**Introduction :**

Gross domestic product (GDP) is the ordinary measurement of the value added created through the production of goods & services in a country during a certain time period. It also measures the income earned from that production(manufacturing, development), or the total amount spent on final goods and services. The proposed dataset for the visualization by R and Tableau by the metadata of Australia, USA and India. We have concluded that log of GDP's of India, Usa and Australiya are integrated of order one that make linear structure for forcasting. We then performed Engle-Granger two step test for cointegration. After fitting the cointegrating regression models, the residuals were tested for stationarity by means of ADF test. The results imply that there is an equilibrating mechanism that keeps the countries' GDP from drifting too far apart from each other.

**Scope of the project:** By using this project we have compare the GDP rate of Aus, India and Usa as well as we forecast the GDP rate of Aus, India and Usa. By Visualization we have compare the trend and found the in which year GDP rate is lowest and Hightest. The project has been integrate the uncomplicated, intermediary, questions that are solved and figerout the trend of GDP in USA, India and Australia and the forecast GDP of USA, India and Australia?

**Data source for Visualization**

A metadata is a dataset that contain all information about other data, thus this dataset has taken by Worldbank. In this data set conatin 265 countries GDP information with 1960 to 2017. The data set describe the

|  |  |  |
| --- | --- | --- |
| NY.GDP.PCAP.CD | GDP per capita (current US$) | GDP per capita is gross domestic product divided by midyear population. GDP is the sum of gross value added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products. It is calculated without making deductions for depreciation of fabricated assets or for depletion and degradation of natural resources. Data are in current U.S. dollars. |

|  |  |
| --- | --- |
| Data Source | World Development Indicators |
| Last Updated Date | 4/24/2019 |
|  |  |

In this data set contains attributes associated with Usa,India, Australia and all other countries.The visualization by R project , I have use only Usa,India, Australia GDP to compare the Gdp and Forcast the GDP . that data include year 1960 to 2017, GDP per capita growth.

The dataset will contain only 3 variables, that all are quantitative variables and The dataset contains information from 1960 to 2017. In data set no missing values. The data Australia ,India, Usa was sourced from the Data World Bank website. **Project Plan**

I have investigated the dataset of GDP by " WorldBank" . I have focused on the GDP per capita and tryed to made a model using the data from the dataset. I l also compare the the total GDPs of Aus,Ind and Usa. I have use of EDA and visualization, build model, forcasting,used Arima model.

**Data Preparation**

The data is sourced by the World Bank database, which in excel format. I have changed excel to csv. It has 3 variables and 57 rows, that contain GDP data yearly wise 1960 to 2016.In this data set has no missing values. So now I have exported the data set to R tool that help to analysis and visualization the data.

**Imported required library in R tool for Visualization:**

library(tseries)

library(vars)

library(urca)

library(forecast)

library(zoo)

library(ggplot2)

**Exploratory Data Analysis:** Data set contains 3 variable. By using summary function we have find the EDA**.**

**Importing the data**

setwd("C:\\Users\\SPEED TAIL\\Desktop\\TableauProject")

gdp<- read.csv("mainDataOFGDP.csv", header = T)

> summary(gdp)

Year India Aus USA

Min. :1960 Min. : 82.38 Min. : 1807 Min. : 3007

1st Qu.:1974 1st Qu.: 158.49 1st Qu.: 6472 1st Qu.: 7242

Median :1988 Median : 311.55 Median :14257 Median :21483

Mean :1988 Mean : 476.63 Mean :20100 Mean :24205

3rd Qu.:2002 3rd Qu.: 472.50 3rd Qu.:23437 3rd Qu.:38166

Max. :2016 Max. :1729.71 Max. :68043 Max. :57904

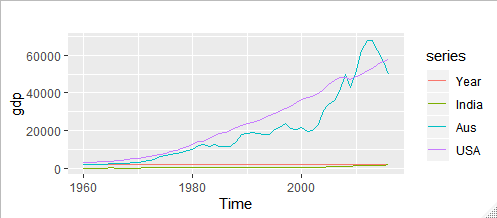
Through exploratory analysis, the data has been summarized that shows the mean, median, quartile in increasing .By using these summarize statics we can analysis of GDP.All 3 variables for the 57 years starting from 1960 to 2017.

