Preliminary Analysis of Small States

Valerie Cheong Took

`r.Sys.Date()`

Executive Summary

The "Choose Your Own Project" is about studying the economic factors that are key drivers of economic growth in small geographical states relying primarily on their human capital as they do not have natural resources such as oil, minerals. To identify the key drivers, correlation tests between the independent variables(economic indicators) and the response, GDP are carried out as well as regression analysis. The study identifies that a virtuous circle of economic development can be created as is the case of Singapore which has continuously invested in its economic development and has capitalised on its geographical location. It is critical for a country to well manage its finances and not allow a negative disequilibrium between its exports and its imports. A recurrent negative Current Account Balance in the case of Mauritius, has restrained to a large extent its economic development.

1.Introduction

Based on the book, *Breakout Nations In Pursuit Of The Next Economic Miracles*, written by *Ruchir Sharma*, there are key economic indicators used to assess whether a nation is on the path of long term economic prosperity.

According to him, there are four key economic indicators namely:

- **Investment** as a share of GDP, when it stands between 25% and 35%, the economy fully benefits from its positive effects. When investment as a share of GDP falls below 20%, it no longer has a significant positive impact on the economic growth of the country.
- **Population Growth** , if it exceeds the 2% threshold which is the rate of replacement, the economy can potentially grow over the long term as it continues to benefit from a pool of locally available workforce.
- **Manufacturing Sector** needs to account for 20% or more of its GDP to be able to make a significant positive contribution to national economic growth.
- **Current Account** as a share of GDP, should be positive or a deficit of less than 5%. If the current account deficit persistently averages 5% or more of GDP over a number of years, it is a clear indication that the economy is going into an economic recession.

Another economic indicator, not mentioned by Ruchir Sharma, is *the level of economic complexity* . This determines how far can a country prosper whenever it experiences an economic boom.The more an economy is complex, it implies that its economic linkages are strong and can facilitate knowledge transfer, allowing for higher chances of a better distribution of economic growth. Messrs Yuriy Gorodnichenko and Dominick Bartelme in their study, *Linkages and Economic Development* (2015) were able to demonstrate that the strength of the economic linkages was positively correlated to measured output per worker and total factor productivity. Economic linkages allow for a more sustained economic development as industrial and technical know-how spreads to various parts of the economy facilitating the country's process of learning on how to produce (and export) more sophisticated and complex products. New literature has revealed the role of accumulating capabilities (that is, knowledge and the ability to diffuse this knowledge across various economic sectors) in the production of more complex goods.

To verify these assumptions, 3 countries of the same income class (high income) have been chosen: Singapore, Mauritius and Malta. These three countries share other similarities as well. They are all islands capitalising on their human resources and geographical location, with little natural resources. All three islands are of relatively small geographical size. Mauritius is a newcomer to this income class and it has been chosen to be compared to the other two countries (Singapore and Malta) who have proven their ability to avoid the "middle income trap".

It would be interesting to understand whether the key drivers of economic growth as explained by Mr Sharma have all contributed to the economic development of these three countries, especially Singapore. Singapore ranked 6th out of 133 countries in 2022 according to the Economic Complexity Index. During the last 20 years, the country's economy has improved its economic complexity, as its ranking went up from 20th to 6th. Mauritius ranked 90 out of 96 in 2021 in terms of economic complexity. Malta is, unfortunately, not included in this ranking of 133 countries.

2.Sourcing of Economic Data

The economic data from World Bank have been used for this study.The data used relates to the period 1979 to 2019. More recent years have been excluded due to the impact of COVID which would have distorted the analysis. The following economic indicators have been selected.

- **GDP in constant 2015 USD** have been used so as to be comparable across the three different countries. Its code is *NY.GDP.MKTP.KD*

- **Gross Capital Formation** as a % of GDP have been taken as the proxy for the level of domestic investment. Its code is *NE.GDI.TOTL.ZS*
- **Foreign direct investment, net inflows** as a % of GDP. Its code is *BX.KLT.DINV.WD.GD.ZS*
- **Manufacturing** as a % of GDP. Its code is *NV.IND.MANF.ZS*
- **Current Account** as a % of GDP. Its code is *BN.CAB.XOKA.GD.ZS*
- **Working Population**, population ages 15 to 64 as a % of total population. Its code is *SP.POP.1564.TO.ZS*

3.Methodology

Two approaches will be used in this study to test whether the chosen economic indicators have any significant influence on the economic growth of the selected sample of countries

A.Correlation Tests

The first and simpler one is to carry out correlation tests between the selected indicators and GDP.If the correlation tests are positive, it implies that they contribute positively to GDP growth. The **Spearman** correlation test has been chosen.It is used to measure the strength of the relationship and does not assume a linear relationship. It is appropriate for both continuous and ordinal data and is robust to outliers. It calculates the correlation based on the ranks of the data points.

B.Regression Analysis

Correlation quantifies the association between a single pair of variables rather than a larger number of variables. It may be the case that the synergies between various economic indicators drives economic growth rather than it be the result of one single economic indicator. So, effecting a regression analysis with all economic indicators and finding out which subset works best for each country will also be part of this study.

B.1 Method Of Selection Of The Economic Indicators For The Regression Analysis

Backward selection will be applied to select the subset of economic indicators that are the right fit. All the four economic indicators are, therefore, included in the first model and then the economic indicator with the largest p value is removed – that is, the variable that is the least statistically significant. This procedure continues until all remaining variables have a p value of 1% or less.

B.2 Interpretation Of The Regression Analysis

The model is the best fit when the p value is small - 1% and 5%.

Another way to counter verify is to check the Adjusted R Squared. It is a more robust indicator than R Squared. R Squared will always increase when more variables are added to the model, even if those variables are only weakly associated with the response. This is why we focus on the adjusted R squared. It looks at whether additional input variables are contributing to the model. The adjusted R square can decrease if a new predictor does not improve the model's fit.

We can also plot the residuals versus the predicted values of GDP. Ideally, the residual plot will show no discernible pattern. The presence of a pattern may indicate a problem with some aspect of the model.

Another indicator that we can use to assess the regression model is the Akaike Information Criterion (AIC) for small datasets. According to the AIC, the best fit model is the one that explains the greatest amount of variation using the fewest possible independent variables. Lower AIC scores are better. AIC is a better indicator when it comes to model comparison.

4.Exploratory Analysis

A. Dataset Creation

The package *wbstats* is used to download the data from the World Bank Database. We first define the list of countries and the set of economic indicators (independent variables) for this study.

```
countrylist <- c("Mauritius" = "MUS", "Malta" = "MLT", "Singapore" = "SGP")
# Grouping all the four economic indicators into one single vector
my_indicators <- c("GDP" = "NY.GDP.MKTP.KD", "Domestic Investment" = "NE.GDI.TOTL.ZS", "FDI" = "BX.KLT.I

# downloading the data
df <- wb_data(indicator = my_indicators, country = countrylist, start_date = 1976, end_date = 2019)
```

B.Data Wrangling

All data relevant to each country is selected and saved into one vector for each country.

```
# Singapore Data
df_singapore <- df %>%
  filter(country == "Singapore")
# Malta Data
df_malta <- df %>%
  filter(country == "Malta")
# Mauritius
df_mauritius <- df %>%
  filter(country == "Mauritius")
```

5.Data Pre processing and Results of the Analysis

We proceed country by country starting with Singapore. The various R coding for the correlation tests for Singapore are also included in the following section.

A.Singapore

A.1 Exploratory Analysis

```
summary(df_singapore)
```

```
##     iso2c              iso3c              country                 date     
##  Length:44          Length:44          Length:44          Min.   :1976  
##  Class :character   Class :character   Class :character   1st Qu.:1987  
##  Mode  :character   Mode  :character   Mode  :character   Median :1998  
##                                                           Mean   :1998  
##                                                           3rd Qu.:2008  
##                                                           Max.   :2019  
##  Current Account        FDI          Domestic Investment Manufacturing  
##  Min.   :-13.136   Min.   : 3.646   Min.   :17.22       Min.   :17.47
```

```
##  1st Qu.:  1.281   1st Qu.: 8.678   1st Qu.:27.13       1st Qu.:20.69
##  Median : 14.980   Median :13.579   Median :33.16       Median :23.01
##  Mean   : 10.652   Mean   :14.315   Mean   :32.64       Mean   :22.72
##  3rd Qu.: 17.972   3rd Qu.:20.599   3rd Qu.:36.72       3rd Qu.:24.74
##  Max.   : 27.143   Max.   :29.764   Max.   :46.93       Max.   :27.12
##       GDP               Working Pop
##  Min.   :2.425e+10   Min.   :64.61
##  1st Qu.:5.140e+10   1st Qu.:72.07
##  Median :1.233e+11   Median :74.72
##  Mean   :1.450e+11   Mean   :74.07
##  3rd Qu.:2.159e+11   3rd Qu.:77.05
##  Max.   :3.498e+11   Max.   :78.82
```

We can note that the level of domestic investment for Singapore has largely been above the minimum threshold of 25% given that the mean level is 32.64% of GDP whilst it has reached a maximum level of 46.93%. Its manufacturing sector has also exceeded the required 20% level with its mean to be at 22.72%. On the other hand, FDI as a % of GDP was not as high as one would expect given that Singapore is one of the most open economies and has adopted a low tax strategy.

It would be interesting to know for how many years has Singapore been able to maintain its level of domestic investment and of the manufacturing sector above the minimum required threshold.

Deep Dive Analysis

```r
# Number of years that the manufacturing sector's contribution is equal to or exceeds 20%
df_singapore_manf <- df_singapore %>%
  filter(Manufacturing >= 20)

nrow(df_singapore_manf)
```

```
## [1] 36
```

```r
# Number of years the level of domestic investment equals to or exceeds 35%
df_singapore_dominv <- df_singapore %>%
  filter(`Domestic Investment` >= 35)

nrow(df_singapore_dominv)
```

```
## [1] 18
```

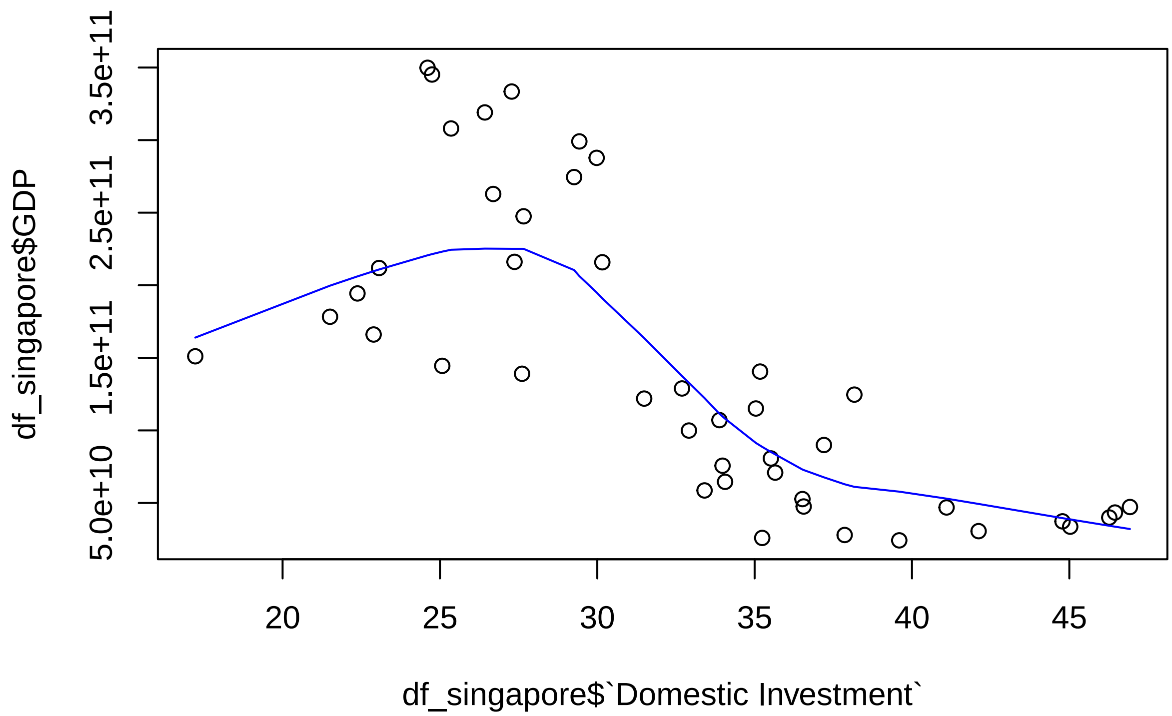
```r
# Number of years current account deficit exceeds or equals to 5%
df_singapore_currentacc <- df_singapore %>%
  filter(`Current Account` <= -5)

nrow(df_singapore_currentacc)
```

```
## [1] 6
```

Further investigation shows that 36 years out of the 43 years under review, Singapore's manufacturing sector contributed more than 20% to the country's GDP. Similarly, the level of domestic investment has exceeded the 35% threshold level for 18 years. In addition, its current account deficit exceeded 5% only in 11 years out of the 43 year period. These results indicate that the country has, most of the time, exceeded or attained the different threshold levels. This may, therefore, explain the economic performance of Singapore being consistently one of the world's top performers.

Exploratory Graphics of Independent Variables with GDP

```
plot(df_singapore$`Domestic Investment`, df_singapore$GDP)
lines(lowess(df_singapore$`Domestic Investment`, df_singapore$GDP), col = "blue")
```



The plot above seems to indicate a relationship between Domestic Investment and GDP, though it is not linear. The graphic slightly resembles a bell shaped curve skewed towards the right. In this case, if the variable, Domestic Investment, remains as is, its largest values may have a larger leverage in determining the regression coefficient than those closer to the mean. When the values have been log transformed, the leverage of the points with largest values remain large but are less dominating.

```
plot(df_singapore$FDI, df_singapore$GDP)
lines(lowess(df_singapore$FDI, df_singapore$GDP), col = "red")
```

```
plot(df_singapore$`Current Account`, df_singapore$GDP)
lines(lowess(df_singapore$`Current Account`, df_singapore$GDP), col = "green")
```

```
plot(df_singapore$`Working Pop`, df_singapore$GDP)
lines(lowess(df_singapore$`Working Pop`, df_singapore$GDP), col = "grey")
```

