04778	0.953342
12	Local_Government_countEconomic_Growth_count	-0.04332	0.957601

From the table above, USAID should not invest in more projects whose sector is either Local_Government, Agriculture_and_Food_Security, Economic_Growth, Environment, Public_Finance, etc. Moreover, USAID should not invest in projects whose dominant theme is Technical support for public and private investment.

Logistic Regression: The case with no outlier countries

Since removing outliers is an appropriate step in data preprocessing, we decided to execute Ridge and Lasso logistic regression modeling without the outlier countries pointed out by our clustering model in section 6.1. The list of outlier countries is below. These countries have such distinct characteristics that we cannot group them into any clusters.

Country Name	Result of investments
Afghanistan	No Positive Change
Angola	No Positive Change
Azerbaijan	No Positive Change
Democratic Republic of Congo	Positive Change
Djibouti	No Positive Change
East Timor	No Positive Change

Iraq	Positive Change
Liberia	No Positive Change
Mongolia	Positive Change
Sierra Leone	No Positive Change
Turkmenistan	Positive Change

The output of the models without the country outliers will give us more specific results while the models with all countries will give more general results.

Regarding performance, the models also have 100% accuracy with precision and recall both at 100% (perfect accuracy). We will focus more on the Lasso result interpretation as the model performs feature selection.

Ridge with no outlier countries - Result Interpretation:

Top ten countries where USAID should invest in future anti-corruption projects:

Country	Coefficient	odds
Ghana	0.449834	1.568051
Bulgaria	0.403451	1.496982
Ecuador	0.395949	1.485794
Jordan	0.363205	1.437931
Mexico	0.359554	1.43269
Kosovo	0.308133	1.360882
Thailand	0.296537	1.345192
Georgia	0.293328	1.340883
Bangladesh	0.250768	1.285012
Guatemala	0.250595	1.28479

Top Features that have the most positive effects on anti-corruption projects' outcome:

Variable	Coefficient	Odds
awd_amount_median	0.41155	1.509155
Rule_of_Law_countLocal_Government_countGovernance_count	0.272617	1.313397
Civil_Society_count	0.267754	1.307025
Media_count	0.260213	1.297206
Rule_of_Law_countLocal_Government_countCivil_Society_count	0.247705	1.281082
Independent_Agencies_count	0.227936	1.256005
dominant_topic_2_count	0.194899	1.215188
Rule_of_Law_countEconomic_Growth_count	0.194551	1.214765
topic2_mean	0.193207	1.213134
change_stability	0.157676	1.170787

If USAID would like to better the outcome of anti-corruption project, this model tells us that they should increase the number of projects whose award amount is between $10M to $25M, and invest in projects whose sector is "Rule_of_Law, Local_Government, Governance", " Civil Society", "Media", for example. USAID should also invest in more projects that fall into the theme of "Corruption reduction in the health sector", as well as invest in countries with the "Very Unstable" initial stability.

Lasso with no outlier countries - Result Interpretation:

Top countries where USAID should invest in future anti-corruption projects:

In the table below, we listed all countries that have a higher probability of getting a Positive outcome than the reference countries if we invested. Reference countries are countries whose coefficients are 0. The list of reference countries are Albania, Armenia, Bangladesh, Bosnia and Herzegovina, Burundi, Cambodia, Colombia, Dominican Republic, Egypt, Georgia, Guinea, Honduras, Indonesia, Jamaica, Jordan, Kenya, Lebanon, Madagascar, Malawi, Moldova, Montenegro, Nepal, Nicaragua, Pakistan, Panama, Philippines, Rwanda, Serbia, Sri Lanka, Uganda, Ukraine, and West Bank and Gaza.

	Country	Coefficient	Odds
1	Ghana	0.382394	1.465789
2	Ecuador	0.35256	1.422705
3	Kosovo	0.3056	1.357439
4	Bulgaria	0.29286	1.340255
5	Mexico	0.091384	1.09569
6	Thailand	0.029502	1.029941
7	Guatemala	0.017207	1.017356

Explanation of Top three countries:

Ghana:

If we invested in Ghana instead of any of the reference countries, the probability of having a Positive outcome will increase by 46.58%, all else equal.

Ecuador:

If we invested in Ecuador instead of any of the reference countries, the probability of having a Positive outcome will increase by 42.27%, all else equal.

Kosovo:

If we invested in Kosovo instead of any reference countries, the probability of having a Positive outcome will increase by 35.74%, all else equal.

All features that have positive effects on anti-corruption projects' outcome:

	Variable	Coefficient	Odds
1	awd_amount_median	0.553227	1.738855
2	topic2_mean	0.162257	1.176162
3	Media_count	0.103505	1.109051
4	Civil_Society_count	0.07059	1.073141

5	Rule_of_Law_countLocal_Government_countGovernance_count	0.064787	1.066932
6	Rule_of_Law_countLocal_Government_countCivil_Society_count	0.019442	1.019632

Explanation of these variables:

1. **awd_amount_median:**

If we increase the number of projects that fall into the $10M - 25M bucket by 1 project, the probability of having a Positive outcome will increase by 73.89%, all else equal.

2. **topic2_mean**

If USAID executes more projects that fall into topic 2 - Corruption reduction in the health sector, then the probability of having a Positive outcome will increase, all else equal.

3. **Media_count**

If we increase the number of projects whose sector is Media by 1 project, the probability of having a Positive outcome will increase by 19.14%, all else equal.

4. **Civil_Society_count**

If the most common sector of anti-corruption projects in a country is Civil Society, the probability of having a Positive outcome will increase by 7.31% compared to countries where the most common sector is a reference sector, all else equal.

5. **Rule_of_Law_countLocal_Government_countGovernance_count**

If the most common sector of anti-corruption projects in a country is "Rule of Law, Local Government, and Governance", the probability of having a Positive outcome will increase by 6.69% compared to countries where the most common sector is a reference sector (including Independent_Agencies, Rule_of_Law, Elections_and_Political_Processes, Local_Government, Legislature, Economic_Governance, Disaster_Recovery, Economic_Growth, Civil_Society, Governance, Healthcare, Environment, Private_Sector, Democracy_and_Governance, Gender_Equality, Natural_Resource, Education, Economic_Growth, Healthcare, Independent_Agencies, Local_GovernmentMediaGovernance, Rule_of_Law, Rule_of_LawEconomic_Growth, Rule_of_LawLocal_Government), all else equal.

6. **Rule_of_Law_countLocal_Government_countCivil_Society_count**

If the most common sector of anti-corruption projects in a country is "Rule of Law, Local Government, and Civil Society", the probability of having a Positive outcome will increase by 1.96% compared to countries where the most common sector is a reference sector, all else equal.

Features that most NEGATIVELY affects a project's success:

	Variable	Coefficient	Odds
1	Local_Government_count	-1.18255	0.306495
2	Governance_count	-0.83614	0.433381
3	awd_amount_very_high	-0.83509	0.433836

4	Public_Finance_count	-0.73003	0.481893
5	intervention_type_Citizen_Media_Business_Engagement_Participation_count	-0.31935	0.72662
6	awd_amount_median_high	-0.27421	0.760176
7	dominant_topic_1_count	-0.22816	0.795995
8	Local_Government_countEconomic_Growth_count	-0.18391	0.832009
9	Rule_of_Law_countGovernance_count	-0.15224	0.858779
10	Agriculture_and_Food_Security_countGovernance_countPublic_Finance_countDemocracy_and_Governance_count	-0.09572	0.908714

Explanations of 4 of the worst features:

1. **Local_Government_count:**

If we increase the number of projects whose sector is Local Government alone by 1 project, the probability of having a Positive outcome will decrease by 69.35%, all else equal.

2. **Governance_count**

If we increase the number of projects whose sector is Governance alone by 1 project, the probability of having a Positive outcome will decrease by 56.66%, all else equal.

3. **awd_amount_very_high**

If we increase the number of projects that fall into the > $100M bucket by 1 project, the probability of having a Positive outcome will decrease by 56.62%, all else equal.

4. **dominant_topic_1_count**

If USAID executes 1 more project that mostly falls into the theme of Technical support for public and private investment, then the probability of having a Positive outcome will decrease by 20.4%, all else equal.

6.3 Temporal analysis for WDI database.

To explore the WDI (World Bank Development Indicator) dataset in detail, we created line plots for each of the 20 indicators by country level and region level in Tableau. By checking the trend of each plot, finding outliers, we analyzed how these indicators changed over years in different countries and regions, and how USAID anti-corruption projects impacted the development indicators in these countries.

The shaded area is the overlapping period between the cleaned USAID project dataset and the World Bank's Development Indicator dataset.

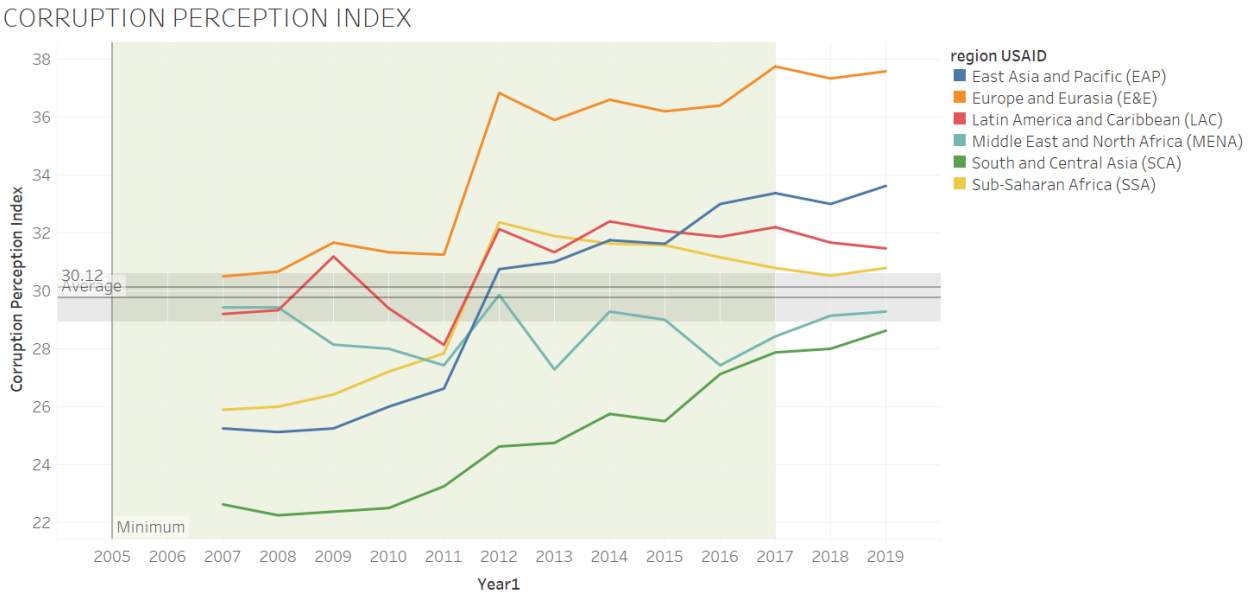
Corruption Perception Index (CPI):

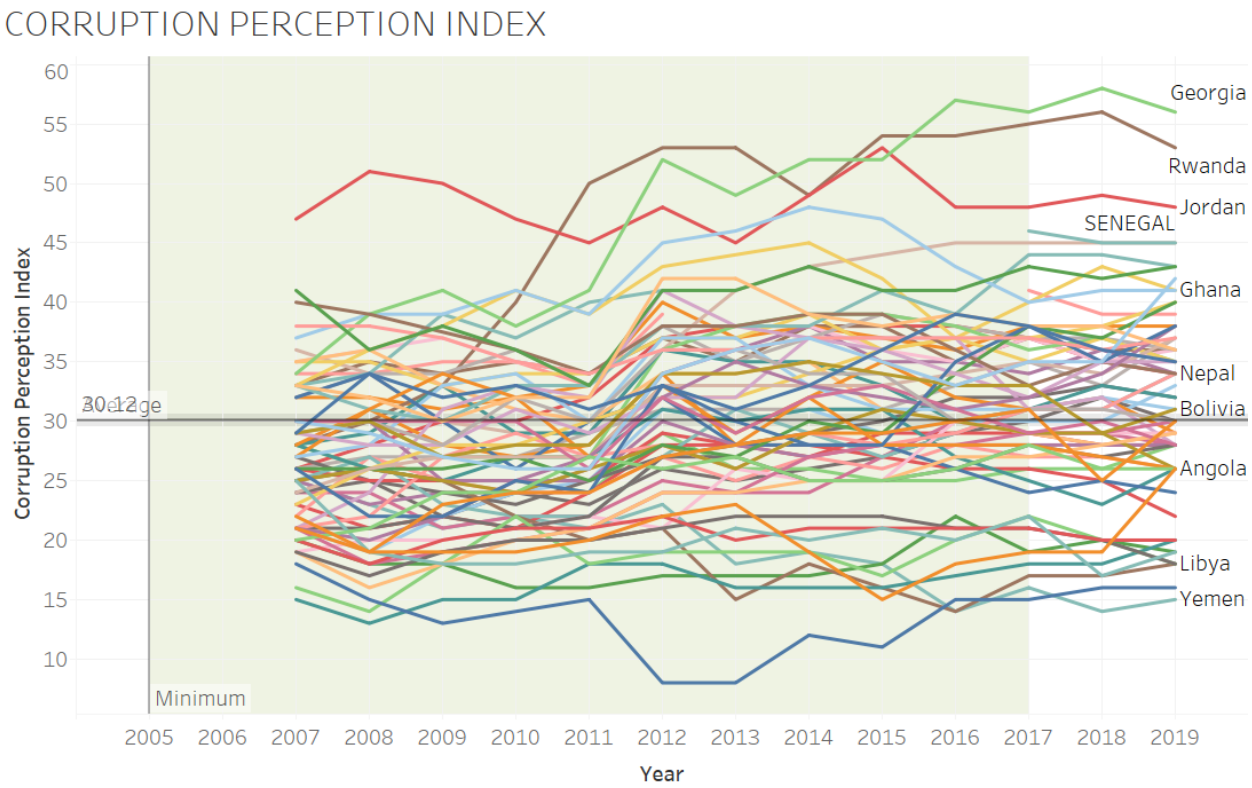
CPI shows the long-term trend in corruption level. The higher the index, the cleaner a country regarding corruption. The scale is 1-100. Since CPI is long-term and USAID's projects focus more on immediate changes, CPI may not reflect the success of USAID anti-corruption project.

Globally

(from now on, when we say "globally", we mean "countries where USAID's anti-corruption projects took place".)

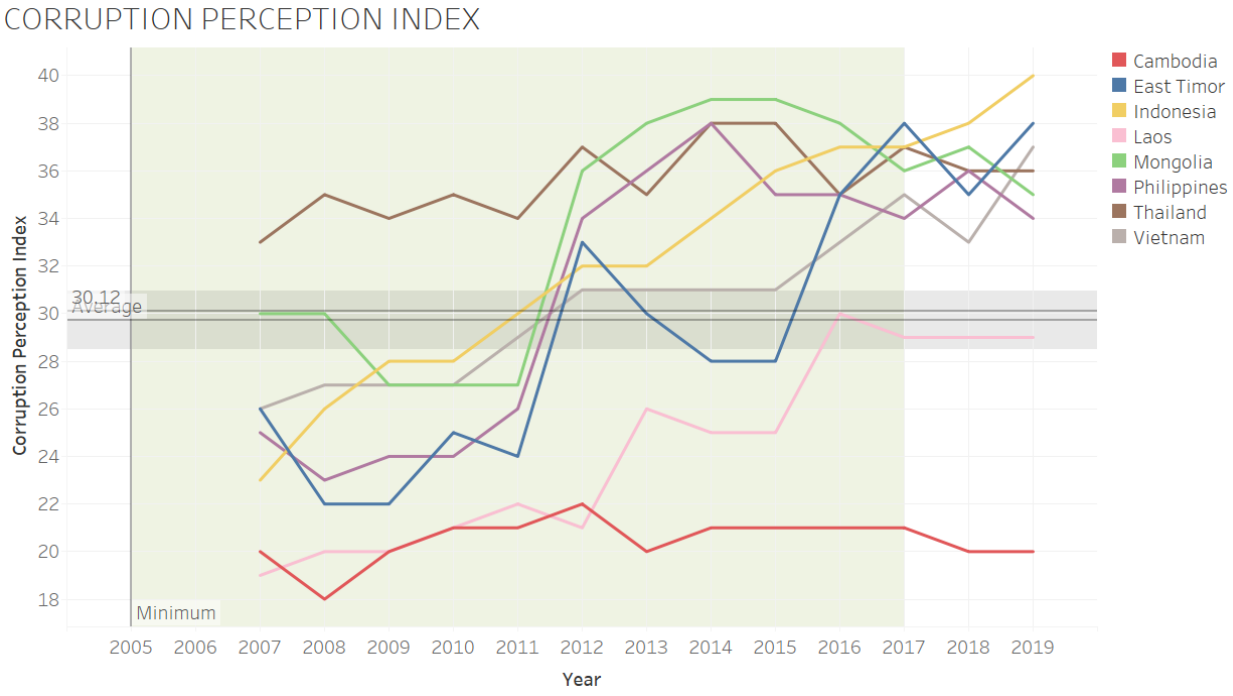
The global average of CPI is 30.12. The most corrupt region is SCA and MENA. All regions except MENA have shown considerable improvement. This may indicate that although MENA was invested the most, the region still needs more support in anti-corruption.





The most corrupt country is Yemen, followed by Afghanistan. Yemen may need more help from USAID.

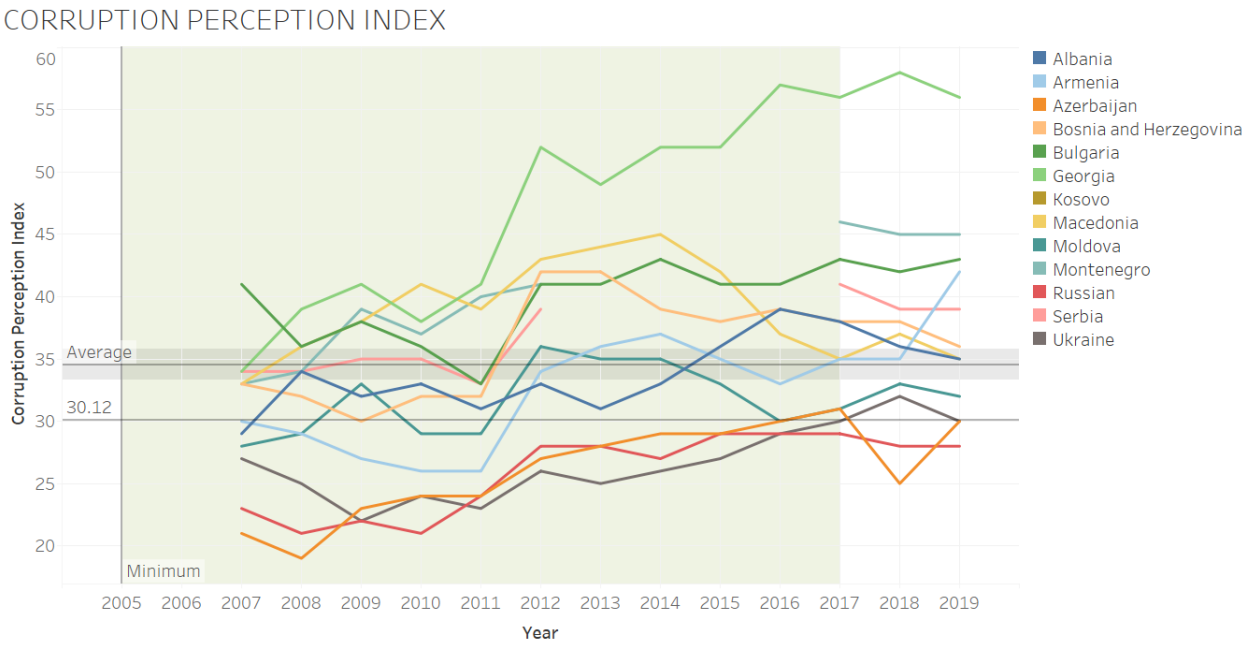
a) East Asia and Pacific (EAP)



The regional average is comparable to the global average. **The improvement in Indonesia, the fifth most invested country by USAID, proved the effectiveness of the projects in this country.** East

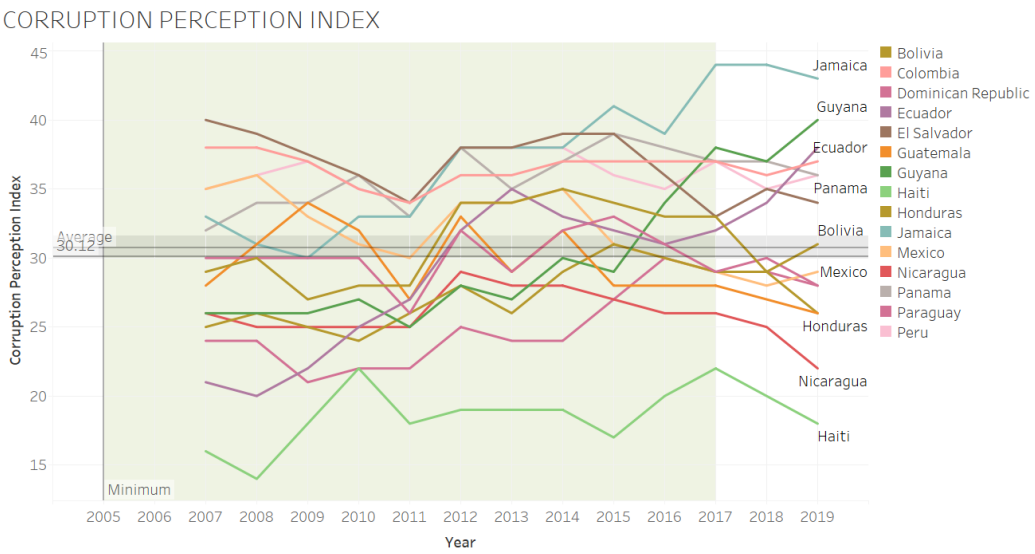
Timor, Vietnam, Mongolia and Laos have also improved considerably. The most worrisome country is Laos, with CPI constantly at a low level, meaning high corruptions.

b) Europe and Eurasia (E&E)



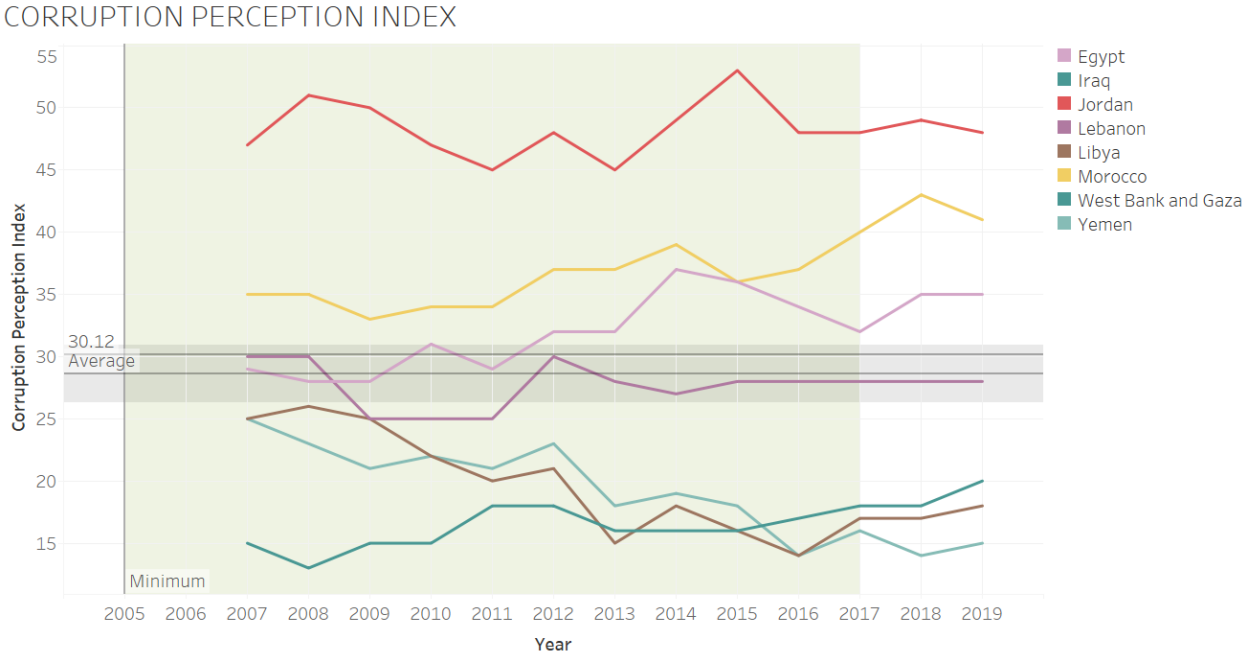
The regional average is 34.53, well above the global average. Most countries have been improving since 2007, except Macedonia. Macedonia's CPI improved during the period of 2007 to 2015, but then has become worse since 2016. The last USAID projects in Macedonia ended in 2014. This may indicate that the **projects did help Macedonia become cleaner, but the effect was not long-lasting after the investment stopped.**

c) Latin America and Caribbean (LAC)



The regional average (30.76) is comparable to the global average. Haiti is the worst regarding corruption, while Jamaica is the cleanest in the region. The improved group include Jamaica, Guyana, Ecuador, and Haiti. The hardly-change group includes Colombia, Panama, Peru, Bolivia, Honduras. Finally, the getting-worse group include El Salvador, Mexico, Dominican Republic, Guatemala, and Nicaragua. The most impressive positive changes are from Jamaica and Guyana. The worst downfalls are from El Salvador and Mexico.

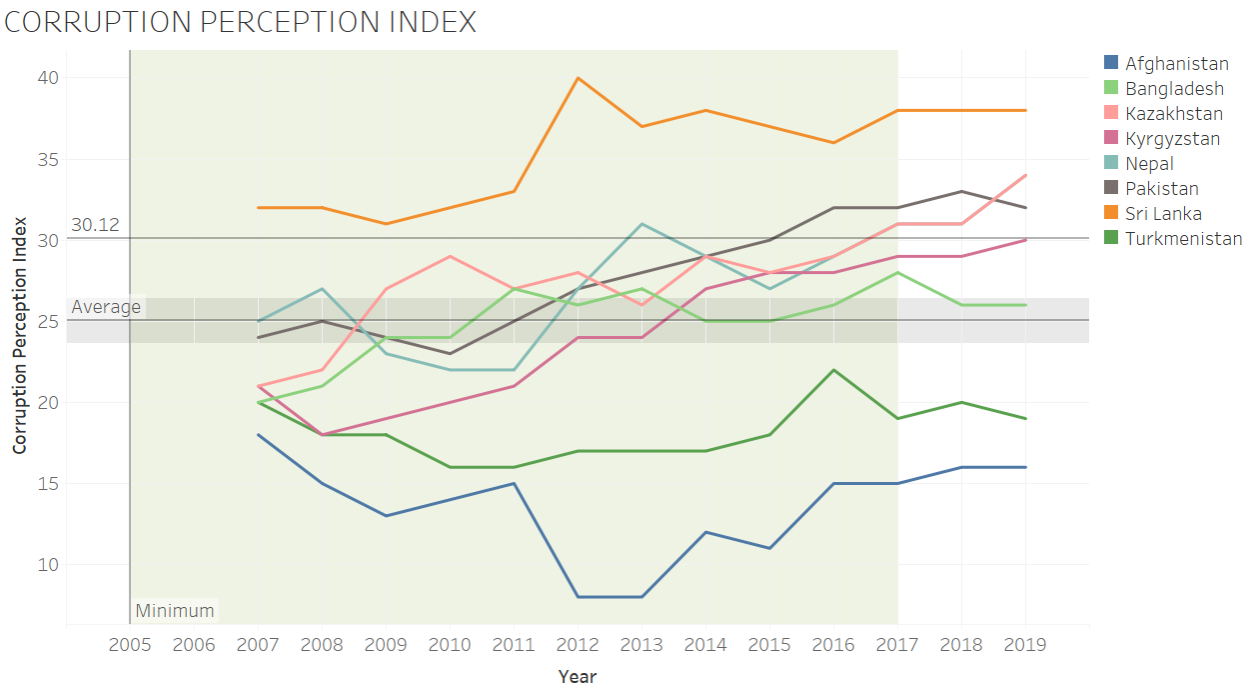
d) Middle East and North Africa (MENA)



The regional average of 28.6 is **below the global average**. MENA is also the most invested region by USAID anti-corruption projects out of the six regions. The most corrupt countries are Yemen, Libya and Iraq. While Yemen and Libya is getting worse, **Iraq (the 2nd most invested country by USAID)**

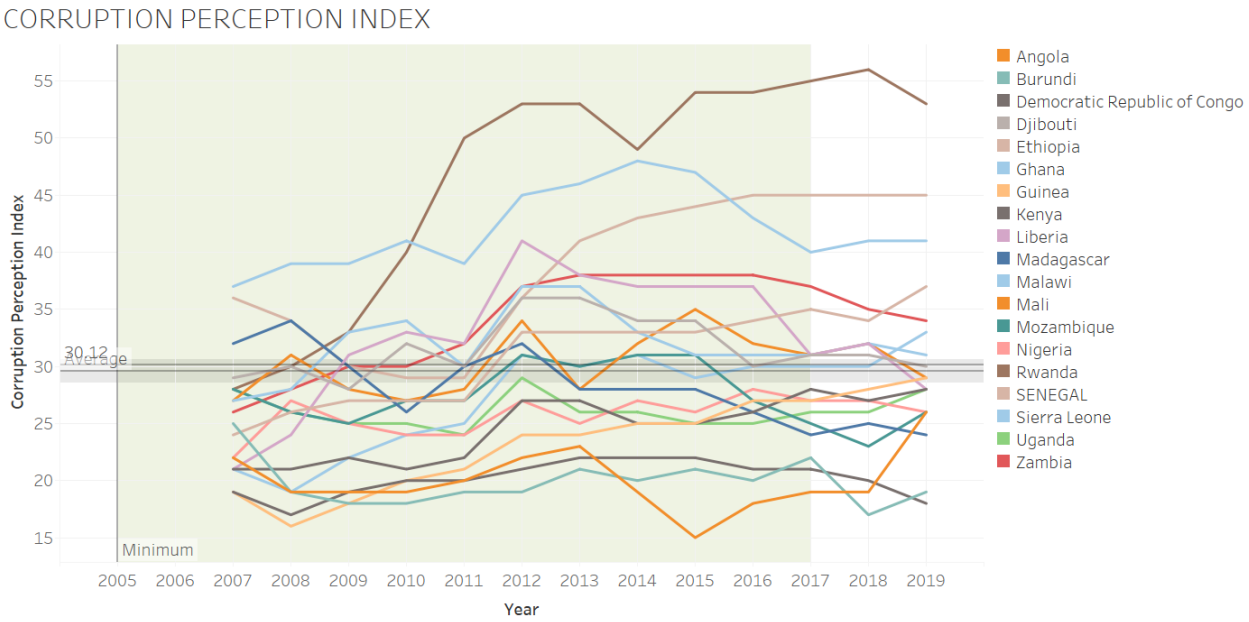
has improved by 5 points since 2007. Jordan is the cleanest country in MENA, leaving Morocco as a far 2nd. Jordan is the 7th most invested country by USAID.

e) South and Central Asia (SAC)



The regional average of 25 is well **below the global average.** All countries have improved compared to 2007, except Afghanistan (the most invested country by USAID) and Turkmenistan. Sri Lanka is the region's cleanest. Afghanistan is the worst one. Afghanistan went through a decline in CPI from 2007 to 20016, but has come back better. The situation could have been worse without USAID's investments.

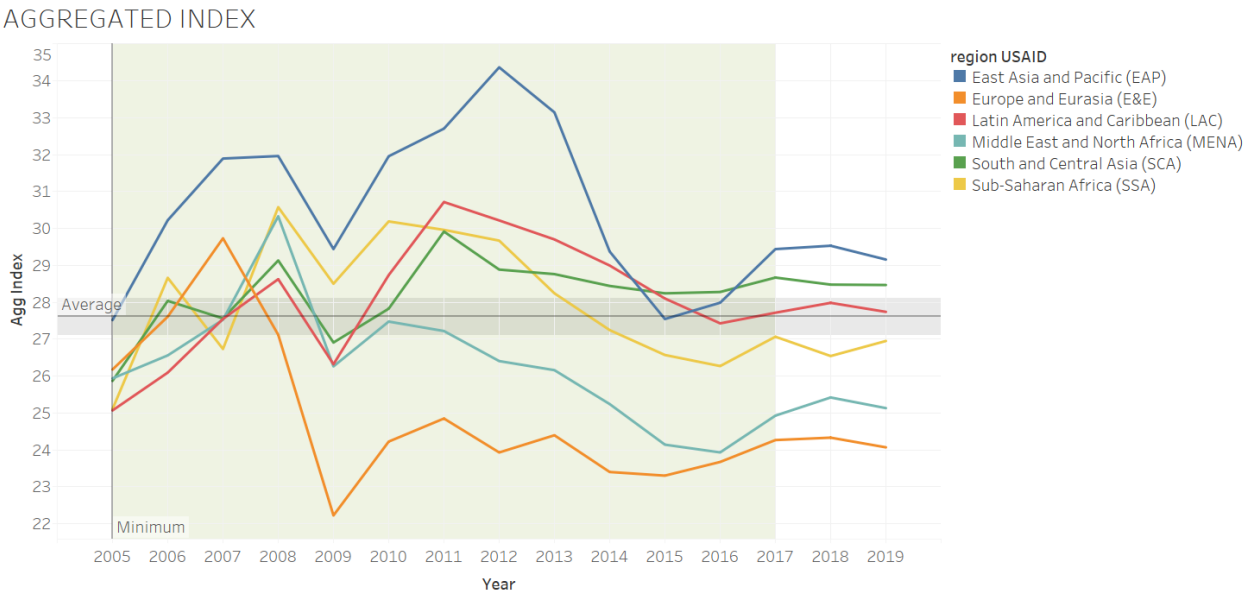
f) Sub-Saharan Africa (SSA)



The regional average of 29.6 is comparable to the global average. Rwanda shows the most improvement and has been the cleanest country of the region since 2010. Senegal has also improved a lot. Democratic Republic of Congo is the most corrupted one in the region in 2019. The bad news is the country has been experiencing a downward trend since 2014. Liberia and Madagascar need more attention as both countries are also trending downward regarding CPI.

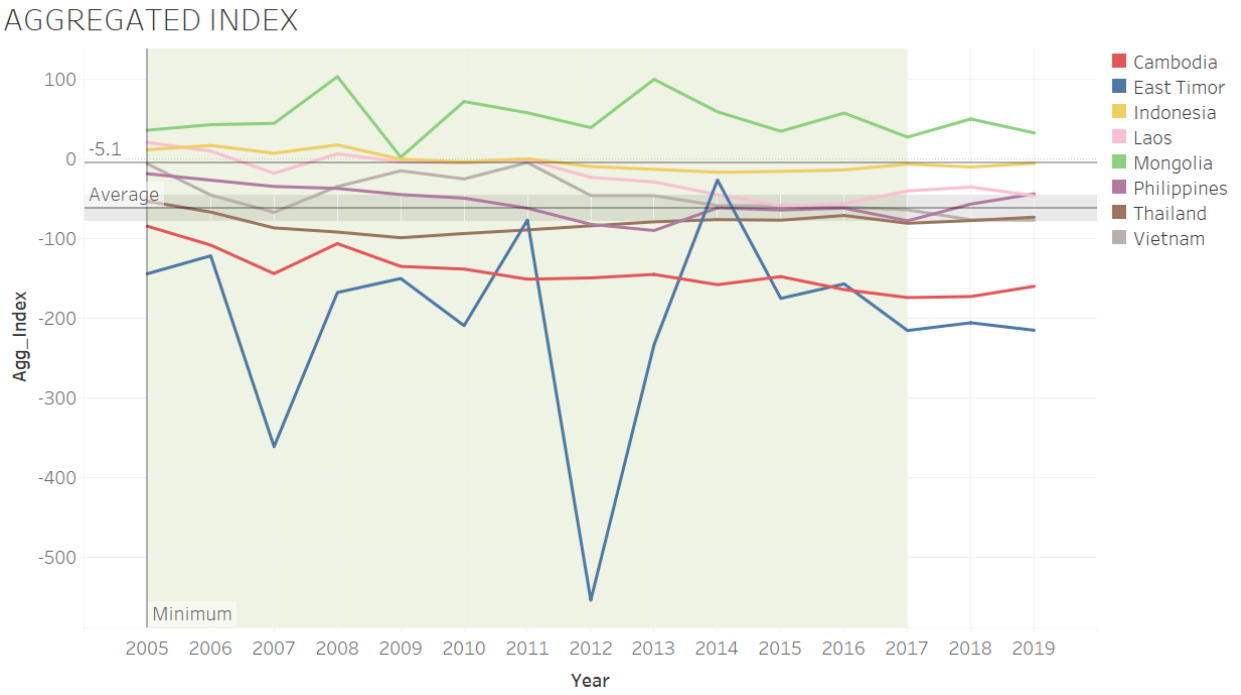
A high-level view: Our aggregated index

The aggregated index is the weighted average (with direction) of the top 20 indicators that were selected by our XGBoost model based on their relevance and importance (F-score) to the anti-corruption topic. **The higher this index, the better developed a country is. This is the clearest evidence for whether USAID anti-corruption projects were successful as the index reflects the immediate impacts on development indicators.**



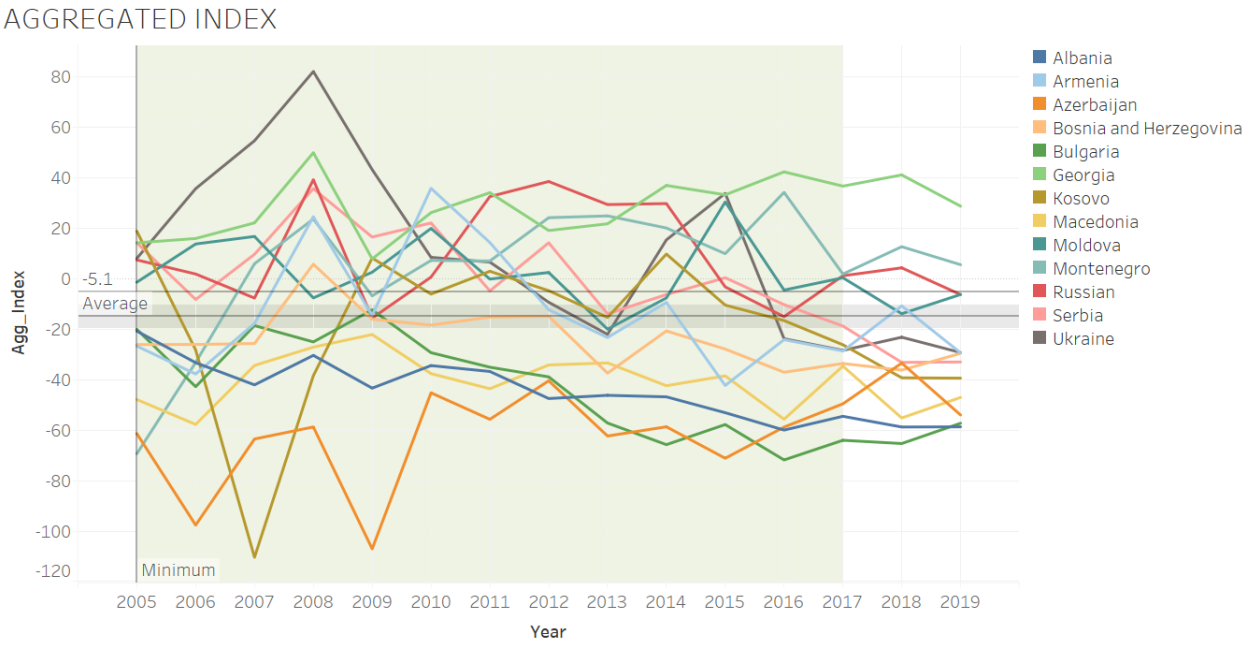
The least developed regions that need attention and support are E&E, MENA, and SSA. Based on this index, SCA has been doing quite well.

a) East Asia and Pacific (EAP)



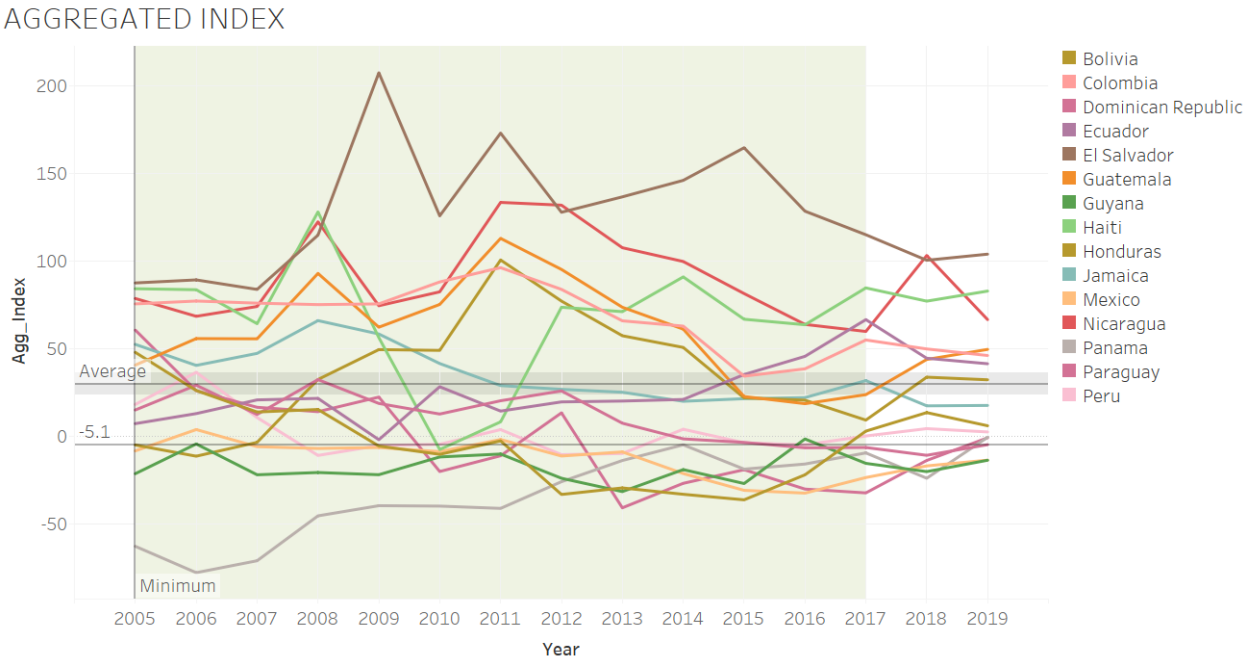
The countries that need attention are East Timor and Cambodia. The development trend has been quite stable for this region.

b) Europe and Eurasia (E&E)



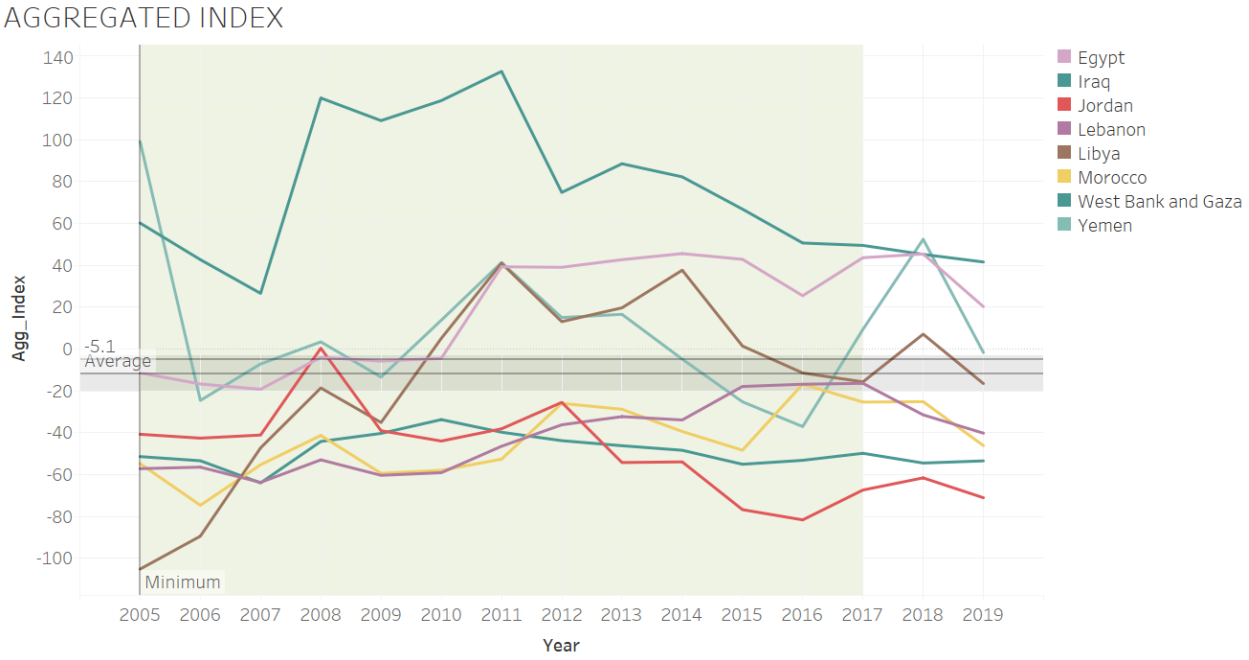
The most worrisome countries are Bulgaria and Albania.

c) Latin America and Caribbean (LAC)



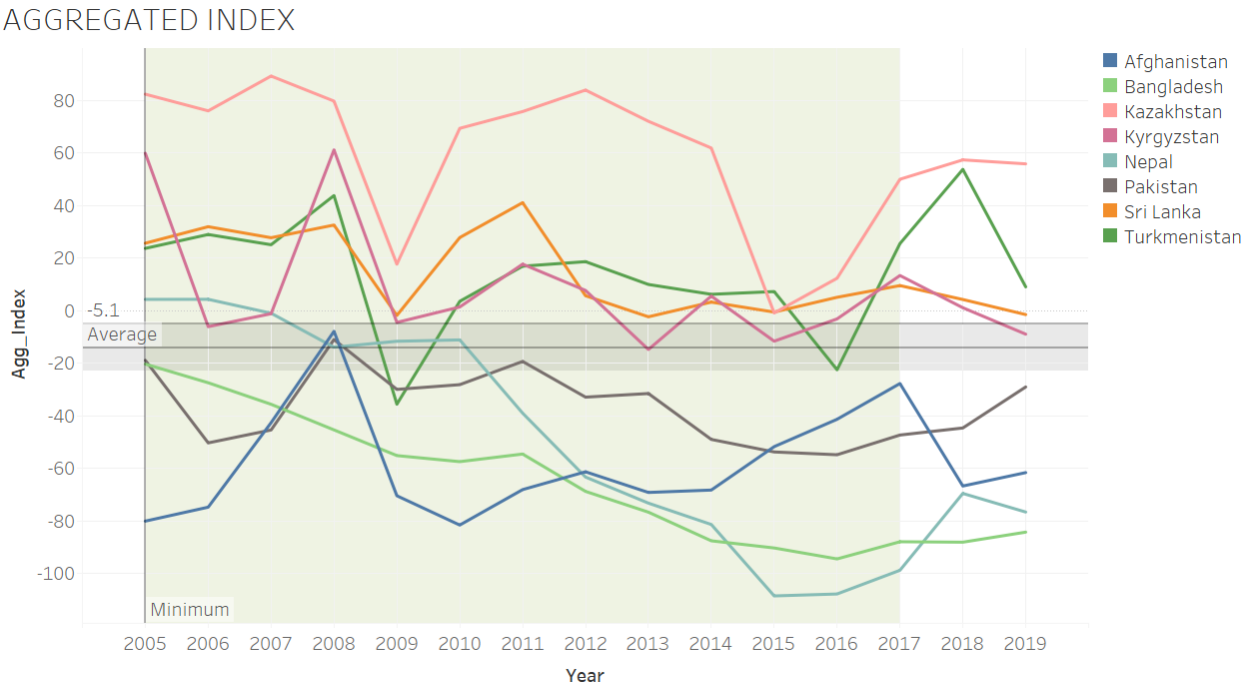
The regional average is above the global average, with Mexico and Guyana the most worrisome ones.

d) Middle East and North Africa (MENA)



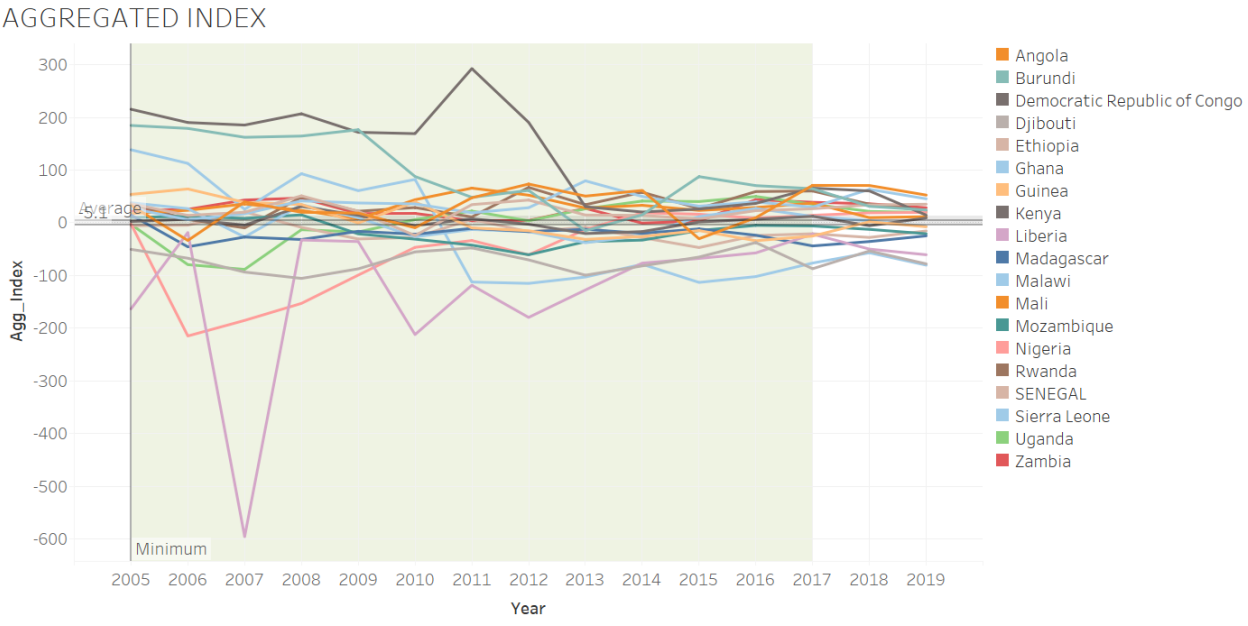
Interestingly, our aggregated index shows Iraq is the leader in terms of overall development, while Jordan is the worst. This is the reverse order when comparing to CPI ranking.

e) South and Central Asia (SAC)



The most worrisome countries are Bangladesh, Nepal, and Afghanistan. Afghanistan has consistently underperformed in both our aggregated index and the CPI.

f) Sub-Saharan Africa (SSA)



Democratic Republic of Congo, Burundi and Sierra Leone have gotten worse. The most worrisome countries are Sierra Leone and Djibouti.

The temporal analysis for each of the 20 indicators that make up the Aggregated index is presented under the Appendix.

Section 7 - Conclusions and Recommendations

We decided to base our conclusions and recommendations on the results of our Lasso Regression model with no outlier countries, since the variables' effects on project success are more reasonable.

Our recommendations will respond to the following core business questions that our team and Deloitte have agreed upon since the beginning of our project.

1. Where should USAID invest more resources to fight corruption?

According to the results of our model, the countries that are most likely to improve from USAID's investment are **Ghana (Cluster 7), Ecuador (Cluster 0), Kosovo (Cluster 0), Bulgaria (Cluster 0), Mexico (Cluster 5), Thailand (Cluster 4), and Guatemala (Cluster 9).** The improvement here is measured by the investment's impact on a country's development. Therefore, **USAID should allocate more resources to these countries in future anti-corruption projects.**

There are 3 out 7 countries in the list that belong to Cluster 0. Ecuador, Kosovo and Bulgaria share common characteristics such as the dominant topic of projects is "Rule of law strengthening for civil society", their most common sector is Local Government and most of the projects falling into the $10M-$25M award amount bucket.

Since countries in the same cluster share many in common, USAID should also look into other countries in Cluster 0, 4, 5, 7, and 9 to see whether there are other potential countries for investment.

2. What makes an anti-corruption project successful? How to help anti-corruption projects become more successful in the future?

To gain more success in anti-corruption projects across countries and regions, USAID should invest in projects with award amounts from $10M to $25M and focusing on the topic of "Corruption reduction in the healthcare sector". Furthermore, USAID should invest more in the following sectors: "Media", "Civil Society" , "Rule of Law, Local Government, Governance", and "Rule of Law, Local Government, and Civil Society".

On the other hand, **what makes an anti-corruption project less likely to achieve success?**

USAID should not invest in projects whose sectors are either "Local Government", "Governance", "Public Finance", "Local Government, Economic Growth", "Rule of Law, Governance", or "Agriculture and Food Security, Governance, Public Finance, Democracy and Governance". Moreover, USAID should not invest >$50M for a project. Regarding intervention type, USAID should avoid investing in "Citizen, Media, Business Engagement, Participation" and projects that are focusing on the theme of "Technical support for public and private investment."

The interesting findings from our research are below:

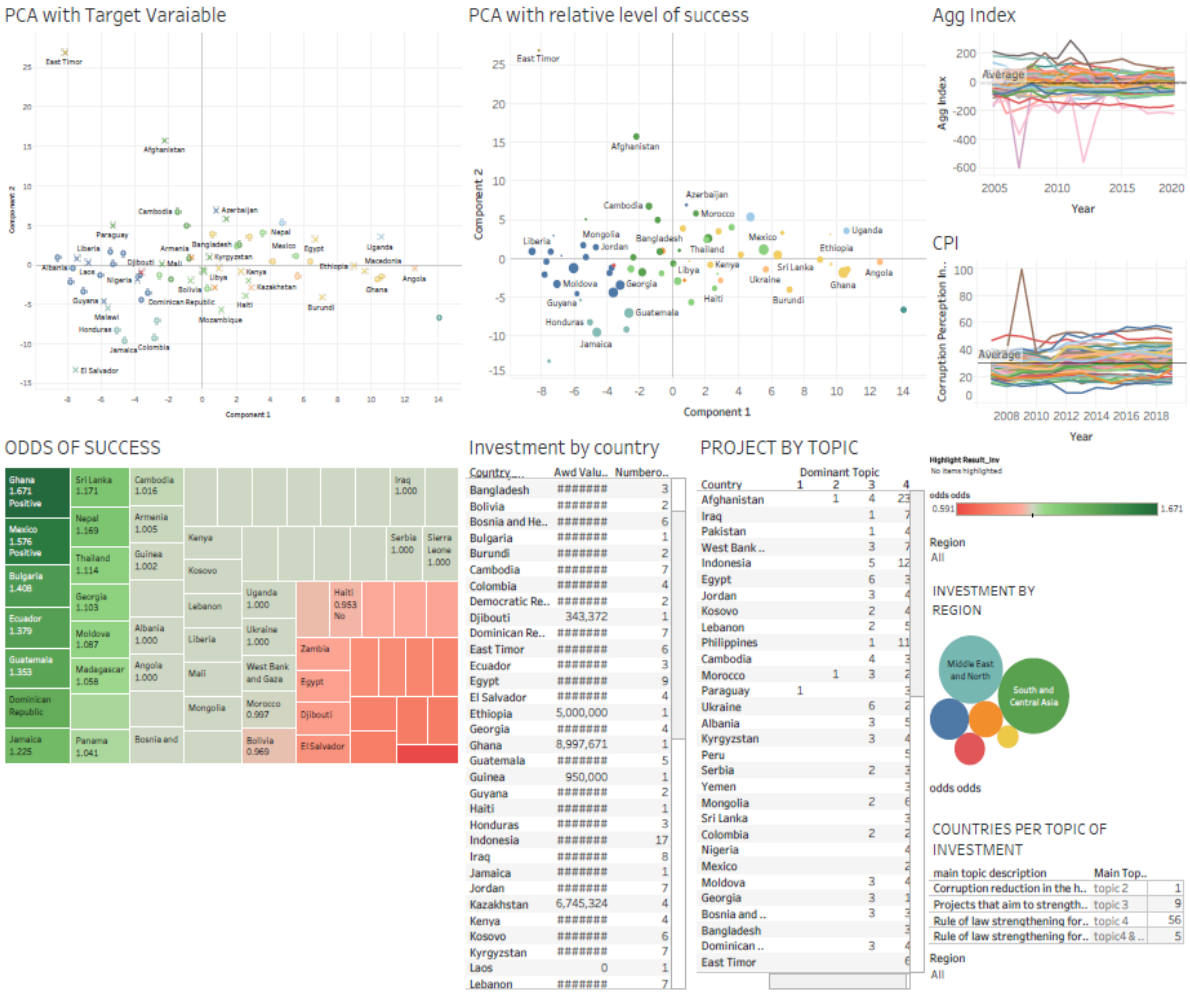
1. Total investment value does not play a significant role in the improvement of a country's corruption level.
2. Countries with higher average project result's and project descriptions' sentiment scores are less likely to be successful, meaning the more positive the overall project results/description sound, the less likely for a country to improve from corruption.

3. What development indicators should USAID monitor to assess the success of anti-corruption projects?

The USAID should monitor the following indicators in order of decreasing importance:

1. Gni Growth (Annual %)
2. Claims On Central Government (Annual Growth As % Of Broad Money)
3. Claims On Private Sector (Annual Growth As % Of Broad Money)
4. Inflation, Gdp Deflator: Linked Series (Annual %)
5. Foreign Direct Investment, Net Inflows (% Of Gdp)
6. Broad Money Growth (Annual %)
7. Foreign Direct Investment, Net Outflows (% Of Gdp)
8. Intentional Homicides (Per 100,000 People)
9. Food Production Index (2004-2006 = 100)
10. Inflation, Consumer Prices (Annual %)
11. Net Barter Terms Of Trade Index (2000 = 100)
12. Exports Of Goods And Services (% Of Gdp)
13. Price Level Ratio Of Ppp Conversion Factor (Gdp) To Market Exchange Rate
14. Total Tax And Contribution Rate (% Of Profit)
15. Bank Liquid Reserves To Bank Assets Ratio (%)
16. Total Reserves (% Of Total External Debt)
17. Gross Savings (% Of Gdp)
18. Merchandise Exports To Low- And Middle-Income Economies Within Region (% Of Total Merchandise Exports)
19. Merchandise Exports To Low- And Middle-Income Economies Outside Region (% Of Total Merchandise Exports)
20. Domestic Private Health Expenditure (% Of Current Health Expenditure)

Section 8- Dashboard



Section 9- Potential Next Steps

For future recommendations, we would need more recent data from USAID, as that would affect our analysis. Moreover, we will do similar analysis for the **regional level**, and also use deep learning to do text mining.

For the current USAID dataset, the latest projects ended in 2017, which have several year's gap with our current year. So our analysis doesn't include the latest project information. If more recent data could be available, the analysis as well as the recommendations concluded would be more effective and applicable. Especially, COVID-19 would have significant impacts to the countries and regions' development indicator performance, which would impact the corruption level to some extent as well. So the further analysis based on the recent years would be helpful to USAID make better decisions when investing in anti-corruption projects.

Furthermore, we mainly focused on country level analysis so far for this project. Though countries are contained in different regions, the similar analysis for the regional level would also result in some interesting findings, which could help USAID make the more general decisions on region level.

For text mining, we can apply deep learning algorithms and techniques such as LSTM,RNN,GRU for topic modelling and sentiment analysis. Additionally, for next steps, we could compare different clusters' characteristics and composition more clearly

References:

Gross National Income (GNI)
https://www.investopedia.com/terms/g/gross-national-income-gni.asp

Glossary:Intentional homicide
https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Intentional_homicide#:~:text=Intentional%20homicide%20means%20killing%20a,and%20other%20criteria%20are%20considered

Metadata Glossary - Food Production Index
https://databank.worldbank.org/metadataglossary/wdi-database-archives-(beta)/series/AG.PRD.FOOD.XD#:~:text=Food%20production%20index%20covers%20food,they%20have%20no%20nutritive%20value.&text=The%20category%20of%20food%20production,edible%20and%20that%20contain%20nutrients.

