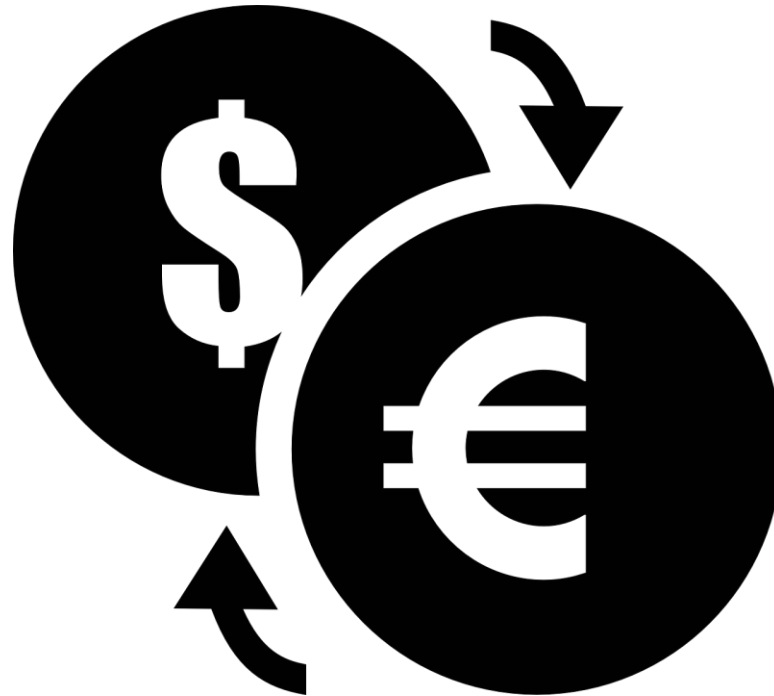


المدرسة الوطنية العليا للذكاء الاصطناعي وعلوم المعطيات - تارودانت
ÉCOLE NATIONALE SUPÉRIEURE DE L'INTELLIGENCE ARTIFICIELLE
ET SCIENCES DES DONNÉES - TAROUDANT



Crée par :

Hicham MIMOUNI

Supervisée par :

M.Mohammed AZIZI

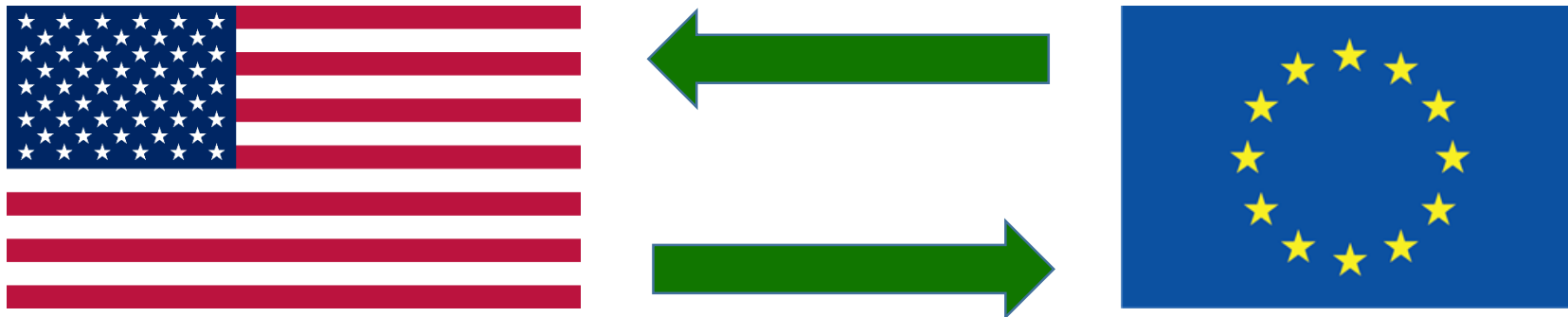
taux de change prédiction

Plan

- Introduction
- Source des Données
- Analyse Descriptive des Données
- Régression Linéaire
- Conclusion

Introduction

Le taux de change entre l'euro et le dollar est un indicateur clé dans l'économie mondiale, reflétant la relation entre deux des plus grandes économies du monde. Les fluctuations de ce taux peuvent avoir des impacts significatifs sur le commerce international, les investissements et même l'économie globale.