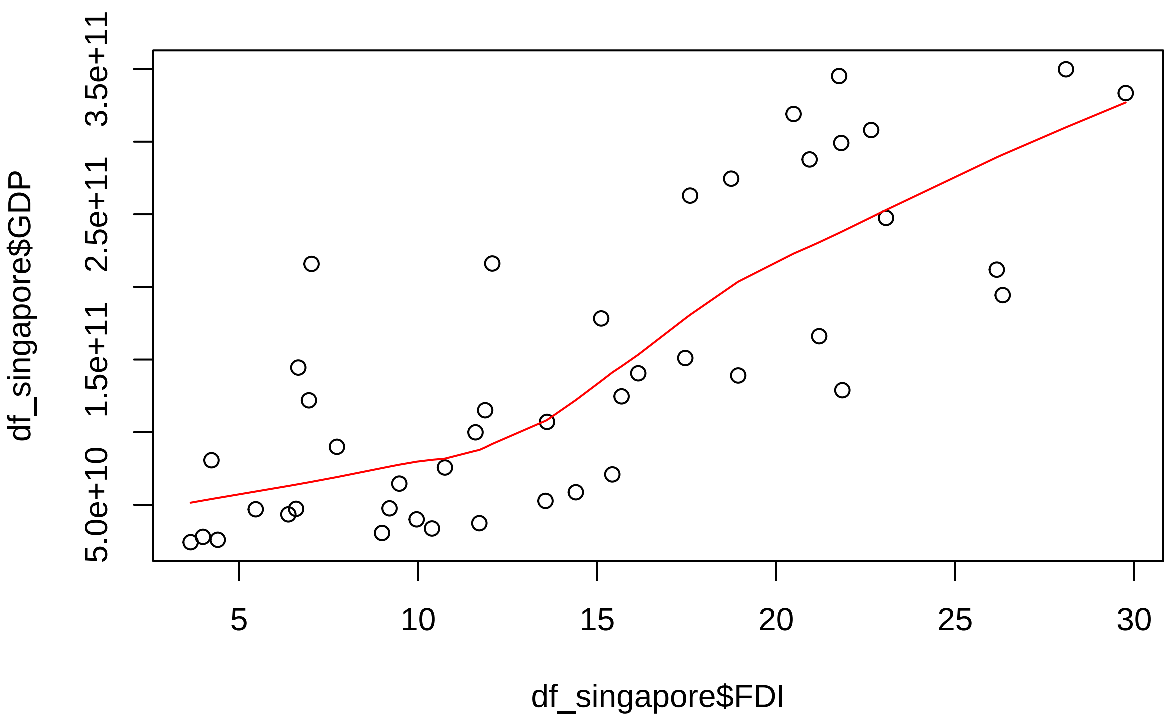


Figure 1: Exploratory Graphics of Independent Variables with GDP

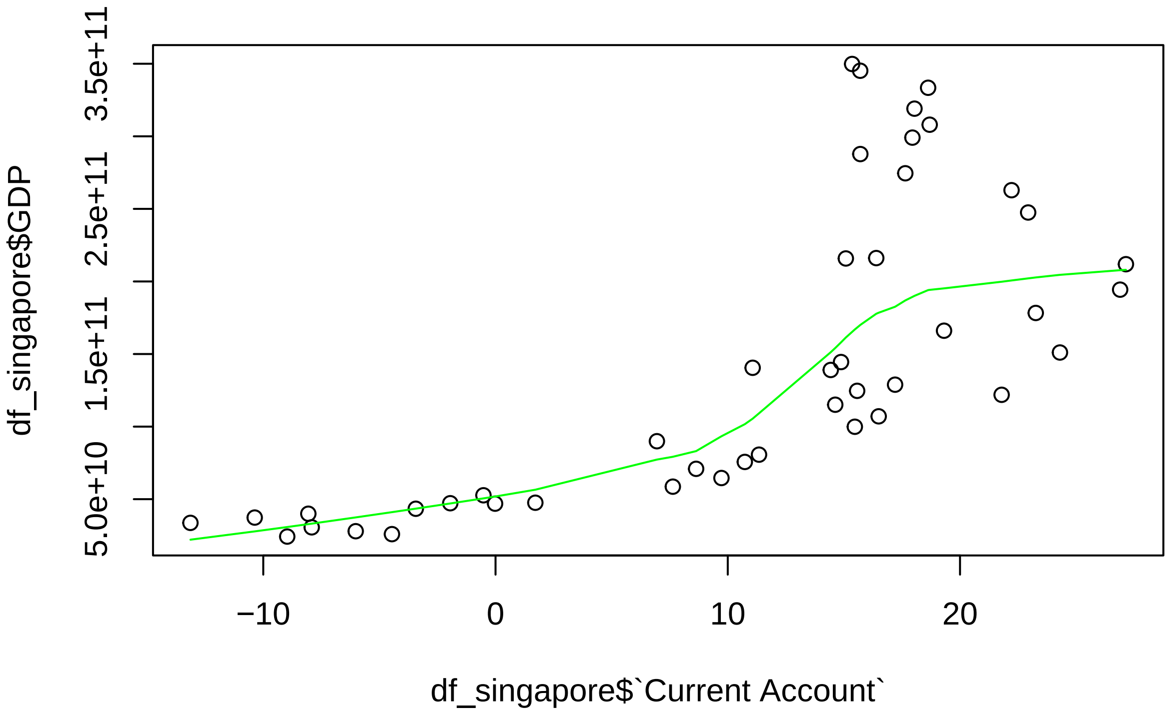
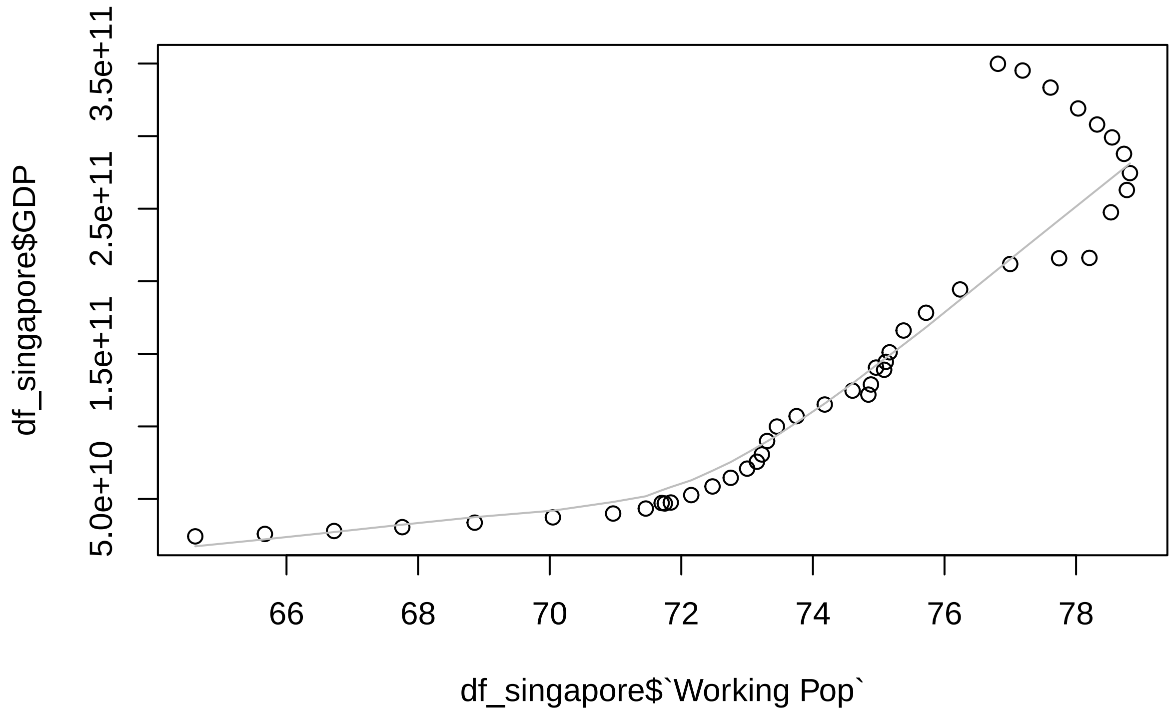
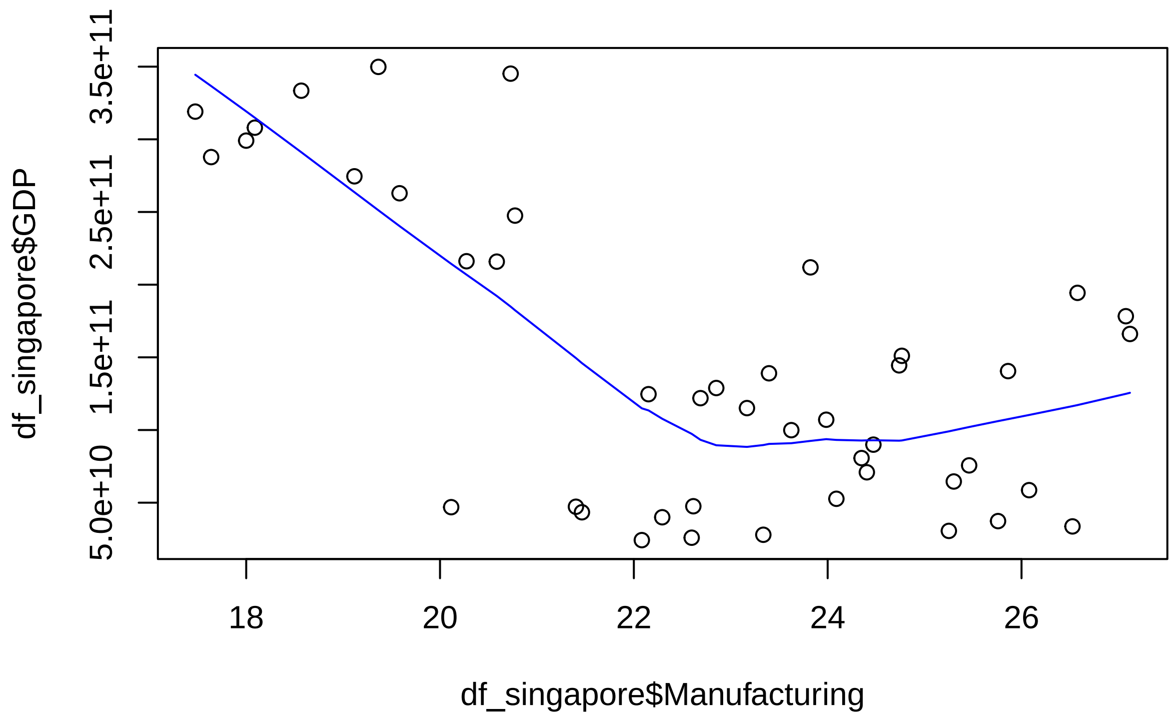


Figure 2: Exploratory Graphics of Independent Variables with GDP



The above graphics indicate that all three predictors (FDI, Current Account and Working Population) have a positive linear relationship with the response, GDP. However, in the case of Working Population, the relationship seems to turn negative at a certain point in time due to the diminishing returns to labour or due to higher strain on resources. Adding more labour to the workforce without investing into more equipment will not generate as much revenue as it is expected. Also, as Singapore is a small country and population is on the rise, there is more pressure on the existing infrastructure and on housing so that this may hinder economic growth. We would need to investigate further to confirm what is affecting negatively output per labour.

```
plot(df_singapore$Manufacturing, df_singapore$GDP)
lines(lowess(df_singapore$Manufacturing, df_singapore$GDP), col = "blue")
```



In regards to Manufacturing, the relationship with GDP seems to be cyclical. This is expected as businesses tend to expand when the economy is growing and maintain status quo or reduce their production capacity when the economy contracts. The variable Manufacturing will also need to be transformed when we come to the regression analysis.

We now proceed with the correlation tests.

A.2 Correlation Tests

Below are a series of correlation tests between GDP and the other five economic indicators.

```
attach(df_singapore)

# FDI and GDP
cor(FDI, GDP, method = "spearman")
```

```
## [1] 0.7852008
```

```
# Domestic Investment and GDP
cor(`Domestic Investment`, GDP, method = "spearman")
```

```
## [1] -0.8317125
```

```r
# Manufacturing and GDP
cor(Manufacturing, GDP, method = "spearman")
```

```
## [1] -0.4434109
```

```r
# Working Population and GDP
cor(`Working Pop`, GDP, method = "spearman")
```

```
## [1] 0.96716
```

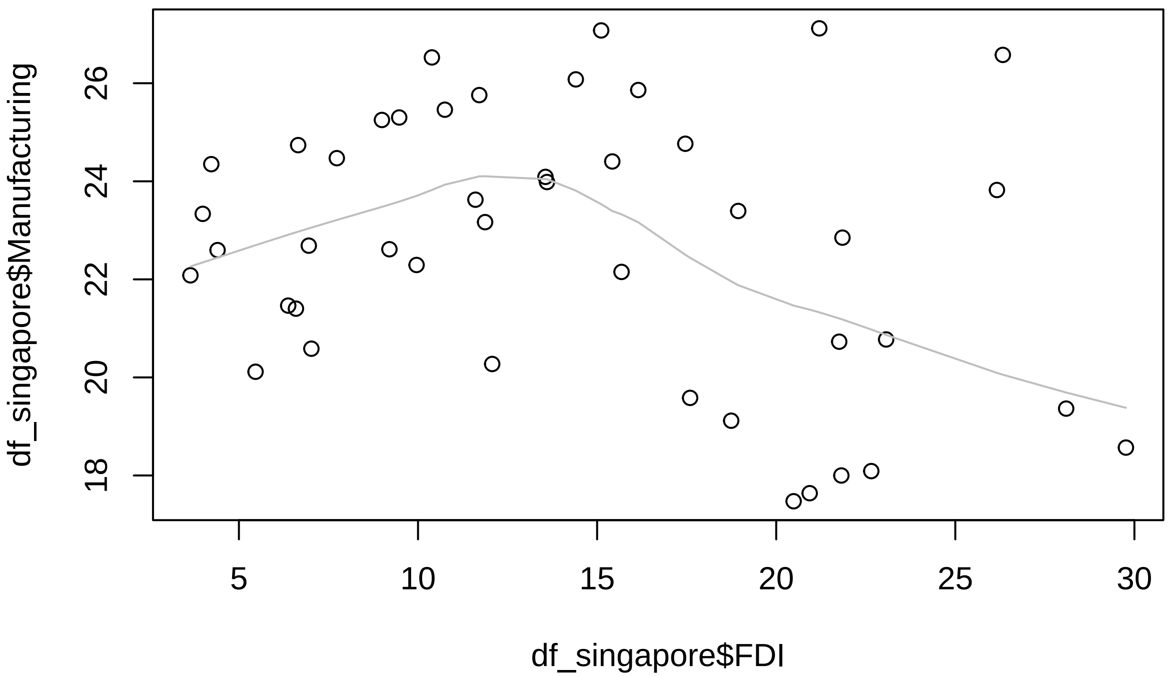
```r
# Current Account and GDP
cor(`Current Account`, GDP, method = "spearman")
```

```
## [1] 0.8274841
```

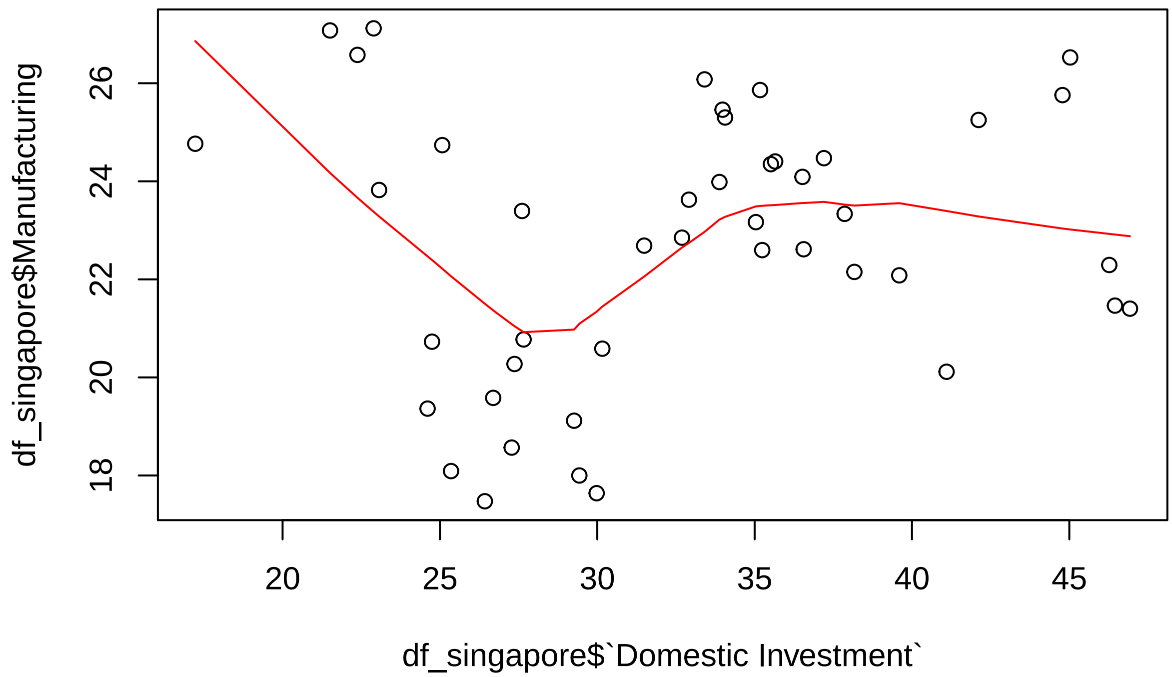
Based on the above, the correlation tests confirmed that only three economic indicators have a positive relationship with GDP namely: FDI, the level of working population and the current account. Out of the three, Working Population has the most positive impact on GDP. It is widely known that Singapore attracts a lot of international talent so as to mitigate the impact of its ageing population. Domestic Investment and Manufacturing, both have a negative relationship with GDP. Also, manufacturing seems to have the weakest effect on GDP based on the value of its correlation coefficient. As their relationship with GDP is not linear, the correlation tests have not been able to capture the *true* nature of their relationship.

We now examine the relationship between the different predictors, more precisely, between FDI and Manufacturing to ascertain whether FDI is being injected in the economy to increase the productive capacity of the country. It is also interesting to verify the relationship between Domestic Investment and Manufacturing - whether local businesses are also contributing to expand the manufacturing capacity of the country.

```r
plot(df_singapore$FDI, df_singapore$Manufacturing)
lines(lowess(df_singapore$FDI, df_singapore$Manufacturing), col = "grey")
```




```r
plot(df_singapore$`Domestic Investment`, df_singapore$Manufacturing)
lines(lowess(df_singapore$`Domestic Investment`, df_singapore$Manufacturing), col = "red")
```



Both graphics displays a cyclical relationship between FDI and Manufacturing and Domestic Investment and Manufacturing. This is also expected as in times of economic boom, investors and businesses are incentivised to invest due to the increase in the level of demand but in times of economic slowdown, the investors will take a wait and see attitude whilst businesses will tend to reduce their production level so as not to inflate the level of inventories.

Additional correlation tests now are carried out to confirm these relationships. However, the correlation tests may not be conclusive as they are not meant to assess cyclical relationships.

```r
attach(df_singapore)
```

```
## The following objects are masked from df_singapore (pos = 3):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop
```

```r
# FDI and Manufacturing
cor(FDI, Manufacturing, method = "spearman")
```

```
## [1] -0.1866103
```

```r
# Domestic Investment and Manufacturing
cor(`Domestic Investment`, Manufacturing, method = "spearman")
```

```
## [1] 0.08597604
```

Based on the value of their correlation coefficient, it seems that the level of FDI moves in opposite direction to the manufacturing sector. It is also noted that the relationship between domestic investment and manufacturing is weak. Thus, the correlation tests have not been able to depict the relationship between these variables.

A.3 Regression Analysis

The loess() function is usually used to to capture non linear relationships between variables. It can be used for a maximum of 4 explanatory variables. However, in this study we have 5 predictors. Therefore, we are going to use the lm() function using all 5 indicators – fit_singapore. Two explanatory variables namely Manufacturing and Domestic Investment will be log transformed due to their non linear relationship with GDP.

Model 1 - All Indicators
From the output of fit_singapore, most of the predictors have significant coefficients at 0.1% and the p value is quite small. However, the intercept is significant at 0.5% whilst the coefficient of the Current Account is not significant. Adjusted R Squared is quite close to 1.

```
# Model 1 - With all indicators
fit_singapore <- lm(GDP ~ FDI + log(`Domestic Investment`) + log(Manufacturing) + `Current Account` + `W

summary(fit_singapore)
```

```
##
## Call:
## lm(formula = GDP ~ FDI + log(‘Domestic Investment‘) + log(Manufacturing) +
##     ‘Current Account‘ + ‘Working Pop‘, data = df_singapore)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -6.796e+10 -1.742e+10  3.663e+09  1.764e+10  7.630e+10
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 7.536e+11  3.348e+11   2.251 0.030237 *
## FDI                         3.666e+09  9.294e+08   3.944 0.000333 ***
## log(‘Domestic Investment‘) -1.716e+11  3.745e+10  -4.584 4.83e-05 ***
## log(Manufacturing)         -3.047e+11  4.557e+10  -6.687 6.54e-08 ***
## ‘Current Account‘          -2.026e+09  1.130e+09  -1.793 0.081007 .
## ‘Working Pop‘               1.220e+10  3.128e+09   3.900 0.000379 ***
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
##
## Residual standard error: 2.91e+10 on 38 degrees of freedom
## Multiple R-squared:  0.929,  Adjusted R-squared:  0.9197
## F-statistic: 99.51 on 5 and 38 DF,  p-value: < 2.2e-16
```

Assessment Of Multicollinearity
To assess the quality of fit, we are going to calculate a number of indicators. * *The Variance Inflation Factor* (VIF) is calculated to ensure that that the economic indicators are not related to each other – the presence of multicollinearity. VIF value that is greater than 5 or 10 can indicate a problematic amount of collinearity. The economic indicators: Current Account and Working Population have a VIF value greater than 5. * A *correlation matrix* involving all five indicators is computed to further check the existence of multicollinearity. * The AIC value which will be compared to the other models to assess which is the best fit.

```r
broom::glance(fit_singapore)
```

```
## # A tibble: 1 x 12
##   r.squared adj.r.squared      sigma statistic  p.value    df logLik   AIC   BIC
##       <dbl>         <dbl>      <dbl>     <dbl>    <dbl> <dbl>  <dbl> <dbl> <dbl>
## 1     0.929         0.920    2.91e10      99.5 9.11e-21     5 -1119. 2253. 2265.
## # i 3 more variables: deviance <dbl>, df.residual <int>, nobs <int>
```

```r
# Calculation Variance Inflation Factor
library(car)
vif(fit_singapore)
```

```
##                      FDI log('Domestic Investment')
##                 2.252596                   3.868980
##       log(Manufacturing)          'Current Account'
##                 1.613176                   7.918349
##            'Working Pop'
##                 6.530427
```

```r
# Correlation Matrix
cor_matrix_sing <- cor(df_singapore[c("FDI", "Domestic Investment", "Manufacturing", "Working Pop", "Cu

cor(cor_matrix_sing)
```

```
##                            FDI Domestic Investment Manufacturing Working Pop
## FDI                  1.0000000          -0.9393119    -0.6995071   0.9521491
## Domestic Investment -0.9393119           1.0000000     0.5340735  -0.9470880
## Manufacturing       -0.6995071           0.5340735     1.0000000  -0.7687718
## Working Pop          0.9521491          -0.9470880    -0.7687718   1.0000000
## Current Account      0.9362283          -0.9929996    -0.6089197   0.9720950
##                     Current Account
## FDI                       0.9362283
## Domestic Investment      -0.9929996
## Manufacturing            -0.6089197
## Working Pop               0.9720950
## Current Account           1.0000000
```

It is clear that all the economic indicators are related to each other. Thus, the model, fit_singapore, may not be accurate.

Model 2 - With Two Indicators: Domestic Investment and Manufacturing
To reduce the effects of multicollinearity, the collinear variables can either be combined into a single predictor or the redundant variables are removed. In the first place, the redundant variables are removed: FDI, Working Population and Current Account. Manufacturing and Domestic Investment have a weaker positive relationship and are, therefore kept in the new model.

```r
# Model 2
fit_singapore.model2 <- lm(GDP ~ log(`Domestic Investment`) + log(Manufacturing), data = df_singapore)

summary(fit_singapore.model2)
```

```
##
## Call:
## lm(formula = GDP ~ log(`Domestic Investment`) + log(Manufacturing),
##     data = df_singapore)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.220e+11 -1.937e+10  1.617e+09  2.274e+10  8.951e+10
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                2.614e+12  1.858e+11  14.069  < 2e-16 ***
## log(`Domestic Investment`) -2.820e+11  2.873e+10  -9.815 2.53e-12 ***
## log(Manufacturing)         -4.795e+11  5.414e+10  -8.855 4.59e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.362e+10 on 41 degrees of freedom
## Multiple R-squared:  0.8279, Adjusted R-squared:  0.8196
## F-statistic: 98.65 on 2 and 41 DF,  p-value: < 2.2e-16
```

```
broom::glance(fit_singapore.model2)
```

```
## # A tibble: 1 x 12
##   r.squared adj.r.squared      sigma statistic  p.value    df logLik   AIC   BIC
##       <dbl>         <dbl>      <dbl>     <dbl>    <dbl> <dbl>  <dbl> <dbl> <dbl>
## 1     0.828         0.820    4.36e10      98.6 2.14e-16     2 -1139. 2286. 2293.
## # i 3 more variables: deviance <dbl>, df.residual <int>, nobs <int>
```

The new model, fit_singapore.model2, has a lower adjusted R squared despite the fact that all coefficients are significant at 1% level including the intercept. The p value is quite small, implying that the relationship is statistically significant. However, in terms of the AIC, its value is higher so it means that this model is not as good as the first model, fit_singapore.