World Bank's Data Catalog - Claims on Central Government
https://datacatalog.worldbank.org/claims-central-government-etc-gdp-0

United Nations' Indicators definition - FDI
https://www.un.org/esa/sustdev/natlinfo/indicators/methodology_sheets/global_econ_partnership/fdi.pdf

Guyana's Foreign Direct Investment quadrupled in 2017
https://dpi.gov.gy/guyanas-foreign-direct-investment-quadrupled-in-2017/

Inflation,consumer prices index(annual %)
https://fred.stlouisfed.org/series/FPCPITOTLZGUSA

Claims On Private Sector (Annual Growth As % Of Broad Money)
https://www.worldbank.org/en/publication/gfdr/gfdr-2016/background/financial-depth

Inflation, Gdp Deflator: Linked Series (Annual %)
https://datacatalog.worldbank.org/inflation-gdp-deflator-linked-series-annual

Broad Money Growth (Annual %)
https://www.investopedia.com/terms/b/broad-money.asp

Foreign Direct Investment, Net Outflows (% Of Gdp)
https://www.investopedia.com/terms/f/fdi.asp
https://www.pucsp.br/icim/ingles/downloads/papers_2011/part_7/part_7_proc_43.pdf

Appendices including code developed for the project

K-means clustering

component_1	component_2	predicted_cluster	true_label
-7.70733	-0.3647	0	Albania
0.807322	6.95894	0	Azerbaijan
-7.81628	-2.04031	0	Bosnia and Herzegovina
-6.05265	-1.19972	0	Bulgaria
-3.21939	-3.39978	0	Dominican Republic
-3.63575	-4.34966	0	Ecuador
-5.84346	-4.50609	0	Guyana
-4.69674	1.490624	0	Jordan
-5.30297	0.209908	0	Kosovo
-6.79562	0.402788	0	Laos
-7.44522	0.914147	0	Liberia
-7.06258	-3.2504	0	Moldova
-5.46153	1.757838	0	Mongolia
-3.86384	-1.83439	0	Nigeria
-3.73629	-1.15442	0	Philippines
-7.27	2.97403	0	SENEGAL
-8.57801	0.958862	0	Turkmenistan
4.710373	5.401184	1	Nepal
10.55359	3.613662	1	Uganda
-0.65846	1.013873	2	Lebanon
0.694333	-2.7803	2	Montenegro
12.58248	-0.38947	3	Angola

2.891629	-2.78451	3	Kazakhstan
5.648553	-1.36458	3	Ukraine
-2.23518	15.78775	4	Afghanistan
-0.80818	0.924318	4	Armenia
-1.46898	6.790897	4	Cambodia
-1.87627	-1.73602	4	Georgia
-0.9348	5.004801	4	Indonesia
0.377791	1.086128	4	Kyrgyzstan
-2.41732	0.189628	4	Mali
1.392329	5.859418	4	Morocco
-5.30705	5.096306	4	Paraguay
0.015143	-0.60697	4	Sierra Leone
2.139919	2.640038	4	Thailand
-0.69983	-1.87726	5	Bolivia
2.526388	-3.81086	5	Haiti
2.027515	2.445502	5	Iraq
0.280654	-2.88035	5	Madagascar
5.512846	1.182573	5	Mexico
1.106192	-5.61354	5	Mozambique
0.088139	-0.75672	5	Pakistan
-2.56452	-1.31072	5	Panama
2.711152	-1.92867	5	Peru
3.564299	4.074075	5	Serbia
-8.16799	26.92011	6	East Timor
0.599793	3.915277	7	Bangladesh
7.087401	-3.98165	7	Burundi
6.64877	3.347239	7	Egypt
8.947594	-0.09084	7	Ethiopia
10.37038	-1.77223	7	Ghana
10.61278	-1.38632	7	Guinea
2.263593	-0.77545	7	Kenya
0.978798	-0.28501	7	Libya

9.614696	-0.68146	7	Macedonia
2.770701	3.547274	7	Nicaragua
9.574789	3.141201	7	Russian
4.123301	0.524204	7	Rwanda
6.373564	0.501642	7	Sri Lanka
14.02701	-6.58424	8	Democratic Republic of Congo
-2.84055	-9.12619	9	Colombia
-7.53745	-13.2082	9	El Salvador
-2.68533	-6.98926	9	Guatemala
-5.05689	-8.19827	9	Honduras
-4.63984	-9.48419	9	Jamaica
-5.63577	-5.32818	9	Malawi
-3.60175	-0.84327	10	Djibouti

Logistics Regression - Ridge (with all countries)

	Variable	Coefficient	Odds
0	duration_mean	0.147173	1.15855438
1	topic0_mean	0.155051	1.16771751
2	topic1_mean	-0.072075	0.93046111
3	topic2_mean	0.216121	1.24125256
4	topic3_mean	-0.01245	0.98762718
5	topic4_mean	-0.067358	0.93486046
6	dominant_topic_4_count	-0.004393	0.99561664
7	dominant_topic_3_count	0.110527	1.1168665
8	dominant_topic_2_count	0.016366	1.01650066
9	dominant_topic_1_count	-0.223881	0.79941026

10	dominant_topic_0_count	0	1
11	Independent_Agencies_count	0.159727	1.17319055
12	Rule_of_Law_count	0.08731	1.09123491
13	Elections_and_Political_Processes_count	-0.007391	0.99263625
14	Local_Government_count	0.001018	1.00101852
15	Legislature_count	0.131595	1.14064626
16	Economic_Governance_count	-0.01247	0.98760743
17	Disaster_Recovery_count	0.053727	1.05519649
18	Economic_Growth_count	-0.156837	0.85484339
19	Media_count	0.226258	1.25389913
20	Civil_Society_count	0.006446	1.00646682
21	Agriculture_and_Food_Security_count	-0.203166	0.81614275
22	Governance_count	0.113287	1.11995331
23	Healthcare_count	0.169739	1.18499553
24	Public_Finance_count	-0.163138	0.84947395
25	Environment_count	-0.195522	0.82240525
26	Private_Sector_count	-0.04663	0.95444048
27	Democracy_and_Governance_count	-0.050093	0.95114096
28	Gender_Equality_count	0	1
29	Natural_Resource_count	-0.113433	0.89276401
30	Education_count	0.152396	1.16462134
31	desc_senti_score_mean	-0.255488	0.77453843

32	desc_sub_score_mean	0.078005	1.08112806
33	res_senti_score_mean	-0.179721	0.83550328
34	res_sub_score_mean	-0.028482	0.97191979
35	intervention_type_Awareness_count	-0.009051	0.99098984
36	intervention_type_Transparency_count	0.179308	1.19638918
37	intervention_type_Assessment_count	-0.143984	0.8659016
38	intervention_type_Accountability_count	0.113898	1.12063781
39	intervention_type_Participation_count	-0.160047	0.85210374
40	intervention_type_Enforcement_count	0.303623	1.35475822
41	intervention_type_Study_count	0	1
42	intervention_type_Citizen_Media_Business_Engag...	-0.189036	0.82775671
43	awd_value_sum	-0.025006	0.97530406
44	awd_value_mean	-0.028307	0.97208989
45	awd_value_median	-0.035458	0.96516327
46	awd_amount_very_low	-0.085647	0.9179182
47	awd_amount_low	-0.056054	0.94548808
48	awd_amount_median_low	-0.043136	0.95778112
49	awd_amount_median	0.426212	1.53144541
50	awd_amount_median_high	-0.120899	0.88612345
51	awd_amount_high	-0.012964	0.98711967
52	awd_amount_very_high	0.008565	1.00860178
53	corruption_initial	0.412166	1.51008509

54	stability_initial	0.011974	1.01204598
55	numberofprojects	0.032574	1.03311034
56	change_stability	0.143775	1.15462429
57	change_corruption	0.139242	1.14940222
58	dominant_topic_2_count	0.214914	1.23975527
59	dominant_topic_3_count	-0.134897	0.87380591
60	dominant_topic_4_count	-0.059821	0.94193312
61	dominant_topic_4_countdominant_topic_3_count	0.092425	1.09683088
62	dominant_topic_4_countdominant_topic_3_countdo...	0.139877	1.15013232
63	Agriculture_and_Food_Security_countGovernance_...	-0.124031	0.88335246
64	Agriculture_and_Food_Security_countNatural_Res...	-0.113433	0.89276401
65	Civil_Society_count	0.23865	1.26953412
66	Civil_Society_countGovernance_countHealthcare_...	-0.189036	0.82775671
67	Economic_Growth_count	-0.2324	0.79262901
68	Governance_count	-0.139884	0.86945909
69	Healthcare_count	0.185348	1.20363723
70	Independent_Agencies_count	0.152962	1.1652807
71	Independent_Agencies_countRule_of_Law_countLoc...	0.130079	1.13891835
72	Local_Government_count	-0.222591	0.80044216
73	Local_Government_countEconomic_Growth_count	-0.213754	0.80754702
74	Local_Government_countMedia_countGovernance_count	0.176315	1.19281374
75	Public_Finance_count	-0.236635	0.78927932

76	Rule_of_Law_count	0.113399	1.12007876
77	Rule_of_Law_countEconomic_Growth_count	0.196926	1.21765393
78	Rule_of_Law_countGovernance_count	-0.285637	0.75153537
79	Rule_of_Law_countLegislature_countEconomic_Gro...	-0.213818	0.80749534
80	Rule_of_Law_countLocal_Government_count	0.182839	1.20062109
81	Rule_of_Law_countLocal_Government_countCivil_S...	0.251954	1.28653686
82	Rule_of_Law_countLocal_Government_countGoverna ...	0.261632	1.2990484
83	Rule_of_Law_countPublic_Finance_count	0.050311	1.05159809
84	Afghanistan	-0.109057	0.89667931
85	Albania	0.18432	1.20240053
86	Angola	-0.294246	0.74509318
87	Armenia	0.277331	1.31960309
88	Azerbaijan	-0.213818	0.80749534
89	Bangladesh	0.227312	1.25522144
90	Bolivia	-0.264114	0.767886
91	Bosnia and Herzegovina	0.196977	1.21771603
92	Bulgaria	0.428475	1.53491499
93	Burundi	-0.212823	0.80829919
94	Cambodia	0.252563	1.28732059
95	Colombia	0.179914	1.19711441
96	Democratic Republic of Congo	0.089346	1.09345893
97	Djibouti	-0.351837	0.70339477

98	Dominican Republic	0.239791	1.27098349
99	East Timor	-0.348801	0.70553352
100	Ecuador	0.4221	1.52516103
101	Egypt	-0.195522	0.82240525
102	El Salvador	-0.360965	0.69700339
103	Ethiopia	-0.265285	0.76698733
104	Georgia	0.27214	1.31277078
105	Ghana	0.475596	1.60897286
106	Guatemala	0.271635	1.31210799
107	Guinea	0.214914	1.23975527
108	Guyana	-0.266235	0.76625904
109	Haiti	-0.255726	0.77435411
110	Honduras	0.139877	1.15013232
111	Indonesia	0.053727	1.05519649
112	Iraq	0.23361	1.26315177
113	Jamaica	0.246176	1.27912468
114	Jordan	0.336844	1.40052057
115	Kazakhstan	-0.356203	0.70033044
116	Kenya	-0.170079	0.84359817
117	Kosovo	0.278918	1.32169896
118	Kyrgyzstan	-0.391452	0.6760745
119	Lebanon	-0.277739	0.7574945

120	Liberia	-0.221593	0.8012414
121	Macedonia	-0.491341	0.61180541
122	Madagascar	0.233051	1.26244586
123	Malawi	-0.331355	0.71795025
124	Mali	-0.265039	0.76717603
125	Mexico	0.346703	1.41439659
126	Moldova	0.196926	1.21765393
127	Mongolia	0.130079	1.13891835
128	Montenegro	-0.287268	0.75031062
129	Morocco	-0.213754	0.80754702
130	Mozambique	-0.24053	0.78621106
131	Nepal	0.251954	1.28653686
132	Nicaragua	0.244776	1.27733516
133	Nigeria	-0.124031	0.88335246
134	Pakistan	-0.366312	0.69328646
135	Panama	0.266614	1.30553641
136	Paraguay	-0.223881	0.79941026
137	Peru	-0.419352	0.65747272
138	Philippines	0.1715	1.18708414
139	Rwanda	0.185348	1.20363723
140	Serbia	0.262212	1.29980207
141	Sierra Leone	-0.113433	0.89276401

142	Sri Lanka	0.176315	1.19281374
143	Thailand	0.310531	1.36414929
144	Turkmenistan	0.380037	1.46233869
145	Uganda	-0.162028	0.85041739
146	Ukraine	0.152396	1.16462134
147	Vietnam	-0.251914	0.77731158
148	West Bank and Gaza	0.1904	1.20973339
149	Yemen	-0.189036	0.82775671
150	Zambia	-0.285637	0.75153537

Lasso Regression (with all countries)

	Variable	Coefficient	Odds
0	duration_mean	0	1
1	topic0_mean	0	1
2	topic1_mean	0	1
3	topic2_mean	0.647296	1.9103682
4	topic3_mean	0	1
5	topic4_mean	0	1
6	dominant_topic_4_count	0	1
7	dominant_topic_3_count	0	1
8	dominant_topic_2_count	0	1
9	dominant_topic_1_count	-0.181638	0.83390316
10	dominant_topic_0_count	0	1

11	Independent_Agencies_count	0	1
12	Rule_of_Law_count	0	1
13	Elections_and_Political_Processes_count	0	1
14	Local_Government_count	0	1
15	Legislature_count	0	1
16	Economic_Governance_count	0	1
17	Disaster_Recovery_count	0	1
18	Economic_Growth_count	0	1
19	Media_count	0.175152	1.1914273
20	Civil_Society_count	0	1
21	Agriculture_and_Food_Security_count	-0.341318	0.71083283
22	Governance_count	0	1
23	Healthcare_count	0	1
24	Public_Finance_count	0	1
25	Environment_count	-0.277482	0.7576892
26	Private_Sector_count	0	1
27	Democracy_and_Governance_count	0	1
28	Gender_Equality_count	0	1
29	Natural_Resource_count	0	1
30	Education_count	0	1
31	desc_senti_score_mean	-0.701895	0.49564517
32	desc_sub_score_mean	0	1

33	res_senti_score_mean	0	1
34	res_sub_score_mean	0	1
35	intervention_type_Awareness_count	0	1
36	intervention_type_Transparency_count	0	1
37	intervention_type_Assessment_count	0	1
38	intervention_type_Accountability_count	0	1
39	intervention_type_Participation_count	0	1
40	intervention_type_Enforcement_count	0.623627	1.86568262
41	intervention_type_Study_count	0	1
42	intervention_type_Citizen_Media_Business_Engag ...	-0.233089	0.79208307
43	awd_value_sum	0	1
44	awd_value_mean	0	1
45	awd_value_median	0	1
46	awd_amount_very_low	0	1
47	awd_amount_low	0	1
48	awd_amount_median_low	0	1
49	awd_amount_median	1.303142	3.68084373
50	awd_amount_median_high	0	1
51	awd_amount_high	0	1
52	awd_amount_very_high	0	1
53	corruption_initial	0.829138	2.29134275
54	stability_initial	0	1

55	numberofprojects	0	1
56	change_stability	0.196529	1.21717062
57	change_corruption	0	1
58	dominant_topic_2_count	0.028854	1.02927431
59	dominant_topic_3_count	0	1
60	dominant_topic_4_count	0	1
61	dominant_topic_4_countdominant_topic_3_count	0	1
62	dominant_topic_4_countdominant_topic_3_countdo...	0	1
63	Agriculture_and_Food_Security_countGovernance_...	0	1
64	Agriculture_and_Food_Security_countNatural_Res...	0	1
65	Civil_Society_count	0.228148	1.25627124
66	Civil_Society_countGovernance_countHealthcare_...	-0.033904	0.9666643
67	Economic_Growth_count	-0.323552	0.72357433
68	Governance_count	0	1
69	Healthcare_count	0	1
70	Independent_Agencies_count	0	1
71	Independent_Agencies_countRule_of_Law_countLoc...	0	1
72	Local_Government_count	-0.494174	0.61007462
73	Local_Government_countEconomic_Growth_count	-0.167869	0.84546458
74	Local_Government_countMedia_countGovernance_count	0.156955	1.16994297
75	Public_Finance_count	-0.249215	0.77941238
76	Rule_of_Law_count	0	1

77	Rule_of_Law_countEconomic_Growth_count	0.104892	1.11059066
78	Rule_of_Law_countGovernance_count	-0.242268	0.78484581
79	Rule_of_Law_countLegislature_countEconomic_Gro...	-0.039176	0.96158146
80	Rule_of_Law_countLocal_Government_count	0	1
81	Rule_of_Law_countLocal_Government_countCivil_S...	0.185308	1.20358909
82	Rule_of_Law_countLocal_Government_countGoverna...	0.329737	1.39060235
83	Rule_of_Law_countPublic_Finance_count	0	1
84	Afghanistan	0	1
85	Albania	0	1
86	Angola	0	1
87	Armenia	0.004681	1.00469197
88	Azerbaijan	-0.237777	0.78837848
89	Bangladesh	0	1
90	Bolivia	-0.032025	0.96848237
91	Bosnia and Herzegovina	0	1
92	Bulgaria	0.342294	1.40817424
93	Burundi	0	1
94	Cambodia	0.015383	1.01550193
95	Colombia	0	1
96	Democratic Republic of Congo	0	1
97	Djibouti	-0.127187	0.88056899
98	Dominican Republic	0.26336	1.3012951

99	East Timor	-0.272544	0.76143993
100	Ecuador	0.321615	1.37935362
101	Egypt	-0.18917	0.82764579
102	El Salvador	-0.171209	0.84264544
103	Ethiopia	0	1
104	Georgia	0.097731	1.10266613
105	Ghana	0.513189	1.67061029
106	Guatemala	0.301838	1.35234213
107	Guinea	0.102459	1.10789188
108	Guyana	-0.088858	0.91497549
109	Haiti	-0.047983	0.95314999
110	Honduras	0	1
111	Indonesia	0	1
112	Iraq	0	1
113	Jamaica	0.201935	1.22376846
114	Jordan	0	1
115	Kazakhstan	-0.039139	0.96161704
116	Kenya	0	1
117	Kosovo	0	1
118	Kyrgyzstan	-0.207076	0.81295786
119	Lebanon	0	1
120	Liberia	0	1

121	Macedonia	-0.253977	0.77570964
122	Madagascar	0.055745	1.05732803
123	Malawi	-0.24073	0.78605383
124	Mali	0	1
125	Mexico	0.453834	1.57433664
126	Moldova	0.014343	1.01444635
127	Mongolia	0	1
128	Montenegro	-0.196775	0.82137542
129	Morocco	-0.064798	0.93725677
130	Mozambique	0	1
131	Nepal	0.077016	1.08005936
132	Nicaragua	0	1
133	Nigeria	0	1
134	Pakistan	-0.278152	0.75718172
135	Panama	0.040032	1.04084408
136	Paraguay	-0.100089	0.90475689
137	Peru	-0.525161	0.59146013
138	Philippines	0	1
139	Rwanda	0	1
140	Serbia	0	1
141	Sierra Leone	0	1
142	Sri Lanka	0.019561	1.01975357

143	Thailand	0.107474	1.11346191
144	Turkmenistan	0.041618	1.04249617
145	Uganda	0	1
146	Ukraine	0	1
147	Vietnam	-0.080378	0.92276747
148	West Bank and Gaza	0	1
149	Yemen	-0.086162	0.91744559
150	Zambia	-0.154068	0.85721373

Lasso Regression with No country outliers:

Country table:

Variable	Coefficient	Odds
Ghana	0.382394	1.465789
Ecuador	0.35256	1.422705
Kosovo	0.3056	1.357439
Bulgaria	0.29286	1.340255
Mexico	0.091384	1.09569
Thailand	0.029502	1.029941
Guatemala	0.017207	1.017356
Albania	0	1
Armenia	0	1
Bangladesh	0	1
Bosnia and Herzegovina	0	1
Burundi	0	1
Cambodia	0	1
Colombia	0	1
Dominican Republic	0	1
Egypt	0	1

Georgia	0	1
Guinea	0	1
Honduras	0	1
Indonesia	0	1
Jamaica	0	1
Jordan	0	1
Kenya	0	1
Lebanon	0	1
Madagascar	0	1
Malawi	0	1
Moldova	0	1
Montenegro	0	1
Nepal	0	1
Nicaragua	0	1
Pakistan	0	1
Panama	0	1
Philippines	0	1
Rwanda	0	1
Serbia	0	1
Sri Lanka	0	1
Uganda	0	1
Ukraine	0	1
West Bank and Gaza	0	1
High Corruption	0	1
Same	0	1
Ethiopia	-0.00702	0.993009
Haiti	-0.01698	0.983163
Bolivia	-0.02914	0.971278
Yemen	-0.05496	0.946524
Mali	-0.07786	0.925096
Paraguay	-0.10239	0.902673

Mozambique	-0.10512	0.900215
Morocco	-0.22358	0.799651
Nigeria	-0.2389	0.787492
Peru	-0.29911	0.741479
Kazakhstan	-0.31657	0.728641
Guyana	-0.3743	0.687768
El Salvador	-0.39351	0.674681
Kyrgyzstan	-0.39898	0.671006
Macedonia	-0.44122	0.643249
Vietnam	-0.52547	0.591277
Zambia	-0.57553	0.562404

Other variables table:

	Variable	Coefficient	Odds
1	awd_amount_median	0.553227	1.738855
2	topic2_mean	0.162257	1.176162
3	Media_count	0.103505	1.109051
4	Civil_Society_count	0.07059	1.073141
5	Rule_of_Law_countLocal_Government_countGovernance_count	0.064787	1.066932
6	Rule_of_Law_countLocal_Government_countCivil_Society_count	0.019442	1.019632
7	duration_mean	0	1
8	topic0_mean	0	1
9	topic1_mean	0	1
10	topic3_mean	0	1
11	topic4_mean	0	1
12	dominant_topic_4_count	0	1
13	dominant_topic_3_count	0	1
14	dominant_topic_2_count	0	1

15	dominant_topic_0_count	0	1
16	Independent_Agencies_count	0	1
17	Rule_of_Law_count	0	1
18	Elections_and_Political_Processes_count	0	1
19	Local_Government_count	0	1
20	Legislature_count	0	1
21	Economic_Governance_count	0	1
22	Disaster_Recovery_count	0	1
23	Economic_Growth_count	0	1
24	Civil_Society_count	0	1
25	Governance_count	0	1
26	Healthcare_count	0	1
27	Environment_count	0	1
28	Private_Sector_count	0	1
29	Democracy_and_Governance_count	0	1
30	Gender_Equality_count	0	1
31	Natural_Resource_count	0	1
32	Education_count	0	1
33	desc_sub_score_mean	0	1
34	res_sub_score_mean	0	1
35	intervention_type_Awareness_count	0	1
36	intervention_type_Transparency_count	0	1
37	intervention_type_Assessment_count	0	1
38	intervention_type_Accountability_count	0	1
39	intervention_type_Participation_count	0	1
40	intervention_type_Enforce	0	1

	ment_count		
41	intervention_type_Study_count	0	1
42	awd_value_sum	0	1
43	awd_value_mean	0	1
44	awd_value_median	0	1
45	awd_amount_very_low	0	1
46	awd_amount_low	0	1
47	awd_amount_median_low	0	1
48	awd_amount_high	0	1
49	stability_initial	0	1
50	numberofprojects	0	1
51	change_stability	0	1
52	dominant_topic_2_count	0	1
53	dominant_topic_3_count	0	1
54	dominant_topic_4_count	0	1
55	dominant_topic_4_countdominant_topic_3_count	0	1
56	dominant_topic_4_countdominant_topic_3_countdominant_topic_2_count	0	1
57	Economic_Growth_count	0	1
58	Healthcare_count	0	1
59	Independent_Agencies_count	0	1
60	Local_Government_count Media_countGovernance_count	0	1
61	Rule_of_Law_count	0	1
62	Rule_of_Law_countEconomic_Growth_count	0	1
63	Rule_of_Law_countLocal_Government_count	0	1
64	Agriculture_and_Food_Security_count	-0.04864	0.952525

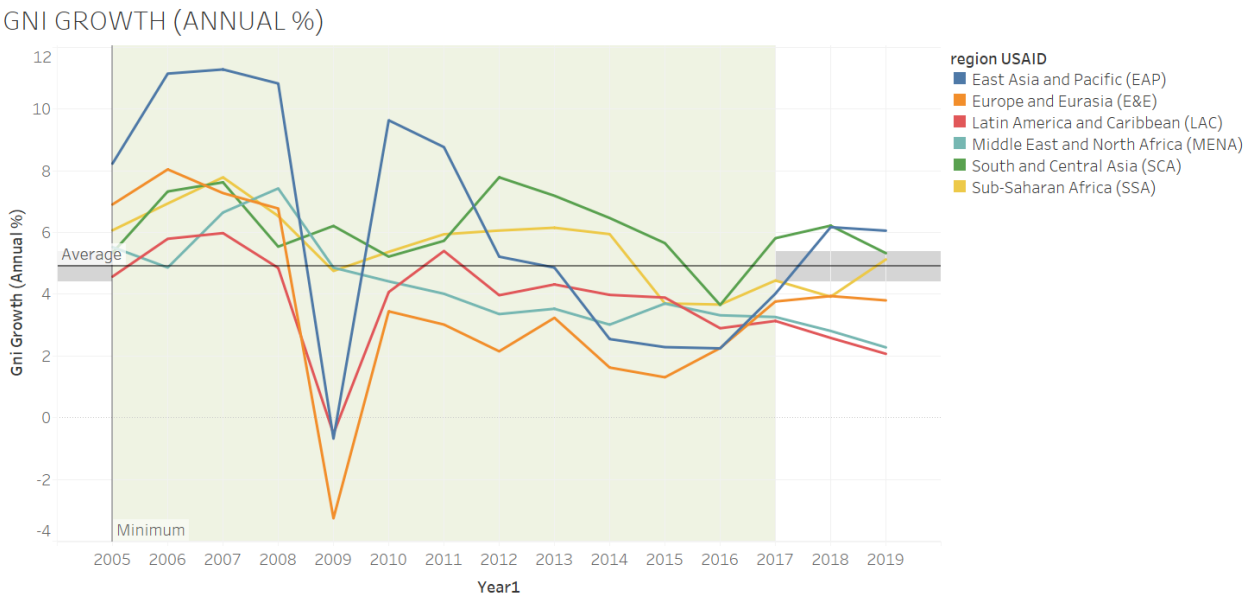
65	Public_Finance_count	-0.05178	0.94954
66	Civil_Society_countGovernance_countHealthcare_count	-0.07412	0.928564
67	Agriculture_and_Food_Security_countGovernance_countPublic_Finance_countDemocracy_and_Governance_count	-0.09572	0.908714
68	Rule_of_Law_countGovernance_count	-0.15224	0.858779
69	desc_senti_score_mean	-0.1767	0.838028
70	Local_Government_countEconomic_Growth_count	-0.18391	0.832009
71	dominant_topic_1_count	-0.22816	0.795995
72	awd_amount_median_high	-0.27421	0.760176
73	intervention_type_Citizen_Media_Business_Engagement_Participation_count	-0.31935	0.72662
74	res_senti_score_mean	-0.49664	0.60857
75	Public_Finance_count	-0.73003	0.481893
76	awd_amount_very_high	-0.83509	0.433836
77	Governance_count	-0.83614	0.433381
78	Local_Government_count	-1.18255	0.306495

Temporal Analysis of Top 20 World Bank Development Indicators:

Now, we would present our analysis with regards to the top 20 indicators that we deem to be the most relevant to a country's corruption status. We will present them from the most important one to the least important one.

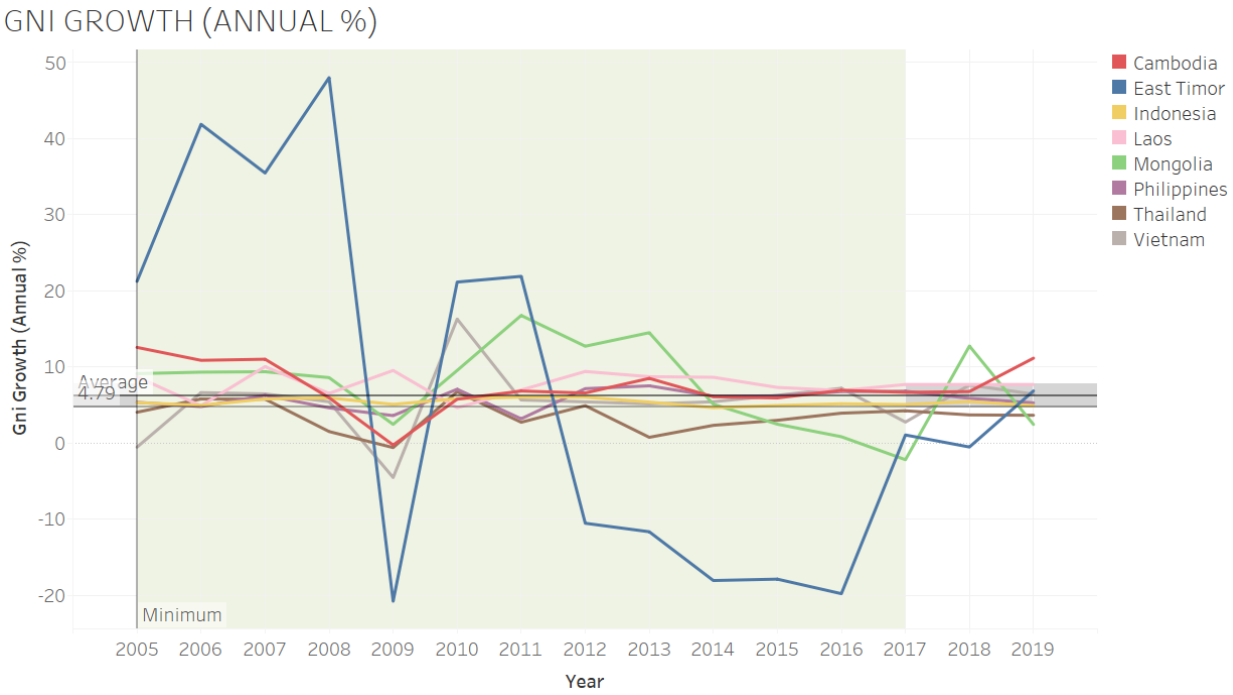
1. **Gni Growth (Annual %)**

"GNI (Gross National Income) is the total amount of money earned by a nation's people and businesses. It is used to measure and track a nation's wealth from year to year." (Investopedia, 2020).The higher annual GNI Growth, the better for a country. However, we should note here that **GNI growth has a low negative correlation with Corruption Perception Index (CPI) at -0.1.** Global average (with 95% CI) is 4.79% over the period of 2005 to 2019.



Currently, the fastest growth region is EAP, followed by SCA, then SSA. The slowest growth belongs to LAC. All regions have seen a slow down in wealth growth rate since 2005. Meaning the economies have grown to the point where there is less room for significant growth.

a) East Asia and Pacific (EAP)

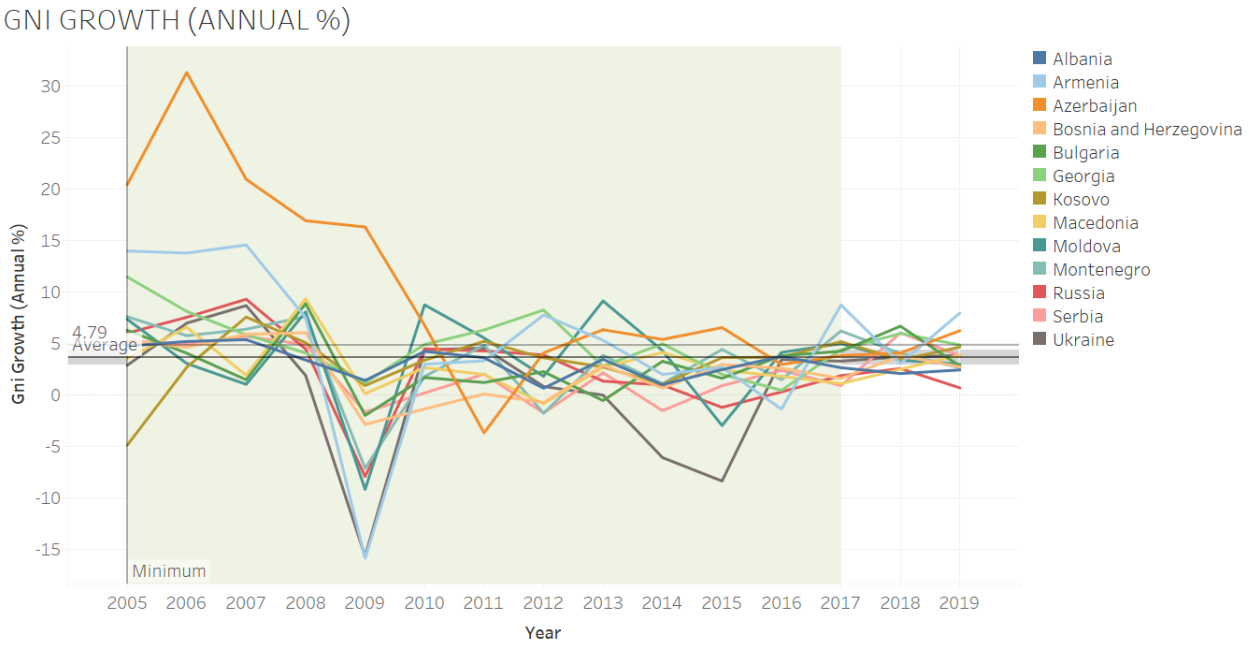


The average of the region (6.17%) is higher than the global average. During the year 2009, all countries suffered a slowdown in GNI growth or even contraction. The reason is likely to be the spillover of the Great Recession in the US. Laos and Indonesia are the only ones that appear to be more or less immune from the Great Recession event. In fact, during the period of 2005 to 2019, Indonesia's annual growth remained around 5%. **This may indicate that the USAID projects did not have an effect on Indonesia's GNI growth**. East Timor, Vietnam, and Mongolia saw a big leap in GNI growth from

2009. **The cases of GNI growth after 2009 in Vietnam and Mongolia could be the results of the USAID projects** as there were projects whose sector of focus was Economics Growth. After 2017, GNI growth in Cambodia, Mongolia, and East Timor improved.

The outliers:
East Timor observed the most significant fluctuations. One year it reached almost 50% GNI annual growth (2008) but the number quickly dropped to -20% the next year (2009). The bad harvest in 2007, which led to many deaths, together with the Great Recession effects, resulted in huge consequences for this country in 2009. After the impressive growth period of 2010-2011, East Timor fell into a long and deep wealth contraction (negative GNI growth) for 5 years, which could be the result of the UN ending its peacekeeping mission there. **East Timor coming back to positive in 2017 could indicate the effectiveness of USAID projects as well, as the economy has been relying on international aid.**

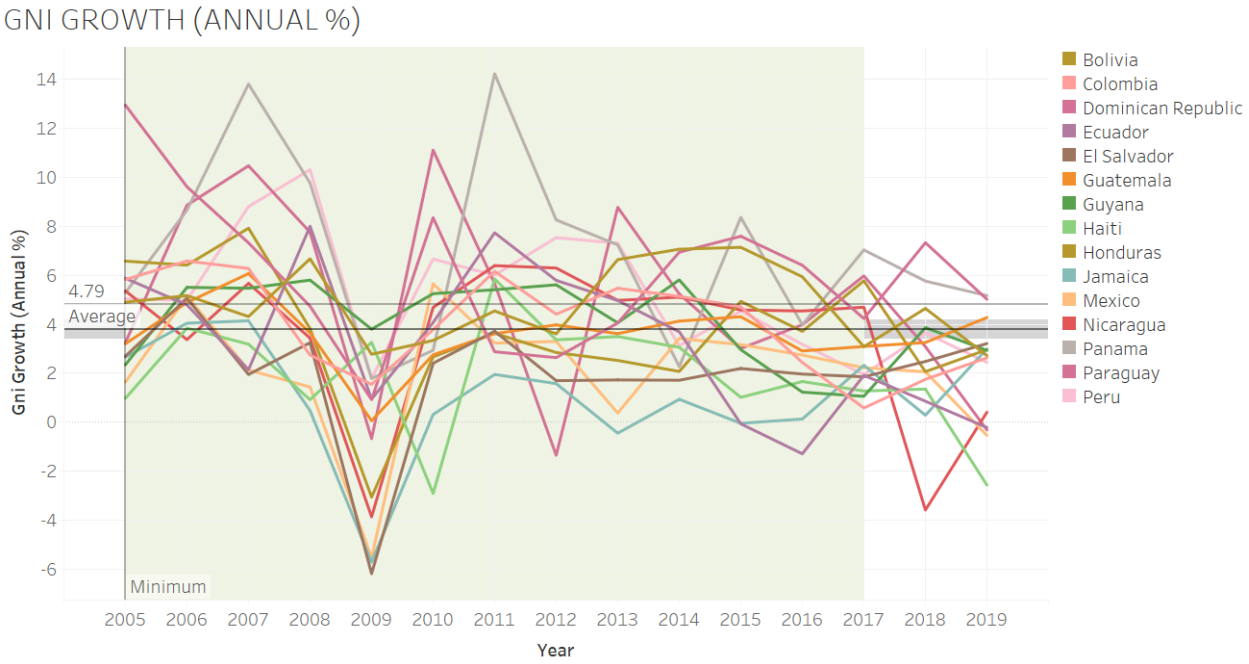
b) Europe and Eurasia (E&E)



The regional average of 3.62% is below the global average. **All countries showed less fluctuation in GNI growth as time went by, which may indicate that USAID projects helped stabilize those economies.** A prime example of that is Ukraine. Moreover, all countries suffered from a drop in GNI growth in 2009 with the worst situation in Ukraine and Armenia - GNI growth dropped to negative 15% and to negative 16%.

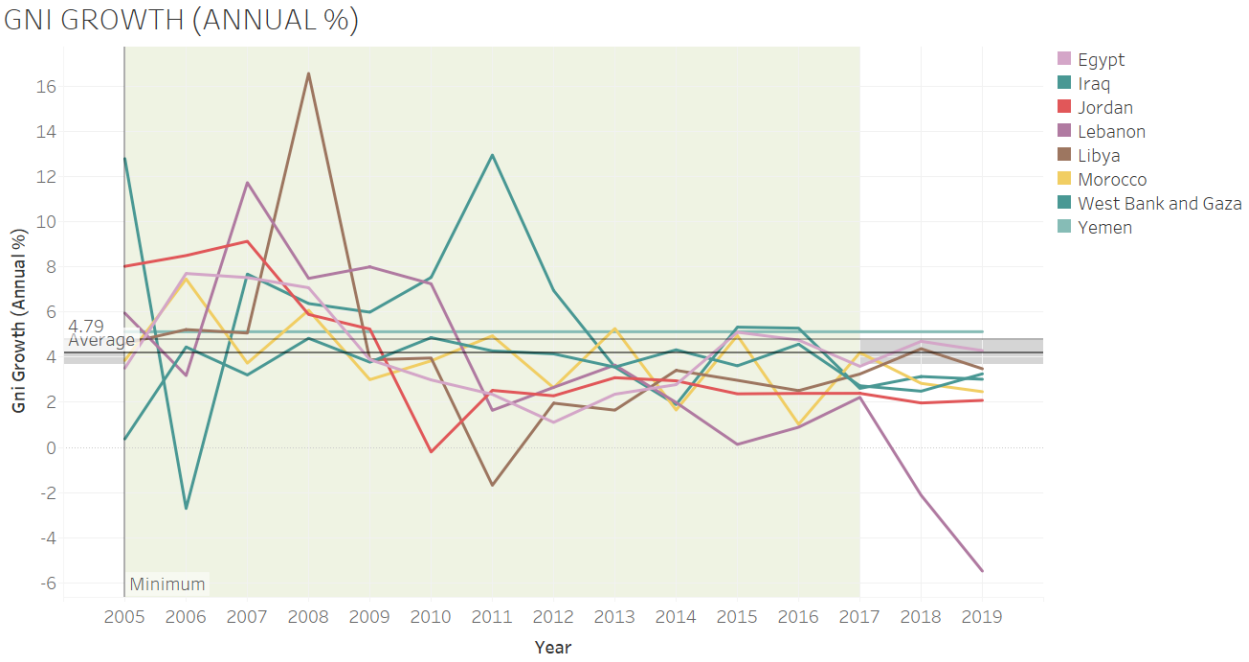
The outliers:
Azerbaijan experienced a tremendous growth rate from 2005 to 2009, with the peak in 2006 reaching 30%+. Even after the drop in 2009, the country still maintained a GNI growth rate of 16%. However, in 2011, the growth rate dropped sharply to negative territory. The aftermath of the Great Recession lasted longer in Azerbaijan than other countries in the region. However, after 2011, the country bounced back and has been growing around 5% each year.

c) Latin America and Caribbean (LAC)



The region average is 3.8%, below the global average. The GNI growth rate has slowed for almost all countries in the region, as the majority of countries starting out in 2005 with growth rates higher than 3.8%, but most of them in 2019 had GNI growth rates lower than 3.8%. All were affected by the Great Recession spillover in 2009. Countries in this region have not experienced stable GNI growth rates.

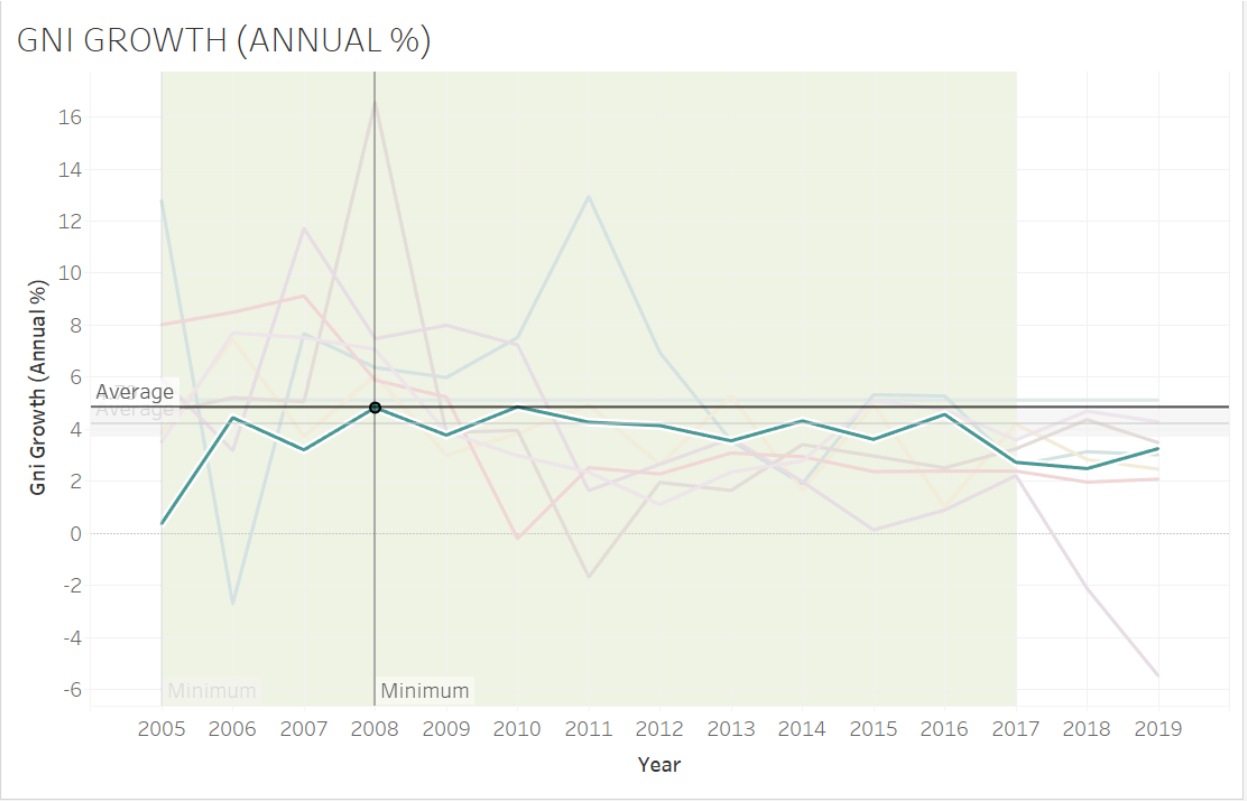
d) Middle East and North Africa (MENA)



The regional average GNI growth rate is 4.2%, lower than the global average. The impact of the Great Recession is not reflected through the overall regional trend. The common theme for most countries in

the MENA region is that since 2014, GNI growth rate has become more stable, which can be resulted from USAID projects.

The case of Iraq:

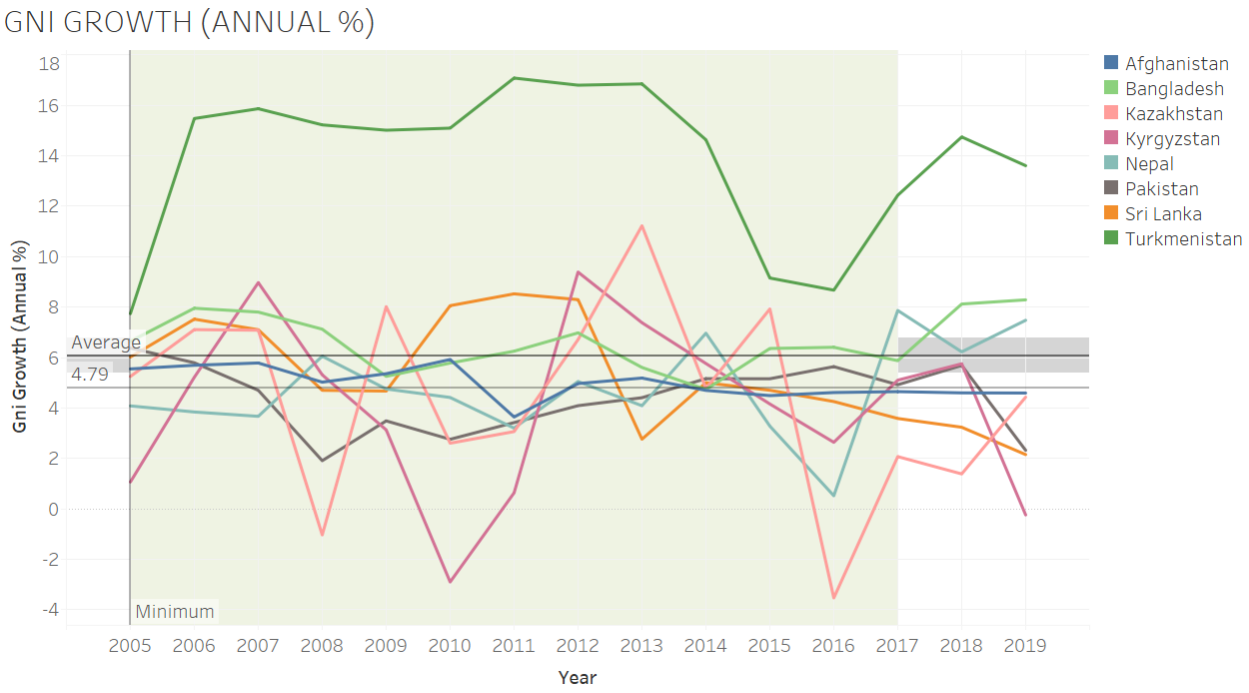


Iraq received the 2nd highest amount of investment from USAID, only behind Afghanistan. **Overall, the GNI growth rate trend for the country has been quite stable**, increasing from 0.4% in 2005 to 4.5% in 2005, then staying around 3-4% up to now, although we did spot lower growth rates after 2017, the year when the last USAID projects in our dataset ended. **This can be a positive outcome of USAID's effort.**

The outliers:

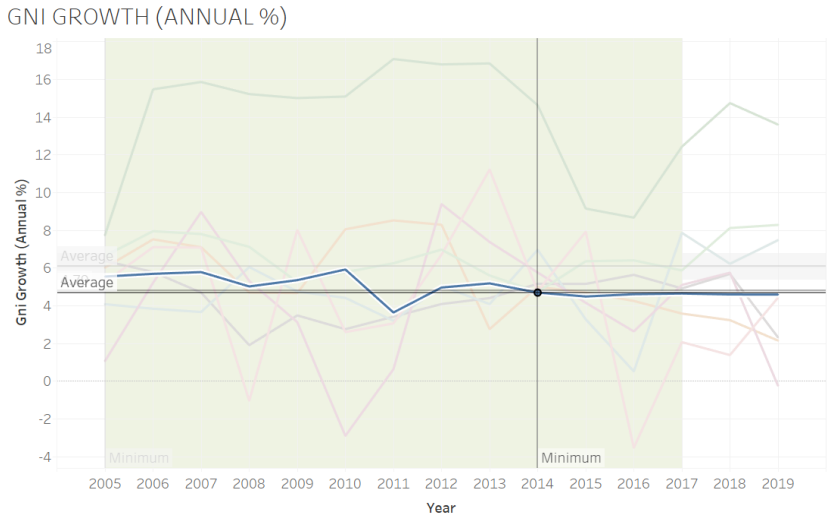
Lebanon has experienced a downward trend in GNI annual growth rate since 2007. While the other countries in the region have stabilized since 2014, Lebanon has struggled from weath contraction since 2018, after 10 consecutive years of slowdown in growth. It is the only country in the region whose GNI annual growth rate is in negative territory in 2018 and 2019.

e) South and Central Asia (SAC)



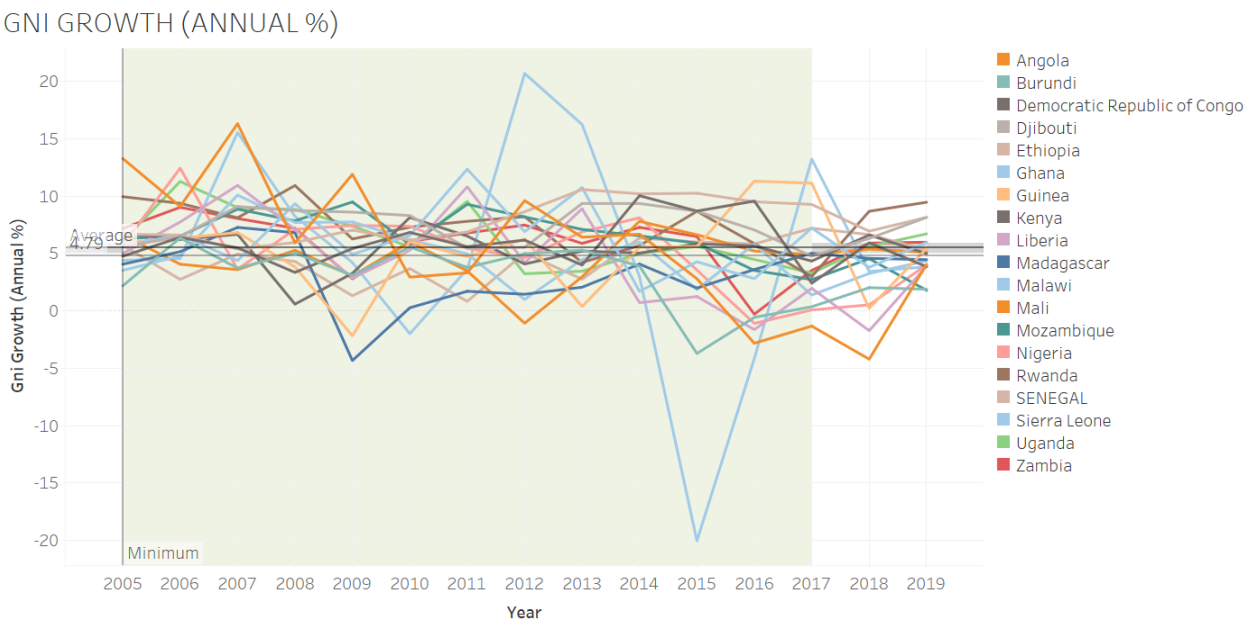
The regional average is 6.07%, above the global average, and the impressive growth rates of Turkmenistan is the major contributor, followed by those of Bangladesh at a distance. Turkmenistan, Bangladesh, Nepal and Kazakhstan have been trending upward since 2016 while Afghanistan stabilizes and Pakistan, Sri Lanka, and Kyrgyzstan have been trending downward. In 2016, the region suffered slowdown and even contraction due to an increase in violence and regional conflict.

The case of Afghanistan:



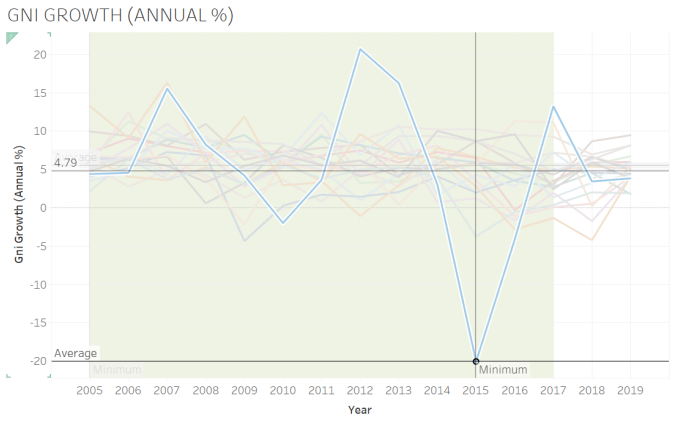
As the country who received the most investment from USAID, Afghanistan observed **stable GNI growth rate from 2005 to 2019**, while other countries in the region have highly fluctuating rates, indicating **potential proof of USAID projects' success.**

f) Sub-Saharan Africa (SSA)



The regional average is 5.49, higher than the global average.

The outliers: Sierra Leone

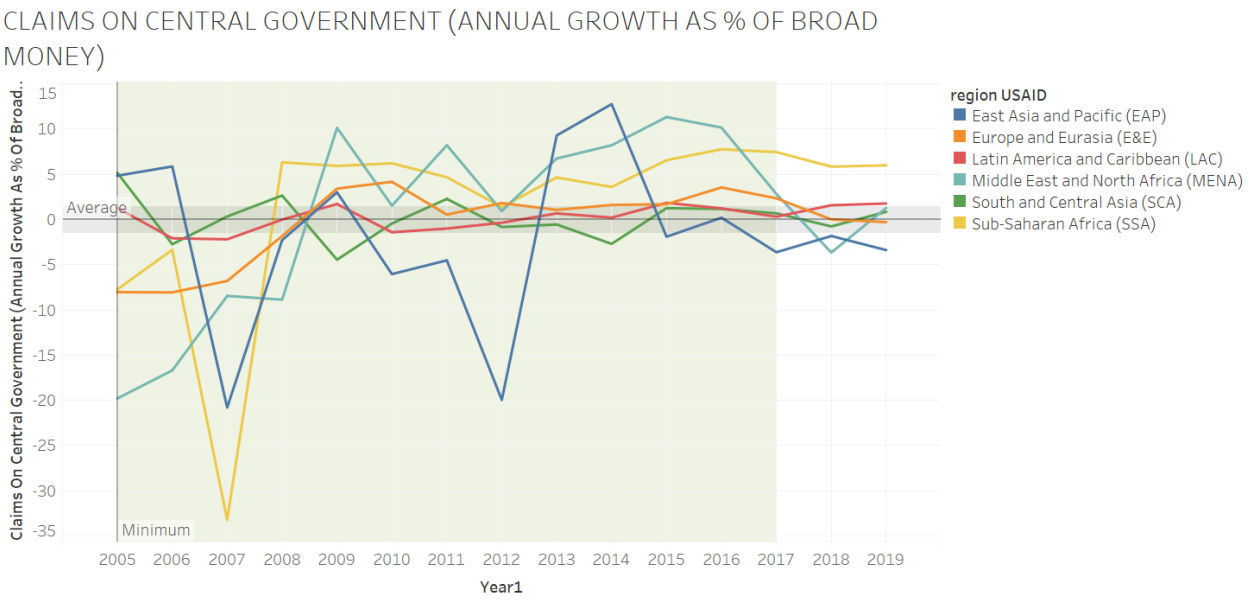


The country has experienced significant fluctuations: the highest growth rate (~20%) and the lowest growth rate (-20%) in a year for the region. It had a great wealth contraction in 2015 then came back quickly in 2016. The country's growth rate is now below the regional average.

2. Claims On Central Government (Annual Growth As % Of Broad Money)

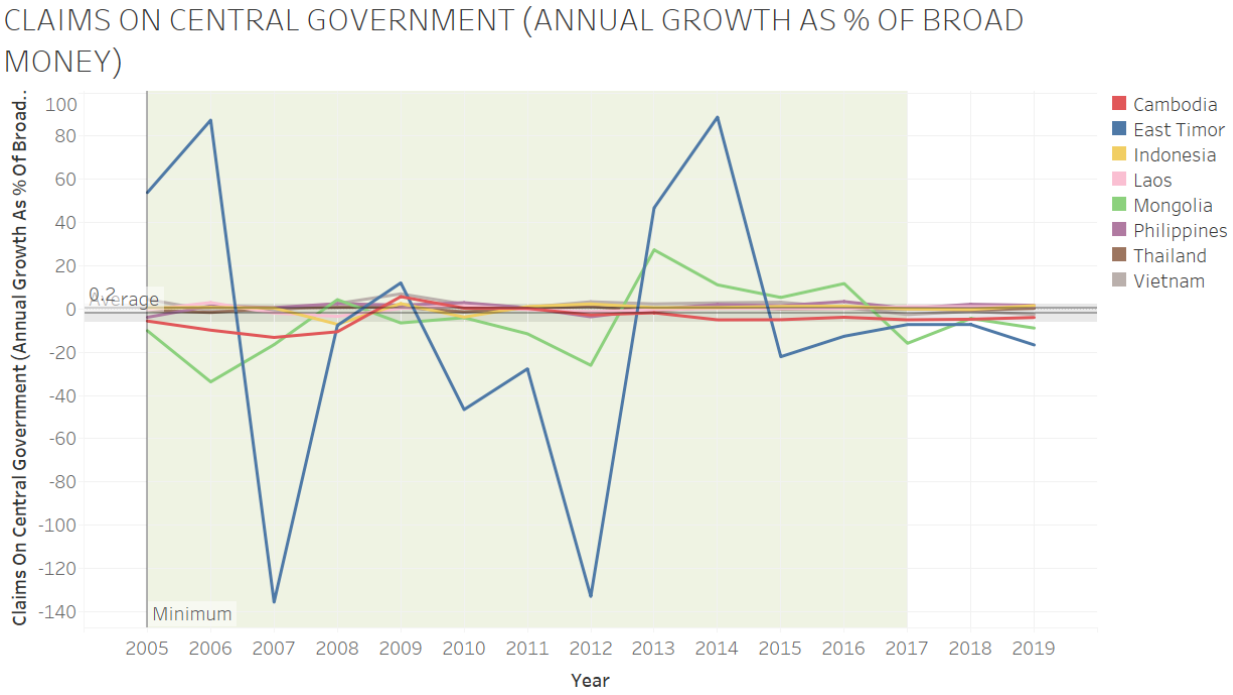
"Claims on central government (IFS line 52AN or 32AN) include loans to central government institutions net of deposits." (World Bank, 2020). Often, **the lower the index, the better** because it means that the government does not have enough money to spend and have to borrow from other entities. Too much debt can be very risky to the future of the economy as we have seen cases of bankruptcy.

The global average is 0.2%.



SSA has the biggest portion of claims on central government as % of Broad money and is the only region with a stably high level of this indicator since 2014.

a) East Asia and Pacific (EAP)

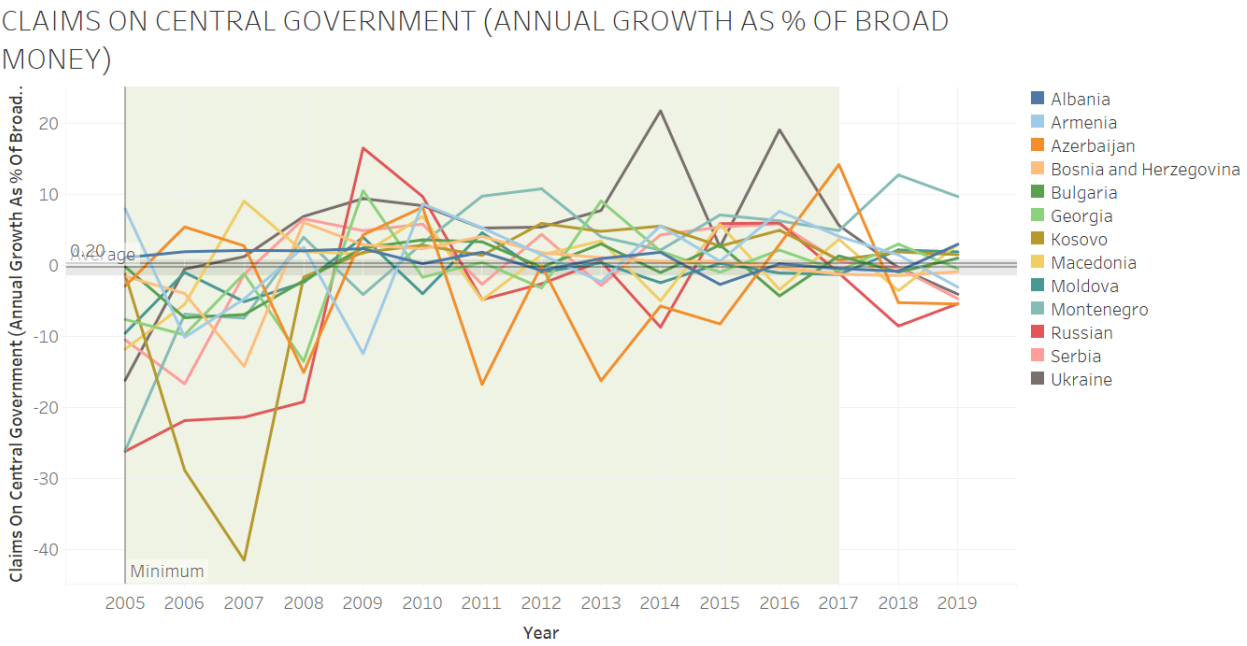


The regional average is in line with the global average. Most countries' claims on central government did not change as % of broad money, except Mongolia and East Timor. However, both countries have stabilized since 2015, especially East Timor.