Data source



Le fichier Excel **data.xsl** a été fourni par **Hana Jamali**, doctorante à l'Université Cadi Ayyad. Hana a réalisé un projet de fin d'études intitulé :

“Hybrid Forex prediction model using multiple regression, simulated annealing, reinforcement learning and technical analysis”

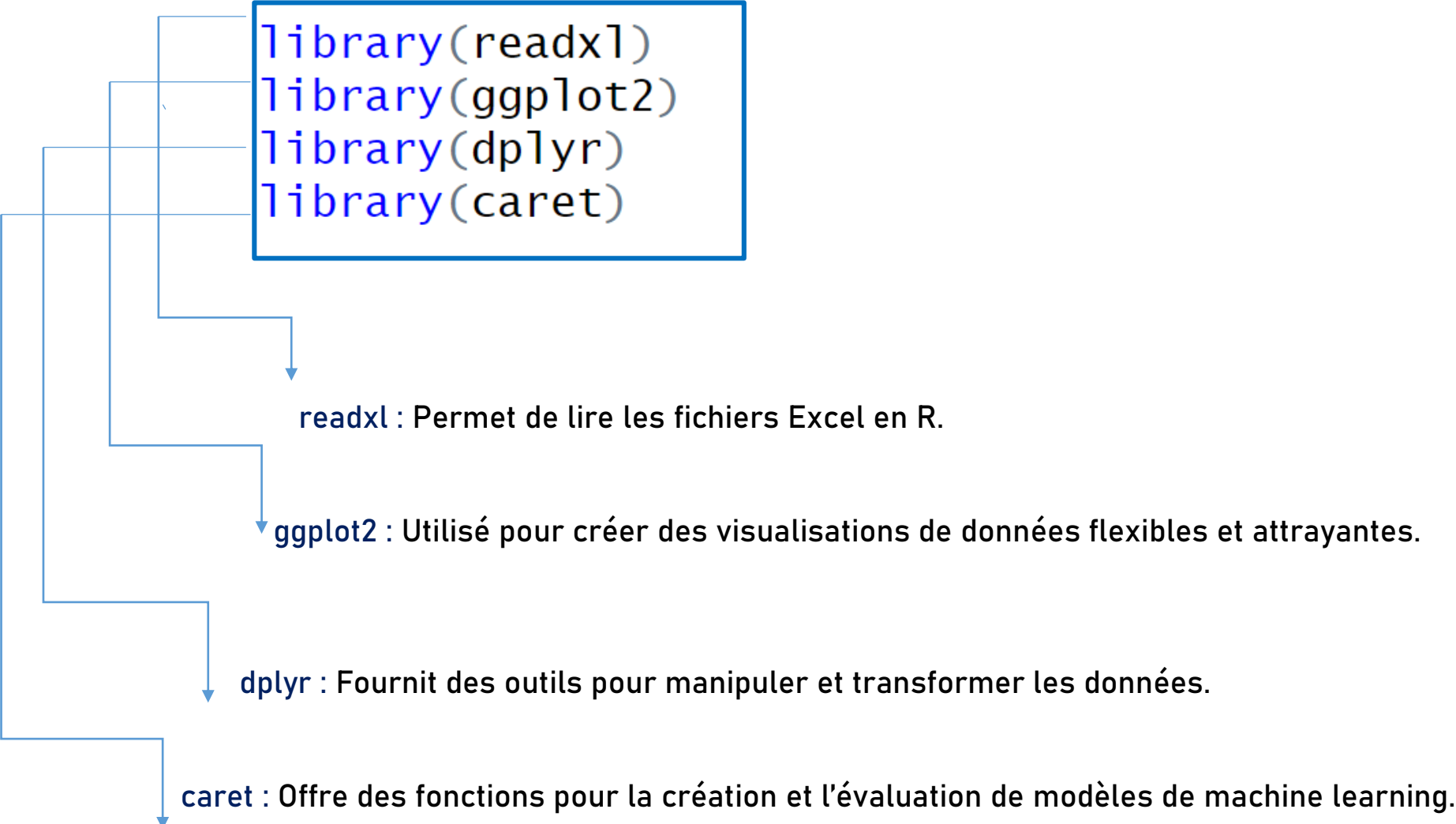
Notre Data

Open	High	Low	J (Close)	J.1	J.2	J.3	J.4	J.5	J.6	Moy2W	MoyMonth
1.0886	1.09464	1.08744	1.09095	1.08861	1.08729	1.088	1.08818	1.09293	1.09744	1.09552	1.10908
1.09095	1.09422	1.08826	1.0897	1.09095	1.08861	1.08729	1.088	1.08818	1.09293	1.09437	1.10804
1.08971	1.09917	1.0895	1.09824	1.0897	1.09095	1.08861	1.08729	1.088	1.08818	1.09417	1.10724
1.09896	1.09912	1.09756	1.0982	1.09824	1.0897	1.09095	1.08861	1.08729	1.088	1.09370	1.10639

1230	9/29/2020	1.167680	1.174570	1.166520	1.174140	1.167680	1.163410	1.162770	1.167350	1.165680	1.170590	1.175166	1.180395
1231	9/30/2020	1.174140	1.175480	1.168470	1.172320	1.174140	1.167680	1.163410	1.162770	1.167350	1.165680	1.174138	1.180005