Model 3 - Interaction Term

We now create a third model by introducing an interaction term namely the multiplication of FDI and Current Account. Both variables have a strong positive relationship and they are also correlated with Working population, Domestic Investment and Manufacturing. In essence, there is no need to add any other variable in the model.

```
attach(df_singapore)
```

```
## The following objects are masked from df_singapore (pos = 3):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_singapore (pos = 4):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop
```

```r
# Model 3 - Interaction Term
# Introducing an interaction term: FDI and Current Account
df_singapore$FDI.CurrAcc <- interaction(FDI, `Current Account`)

# Model with the interaction term
fit_singapore_interaction <- lm(df_singapore$GDP ~ df_singapore$FDI + df_singapore$`Current Account` + d

summary(fit_singapore_interaction)
```

```
##
## Call:
## lm(formula = df_singapore$GDP ~ df_singapore$FDI + df_singapore$`Current Account` +
##     df_singapore$FDI.CurrAcc, data = df_singapore, na.action = na.exclude)
##
## Residuals:
## ALL 44 residuals are 0: no residual degrees of freedom!
##
## Coefficients: (2 not defined because of singularities)
##                                                              Estimate
## (Intercept)                                                 1.062e+12
## df_singapore$FDI                                           -6.245e+10
## df_singapore$`Current Account`                              2.888e+10
## df_singapore$FDI.CurrAcc11.7103417060639.-10.368009979278   6.304e+09
## df_singapore$FDI.CurrAcc3.64608469550863.-8.9680890229879  -5.509e+11
## df_singapore$FDI.CurrAcc9.95912280287286.-8.05910345833412 -1.671e+11
## df_singapore$FDI.CurrAcc8.99256617445563.-7.9142496584868  -2.410e+11
## df_singapore$FDI.CurrAcc3.98964635197497.-6.01956905598569 -6.109e+11
## df_singapore$FDI.CurrAcc4.40373752423198.-4.45948103719695 -6.321e+11
## df_singapore$FDI.CurrAcc6.37594925354328.-3.43279404719149 -5.211e+11
## df_singapore$FDI.CurrAcc6.59196673143056.-1.94886737000521 -5.466e+11
## df_singapore$FDI.CurrAcc13.5578926036223.-0.519344689052203 -1.474e+11
## df_singapore$FDI.CurrAcc5.46417677336253.-0.0189810743321321 -6.731e+11
## df_singapore$FDI.CurrAcc9.20162995625447.1.71480429367363  -4.891e+11
## df_singapore$FDI.CurrAcc7.73270125338355.6.94851439736299  -6.897e+11
## df_singapore$FDI.CurrAcc14.4051395946458.7.63423434751461  -3.241e+11
## df_singapore$FDI.CurrAcc15.4235472379856.8.63724140435271  -2.771e+11
## df_singapore$FDI.CurrAcc9.47499553012531.9.7279605415254   -6.865e+11
## df_singapore$FDI.CurrAcc10.7488600422674.10.7339641366336  -6.249e+11
## df_singapore$FDI.CurrAcc16.148890451527.11.0715031927897   -2.325e+11
## df_singapore$FDI.CurrAcc4.22844121278652.11.3460660380192  -1.045e+12
## df_singapore$FDI.CurrAcc18.9398606857353.14.4329302808565  -1.568e+11
## df_singapore$FDI.CurrAcc11.8724659997364.14.627347115363   -6.278e+11
## df_singapore$FDI.CurrAcc6.65366588247058.14.8770537568056  -9.315e+11
## df_singapore$FDI.CurrAcc7.02328608228291.15.0834356181601  -8.430e+11
## df_singapore$FDI.CurrAcc28.0949935764636.15.3540131622503   5.992e+11
## df_singapore$FDI.CurrAcc11.6030848586318.15.4701578956137  -6.841e+11
## df_singapore$FDI.CurrAcc15.6823709085457.15.5698355478412  -4.075e+11
## df_singapore$FDI.CurrAcc21.7592915016982.15.7021310963339   1.888e+11
## df_singapore$FDI.CurrAcc20.9344810446532.15.707187590209    7.970e+10
## df_singapore$FDI.CurrAcc12.0710944143275.16.3930632928178  -5.653e+11
## df_singapore$FDI.CurrAcc13.6003946544885.16.4967504477224  -5.818e+11
## df_singapore$FDI.CurrAcc21.8492034558928.17.2038802401523  -6.528e+10
## df_singapore$FDI.CurrAcc18.7435244231802.17.6432739045995  -1.263e+11
```

```
## df_singapore$FDI.CurrAcc21.8184880782679.17.9502135491878     8.145e+10
## df_singapore$FDI.CurrAcc20.4865236102582.18.0392463471499     1.565e+10
## df_singapore$FDI.CurrAcc29.7636517525701.18.6260950135482     5.925e+11
## df_singapore$FDI.CurrAcc22.6541826890492.18.6930222903094     1.211e+11
## df_singapore$FDI.CurrAcc21.2027237505542.19.3111213606956    -1.293e+11
## df_singapore$FDI.CurrAcc6.95062480351291.21.7915511543613    -1.135e+12
## df_singapore$FDI.CurrAcc17.5960314065855.22.2184518342372    -3.417e+11
## df_singapore$FDI.CurrAcc23.0694716770803.22.9333472183785    -3.594e+10
## df_singapore$FDI.CurrAcc15.1133588478776.23.2604773041195    -6.115e+11
## df_singapore$FDI.CurrAcc17.4624153696896.24.3020823146558    -5.221e+11
## df_singapore$FDI.CurrAcc26.3271523046736.26.8933813503311           NA
## df_singapore$FDI.CurrAcc26.1619888768056.27.1433325306869           NA
##                                                              Std. Error t value
## (Intercept)                                                         NaN     NaN
## df_singapore$FDI                                                    NaN     NaN
## df_singapore$`Current Account`                                      NaN     NaN
## df_singapore$FDI.CurrAcc11.7103417060639.-10.368009979278           NaN     NaN
## df_singapore$FDI.CurrAcc3.64608469550863.-8.9680890229879           NaN     NaN
## df_singapore$FDI.CurrAcc9.95912280287286.-8.05910345833412          NaN     NaN
## df_singapore$FDI.CurrAcc8.99256617445563.-7.9142496584868           NaN     NaN
## df_singapore$FDI.CurrAcc3.98964635197497.-6.01956905598569          NaN     NaN
## df_singapore$FDI.CurrAcc4.40373752423198.-4.45948103719695          NaN     NaN
## df_singapore$FDI.CurrAcc6.37594925354328.-3.43279404719149          NaN     NaN
## df_singapore$FDI.CurrAcc6.59196673143056.-1.94886737000521          NaN     NaN
## df_singapore$FDI.CurrAcc13.5578926036223.-0.519344689052203         NaN     NaN
## df_singapore$FDI.CurrAcc5.46417677336253.-0.0189810743321321        NaN     NaN
## df_singapore$FDI.CurrAcc9.20162995625447.1.71480429367363           NaN     NaN
## df_singapore$FDI.CurrAcc7.73270125338355.6.94851439736299           NaN     NaN
## df_singapore$FDI.CurrAcc14.4051395946458.7.63423434751461           NaN     NaN
## df_singapore$FDI.CurrAcc15.4235472379856.8.63724140435271           NaN     NaN
## df_singapore$FDI.CurrAcc9.47499553012531.9.7279605415254            NaN     NaN
## df_singapore$FDI.CurrAcc10.7488600422674.10.7339641366336           NaN     NaN
## df_singapore$FDI.CurrAcc16.148890451527.11.0715031927897            NaN     NaN
## df_singapore$FDI.CurrAcc4.22844121278652.11.3460660380192           NaN     NaN
## df_singapore$FDI.CurrAcc18.9398606857353.14.4329302808565           NaN     NaN
## df_singapore$FDI.CurrAcc11.8724659997364.14.627347115363            NaN     NaN
## df_singapore$FDI.CurrAcc6.65366588247058.14.8770537568056           NaN     NaN
## df_singapore$FDI.CurrAcc7.02328608228291.15.0834356181601           NaN     NaN
## df_singapore$FDI.CurrAcc28.0949935764636.15.3540131622503           NaN     NaN
## df_singapore$FDI.CurrAcc11.6030848586318.15.4701578956137           NaN     NaN
## df_singapore$FDI.CurrAcc15.6823709085457.15.5698355478412           NaN     NaN
## df_singapore$FDI.CurrAcc21.7592915016982.15.7021310963339           NaN     NaN
## df_singapore$FDI.CurrAcc20.9344810446532.15.707187590209            NaN     NaN
## df_singapore$FDI.CurrAcc12.0710944143275.16.3930632928178           NaN     NaN
## df_singapore$FDI.CurrAcc13.6003946544885.16.4967504477224           NaN     NaN
## df_singapore$FDI.CurrAcc21.8492034558928.17.2038802401523           NaN     NaN
## df_singapore$FDI.CurrAcc18.7435244231802.17.6432739045995           NaN     NaN
## df_singapore$FDI.CurrAcc21.8184880782679.17.9502135491878           NaN     NaN
## df_singapore$FDI.CurrAcc20.4865236102582.18.0392463471499           NaN     NaN
## df_singapore$FDI.CurrAcc29.7636517525701.18.6260950135482           NaN     NaN
## df_singapore$FDI.CurrAcc22.6541826890492.18.6930222903094           NaN     NaN
## df_singapore$FDI.CurrAcc21.2027237505542.19.3111213606956           NaN     NaN
## df_singapore$FDI.CurrAcc6.95062480351291.21.7915511543613           NaN     NaN
## df_singapore$FDI.CurrAcc17.5960314065855.22.2184518342372           NaN     NaN
```

```
## df_singapore$FDI.CurrAcc23.0694716770803.22.9333472183785           NaN     NaN
## df_singapore$FDI.CurrAcc15.1133588478776.23.2604773041195           NaN     NaN
## df_singapore$FDI.CurrAcc17.4624153696896.24.3020823146558           NaN     NaN
## df_singapore$FDI.CurrAcc26.3271523046736.26.8933813503311            NA      NA
## df_singapore$FDI.CurrAcc26.1619888768056.27.1433325306869            NA      NA
##                                                               Pr(>|t|)
## (Intercept)                                                        NaN
## df_singapore$FDI                                                   NaN
## df_singapore$`Current Account`                                     NaN
## df_singapore$FDI.CurrAcc11.7103417060639.-10.368009979278          NaN
## df_singapore$FDI.CurrAcc3.64608469550863.-8.9680890229879          NaN
## df_singapore$FDI.CurrAcc9.95912280287286.-8.05910345833412         NaN
## df_singapore$FDI.CurrAcc8.99256617445563.-7.9142496584868          NaN
## df_singapore$FDI.CurrAcc3.98964635197497.-6.01956905598569         NaN
## df_singapore$FDI.CurrAcc4.40373752423198.-4.45948103719695         NaN
## df_singapore$FDI.CurrAcc6.37594925354328.-3.43279404719149         NaN
## df_singapore$FDI.CurrAcc6.59196673143056.-1.94886737000521         NaN
## df_singapore$FDI.CurrAcc13.5578926036223.-0.519344689052203        NaN
## df_singapore$FDI.CurrAcc5.46417677336253.-0.0189810743321321       NaN
## df_singapore$FDI.CurrAcc9.20162995625447.1.71480429367363          NaN
## df_singapore$FDI.CurrAcc7.73270125338355.6.94851439736299          NaN
## df_singapore$FDI.CurrAcc14.4051395946458.7.63423434751461          NaN
## df_singapore$FDI.CurrAcc15.4235472379856.8.63724140435271          NaN
## df_singapore$FDI.CurrAcc9.47499553012531.9.7279605415254           NaN
## df_singapore$FDI.CurrAcc10.7488600422674.10.7339641366336          NaN
## df_singapore$FDI.CurrAcc16.148890451527.11.0715031927897           NaN
## df_singapore$FDI.CurrAcc4.22844121278652.11.3460660380192          NaN
## df_singapore$FDI.CurrAcc18.9398606857353.14.4329302808565          NaN
## df_singapore$FDI.CurrAcc11.8724659997364.14.627347115363           NaN
## df_singapore$FDI.CurrAcc6.65366588247058.14.8770537568056          NaN
## df_singapore$FDI.CurrAcc7.02328608228291.15.0834356181601          NaN
## df_singapore$FDI.CurrAcc28.0949935764636.15.3540131622503          NaN
## df_singapore$FDI.CurrAcc11.6030848586318.15.4701578956137          NaN
## df_singapore$FDI.CurrAcc15.6823709085457.15.5698355478412          NaN
## df_singapore$FDI.CurrAcc21.7592915016982.15.7021310963339          NaN
## df_singapore$FDI.CurrAcc20.9344810446532.15.707187590209           NaN
## df_singapore$FDI.CurrAcc12.0710944143275.16.3930632928178          NaN
## df_singapore$FDI.CurrAcc13.6003946544885.16.4967504477224          NaN
## df_singapore$FDI.CurrAcc21.8492034558928.17.2038802401523          NaN
## df_singapore$FDI.CurrAcc18.7435244231802.17.6432739045995          NaN
## df_singapore$FDI.CurrAcc21.8184880782679.17.9502135491878          NaN
## df_singapore$FDI.CurrAcc20.4865236102582.18.0392463471499          NaN
## df_singapore$FDI.CurrAcc29.7636517525701.18.6260950135482          NaN
## df_singapore$FDI.CurrAcc22.6541826890492.18.6930222903094          NaN
## df_singapore$FDI.CurrAcc21.2027237505542.19.3111213606956          NaN
## df_singapore$FDI.CurrAcc6.95062480351291.21.7915511543613          NaN
## df_singapore$FDI.CurrAcc17.5960314065855.22.2184518342372          NaN
## df_singapore$FDI.CurrAcc23.0694716770803.22.9333472183785          NaN
## df_singapore$FDI.CurrAcc15.1133588478776.23.2604773041195          NaN
## df_singapore$FDI.CurrAcc17.4624153696896.24.3020823146558          NaN
## df_singapore$FDI.CurrAcc26.3271523046736.26.8933813503311           NA
## df_singapore$FDI.CurrAcc26.1619888768056.27.1433325306869           NA
##
## Residual standard error: NaN on 0 degrees of freedom
```

```
## Multiple R-squared:      1,  Adjusted R-squared:    NaN
## F-statistic:   NaN on 43 and 0 DF,  p-value: NA
```

A look at the output of the model, fit_singapore_interaction, indicates that this model has not resolved the multicollinearity problem between the explanatory variables and is, therefore, not reliable.

Model 4 - Polynomial Regression
As the relationship between GDP and FDI is cyclical, a polynomial regression - a quadratic equation may be a better fit. We are also going to include Domestic Investment and Manufacturing as based on the correlation matrix, they both have a weak relationship with the other variables.

```r
# Model 4 - Polynomial Regression
quad.sing.3 <- GDP ~ FDI + I(FDI^2) + log(Manufacturing) + log(`Domestic Investment`)
fit_singapore_quadratic <- lm(quad.sing.3, data = df_singapore)

summary(fit_singapore_quadratic)
```

```
##
## Call:
## lm(formula = quad.sing.3, data = df_singapore)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -8.636e+10 -1.958e+10  5.722e+08  1.756e+10  8.044e+10
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 1.985e+12  1.945e+11  10.209 1.42e-12 ***
## FDI                         6.469e+09  3.413e+09   1.896   0.0654 .
## I(FDI^2)                   -4.299e+07  1.078e+08  -0.399   0.6921
## log(Manufacturing)         -4.126e+11  4.654e+10  -8.867 6.87e-11 ***
## log(`Domestic Investment`) -1.839e+11  2.995e+10  -6.140 3.31e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.464e+10 on 39 degrees of freedom
## Multiple R-squared:  0.8968, Adjusted R-squared:  0.8862
## F-statistic: 84.69 on 4 and 39 DF,  p-value: < 2.2e-16
```

```r
broom::glance(fit_singapore_quadratic)
```

```
## # A tibble: 1 x 12
##   r.squared adj.r.squared      sigma statistic  p.value    df logLik   AIC   BIC
##       <dbl>         <dbl>      <dbl>     <dbl>    <dbl> <dbl>  <dbl> <dbl> <dbl>
## 1     0.897         0.886    3.46e10      84.7 1.09e-18     4 -1128. 2267. 2278.
## # i 3 more variables: deviance <dbl>, df.residual <int>, nobs <int>
```

Based on the output of the polynomial regression, it can be inferred that this model is the best fit based on its AIC value, 2267 which is the lowest amongst all the existing models even though its adjusted R squared is not as high as the first model, fit_singapore. We can also note that the coefficients for FDI and FDI^2 are not significant. However, their presence do contribute to a better quality of the regression model since the adjusted R squared for this model is higher than fit_singapore_model2.

Assessment of Sequential Correlation
It is surprising to note that the coefficients of both Manufacturing and Domestic Investment are negative. It would have been more logical that they are positive as expanding the productive capacity of the country leads to more exports and therefore, revenue for the country. This may be due to **sequential correlation** whereby observations that are close together in time are highly correlated. Sequential correlation violates the assumption that error terms are independent and identically distributed. In essence, the presence of sequential correlation leads to incorrect conclusions about the relationship of GDP and the other economic indicators.

To assess whether there is sequential correlation:

```
acf(residuals(fit_singapore_quadratic))
```

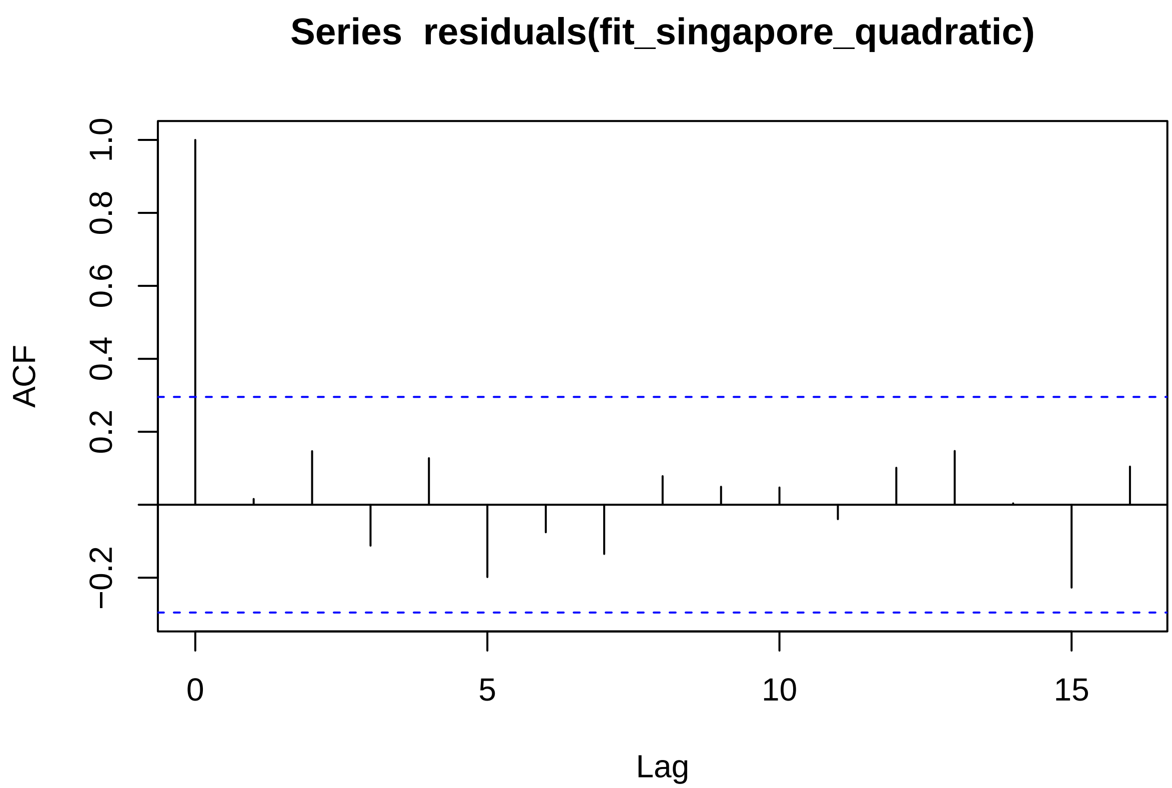


Figure 3: Assessment for Sequential Correlation

```
plot(predict(fit_singapore_quadratic), residuals(fit_singapore_quadratic))
```

The acf() function is used. In the plot, the blue dashed lines represent the confidence intervals at 95%. The vertical axis represents the autocorrelation and the horizontal axis represents the lag. The autocorrelation at lag0 is included by default. This always takes the value of 1, since it is the correlation between the data series and itself. Interest is the autocorrelations at lag1 and later lags. There is no significant autocorrelation as none of the spikes go beyond the confidence bounds. However, the bar heights indicate a seasonal trend.