The outliers:

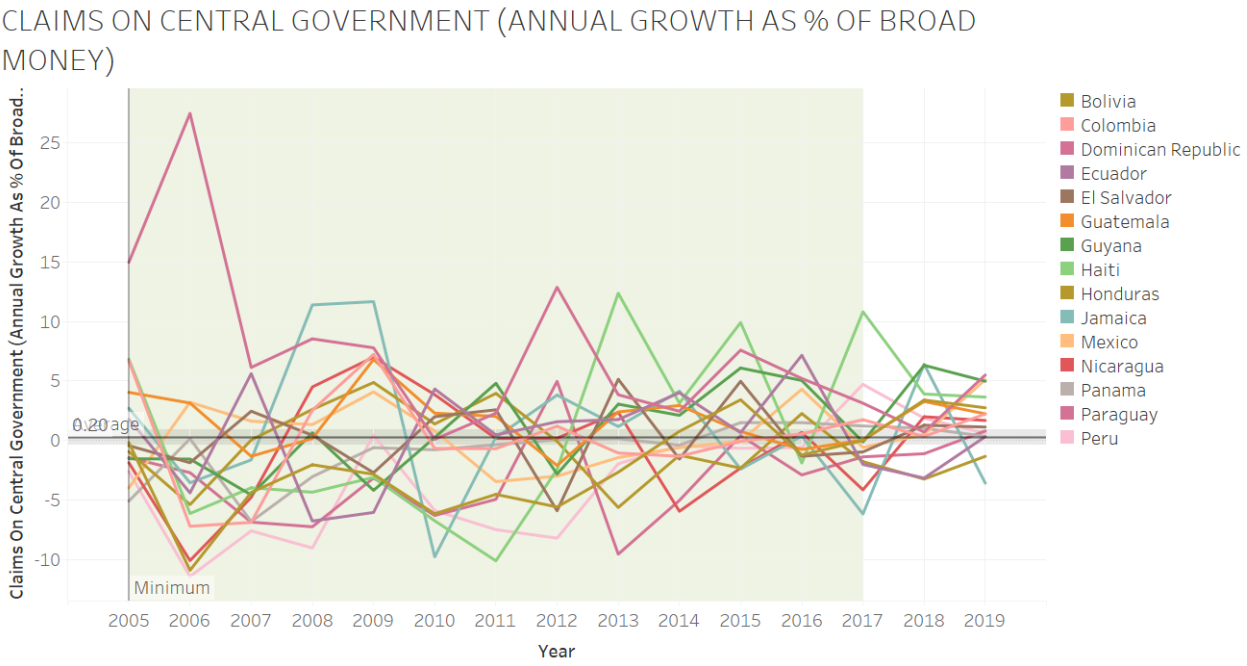
After high increases in the index in 2005 and 2006, **East Timor** saw extreme reduction in the portion of loans to the central government. In 2007 and 2012. The reduction in 2007 lasted to 2013.

b) Europe and Eurasia (E&E)



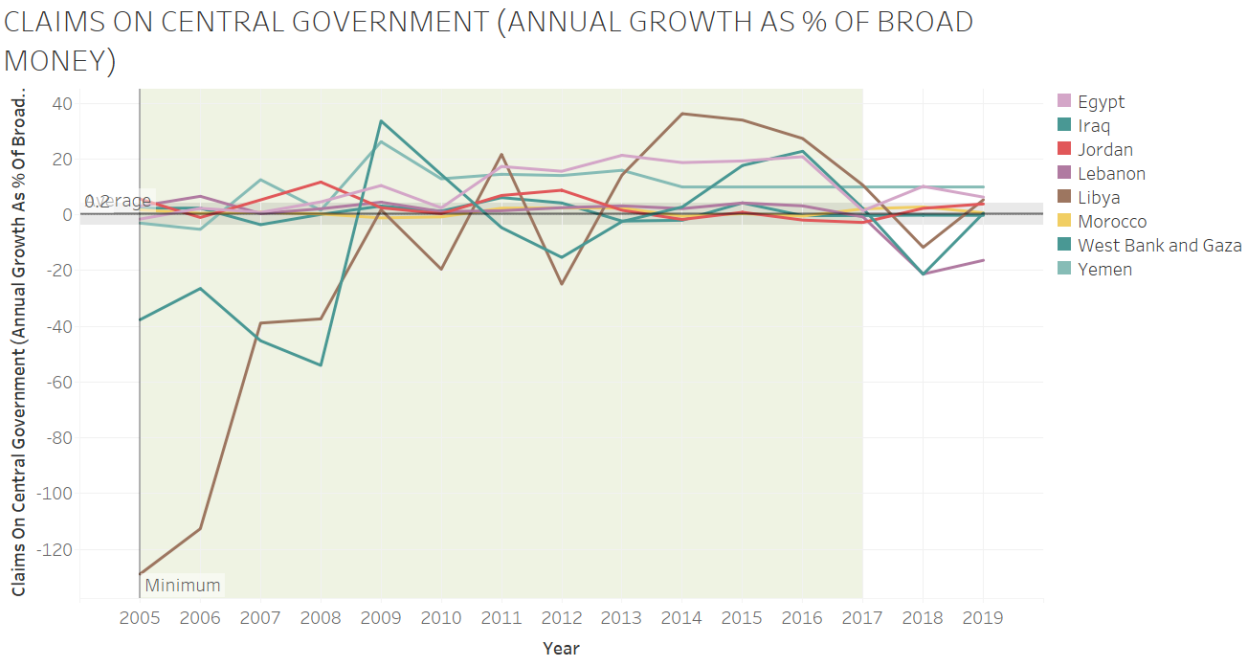
Although the regional average is also inline with the global average, E&E has experienced much more fluctuations in the growth of the portion of the loans to central governments than EAP.

c) Latin America and Caribbean (LAC)



LAC experiences a similar case to E&E: a lot of fluctuations.

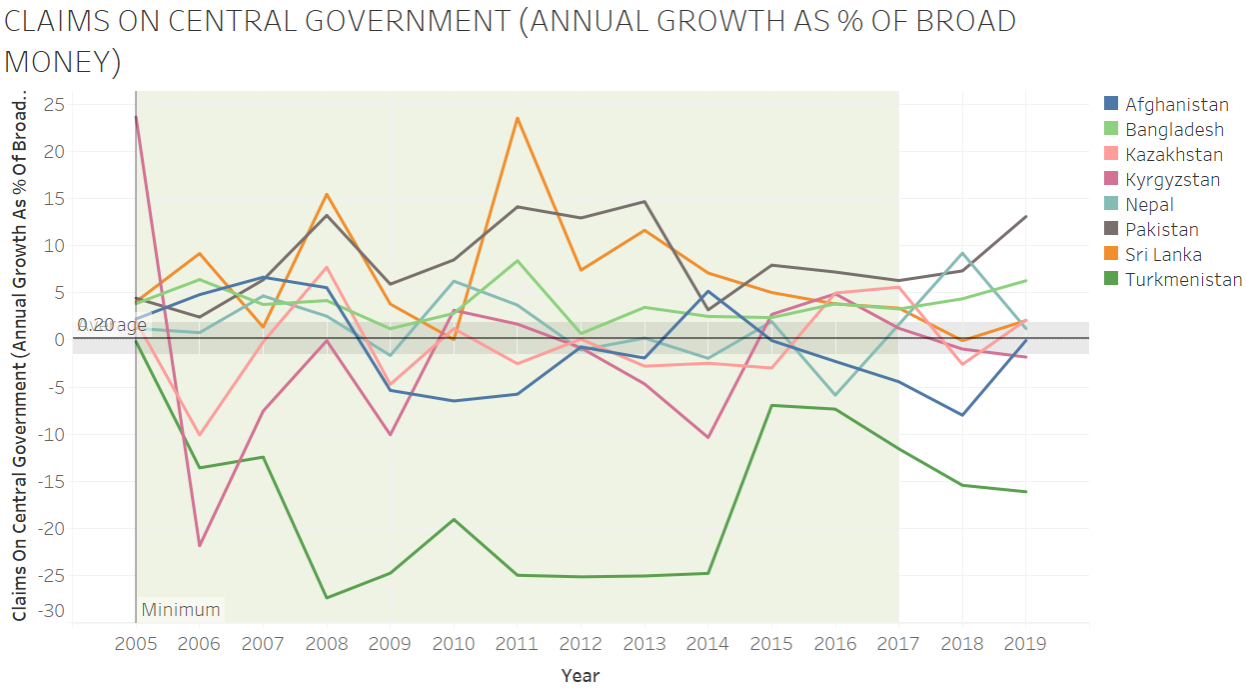
d) Middle East and North Africa (MENA)



The outliers:

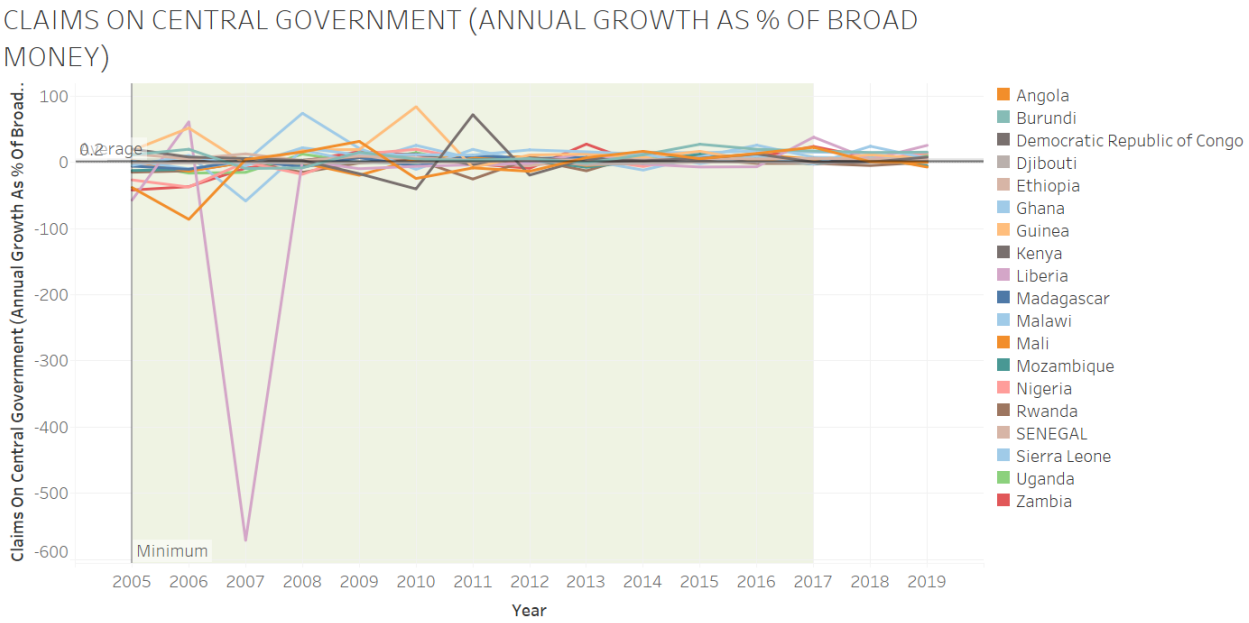
Libya and Iraq have seen huge fluctuations throughout the years. Libya reduced the loans to the central government by 120% in 2005, then continued to reduce but at a lower rate, then finally came to the positive territory in 2014. The period of large loaning to the central government happened between 2014 and 2017. The fluctuations in Iraq have become smaller in magnitude.

e) South and Central Asia (SAC)



The outlier is Turkmenistan. It has been reducing the claims on central government as % of Broad money ever since 2005, but at a lower rate since 2015. The country seems to have been doing well in terms of government assets.

f) Sub-Saharan Africa (SSA)



In this region, we see the most extreme case globally in Liberia. In 2007, the country cut the loans to the government immensely. Other than that outlier data point in 2007, the region overall has become more stable since 2012, a good sign of stable economies.

3. Claims On Private Sector (Annual Growth As % Of Broad Money)

definition: include gross credit from the financial system to individuals, enterprises, nonfinancial public entities not included under net domestic credit, and financial institutions not included elsewhere.

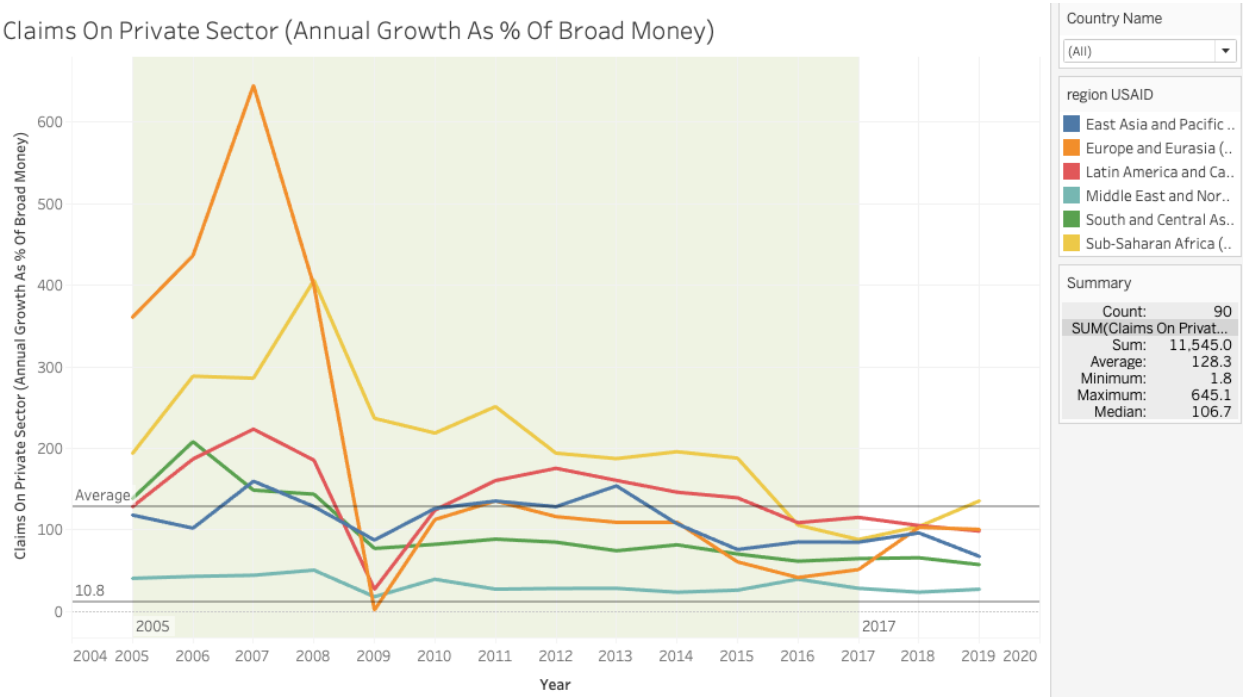
Influence:

Recent studies find that unregulated **financial** sector **growth can** (and often **does**) have an adverse **effect** on the economies studied, wherein the increased strength of the **financial** system corresponds to a lowering of gross domestic product (GDP) **growth**

The higher the Claims On Private Sector is, **the lower** of economy

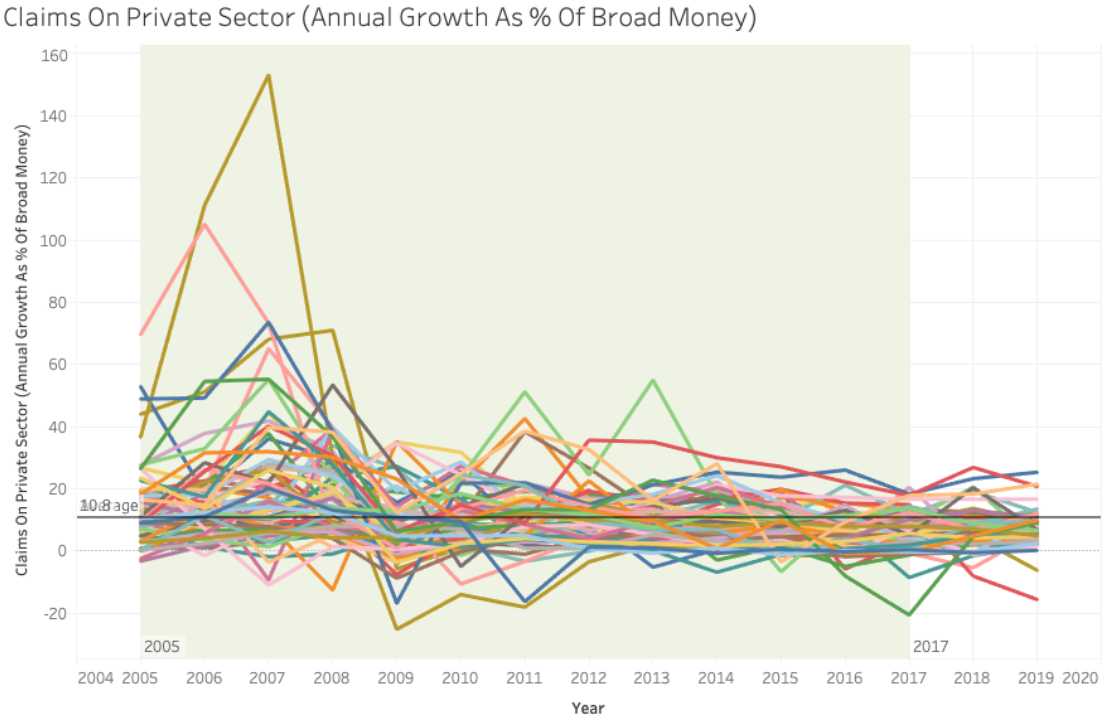
Correlation between inflation and CPI: -0.03

Regionally:



By looking at the regional level, the average is very high, which is 128.3%. he regional average is 10 times higher than the country-level. E&E and SSA have had higher numbers over the years. E&E peak over 600% in 2007, which can also explain the cause of the great recession

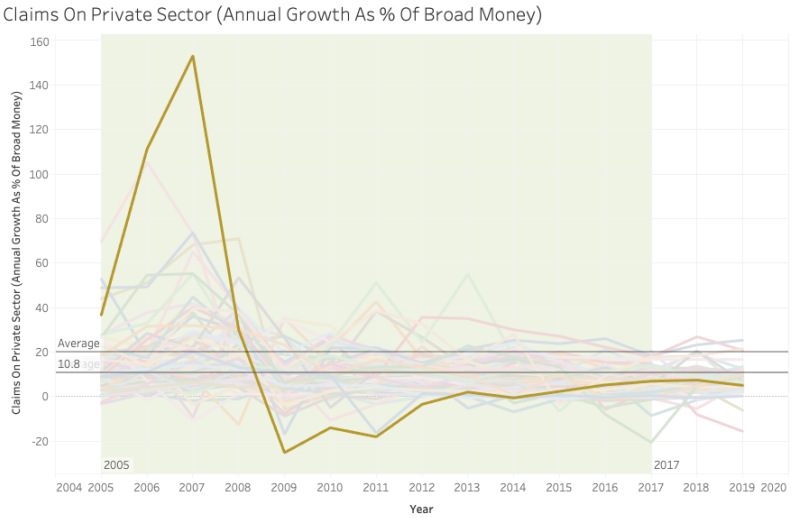
Globally



Average globally: 10.8%

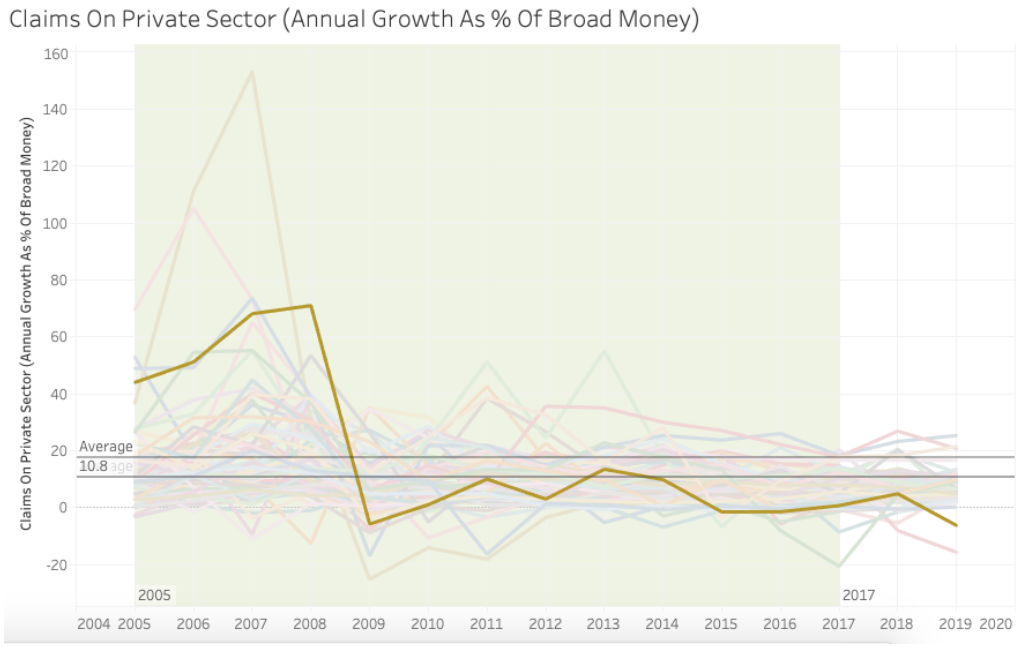
Iraq average: 3.2% (flat)
Afghanistan average: 3.9%(flat, drops to lowest(-16.3) on 2011)
outliers:Montenegro,Ukraine,Kazakhstan (fluctuated greatly)

The case of Montenegro:



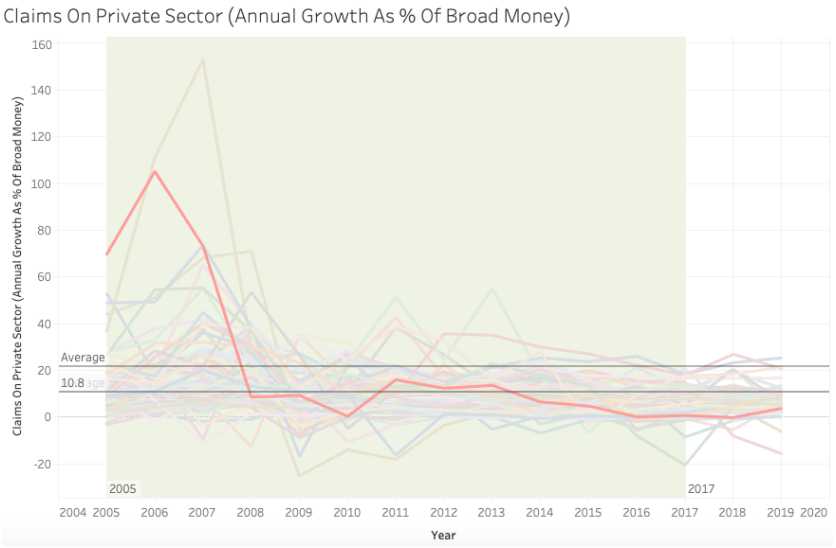
The average is 19.9% which is above average globally
The inflation peaks on 2007 then drops quickly and then keeps flat below the global average

The case of Ukraine:



The average is 17.5% which is above average globally
The inflation peaks on 2007 then drops quickly and then keeps flat below the global average

The case of Kazakhstan:



The average is 21.9% which is above average globally
The inflation peaks on 2006 then drops quickly and then keeps flat below the global average

4. **Inflation, Gdp Deflator: Linked Series (Annual %)**

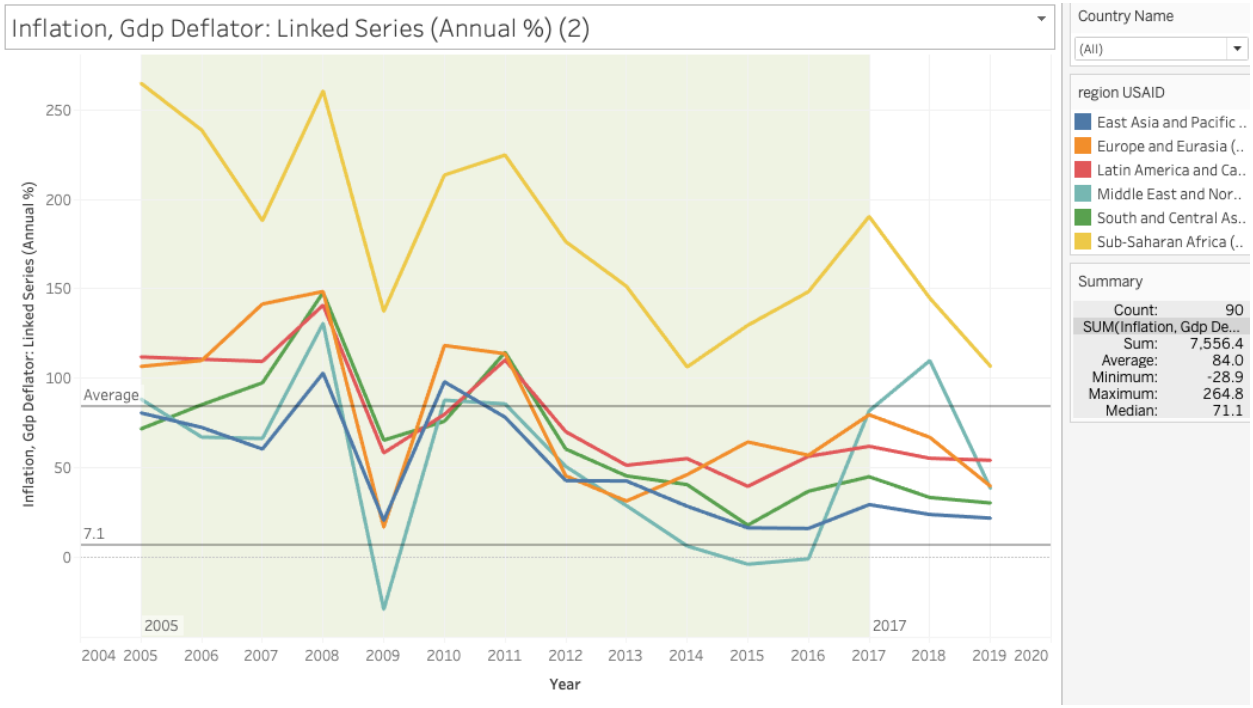
definition:

Inflation as measured by the annual growth rate of the GDP implicit deflator shows the rate of price change in the economy as a whole. This series has been linked to produce a consistent time series to counteract breaks in series over time due to changes in base years, source data and methodologies.

Influence:

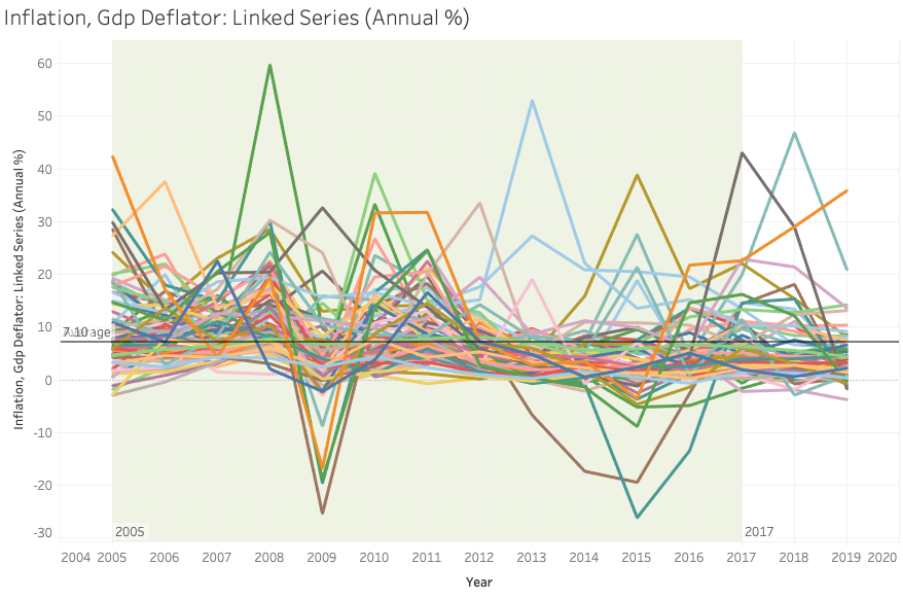
the **higher** the inflation is not good economy since it reduces the money value and cost of the living rises
the **lower** the inflation is **good** for GDP or growth of the economy overall

Correlation between inflation and CPI: -0.0196
Regionally:



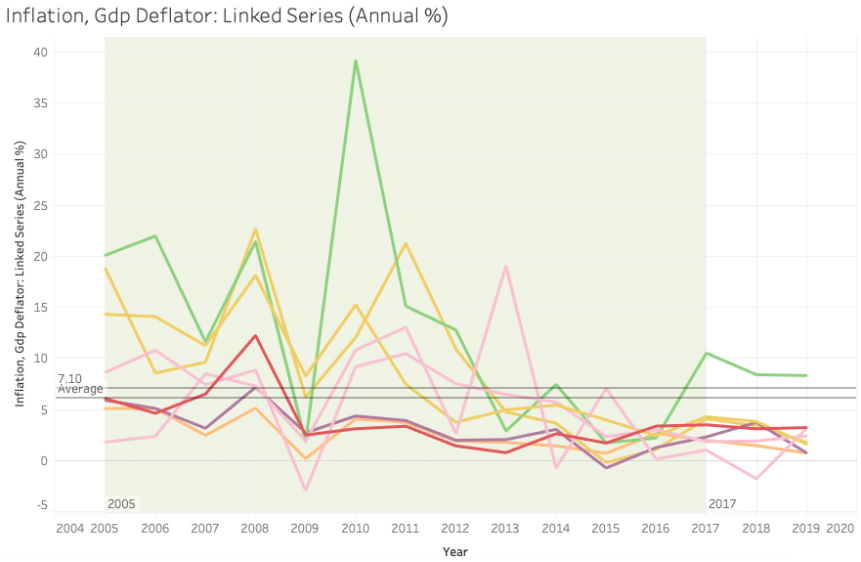
By looking at the regional level, the average is very high, which is 84%. the regional average is 10 times higher than the country-level. The SSA has very high inflation over the years and also fluctuates tremendously. SSA drives the regional average 10 times higher than the country-level

Globally



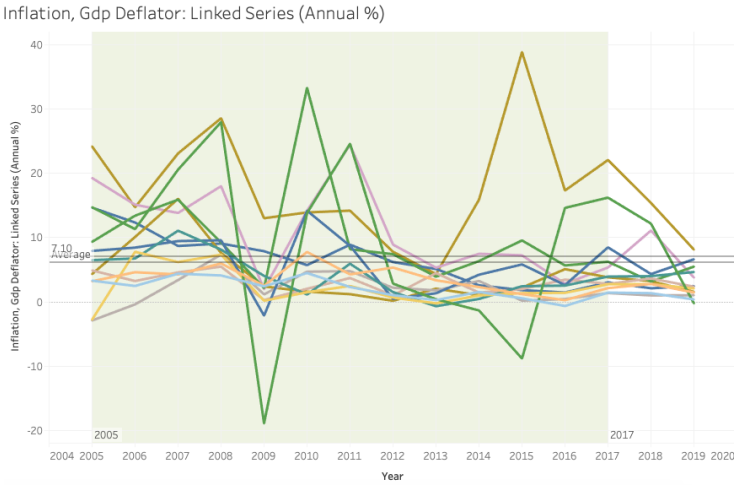
Average globally: 7.1%
Iraq average: 24.69% (fluctuate greatly)
Afghanistan average: 16.59%(fluctuate greatly)
Overall, many countries fluctuate greatly around the world average so we look at regional level

a) East Asia and Pacific (EAP)



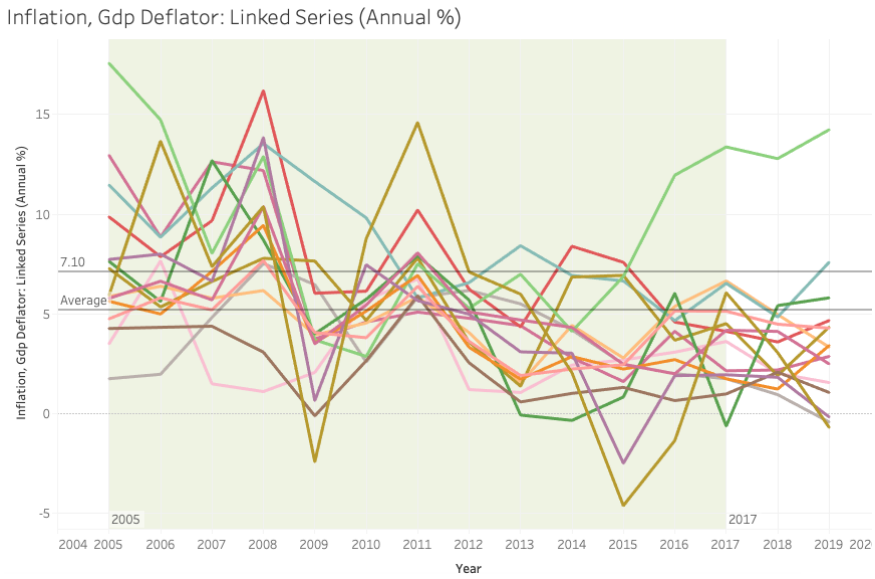
The regional average is 6.15% slightly above global.
The outliers are Mongolia which peaks in 2010(39.18%)

b) Europe and Eurasia (E&E)



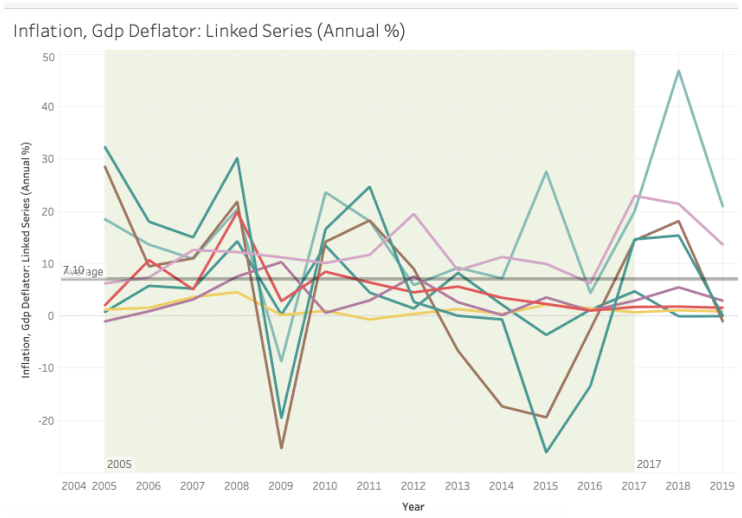
The regional average is 6.09% below global.
The outliers are Ukraine which peaks on 2016(38.88%) and Azerbaijan drops to the lowest in 2009(-18.9%)

c) Latin America and Caribbean (LAC)



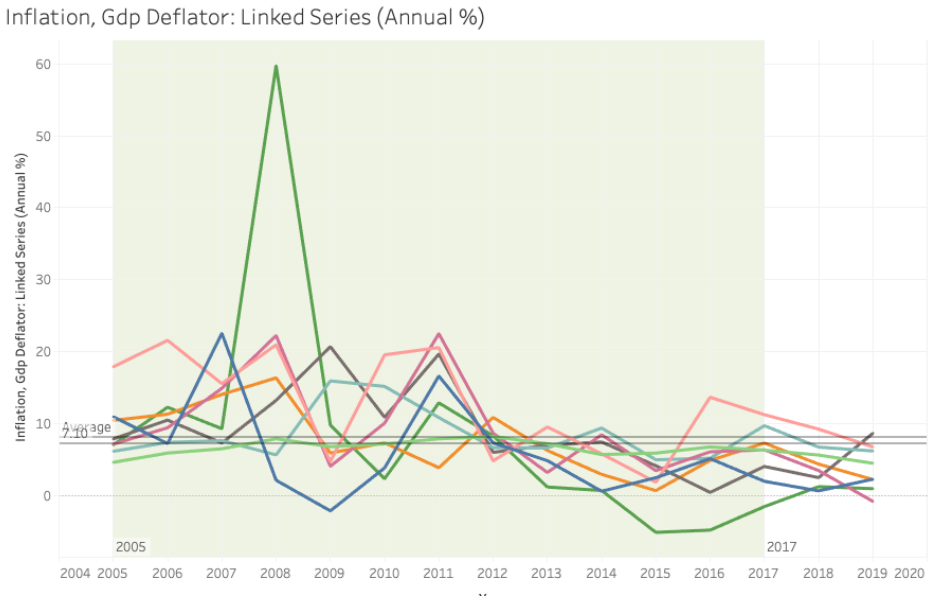
The regional average is 5.19% below global.

d) Middle East and North Africa (MENA)



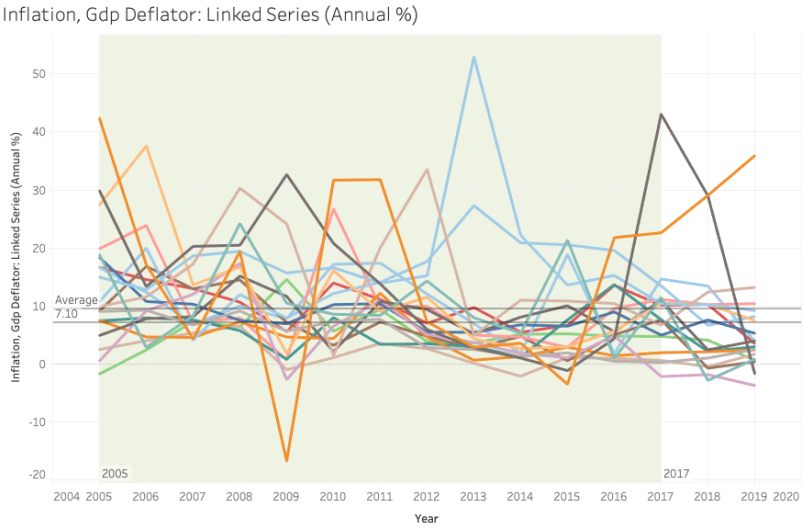
The regional average is 6.75% below global.

e) South and Central Asia (SAC)



The regional average is 8.09% above global.
The outlier is Turkmenistan which it peak to 59.74% in 2008

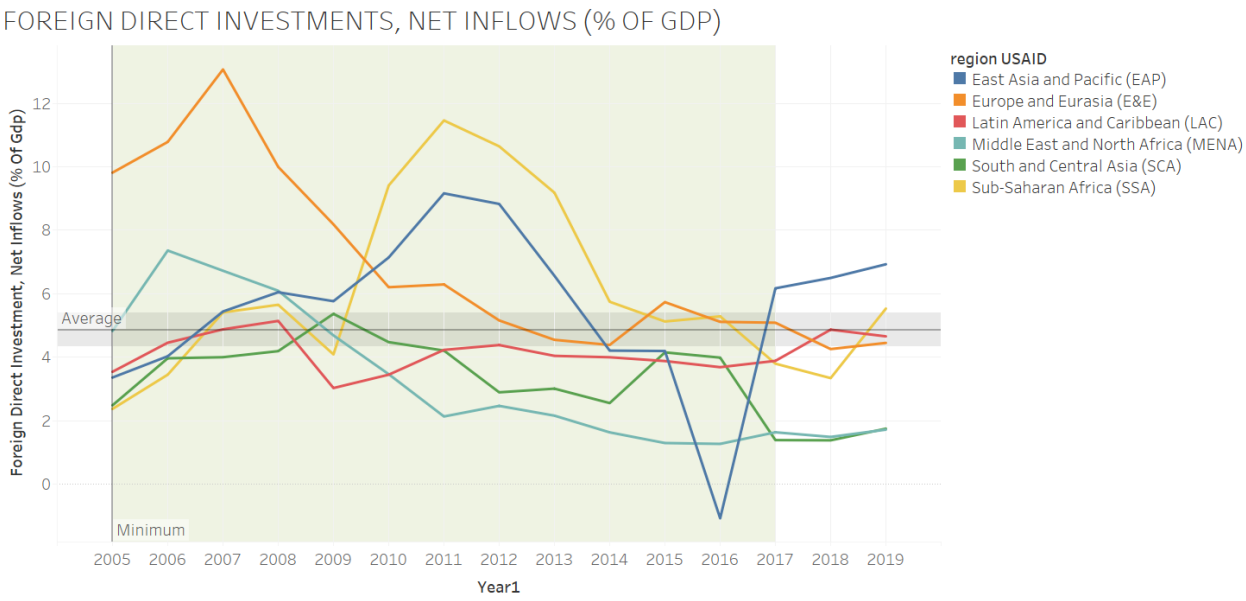
f) Sub-Saharan Africa (SSA)



The regional average is 9.41% above global.
The outliers are Angola, Ghana and Congo that fluctuate greatly.

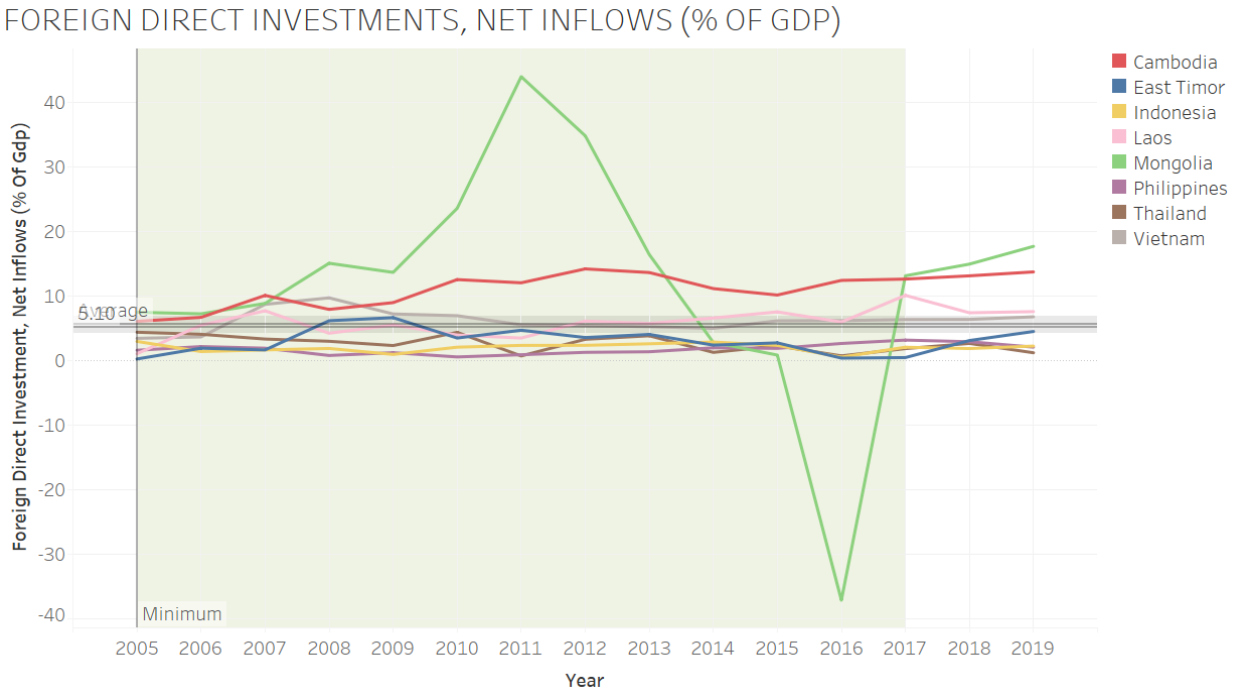
5. Foreign Direct Investment, Net Inflows (% Of Gdp)

"FDI net inflows are the value of inward direct investment made by non-resident investors in the reporting economy, including reinvested earnings and intra-company loans, net of repatriation of capital and repayment of loans."(United Nations) The higher the better for countries as well as for anti-corruption progress based on our positive correlation between this indicator and CPI.
The global average is 5.10% of GDP.



EAP is the most attractive region to foreign investors currently, while MENA and SCA have been decreasing regarding FDI net inflows, and are now the least attractive regions to foreign investors.

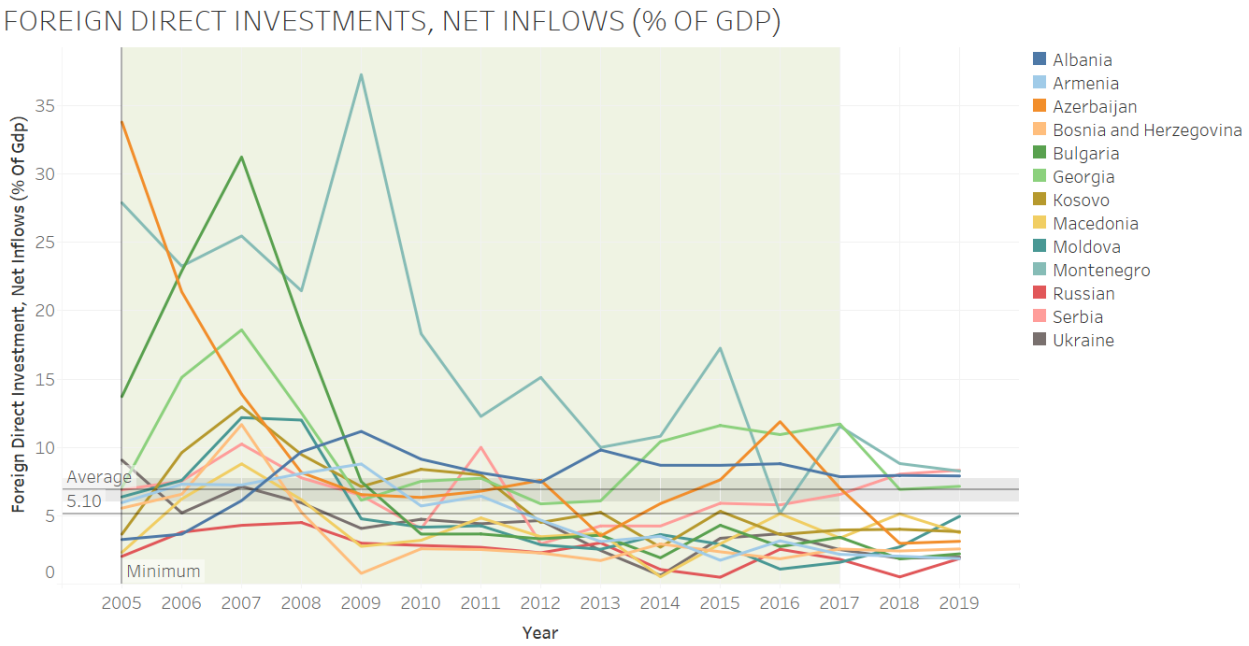
a) East Asia and Pacific (EAP)



The regional average is inline with the global average. All countries, except **Mongolia (outlier)** have experienced FDI inflows. FDI has become more and more significant to the economies of Cambodia, Laos, and Vietnam, meaning these countries are attractive places for foreign investors.

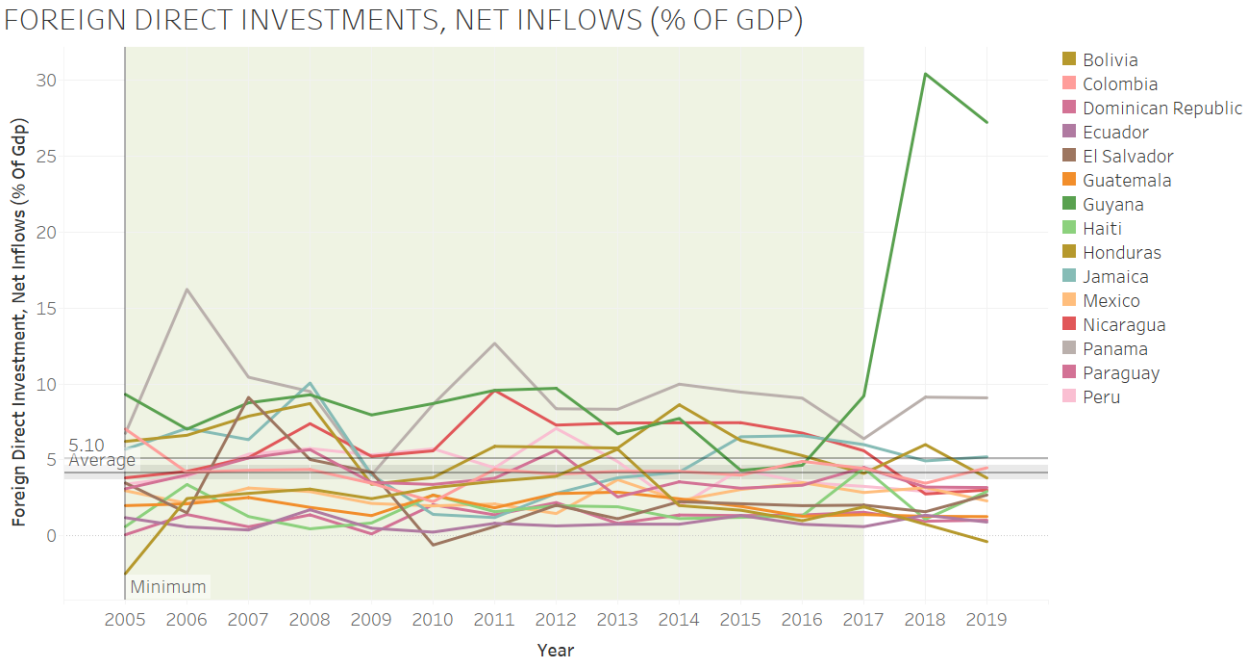
The outlier, Mongolia, saw a jump in FDI as % of GDP in 2011 and 2012, but dropped to only 2.8% in 2014. Then it dropped again in 2016 with tremendous FDI outflows (negative inflows). However, the country came back strong in 2017 and has stayed high since.

b) Europe and Eurasia (E&E)



The regional average of 6.87% is above the global average, with major thanks to Montenegro. All countries had positive FDI inflows from 2005 to 2019.

c) Latin America and Caribbean (LAC)

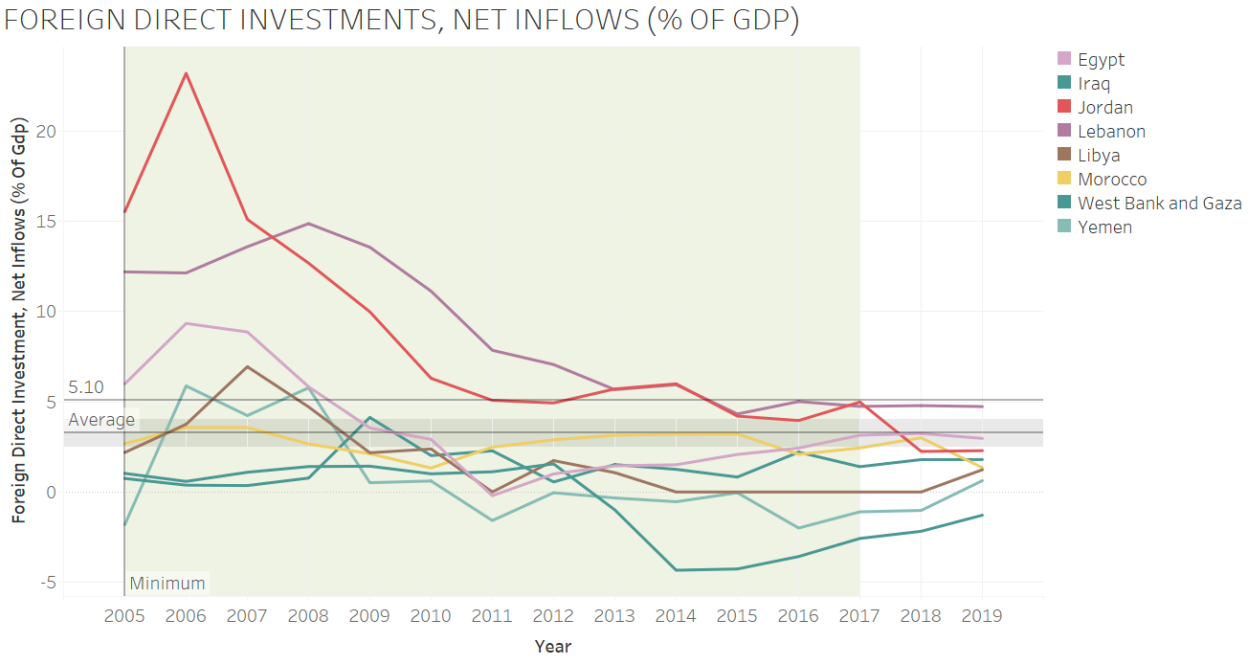


The regional average of 4.15 is below the global average. The FDI share of GDP is quite stable throughout the region.

The outlier:

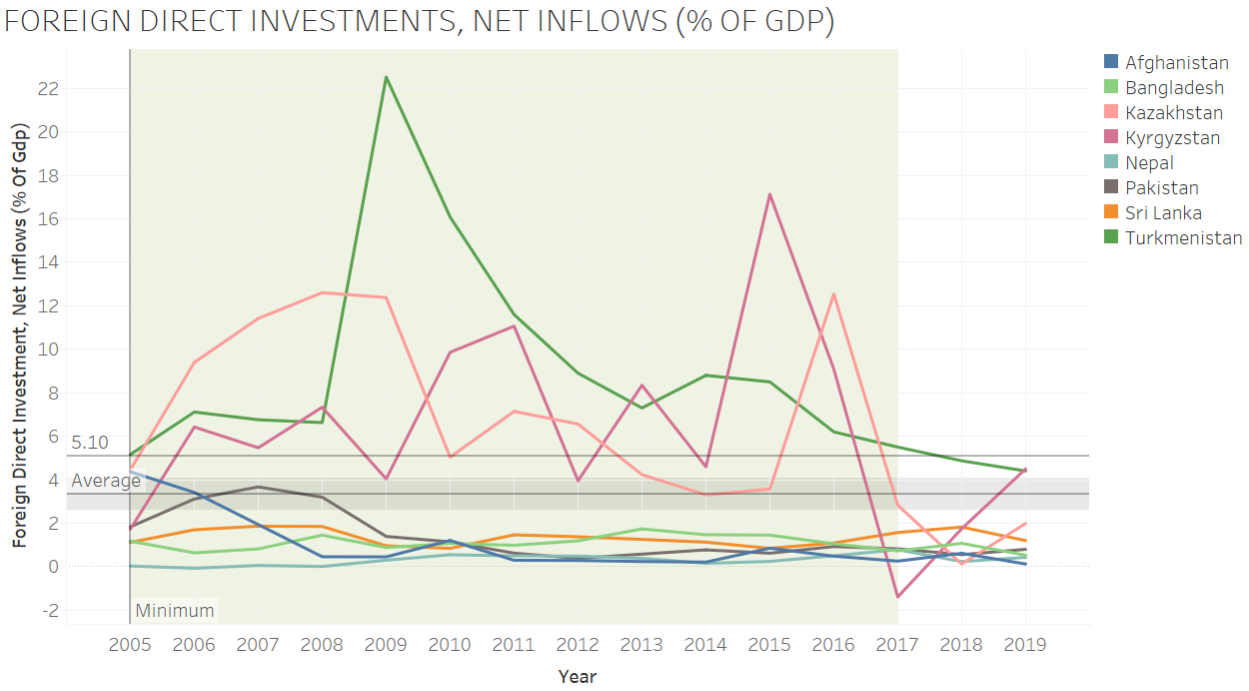
Guyana experienced a big jump in this indicator in 2017 and has stayed high ever since. This is thanks to ExxonMobil's discovery of major oil reserves off Guyana's coast.(Guyana's Department of Public Information, 2018)

d) Middle East and North Africa (MENA)



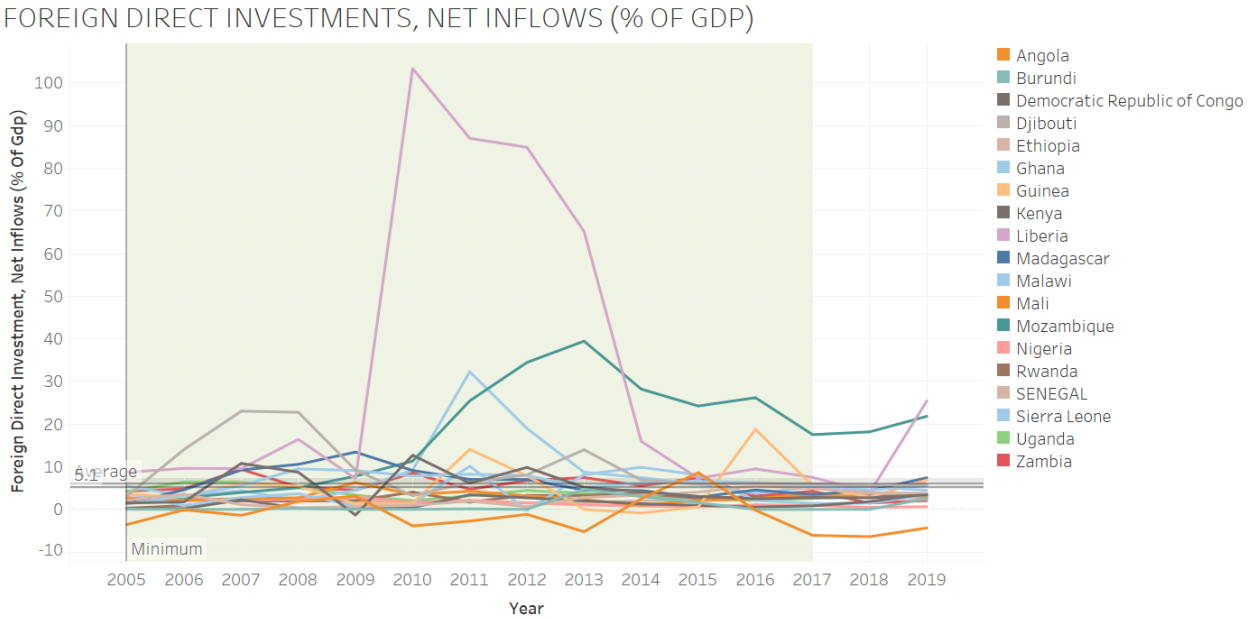
The regional average of 3.26 is below the global average. Iraq, despite being highly invested by USAID, has experienced negative FDI net inflows since 2014 due to ISIS' activities. Yemen has constantly seen near 0 FDI inflows due to the continued war and violence. Overall, the region is not as attractive to foreign investors as the past period from 2005 to 2008.

e) South and Central Asia (SAC)



The regional average of 3.32 % of GDP is below the global average. The top 3 most attractive countries to investors are Turkmenistan, Kyrgyzstan, and Kazakhstan. These countries also experienced the most fluctuations in this indicator. Despite USAID's investments, FDI investments have not grown in Afghanistan, after decreasing in 2005.

f) Sub-Saharan Africa (SSA)



The regional average of 6% is above the global average, thanks to Liberia and Mozambique. There is a huge FDI inflow to Liberia from 2010 to 2014. The situation is better for Mozambique as the country has observed stable high FDI net inflow at around 20% of its GDP since 2011.

6. Broad Money Growth (Annual %)

definition:

Broad money is the most flexible method for measuring an economy's money supply, accounting for cash and other assets easily converted into currency. ... Central banks tend to keep tabs on broad money growth to help forecast inflation.