1232	10/1/2020	1.172340	1.176960	1.171710	1.174390	1.172320	1.174140	1.167680	1.163410	1.162770	1.167350	1.173438	1.179750
1233	10/2/2020	1.174390	1.174450	1.169560	1.171570	1.174390	1.172320	1.174140	1.167680	1.163410	1.162770	1.172804	1.179130
1234	10/4/2020	1.171100	1.172010	1.170530	1.171950	1.171570	1.174390	1.172320	1.174140	1.167680	1.163410	1.171894	1.178471
1235	10/5/2020	1.171950	1.179750	1.171760	1.178910	1.171950	1.171570	1.174390	1.172320	1.174140	1.167680	1.171541	1.177979
1236	10/6/2020	1.178890	1.180770	1.173020	1.173710	1.178910	1.171950	1.171570	1.174390	1.172320	1.174140	1.170766	1.177361
1237	10/7/2020	1.173700	1.178180	1.172490	1.176570	1.173710	1.178910	1.171950	1.171570	1.174390	1.172320	1.170789	1.177094
1238	10/8/2020	1.176570	1.178160	1.173260	1.176680	1.176570	1.173710	1.178910	1.171950	1.171570	1.174390	1.171224	1.176846

Importation des bibliothèques

```
library(readxl)  
library(ggplot2)  
library(dplyr)  
library(caret)
```



readxl : Permet de lire les fichiers Excel en R.

ggplot2 : Utilisé pour créer des visualisations de données flexibles et attrayantes.