This is confirmed in the plot of fitted values against the residuals. The scatterplot indicates that there is some kind of trend. Thus, it is confirmed that the model, fit_singapore_quadratic is not the best fit. Given

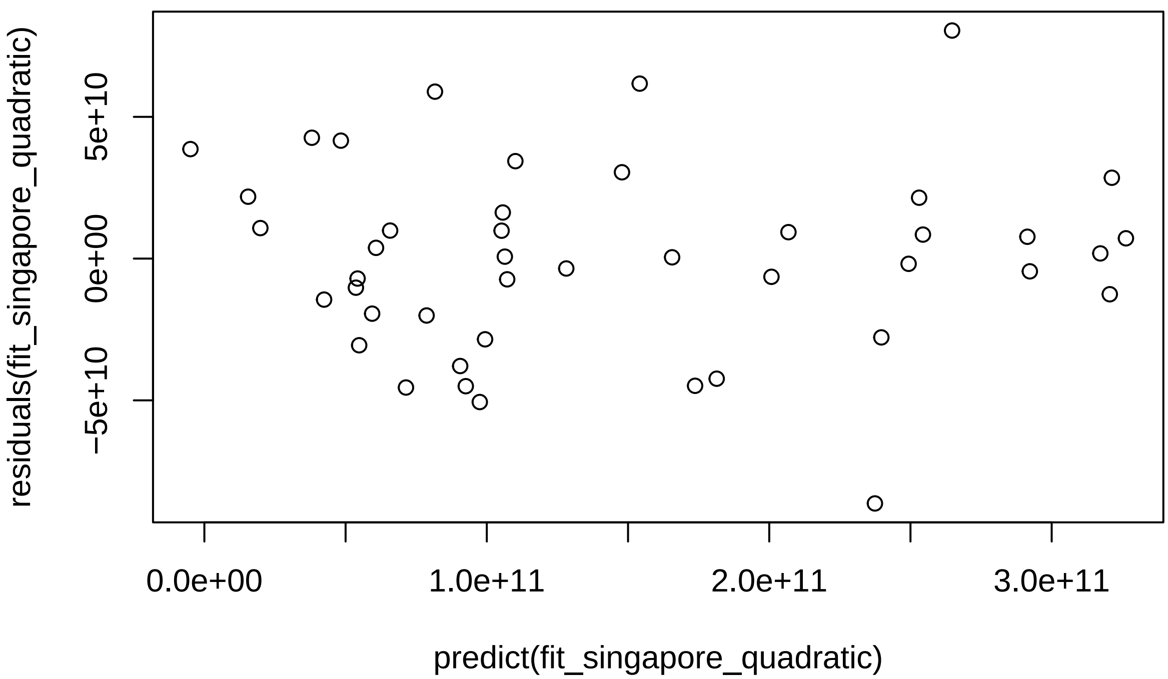


Figure 4: Assessment for Sequential Correlation

the presence of seasonality, the ARIMA model would need to be used to find a better model. However, this is beyond the scope of this study.

We now proceed to have a closer look at the data related to Malta.

B.Malta

B.1 Exploratory Analysis and Data Wrangling

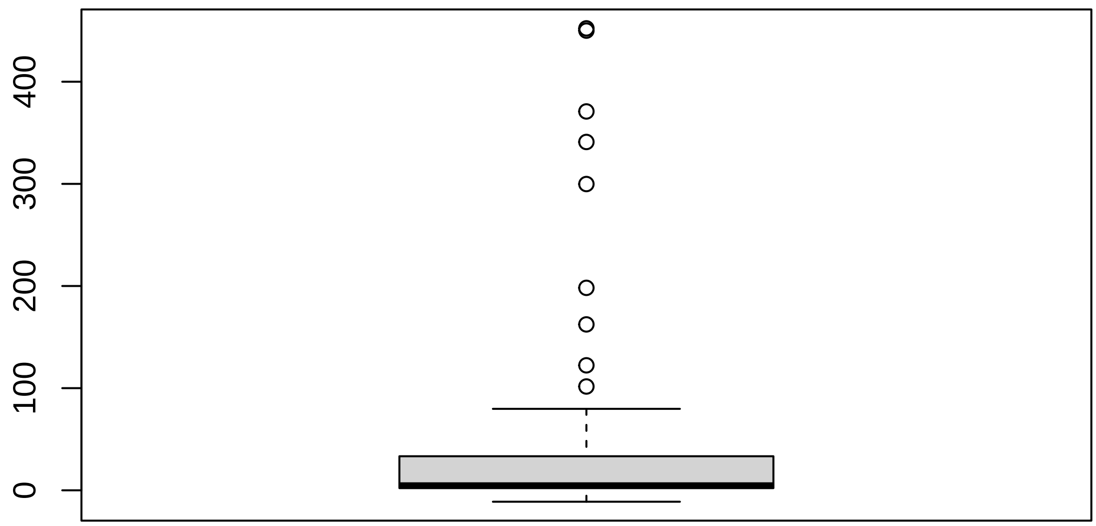
```
summary(df_malta)
```

```
##     iso2c              iso3c              country                date     
##  Length:44          Length:44          Length:44          Min.   :1976  
##  Class :character   Class :character   Class :character   1st Qu.:1987  
##  Mode  :character   Mode  :character   Mode  :character   Median :1998  
##                                                           Mean   :1998  
##                                                           3rd Qu.:2008  
##                                                           Max.   :2019  
##  Current Account         FDI          Domestic Investment Manufacturing   
##  Min.   :-11.9011   Min.   :-11.174   Min.   :15.91       Min.   : 6.431  
##  1st Qu.: -3.7062   1st Qu.:  2.123   1st Qu.:21.44       1st Qu.:12.084  
##  Median : -0.3423   Median :  4.892   Median :23.59       Median :18.110  
##  Mean   :  0.3490   Mean   : 64.006   Mean   :23.74       Mean   :17.545  
##  3rd Qu.:  2.7434   3rd Qu.: 32.724   3rd Qu.:26.58       3rd Qu.:23.061  
##  Max.   : 19.1577   Max.   :452.221   Max.   :30.63       Max.   :28.095  
##       GDP              Working Pop   
##  Min.   :1.649e+09   Min.   :66.10  
##  1st Qu.:2.832e+09   1st Qu.:66.51  
##  Median :5.216e+09   Median :67.17  
##  Mean   :6.042e+09   Mean   :67.50  
##  3rd Qu.:8.128e+09   3rd Qu.:68.27  
##  Max.   :1.488e+10   Max.   :70.20
```

Based on the detailed information of the data related to Malta, it is noted that its manufacturing sector has definitely reached the 20% threshold for the period under review given that its maximum stands at 28.1% and its mean is at 17.54%. In regards to its level of domestic investment, its maximum level reaches 30.63% indicating that it has reached the minimum threshold of 25% sometime during the period under review. It is also interesting to note that Malta's level of FDI is quite substantial and far surpasses the rest of the sample of countries with its mean at 67.47% reaching a maximum of 452.2%. It can be safely assumed that these are monies transiting through the offshore centre of the country as Malta is well known to be a tax haven. This also implies that the variable, FDI, is largely skewed to the right and we would need to apply a logarithmic transformation. However, transforming negative values to logarithmic values would result in NaN values.

In this case, it is best to remove the outlier value(s) so that they do not skew the results and lead to inaccurate interpretation. To do so, a boxplot of the FDI's data is plotted as shown the figure below.

```
# Identifying outliers
boxplot(df_malta$FDI)
```



From the boxplot, it is noted that there are a few outlier points above the 100% level in regards to the Maltese FDI data. Removing these outlier points will lead to better interpretation results.

```r
# Removing outliers
df_malta_clean <- df_malta %>%
  filter(FDI <= 100)

knitr::kable(summary(df_malta_clean), caption = "Summary of Maltese cleaned dataset")
```

Table 1: Summary of Maltese cleaned dataset

	iso2c	iso3c	country	date	Current Account	FDI	Domestic Investment	Manufacturing	GDP	Working Pop
	Length:35	Length:35	Length:35	Min. :1976	Min. :-11.9011	Min. :-11.174	Min. :15.91	Min. : 6.997	Min. :1.649e+09	Min. :66.10
	Class :character	Class :character	Class :character	1st Qu.:1984	1st Qu.:-3.1909	1st Qu.: 2.052	1st Qu.:22.26	1st Qu.:15.735	1st Qu.:2.611e+09	1st Qu.:66.35
	Mode :character	Mode :character	Mode :character	Median :1993	Median : -0.1329	Median : 2.953	Median :25.09	Median :19.849	Median :4.144e+09	Median :66.80
	NA	NA	NA	Mean :1994	Mean : -0.2668	Mean : 9.076	Mean :24.20	Mean :19.334	Mean :5.050e+09	Mean :67.15

iso2c	iso3c	country	date	Current Account	FDI	Domestic Investment	Manufacturing	GDP	Working Pop
NA	NA	NA	3rd Qu.:2002	3rd Qu.: 2.5769	3rd Qu.: 10.407	3rd Qu.:27.09	3rd Qu.:24.250	3rd Qu.:6.767e+09	3rd Qu.:67.76
NA	NA	NA	Max. :2016	Max. : 10.4257	Max. : 79.749	Max. :30.63	Max. :28.095	Max. :1.180e+10	Max. :70.08

We can note a few changes in the *clean* database. The mean level of FDI is now 9.1% and its maximum is 79.75% - relatively high. Its median value is 2.9%. In this case and if there is need be, it is better that we use the median value given the large range of values. We can also note that the mean of the Manufacturing sector has changed to 19.33% and that of the Current Account now stands at a deficit of 0.27%. We can safely conclude that removing the outliers did not have a major impact on the remaining data.

Deep Dive Analysis

A deeper dive in reveals that the Maltese manufacturing sector's contribution was at least 20% of GDP for 17 years. Its level of domestic investment also exceeded the 25% threshold for almost the same amount of time:18 years. Malta had a current account deficit exceeding or equal to 5% for 6 years out of the 43 years, indicating that the country was able to well manage its expenditure for most of the time. These results indicate that the country was able to create a virtuous circle of economic growth resulting into a long term economic prosperity.

```r
# Number of years the Maltese Manufacturing sector's contribution equals or exceeds 20%
df_malta_clean_manf <- df_malta_clean %>%
  filter(Manufacturing >= 20)

nrow(df_malta_clean_manf)
```

```
## [1] 17
```

```r
# Number of years the level of the Maltese Domestic Investment equals or exceeds 25%
df_malta_clean_dominv <- df_malta_clean %>%
  filter(`Domestic Investment` >= 25)

nrow(df_malta_clean_dominv)
```

```
## [1] 18
```

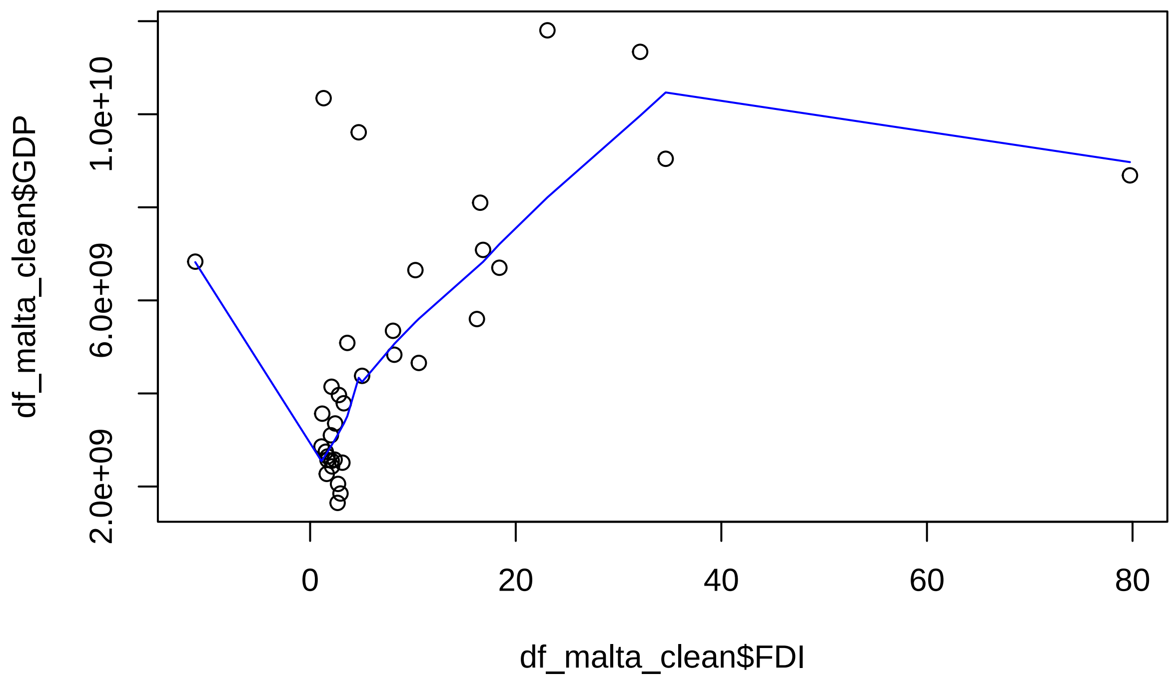
```r
# Number of years that the Maltese current account deficit exceeded 5%
df_malta_clean_currentacc <- df_malta_clean %>%
  filter(`Current Account` <= -5)

nrow(df_malta_clean_currentacc)
```

```
## [1] 6
```

Exploratory Graphics of Independent Variables With GDP

We shall now explore the relationship of GDP with the other economic indicators.

```r
# Plot of FDI against GDP
plot(df_malta_clean$FDI, df_malta_clean$GDP)
lines(lowess(df_malta_clean$FDI, df_malta_clean$GDP), col = "blue")
```



The graphic demonstrates that the relationship is cyclical. This is more typical as investors are attracted to economies which are experiencing an economic boom and will recall their monies when the economy is slowing down. Given that Malta is a tax haven, it may be easier for investors to recall their monies and invest elsewhere when the country's economy is in recession.

```r
attach(df_malta_clean)
```

```
## The following objects are masked from df_singapore (pos = 3):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_singapore (pos = 4):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_singapore (pos = 5):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop
```

```r
# Plot Manufacturing against GDP
plot(Manufacturing, GDP)
lines(lowess(Manufacturing, GDP), col = "darkblue")
```

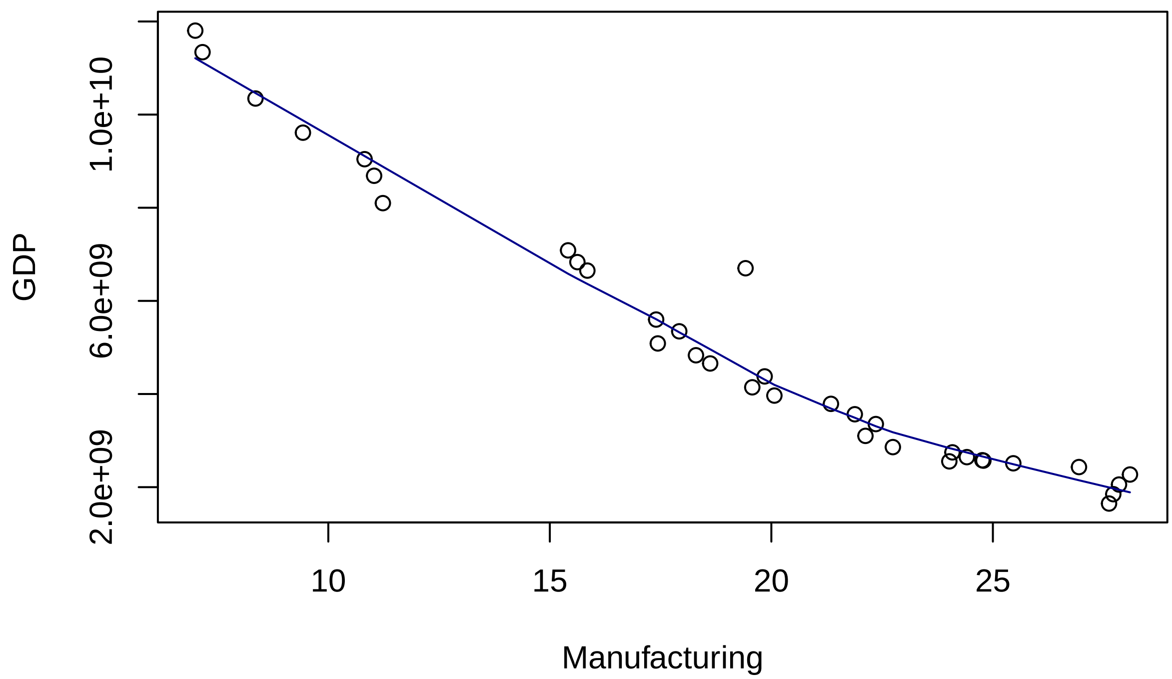


Figure 5: Exploratory Graphics of Independent Variables with GDP

```r
# Plot Domestic Investment against GDP
plot(`Domestic Investment`, GDP)
lines(lowess(`Domestic Investment`, GDP), col = "red")
```

```r
# Plot Current Account against GDP
plot(`Current Account`, GDP)
lines(lowess(`Current Account`, GDP), col = "green")
```