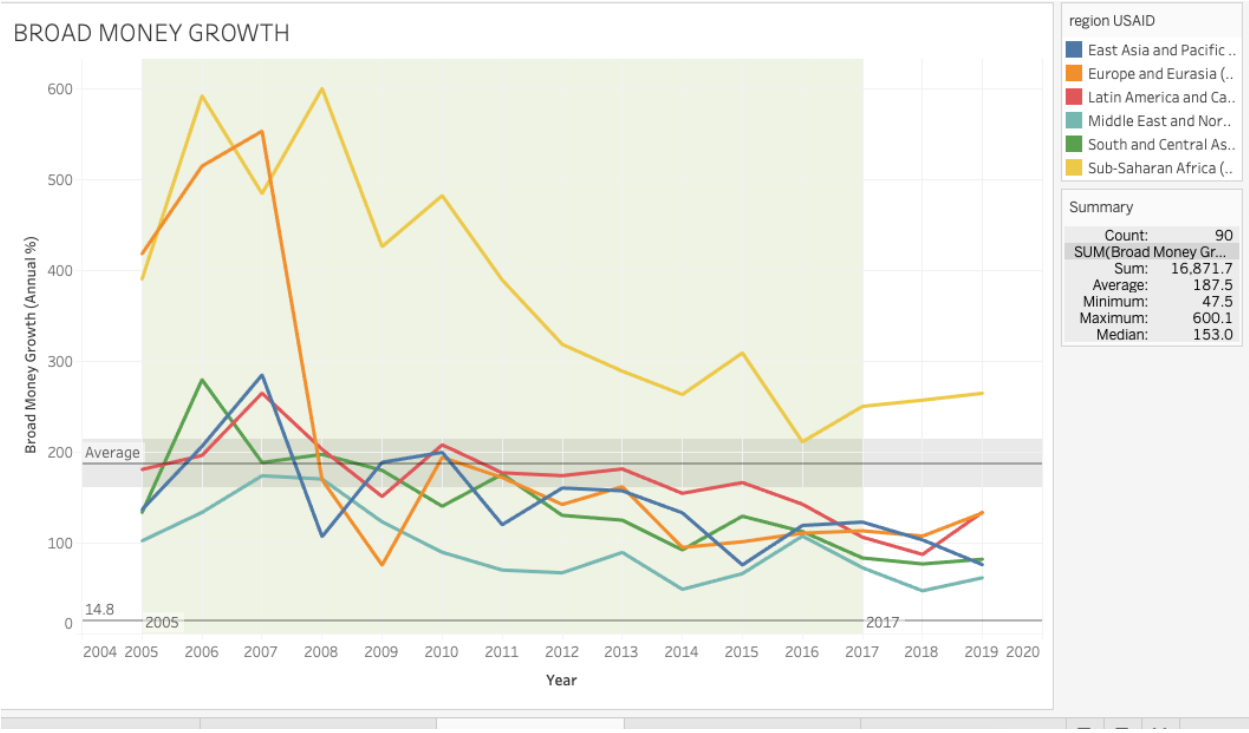
Influence:

The Federal Reserve uses lower interest rates to **increase** the **money** supply when the goal is to stimulate the economy.

the **higher** the broad money growth will help to **stimulate** the economy

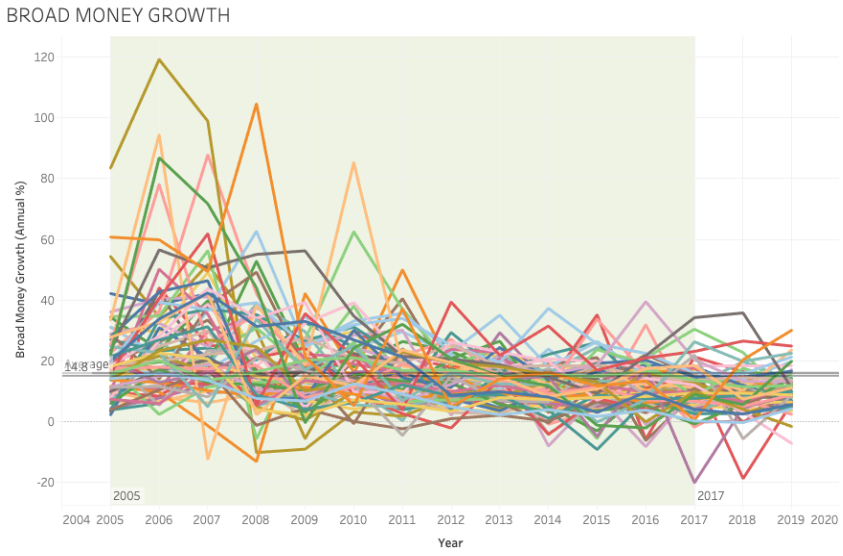
Correlation between inflation and CPI: -0.22

regionally:



By looking at the regional level, the average is very high, which is 187.5%. the regional average is 10 times higher than the country-level. The SSA has very high broad money growth over the years and also fluctuates tremendously. SSA drives the regional average 10 times higher than the country-level. Although the rate keeps in a high level, but all gradually declines.

Globally



Average globally: 14.8%

Generally have high broad money growth from 2006 to 2008, and then keeps decreasing

Iraq average: 14.71% (peak in 2007(37.11%) then keeps dropping)

Afghanistan average: 17.39% (peak in 2007(42.4%) then keeps dropping)

The case of Iraq **:**

The case of Afghanistan**:**

7. **Foreign Direct Investment, Net Outflows (% Of Gdp)**

definition:

A foreign direct investment (FDI) is an investment made by a firm or individual in one country into business interests located in another country. Generally, FDI takes place when an investor establishes foreign business operations or acquires foreignbusiness assets in a foreign company.

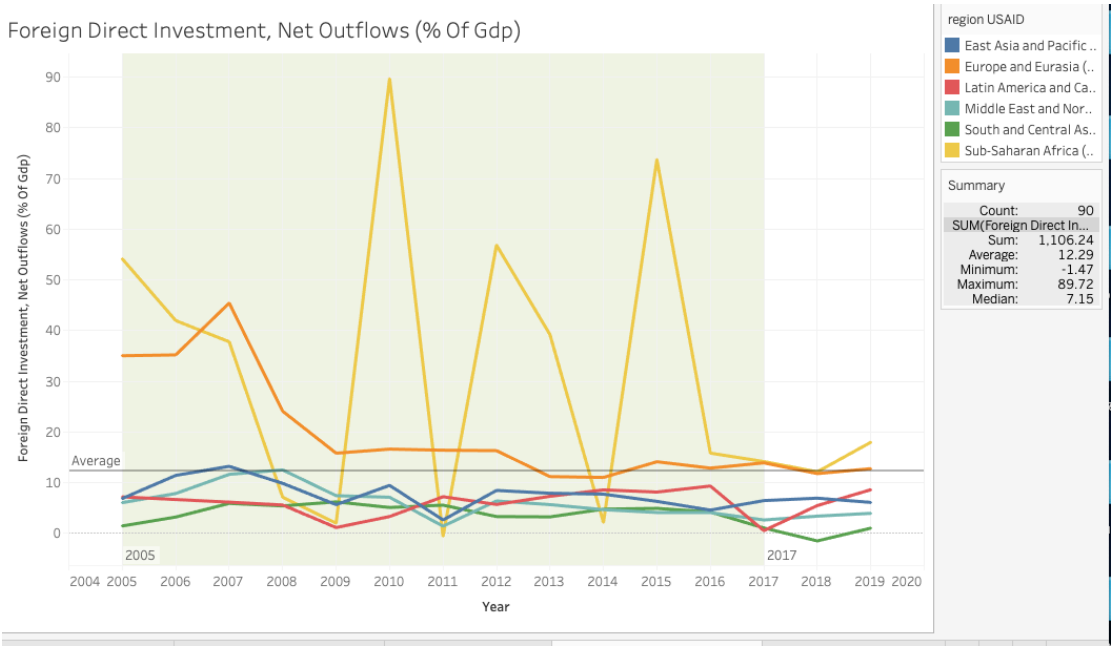
Influence:

FDI and economic growth are positively interdependent. Large economic growth provides high profit opportunities attracting higher domestic and foreign direct investments.

the **higher** the **FDI** is **good** for GDP or growth of the economy overall

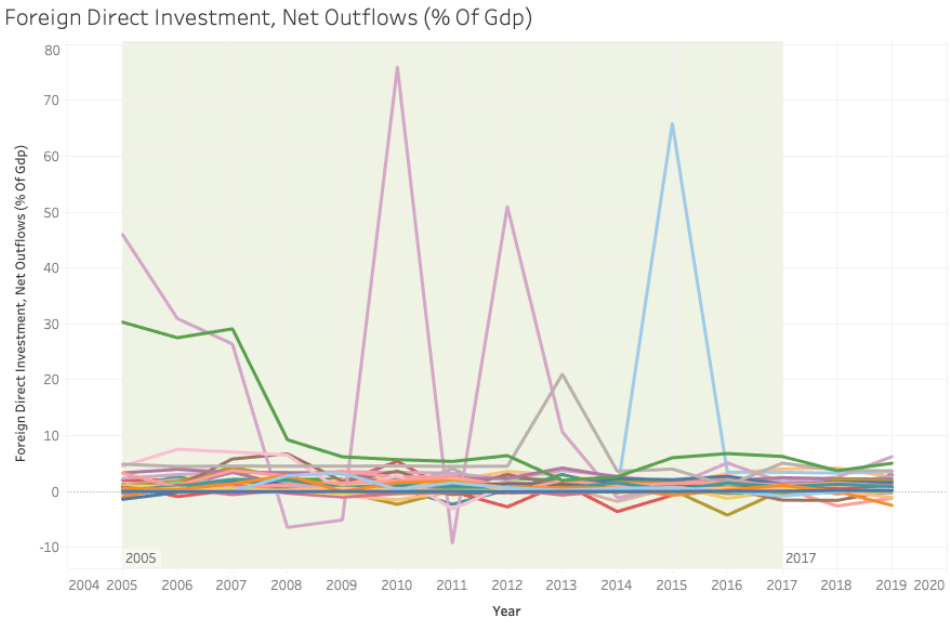
Correlation between inflation and CPI: 0.014

Regionally:

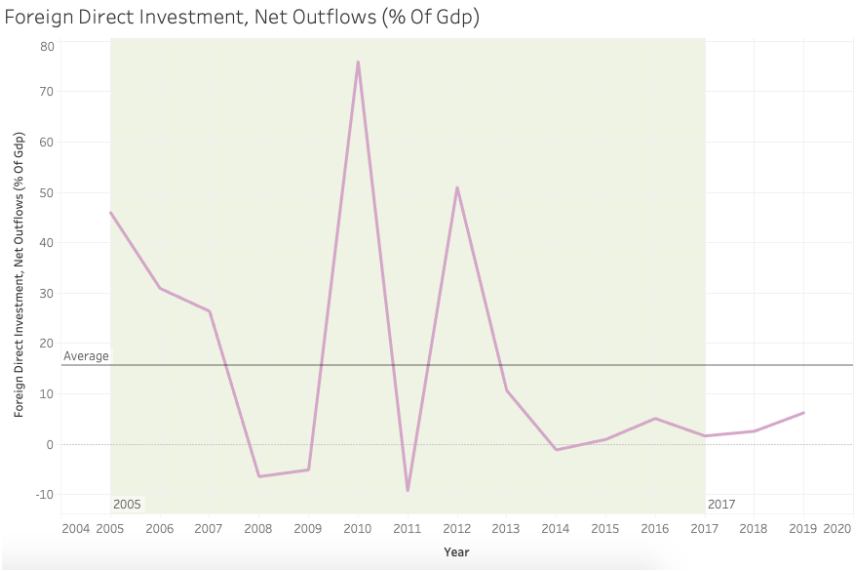


By looking at the regional level, the average is 12.29%. the regional average is 10 times higher than the country-level. The SSA fluctuates greatly over the years and peak in 2010 which is around 90%

Globally

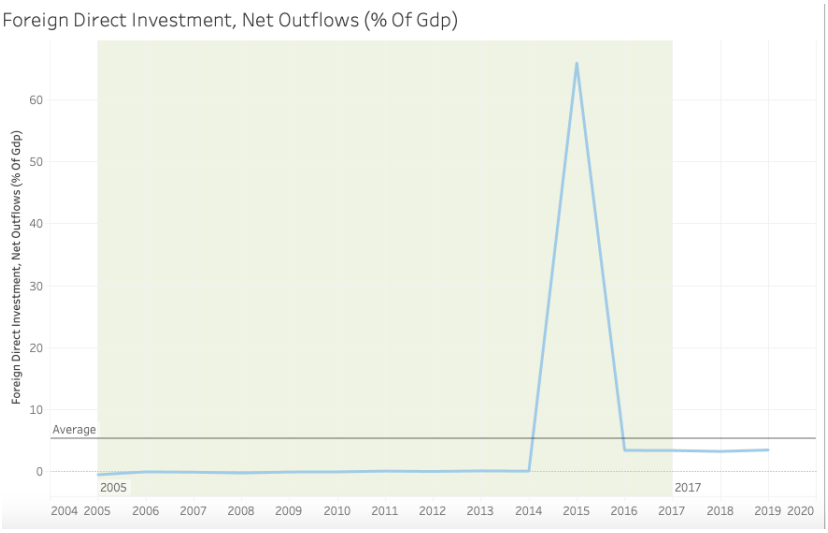


Average globally: 1.04%
Iraq average: 0.128% (flat)
Afghanistan average: 0.0283%(flat)
Outliers:Liberia,Sierra Leone,Azerbaijan (fluctuated greatly)
The case of Liberia:



The average is 15.71% which is above average globally
The inflation peaks in 2010(76%),2012(51.02%)

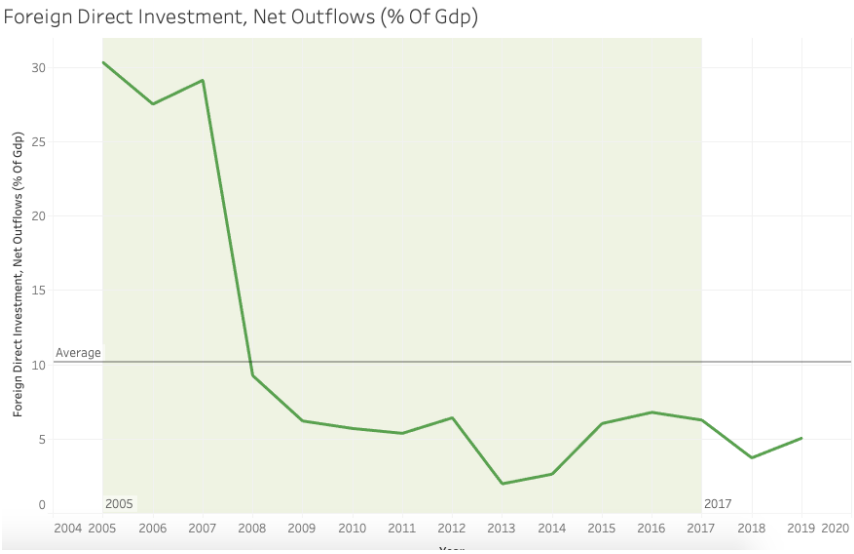
The case of Sierra Leone:



The average is 5.29% which is above average globally
The inflation peaks in 2015(65.89%) and normal for other years

The case of Azerbaijan:

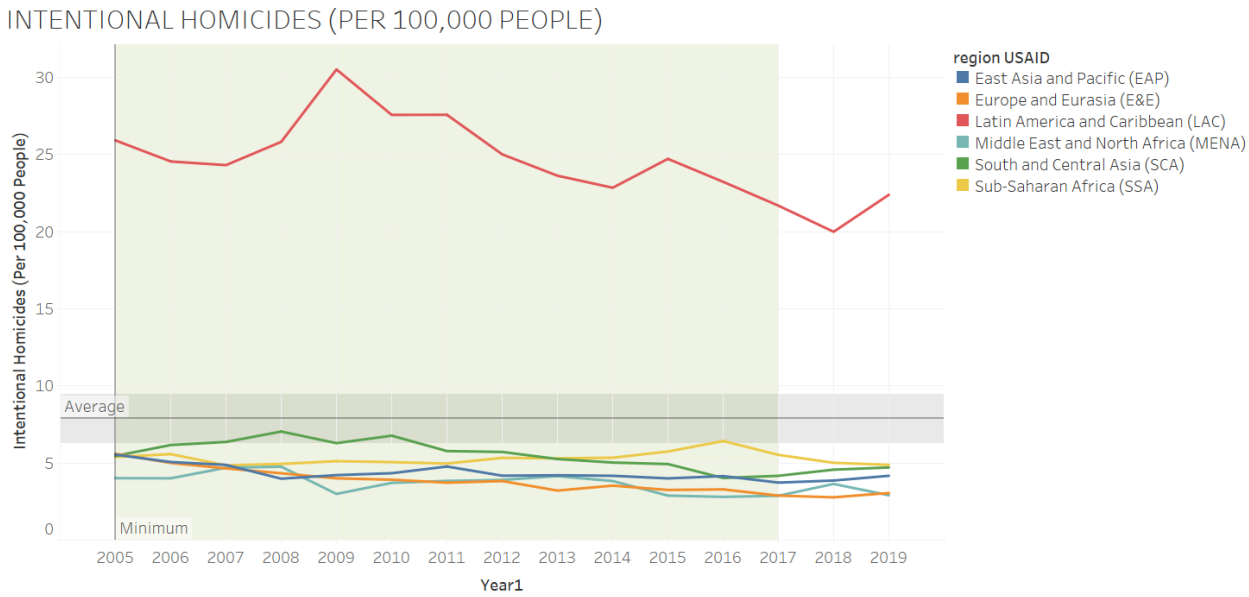
The average is 10.17% which is above average globally
The inflation peaks in 2005(30.33%) and keeps decreasing



8. Intentional Homicides (Per 100,000 People)

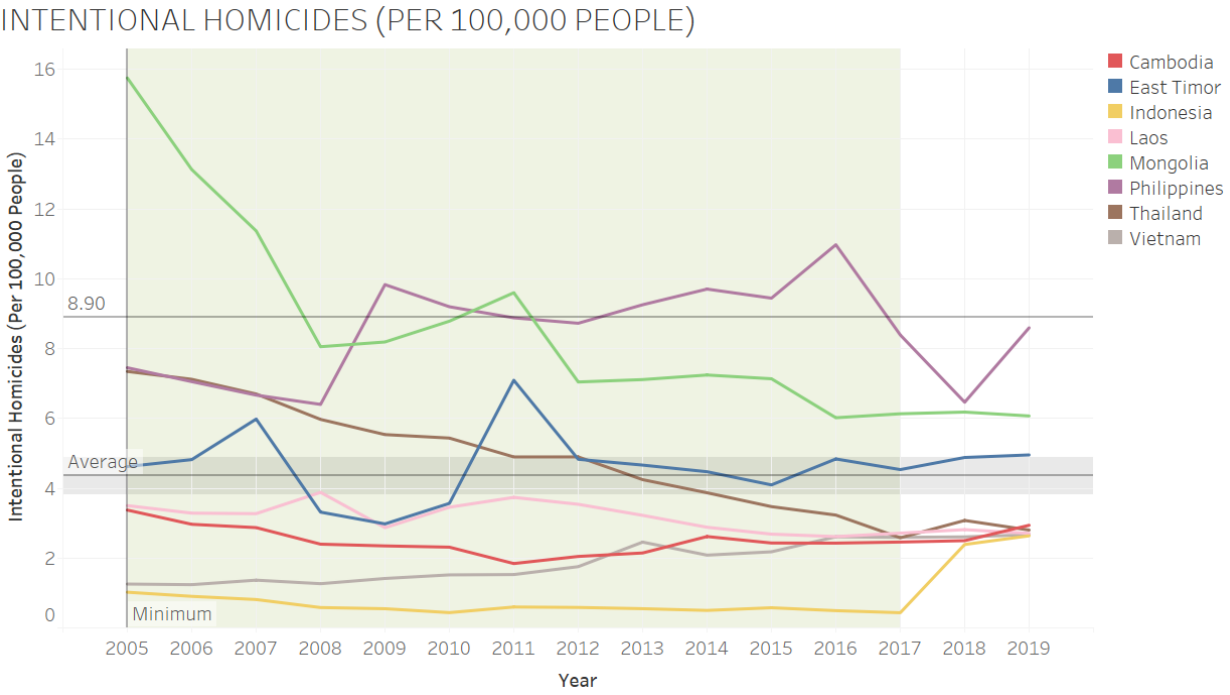
"Intentional homicide means killing a human being willfully and illegally. That means the intent was to cause death or serious injury, but not necessarily that it was planned beforehand." (Eurostat, 2018). The lower this rate the better.

The global average is 8.9 deaths per 100,000 people.



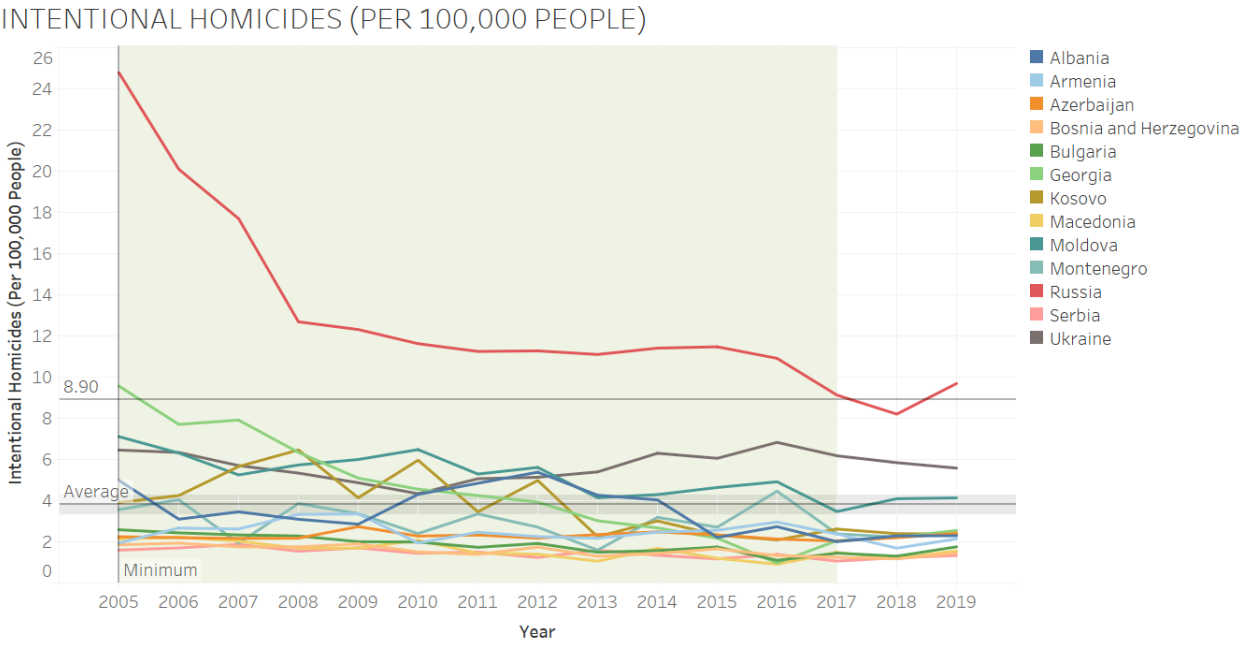
LAC is the most worrisome case, an outlier to the world. USAID may implement projects focusing on Rule of Law as well as enforcement to help LAC.

a) East Asia and Pacific (EAP)



The regional average is 4.35 deaths, well below the global average. Indonesia, 5th place regarding investment amount, shows the lowest number of deaths, especially during the time window of USAID projects. This is a good sign for the projects' effectiveness. After 2017, the deaths number in Indonesia went up, which may show that the country needs continued support. **Mongolia shows the best improvement with continued decrease in the rate.** Indonesia and the Philippines are the only 2 that showed an upward trend from 2018.

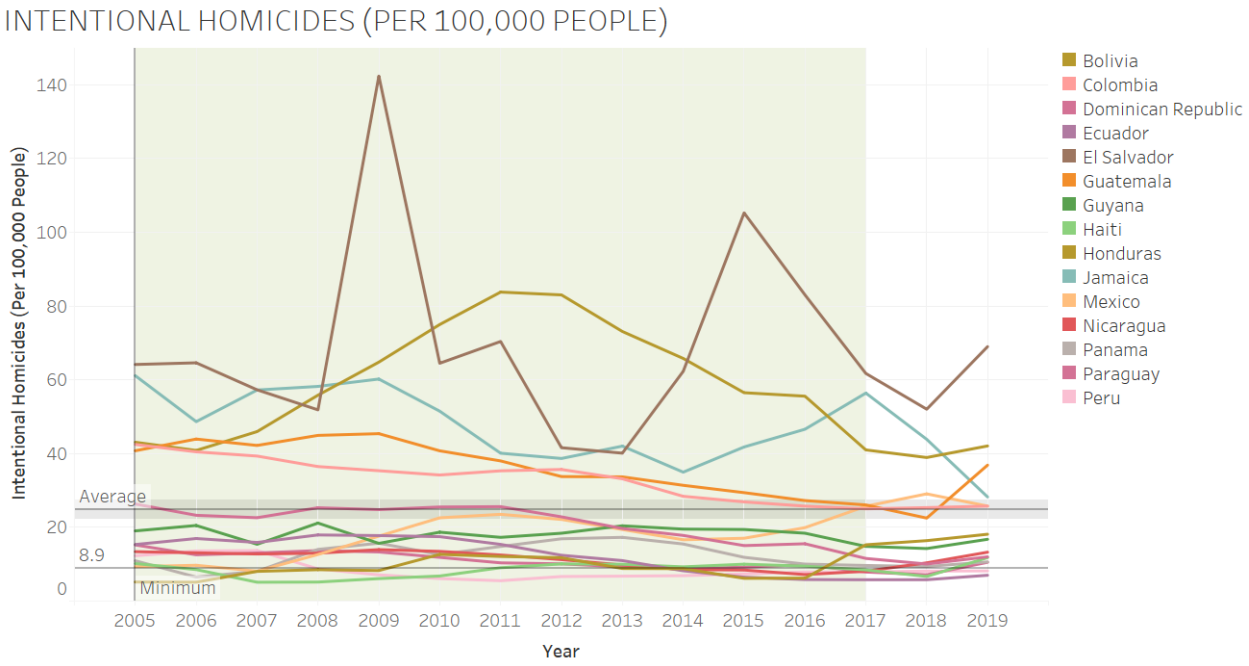
b) Europe and Eurasia (E&E)



The regional average is 4.05 deaths, well below the global average. The countries either have a stable rate or show improvement overtime, which is positive news.

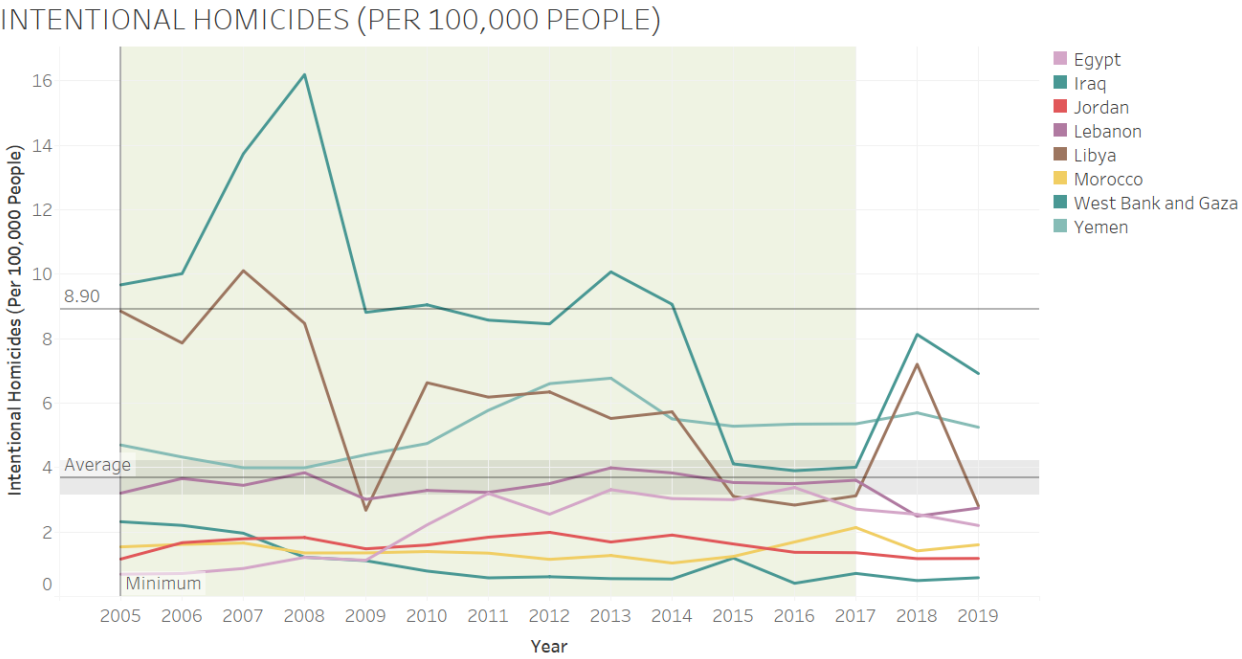
The outliers:
The outstanding case here is **Russia**, going down from 25 deaths in 2005 to 9.7 deaths in 2019. The major drop in Russia happened from 2006 to 2008, the same time with USAID projects focusing on Rule of Law here. **We may say Russia is a successful case of USAID investment.**

c) Latin America and Caribbean (LAC)



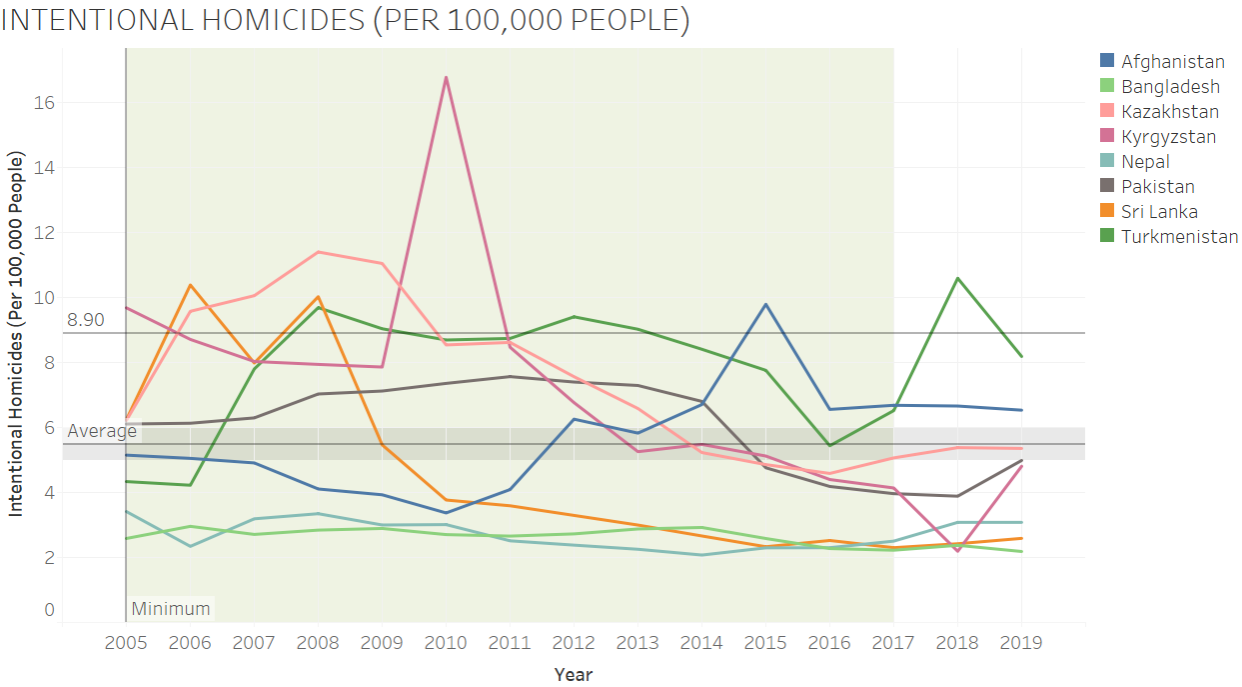
The regional average is 24.8, much higher than the global average, and is the only region whose rate is higher than the global average. LAC has been plagued by crimes and violence for years. El Salvador is the most extreme case. Although the rate has come down a lot from the peak in 2009, it still stays very high. Honduras shows the most improvement, coming down from over 80 deaths in 2011 to over 40 deaths in 2019. Jamaica also shows gradual decrease.

d) Middle East and North Africa (MENA)



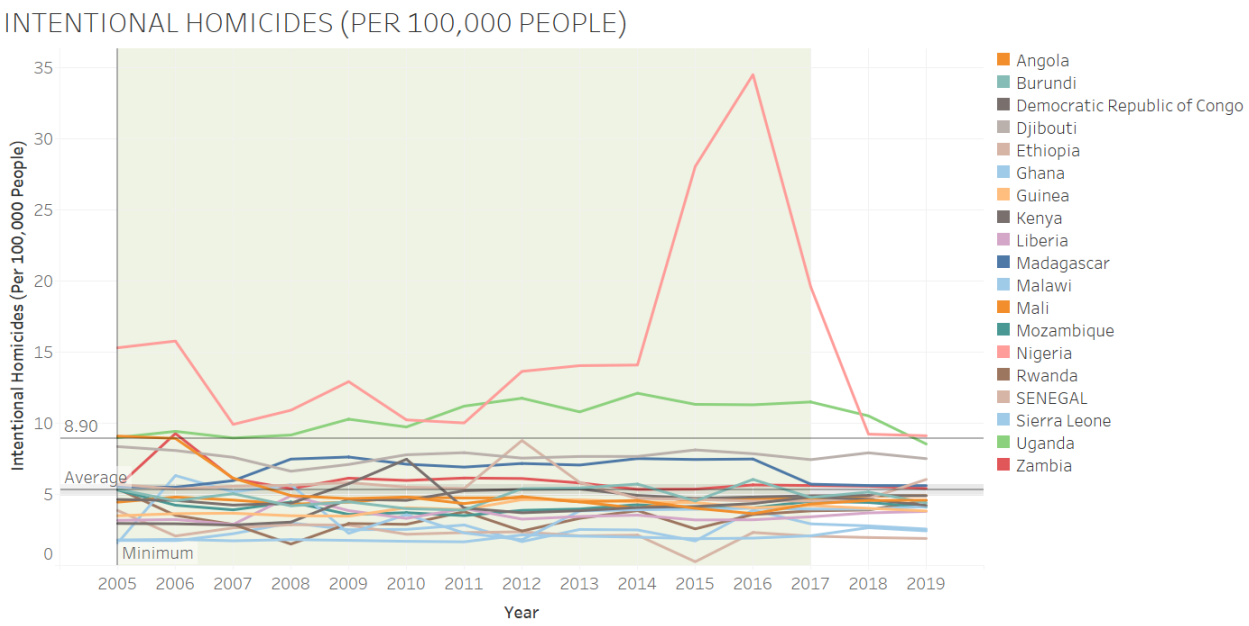
The regional average is 3.67, well below the global average. Iraq and Libya are in the top 3 of countries with the highest intentional homicides rate in the region, but both show a downward trend, which is good news.

e) South and Central Asia (SAC)



The regional average is 5.49 deaths per 100,000 people. Kyrgyzstan, Kazakhstan, Pakistan, and Bangladesh have been improving with lower and lower rates. Meanwhile, despite the large amount of investment, the rate has been trending upward in Afghanistan. Another worsening case is Turkmenistan.

f) Sub-Saharan Africa (SSA)



The regional average is 5.3.

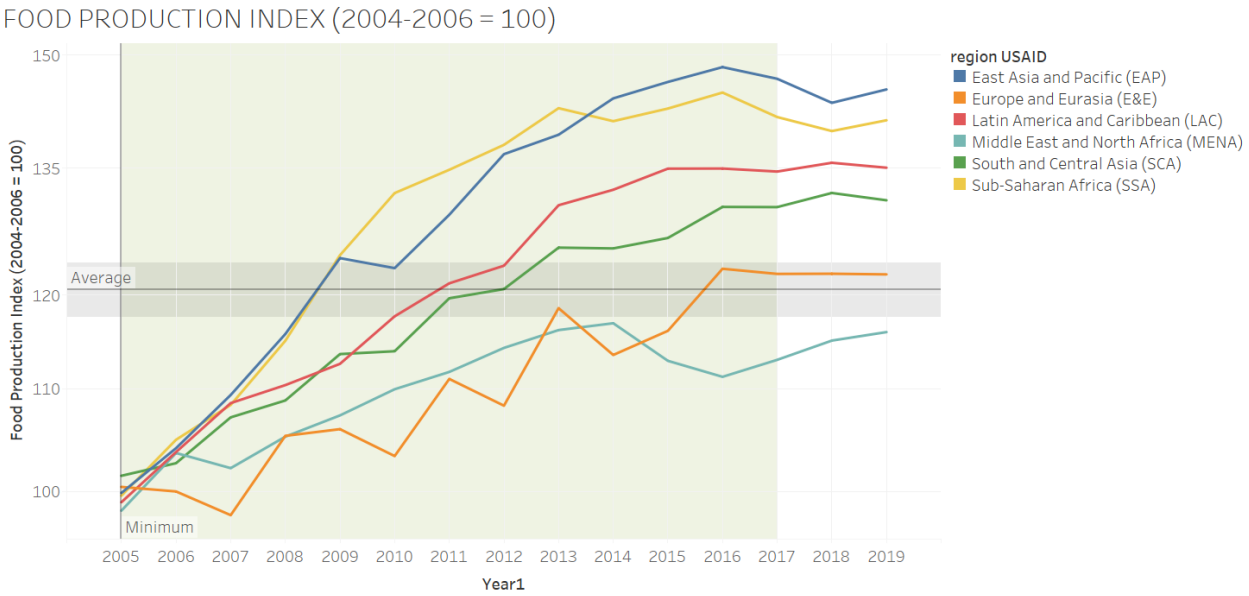
The outliers:

Nigeria, for most of the time in our considered period, stay #1 regarding the highest rates of intentional homicides. This rate skyrocketed in 2015, going even higher in 2016,lowering in 2017, then dropping back to the normal level in 2018. The surge in 2015 and 2016 is due to a series of massacres related to politics and religion.

9. Food Production Index (2004-2006 = 100)

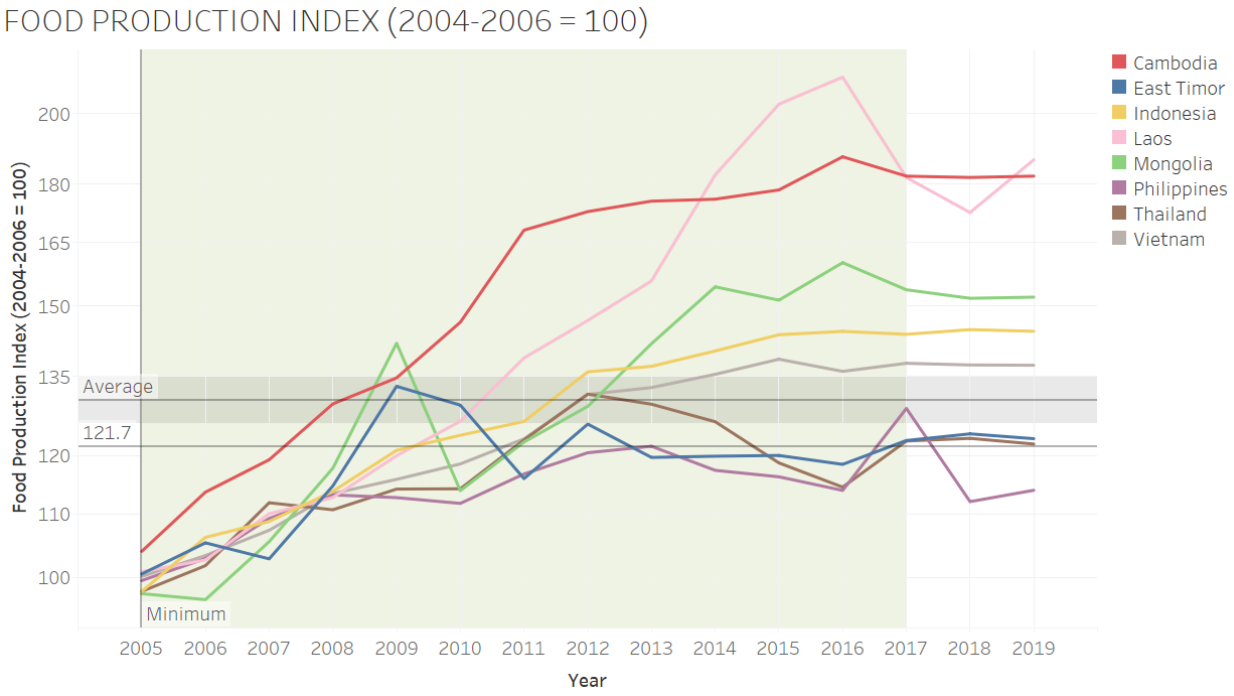
"Food production index covers food crops that are considered edible and that contain nutrients." (World Bank, 2020). For the case of developing countries, the higher the index, the better.

The global average is 121.7, meaning that since the 2004-2006 period, the global food production has grown by 21.7%.



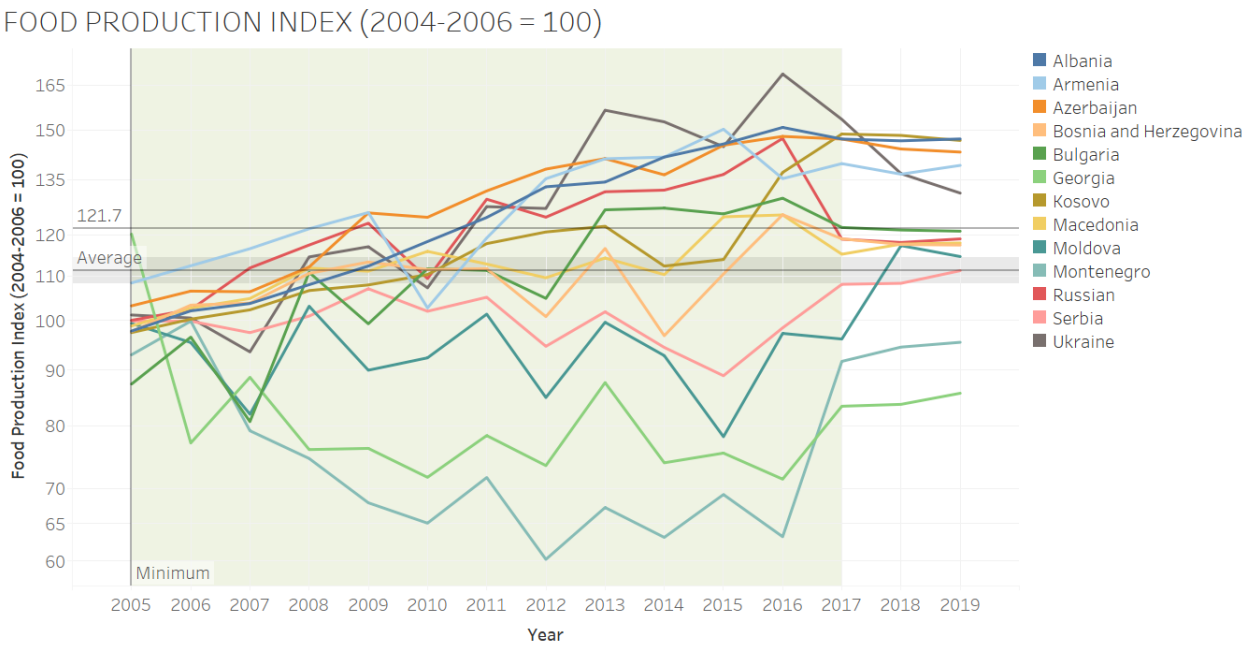
All regions experience an increase but with a diminishing rate. MENA has the slowest growth, with a dip from 2014 to 2017, but has come back in 2017.

a) East Asia and Pacific (EAP)



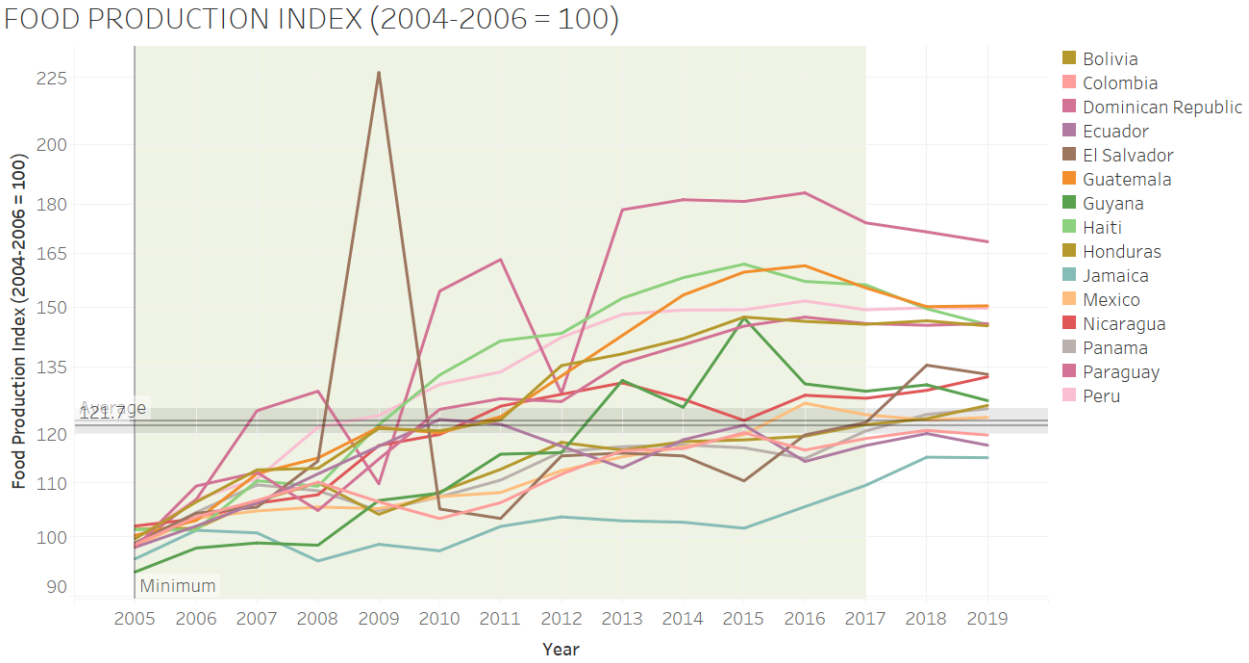
The regional average is 130.4, above the global average. The upward growth trend is stable across the region. The most impressive growth coming from Laos and Cambodia (both around 80% growth since 2004-2006 period), then Mongolia as a far third. **Since agriculture is the major economic driver for Laos and Cambodia, the USAID projects focusing on economic growth from 2007-2012 may have resulted in the incredible increase in Food Production.**

b) Europe and Eurasia (E&E)



The regional average is 111.4, below the global average. Overall, E&E suffered from a drop in growth of food production from 2007 to 2014. All countries except for Russia have experienced an upward trend since 2014. However, Georgia's and Montenegro's food production is still shrinking, though not as much.

c) Latin America and Caribbean (LAC)

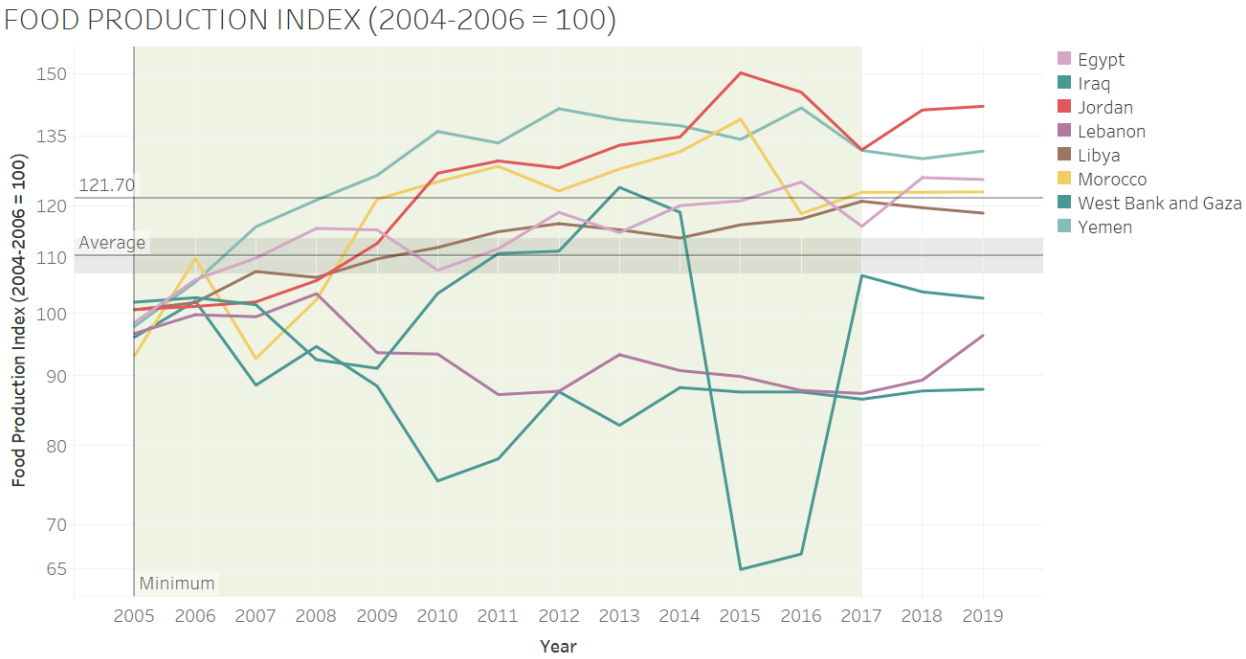


The regional average (122.9) is comparable to the global average. All countries have experienced gradual increases in food production growth rate. The most impressive growth is from Paraguay, increasing 69% since the 2004-2006 period, although the countries had some large fluctuations. The index has stabilized since 2013 in Paraguay.

The outliers:

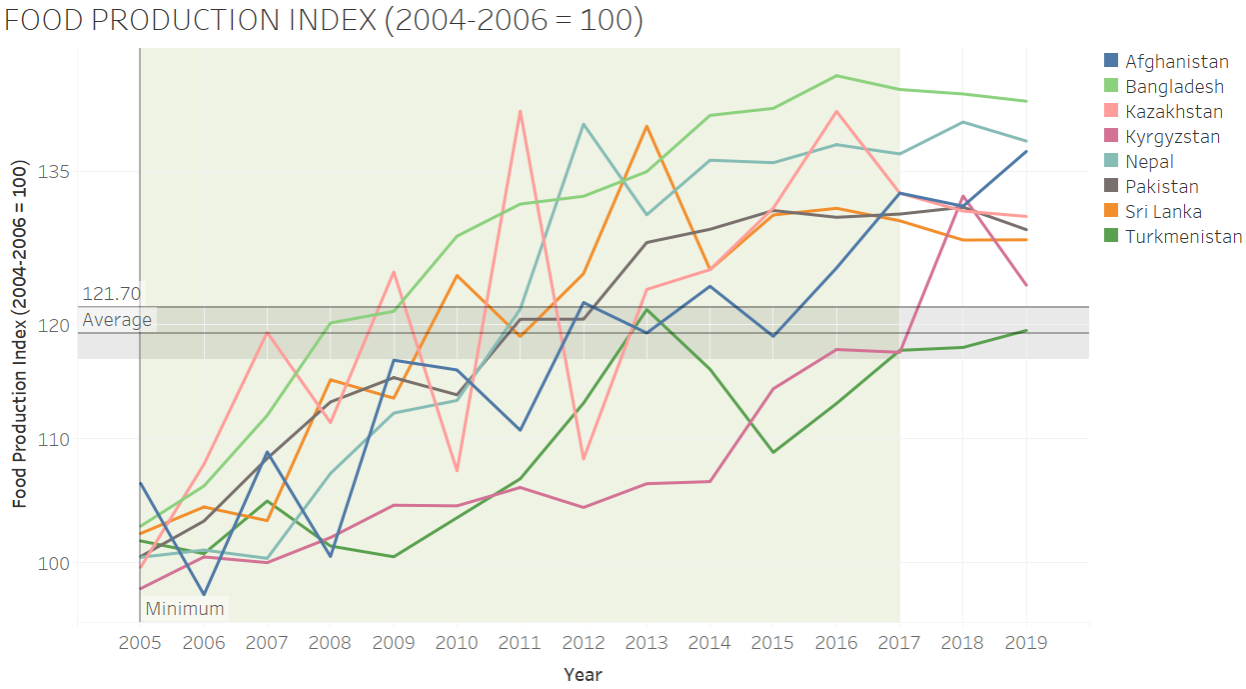
El Salvador saw a significant boost in the index in 2009 but quickly dropped to a much lower level in 2010. This peak is definitely an outlier.

d) Middle East and North Africa (MENA)



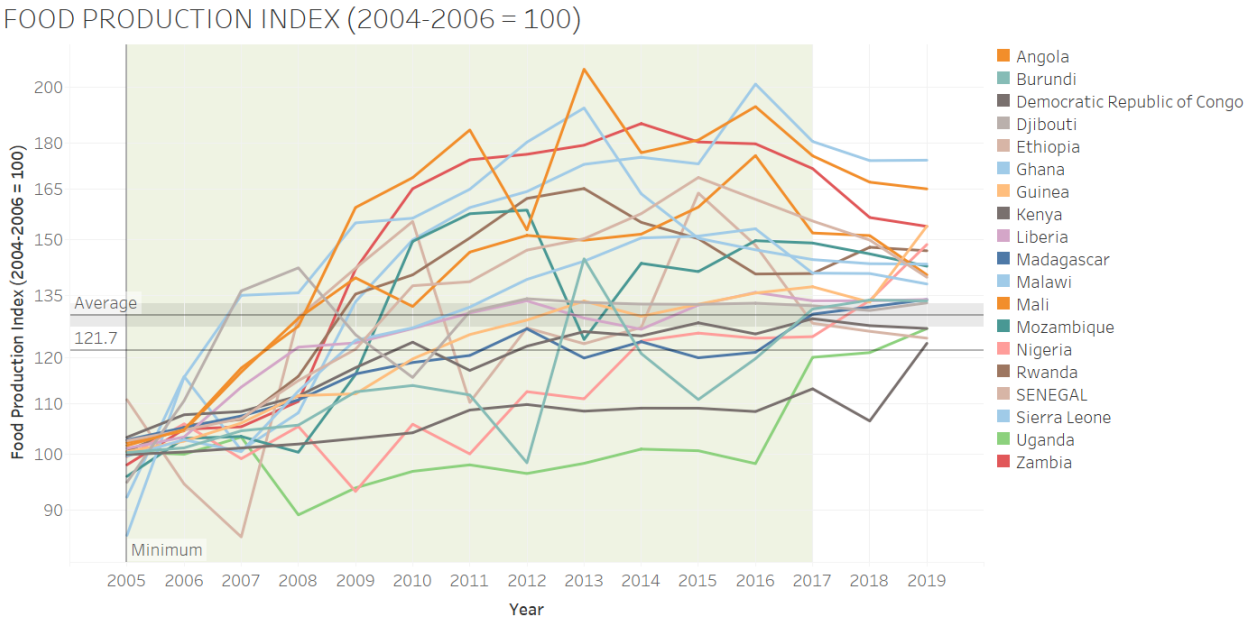
The regional average is 106.72, **well below the global average**. Iraq, Lebanon, and the West Bank and Gaza, despite being on top of the list of countries most invested by USAID, experienced shrinking in food production. These countries have suffered greatly from war, violence and political unrest as well as terrorism (ISIS). While growing rapidly from 2009 to 2014, the index in Iraq saw a huge drop in 2015. This is due to ISIS' activities. After ISIS receded in 2017, the Food Production index went back to normal growth.

e) South and Central Asia (SAC)



The regional average is comparable to the global average. All countries have seen gradual increases in the rate of growth. Kazakhstan fluctuated a lot from 2007 to 2013, but since then has stabilized. The most impressive growth belongs to Bangladesh.

f) Sub-Saharan Africa (SSA)



The regional average of 130.1 is above the global average. The most impressive growth belongs to Sierra Leone. Since 2017, all countries have positive growth rates in food production compared to the 2004-2006 period.

10. Inflation, Consumer Prices (Annual %)

Definition:

Inflation as measured by the consumer price index reflects the annual percentage change in the cost to the average consumer of acquiring a basket of goods and services that may be fixed or changed at specified intervals, such as yearly.