**Visualization of GDP:**

library(ggplot2); library(ggfortify)

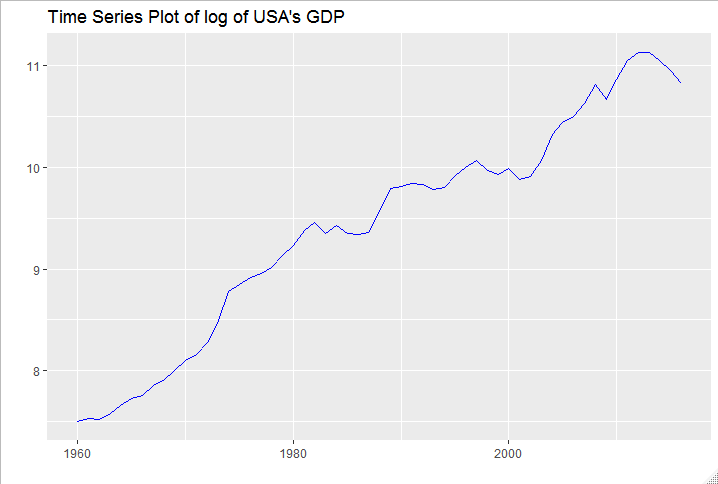
autoplot(gdp)

autoplot(gdp, facets=F)

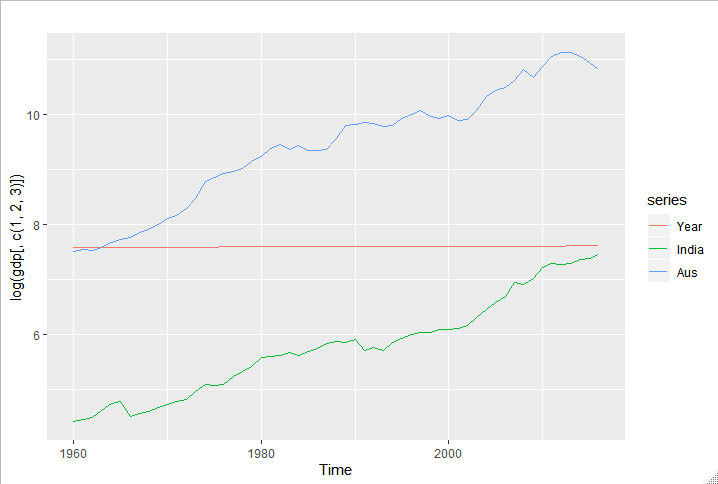


The above plot shows the trend of GDP, which shows an increase trend of GDP.

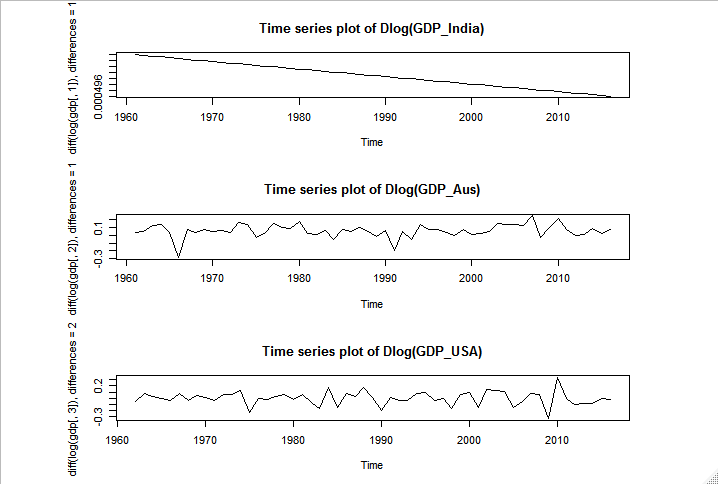
Time series plot for USA GDP:



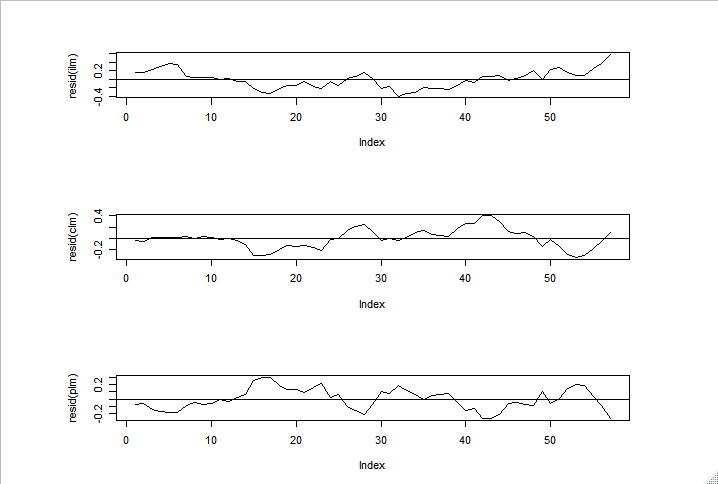
Time series GDP yearly wise between Aus and India:



**Time series Plot:** After log transformation GDP rate



**Residual value**: Actual – predicted value



**Model test:**

> Box.test(residuals(fit), type="Ljung" )

Box-Ljung test

data: residuals(fit)

X-squared = 0.52483, df = 1, p-value = 0.4688

Accuracy test:

> accuracy(fcast, lgdp[,1])

ME RMSE MAE MPE

Training set 1.482908e-05 0.000496759 0.0003705595 0.0001931592

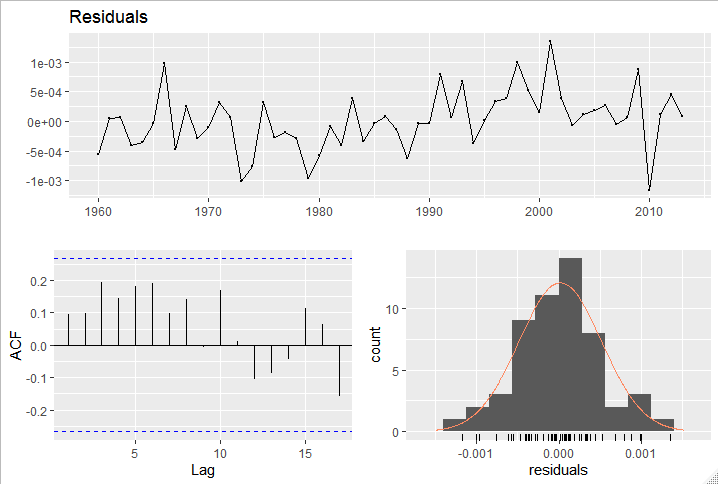
Test set 1.638673e-03 0.001756373 0.0016386730 0.0215373366

MAPE MASE ACF1 Theil's U

Training set 0.004879379 0.7360729 0.09590824 NA

Test set 0.021537337 3.2550308 -0.07224554 4.196847

**Fiitted Model**: By using Arima model I have forcast the GDP and our Arima model gives fitted model



#**deterministic trend**

trend<- seq\_along(train\_p)

model1<- auto.arima(train\_p, d=0, xreg=1960:2014)

summary(model1)

> summary(model1)

Series: train\_p

Regression with ARIMA(1,0,0) errors

Coefficients:

ar1 intercept xreg

0.9799 -53480.81 27.2612

s.e. 0.0183 11444.59 5.7628

sigma^2 estimated as 3031: log likelihood=-298.57

AIC=605.14 AICc=605.94 BIC=613.17

Training set error measures:

ME RMSE MAE MPE MAPE MASE

Training set -4.282003 53.53 36.37953 -7.63386 11.89448 1.054176

ACF1

Training set 0.3045219

**#stochastic trend**

model2<- auto.arima(train\_p, d=1)

summary(model2)

model3<- auto.arima(train\_c, d=1)

summary(model3)

> summary(model2)

Series: train\_p

ARIMA(0,1,4) with drift

Coefficients:

ma1 ma2 ma3 ma4 drift

0.0782 0.1208 0.5816 0.4933 25.6313

s.e. 0.1738 0.1641 0.3069 0.2286 13.2685

sigma^2 estimated as 2152: log likelihood=-283.23

AIC=578.47 AICc=580.25 BIC=590.4

Training set error measures:

ME RMSE MAE MPE MAPE MASE

Training set 0.1897309 43.78554 29.15336 -2.730076 7.95363 0.8447817

ACF1

Training set 0.03750945

**Stochastic trend:**

> summary(model4)

Series: train\_p

ARIMA(3,2,0)

Coefficients:

ar1 ar2 ar3

-0.7423 -0.7354 -0.3754

s.e. 0.1293 0.1301 0.1292

sigma^2 estimated as 2269: log likelihood=-279.06

AIC=566.12 AICc=566.95 BIC=574

Training set error measures:

ME RMSE MAE MPE MAPE MASE

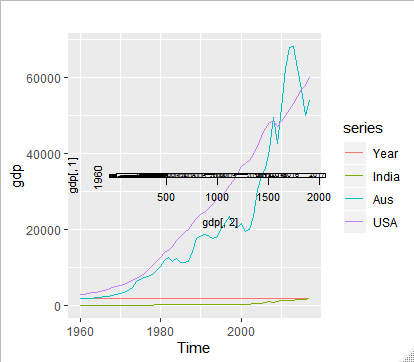
Training set 2.560451 45.41977 28.86137 0.561733 6.990897 0.8363205

ACF1

Training set -0.0384823

The all figures shows the relation between GDP growth with time and year ( y axis ).By R code we can Assess the Relationship between the Economic Growths of Aus, Usa, and India. The statistics indicate disparity in GDP, with Usa having the highest GDP. The standard deviation of Usa GPD is the highest, indicating substantial fluctuation over the time period under study. First step of our analysis consisted of determining the order of integration of the three variables by means of the ADF test. The test was first performed with only a drift term and then with constant and trend for the series in levels and first difference. We have concluded that log of GDP's of India, Usa and Australiya are integrated of order one that make linear structure for forcasting. We then performed Engle-Granger two step test for cointegration. After fitting the cointegrating regression models, the residuals were tested for stationarity by means of ADF test. The results imply that there is an equilibrating mechanism that keeps the countries' GDP from drifting too far apart from each other.

When two variables are cointegrated, there should be granger causality in at least one direction.The results of Toda-Yamamoto Granger causality test show strong evidence of causality running from India to Australiya. But Australiya's GDP does not granger-cause India's GDP. We also found that \*Usa and India together granger-cause Australiya's GDP. One possible reason for this could be the enormous volumes of exports from Usa and India to Australiya. Exports from Australiya to these countries are very small as compared to the imports from these countries.\* Based on the results of the causality analysis, we can conclude that India and Usa affect Australiya's economic growth more than Australiya's economic growth affect those countries'. An VECM (Vector Error Correction Model) was fitted. The fit looks good with a MAPE of 0.64 except that the model is slightly over-predicting.