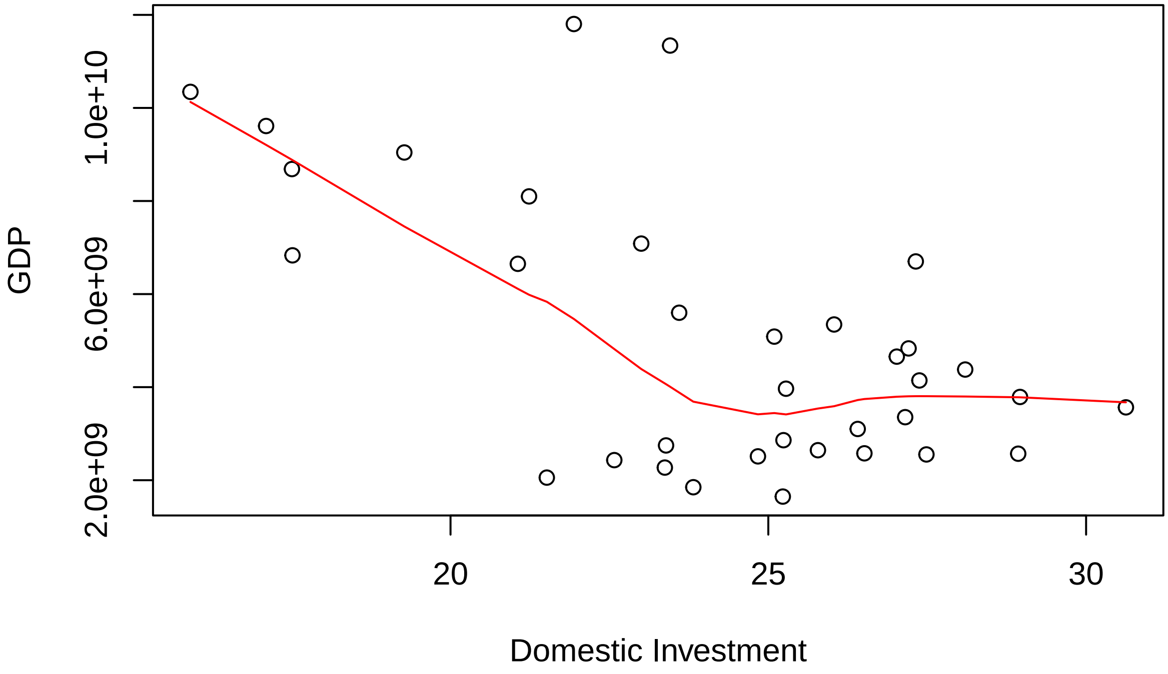
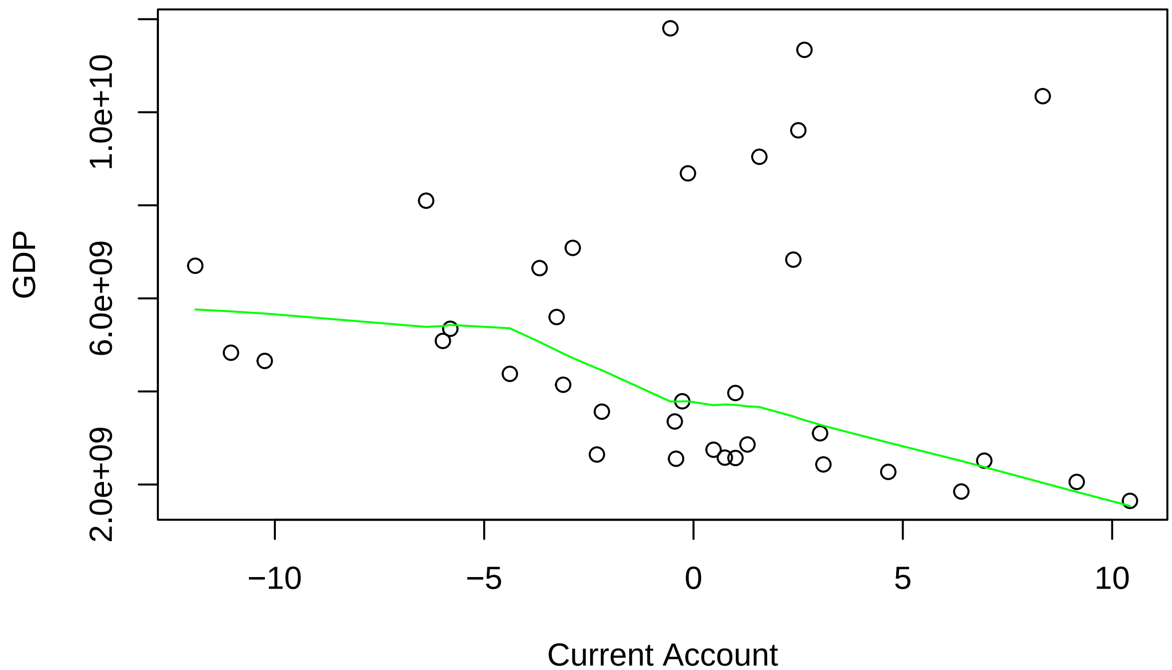
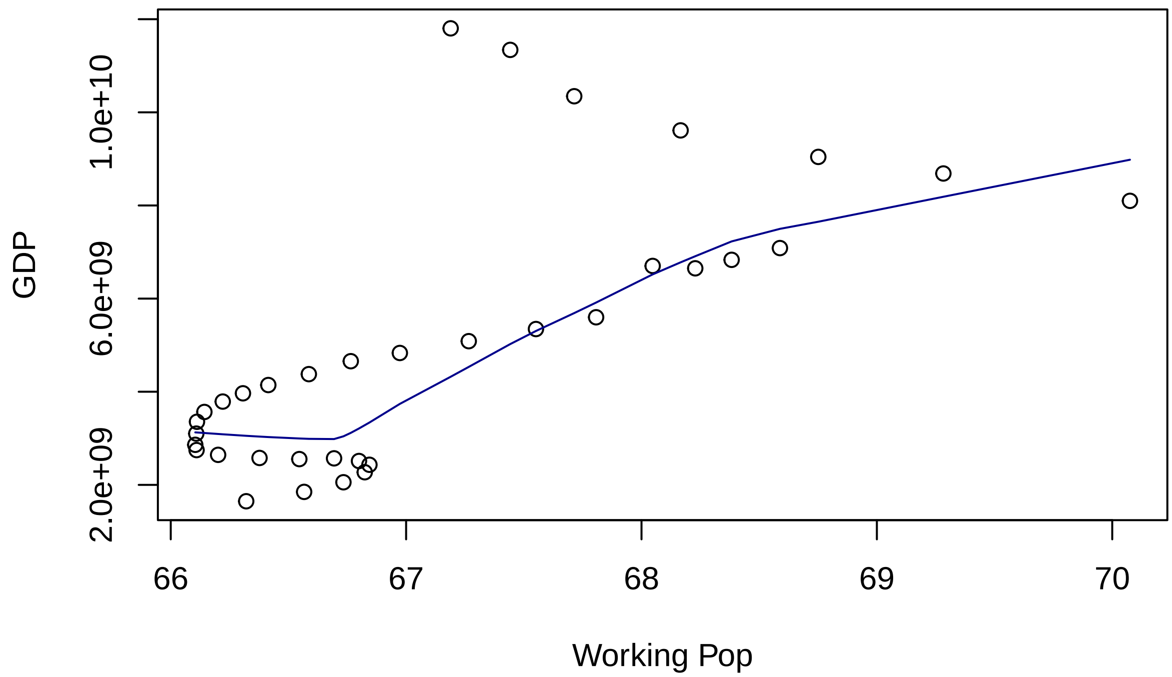


Figure 6: Exploratory Graphics of Independent Variables with GDP



All three graphics detailed above illustrate a linear negative relationship between GDP and the following three explanatory variables: Manufacturing, Domestic Investment and Current Account. It is widely known that many countries are focusing on growing their services sector rather than their manufacturing sector despite the fact the manufacturing sector has a much wider spill over effects on the rest of the economy. When factories are expanding their productive capacities, they need more machinery, tools, better infrastructure (power supply, more road networks) which in turn generate more demand for jobs and better work prospects for the local workforce. Furthermore, as there is less manufacturing activities in times of economic boom, Malta has to import more for the local demand in terms of goods. This may explain why the Current account has a negative relationship with GDP. As for Domestic Investment, with the massive inflows of FDI, there may be a *crowding out effect* of domestic investment especially in case when foreign firms are more productive than local businesses. However, we would need to investigate further for us to confirm these hypotheses.

```
# Plot Working Population against GDP
plot(`Working Pop`, GDP)
lines(lowess(`Working Pop`, GDP), col = "darkblue")
```



The shape of the graphic is similar to that of the Singaporean data. Malta is also a small country and if the available pool of resources (housing, infrastructure) is not well managed, labour output per worker may decline. As the share of manufacturing in GDP is on the decline, we can also infer that insufficient investment is made in the productive capacity, impacting adversely labour output per worker.

We now start the correlation tests.

B.2 Correlation Tests

```r
attach(df_malta_clean)
```

```
## The following objects are masked from df_malta_clean (pos = 3):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_singapore (pos = 4):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_singapore (pos = 5):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop
```

```
## The following objects are masked from df_singapore (pos = 6):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop
```

```r
# Correlation between GDP and FDI
cor(FDI, GDP, method = "spearman")
```

```
## [1] 0.5733894
```

```r
# Domestic Investment and GDP
cor(`Domestic Investment`, GDP, method = "spearman")
```

```
## [1] -0.4042017
```

```r
# Manufacturing and GDP
cor(Manufacturing, GDP, method = "spearman")
```

```
## [1] -0.9876751
```

```r
# Working Population and GDP
cor(`Working Pop`, GDP, method = "spearman")
```

```
## [1] 0.6851541
```

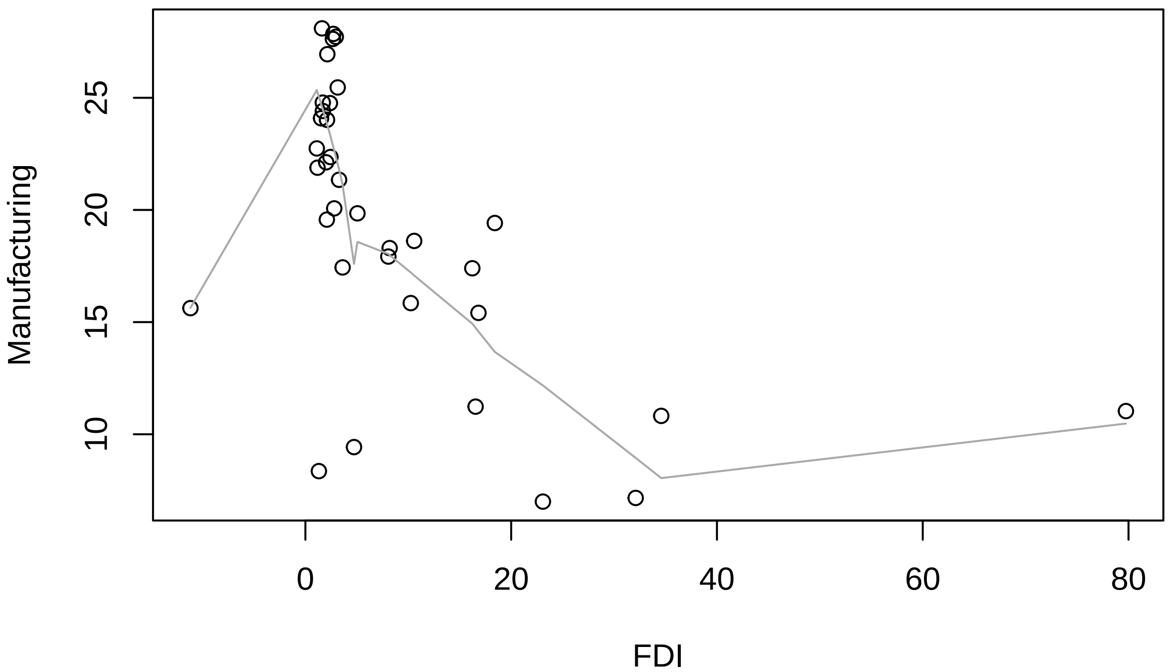
```r
# Current Account and GDP
cor(`Current Account`, GDP, method = "spearman")
```

```
## [1] -0.3910364
```

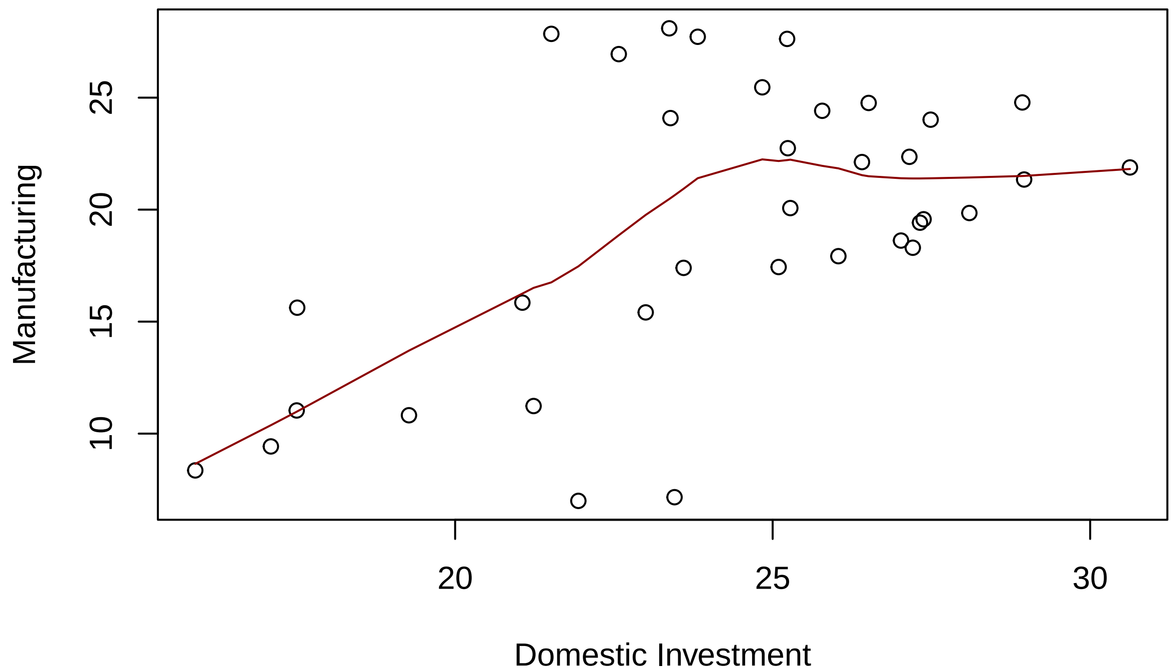
The correlation tests confirm that Manufacturing, Current Account and Domestic Investment all have a negative linear relationship with GDP. Working Population as well as FDI has a positive relationship with GDP. As was the case for Singapore, non linear relationship are not able to be detected by the correlation tests. It is interesting to note that the manufacturing sector moves in opposite direction almost 100% of the time as its correlation coefficient is almost negative one. This confirms what we have noted in the previous exploratory analysis.

Similar to Singapore, we are going to explore the relationship between FDI and Manufacturing and Domestic Investment and Manufacturing respectively.

```r
# Plot FDI against Manufacturing
plot(FDI, Manufacturing)
lines(lowess(FDI, Manufacturing), col = "darkgrey")
```




```
# Plot Domestic Investment against Manufacturing
plot(`Domestic Investment`, Manufacturing)
lines(lowess(`Domestic Investment`, Manufacturing), col = "darkred")
```



The relationship between FDI and Manufacturing is altogether negative. It was initially a positive linear relationship but as Malta starts to focus on its services sector, the role of the Manufacturing sector took a back seat. The relationship with FDI consequently became a linear negative relationship. On the other hand, the relationship between Domestic Investment and Manufacturing is initially a linear positive relationship but then gradually plateaued. The level of economic complexity for Malta is relatively low, possibly implying that the manufacturing sector may be plagued by outdated technology, limited capacity and reduced competitiveness. This may be why the relationship has plateaued.

We are now going to do the correlation tests between these explanatory variables.

```r
# Correlation between FDI and Manufacturing
cor(FDI, Manufacturing, method = "spearman")
```

```
## [1] -0.5689076
```

```r
# Correlation between Domestic Investment and Manufacturing
cor(`Domestic Investment`, Manufacturing, method = "spearman")
```

```
## [1] 0.4064426
```

It is confirmed that the level of FDI has a negative relationship with the manufacturing sector, thereby confirming that most of the foreign monies goes to the Maltese offshore sector or to the services sector. The level of Domestic Investment moves in the same direction as the level of Manufacturing's contribution as their correlation coefficient is positive but the strength of their relationship is relatively weak, 0.4. Most probably, this is because their relationship has plateaued.

B.3 Regression Analysis

A regression analysis is now performed to find out more about the key drivers of economic growth for Malta. A model is first determined using all the five independent variables – fit_malta.

Model 1 - With All Indicators
Though Adjusted R Squared is close to 1, only one variable has a significant coefficient at 0.1% namely: Manufacturing.

```
# Model 1 - With all indicators
fit_malta <- lm(GDP ~ Manufacturing + `Domestic Investment` + FDI + `Current Account` + `Working Pop`, d

summary(fit_malta)
```

```
##
## Call:
## lm(formula = GDP ~ Manufacturing + `Domestic Investment` + FDI +
##     `Current Account` + `Working Pop`, data = df_malta_clean)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -781569860 -266076218  -89909429   61324720 1991327706
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)           4.939e+09  1.493e+10   0.331    0.743
## Manufacturing        -4.295e+08  2.629e+07 -16.341 3.62e-16 ***
## `Domestic Investment` -2.133e+07  6.086e+07  -0.350    0.729
## FDI                   5.992e+06  8.736e+06   0.686    0.498
## `Current Account`     3.545e+07  3.142e+07   1.128    0.268
## `Working Pop`         1.323e+08  2.070e+08   0.639    0.528
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 612800000 on 29 degrees of freedom
## Multiple R-squared:  0.9618, Adjusted R-squared:  0.9552
## F-statistic: 145.9 on 5 and 29 DF,  p-value: < 2.2e-16
```

Assessment Of Multicollinearity
Additional analysis is done with the calculation of VIF, correlation matrix and the AIC. They may explain why most economic indicators do not have significant coefficients in the model, fit_malta.

```
# Variance Inflation Factor
vif(fit_malta)
```

```
##         Manufacturing `Domestic Investment`                   FDI
##              2.449580              4.427938              1.630874
##     `Current Account`         `Working Pop`
##              2.540605              3.974496
```

Based on the above, we can note that all five variables have VIF values below 5.

```r
# Correlation Matrix
cor_matrix_malta <- cor(df_malta_clean[c("FDI", "Manufacturing", "Domestic Investment", "Current Accoun

cor(cor_matrix_malta)
```

```
##                           FDI Manufacturing Domestic Investment
## FDI                  1.0000000    -0.9389814          -0.7173848
## Manufacturing       -0.9389814     1.0000000           0.8040116
## Domestic Investment -0.7173848     0.8040116           1.0000000
## Current Account     -0.3540759     0.3017330          -0.2655892
## Working Pop          0.9025688    -0.9681367          -0.8510681
##                     Current Account Working Pop
## FDI                      -0.3540759   0.9025688
## Manufacturing             0.3017330  -0.9681367
## Domestic Investment      -0.2655892  -0.8510681
## Current Account           1.0000000  -0.2795443
## Working Pop              -0.2795443   1.0000000
```

```r
broom::glance(fit_malta)
```

```
## # A tibble: 1 x 12
##   r.squared adj.r.squared      sigma statistic  p.value    df logLik   AIC   BIC
##       <dbl>         <dbl>      <dbl>     <dbl>    <dbl> <dbl>  <dbl> <dbl> <dbl>
## 1     0.962         0.955 612824085.      146. 1.25e-19     5  -755. 1523. 1534.
## # i 3 more variables: deviance <dbl>, df.residual <int>, nobs <int>
```

Based on the above results, we can note that the variable, Current Account is the only one weakly correlated with all other variables as the values of its correlation coefficient are all below 0.5. The AIC value for this model is 1523. It needs to be compared with the other models for the Maltese data.

Model 2 - With One Indicator: Manufacturing
So, we now apply a forward selection approach. We determine a new model with only one variable: Manufacturing. According to the correlation tests, it has almost 100% impact on GDP. Based on the output of the model, fit_malta_1, both the intercept and the coefficient of Manufacturing are significant at 0.1%. As expected, Manufacturing has a negative coefficient. The p value is relatively small and the adjusted R Squared is close to 1. The AIC has a lower value implying that this model, fit_malta_1 is better than fit_malta.

```r
# Model 2 with one indicator - Manufacturing
fit_malta_1 <- lm(GDP ~ Manufacturing, data = df_malta_clean)

summary(fit_malta_1)
```

```
##
## Call:
## lm(formula = GDP ~ Manufacturing, data = df_malta_clean)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -820101467 -410463196 -109949528  302490383 1688795959
##
## Coefficients:
```

```
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.378e+10  3.555e+08   38.76   <2e-16 ***
## Manufacturing -4.514e+08  1.752e+07  -25.77   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 639100000 on 33 degrees of freedom
## Multiple R-squared:  0.9527, Adjusted R-squared:  0.9512
## F-statistic: 664.2 on 1 and 33 DF,  p-value: < 2.2e-16
```

```r
broom::glance(fit_malta_1)
```

```
## # A tibble: 1 x 12
##   r.squared adj.r.squared      sigma statistic  p.value    df logLik   AIC   BIC
##       <dbl>         <dbl>      <dbl>     <dbl>    <dbl> <dbl>  <dbl> <dbl> <dbl>
## 1     0.953         0.951 639131018.      664. 1.95e-23     1  -758. 1523. 1527.
## # i 3 more variables: deviance <dbl>, df.residual <int>, nobs <int>
```

Model 3 - With Manufacturing and Current Account

As the variable, Current Account is weakly correlated to the other variables, we add it to Manufacturing and create a new model, fit_malta_2 to check if Current Account improves the linear model.

```r
# Model 3 - With Manufacturing and Current Account
fit_malta_2 <- lm(GDP ~ Manufacturing + `Current Account`, data = df_malta_clean)

summary(fit_malta_2)
```

```
##
## Call:
## lm(formula = GDP ~ Manufacturing + `Current Account`, data = df_malta_clean)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -800981211 -303193424 -111894627   76710312 2170485271
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)         1.396e+10  3.510e+08  39.767   <2e-16 ***
## Manufacturing      -4.602e+08  1.728e+07 -26.632   <2e-16 ***
## `Current Account`   4.134e+07  2.028e+07   2.038   0.0499 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 610600000 on 32 degrees of freedom
## Multiple R-squared:  0.9581, Adjusted R-squared:  0.9555
## F-statistic: 365.9 on 2 and 32 DF,  p-value: < 2.2e-16
```

```r
broom::glance(fit_malta_2)
```

```
## # A tibble: 1 x 12
##   r.squared adj.r.squared      sigma statistic  p.value    df logLik   AIC   BIC
##       <dbl>         <dbl>      <dbl>     <dbl>    <dbl> <dbl>  <dbl> <dbl> <dbl>
## 1     0.958         0.955 610613951.      366. 9.01e-23     2  -756. 1520. 1527.
## # i 3 more variables: deviance <dbl>, df.residual <int>, nobs <int>
```

The model, fit_malta_2, seems to be slightly better with an improvement in adjusted R Squared. The coefficients for both Manufacturing and the intercept are significant at 0.1% and the p value is relatively small. The coefficient of the Current Account is significant at 5%. The AIC value has decreased once more. Fit_malta_2 is a better model.

