Compare

# DoDownload the data

setwd('C:/Users/VerkhovodTS/Desktop/R')  
f1 <- read.csv2('gdp\_test2.csv', header = TRUE, encoding = 'UNICOD')  
f2 <- read.csv2('GDP\_fit.csv', header = TRUE, encoding = 'UNICOD')  
f2 <- f2[-1]  
f <- dplyr::bind\_cols(f1, f2)  
f <- f[,-1]  
head(f)

## Country GDP Population Unemployment Area..km.2  
## 1 Armenia 13672802158 2957731 16.99 29800  
## 2 Austria 446314739528 8877067 4.67 83871  
## 3 Belgium 529606710418 11484055 5.59 32545  
## 4 Bahrain 38574069149 1641172 0.71 695  
## 5 Bahamas, The 12827000000 389482 10.36 13940  
## 6 Belarus 63080457023 9466856 4.59 207600  
## Average.life.expectancy..years Constitutional.form p\_sr  
## 1 72.4 Republic 46436244085  
## 2 79.3 Republic 103784685305  
## 3 79.0 Constitutional monarchy 129042028034  
## 4 74.8 Constitutional monarchy 33680994776  
## 5 65.7 Constitutional monarchy 21554217315  
## 6 70.2 Republic 109498751849  
## p\_mr p\_pr p\_dt p\_rf  
## 1 98042633247 136800418159 84027302363 15327125860  
## 2 275624979933 141028028181 84027302363 80138440858  
## 3 281828223536 144205207894 84027302363 31747431133  
## 4 125878688709 136431871990 84027302363 18646311719  
## 5 -43838422475 136276759010 84027302363 5940473558  
## 6 138411791713 141677081440 84027302363 80138440858

# Compare models

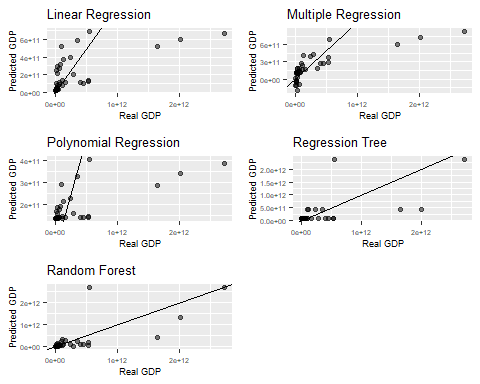
library(gridExtra)

## Warning: package 'gridExtra' was built under R version 3.6.3

library(ggplot2)

## Warning: package 'ggplot2' was built under R version 3.6.3

g\_sr <- ggplot(f, aes(x=GDP, y=p\_sr)) +   
 geom\_abline(intercept=0, slope=1) +  
 geom\_point(alpha=0.5) + labs(title="Linear Regression", x="Real GDP", y="Predicted GDP") +   
 theme(plot.title=element\_text(size=10), axis.title.x=element\_text(size=7), axis.title.y=element\_text(size=7), axis.text.x=element\_text(size=5), axis.text.y=element\_text(size=5)) + theme(legend.position="none")  
  
g\_mr <- ggplot(f, aes(x=GDP, y=p\_mr)) +   
 geom\_abline(intercept=0, slope=1) +  
 geom\_point(alpha=0.5) + labs(title="Multiple Regression", x="Real GDP", y="Predicted GDP") +   
 theme(plot.title=element\_text(size=10), axis.title.x=element\_text(size=7), axis.title.y=element\_text(size=7), axis.text.x=element\_text(size=5), axis.text.y=element\_text(size=5)) + theme(legend.position="none")  
  
g\_pr <- ggplot(f, aes(x=GDP, y=p\_pr)) +   
 geom\_abline(intercept=0, slope=1) +  
 geom\_point(alpha=0.5) + labs(title="Polynomial Regression", x="Real GDP", y="Predicted GDP") +   
 theme(plot.title=element\_text(size=10), axis.title.x=element\_text(size=7), axis.title.y=element\_text(size=7), axis.text.x=element\_text(size=5), axis.text.y=element\_text(size=5)) + theme(legend.position="none")   
  
g\_dt <- ggplot(f, aes(x=GDP, y=p\_dt)) +   
 geom\_abline(intercept=0, slope=1) +  
 geom\_point(alpha=0.5) + labs(title="Regression Tree", x="Real GDP", y="Predicted GDP") +   
 theme(plot.title=element\_text(size=10), axis.title.x=element\_text(size=7), axis.title.y=element\_text(size=7), axis.text.x=element\_text(size=5), axis.text.y=element\_text(size=5)) + theme(legend.position="none")  
  
g\_rf <- ggplot(f, aes(x=GDP, y=p\_rf)) +   
 geom\_abline(intercept=0, slope=1) +  
 geom\_point(alpha=0.5) + labs(title="Random Forest", x="Real GDP", y="Predicted GDP") +   
 theme(plot.title=element\_text(size=10), axis.title.x=element\_text(size=7), axis.title.y=element\_text(size=7), axis.text.x=element\_text(size=5), axis.text.y=element\_text(size=5)) + theme(legend.position="none")  
  
gridExtra::grid.arrange(g\_sr,g\_mr,g\_pr,g\_dt,g\_rf,ncol=2)



# Calc prediction error and visualize it

sr <- mean ((f$GDP - f$p\_sr) ^ 2)   
mr <- mean ((f$GDP - f$p\_mr) ^ 2)   
pr <- mean ((f$GDP - f$p\_pr) ^ 2)   
dt <- mean ((f$GDP - f$p\_dt) ^ 2)   
rf <- mean ((f$GDP - f$p\_rf) ^ 2)   
mse <- data.frame(sr,mr,pr,dt,rf)   
head(mse)

## sr mr pr dt rf  
## 1 2.26201e+23 1.807404e+23 2.869689e+23 2.201086e+23 1.965748e+23

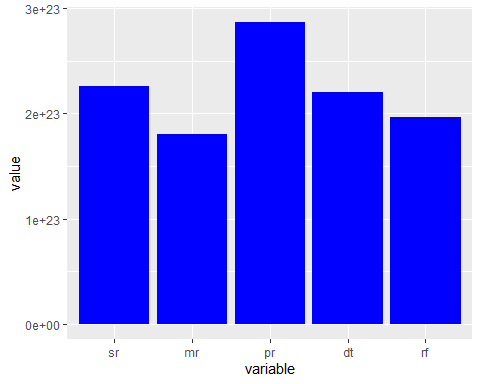
mse1 <- reshape::melt.data.frame(mse)

## Using as id variables

head(mse1)

## variable value  
## 1 sr 2.262010e+23  
## 2 mr 1.807404e+23  
## 3 pr 2.869689e+23  
## 4 dt 2.201086e+23  
## 5 rf 1.965748e+23

b1 <- ggplot(mse1, aes(x=variable, y=value)) + geom\_bar(stat="summary", fun=mean, fill = 'blue')  
b1



# Save results

ggsave("plot.jpg", plot=b1 + theme\_classic(), width=20, height=15, units="cm", dpi=600)