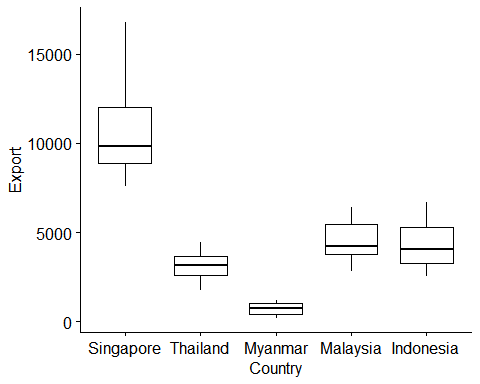
1. **Find the statistical difference between exports of all countries**

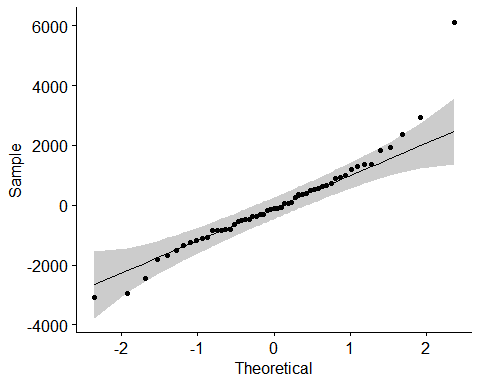
The ANOVA test is used to compare the mean of multiple group.

* **Assumptions**
* Independence of observation
* No significant outliers
* Normality
* Homogeneity of variance
* **Outliers**



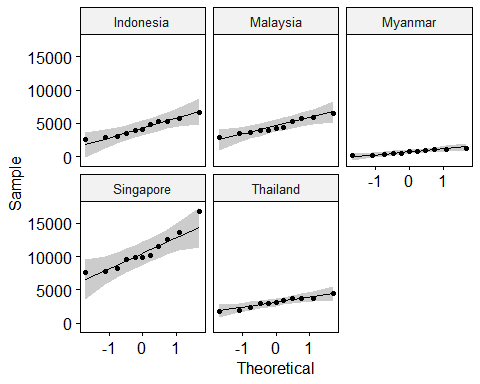
**From the above boxplot we can say that there are not significant outliers.**

* **Normality (Individual)**



**From the above QQplot since almost all the points are lying on the straight line we can say the data is normally distributed.**

* **Normality (Group)**



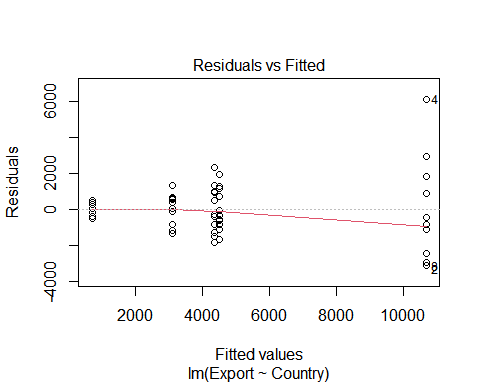
**All the points falls approximately along the reference line, for each cell. So, we can assume normality of the data.**

**Shapiro-Wilk test**

Country variable statistic p value  
## <chr> <chr> <dbl> <dbl>  
## 1 Indonesia Export 0.960 0.772  
## 2 Malaysia Export 0.952 0.664  
## 3 Myanmar Export 0.925 0.366  
## 4 Singapore Export 0.910 0.246  
## 5 Thailand Export 0.946 0.592

**The data is normally distributed (p>0.05) for each Country, as assessed by Shapiro-Wilk test of normality.**

* **Homogeneity of Variance assumption**



**In the above plot, there is no evident relationships between residuals and fitted values. So, we can assume the homogeneity of variances.**

* **One Way ANOVA**

**H0:** The mean of all groups are equal.

**H1:** At least on mean is not equal to other.

* **ANOVA Table**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Effect | DFn | DFd | F | p p<.05 |
| Country | 4 | 50 | 64.439 | 4.08e-19 |

**Since p-value < 0.05 we conclude that there are significant differences between the groups.**

1. **Find the statistical difference between population of Singapore and Thailand**

To compare the mean of two group we use t test.

**H0:** The mean of two group are equal.

**H1:** Not H0

## Welch Two Sample t-test  
##   
## data: data\_task\_sing$Population and data\_task\_thai$Population  
## t = -201.34, df = 11.528, p-value < 2.2e-16  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -63397170 -62033623  
## sample estimates:  
## mean of x mean of y   
## 5332902 68048299

**Since p- value < 0.05, we conclude that the mean of two group is not equal.**

1. **Find the impact of population, GDP on total trade of the countries**

* **Model: total\_trade = 1.321100e+03 -2.313503e+01\*`Population\_in \_million` -5.386757e-03\*`GDP \_(current \_USMillion$)`**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 1.321100e+03 | 6.455400e+02 | 2.046504 | 0.045777442 |
| `Population\_in \_million` | -2.313503e+01 | 7.850921e+00 | -2.946792 | 0.004796835 |
| `GDP \_(current \_USMillion$)` | -5.386757e-03 | 2.534283e-03 | -2.125554 | 0.038309677 |

Residual standard error: 2725 on 52 degrees of freedom

Multiple R-squared: 0.6177, Adjusted R-squared: 0.603

F-statistic: 42 on 2 and 52 DF, p-value: 1.391e-11

* **Interpretation :**

p-value of F-Statistics is 1.391e-11, which is highly significant. This means that, at least one of the predictor variable is highly related to the outcome variable.

Adjusted R-squared: 0.603 meaning that 60% of the variance in the measure of total trade can be predicted by Population & GDP.

**Total trade of Countries is highly impacted by GDP followed by Population.**

1. **Create a time series chart for total trade**

* **From the below time series plot,**

1. **Total trade increases for country Myanmar**
2. **For Singapore till 2011-12 there was increase in total trade and Total trade decreases for countries Thailand, Indonesia, Malaysia.**
3. **decreases afterwards.**