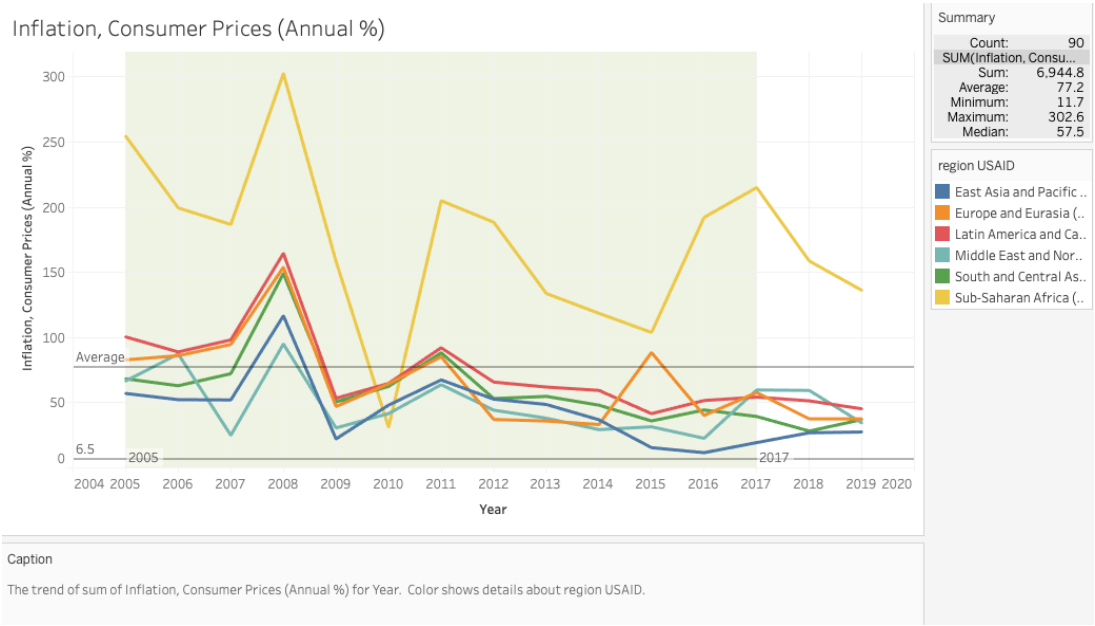
Influence:

the **higher** the inflation is not good economy since it reduces the money value and cost of the living rises

the **lower** the inflation is **good** for GDP or growth of the economy overall

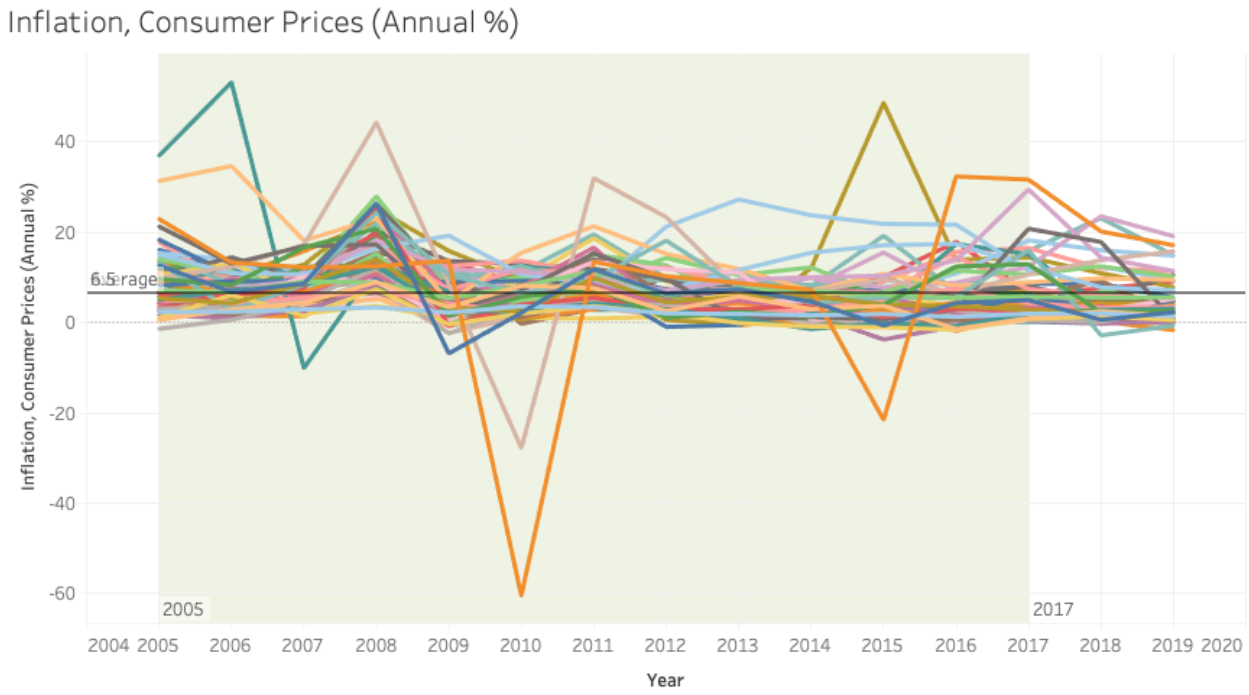
Correlation between inflation and CPI: -0.1998

Regionally:



By looking at the regional level, the average is very high, which is 77.2%, and it peaked in 2008 which is due to the great recession. The SSA has very high inflation over the years and also fluctuates tremendously. SSA drives the regional average 10 times higher than the country-level

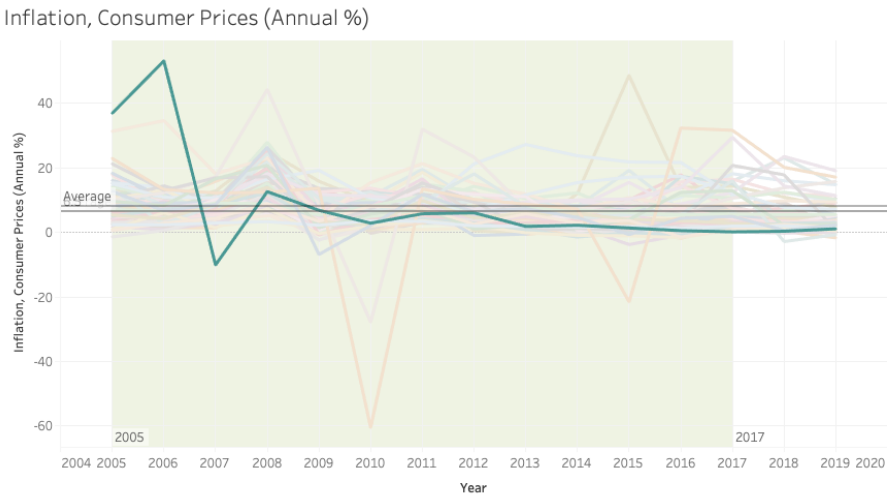
Globally:



Average globally: 6.5%

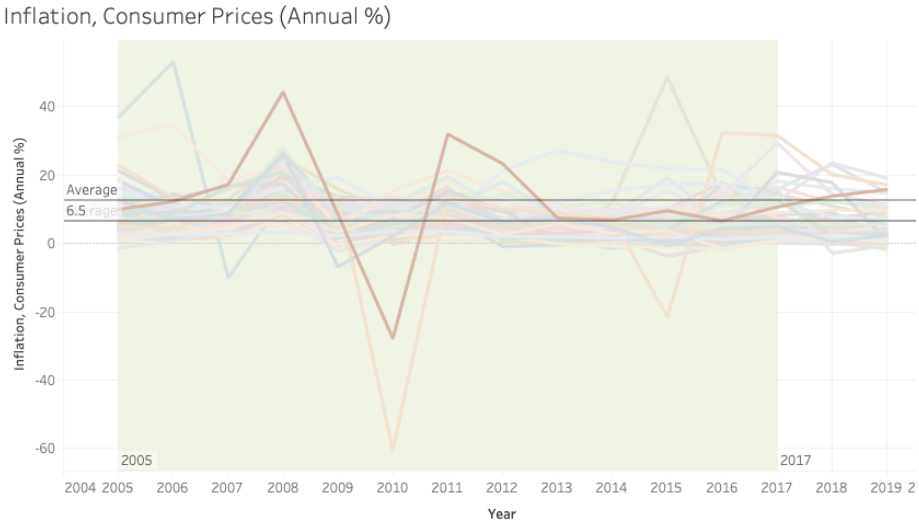
outliers:Iraq, Ethiopia, Angola, Ukraine, Afghanistan (fluctuated greatly)

The case of Iraq:



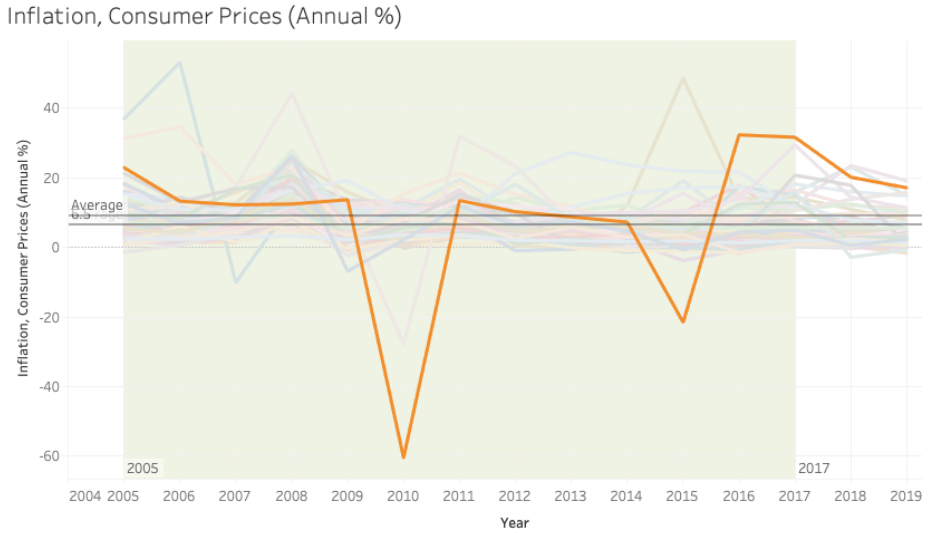
The average is 8.15% which is above average globally
The inflation is high from 2005 to 2006 but declines after 2006 and flatten afterwards that might because the anti-corruption projects work.

The case of Ethiopia:



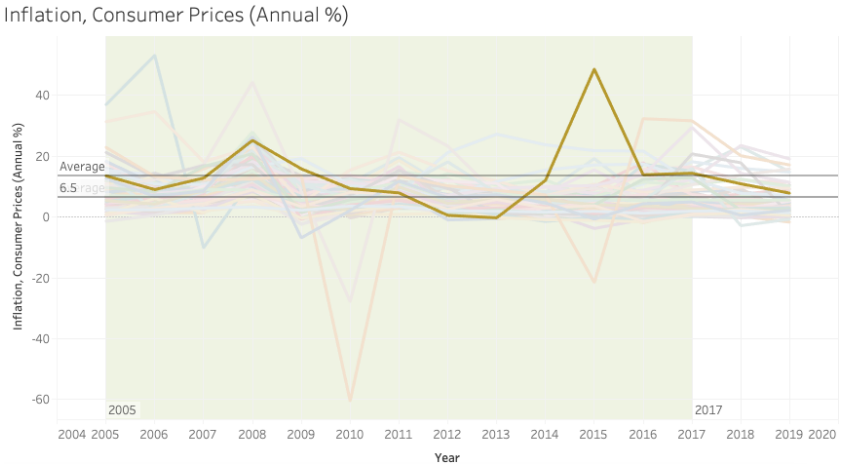
The average is 12.7% which is above average globally
The inflation is high from 2007 to 2013 and flatten around 6.5%(average globally) after 2013 that might be because the anti-corruption projects work.

The case of Angola:



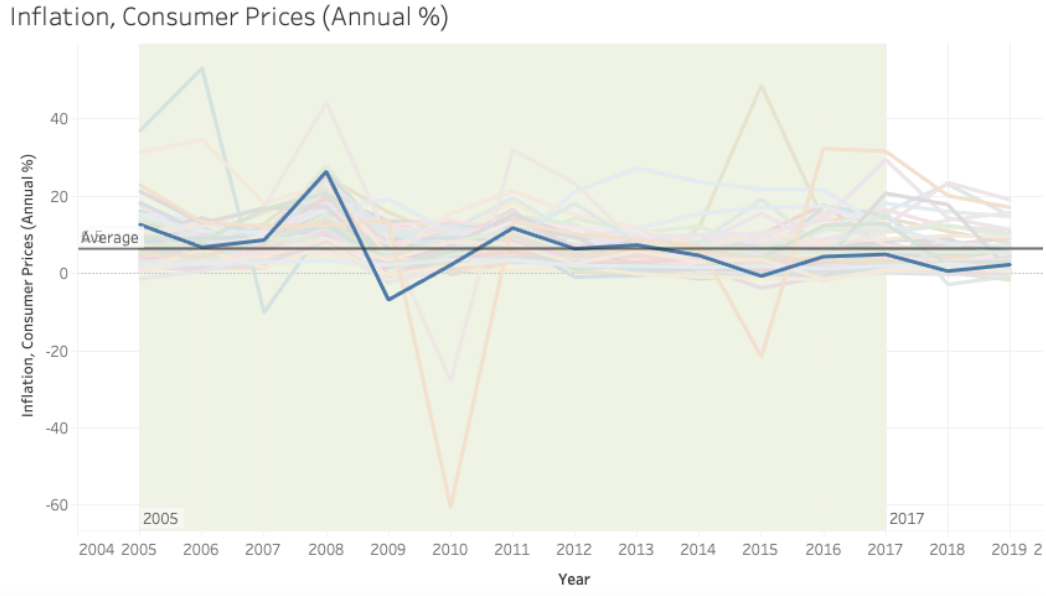
The average is 8.9% which is above average globally
The inflation dropped greatly in 2010 and flattened around 6.5%(average globally) from 2011 to 2014. Then went back to fluctuate greatly afterwards

The case of Ukraine:



The average is 13.5% which is above average globally
The inflation peaked in 2008(world recession crisis) and 2015(From 2014 to 2015, the Ukrainian economy suffered a downturn)

The case of Afghanistan:

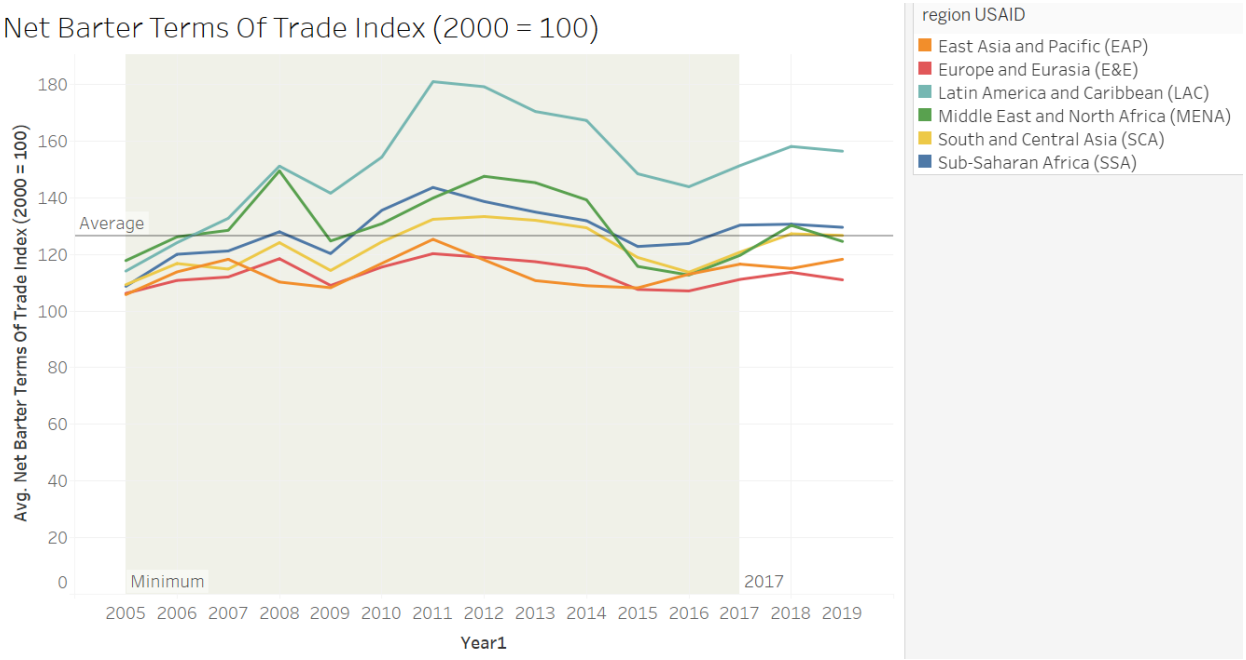


The average is 6.1% which is below average globally
The inflation is high due to the World Recession and flatten around 6.5%(average globally) afterwards.
Anti-corruption projects might help the economic recovery .

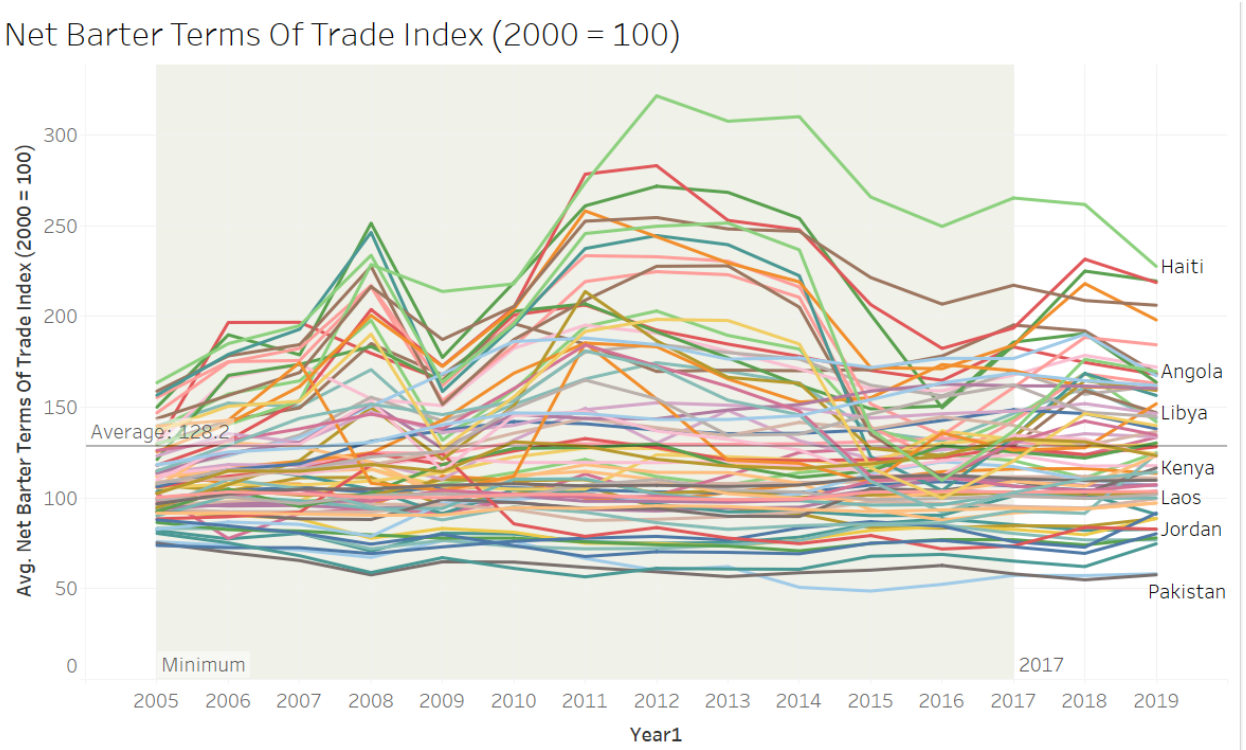
11. Net Barter Terms Of Trade Index (2000 = 100)

Net barter terms of trade index is calculated as the percentage ratio of the export unit value indexes to the import unit value indexes, measured relative to the base year 2000. Unit value indexes are based on data reported by countries that demonstrate consistency under UNCTAD quality controls, supplemented by UNCTAD's estimates using the previous year_s trade values at the Standard International Trade Classification three-digit level as weights. To improve data coverage, especially for the latest periods, UNCTAD constructs a set of average prices indexes at the three-digit product classification of the Standard International Trade Classification revision 3 using UNCTAD _s Commodity Price Statistics, interna _tional and national sources, and UNCTAD secretariat estimates and calculates unit value indexes at the country level using the current year's trade values as weights. Collecting and tabulating trade statistics are difficult. Net Barter Terms Of Trade Index (2000 = 100) has a negative correlation with Corruption Perception Index (CPI) at -0.13.

Globally

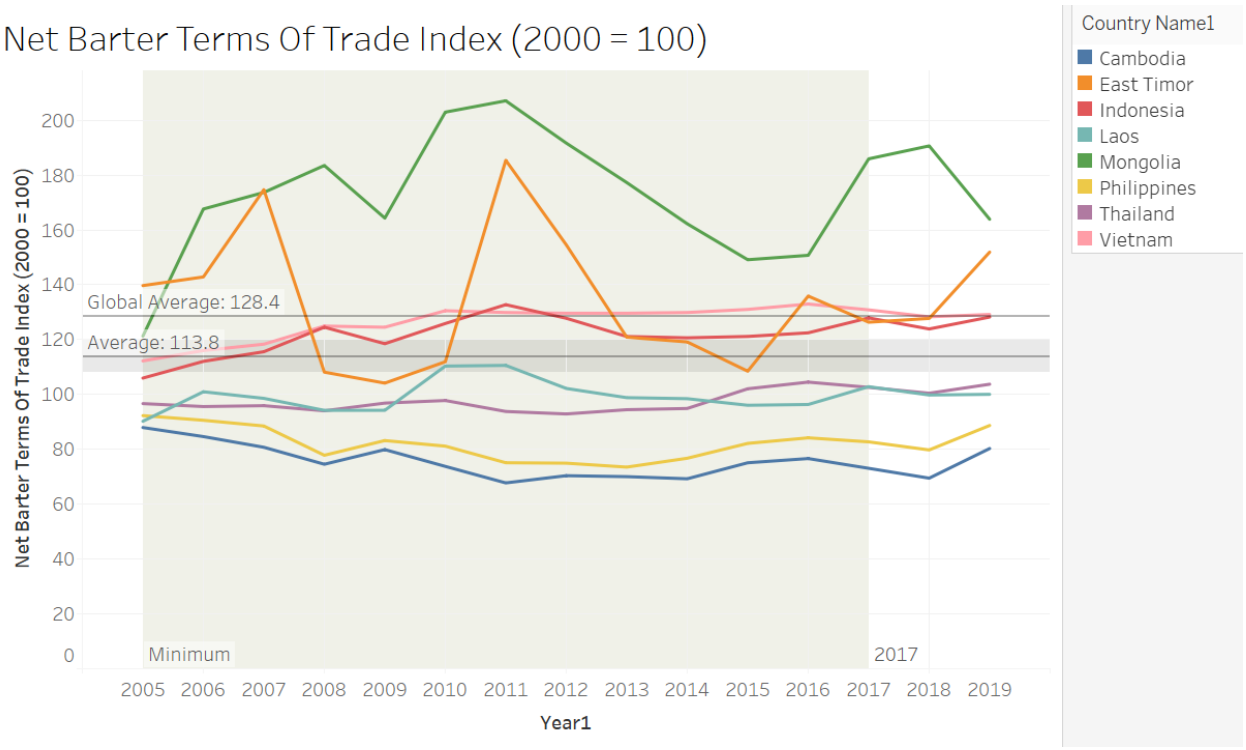


The region with the largest Net Barter Terms Of Trade Index (2000 = 100) is Latin America and Caribbean. All the regions show similar trends over the years, with a decrease from 2008 to 2009, an increase from 2009 to 2011, and a decrease from 2011 to 2015.



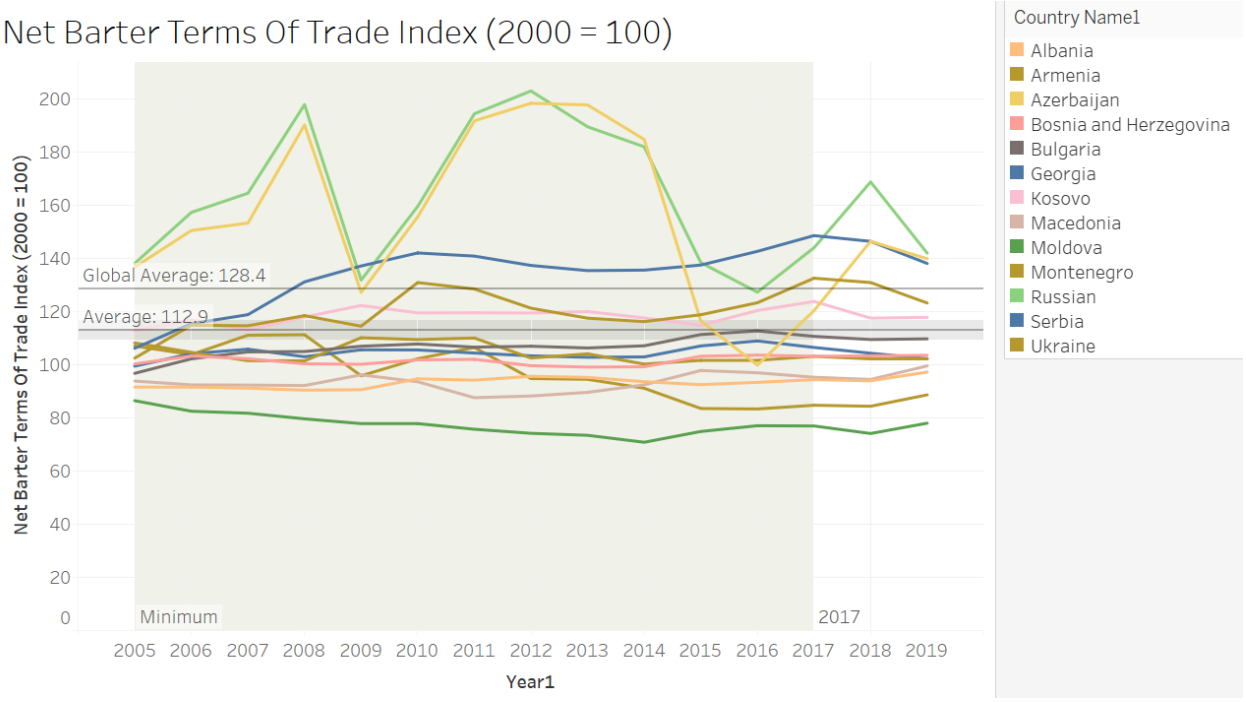
The Net Barter Terms Of Trade Index (2000 = 100) for the most countries are between 70 to 150.

a) East Asia and Pacific (EAP)



The regional average of 113.8 is slightly lower than the global average value, 128.4. The Net Barter Terms Of Trade Index for most of the countries are stable over the years (except East Timor and Mongolia). East Timor and Mongolia had large variations in Net Barter Terms Of Trade Index over the years.

b) Europe and Eurasia (E&E)

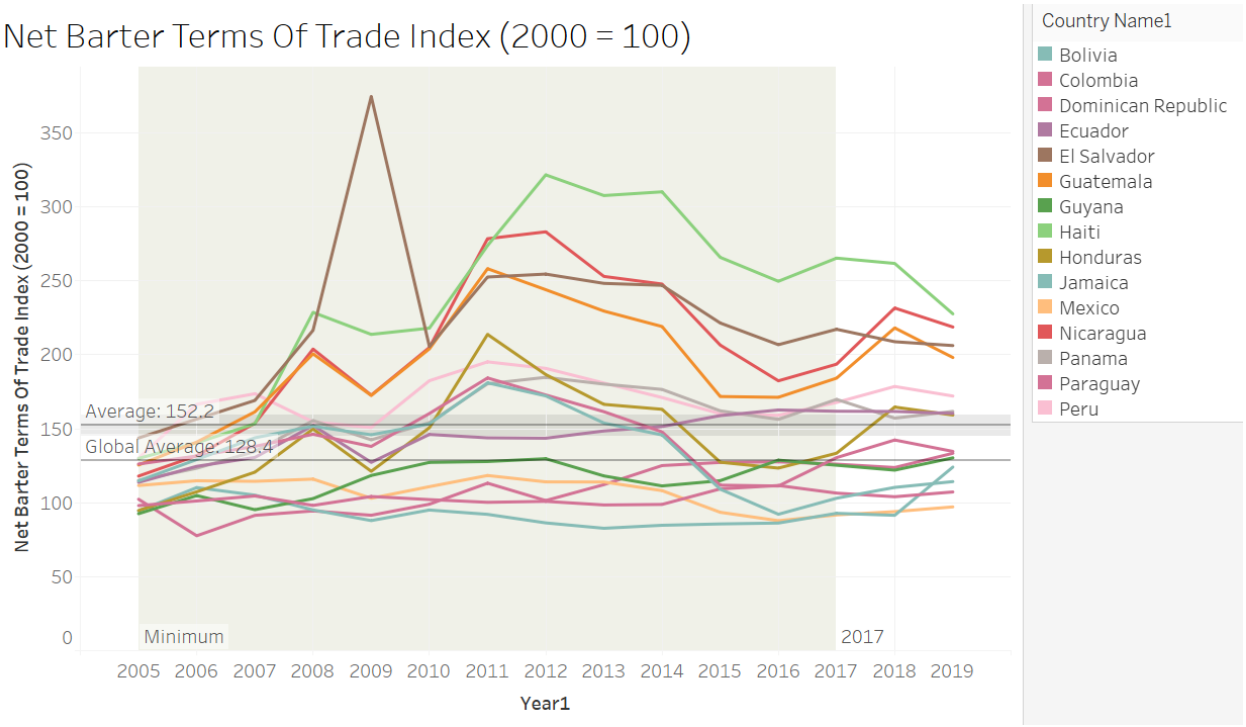


The regional average of 112.9 is slightly lower than the global average value, 128.4. The Net Barter Terms Of Trade Index for most of the countries are stable over the years (except Russian and Azerbaijan).

The outliers:

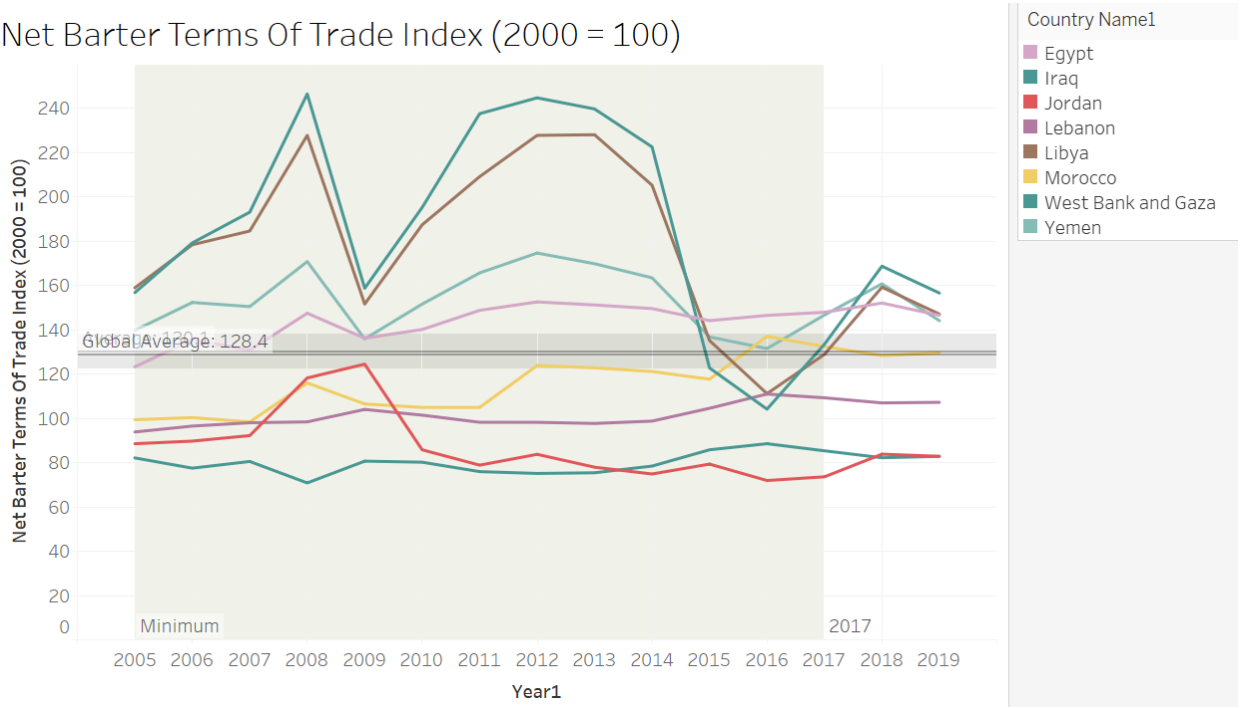
Russian and **Azerbaijan** had very similar trends over the years. They both had a sharp drop from 2008 to 2009, a sharp increase from 2009 to 2012, and another tremendous decrease from 2012 to 2016. It is supposed that the two countries had similar related policies, which result in the very similar trends.

c) Latin America and Caribbean (LAC)



The regional average of 152.2 is higher than the global average value, 128.4. Most of the countries show a consistent trend over the years, with an increase from 2007 to 2012, and a decrease from 2012 to 2016.

d) Middle East and North Africa (MENA)

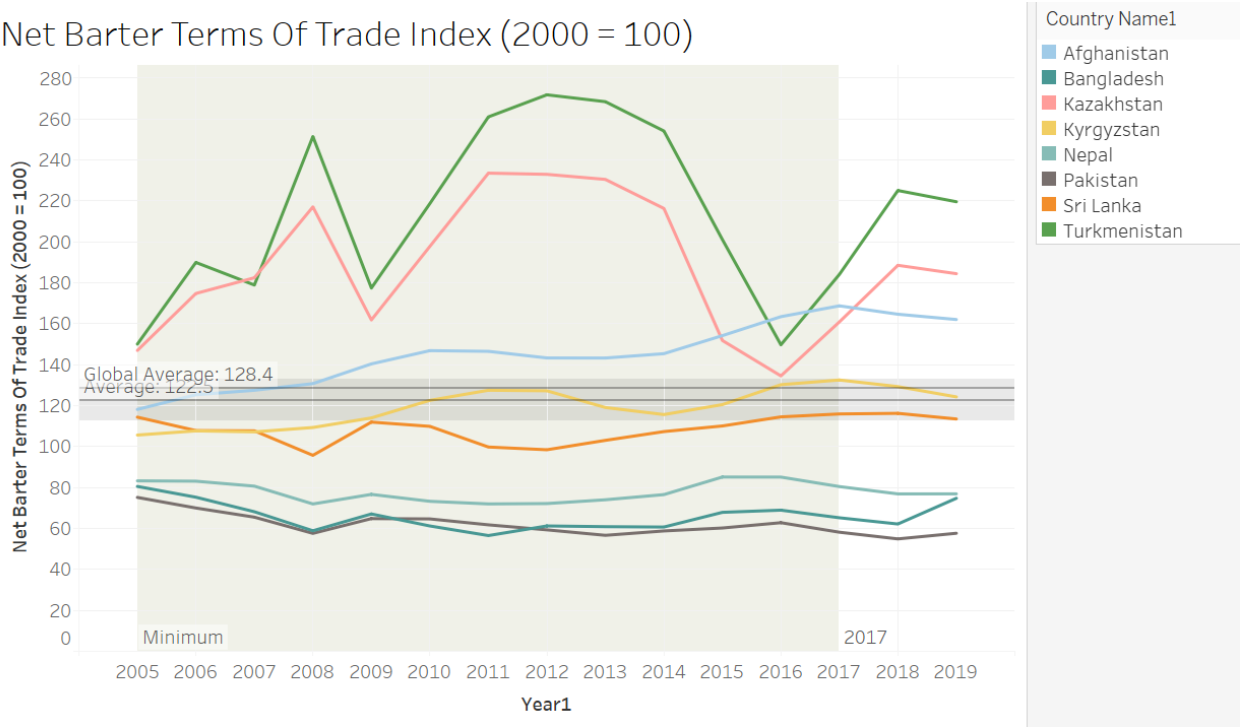


The regional average of 138.0 is higher than the global average value, 128.4. Most of the countries show a stable trend over the years.

The outliers:

Iraq and **Libya** had very similar trends over the years. They both had a sharp increase from 2006 to 2007 and from 2009 to 2012, a sharp increase from 2007 to 2008 and from 2013 to 2015.

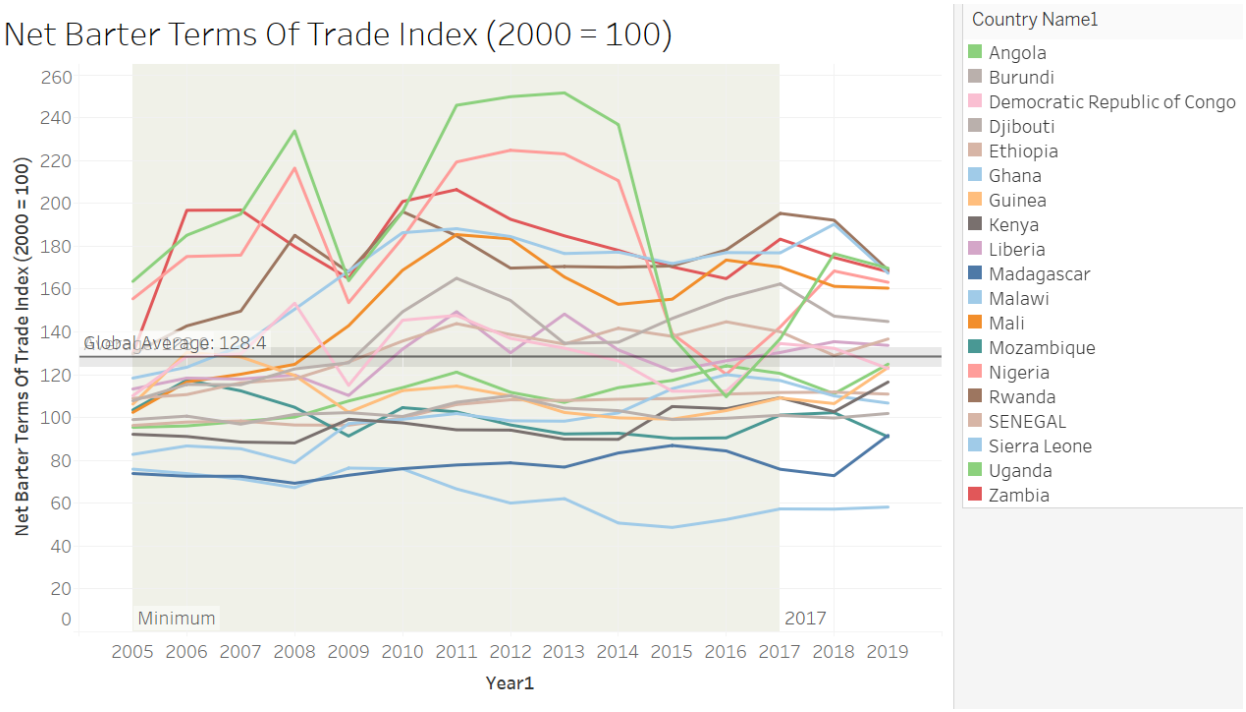
e) South and Central Asia (SAC)



The regional average of 122.5 is close to the global average value, 128.4. Most of the countries show a stable trend over the years (except Turkmenistan and Kazakhstan).

The outliers:
Turkmenistan and **Kazakhstan** had very similar trends over the years. They both had a sharp increase from 2007 to 2008 and from 2009 to 2012, a sharp increase from 2008 to 2009 and from 2013 to 2016.

f) Sub-Saharan Africa (SSA)

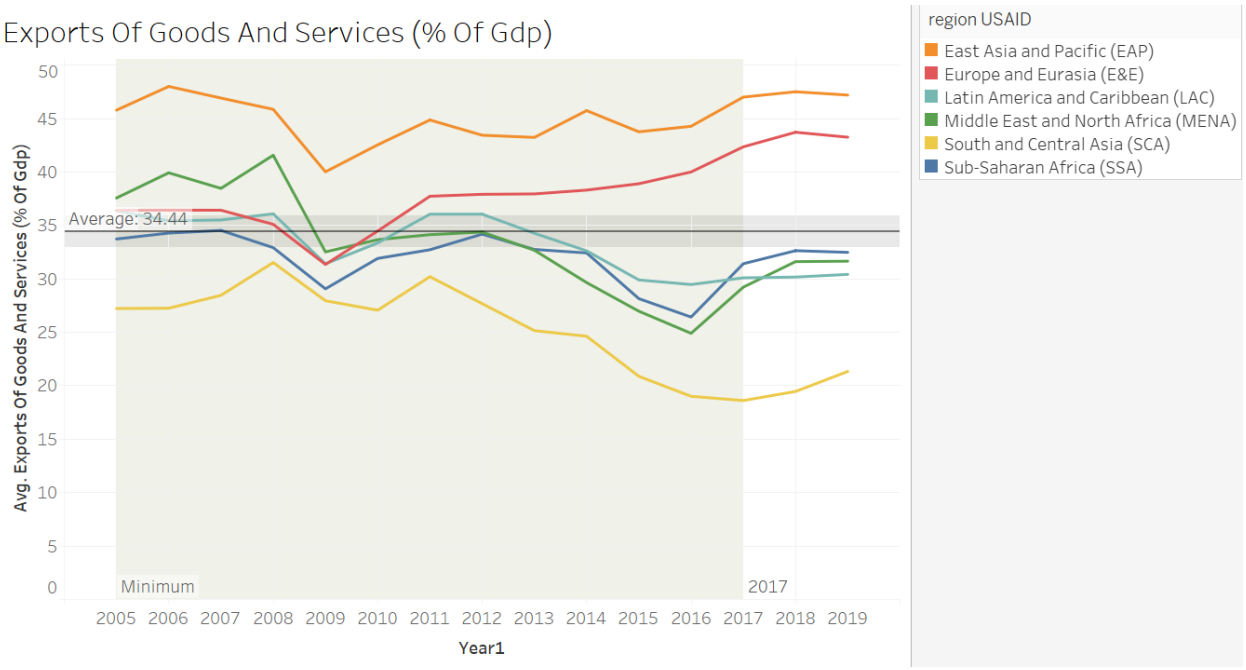


The regional average of 128.4 is close to the global average value. Both Angola and Nigeria experienced tremendous drop from 2014 to 2016.

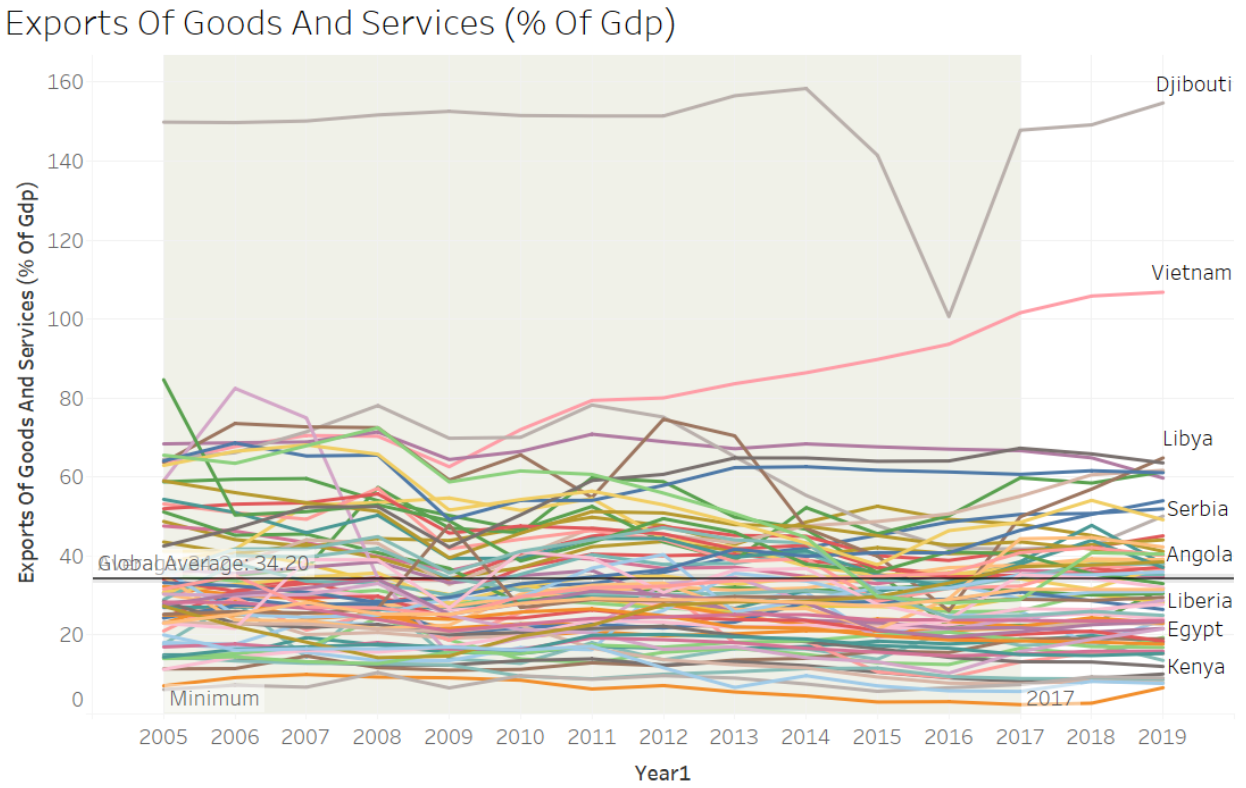
12. Exports Of Goods And Services (% Of Gdp)

Trade in goods and services is defined as the transactions in goods and services between residents and non-residents. It is measured in million USD, as a percentage of GDP for net trade, and also in annual growth for exports and imports. All OECD countries compile their data according to the 2008 System of National Accounts (SNA). Exports Of Goods And Services (% Of Gdp) has a slight positive correlation with Corruption Perception Index (CPI) at 0.072.

Globally

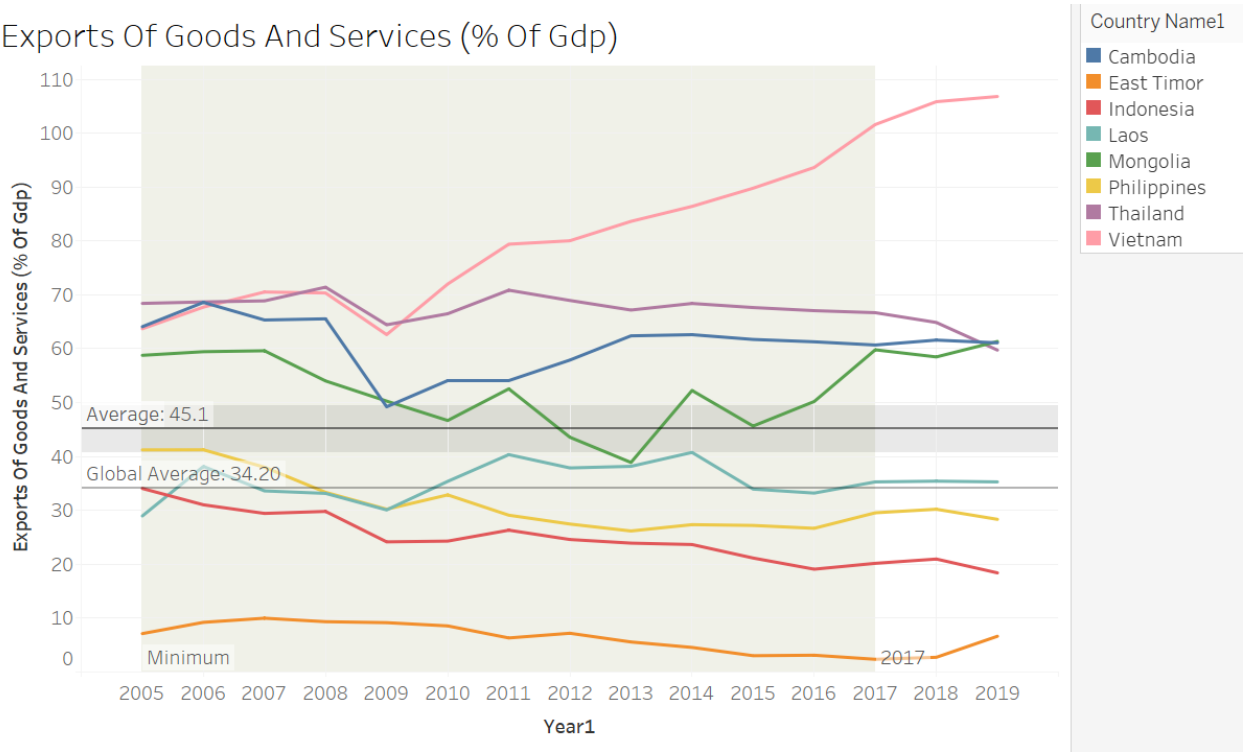


The region-level global average of Exports Of Goods And Services (% Of Gdp) is 34.44. The region with the largest Exports Of Goods And Services (% Of Gdp) is the East Asia and Pacific region. And South and Central Asia has the lowest value. All the regions show a decrease from 2008 to 2009.



The Exports Of Goods And Services (% Of Gdp) for most countries are between 0 to 50%. The country with the largest Exports Of Goods And Services (% Of Gdp) is Djibouti.

a) East Asia and Pacific (EAP)

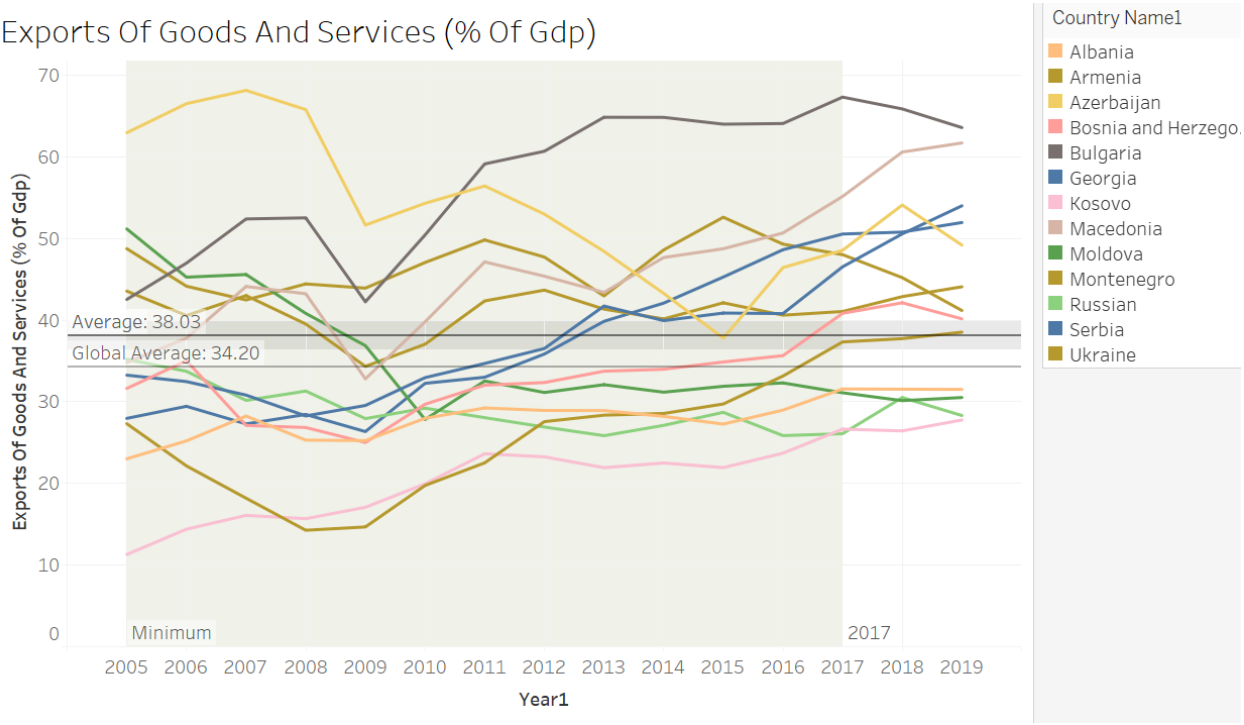


The regional average of 45.1 is larger than the global average value, 34.20. The Exports Of Goods And Services (% Of Gdp) for most of the countries are stable over the years. East Timor has very low Exports Of Goods And Services (% Of Gdp) values of around 5% over the years.

The outliers:

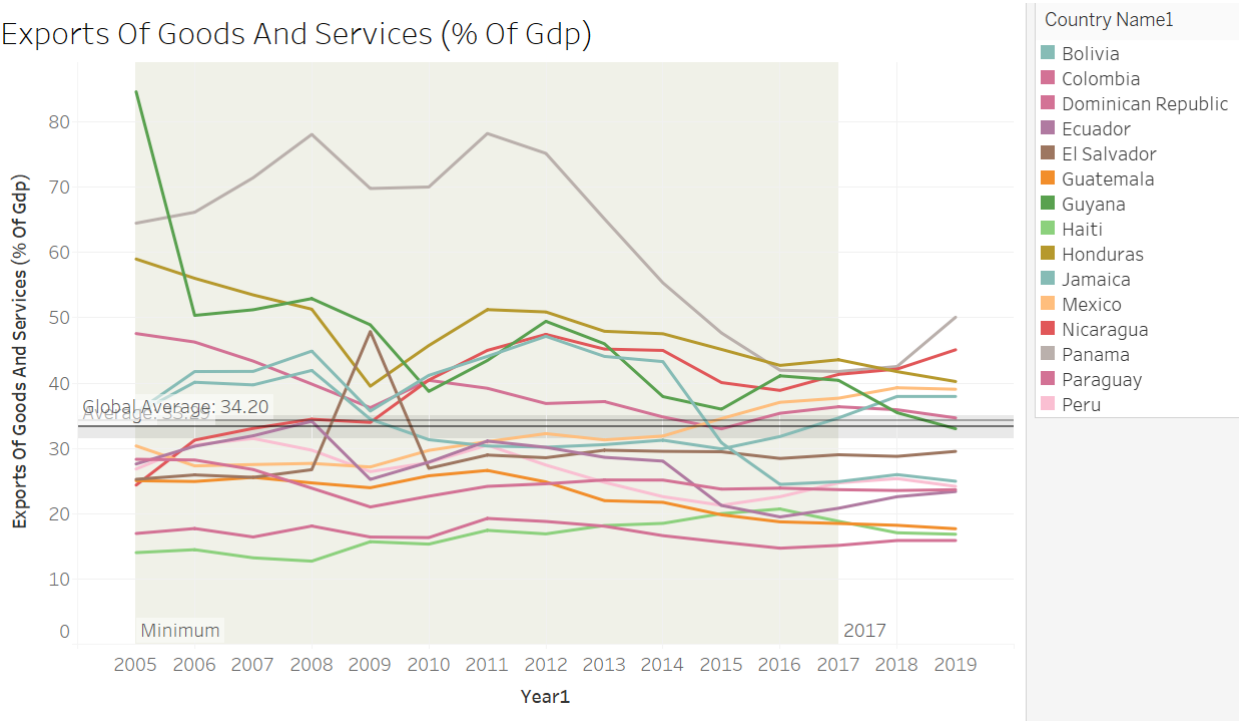
Vietnam experienced a continuous increase from 2009 to 2019, with the peak of 117%, which could indicate the effectiveness of USAID projects.

b) Europe and Eurasia (E&E)



The regional average of 38.03 is slightly larger than the global average value, 34.20. Almost all the countries show a consistent trend over the years, with a steady increase from 2009 to 2017.

c) Latin America and Caribbean (LAC)

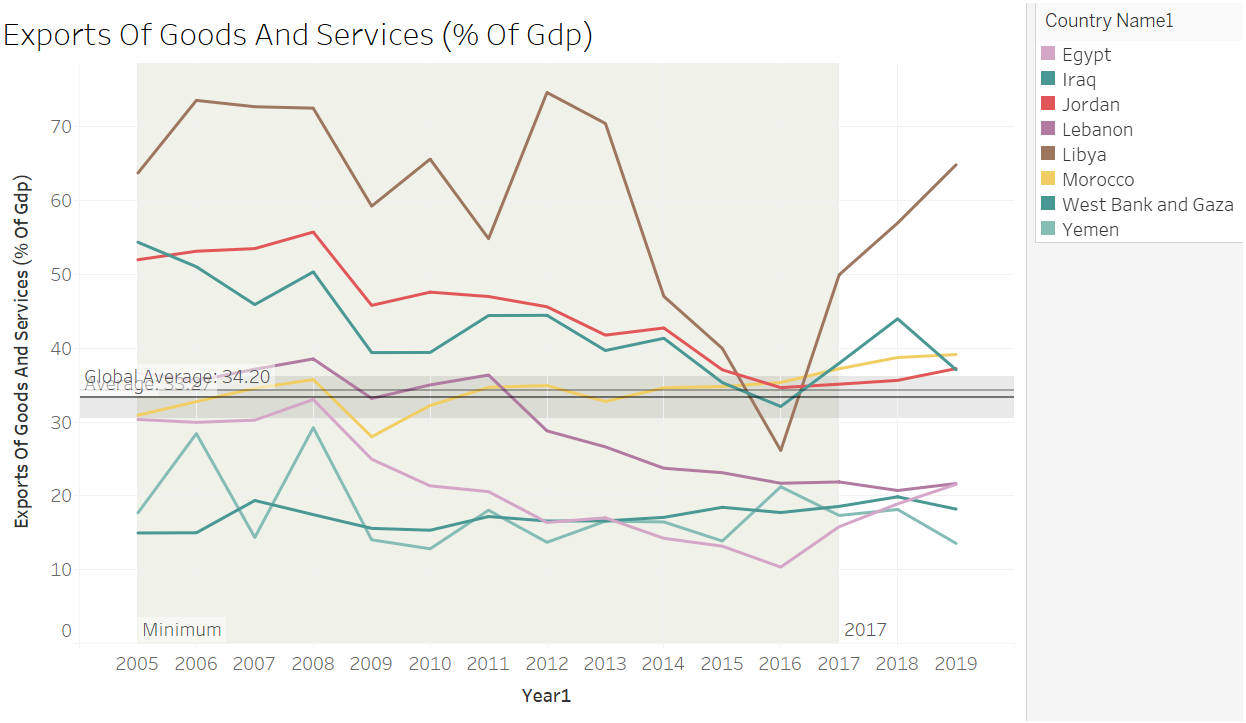


The regional average of 33.29 is close to the global average value, 34.20. The Exports Of Goods And Services (% Of Gdp) for most of the countries are stable over the years.

The outliers:

Panama had a larger Exports Of Goods And Services (% Of Gdp) from 2005 to 2012, with a peak of 75%, which could indicate the effectiveness of USAID projects.

d) Middle East and North Africa (MENA)

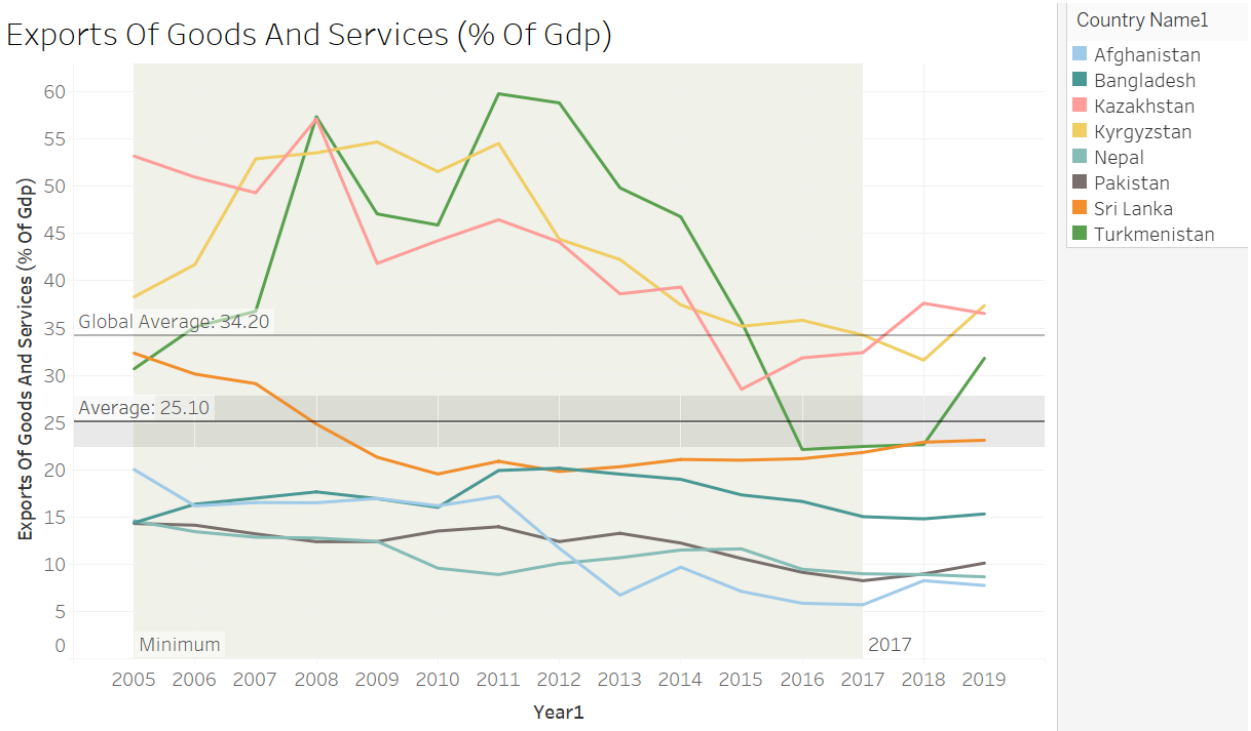


The regional average of 33.27 is close to the global average value, 34.20. Libya, Jordan, and Lebanon have large variations over the years for The Exports Of Goods And Services (% Of Gdp).

The outliers:

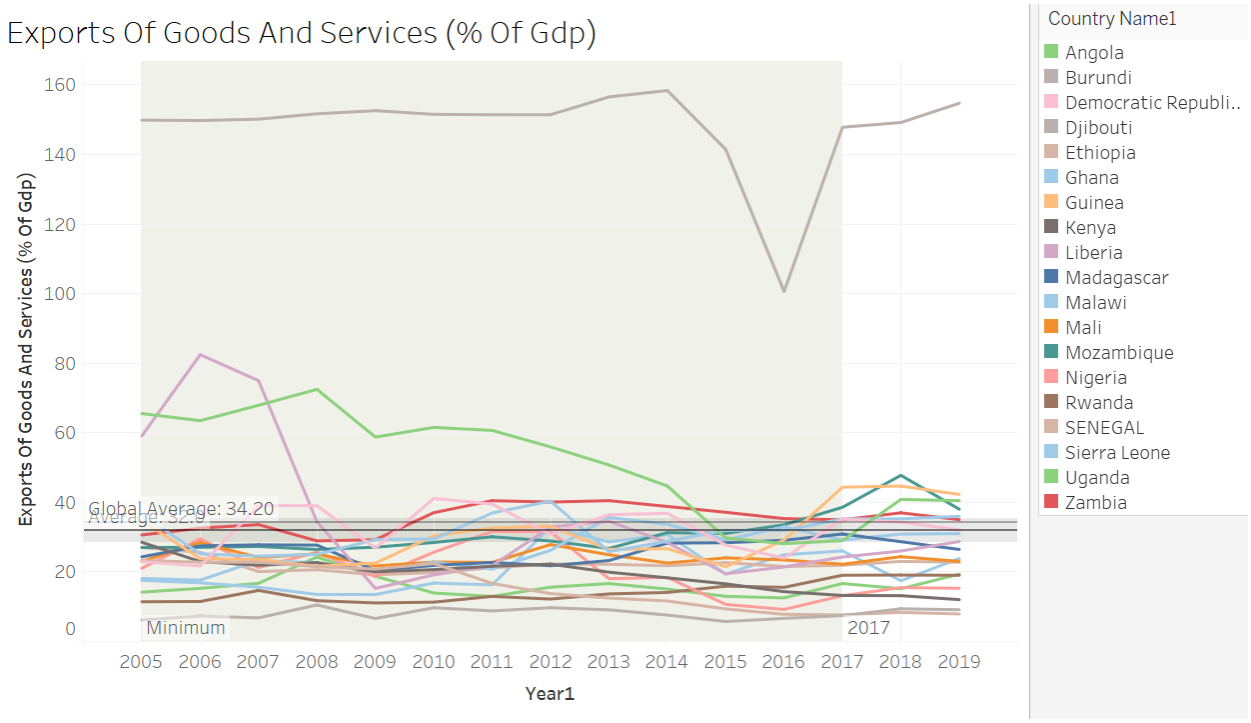
Libya had a tremendous decrease from 2012 to 2016, dropping from 70%+ to 20%, and then had a tremendous increase from 2016 to 2017.

e) South and Central Asia (SAC)



The regional average of 25.10 is lower than the global average value, 34.20. The Exports Of Goods And Services (% Of Gdp) for Turkmenistan, Kazakhstan, and Kyrgyzstan all had an increase from 2005 to 2011 and a decrease from 2011 to 2016.

f) Sub-Saharan Africa (SSA)



The regional average of 32.0 is close to the global average value, 34.20. The Exports Of Goods And Services (% Of Gdp) for most of the countries are stable over the years (except Djibouti, Angola, and Liberia).