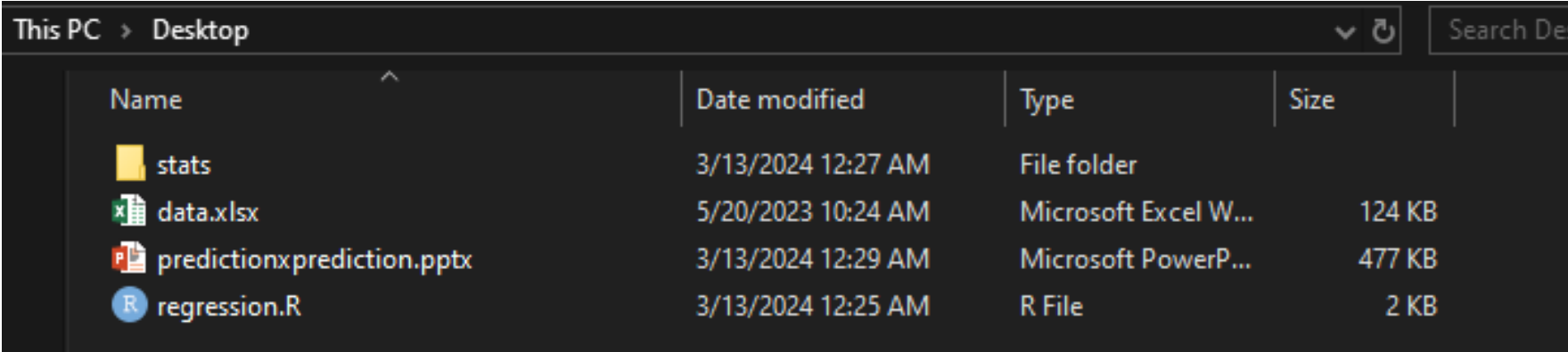
dplyr : Fournit des outils pour manipuler et transformer les données.

caret : Offre des fonctions pour la création et l'évaluation de modèles de machine learning.

Importation des données

```
data <- read_excel('C:/Users/LENOVO/Desktop/data.xlsx')
```

Cette ligne de code utilise la fonction `read_excel` de la bibliothèque `readxl` pour lire le fichier Excel nommé `data.xlsx` et le stocker dans une variable appelée `data`.



Name	Date modified	Type	Size
stats	3/13/2024 12:27 AM	File folder	
data.xlsx	5/20/2023 10:24 AM	Microsoft Excel W...	124 KB
predictionxprediction.pptx	3/13/2024 12:29 AM	Microsoft PowerP...	477 KB
regression.R	3/13/2024 12:25 AM	R File	2 KB