Assessment of Model 2
As an additional verification, the plot of residuals versus predicted values of GDP is calculated. A closer look tells us that there seems to be a pattern in the plot of residuals against the fitted values of GDP, implying that there is still room for improvement.

```
plot(predict(fit_malta_2), residuals(fit_malta_2))
```

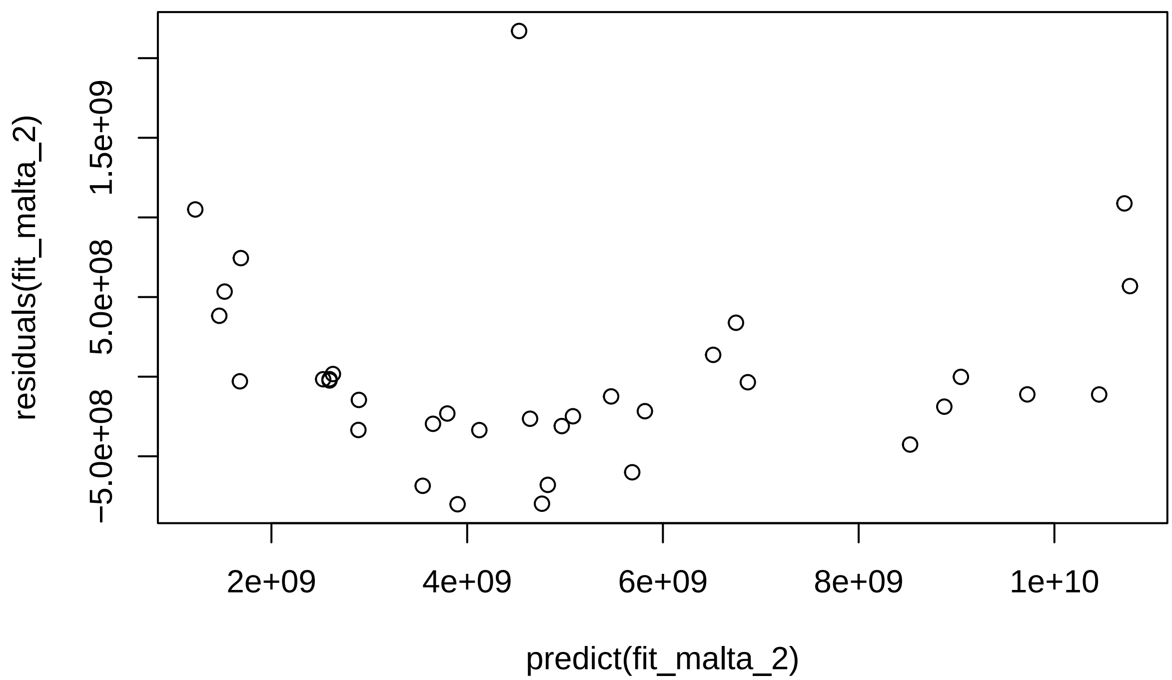


Figure 7: Plot of Fitted Values against Residuals

Model 4 - Interaction Term
The same approach as in the case of Singapore is also used for Malta. An interaction term is determined and used to create a new model. Given that Malta is a well-known tax haven and the level of FDI is well above the norm for the selected sample of countries, FDI, as economic indicator, is deemed to have an impact on the Maltese economy. Furthermore, Malta is a highly industrialised economy though its economy complexity is quite low. FDI is, therefore, used together with the variable, Manufacturing to create a new term.

```
# Model 4 - Interaction Term
# Introducing an interaction term - Manufacturing and FDI
df_malta_clean$FDIManf <- interaction(FDI, Manufacturing)

# Model with the interaction term
```

```r
fit_malta_interaction <- lm(GDP ~ FDIManf + FDI + Manufacturing, data = df_malta_clean)

summary(fit_malta_interaction)
```

```
##
## Call:
## lm(formula = GDP ~ FDIManf + FDI + Manufacturing, data = df_malta_clean)
##
## Residuals:
## ALL 35 residuals are 0: no residual degrees of freedom!
##
## Coefficients: (2 not defined because of singularities)
##                                               Estimate Std. Error t value
## (Intercept)                                  1.180e+10        NaN     NaN
## FDIManf32.1027329502695.7.16143634059413    -4.624e+08        NaN     NaN
## FDIManf1.31122389888756.8.35674965810255    -1.458e+09        NaN     NaN
## FDIManf4.72416455167554.9.42752915663499    -2.191e+09        NaN     NaN
## FDIManf34.586489580652.10.8189570991513     -2.760e+09        NaN     NaN
## FDIManf79.7489301739108.11.0339893098324    -3.117e+09        NaN     NaN
## FDIManf16.5416580886968.11.2315491501735    -3.704e+09        NaN     NaN
## FDIManf16.8221090205348.15.412192024912     -4.718e+09        NaN     NaN
## FDIManf-11.1735603550901.15.6223478328421 -4.971e+09        NaN     NaN
## FDIManf10.2407937772704.15.8474963153452    -5.153e+09        NaN     NaN
## FDIManf16.2220473163437.17.3999528882142    -6.204e+09        NaN     NaN
## FDIManf3.61741308862879.17.437147556391     -6.717e+09        NaN     NaN
## FDIManf8.06178133638723.17.921512429036     -6.457e+09        NaN     NaN
## FDIManf8.18828260585033.18.298800436205     -6.971e+09        NaN     NaN
## FDIManf10.5729391307934.18.6192880089163    -7.147e+09        NaN     NaN
## FDIManf18.4102128023528.19.4173407022303    -5.102e+09        NaN     NaN
## FDIManf2.08333282504.19.5698337918643       -7.659e+09        NaN     NaN
## FDIManf5.05895699511925.19.8490264567386    -7.425e+09        NaN     NaN
## FDIManf2.80264350439879.20.068774866345     -7.837e+09        NaN     NaN
## FDIManf3.27384097323115.21.3452295201444    -8.014e+09        NaN     NaN
## FDIManf1.17848185469097.21.8838264608161    -8.238e+09        NaN     NaN
## FDIManf2.02045509383544.22.1239239973457    -8.702e+09        NaN     NaN
## FDIManf2.43791619975848.22.3582049458706    -8.449e+09        NaN     NaN
## FDIManf1.10733328099375.22.7428015327328    -8.943e+09        NaN     NaN
## FDIManf2.10267793750403.24.0169925184038    -9.249e+09        NaN     NaN
## FDIManf1.5249330770729.24.0869958737782     -9.056e+09        NaN     NaN
## FDIManf1.69711244207837.24.4130260523667    -9.159e+09        NaN     NaN
## FDIManf2.38202662054004.24.7621706148056    -9.225e+09        NaN     NaN
## FDIManf1.69031972390541.24.7858602724039    -9.234e+09        NaN     NaN
## FDIManf3.14016589676441.25.461103534523     -9.291e+09        NaN     NaN
## FDIManf2.13035896867435.26.9434267168536    -9.372e+09        NaN     NaN
## FDIManf2.67326179599224.27.6222036050924    -1.015e+10        NaN     NaN
## FDIManf2.95251971902326.27.7176817769284    -9.954e+09        NaN     NaN
## FDIManf2.71358347029766.27.8470535825316    -9.747e+09        NaN     NaN
## FDIManf1.61590583741528.28.094674927059     -9.532e+09        NaN     NaN
## FDI                                                 NA         NA      NA
## Manufacturing                                       NA         NA      NA
##                                             Pr(>|t|)
## (Intercept)                                      NaN
## FDIManf32.1027329502695.7.16143634059413         NaN
```

```
## FDIManf1.31122389888756.8.35674965810255       NaN
## FDIManf4.72416455167554.9.42752915663499       NaN
## FDIManf34.586489580652.10.8189570991513        NaN
## FDIManf79.7489301739108.11.0339893098324       NaN
## FDIManf16.5416580886968.11.2315491501735       NaN
## FDIManf16.8221090205348.15.412192024912        NaN
## FDIManf-11.1735603550901.15.6223478328421      NaN
## FDIManf10.2407937772704.15.8474963153452       NaN
## FDIManf16.2220473163437.17.3999528882142       NaN
## FDIManf3.61741308862879.17.437147556391        NaN
## FDIManf8.06178133638723.17.921512429036        NaN
## FDIManf8.18828260585033.18.298800436205        NaN
## FDIManf10.5729391307934.18.6192880089163       NaN
## FDIManf18.4102128023528.19.4173407022303       NaN
## FDIManf2.08333282504.19.5698337918643          NaN
## FDIManf5.05895699511925.19.8490264567386       NaN
## FDIManf2.80264350439879.20.068774866345        NaN
## FDIManf3.27384097323115.21.3452295201444       NaN
## FDIManf1.17848185469097.21.8838264608161       NaN
## FDIManf2.02045509383544.22.1239239973457       NaN
## FDIManf2.43791619975848.22.3582049458706       NaN
## FDIManf1.10733328099375.22.7428015327328       NaN
## FDIManf2.10267793750403.24.0169925184038       NaN
## FDIManf1.5249330770729.24.0869958737782        NaN
## FDIManf1.69711244207837.24.4130260523667       NaN
## FDIManf2.38202662054004.24.7621706148056       NaN
## FDIManf1.69031972390541.24.7858602724039       NaN
## FDIManf3.14016589676441.25.461103534523        NaN
## FDIManf2.13035896867435.26.9434267168536       NaN
## FDIManf2.67326179599224.27.6222036050924       NaN
## FDIManf2.95251971902326.27.7176817769284       NaN
## FDIManf2.71358347029766.27.8470535825316       NaN
## FDIManf1.61590583741528.28.094674927059        NaN
## FDI                                             NA
## Manufacturing                                   NA
##
## Residual standard error: NaN on 0 degrees of freedom
## Multiple R-squared:      1,  Adjusted R-squared:    NaN
## F-statistic:   NaN on 34 and 0 DF,  p-value: NA
```

```
broom::glance(fit_malta_interaction)
```

```
## # A tibble: 1 x 12
##   r.squared adj.r.squared sigma statistic p.value    df logLik   AIC   BIC
##       <dbl>         <dbl> <dbl>     <dbl>   <dbl> <dbl>  <dbl> <dbl> <dbl>
## 1         1           NaN   NaN       NaN     NaN    34    Inf  -Inf  -Inf
## # i 3 more variables: deviance <dbl>, df.residual <int>, nobs <int>
```

This new model, fit_malta_interaction, does not solve the multicollinearity problem as well. Similar to Singapore, a polynomial regression need to be tried on. In the case of Malta, the economic indicator, FDI must have a significant impact on GDP. A model with a quadratic form will be created with FDI and Current Account as it is the only explanatory variable having weak correlation with the other variables.

Model 5 - Polynomial Regression

```r
attach(df_malta_clean)
```

```
## The following objects are masked from df_malta_clean (pos = 3):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_malta_clean (pos = 4):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_singapore (pos = 5):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_singapore (pos = 6):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_singapore (pos = 7):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop
```

```r
# Model 5 - Polynomial Regression
quad.malta2 <- GDP ~ FDI + I(FDI^2) + `Current Account`

fit_malta_quadratic <- lm(quad.malta2, data = df_malta_clean)

summary(fit_malta_quadratic)
```

```
##
## Call:
## lm(formula = quad.malta2, data = df_malta_clean)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -2.614e+09 -1.323e+09 -5.028e+08  1.611e+08  6.414e+09
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)       3542705936  530990700   6.672 1.84e-07 ***
## FDI                236192507   67463726   3.501  0.00143 **
## I(FDI^2)            -2033860     940661  -2.162  0.03844 *
## `Current Account`    9882587   80448346   0.123  0.90302
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 2.338e+09 on 31 degrees of freedom
## Multiple R-squared:  0.4048, Adjusted R-squared:  0.3472
## F-statistic: 7.029 on 3 and 31 DF,  p-value: 0.0009698
```

```r
broom::glance(fit_malta_quadratic)
```

```
## # A tibble: 1 x 12
##   r.squared adj.r.squared       sigma statistic p.value    df logLik   AIC   BIC
##       <dbl>         <dbl>       <dbl>     <dbl>   <dbl> <dbl>  <dbl> <dbl> <dbl>
## 1     0.405         0.347 2338266805.      7.03 9.70e-4     3  -803. 1615. 1623.
## # i 3 more variables: deviance <dbl>, df.residual <int>, nobs <int>
```

The quadratic model does not seem to be a better fit. Both the Adjusted R Squared and the AIC value indicates this model is worse than the previous ones.

Assessment Of Sequential Correlation
The model, fit_malta2 is the best one. It is possible that sequential correlation is distorting the results. The acf() is going to be applied to check for autocorrelation.

```r
acf(residuals(fit_malta_2))
```

At lag 1, the spike goes beyond the confidence bound indicating significant auto correlation or sequential correlation. Also, as there are regular spikes at certain lags, there is also a seasonality effect. The ARIMA model would, therefore, be a better model to use for Malta as well.

The following section will cover the analysis on the remaining country, Mauritius. The same approach is going to be used

C.Mauritius

C.1 Exploratory Data Analysis

```r
summary(df_mauritius)
```

```
##     iso2c              iso3c              country                date     
##  Length:44          Length:44          Length:44          Min.   :1976  
##  Class :character   Class :character   Class :character   1st Qu.:1987  
##  Mode  :character   Mode  :character   Mode  :character   Median :1998  
##                                                           Mean   :1998  
##                                                           3rd Qu.:2008  
##                                                           Max.   :2019  
##  Current Account        FDI          Domestic Investment Manufacturing  
##  Min.   :-13.546   Min.   :-0.5919   Min.   :17.23       Min.   :10.52  
##  1st Qu.: -6.600   1st Qu.: 0.4442   1st Qu.:21.61       1st Qu.:13.21  
##  Median : -4.169   Median : 0.8592   Median :24.65       Median :15.65  
##  Mean   : -4.049   Mean   : 1.5454   Mean   :24.50       Mean   :16.26  
##  3rd Qu.: -1.453   3rd Qu.: 2.8607   3rd Qu.:28.06       3rd Qu.:19.77  
##  Max.   :  6.346   Max.   : 5.6207   Max.   :31.36       Max.   :20.75  
##       GDP              Working Pop   
##  Min.   :2.183e+09   Min.   :59.51  
##  1st Qu.:3.305e+09   1st Qu.:64.24  
##  Median :5.889e+09   Median :67.70  
##  Mean   :6.547e+09   Mean   :67.28  
##  3rd Qu.:9.366e+09   3rd Qu.:71.07  
##  Max.   :1.387e+10   Max.   :72.26  
```

In regards to Mauritius, the manufacturing sector's maximum contribution reaches the 20%. Its level of domestic investment has more or less reached the minimum threshold of 25%. However, given the size of the manufacturing sector of Mauritius, we can conclude that most of the local investment is not meant for the manufacturing sector. It is assumed that it is mainly channelled into the services sector. The level of FDI for Mauritius is quite negligible. The current account of Mauritius seems to be, most of the time, in deficit as its mean hovers around -4%, contrary to Malta and Singapore.

Deep Dive Analysis

```r
# Number of years the Mauritian Manufacturing sector contribution equalled or exceeded 20%
df_mauritius_manf <- df_mauritius %>%
  filter(Manufacturing >= 20)

nrow(df_mauritius_manf)
```

```
## [1] 9
```

```r
# Number of years the level of Domestic Investment exceeded or equalled to 25%
df_mauritius_dominv <- df_mauritius %>%
  filter(`Domestic Investment` >= 25)

nrow(df_mauritius_dominv)
```

```
## [1] 21
```

```r
# Number of years the level of Current Account deficit equalled or exceeded 5%
df_mauritius_currentacc <- df_mauritius %>%
  filter(`Current Account` <= -5)

nrow(df_mauritius_currentacc)
```

```
## [1] 15
```

A closer analysis indicates that Mauritius current account deficit exceeded or equalled 5% for 15 years, which is quite worrisome. The level of domestic investment in Mauritius reached or exceeded the minimum threshold of 25% for 21 years. Given that the country spends more than it earns, the positive impact of substantial flow of investment monies did not have the desired effects on the Mauritian economy. Also, its manufacturing sector's contribution to GDP reached the 20% threshold for 9 years only so that we can safely conclude that it was not able to significantly influence the economic growth of the Mauritian economy, over the long term. It may explain why Mauritius is finding it challenging to remain in the "high income" category contrary to Malta and Singapore.

Exploratory Graphics of Independent Variables with GDP

```r
attach(df_mauritius)
```

```
## The following objects are masked from df_malta_clean (pos = 3):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_malta_clean (pos = 4):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_malta_clean (pos = 5):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_singapore (pos = 6):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_singapore (pos = 7):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_singapore (pos = 8):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop
```

```r
# Plot Manufacturing against GDP
plot(Manufacturing, GDP)
lines(lowess(Manufacturing, GDP), col = "blue")
```

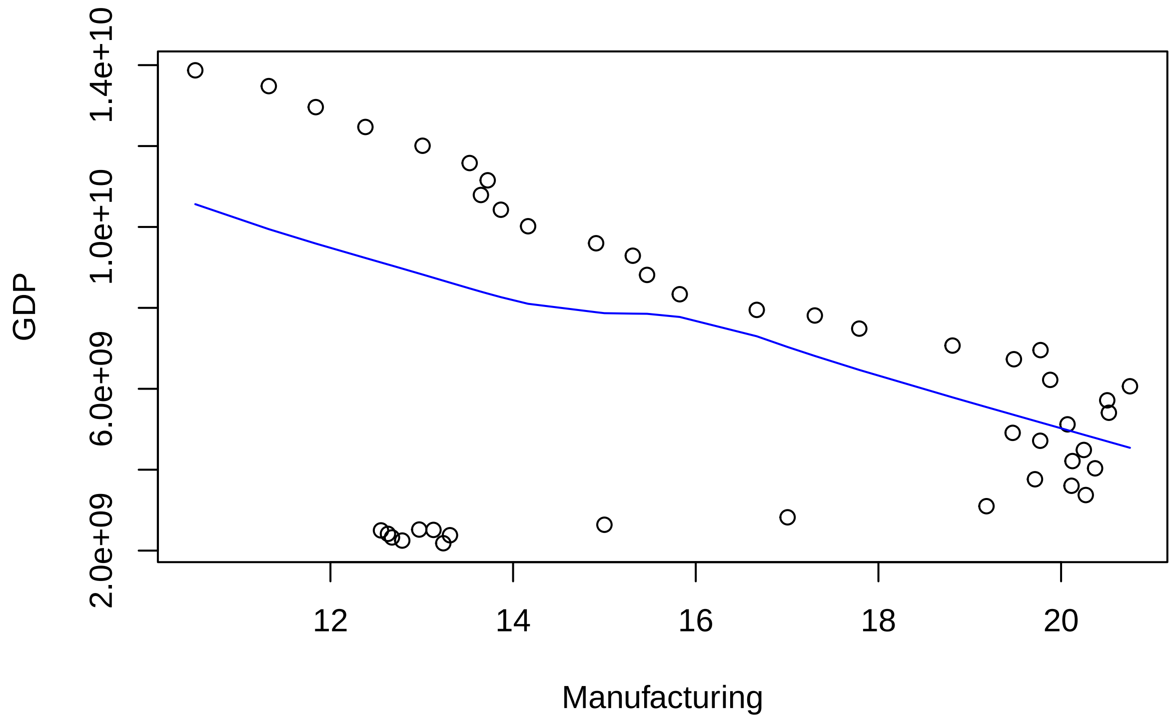


Figure 8: Exploratory Graphics of Independent Variables against GDP

```r
# Plot FDI against GDP
plot(FDI, GDP)
lines(lowess(FDI, GDP), col = "red")
```

```r
# Plot Working Population against GDP
plot(`Working Pop`, GDP)
lines(lowess(`Working Pop`, GDP), col = "green")
```

```r
# Plot Current Account against GDP
plot(`Current Account`, GDP)
lines(lowess(`Working Pop`, GDP), col = "grey")
```

```r
# Plot Domestic Investment against GDP
plot(`Domestic Investment`, GDP)
```