The outliers:

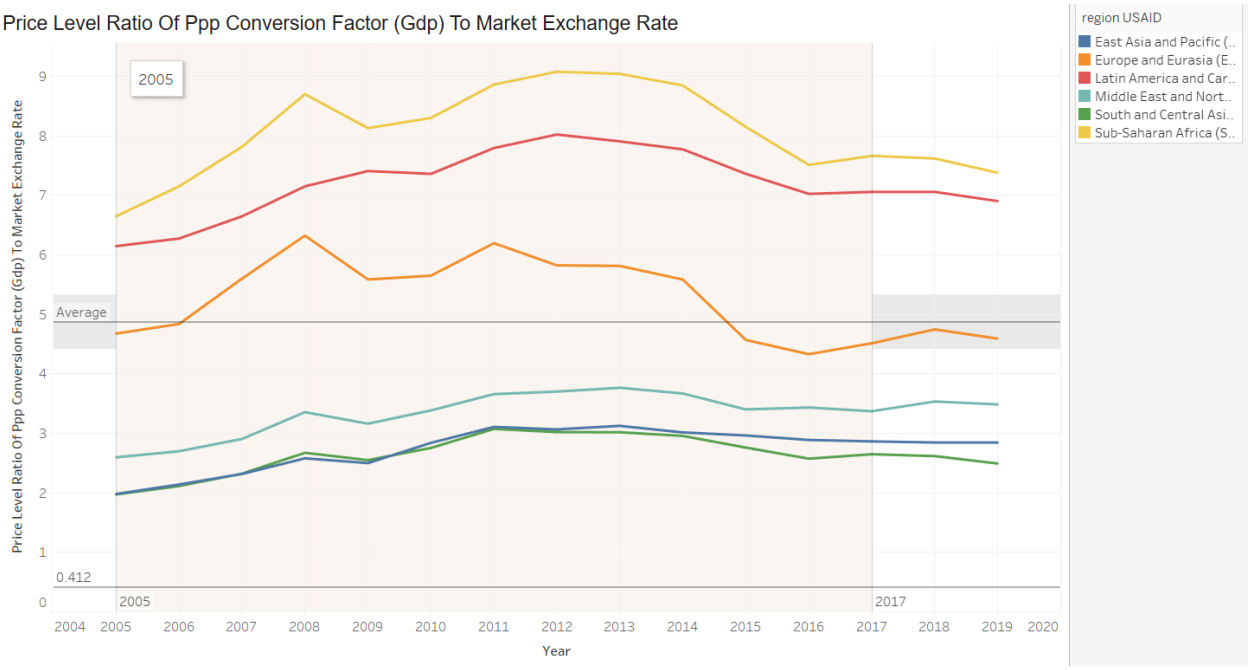
Djibouti had a much larger Exports Of Goods And Services (% Of Gdp) then the other countries with the values of about 150% for the most years. It had a decrease from 2014 to 2016, and an increase from 2016 to 2017.

13. Price Level Ratio Of Ppp Conversion Factor (Gdp) To Market Exchange Rate

Price level ratio is the ratio of a purchasing power parity (PPP) conversion factor to an exchange rate. It provides a measure of the differences in price levels between countries by indicating the number of units of the common currency needed to buy the same volume of the aggregation level in each country. So the higher price level ratio, the better for country's development. And the correlation between cpi and price level ratio is positive but low,,which is only 0.09.So maybe price level ratio indicator can not affect cpi much better.

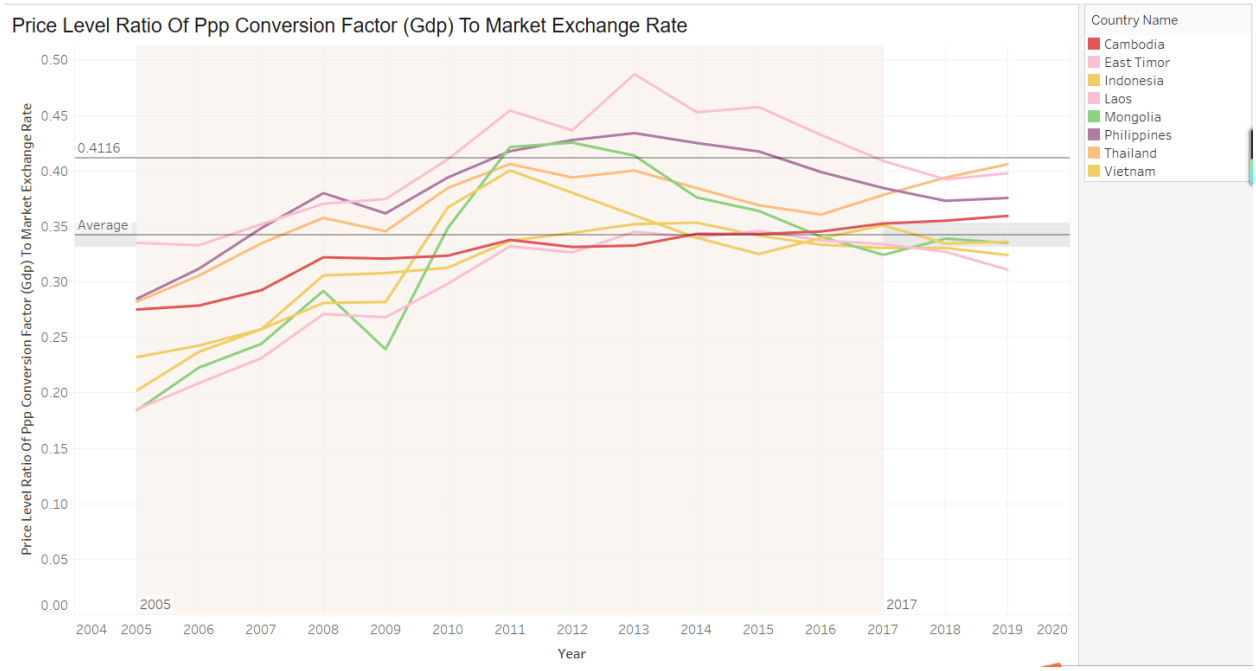
Global average (with 95% CI) is 0.4116 over the period of 2005 to 2019.

Regional graph:



Each region has a very slow growth, but in the end it is not much different from the **Price Level Ratio** value of the first year.They all have a small decline between 2008-2009.SSA maintains the highest **Price Level Ratio** value in the world every year.

a) East Asia and Pacific (EAP)

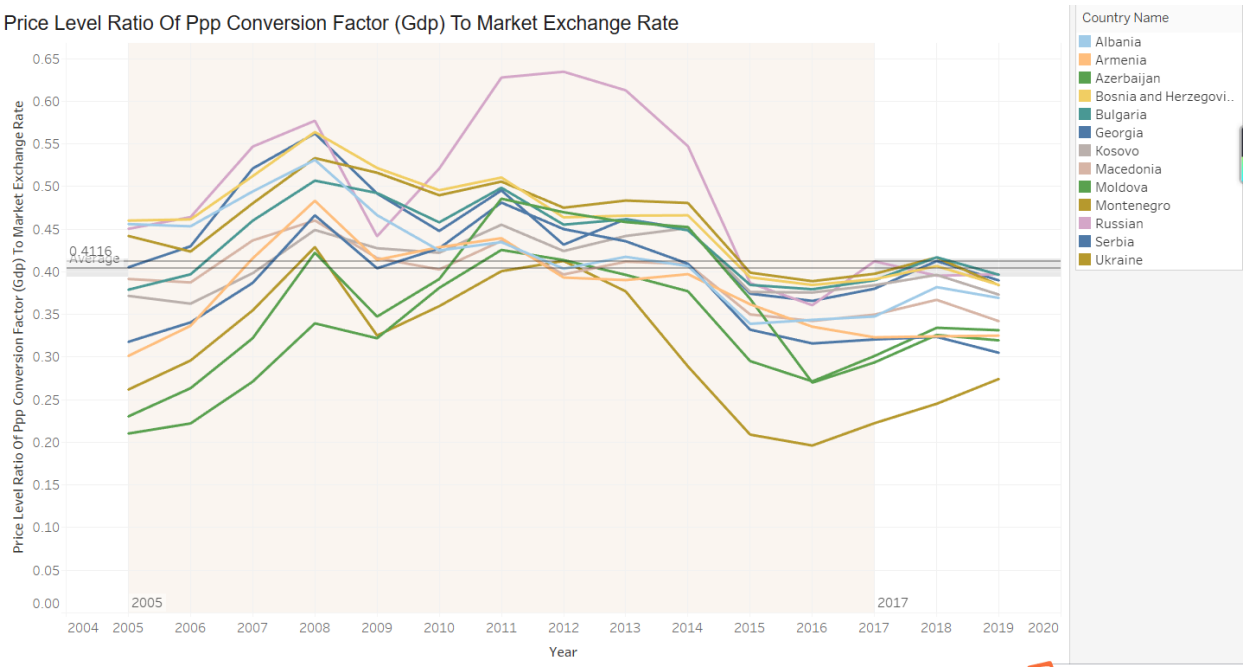


We can see from the graph for EAP region that the average of the region is higher than the global average. Almost all countries in this region show an upward trend before 201。 Half of the countries (such as East Timor, Philippines, Mongolia) fell rapidly after 2011, and almost all countries flattened after 2016. The price level ratio did not rise or fall sharply but fluctuated around the average after2016. All countries showed less fluctuation in price level ratio growth as time went by in 2009.Even the index of Mongolia reached the lowest value of all years in the world in 2009.The reason may be related to the Ukraine and the decrease of countries' GDP.

The outliers:

Laos generally maintains a similar change trend with other countries in the EAP region, but its price level ratio is much higher than other countries in most years

b) Europe and Eurasia (E&E)



The E&E regional average is almost the same as the global average.

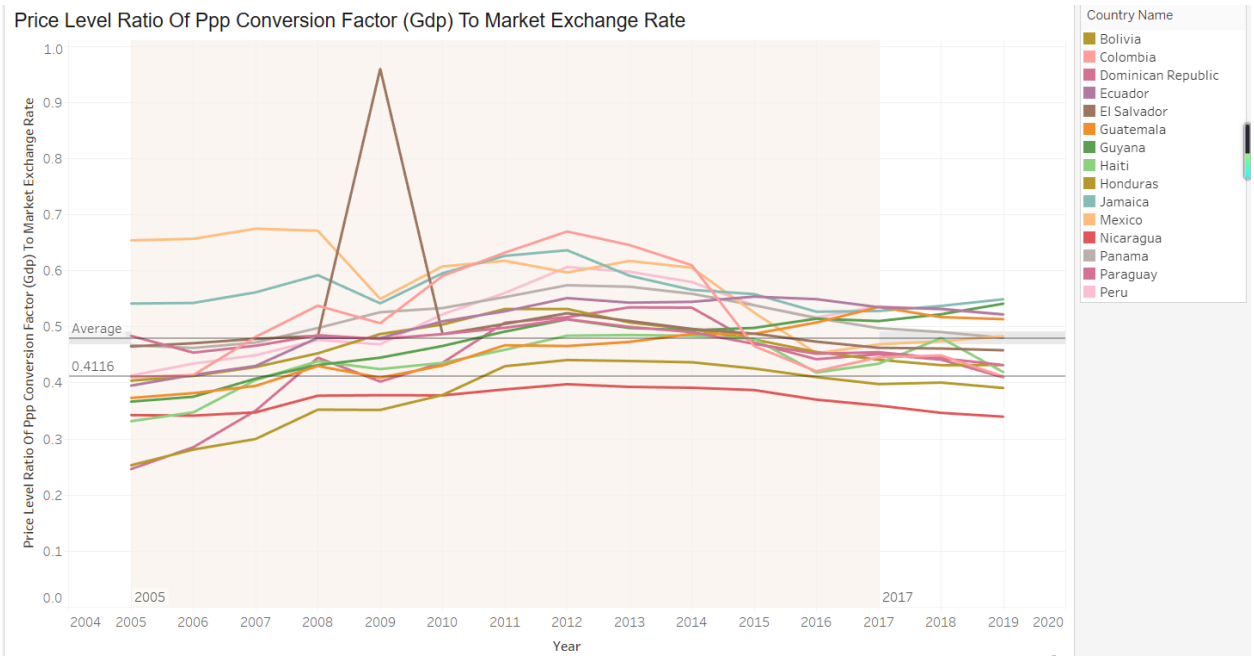
This price level ratio indicator in E&E region has very interesting changes. The trend of each country is almost the same, and the continuous collective rise and fall, which perfectly reflects the changes in the global economy from 2005 to 2017.

All countries in this region rose steadily before 2008, and dropped sharply to the lowest point during 2008-2009 (Ukraine and Moldovia reached a global trough in this year). It recovered in the two years from 2009 to 2011, but fell again after 2012.Finally the price level ratio fell in 2016,rise and stabilize year after year.

The outliers:

In these years ,Compared with other countries, **Russian** has changed the most in the past few years, especially the price level ratio from 2010 to 2015 far exceeded that of other countries.

c) Latin America and Caribean (LAC)

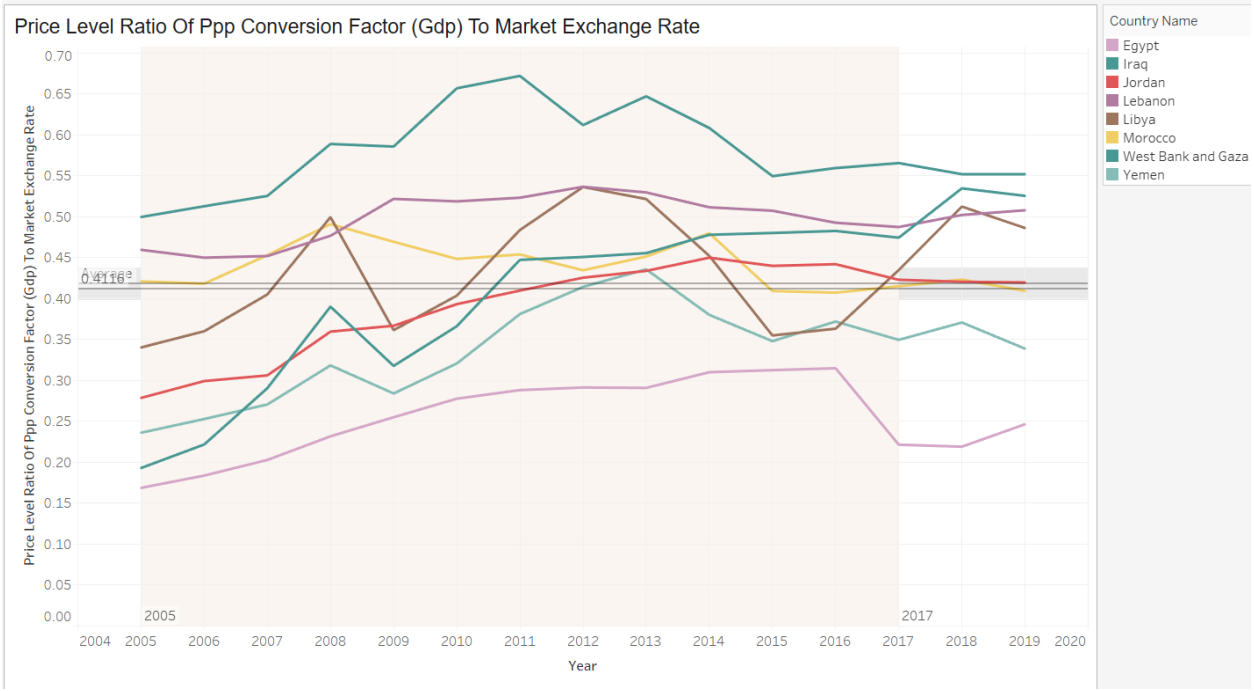


The LAC regional average is higher than the global average. All countries showed less fluctuation in price level ratio growth as time went by,almost unaffected by the world economy and the Great Depression.Only a small decrease in index between 2008-2009.

The outliers:

Different from the small changes in other countries,the price level ratio of **El Salvator** fell sharply and rose sharply from 2008 to 2010. It was greatly affected by the world economy, reflecting the rapid depression and rapid economic recovery.

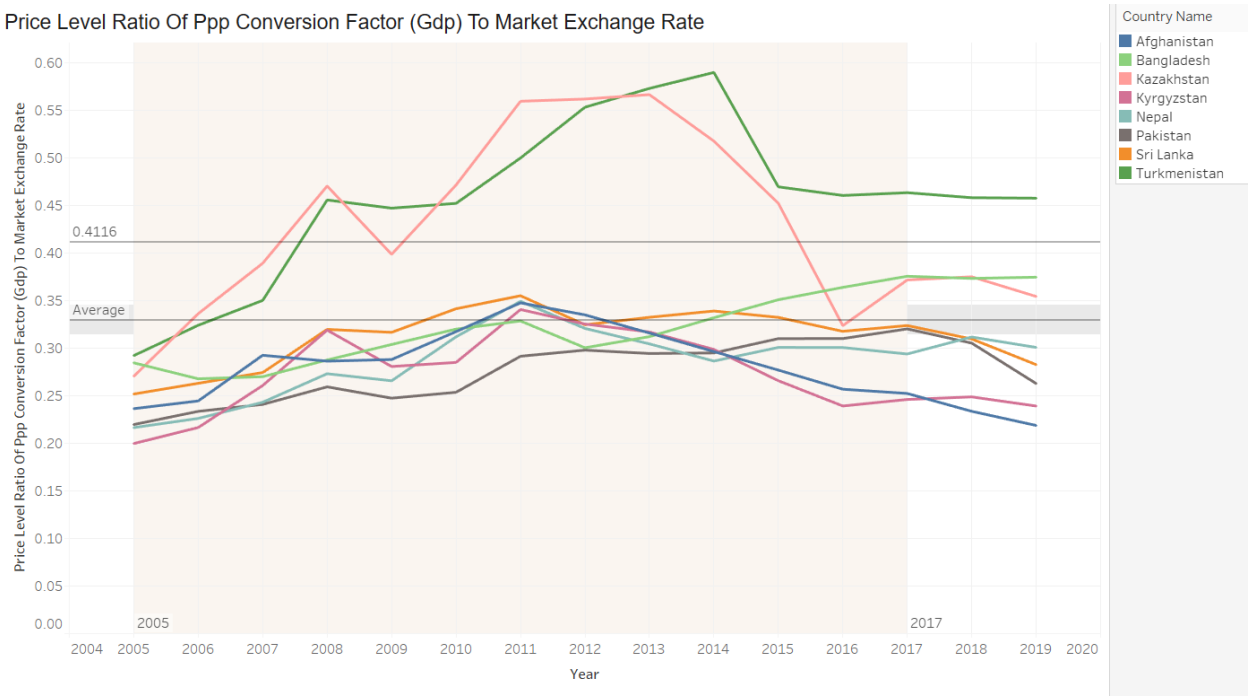
d) Middle East and North Africa (MENA)



The MENA regional average is almost the same as the global average.

The year-on-year change trends in various countries are relatively scattered and not consistent. And the price level ratio value is not concentrated near the average value like other regions.But the overall change is not big.Price level ratios of these countries are either far below the average value each year(like Egypt,Yemen), or higher than the average value(like Lenbanon,West bank and Gaza).

e) South and Central Asia (SAC)



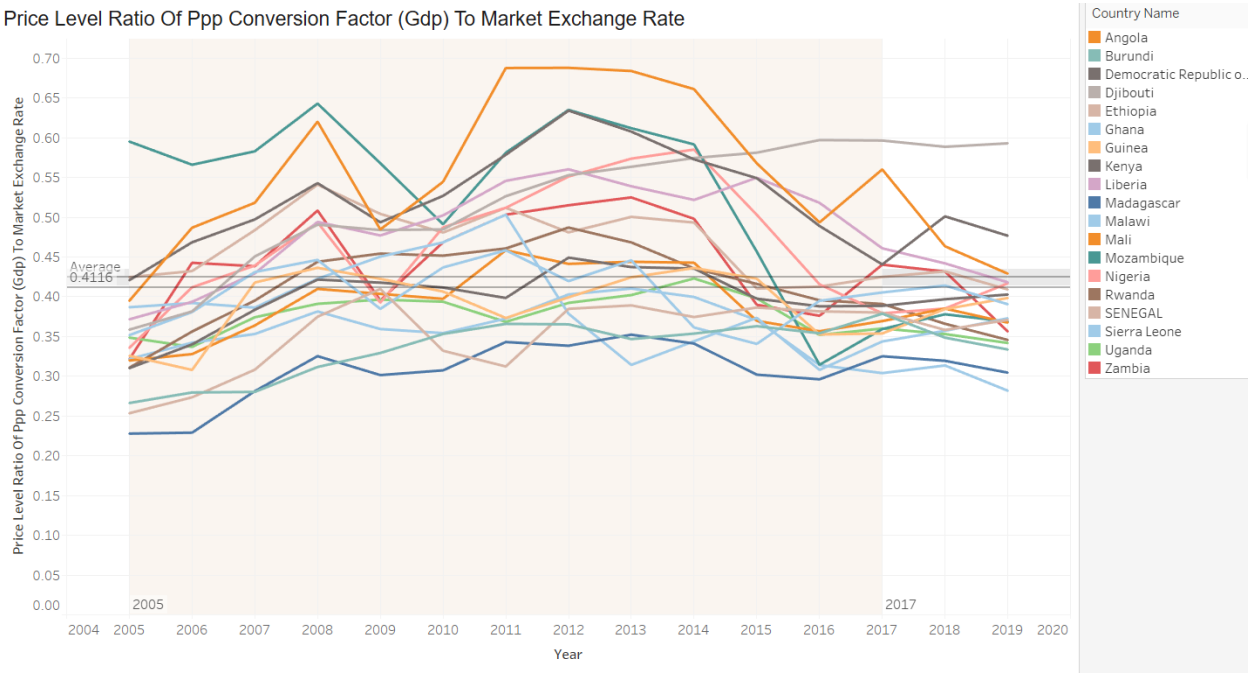
The SAC regional average is lower than the global average.

All countries ,except Kazakhstan and Turkmenistan,showed less fluctuation in price level ratio growth as time went by.

The outliers:

The price level ratios of Kazakhstan and Turkmenistan are much higher than other countries and the average level.And the magnitude of the changes in these two countries is much greater than in other countries, and reached a peak between 2010 and 2014.

f) Sub-Saharan Africa (SSA)



The SSA regional average is almost the same as the global average.
The changes in various countries are relatively disorderly and different, but the overall performance first rises, then declines, then rises and then falls, and finally tends to be flat, similar to the global change trend.

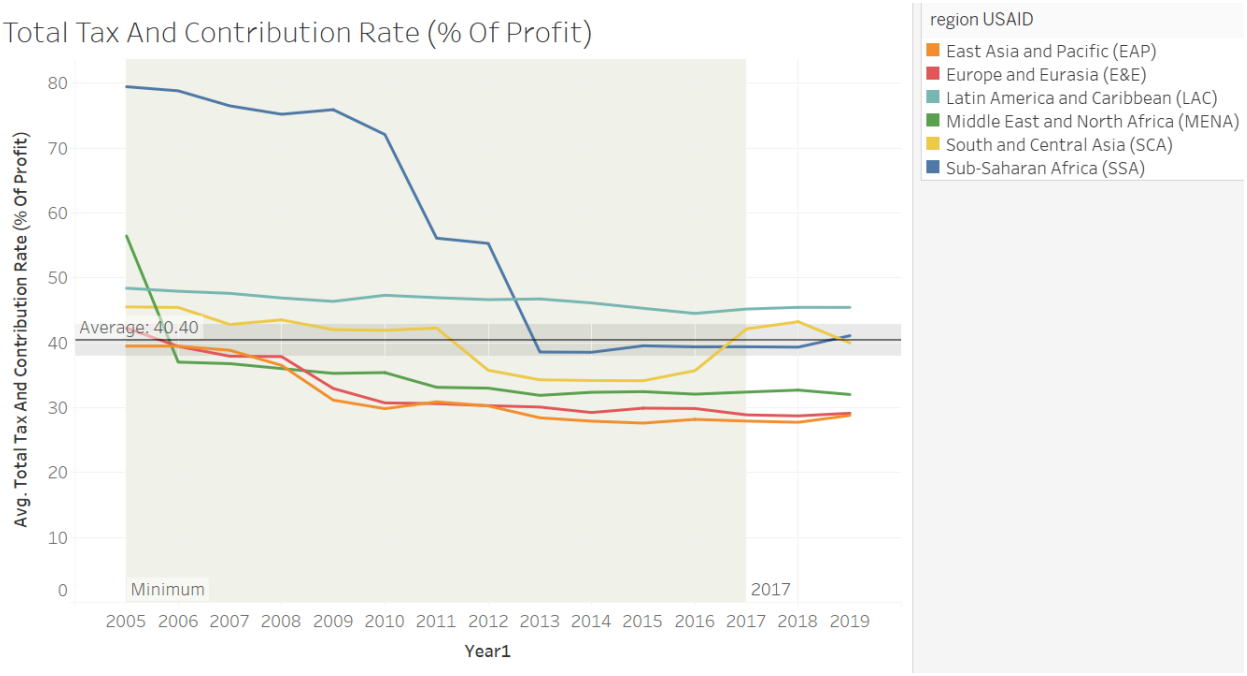
The outliers:

`Djibouti is steadily increasing almost every year, and reached the highest value in the world in 2017.`

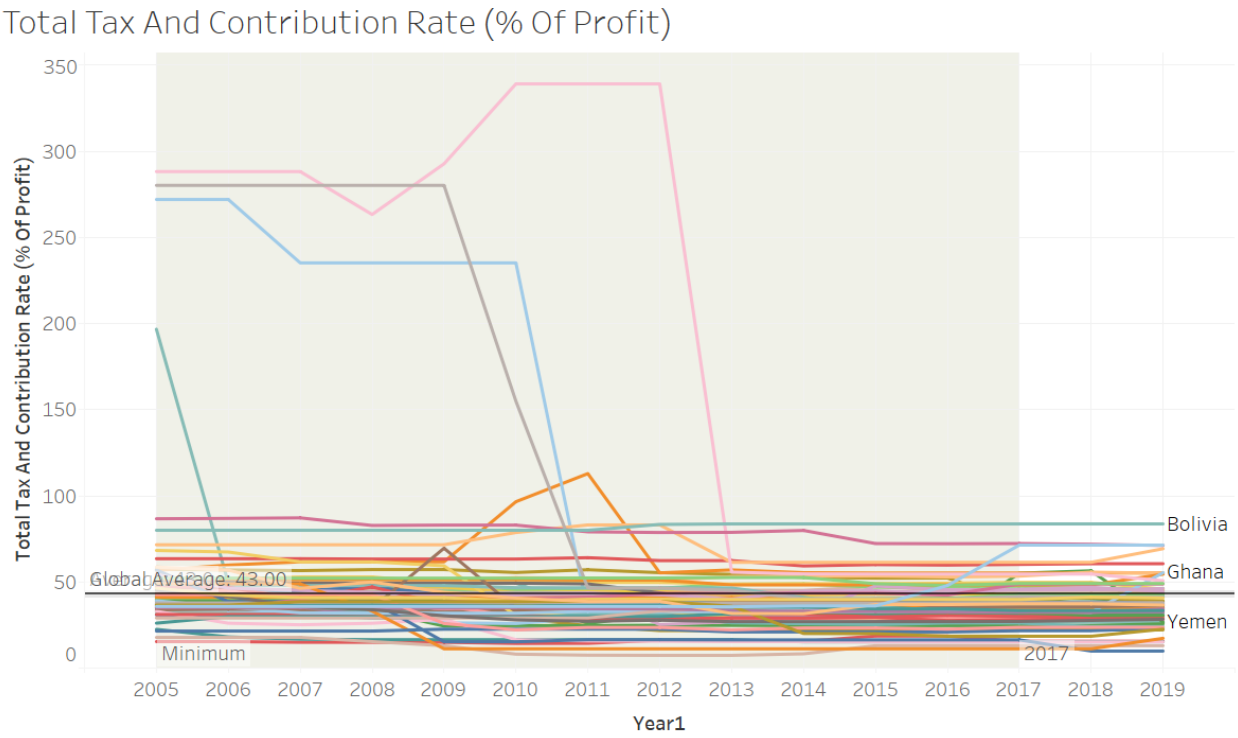
14. Total Tax And Contribution Rate (% Of Profit)

Total tax rate measures the amount of taxes and mandatory contributions payable by businesses after accounting for allowable deductions and exemptions as a share of commercial profits. Taxes withheld (such as personal income tax) or collected and remitted to tax authorities (such as value added taxes, sales taxes or goods and service taxes) are excluded. Total Tax And Contribution Rate (% Of Profit) has a negative correlation with Corruption Perception Index (CPI) at -0.20.

Global

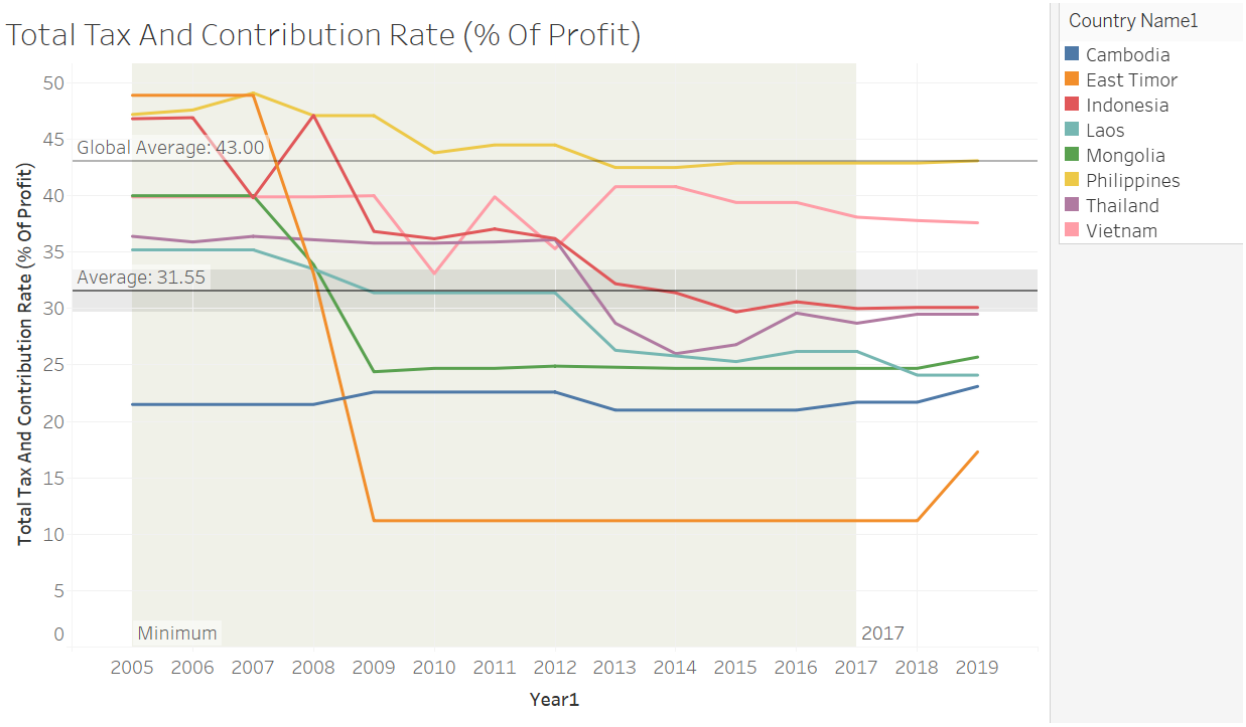


The region-level global average of Total Tax And Contribution Rate (% Of Profit) is 40.40. Sub-Saharan Africa has a much higher Bank Liquid Reserves To Bank Assets Ratio (%) than the other regions from 2005 to 2012.



The Total Tax And Contribution Rate (% Of Profit) for most of the countries are between 0 to 70%.

a) East Asia and Pacific (EAP)

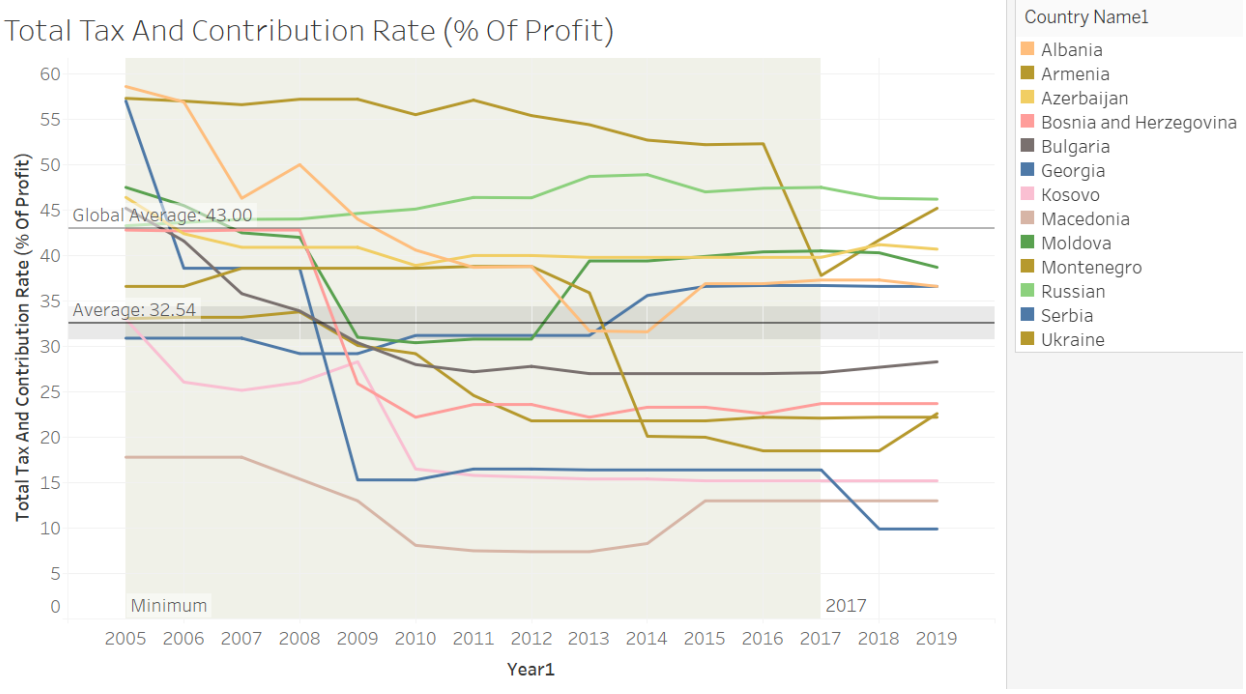


The regional average of 31.55 is larger than the global average value, 43.00. The Total Tax And Contribution Rate (% Of Profit) for most of the countries are stable over the years (except East Timor). But the differences among these countries are very obvious.

The outliers:

East Timor experienced a tremendous decrease from 2007 to 2009, dropping from about 50% to 12%, which could indicate the effectiveness of USAID projects.

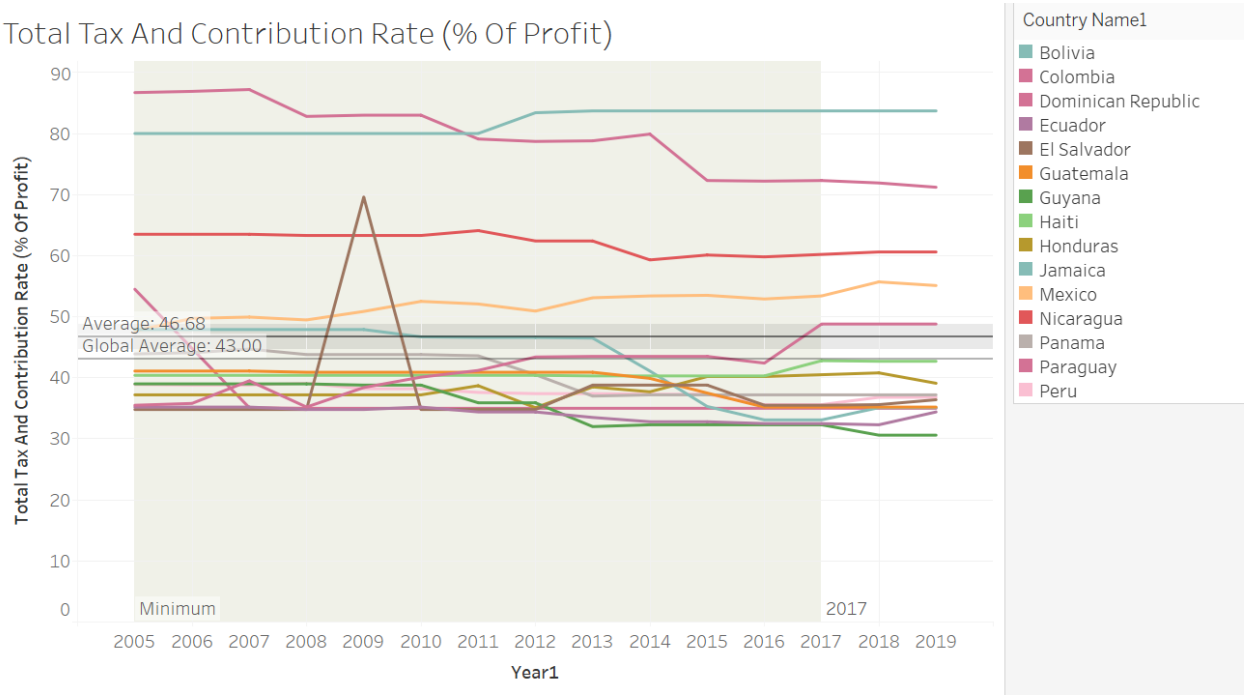
b) Europe and Eurasia (E&E)



The regional average of 32.54 is lower than the global average value, 43.00. The Total Tax And Contribution Rate (% Of Profit) for most of the countries are stable over the years (except Georgia). But the differences among these countries are very obvious.

The outliers:
Georgia experienced a tremendous decrease from 2005 to 2009, dropping from about 55% to 17%, which could indicate the effectiveness of USAID projects.

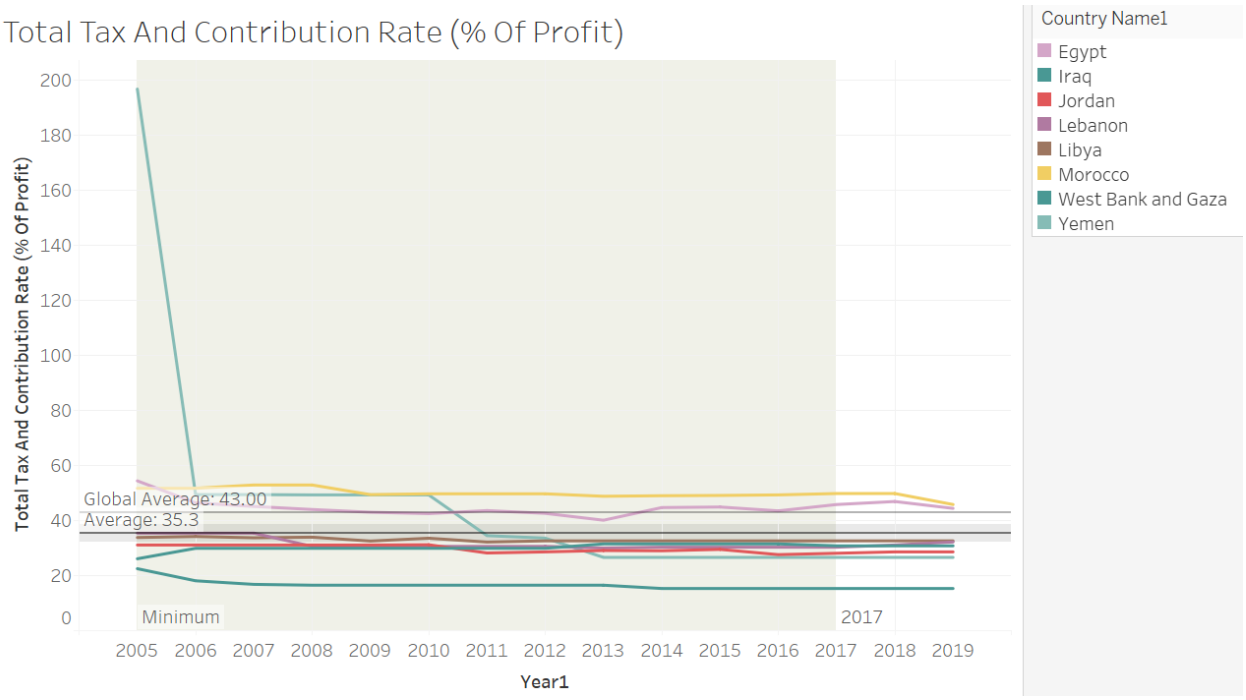
c) Latin America and Caribbean (LAC)



The regional average of 46.68 is close to the global average value, 43.00. The Total Tax And Contribution Rate (% Of Profit) for most of the countries are very stable over the years (except EI Salvador). But the differences among these countries are very obvious.

The outliers:
EI Salvador experienced a tremendous increase from 2008 to 2009 and a tremendous decrease from 2009 to 2010, with the peak of about 70%, which could indicate the effectiveness of USAID projects.

d) Middle East and North Africa (MENA)

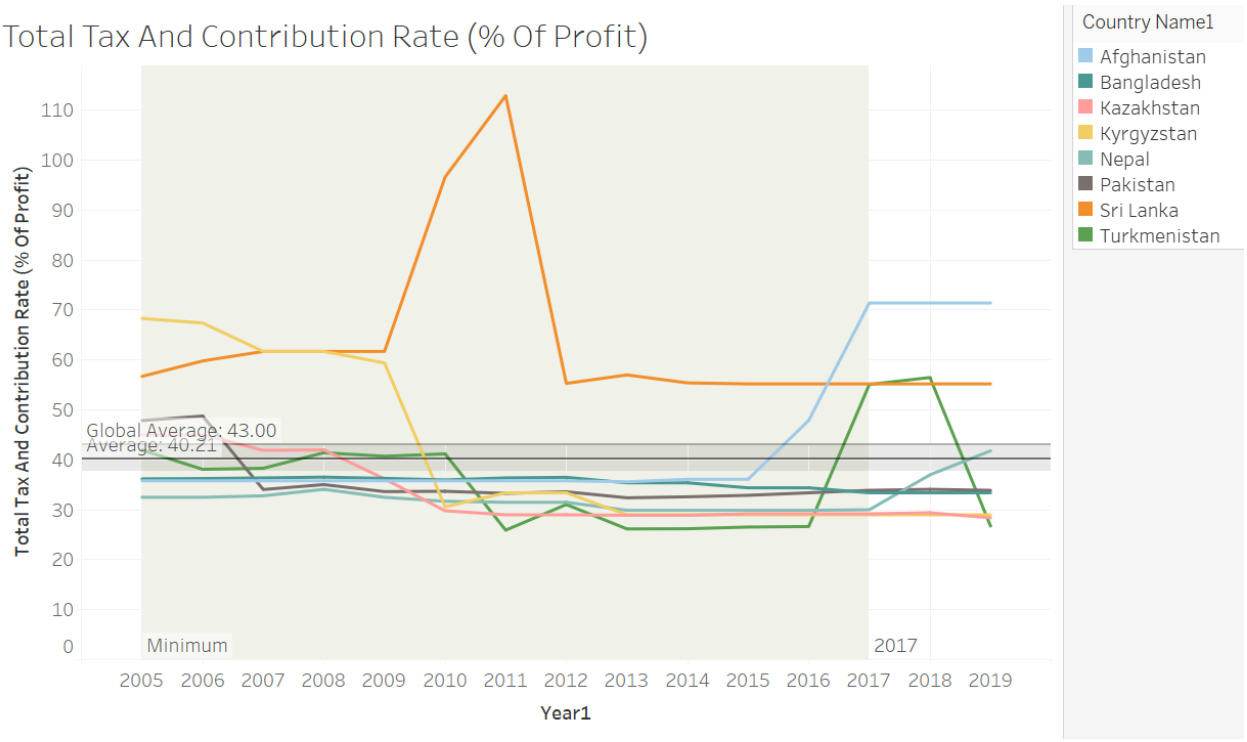


The regional average of 35.3 is slightly lower than the global average value, 43.00. The Total Tax And Contribution Rate (% Of Profit) for most of the countries are very stable over the years (except Yemen). The value differences among these countries are small.

The outliers:

Yemen experienced a tremendous increase from 2005 to 2006 dropping from about 200% to 50% to the average value level.

e) South and Central Asia (SAC)

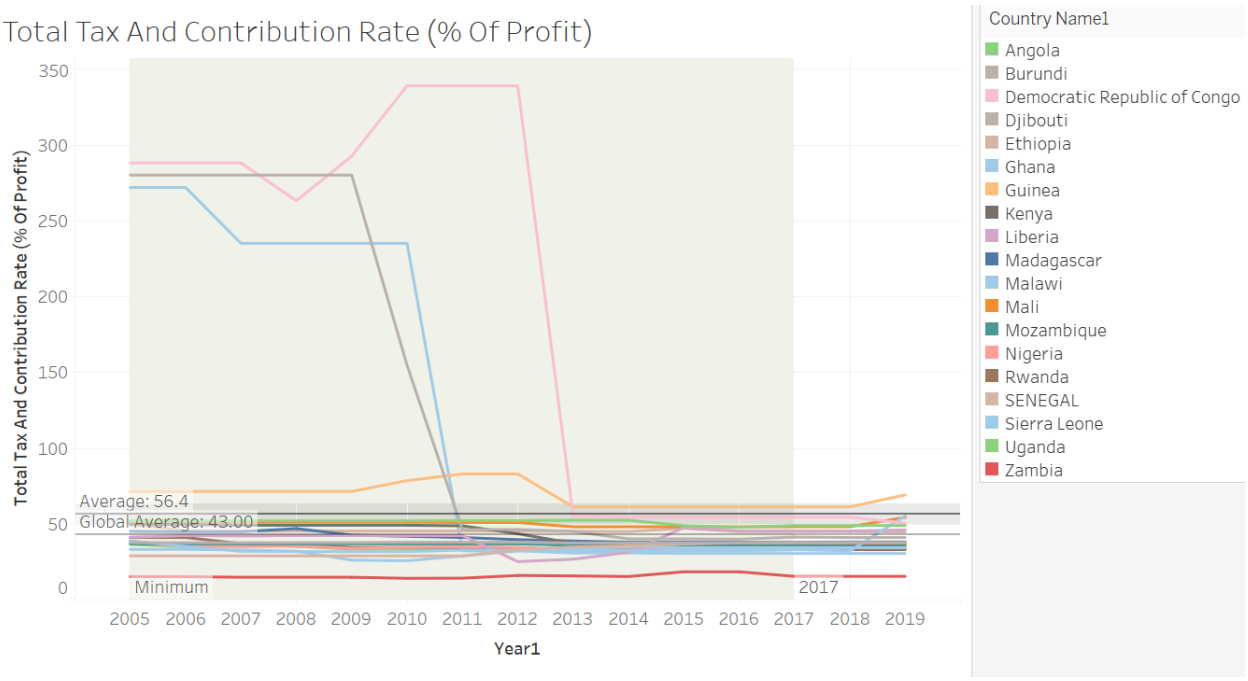


The regional average of 40.21 is slightly larger than the global average value, 43.00. The Total Tax And Contribution Rate (% Of Profit) for most of the countries are very stable over the years (except Sri Lanka, Afghanistan, and Kyrgyzstan). Afghanistan dropped sharply from 2009 to 2010. Kyrgyzstan increased sharply from 2015 to 2017.

The outliers:

Sri Lanka experienced a tremendous increase from 2009 to 2011, with the peak of 110%+, and then experienced a tremendous decrease from 2011 to 2012, dropping to 55%, which could indicate the effectiveness of USAID projects.

f) Sub-Saharan Africa (SSA)



The regional average of 56.4 is larger than the global average value, 43.00. The Total Tax And Contribution Rate (% Of Profit) for most of the countries are very stable over the years (except Demoncratic Republic of Congo, Ghana, and Burundi).

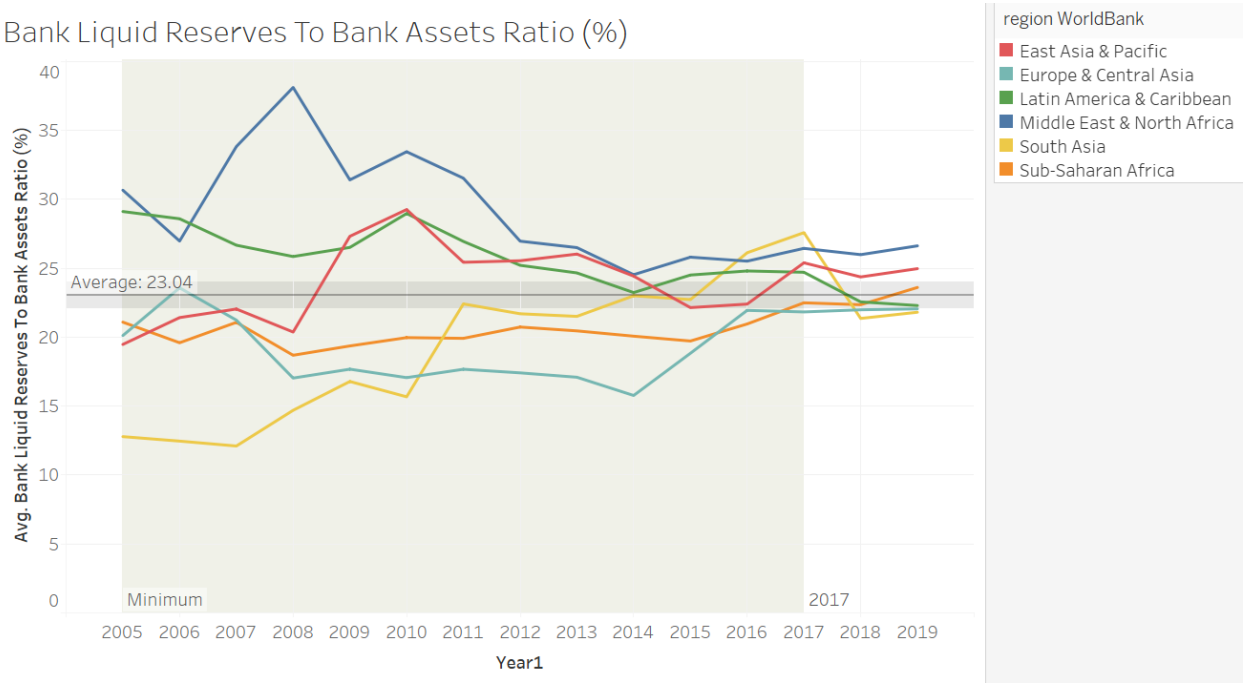
The outliers:

Demoncratic Republic of Congo, Ghana, and **Burundi** all had a tremendous drop from 2009 to 2011 or 2013, which could indicate the effectiveness of USAID projects.

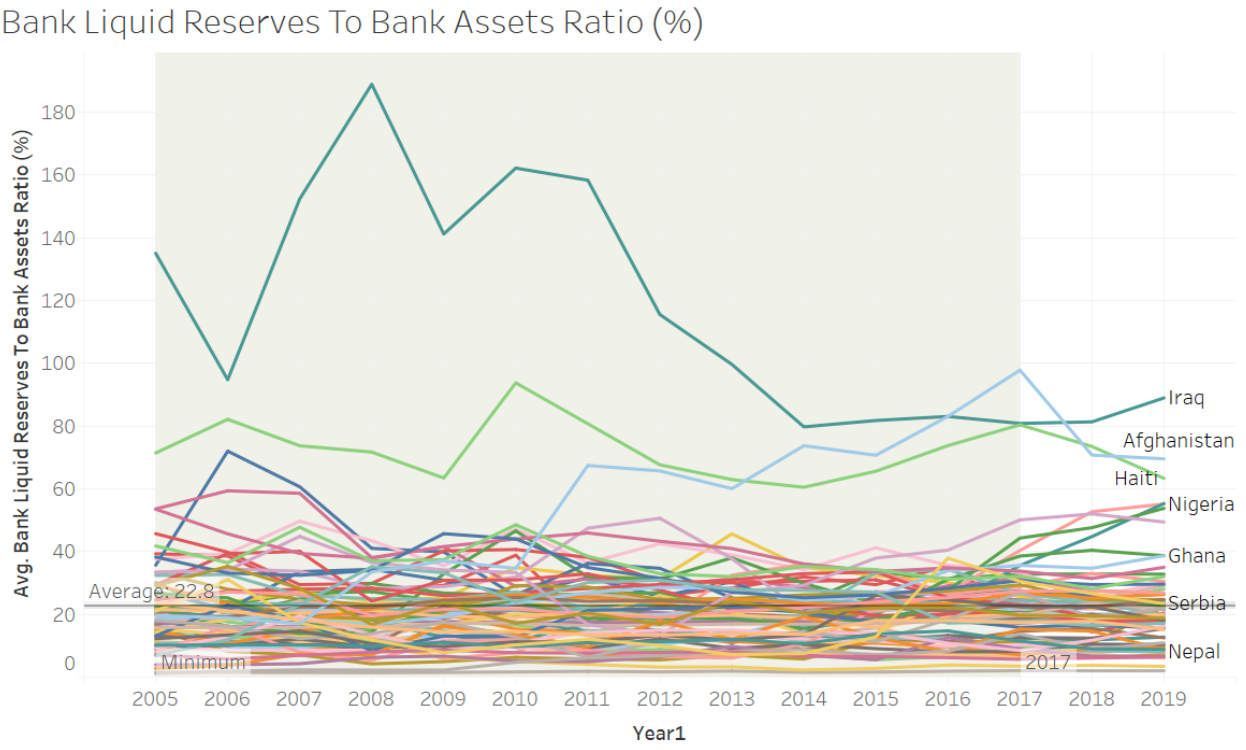
15. Bank Liquid Reserves To Bank Assets Ratio (%)

Ratio of bank liquid reserves to bank assets is the ratio of domestic currency holdings and deposits with the monetary authorities to claims on other governments, nonfinancial public enterprises, the private sector, and other banking institutions. Bank Liquid Reserves To Bank Assets Ratio (%) has a negative correlation with Corruption Perception Index (CPI) at -0.27.

Globally

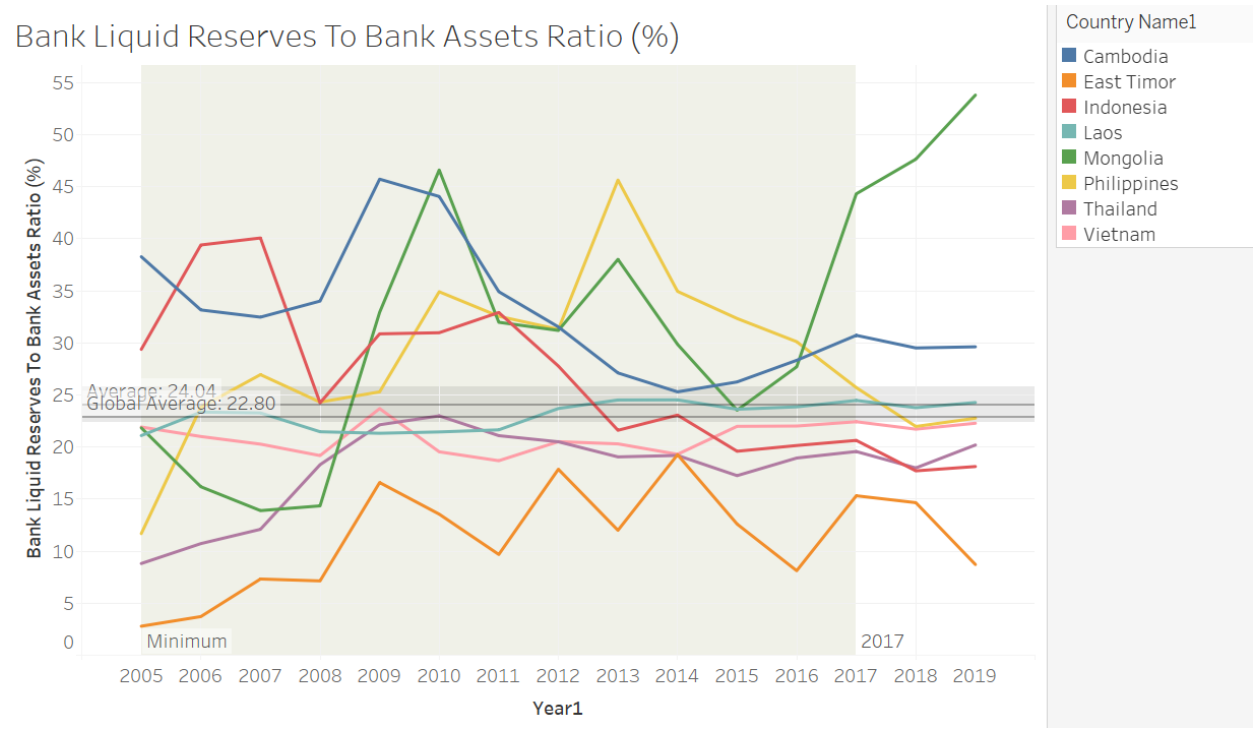


The global average of Bank Liquid Reserves To Bank Assets Ratio (%) is 23.04. The region with the largest Bank Liquid Reserves To Bank Assets Ratio (%) is Middle East & South Africa. The variance among these regions became less and less as the years went on.