**Forecasting GDP:**

**By Arima Model:**

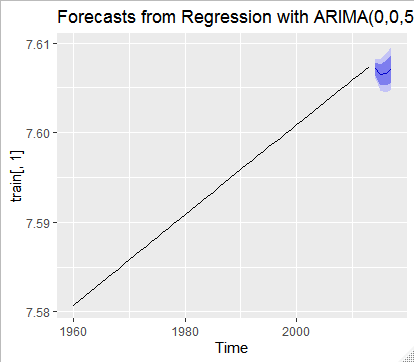
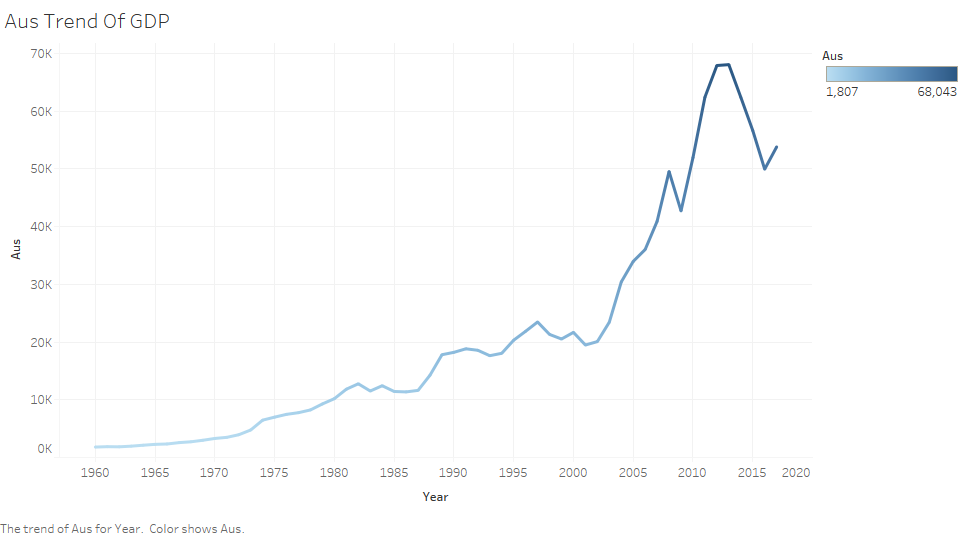
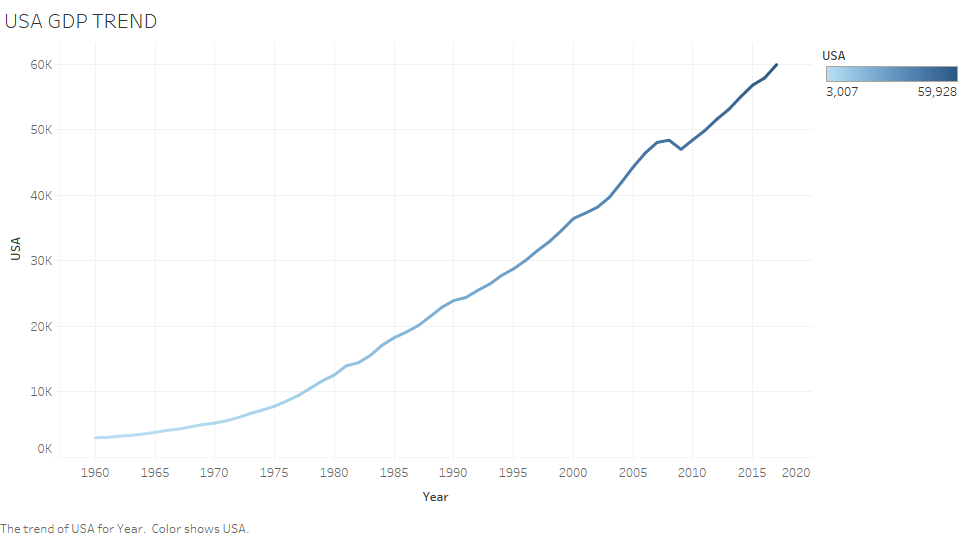