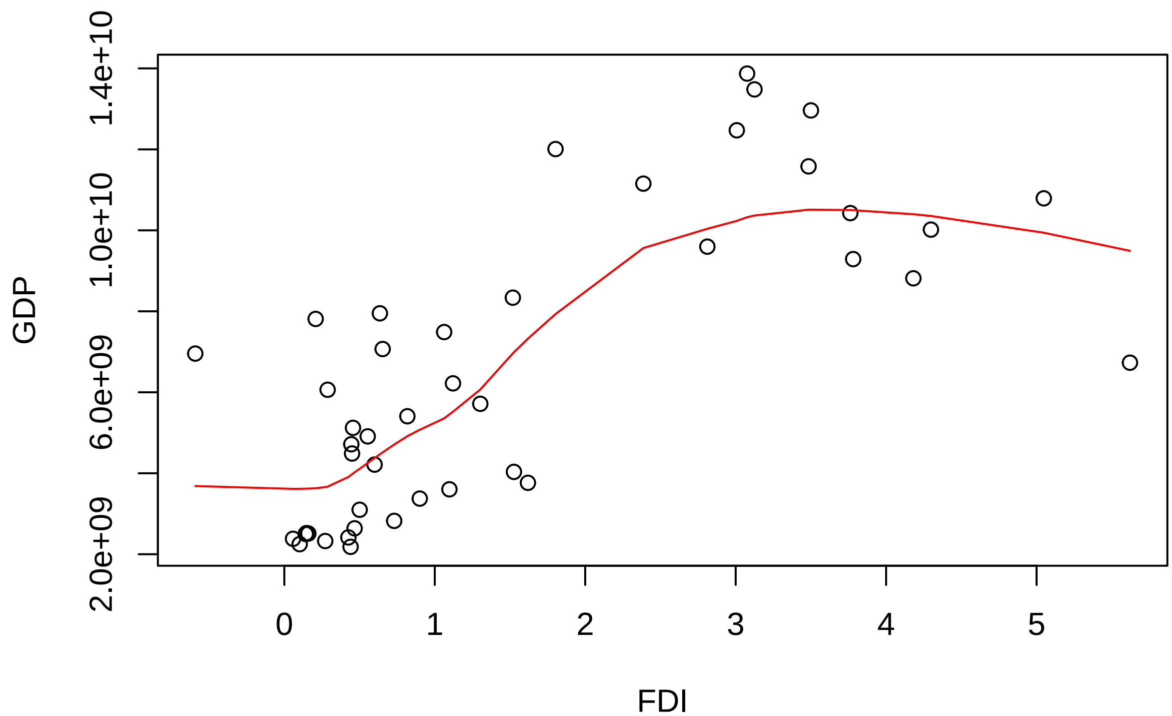


Figure 9: Exploratory Graphics of Independent Variables against GDP

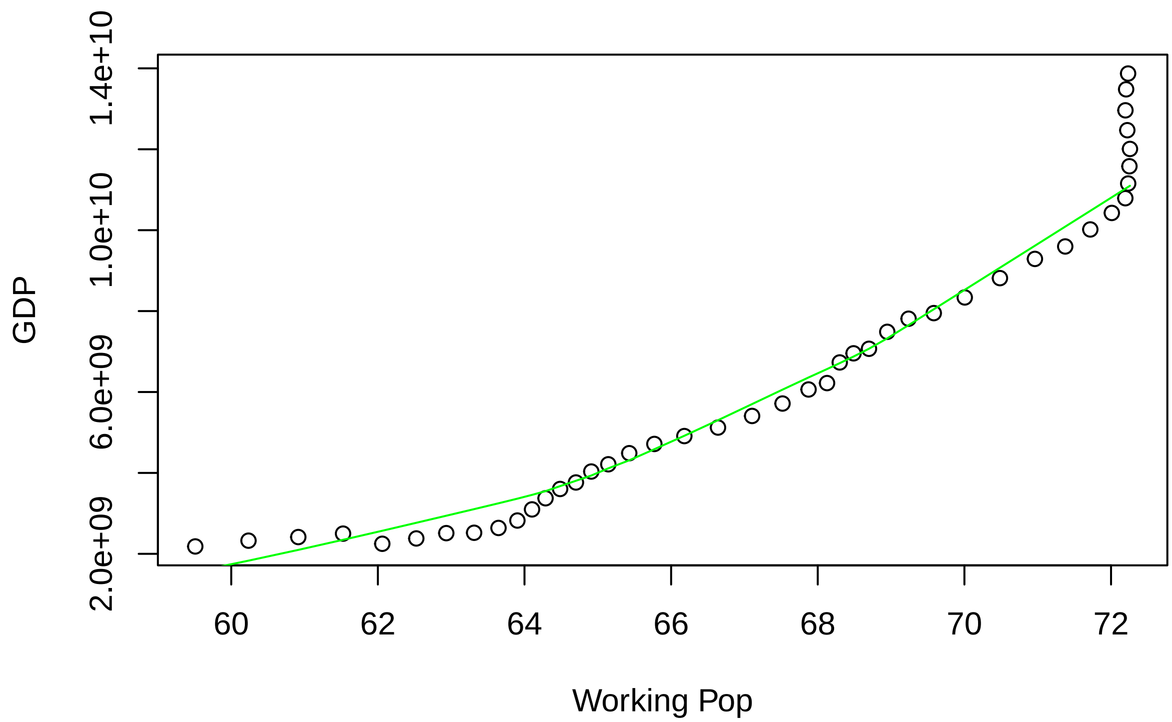


Figure 10: Exploratory Graphics of Independent Variables against GDP

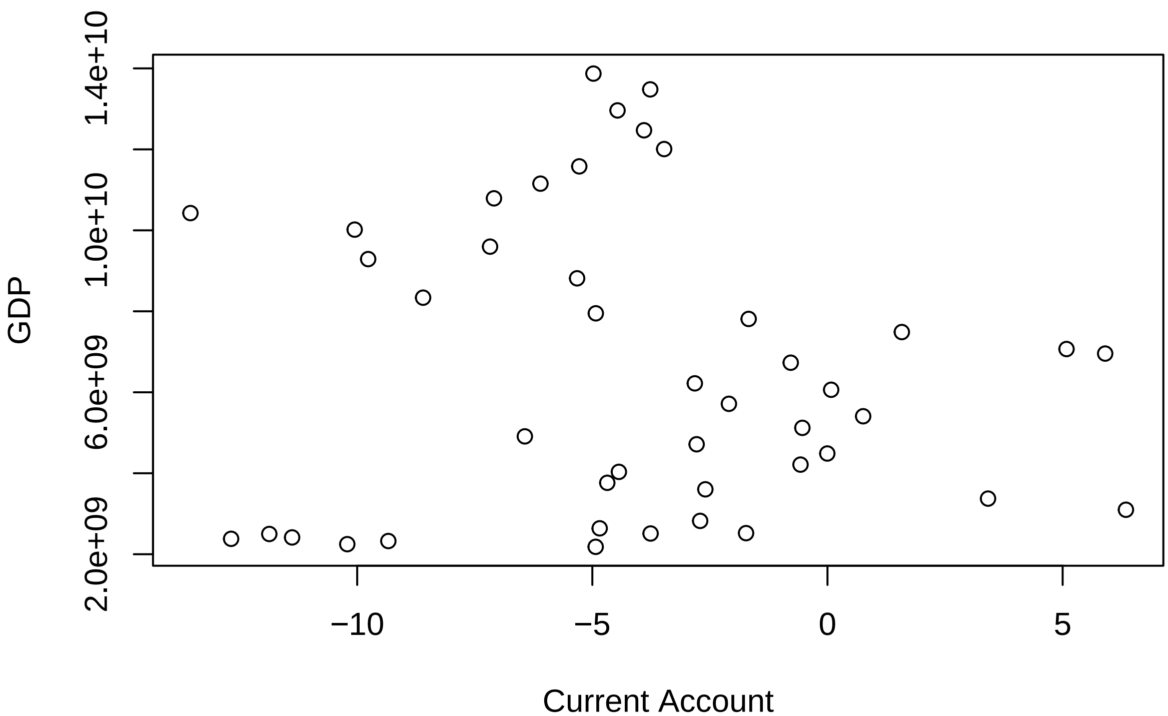


Figure 11: Exploratory Graphics of Independent Variables against GDP

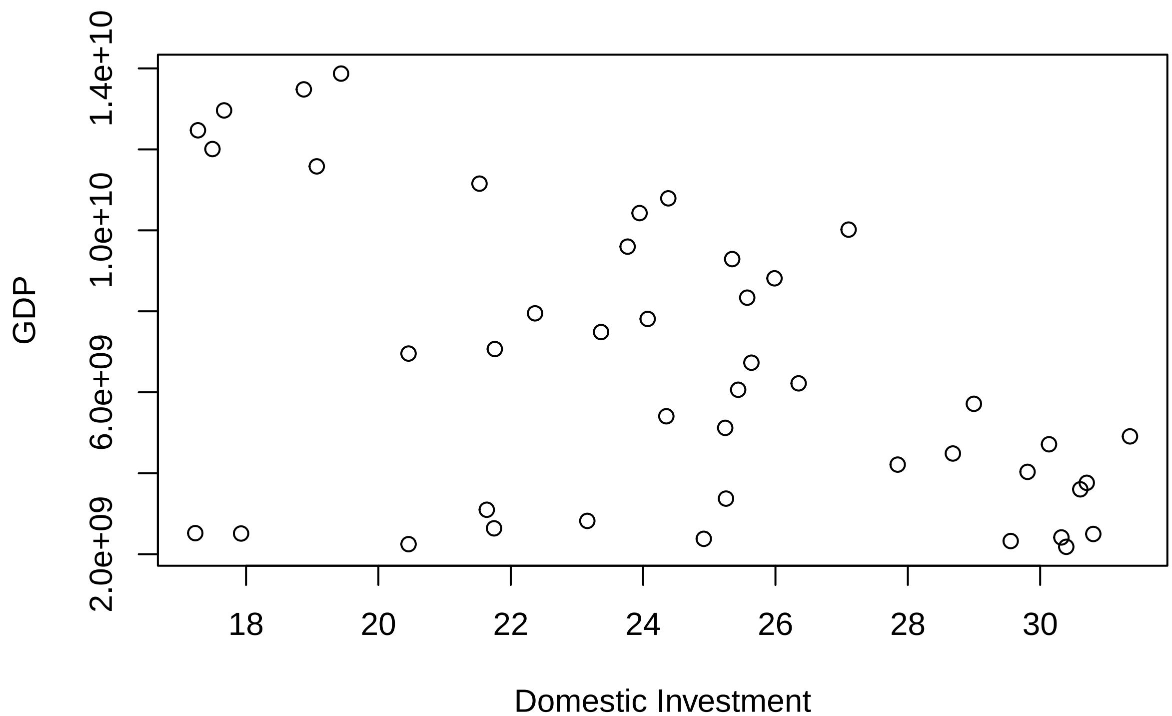
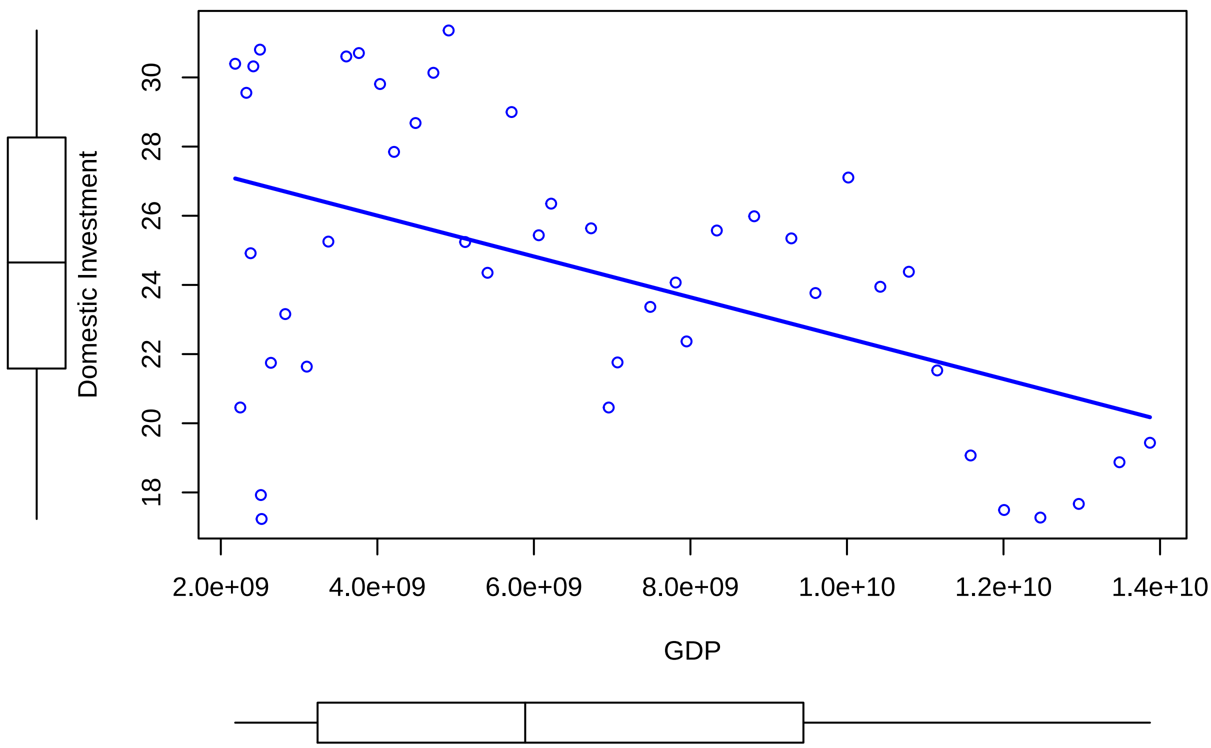


Figure 12: Exploratory Graphics of Independent Variables against GDP

```r
scatterplot(`Domestic Investment` ~ GDP, data = df_mauritius, smooth = FALSE, grid = FALSE)
```



Only Manufacturing has a linear negative relationship with GDP in the case of Mauritius. FDI, Domestic Investment and Current Account all have a cyclical relationship with GDP. Thus, in case of economic slowdown, the level of demand for goods and services also contracts as well as the level of investments (both local and foreign). As for Working Population, it is a positive relationship though the curve starts to go backwards indicating diminishing returns of labour, similar to the other two countries.

C.2 Correlation Tests

```r
attach(df_mauritius)
```

```
## The following objects are masked from df_mauritius (pos = 3):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_malta_clean (pos = 4):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_malta_clean (pos = 5):
```

```
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_malta_clean (pos = 6):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_singapore (pos = 7):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_singapore (pos = 8):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_singapore (pos = 9):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop
```

```r
# Correlation between FDI and GDP
cor(FDI, GDP, method = "spearman")
```

```
## [1] 0.7465821
```

```r
# Correlation between Domestic Investment and GDP
cor(`Domestic Investment`, GDP, method = "spearman")
```

```
## [1] -0.4259338
```

```r
# Correlation between Manufacturing and GDP
cor(Manufacturing, GDP, method = "spearman")
```

```
## [1] -0.16716
```

```r
# Correlation between Working Population and GDP
cor(`Working Pop`, GDP, method = "spearman")
```

```
## [1] 0.9921071
```

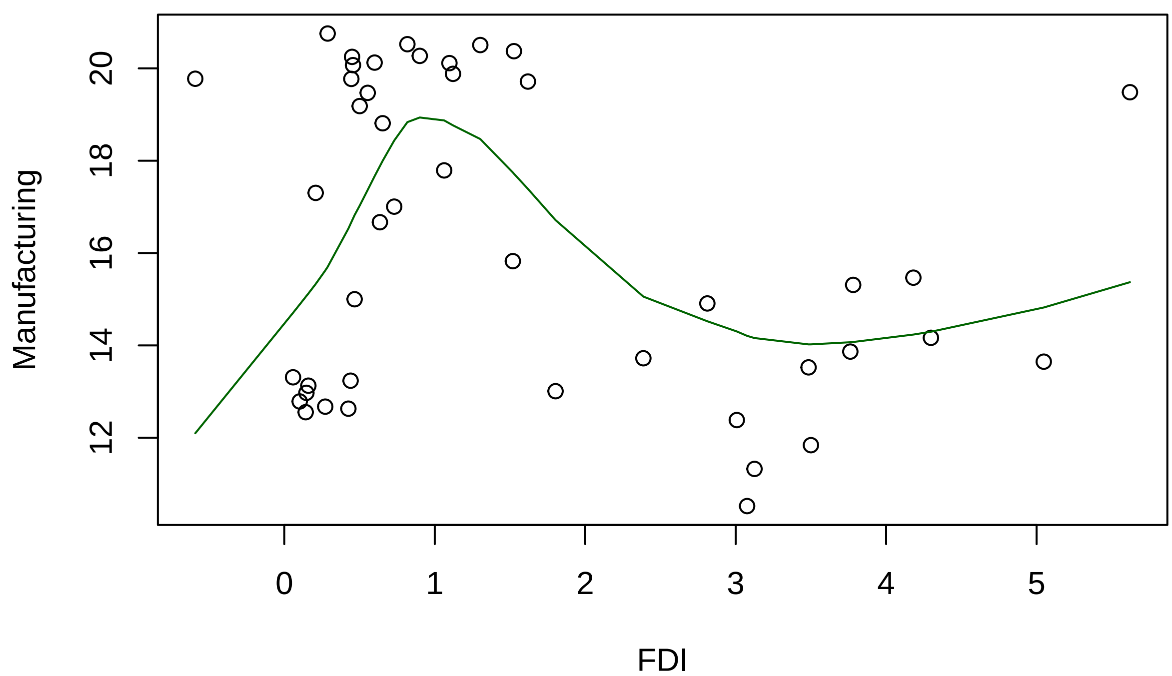
```r
# Correlation between Current Account and GDP
cor(`Current Account`, GDP, method = "spearman")
```

```
## [1] -0.04171952
```

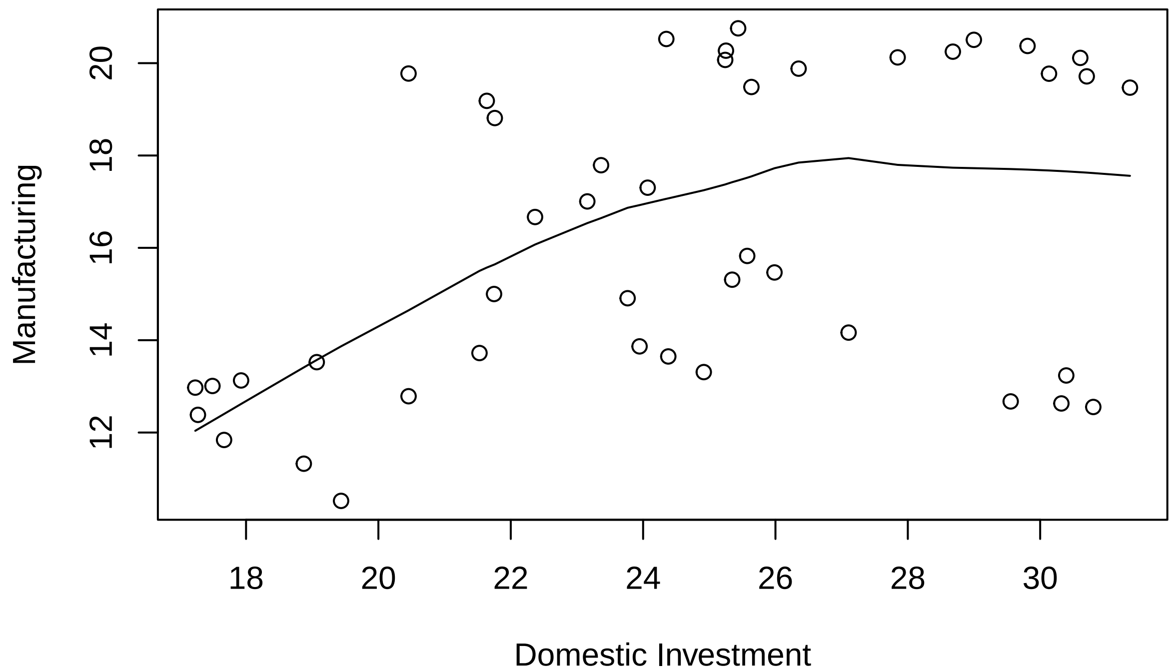
The correlation tests have been able to identify the correct relationship for Working Population and Manufacturing. As the remaining variables have a cyclical relationship, the correlation tests have failed.