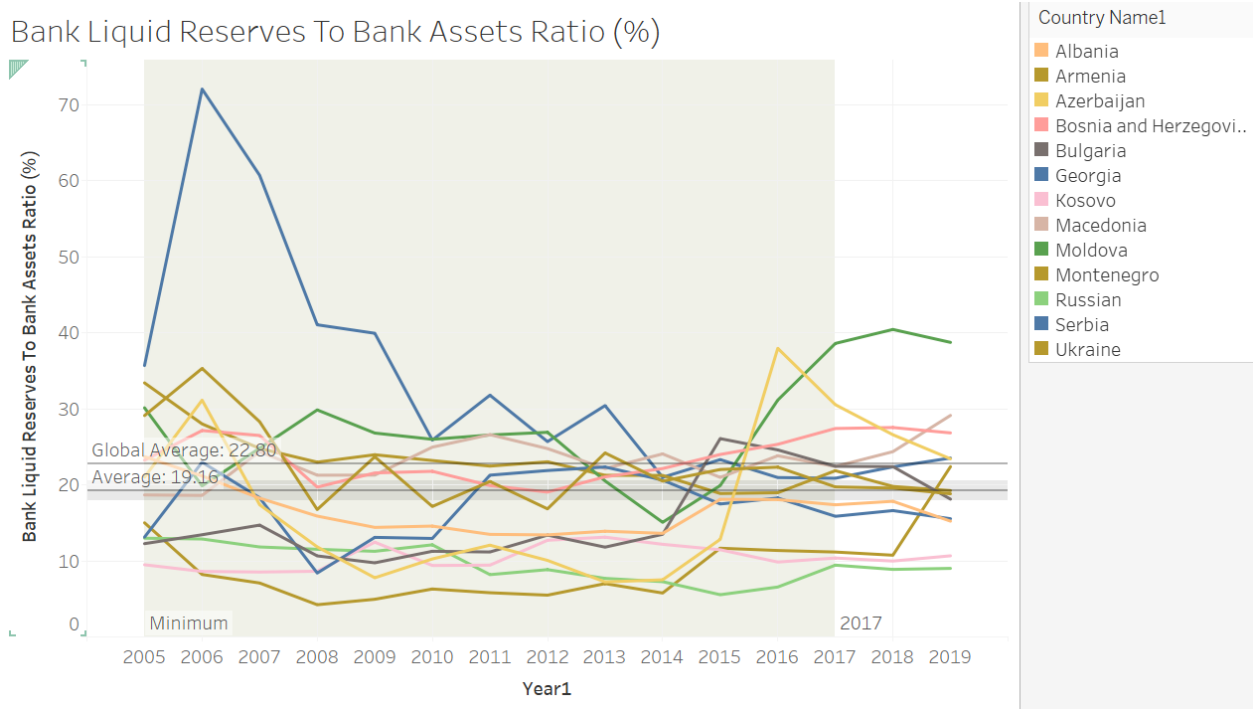
The Bank Liquid Reserves To Bank Assets Ratio (%) for the most countries are between 0 to 40%. The country with the largest Bank Liquid Reserves To Bank Assets Ratio (%) is Iraq.

a) East Asia and Pacific (EAP)



The regional average of 24.04 is slightly higher than the global average (22.80). There's no obvious trend change over time for these countries in general. Mongolia experienced a tremendous growth rate from 2007 to 2010, with the peak in 2010 reaching 45%+.

b) Europe and Eurasia (E&E)

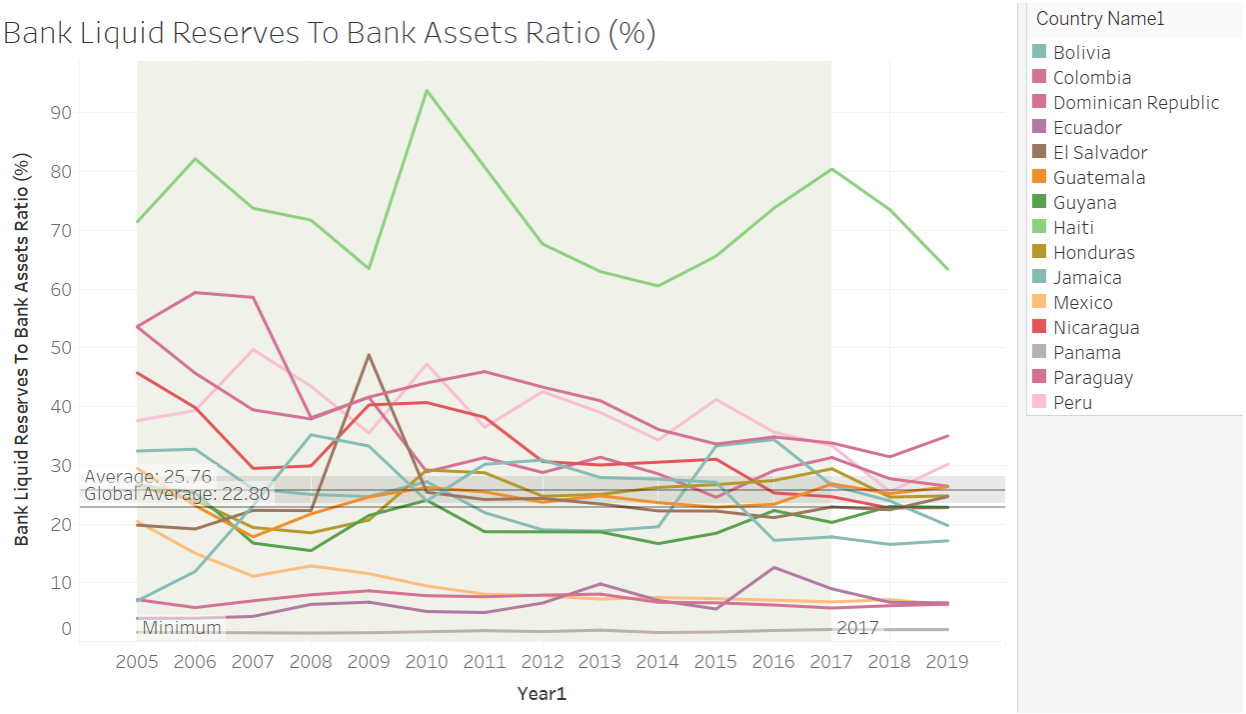


The regional average of 19.16 is slightly lower than the global average (22.80). There's no obvious trend change over time for most countries (except Serbia).

The outliers:

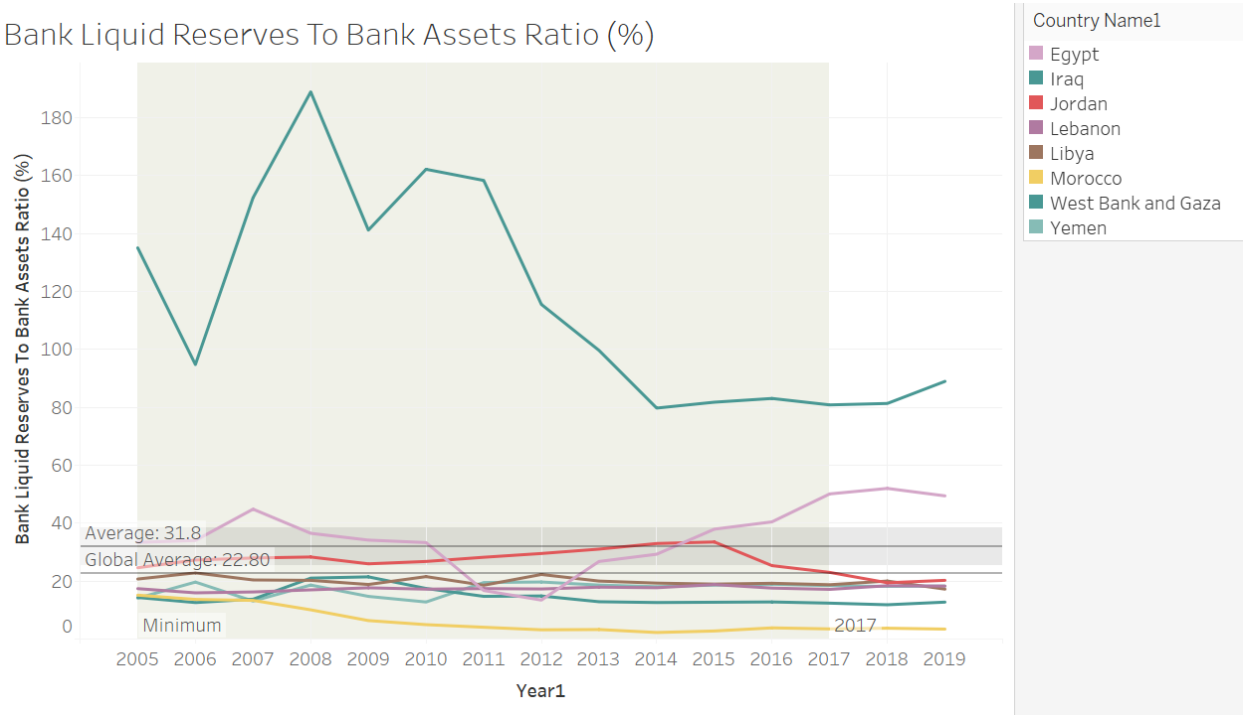
Serbia experienced a tremendous growth rate from 2005 to 2006, with the peak of 70%+. And then it dropped tremendously from 2006 to 2010, and is back to the average level value.

c) Latin America and Caribbean (LAC)



The regional average of 25.76 is slightly higher than the global average (22.80). Haiti, EI Salvador, and Nicaragua all had shape increases in 2009 - 2010. Mexico, Colombia, and Panama have the relatively low Bank Liquid Reserves To Bank Assets Ratio (%) values compared with the other countries, with the values of around 5%. **Haiti** has a relatively higher value than the other countries, with the highest values among all the years.

d) Middle East and North Africa (MENA)

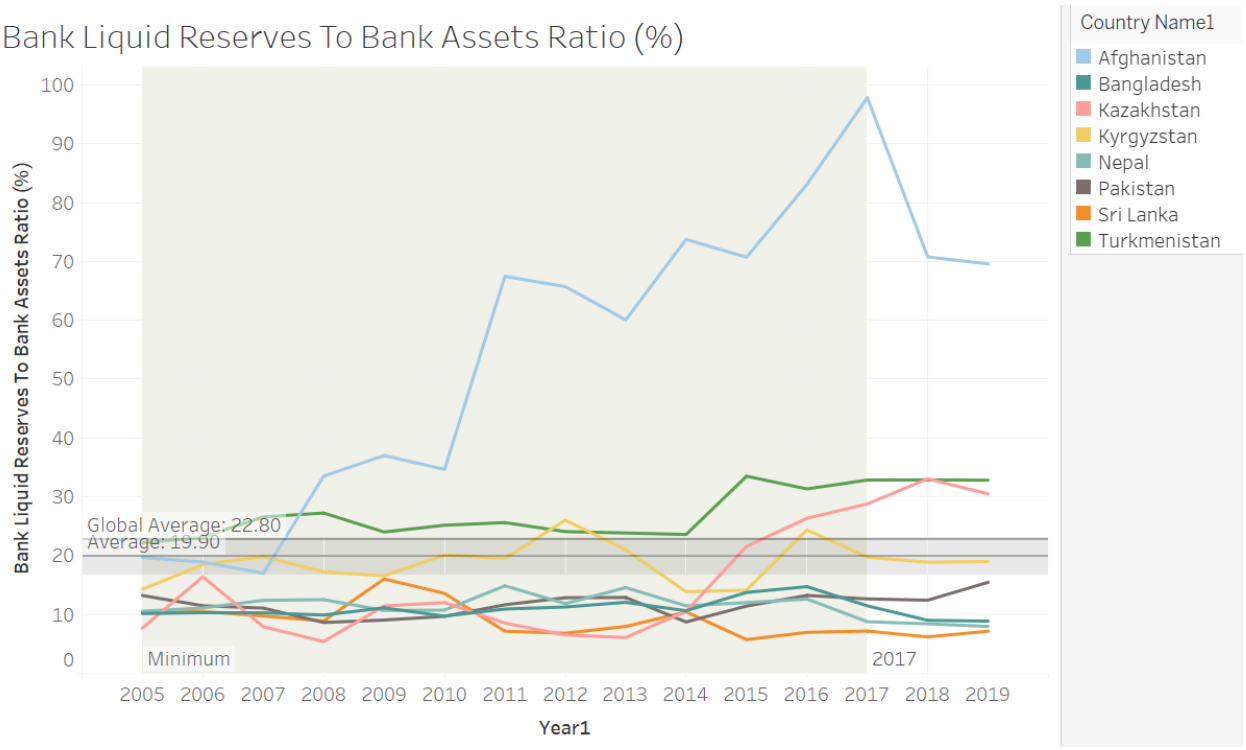


The regional average of 31.8 is higher than the global average value, 22.80. Most of the countries have stable Bank Liquid Reserves To Bank Assets Ratio (%) values over the years, with the values around the average.

The outliers:

Irap has the much higher Bank Liquid Reserves To Bank Assets Ratio (%) values than the other countries for all the years, with the values ranging from 80 to 180. It is supposed that the increase over the year from 2006 to 2011 could be explained by the impact of Irap War.

e) South and Central Asia (SAC)

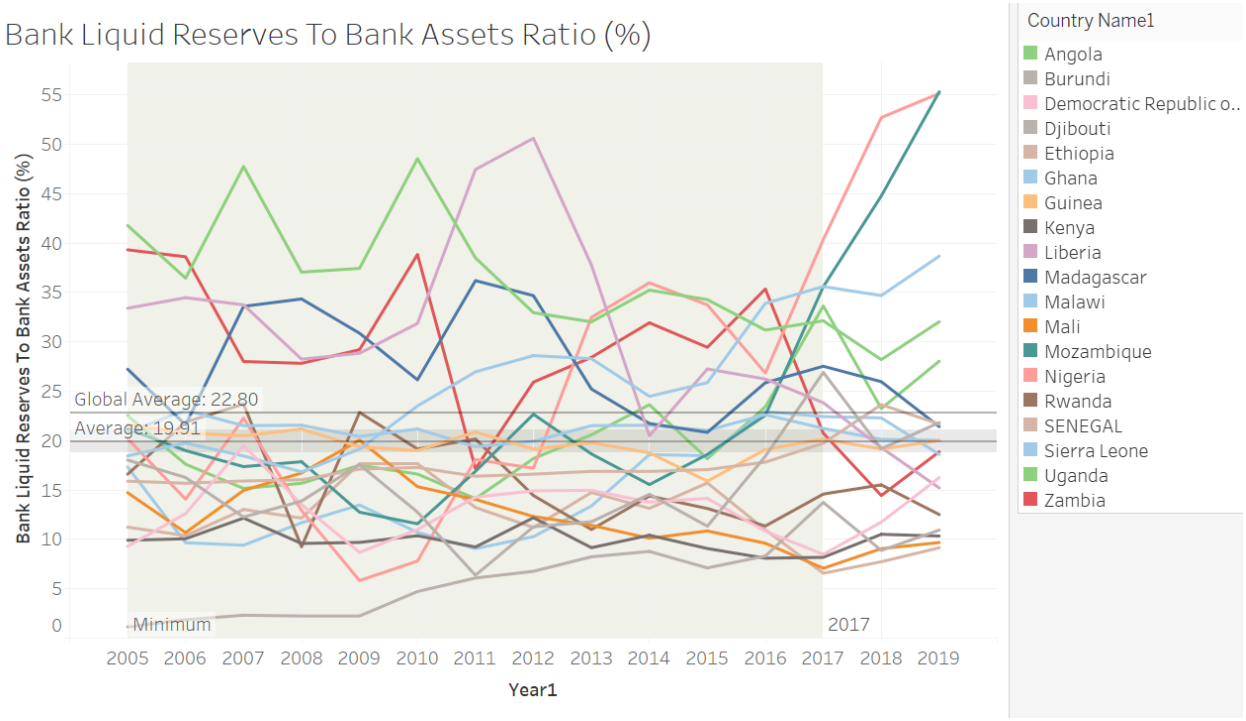


The regional average of 19.90 is slightly lower than the global average value, 22.80. The Bank Liquid Reserves To Bank Assets Ratio (%) values for most of the countries are stable over the years (except Afghanistan), with values around the average.

The outliers:

Afghanistan had continuous growth from 2006 to 2017, with the peak of nearly 100%. It is supposed that the increase over the year from 2006 to 2011 could be explained by the impact of the War.

f) Sub-Saharan Africa (SSA)

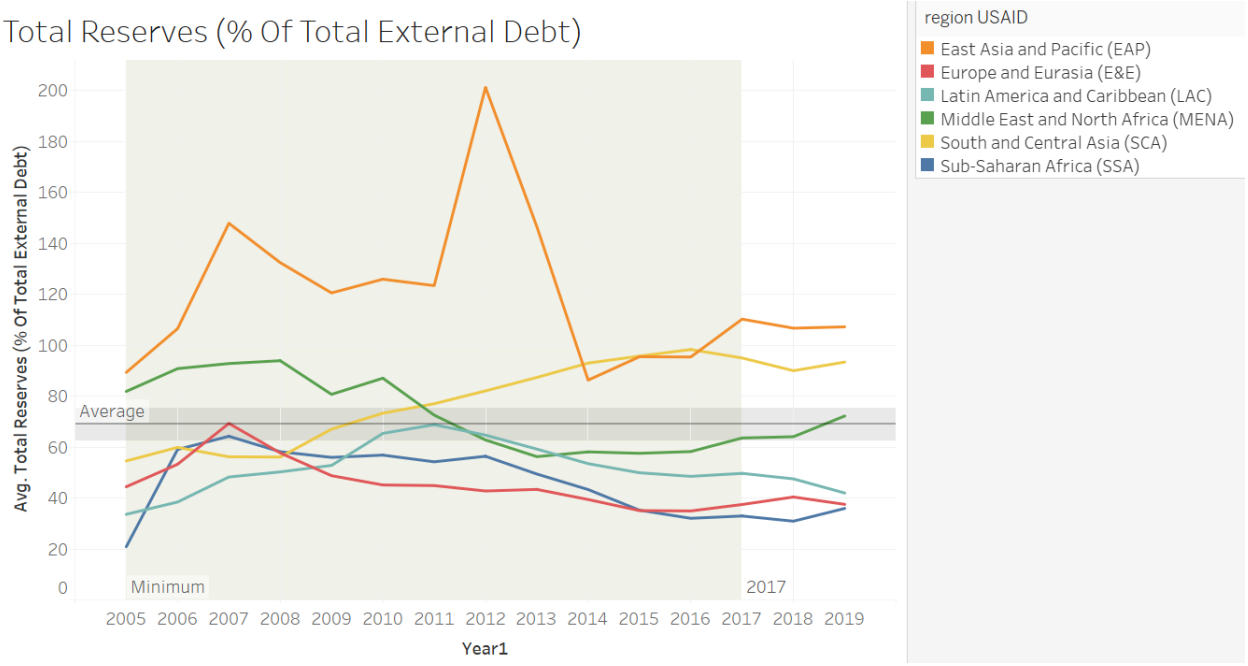


The regional average of 19.91 is slightly lower than the global average (22.80). Most of the countries have stable Bank Liquid Reserves To Bank Assets Ratio (%) values over the years.

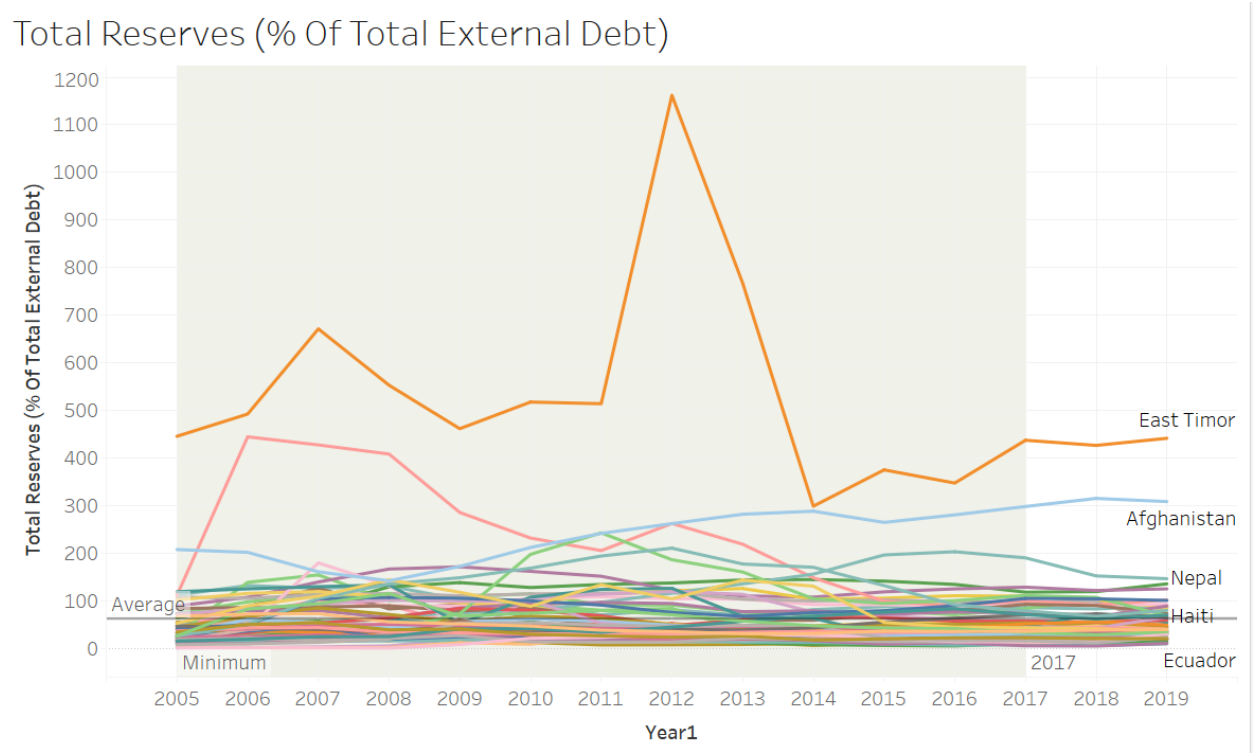
16. Total Reserves (% Of Total External Debt)

Total Reserves (% Of Total External Debt) represents the International reserves to total external debt stocks. Total Reserves (% Of Total External Debt) has a slight positive correlation with Corruption Perception Index (CPI) at 0.072.

Globally

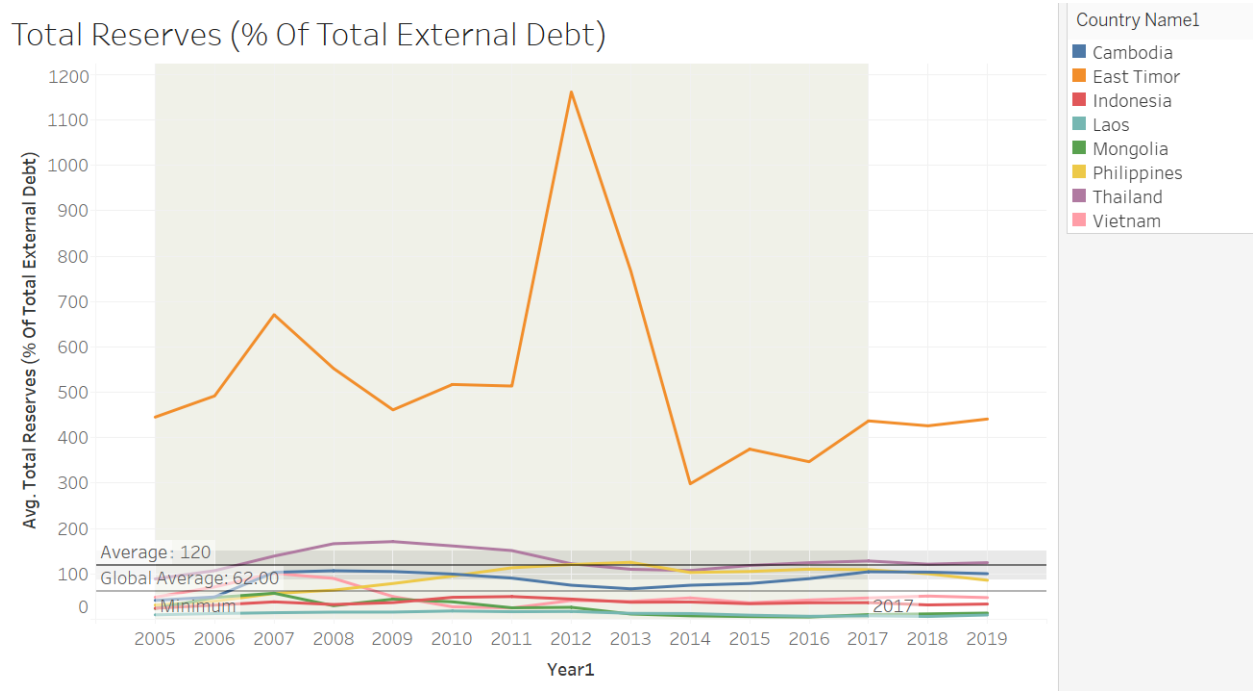


The region with the largest Total Reserves (% Of Total External Debt) is East Asia and Pacific region for the most years.



The Total Reserves (% Of Total External Debt) for most countries are between 0 to 130%. The country with the largest Total Reserves (% Of Total External Debt) is East Timor.

a) East Asia and Pacific (EAP)

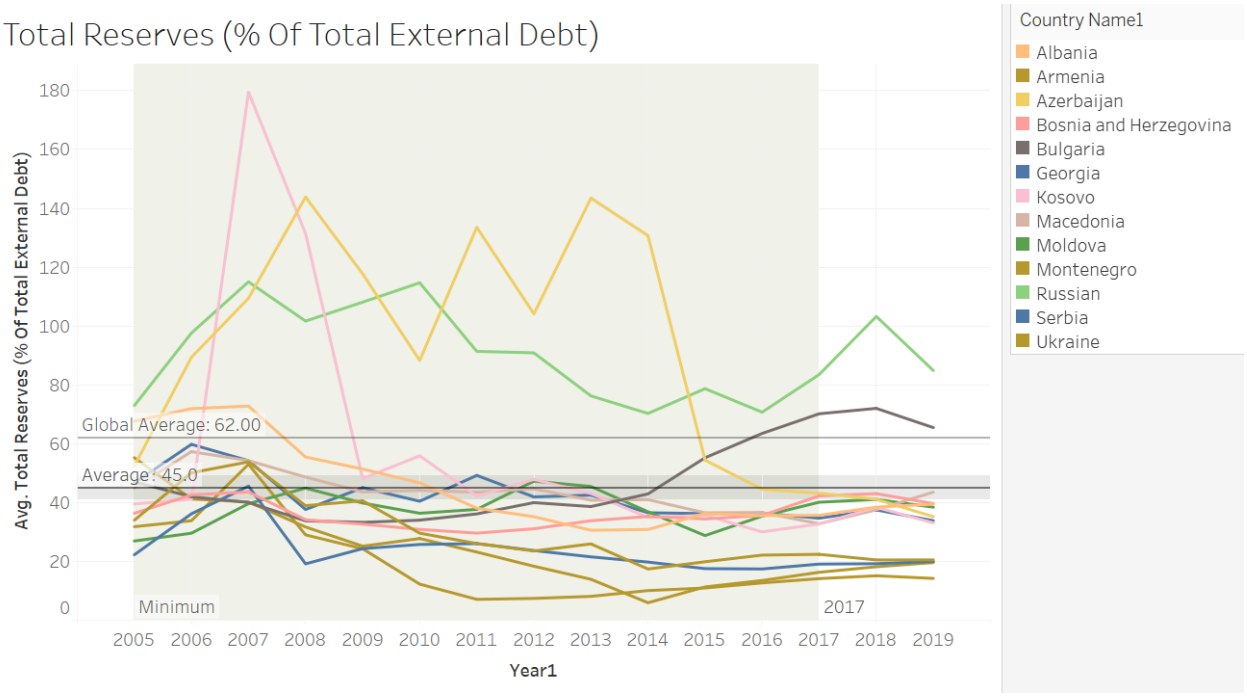


The regional average (120) is larger than the global average value (62). The Total Reserves (% Of Total External Debt) for most of the countries are very stable over the years (except East Timor).

The outliers:

East Timor had a tremendous increase from 2011 to 2012, and a tremendous drop from 2012 to 2014. It has the higher Total Reserves (% Of Total External Debt) than all the other countries.

b) Europe and Eurasia (E&E)

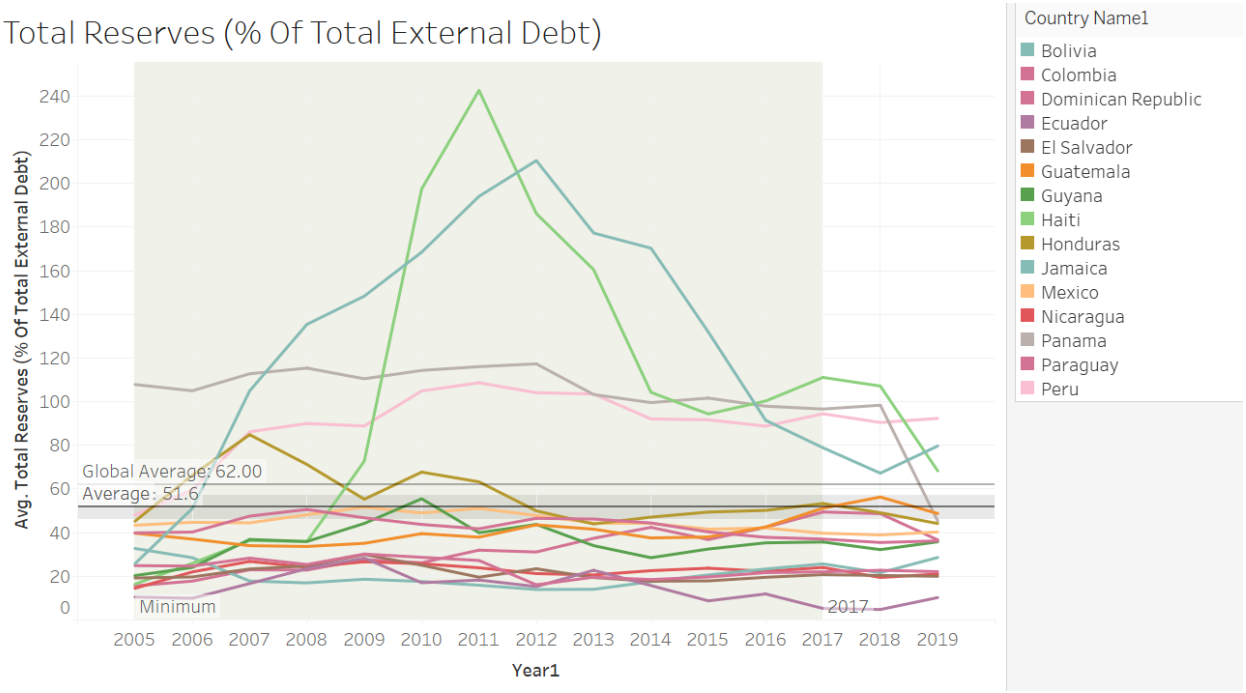


The regional average (45) is lower than the global average value (62). Almost all the countries experienced a decrease from 2007 to 2008. Kosovo, Azerbaijan, and Russian have large variations over the years.

The outliers:

Kosovo experienced a tremendous increase from 2006 to 2007, with the peak of around 180, and a tremendous decrease from 2007 to 2009.

c) Latin America and Caribbean (LAC)

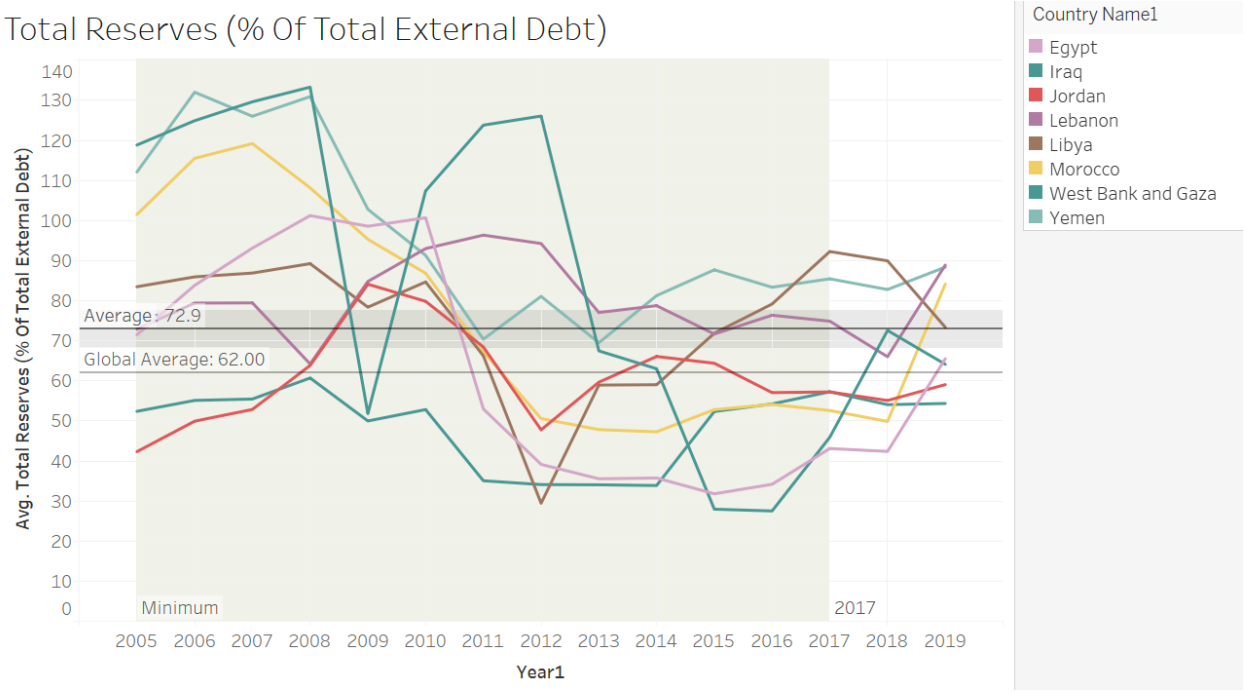


The regional average (51.6) is lower than the global average value (62). Most of the countries have stable Total Reserves (% Of Total External Debt) over the years except Haiti and Bolivia.

The outliers:

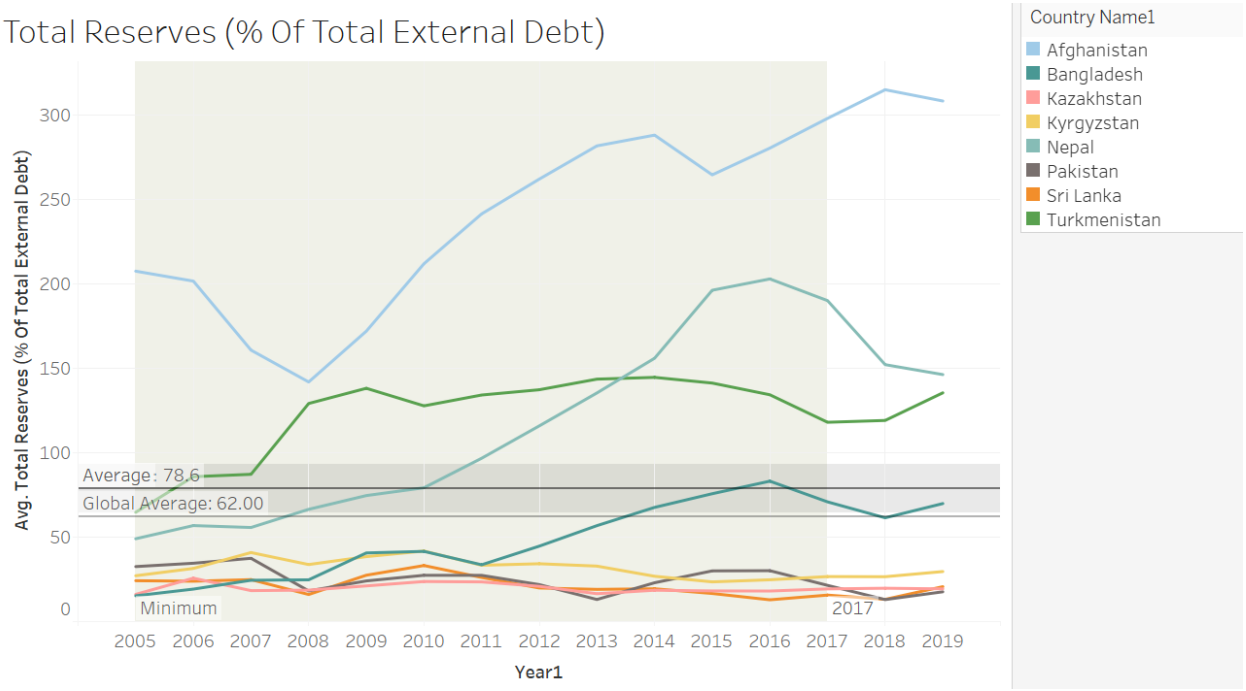
Both **Haiti** and **Bolivia** experienced a tremendous increase from 2007 to 2011(2012), and a tremendous decrease from 2011(2012) to 2015.

d) Middle East and North Africa (MENA)



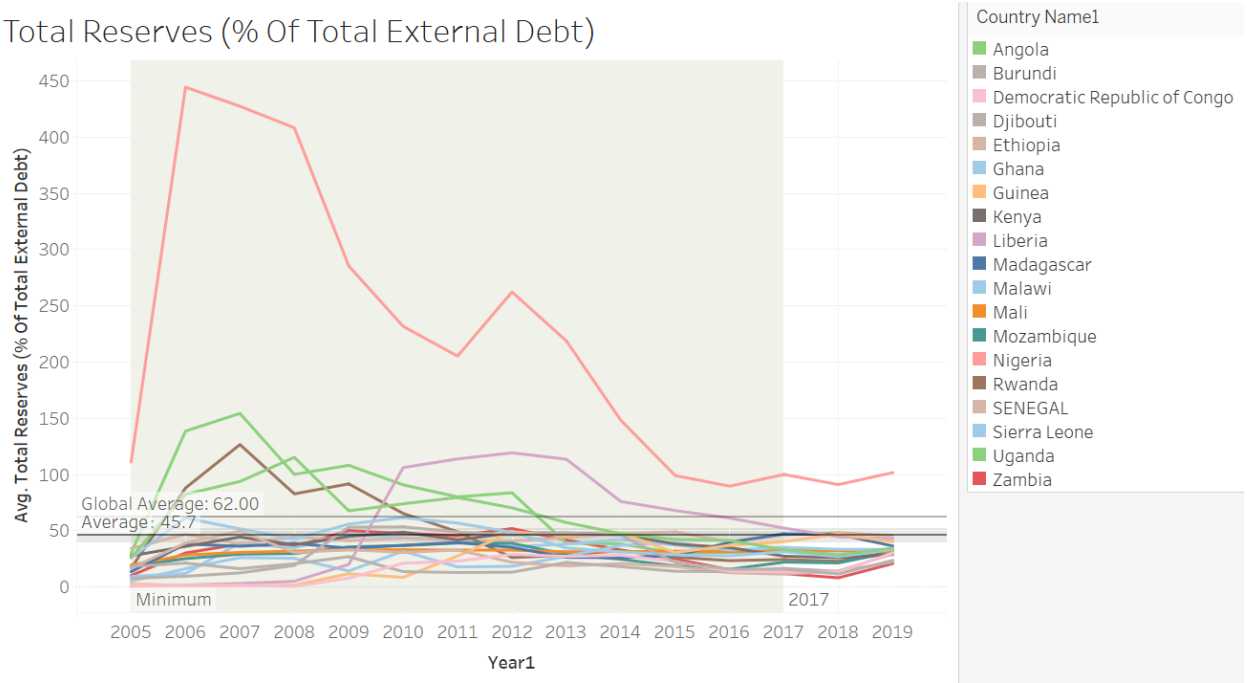
The regional average (72.9) is lower than the global average value (62). All the countries have large variances for Total Reserves (% Of Total External Debt) over the years.

e) South and Central Asia (SAC)



The regional average (78.6) is larger than the global average value (62). Afghanistan, Turkmenistan, and Bangladesh all have an increase for Total Reserves (% Of Total External Debt) from 2011 to 2016.

f) Sub-Saharan Africa (SSA)



The regional average (45.7) is lower than the global average value (62). Most of the countries have stable Total Reserves (% Of Total External Debt) over the years.

The outliers:

Nigeria experienced a tremendous increase from 2005 to 2006, with the peak of around 450, and a tremendous decrease from 2006 to 2011.

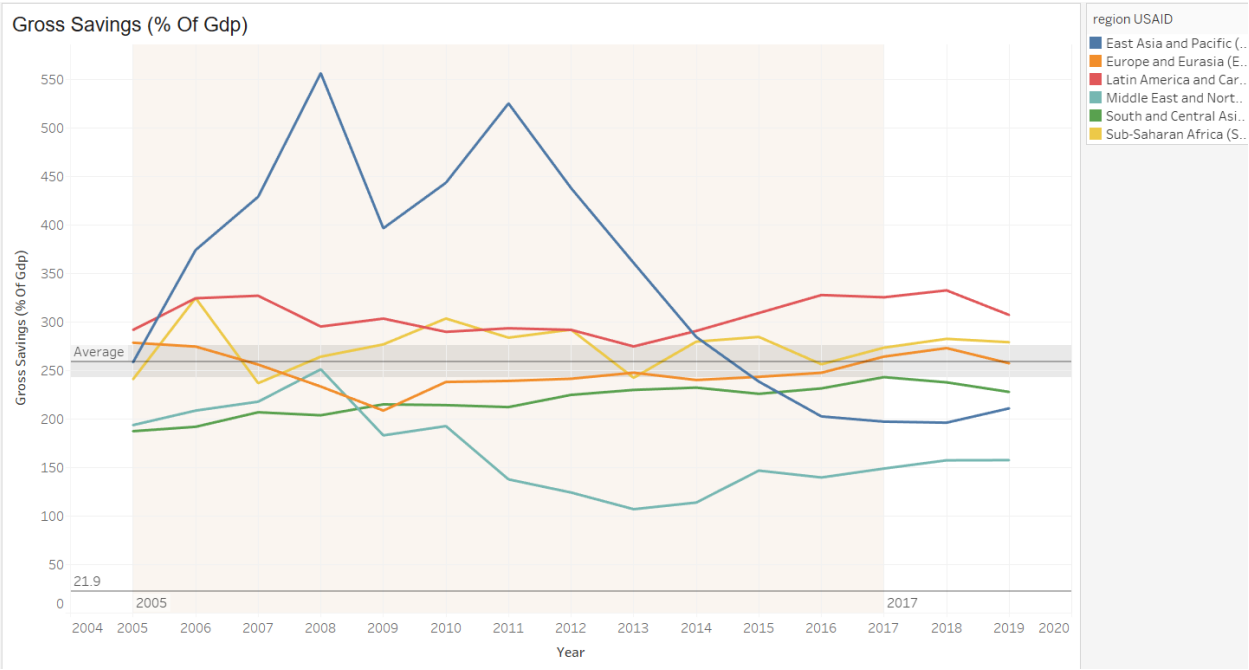
17. Gross Savings (% Of Gdp)

Gross Domestic Saving is GDP minus final consumption expenditure. It is expressed as a percentage of GDP. Description: Gross Domestic Saving consists of savings of household sector, private corporate sector and public sector.

The correlation between cpi and gross savings is lower than -0.1,which is negative.

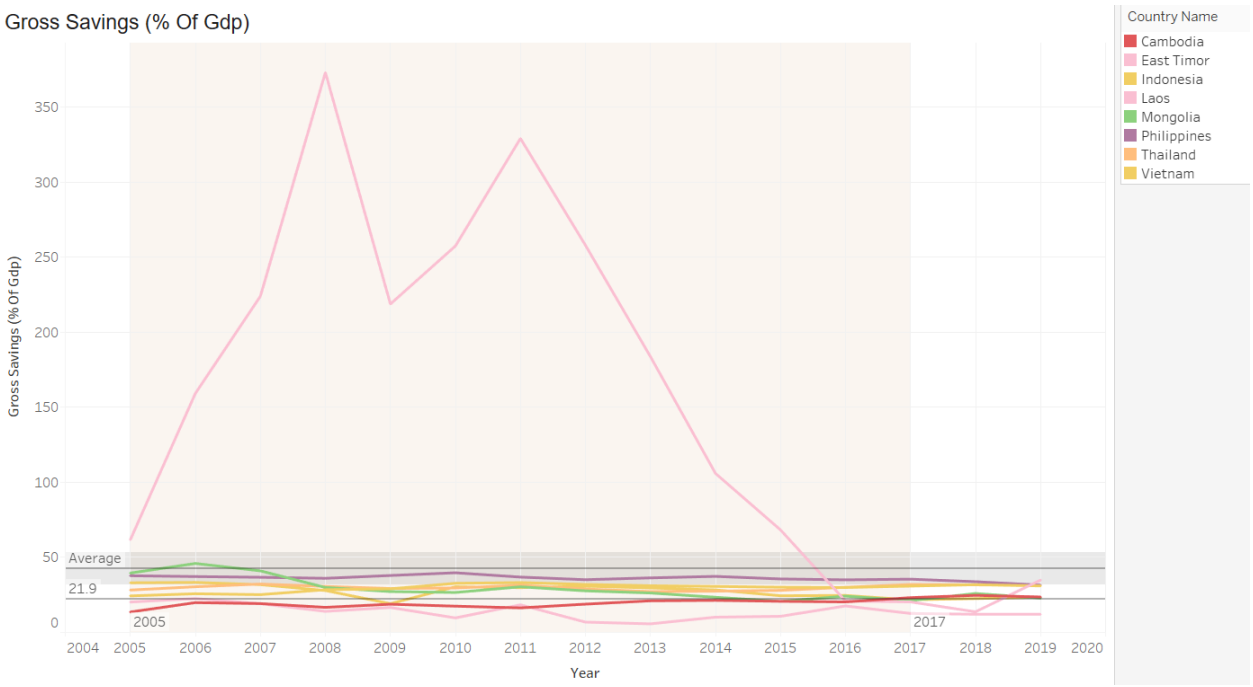
Global average (with 95% CI) is 21.9% over the period of 2005 to 2019.

Regional graph:



Except for EAP, all regions only fluctuated slightly, with little change.MENA even experienced a light decrease from 2005 to 2017.EAP increased first and had two jumps from 2007 to 2009 and from 2009 to 2013.And then EAP sharply decreased to 180 until 2017.

a) East Asia and Pacific (EAP)

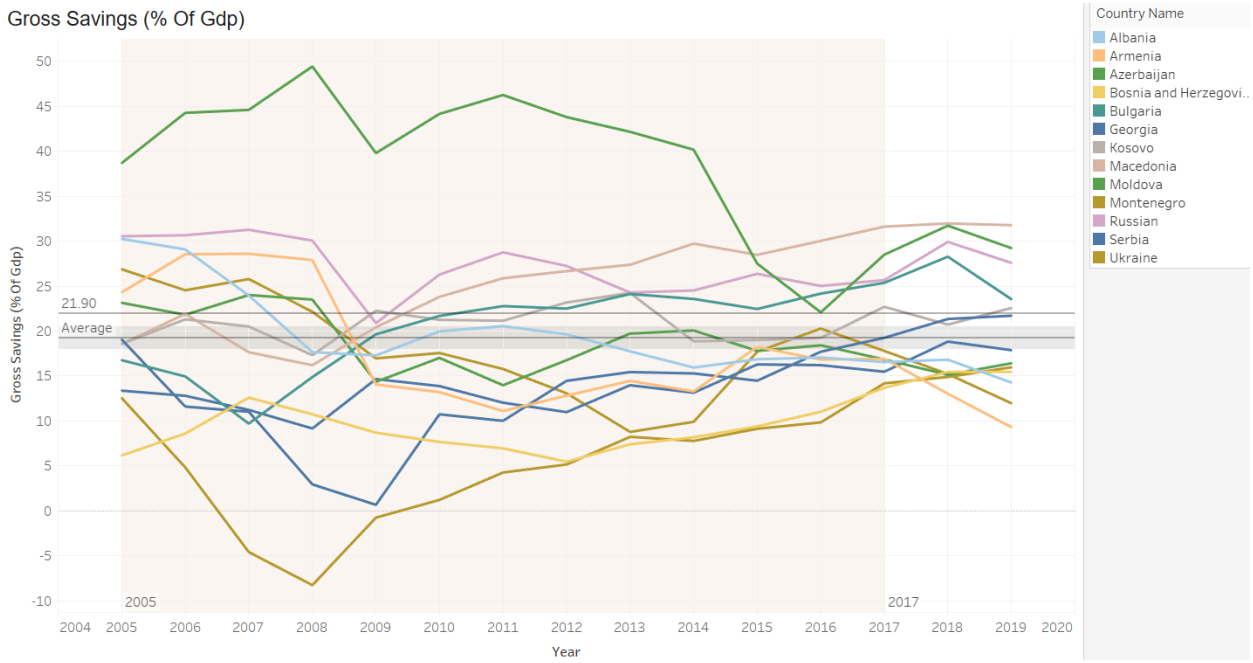


Almost all countries are lower than average and did not change much during the years.

The outliers:

The Gross Savings value of East Timor is higher than average and all countries in the EAP region.And it experienced a rapid rise, fall, rise again, in the process of rapid decline, finally fell to near the average.

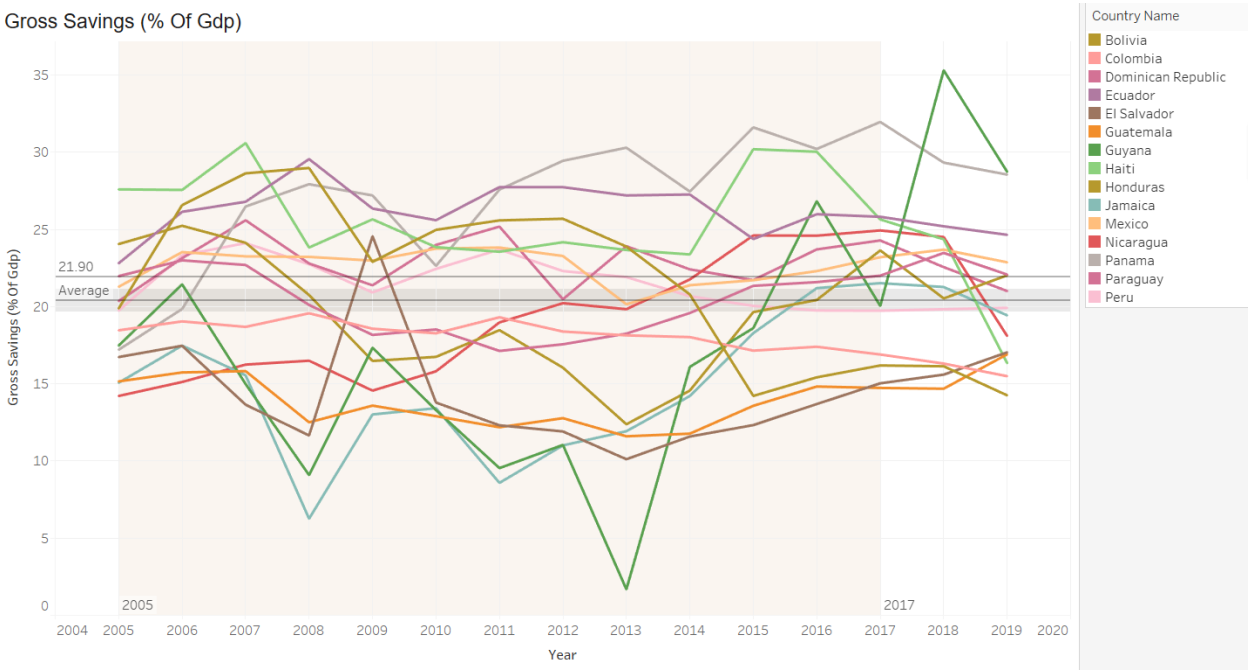
b) Europe and Eurasia (E&E)



The average of E&E is almost similar to(a little lower than) the global average.

Different from the overall trend of other countries(increase),Moldova decreased from 38% to 30% overall.

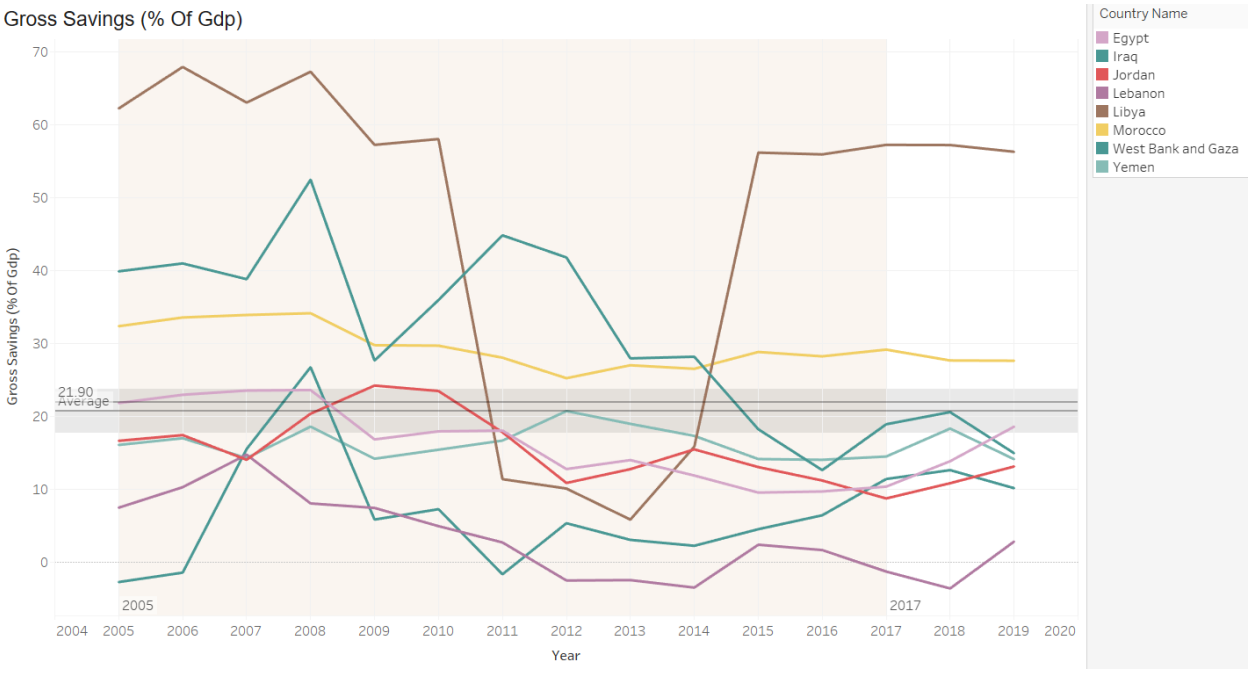
c) Latin America and Caribbean (LAC)



The average of LAC is almost similar to the global average.

The changes in various countries are relatively disorderly and different.Several countries have a sharp increase and then have a sharp decrease during 2008 and 2010.

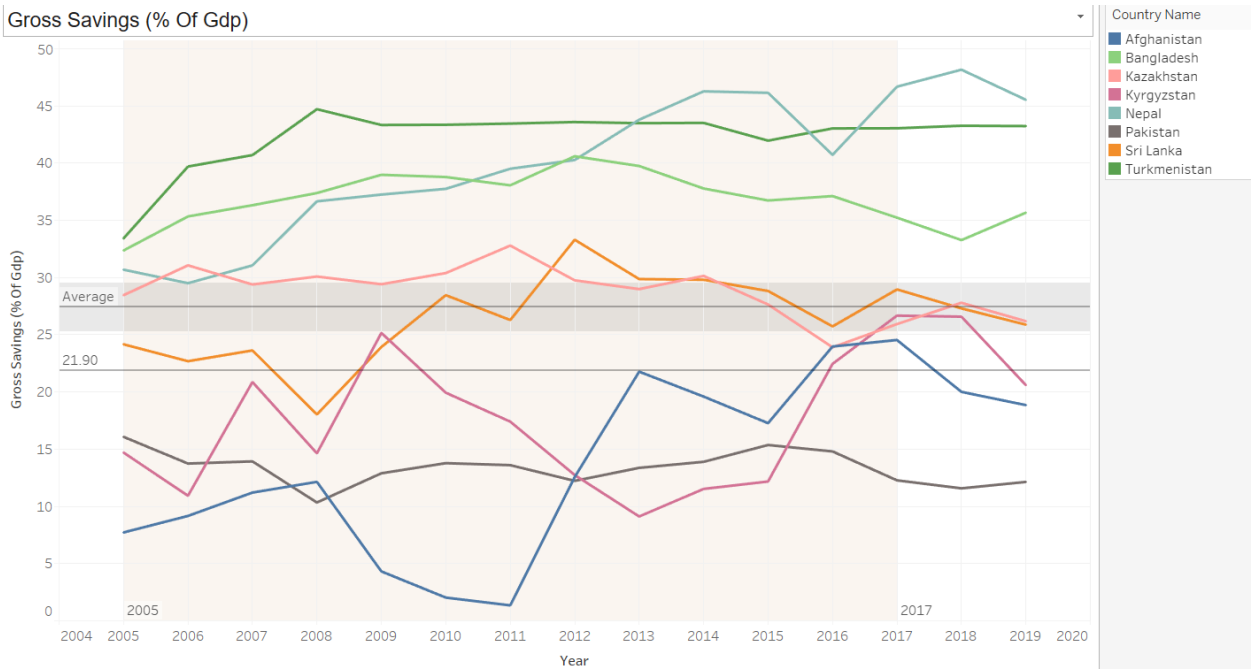
d) Middle East and North Africa (MENA)



The average of MENA is almost similar to the global average.

Libya sharply decreases and then sharply decreases these years and remains the highest value among these countries after 2015.

e) South and Central Asia (SAC)

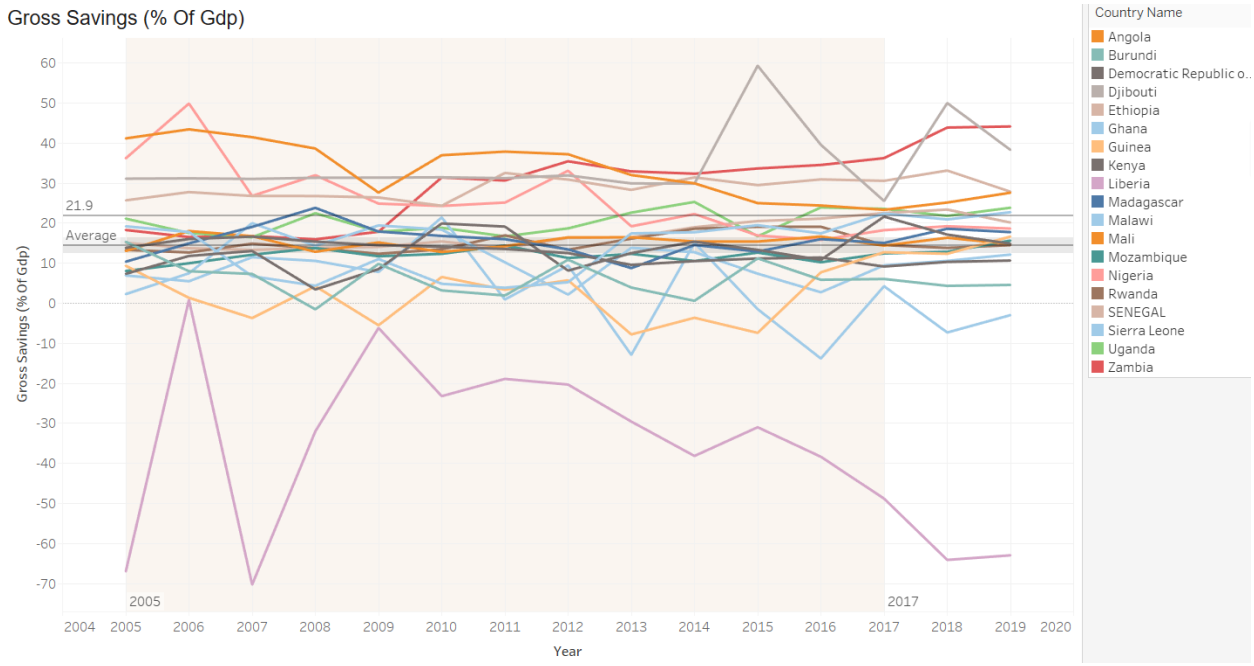


The average of this region is higher than the global average .

The outliers:

After a general increase,**Afghanistan** experienced a sharp decrease from 2008 to 2011,and then sharply increased to the global average.Because of the negative correlation between gross saving and cpi,that maybe can explain why USAID gave so much investment money to Afghanistan before 2017.

f) Sub-Saharan Africa (SSA)



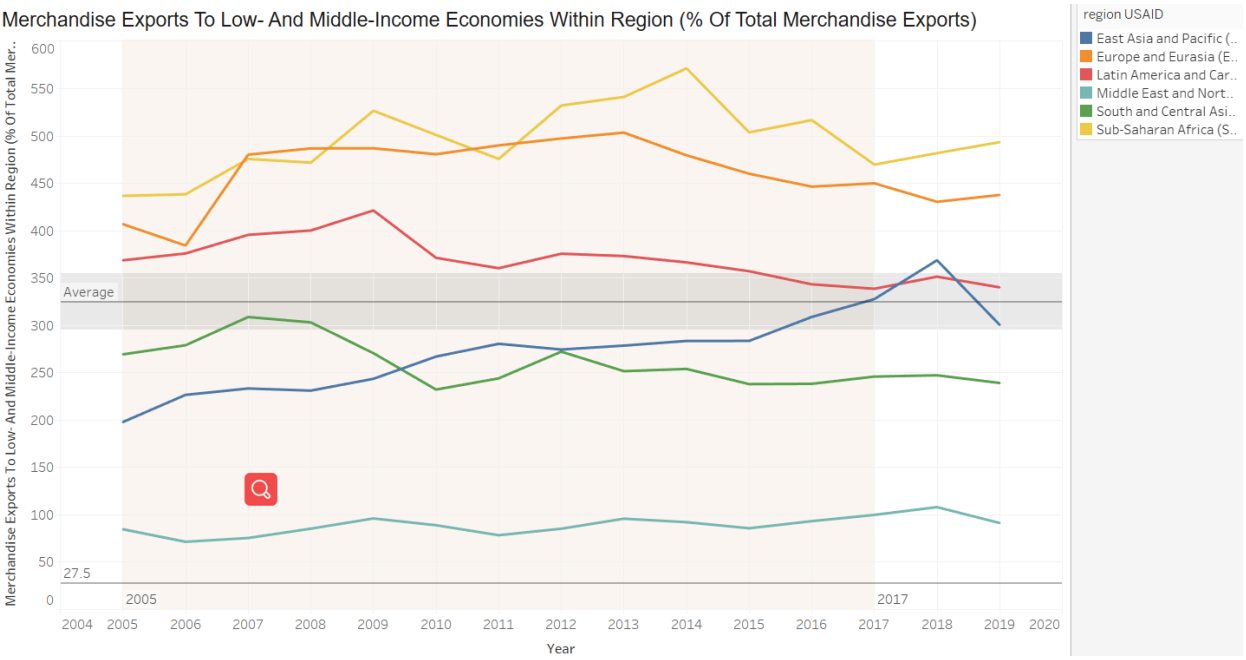
The average of this region is lower than the global average .

Liberia is lower than all countries during 2005 to 2017。Although it rose rapidly in 2006, it is still the lowest value this year.