Fig: Forecasting Aus GDP rate

**The visualization show the in future Aus GDP will decrease.**

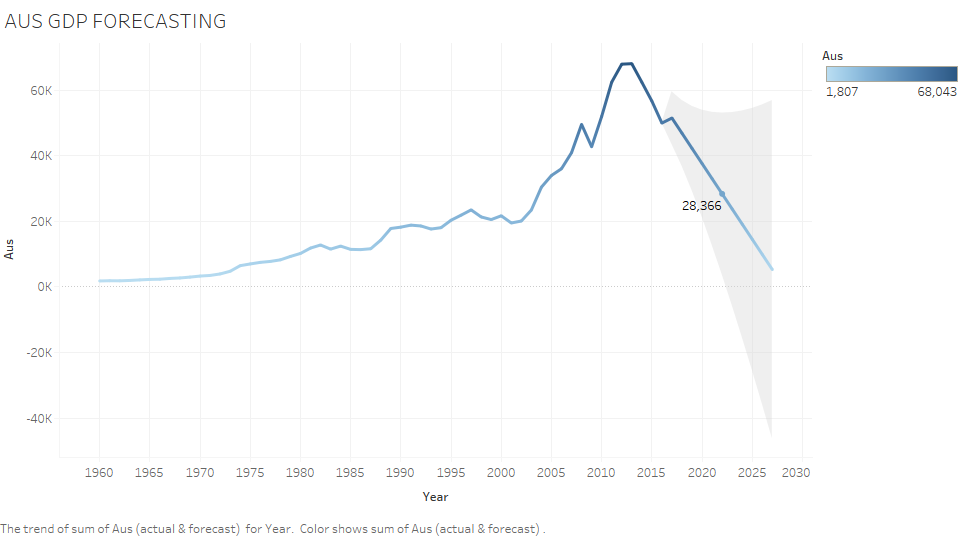
**Visualization By Tableau:**





**Forecasting GDP of Aus and USA:**

**Aus:**



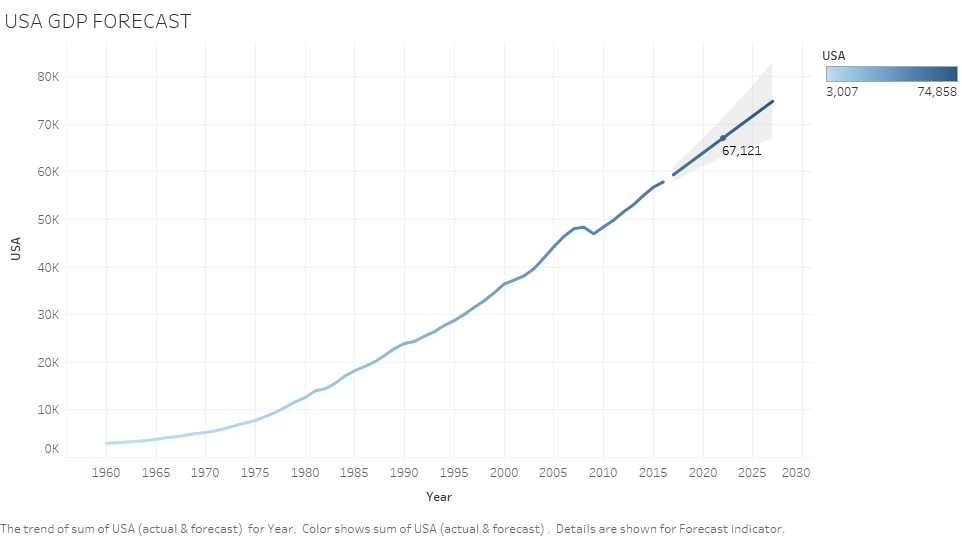
# Options Used to Create Forecasts

|  |  |
| --- | --- |
| **Time series:** | Year |
| **Measures:** | Sum of Aus |
|  |  |
| **Forecast forward:** | 11 periods (2017 – 2027) |
| **Forecast based on:** | 1960 – 2016 |
| **Ignore last:** | 1 period (2017) |
| **Seasonal pattern:** | None (Searched for a seasonal pattern recurring every 1 Periods) |

# Sum of Aus

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Initial** | | |  | **Change From Initial** |  | **Seasonal Effect** | |  | **Contribution** | |  |  |
| **2017** | | |  | **2017 – 2027** |  | **High** | **Low** |  | **Trend** | **Season** |  | **Quality** |
| 51,464 | ± | 8,019 |  | -46,197 |  | None | |  | 100.0% | 0.0% |  | Poor |

GDF forecast for USA:



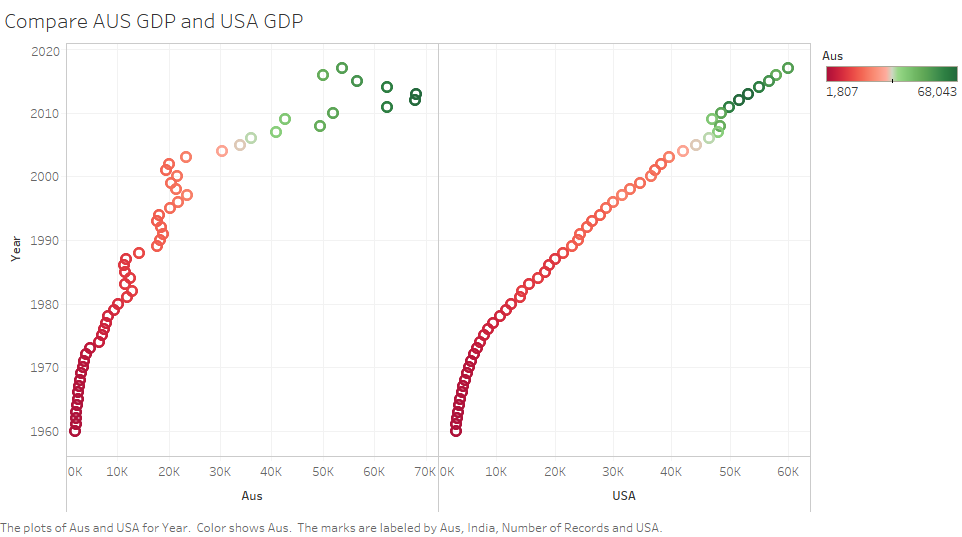
# Options Used to Create Forecasts

|  |  |
| --- | --- |
| **Time series:** | Year |
| **Measures:** | Sum of USA |
|  |  |
| **Forecast forward:** | 11 periods (2017 – 2027) |
| **Forecast based on:** | 1960 – 2016 |
| **Ignore last:** | 1 period (2017) |
| **Seasonal pattern:** | None (Searched for a seasonal pattern recurring every 1 Periods) |

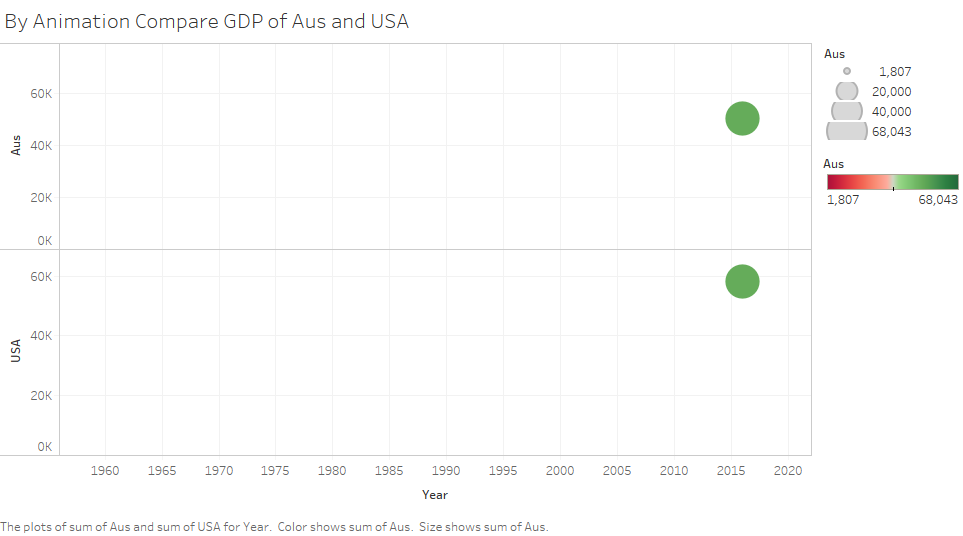
# Sum of USA

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Initial** | | |  | **Change From Initial** |  | **Seasonal Effect** | |  | **Contribution** | |  |  |
| **2017** | | |  | **2017 – 2027** |  | **High** | **Low** |  | **Trend** | **Season** |  | **Quality** |
| 59,385 | ± | 1,614 |  | 15,473 |  | None | |  | 100.0% | 0.0% |  | Ok |

**GDP compare of USA and Aus:**



**Animation of GDP trend of USA:** By tableau we can see yearly wise and can see visual trend of GDP



**References**:

Shaikh, [Anwar](https://www.referenceforbusiness.com/knowledge/Anwar.html) M. *Measuring the Wealth of Nations: The Political Economy of National Accounts.*New York: Cambridge University Press, 1996.

*Survey of Current Business.*Monthly. Washington: Bureau of Economic Analysis, U.S. Department of Commerce.

'Upcoming Comprehensive Revision of the National Income and Product Accounts." *Survey of Current Business,*February 1999, 7.

[Advanced R](https://www.amazon.com/gp/product/1466586966/ref=as_li_tl?ie=UTF8&camp=1789&creative=9325&creativeASIN=1466586966&linkCode=as2&tag=cmdlinetips-20&linkId=3c471c54d8be0d9e74aacdaa6d1cc8de) http://ir-na.amazon-adsystem.com/e/ir?t=cmdlinetips-20&l=am2&o=1&a=1466586966 by Hadley Wickham

### Ggplot2 by Hadley Wickham

### [Data Visualization: A practical introduction](https://amzn.to/2EiCcD3), by Kieran Healy