The relationship between FDI and Manufacturing and the relationship between Domestic Investment and Manufacturing will now be explored.

```
# FDI and Manufacturing
plot(FDI, Manufacturing)
lines(lowess(FDI, Manufacturing), col = "darkgreen")
```




```
# Domestic Investment and Manufacturing
plot(`Domestic Investment`, Manufacturing)
lines(lowess(`Domestic Investment`, Manufacturing), col = "black")
```



FDI and Manufacturing have a cyclical relationship. This is most probably due to the national policies initiated by the government, providing financial,legal and fiscal incentives to investment in specific economic sectors. Mauritius started its economic development by setting up the textile manufacturing sector. More recently, the country has been focusing on the services sector such as tourism, financial services, business process outsourcing. As a result, foreign investment has been focusing on those business activities rather than on the Manufacturing sector which has inevitably shrunk.

Domestic Investment has a positive relationship with Manufacturing though the curve has plateaued. This is probably due to the fact that local businesses are investing to maintain the productive capacity and not to expand the size of their economic activities. We are now going to do the correlation tests.

```
# Correlation between FDI and Manufacturing
cor(FDI, Manufacturing, method = "spearman")
```

```
## [1] -0.07864693
```

```
# Correlation between Domestic Investment and Manufacturing
cor(`Domestic Investment`, Manufacturing, method = "spearman")
```

```
## [1] 0.4369274
```

The correlation tests indicate that the relationship between FDI and manufacturing is weak and negative. Domestic investment and manufacturing has a weak positive relationship as expected, given the reluctance of local businesses to invest into new factories and more modern equipment.

C.3 Regression Analysis

We start by including all indicators in the first model for Mauritius.

Model 1 - With All Indicators
The output of the model, fit_mauritius, indicates that FDI is the only variable that does not have a significant coefficient. All the other variables have significant coefficients at 0.1%. The Adjusted R squared figure is close to 1 and the AIC value is 1908.

```
# Model 1 - With all indicators
fit_mauritius <- lm(GDP ~ `Working Pop` + FDI + `Current Account` + `Domestic Investment` + Manufacturi

summary(fit_mauritius)
```

```
##
## Call:
## lm(formula = GDP ~ `Working Pop` + FDI + `Current Account` +
##     `Domestic Investment` + Manufacturing, data = df_mauritius)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.172e+09 -3.786e+08  1.160e+08  3.401e+08  1.023e+09
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)             -5.123e+10  2.816e+09 -18.190  < 2e-16 ***
## `Working Pop`            9.269e+08  3.994e+07  23.209  < 2e-16 ***
## FDI                     -4.981e+07  9.103e+07  -0.547    0.587
## `Current Account`        1.727e+08  3.518e+07   4.907 1.77e-05 ***
## `Domestic Investment`    1.798e+08  3.953e+07   4.550 5.35e-05 ***
## Manufacturing           -5.051e+08  5.728e+07  -8.818 9.98e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 579500000 on 38 degrees of freedom
## Multiple R-squared:  0.9773, Adjusted R-squared:  0.9743
## F-statistic: 327.1 on 5 and 38 DF,  p-value: < 2.2e-16
```

```
broom::glance(fit_mauritius)
```

```
## # A tibble: 1 x 12
##   r.squared adj.r.squared      sigma statistic  p.value    df logLik   AIC   BIC
##       <dbl>         <dbl>      <dbl>     <dbl>    <dbl> <dbl>  <dbl> <dbl> <dbl>
## 1     0.977         0.974 579536474.      327. 3.88e-30     5  -947. 1908. 1921.
## # i 3 more variables: deviance <dbl>, df.residual <int>, nobs <int>
```

Assessment of Multicollinearity
All variables have a VIF value of less than 5 including that of FDI. However, the correlation matrix indicates that that there is multicollinearity amongst the variables.

```
# Calculating VIF
vif(fit_mauritius)
```

```
##         `Working Pop`                    FDI     `Current Account`
##              3.072313               2.566989              3.548674
## `Domestic Investment`          Manufacturing
##              3.663808               4.571786
```

```
# Correlation Matrix
cor_matrix_mtius <- cor(df_mauritius[c("FDI", "Working Pop", "Current Account", "Domestic Investment",

cor(cor_matrix_mtius)
```

```
##                            FDI Working Pop Current Account Domestic Investment
## FDI                  1.0000000   0.8798700      -0.6655564          -0.5972943
## Working Pop          0.8798700   1.0000000      -0.3084425          -0.8636386
## Current Account     -0.6655564  -0.3084425       1.0000000          -0.1709858
## Domestic Investment -0.5972943  -0.8636386      -0.1709858           1.0000000
## Manufacturing       -0.9505814  -0.8129844       0.7070833           0.5492512
##                     Manufacturing
## FDI                    -0.9505814
## Working Pop            -0.8129844
## Current Account         0.7070833
## Domestic Investment     0.5492512
## Manufacturing           1.0000000
```

Model 2 - Polynomial Regression with FDI

We now have to transform the variables so as to determine a new model. FDI correlates with all the other explanatory variables. Given that it has a cyclical relationship with GDP, a polynomial regression will be calculated as the second model, fit_mauritius_quadratic.

```
# Model 2 - Polynomial Regression with FDI
quad.mtius1 <- GDP ~ FDI + I(FDI^2)

fit_mauritius_quadratic <- lm(quad.mtius1, data = df_mauritius)

summary(fit_mauritius_quadratic)
```

```
##
## Call:
## lm(formula = quad.mtius1, data = df_mauritius)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -4.134e+09 -1.539e+09 -3.832e+08  1.237e+09  6.644e+09
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2819597533  619353424   4.552 4.67e-05 ***
## FDI         3941313583  762005948   5.172 6.43e-06 ***
## I(FDI^2)    -497171857  158607365  -3.135  0.00318 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.325e+09 on 41 degrees of freedom
## Multiple R-squared:  0.6058, Adjusted R-squared:  0.5866
## F-statistic: 31.51 on 2 and 41 DF,  p-value: 5.143e-09
```

```r
broom::glance(fit_mauritius_quadratic)
```

```
## # A tibble: 1 x 12
##   r.squared adj.r.squared       sigma statistic p.value    df logLik   AIC   BIC
##       <dbl>         <dbl>       <dbl>     <dbl>   <dbl> <dbl>  <dbl> <dbl> <dbl>
## 1     0.606         0.587 2324546303.      31.5 5.14e-9     2 -1010. 2028. 2035.
## # i 3 more variables: deviance <dbl>, df.residual <int>, nobs <int>
```

All the coefficients estimates of the independent variables of the quadratic model are significant. FDI and the intercept are significant at 1%. The adjusted R Squared stands at 0.6. However, the AIC value, 2208, is higher implying that this polynomial regression is not as good as the first one.

Model 3 - With 2 Indicators: Current Account and Working Population
Introducing an interaction term as we have seen for the first two countries does not resolve the multicollinear problem. Referring to the multi correlation matrix, we can note that the variable, Current Account is the least correlated with the other variables. Working Population is correlated with most of the variables except for Current Account. A third model is now going to be calculated using these two variables.

```r
# Model 3 - With 2 indicators - Current Account and Working Population
fit_mauritius2 <- lm(GDP ~ `Current Account` + `Working Pop`, data = df_mauritius)

summary(fit_mauritius2)
```

```
##
## Call:
## lm(formula = GDP ~ `Current Account` + `Working Pop`, data = df_mauritius)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.050e+09 -6.866e+08 -3.280e+08  1.066e+08  2.843e+09
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)        -5.371e+10  2.767e+09 -19.411   <2e-16 ***
## `Current Account` -7.544e+07  3.359e+07  -2.246   0.0302 *
## `Working Pop`      8.911e+08  4.098e+07  21.745   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.042e+09 on 41 degrees of freedom
## Multiple R-squared:  0.9208, Adjusted R-squared:  0.9169
## F-statistic: 238.2 on 2 and 41 DF,  p-value: < 2.2e-16
```

```r
broom::glance(fit_mauritius2)
```

```
## # A tibble: 1 x 12
##   r.squared adj.r.squared      sigma statistic  p.value    df logLik   AIC   BIC
##       <dbl>         <dbl>      <dbl>     <dbl>    <dbl> <dbl>  <dbl> <dbl> <dbl>
## 1     0.921         0.917     1.04e9      238. 2.67e-23     2  -975. 1957. 1964.
## # i 3 more variables: deviance <dbl>, df.residual <int>, nobs <int>
```

The third model does not provide much improvement from the first one. The AIC value is higher: 1953 and adjusted R Squared is 0.92. In this case, the coefficient estimate for the Working Population and the intercept are significant at 0.1% and for the Current Account at 5%.

Model 4 - Excluding FDI

It seems that the first model, fit_mauritius is the best amongst the three models. Given that the coefficient estimate for FDI is not significant, it shall be removed and a new model will be calculated with the rest of the variables: fit_mauritius.excFDI

```r
# Model 4 - Exc FDI
fit_mauritius.excFDI <- lm(GDP ~ `Working Pop` + `Current Account` + `Domestic Investment` + Manufacturi

summary(fit_mauritius.excFDI)
```

```
##
## Call:
## lm(formula = GDP ~ 'Working Pop' + 'Current Account' + 'Domestic Investment' +
##     Manufacturing, data = df_mauritius)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.187e+09 -3.494e+08  1.161e+08  3.343e+08  9.985e+08
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)           -5.019e+10  2.054e+09 -24.433  < 2e-16 ***
## 'Working Pop'          9.106e+08  2.652e+07  34.339  < 2e-16 ***
## 'Current Account'      1.728e+08  3.487e+07   4.957 1.43e-05 ***
## 'Domestic Investment'  1.736e+08  3.749e+07   4.630 3.99e-05 ***
## Manufacturing         -4.974e+08  5.502e+07  -9.040 4.12e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 574300000 on 39 degrees of freedom
## Multiple R-squared:  0.9771, Adjusted R-squared:  0.9748
## F-statistic: 416.3 on 4 and 39 DF,  p-value: < 2.2e-16
```

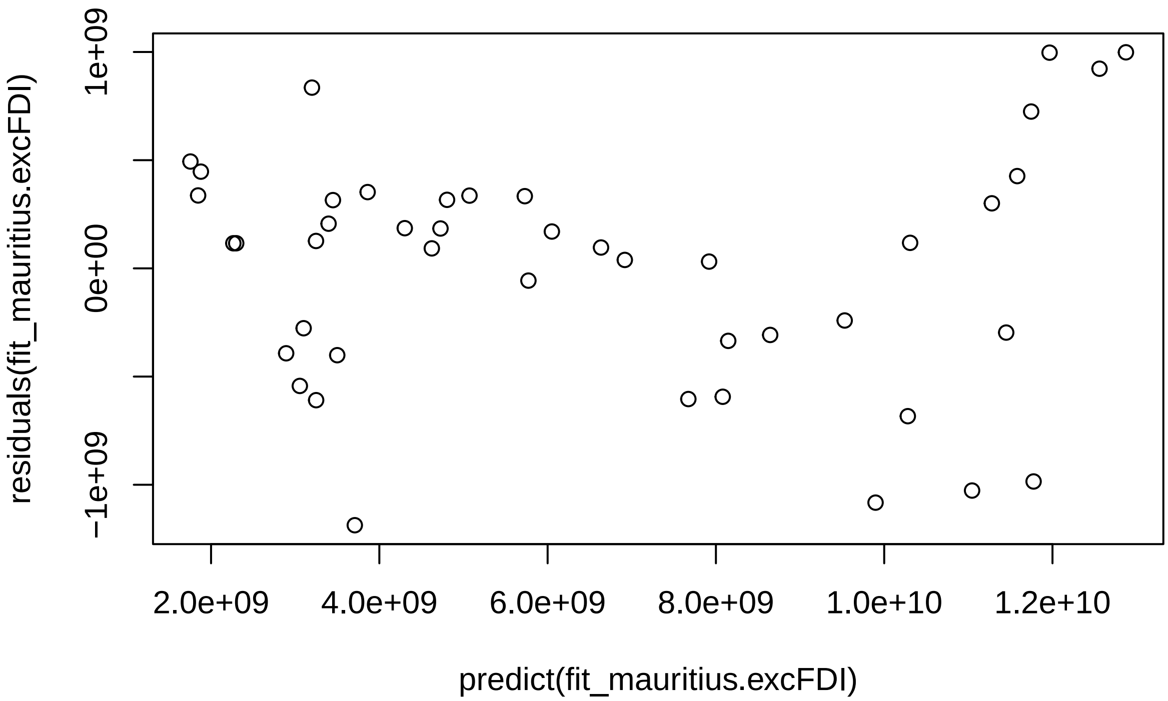
```r
broom::glance(fit_mauritius.excFDI)
```

```
## # A tibble: 1 x 12
##   r.squared adj.r.squared      sigma statistic  p.value    df logLik   AIC   BIC
##       <dbl>         <dbl>      <dbl>     <dbl>    <dbl> <dbl>  <dbl> <dbl> <dbl>
## 1     0.977         0.975 574307498.      416. 2.06e-31     4  -947. 1906. 1917.
## # i 3 more variables: deviance <dbl>, df.residual <int>, nobs <int>
```

This fourth model is an improvement from the first model. Its AIC value is the lowest - 1906. All coefficient estimates for the explanatory variables and the intercept are significant at 0.1% and its Adjusted R Squared is slightly higher. In this model, only Manufacturing has a negative coefficient. Mauritius has embarked on a strategy whereby is has been focusing on growing its services sector and relocating its textile manufacturing activities in other countries namely Bangladesh, China and Madagascar. Thus, as the size of the economy grows, the share of Manufacturing is on the decline. However, in the past two years, there has been attempts at reviving the manufacturing sector and these have not yet produced substantial results.

Assessment Of Sequential Correlation - Model 4

```r
# Plot fitted values against residuals
plot(predict(fit_mauritius.excFDI), residuals(fit_mauritius.excFDI))
```



There seems to be a pattern in the graphic of fitted values against the residuals implying that this may not be the right fit for Mauritius. Using the acf function does not work in this case as the data in fit_mauritius.excFDI is not evenly spaced. Using acf function in the *nlme* package would require that we use the gls or lme functions to create the model. The model will not be comparable with the previous models.

Analysis of Variance

We can use the anova function to compare the three models that have relatively high adjusted R Squared and lower AIC values.

```r
anova(fit_mauritius2, fit_mauritius.excFDI, fit_mauritius)
```

```
## Analysis of Variance Table
##
## Model 1: GDP ~ `Current Account` + `Working Pop`
## Model 2: GDP ~ `Working Pop` + `Current Account` + `Domestic Investment` +
##     Manufacturing
## Model 3: GDP ~ `Working Pop` + FDI + `Current Account` + `Domestic Investment` +
##     Manufacturing
##   Res.Df        RSS Df  Sum of Sq       F    Pr(>F)
## 1     41 4.4533e+19
## 2     39 1.2863e+19  2 3.1670e+19 47.1470 5.089e-11 ***
## 3     38 1.2763e+19  1 1.0056e+17  0.2994    0.5875
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Based on the anova results, it seems that the model excluding the FDI variable is the significant one. However, given that the graphic of fitted values against residuals demonstrate a pattern, it is not the best fit model.

6. Cross Validation of Results

The number of observations for each country is quite limited: 35 - 44. The approach of splitting the sample into two sets: a training set and another one, a test set is not possible. In this case, the bootstrap method can be used to assess the quality of the model chosen for each country.

A function, boot.fn() is first created which takes in the data set for each country as well as a set of indices for the observations, and returns the intercept and the estimates of the coefficients of each chosen predictor. This function is then applied to the full set of observations in order to compute the estimates of the various coefficients. The boot() function is then used to compute the standard errors of the bootstrap estimates for the intercept and coefficient estimates of the chosen predictors.

A. Singapore

```r
attach(df_singapore)
```

```
## The following objects are masked from df_mauritius (pos = 3):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_mauritius (pos = 4):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_malta_clean (pos = 5):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_malta_clean (pos = 6):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_malta_clean (pos = 7):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop
```

```
## The following objects are masked from df_singapore (pos = 8):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_singapore (pos = 9):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop

## The following objects are masked from df_singapore (pos = 10):
##
##     country, Current Account, date, Domestic Investment, FDI, GDP,
##     iso2c, iso3c, Manufacturing, Working Pop
```

```
library(boot)
```

```
##
## Attaching package: 'boot'

## The following object is masked from 'package:car':
##
##     logit
```

```
boot.fn_sing <- function(data, index) {
  return(coef(lm(quad.sing.3, data = df_singapore)))
}

boot.fn_sing(df_singapore, 1:44)
```

```
##                (Intercept)                        FDI
##              1985281691751                 6468994792
##                   I(FDI^2)         log(Manufacturing)
##                  -42988932              -412628053945
## log(`Domestic Investment`)
##              -183891833336
```

```
boot(df_singapore, boot.fn_sing, 44)
```

```
##
## ORDINARY NONPARAMETRIC BOOTSTRAP
##
##
## Call:
## boot(data = df_singapore, statistic = boot.fn_sing, R = 44)
##
##
## Bootstrap Statistics :
##          original  bias    std. error
## t1* 1985281691751       0           0
## t2*    6468994792       0           0
## t3*     -42988932       0           0
## t4* -412628053945       0           0
## t5* -183891833336       0           0
```

We note that for both bias and standard errors values, they are all zero, indicating that this is the right explanatory model for Singapore's GDP.

The same function can be applied to the data related to Mauritius and Malta to cross validate the results. However, given the limitations of the lm() function, it is best that a more in depth analysis is done using the ARIMA models before any cross validation is done.

7.Concluding Remarks

It is quite surprising to find that all three countries regression model demonstrates a negative relationship between the economic indicator Manufacturing and GDP. However, this is a partial analysis. Further analysis is needed to remove the sequential correlation by using the ARIMA model. Also, in terms of the size of the dataset, it may be necessary to use more data such as quarterly instead of yearly. Unfortunately, these data are not available on the World Bank database.

From the series of correlations tests effected for each country, it can be inferred that Singapore has undoubtedly been investing heavily in its economy, strengthening its manufacturing sector as the latter's contribution has exceeded the 20% mark for most of the time, creating a virtuous circle of economic growth which may have contributed to the improvement of its economic complexity for the past few decades. It is also important to note that Singapore is known to have a large community of expatriates – a pool of skilled labour to offset the effects of its ageing population and the lack of local skills for the set up of a number of its existing industries. Though its economy is small, it has capitalised on its geographic location to set up a trading hub for the region targeting neighbouring countries such as Malaysia, Indonesia, Thailand which are bigger markets.

Similarly, Malta has also been investing heavily into its economy though its economic complexity is relatively low as it is not included in the Economic Complexity Index. However, it has taken advantage of its geographic proximity with the European continent and has also been promoting its offshore sector. Its consistent approach of capitalising on its geographic location and attracting foreign funds have helped the country to develop its industrial sectors and its tourism sector. Being of small size, the tourism sector has undoubtedly played an important role in the development of the economy.

Mauritius suffers continuously from the negative impact of its current account deficit which is quite substantial compared to that of Singapore and Malta. This is why though the level of investment is quite high, its economy is unable to grow because most of the country's earnings are spent rather than invested.The coefficient estimate for the Current Account variable is positive implying that a large deficit hinders the economic growth of the country and vice versa. Also, its working population is ageing and the country is also suffering from brain drain. It has only recently addressed this issue by allowing foreign labour. The government is also trying to address the gender gap as there is still for improvement as regards to the role of women in the workforce.

8.References

1. Introduction to Data Science by Rafael A. Izizarry
2. *An Introduction to Statistical Learning: With Applications in R (Springer Texts in Statistics Book 103)* by James Gareth, Daniela Witten, Trevor Hastie, Robert Tibshirani
3. The Role of Manufacturing and Services Sector in Economic Growth: An Empirical Study of Developing Countries by Eman Attiah. *European Research Studies Journal*
4. The Observatory of Economic Complexity
5. Linkages and Economic Development by Dominick Bartelme and Yuriy Gorodnichenko. *National Bureau of Economic Research*

6. Breakout Nations: In Pursuit of the Next Economic Miracles by Ruchir Sharma

7. A Practical Guide to Data Analysis Using R: An Example-Based Approach by John H Maindonald, W. John Braun, Jeffrey L. Andrews

8. R Graphics Cookbook, 2nd Edition, by Winston Chang