18. **Merchandise Exports To Low- And Middle-Income Economies Within Region (% Of Total Merchandise Exports)**

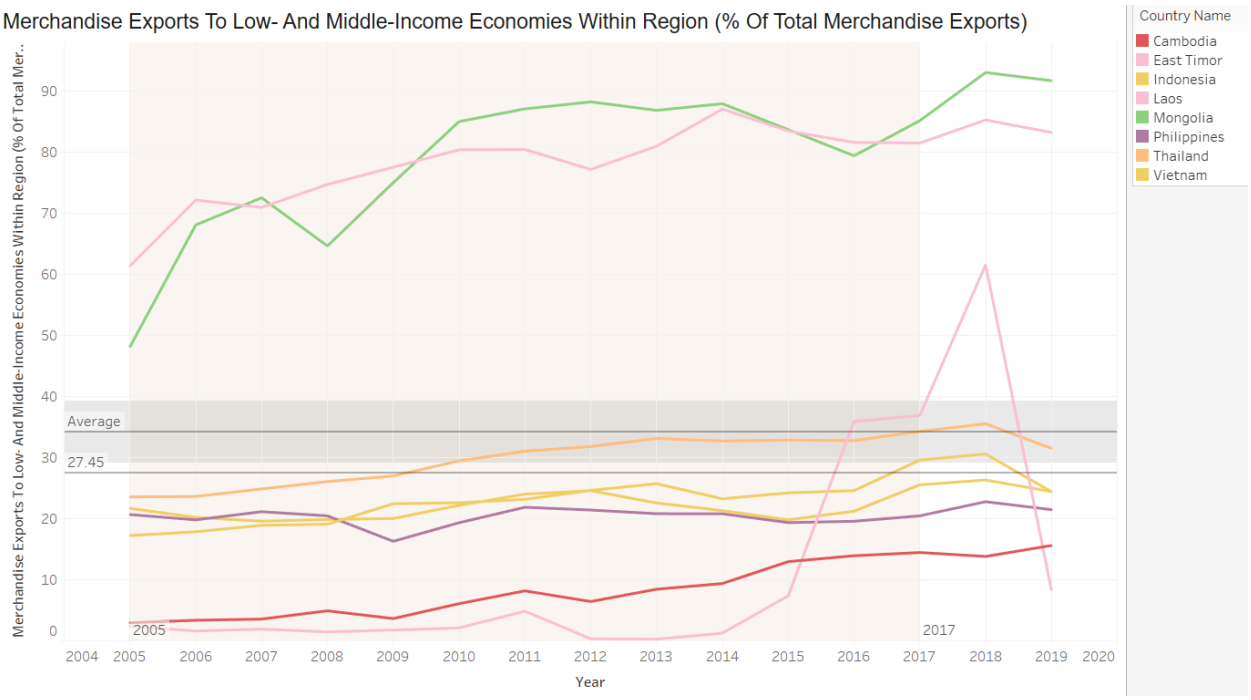
Merchandise exports to low- and middle-income economies within a region are the sum of merchandise exports from the reporting economy to other low- and middle-income economies in the same World Bank region as a percentage of total merchandise exports by the economy.
Global average (with 95% CI) is 27.45% over the period of 2005 to 2019.The correlation between cpi and this indicator is 0.116(which is better than price level ratio,but still not a strong correlation).That means the higher Merchandise Exports To Low- And Middle-Income Economies Within Region ,the earlier for corruption.But the lower this indicator is,the better for countries.

Regional graph:



All regions have not changed much as a whole, with small fluctuations between 2005 and 2017.SSA remains the highest Merchandise Exports within region value after 2011,and MENA remains the lowest value all years.

a) East Asia and Pacific (EAP)

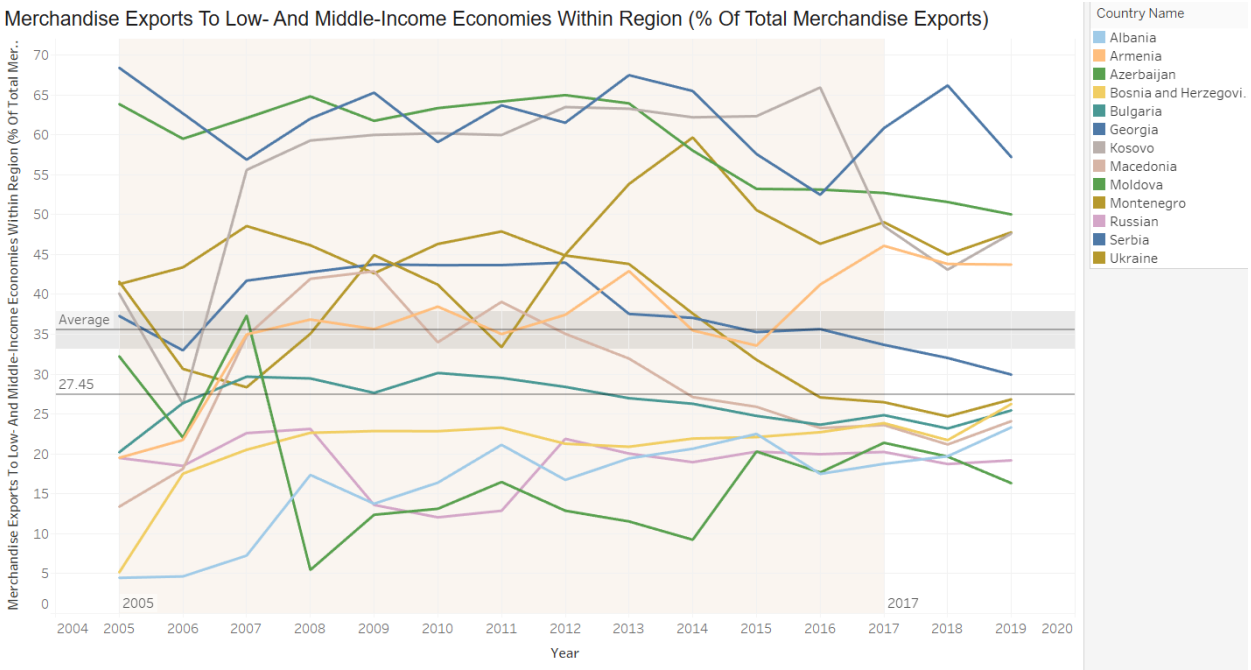


This graph shows that two countries(Loes and Mongolia) are much higher than average and other countries.And the values of these two countries also increase overall.But other countries showed less fluctuation in price level ratio growth as time went by and are all lower than average every year..

The outliers:

East Timor showed a huge increase in 2017-2018, and reached a peak in 2018 (but still lower than average), and then dropped rapidly.Mongolia shows a small decrease during 2007 and 2008.

b) Europe and Eurasia (E&E)

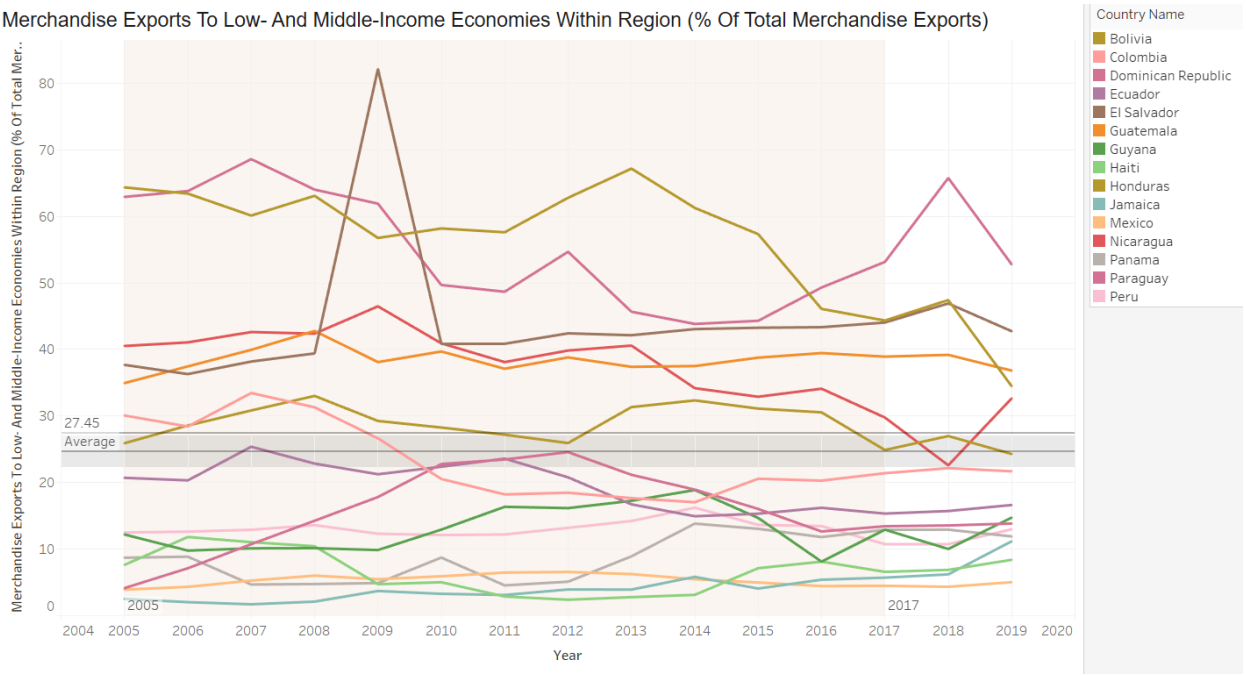


The changes in various countries are relatively disorderly and different.Most countries are below or above the average every year, and there is almost no case around the average line.

The outliers:

Kosovo and Macedonia have a sharp increase between 2006 and 2008.Azerbajian has a sharp decrease between 2007 and 2008.

c) Latin America and Caribbean (LAC)

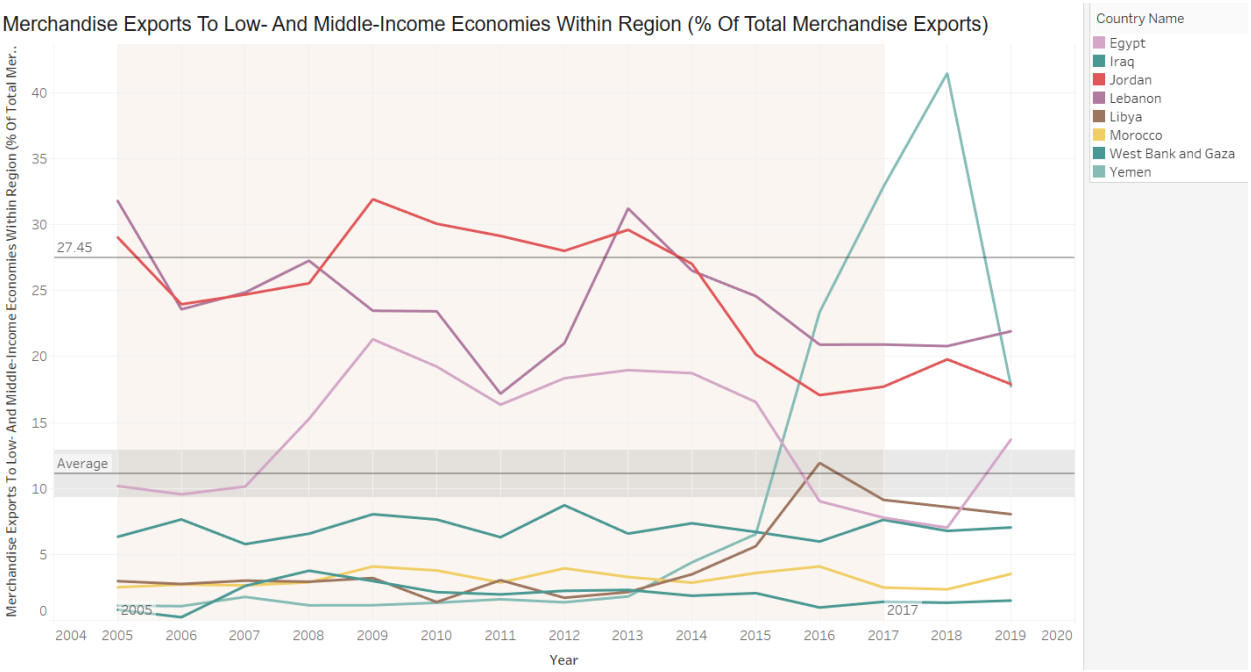


All countries showed less fluctuation in price level ratio growth as time went by and showed little change overall .Most countries are below or above the average every year, and there is almost no case around the average line.

The outliers:

Different from the small changes in other countries,the value of **El Salvator** fell sharply and rose sharply from 2008 to 2010.The value of **El Salvator** in 2009 rose to twice the value in 2008.

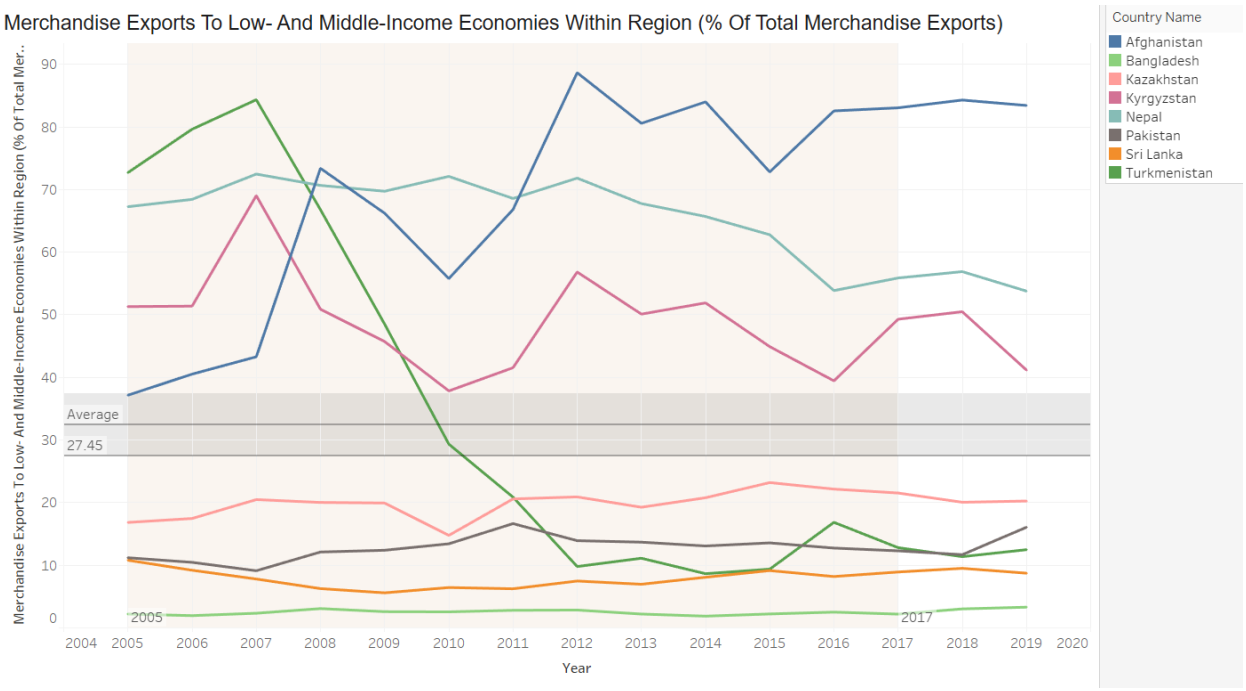
d) Middle East and North Africa (MENA)



The MENA regional average is much higher than the global average.
Four countries have changed a lot, other countries have not changed much.And all countries change very irregularly.

The outliers:
The value of **Yemen** fell sharply and rose sharply from 2015 to 2019.The value of **Yemen** reached its peak in 2018, and was the highest median country in the region that year.And the value of Yemen is higher than all the countries among MENA region after 2016.

e) South and Central Asia (SAC)



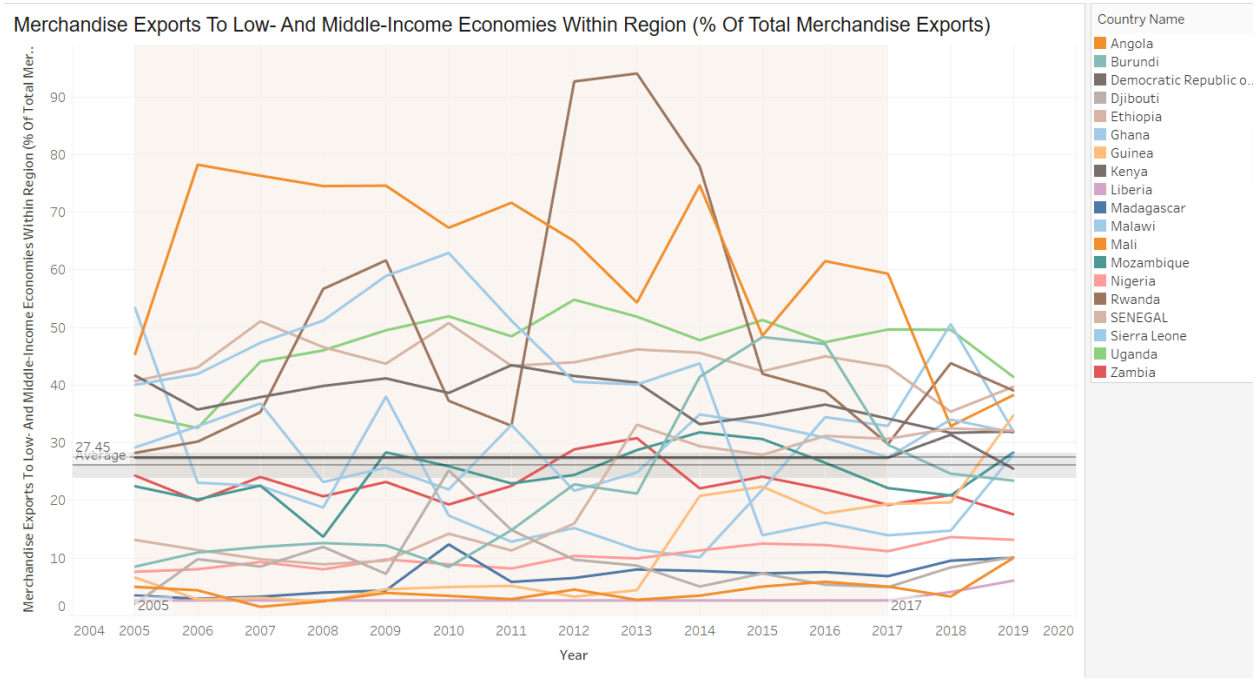
The SAC regional average is much higher than the global average.

Three countries have changed a lot, other countries have not changed much.In particular, the four countries (Kazakhstan,Pakistan,Sri lanka,Bangladesh whose values are below the average have almost no change from year to year, maintaining a similar level.The countries with relatively large changes generally declined first and then increased after 2014 and almost all reached the lowest point in 2016, which may be related to the regional economic environment.

The outliers:

Turkmenistan reached the highest value in 2007, but suffered a sharp decline after 2007, fell below the average after 2009, and reached a trough in 2012, and then it has changed steadily.It may be due to the average economic development of other countries in the regionafter 2007, so Merchandise Exports trade is rarely carried out.

f) Sub-Saharan Africa (SSA)



The changes in various countries are relatively disorderly and different.Several countries below the average line have less change during years.

The outliers:

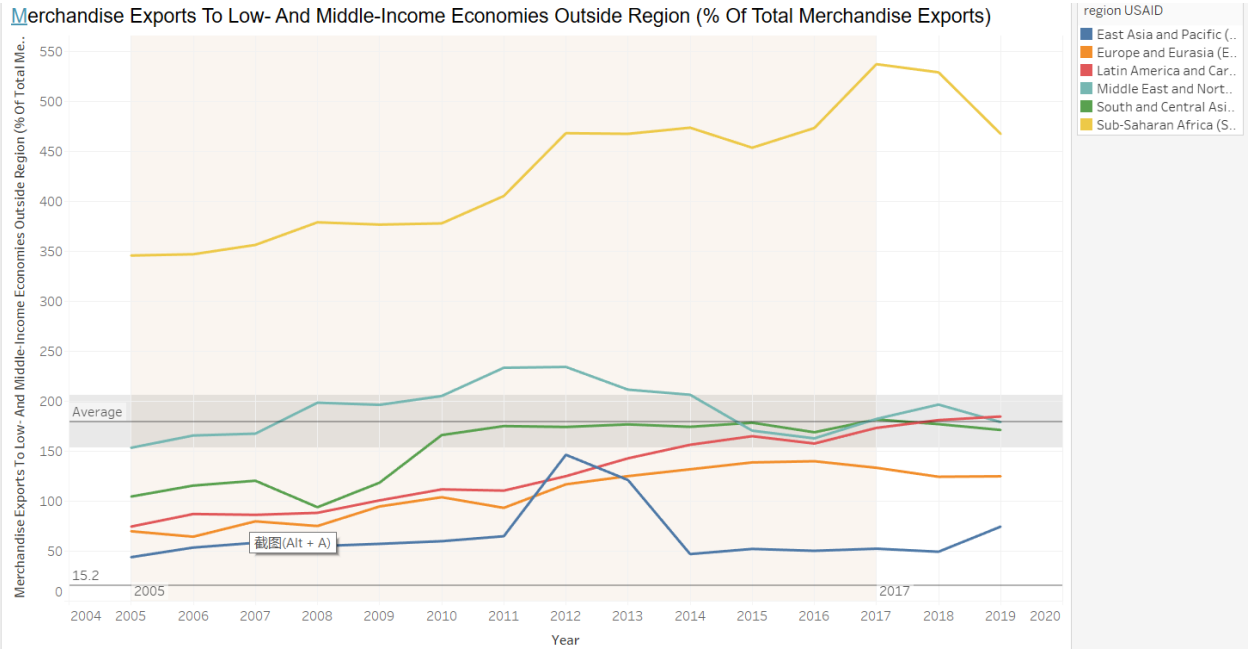
Almost every year, the Merchandise Exports value of Mali is much higher than that of other countries, and Rwanda has risen rapidly after 2019 and has been the highest value in the entire region for two years.

19. Merchandise Exports To Low- And Middle-Income Economies Outside Region (% Of Total Merchandise Exports)

Merchandise exports to low- and middle-income economies outside region are the sum of merchandise exports from the reporting economy to other low- and middle-income economies in other World Bank regions according to the World Bank classification of economies.And the correlation between cpi and price level ratio is negative and low,,which is -0.23. Merchandise Exports To Low- And Middle-Income Economies Outside Region indicator can not affect cpi much better.And the higher the indicator is ,the lower the corruption index is .

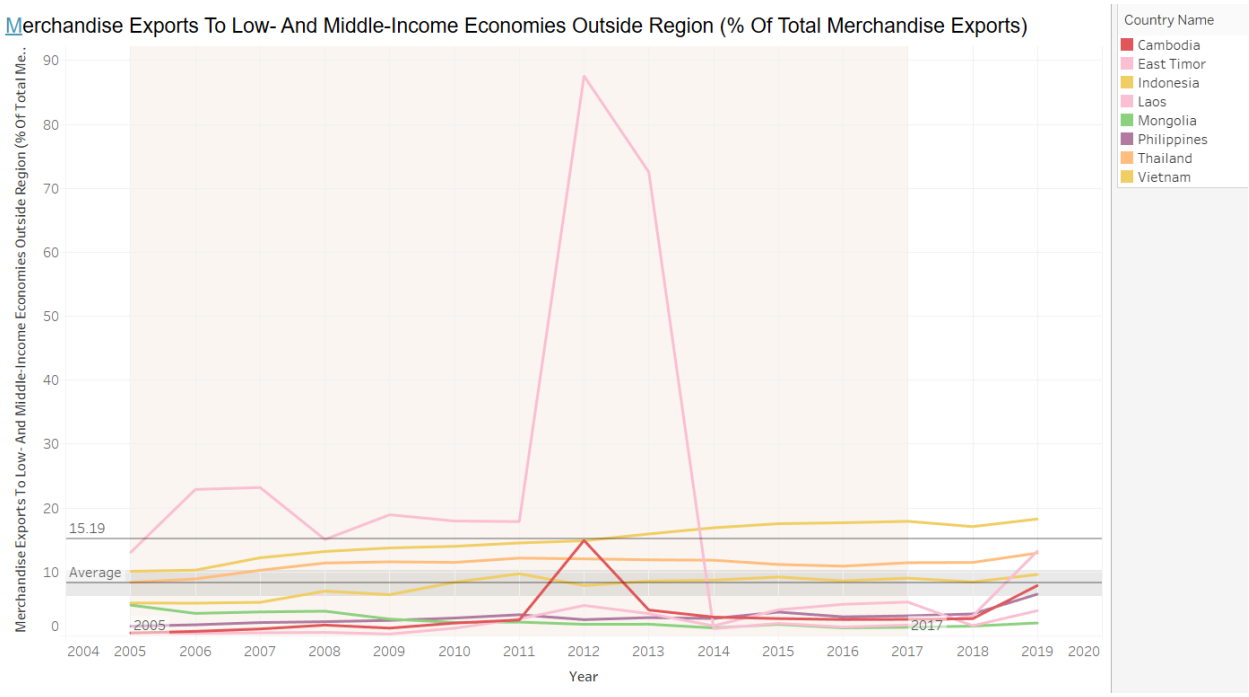
Global average (with 95% CI) is 15.19% over the period of 2005 to 2019.

Regional graph :



Merchandise Exports outside region values of all regions(except SSA) are lower than average from 2005 to 2017.Only SSA is higher than average all years(increase over years) and far exceeds other regions.EAP has a small jump during 2011 and 2014.

a) East Asia and Pacific (EAP)



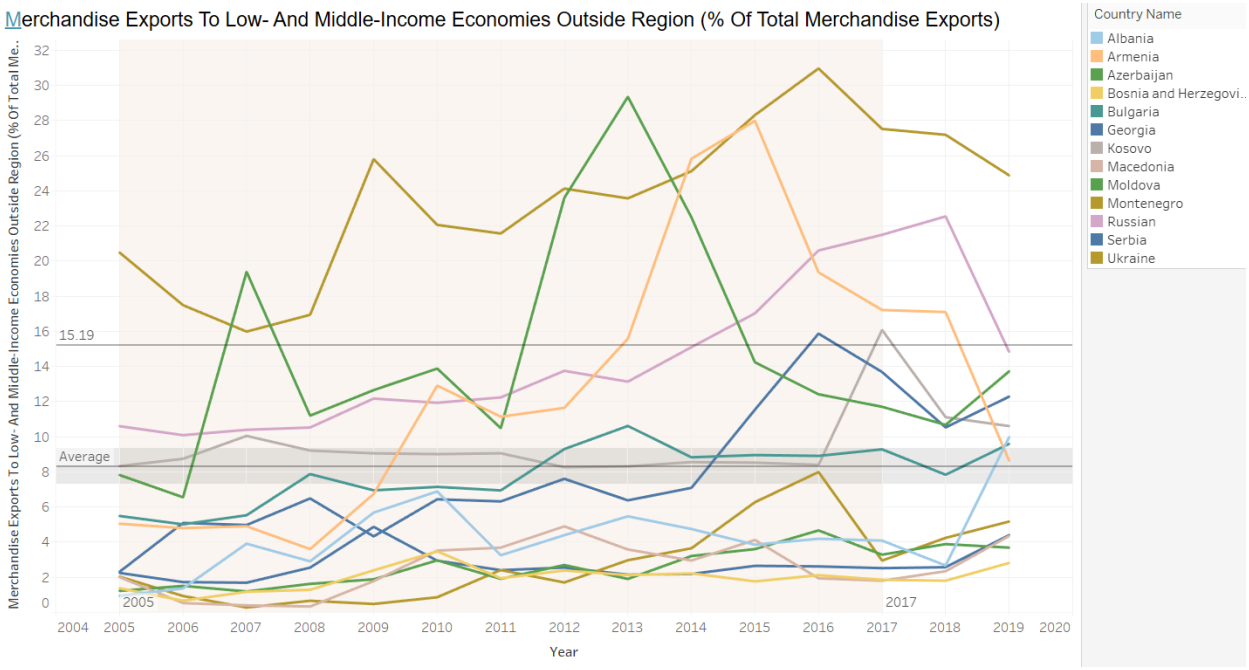
The EAP regional average is a little lower than the global average.

The changes in various countries are very stable, almost unchanged in each year, and almost all are lower than average.

The outliers:

The Merchandise Exports outside region of East Timor is higher than average before 2014. It rose rapidly from 2011 to 2012, reached its peak in 2012, and was higher than all countries in each year, and dropped rapidly after 2012.

b) Europe and Eurasia (E&E)



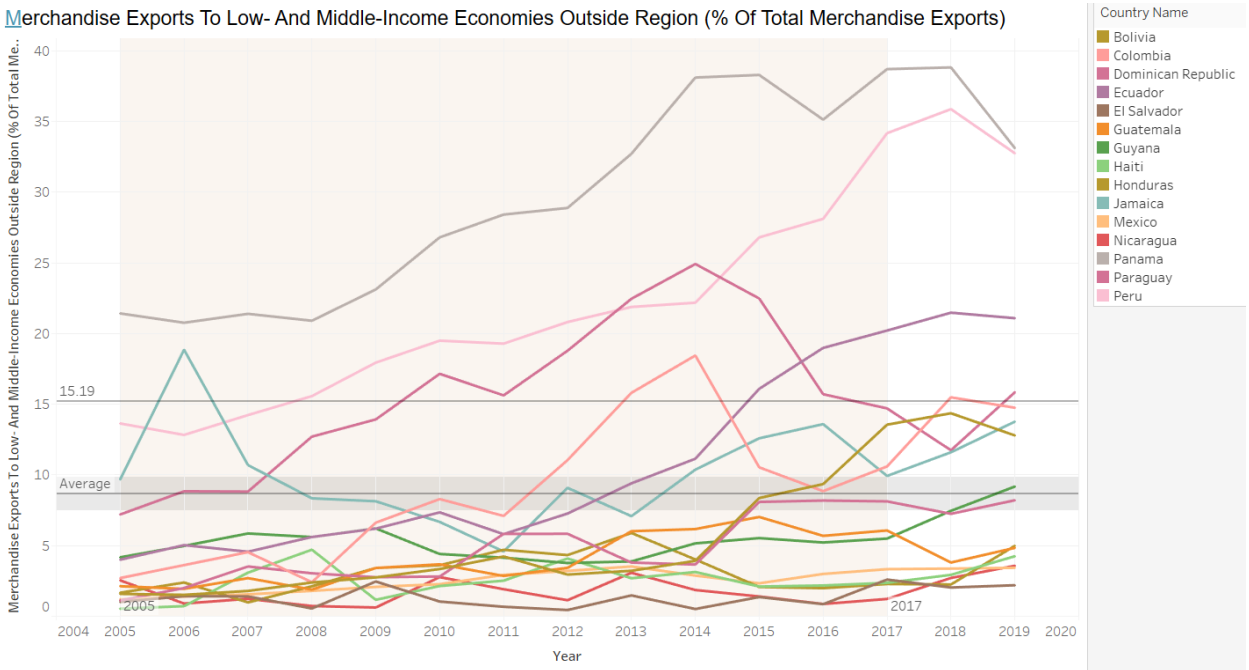
The E&E regional average is much lower than the global average.

The changes in various countries are relatively disorderly and different.Several countries below the average line have less change during years.

The outliers:

The Merchandise Exports outside region value of Ukraine is higher than average value all years.

c) Latin America and Caribbean (LAC)

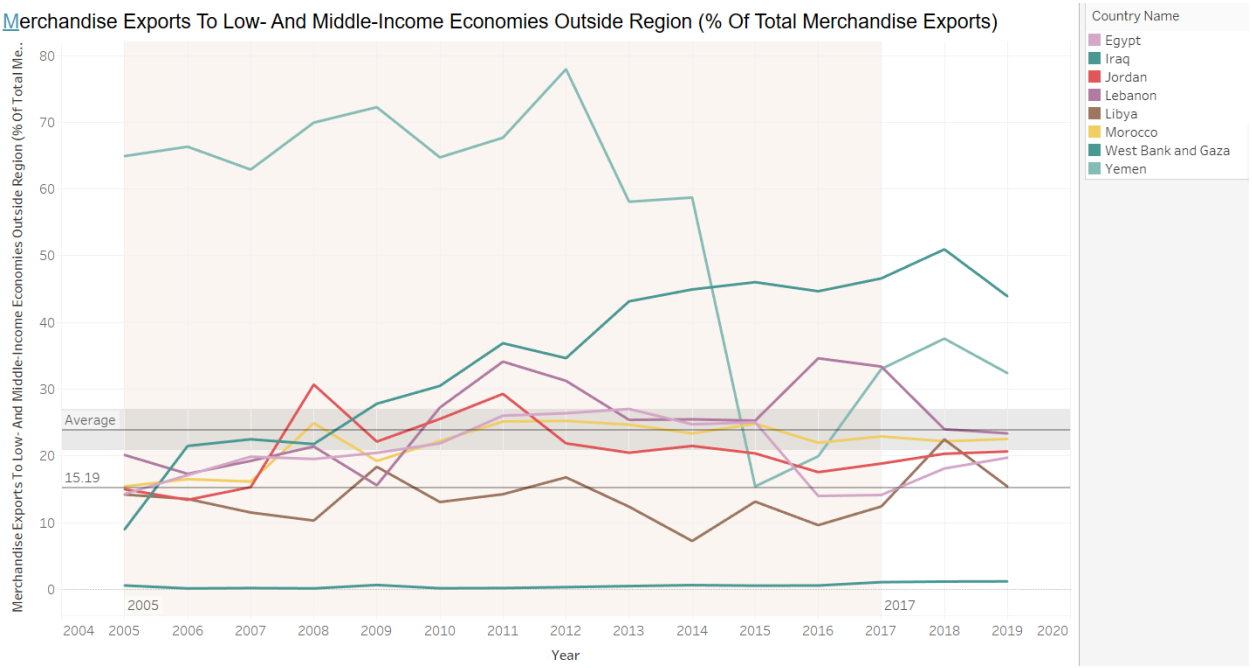


Several countries have increased over years, other countries have not changed much.The overall trend in this area is rising over time.And there are three countries whose values are much higher than the average.

The outliers:

Panama and Peru rose rapidly overall from 2005 to 2017, and the rate of increase was faster than other countries

d) Middle East and North Africa (MENA)



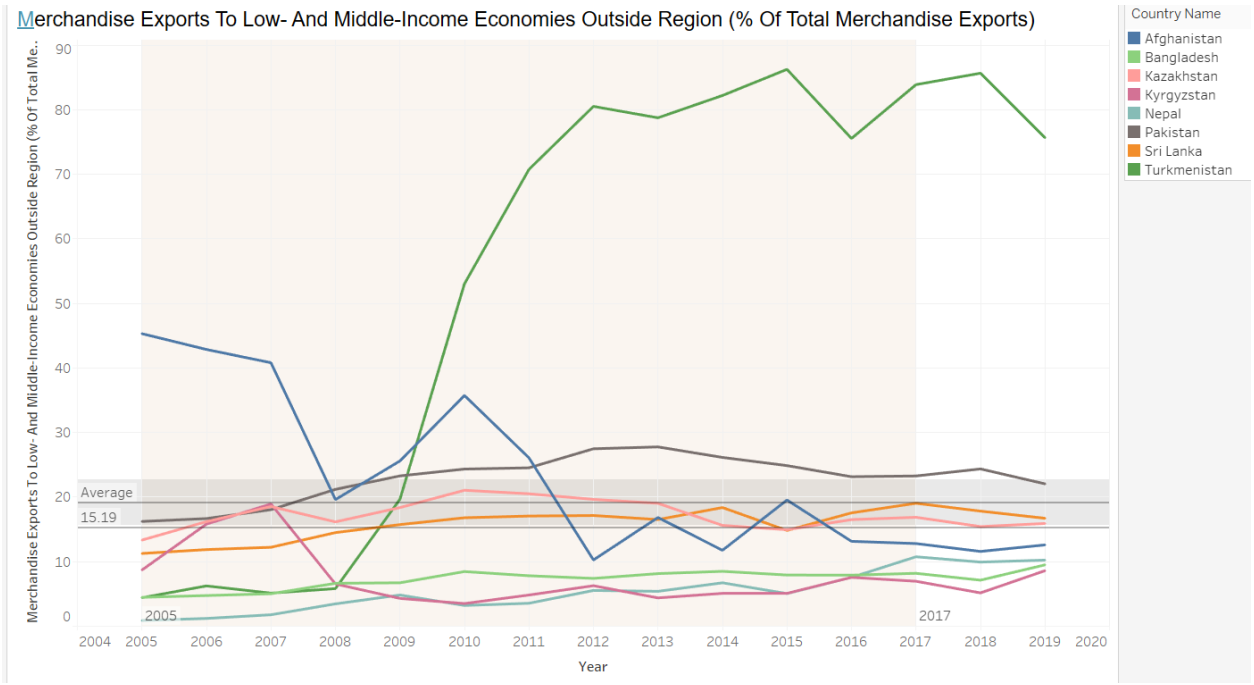
The average of MENA is a little higher than the global average.The countries have not changed much, but have risen slowly in general.

The outliers:

The value outside the region of Yemen is much higher than others before 2013,but sharply decreased after 2013.It shows a little increase after 2015,but not so much.

West bank and Gaza almost did not change during years(near 0)**.Iraq** showed a constant increase during years,which maybe means more corruption (maybe that is why USAID invested much in past report based on the negative correlation between this indicator and cpi)

e) South and Central Asia (SAC)

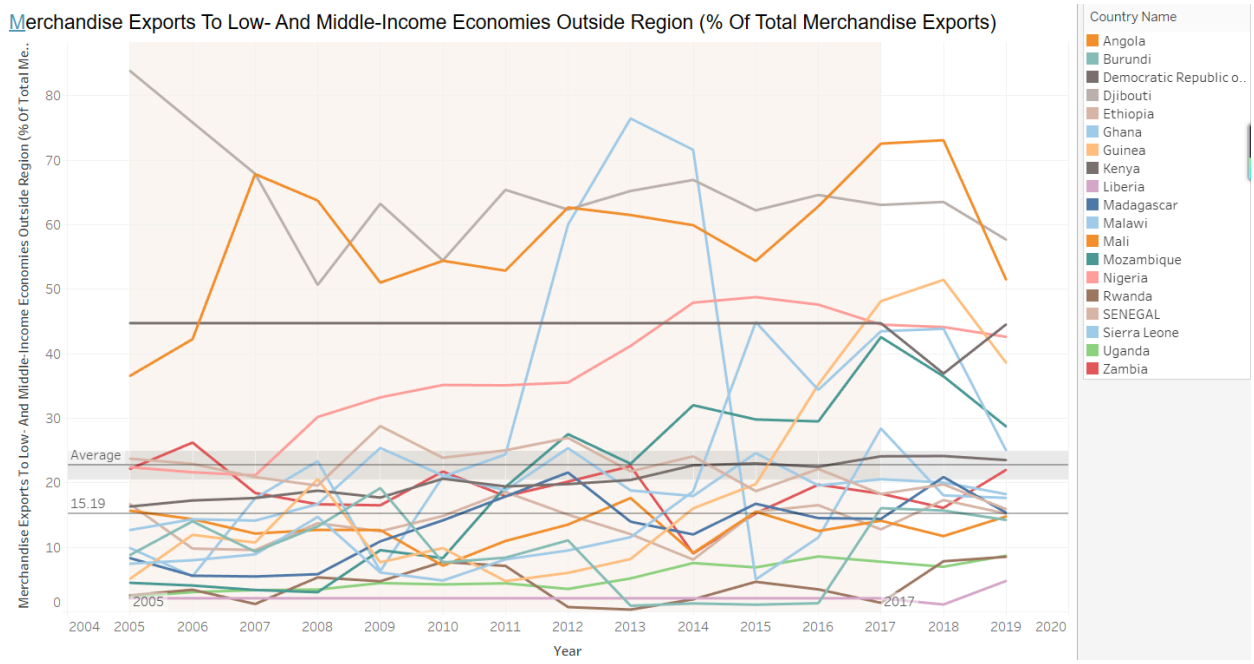


The average of SAC is a little higher than the global average.Almost all countries showed less fluctuation in price level ratio growth as time went by.

The outliers:

Afghanistan and Turkmenistan showed very different changes.Afghanistan dropped from a high value to below the average value before 2012, and Turkmenistan rose rapidly from a very low value to the highest value in the entire region, and in 2012 Has maintained a high export outside region.

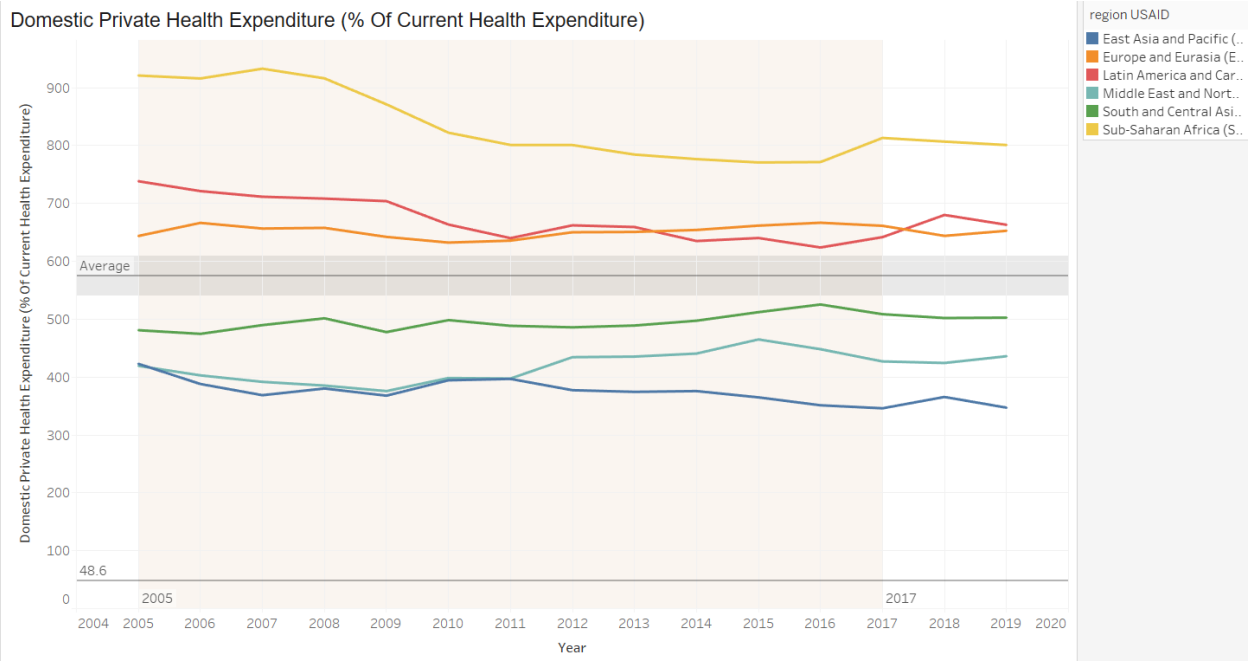
f) Sub-Saharan Africa (SSA)



20. Domestic Private Health Expenditure (% Of Current Health Expenditure)

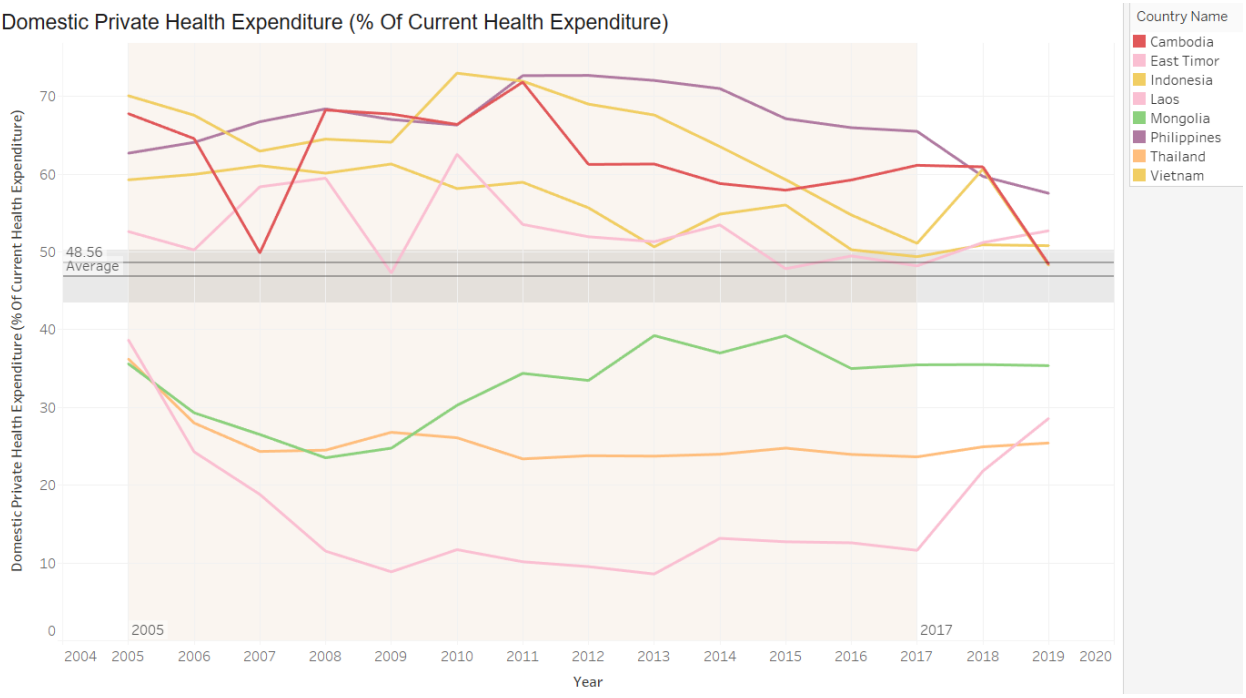
Share of current health expenditures funded from domestic private sources. Domestic private sources include funds from households, corporations and non-profit organizations. Such expenditures can be either prepaid to voluntary health insurance or paid directly to healthcare providers.
The correlation between cpi and gross savings is about -0.26,which is negative.The higher this indicator is ,the lower corruption index is.The higher domestic private health expenditure is ,the better for countries' GDP growth and development.
Global average (with 95% CI) is 48.56% over the period of 2005 to 2019.

Regional graph



All regions have not changed much, and the final value in 2017 is almost equal to the domestic private health expenditure value in 2005.SSA maintains the highest domestic private health expenditure value in the world every year.

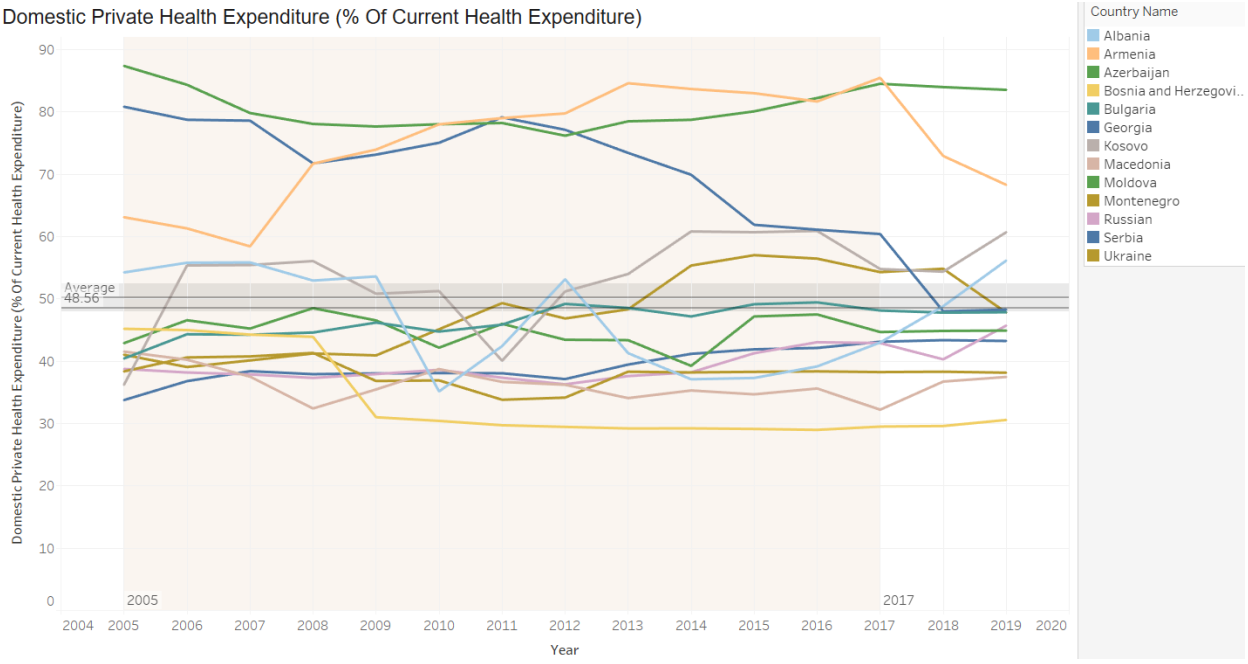
a) East Asia and Pacific (EAP)



The average of EAP is almost similar to the global average.

Three countries are lower than average and others are higher than average over years .And they all have a little change overall.

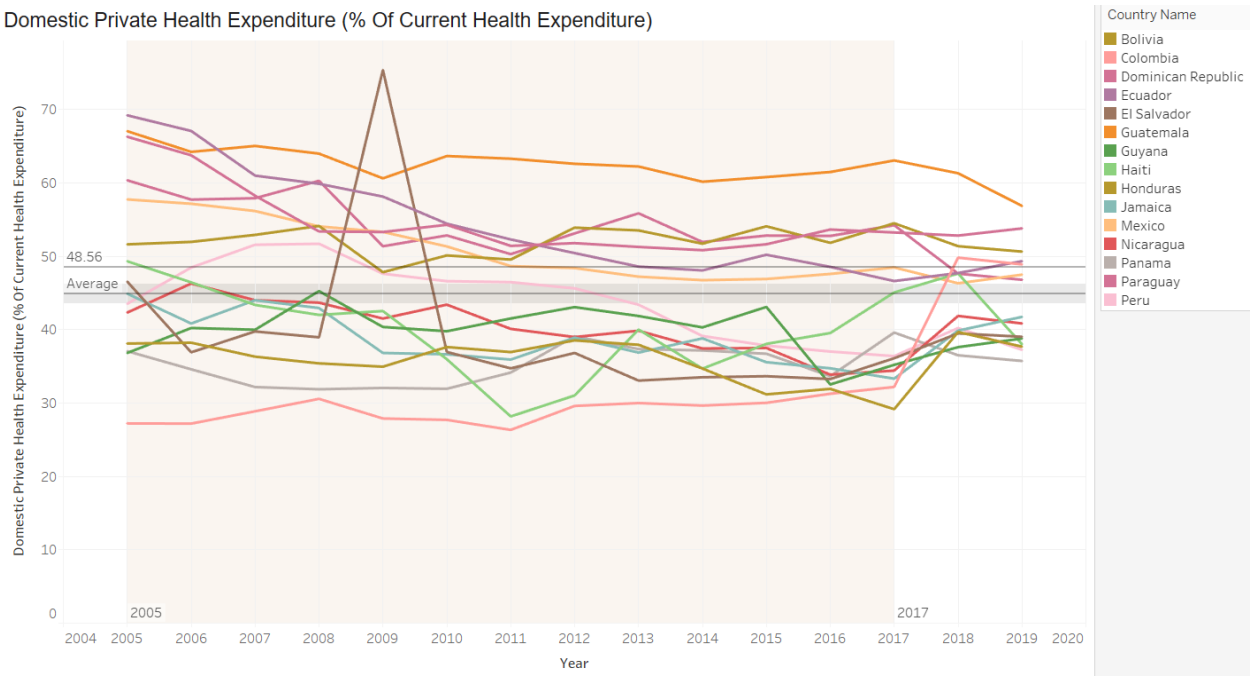
b) Europe and Eurasia (E&E)



The average of this region is almost the same as the global average .

These countries have not changed much as a whole, with small fluctuations between 2005 and 2017

c) Latin America and Caribbean (LAC)

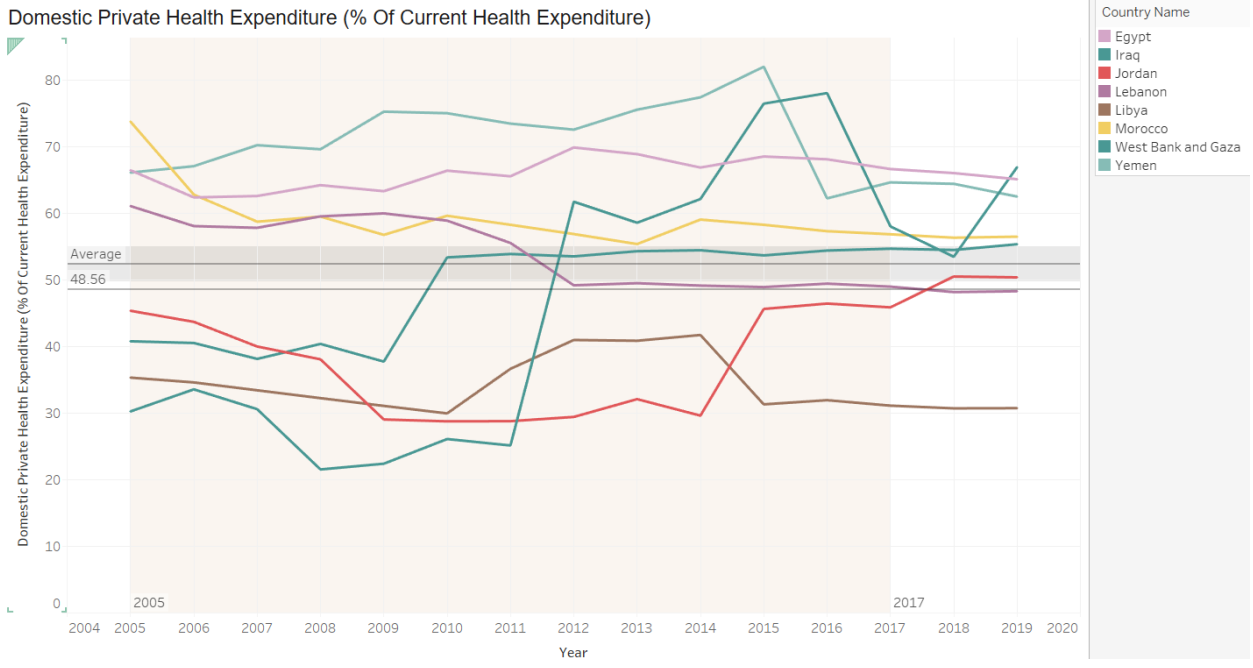


The average of this region is a little lower than the global average .Countries showed less fluctuation in price level ratio growth as time went by.

The outliers:

El Salvador shows a sharp increase and then decrease between 2008 to 2010,which is similar to other indicators.Maybe this country experiences an economic change in these 2 years.

d) Middle East and North Africa (MENA)

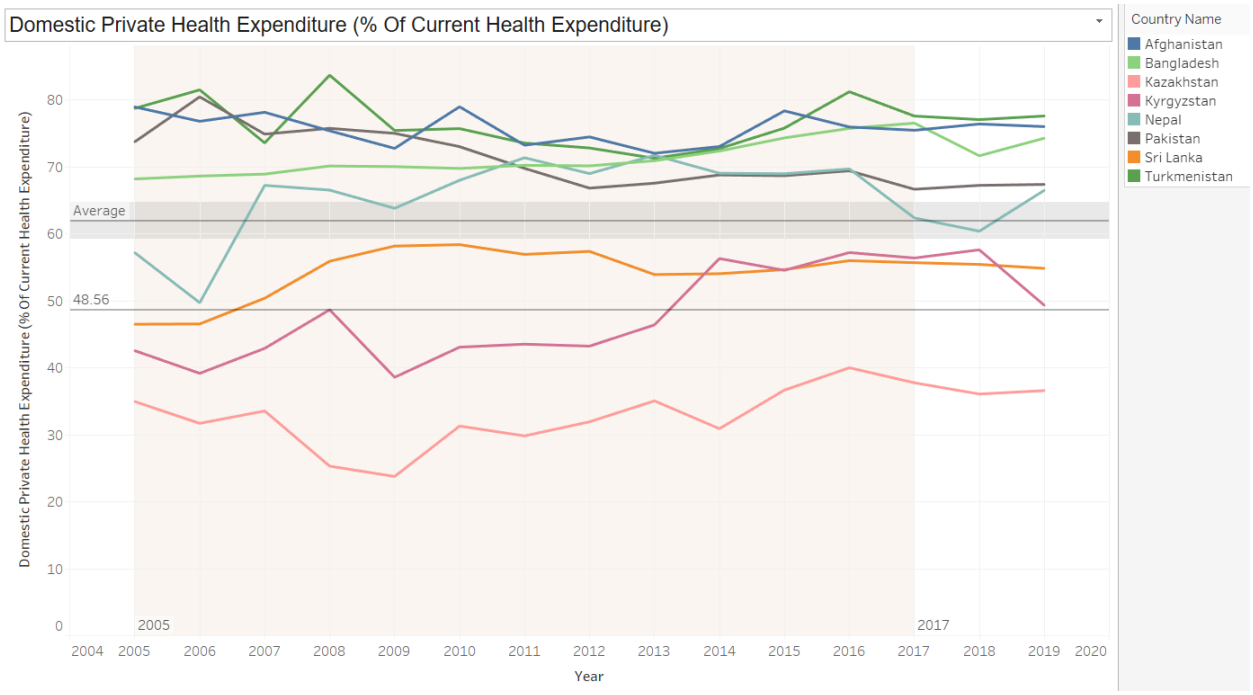


The average of this region is a little higher than the global average .

The outliers:

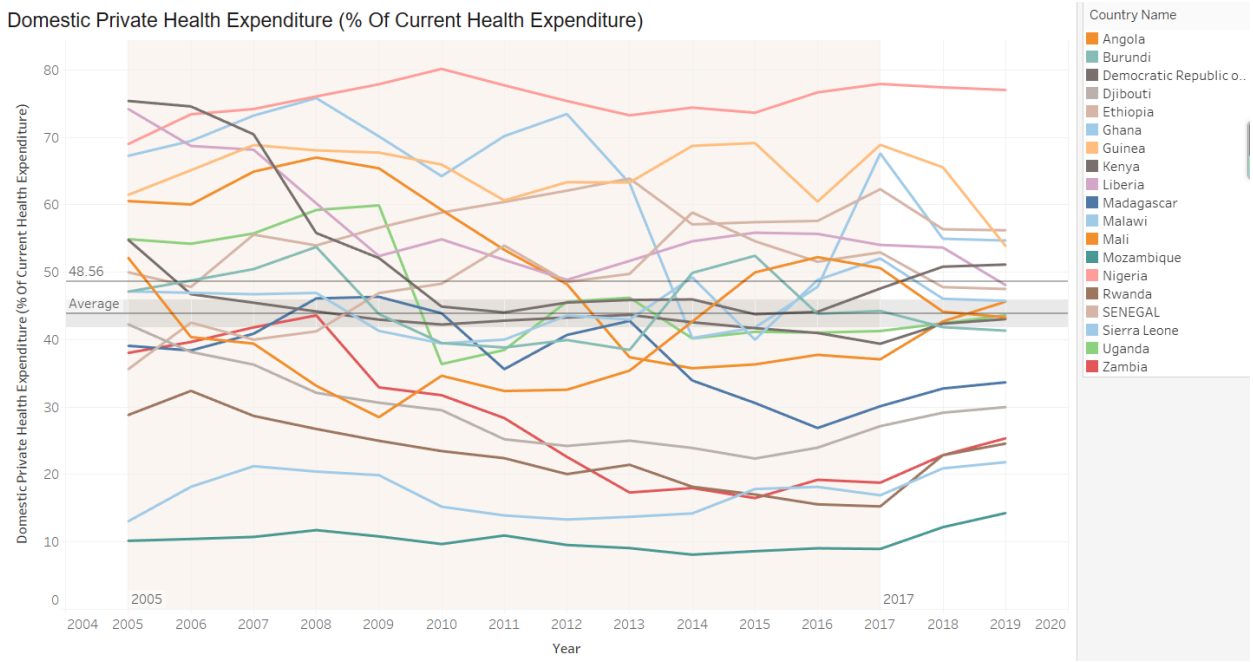
Iraq had a sharp increase during 2011 and 2012,and finally jumped to above average.

e) South and Central Asia (SAC)



The average of SAC region is much higher than the global average.Almost all countries are lower than average and did not change much during the years.

f) Sub-Saharan Africa (SSA)



The average of SSA region is much lower than the global average.The changes in various countries are relatively disorderly and different.Several countries(like Mozambique,Malawi) almost didn't change during years