Analyse descriptive des données

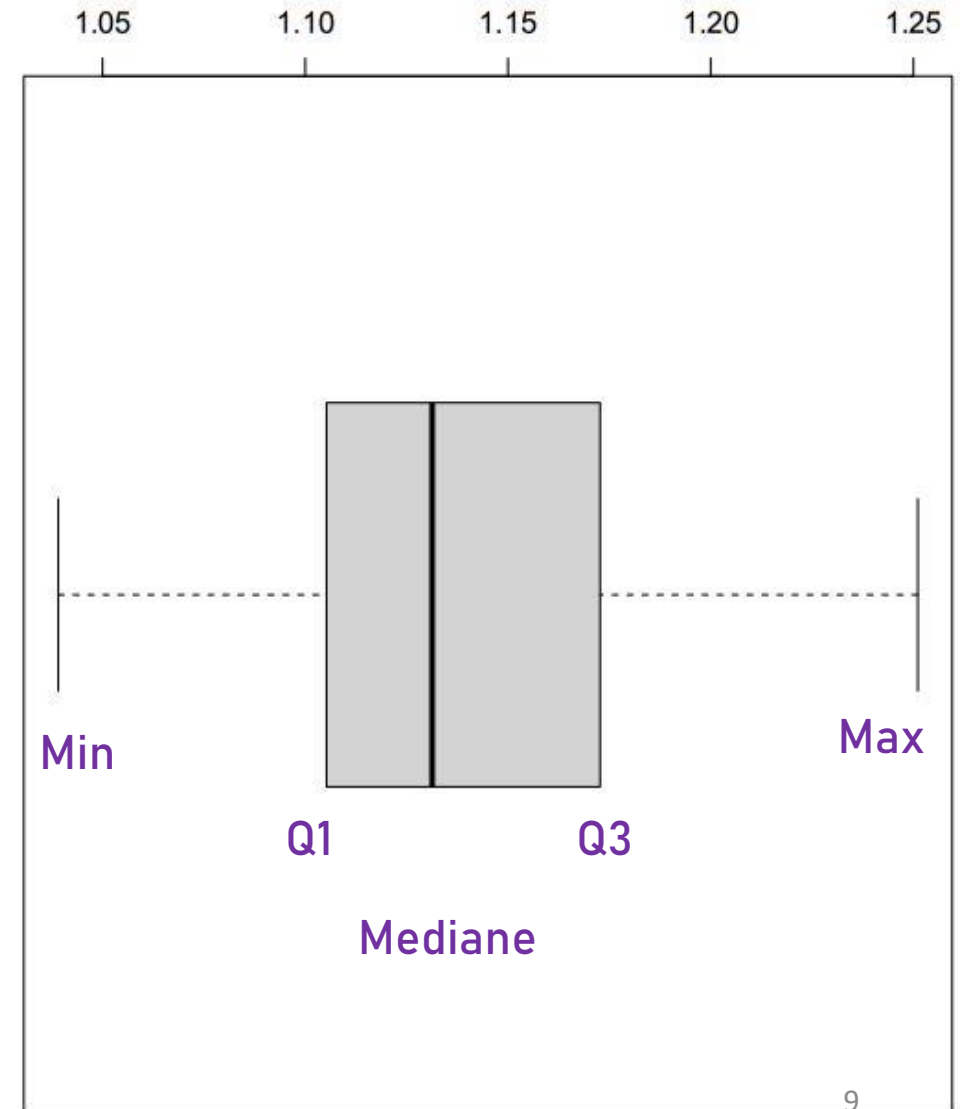
```
> summary(data)
```

Time	Open	High	Low	J (Close)	J.1	J.2
Min. :2016-10-26 00:00:00.00	Min. :1.039	Min. :1.042	Min. :1.034	Min. :1.039	Min. :1.039	Min. :1.039
1st Qu.:2017-10-22 00:00:00.00	1st Qu.:1.105	1st Qu.:1.108	1st Qu.:1.102	1st Qu.:1.105	1st Qu.:1.105	1st Qu.:1.105
Median :2018-10-18 00:00:00.00	Median :1.131	Median :1.134	Median :1.128	Median :1.131	Median :1.131	Median :1.131
Mean :2018-10-17 21:37:58.74	Mean :1.137	Mean :1.140	Mean :1.133	Mean :1.137	Mean :1.137	Mean :1.136
3rd Qu.:2019-10-14 00:00:00.00	3rd Qu.:1.173	3rd Qu.:1.176	3rd Qu.:1.168	3rd Qu.:1.173	3rd Qu.:1.173	3rd Qu.:1.173
Max. :2020-10-08 00:00:00.00	Max. :1.251	Max. :1.256	Max. :1.245	Max. :1.251	Max. :1.251	Max. :1.251
J.3	J.4	J.5	J.6	Moy2W	MoyMonth	
Min. :1.039	Min. :1.039	Min. :1.039	Min. :1.039	Min. :1.045	Min. :1.050	
1st Qu.:1.105	1st Qu.:1.105	1st Qu.:1.105	1st Qu.:1.105	1st Qu.:1.104	1st Qu.:1.104	
Median :1.131	Median :1.131	Median :1.131	Median :1.131	Median :1.130	Median :1.129	
Mean :1.136	Mean :1.136	Mean :1.136	Mean :1.136	Mean :1.136	Mean :1.136	
3rd Qu.:1.172	3rd Qu.:1.172	3rd Qu.:1.172	3rd Qu.:1.172	3rd Qu.:1.170	3rd Qu.:1.169	
Max. :1.251	Max. :1.251	Max. :1.251	Max. :1.251	Max. :1.240	Max. :1.237	

Diagramme en boîte

Open	
Min.	:1.039
1st Qu.:	1.105
Median	:1.131
Mean	:1.137
3rd Qu.:	1.173
Max.	:1.251

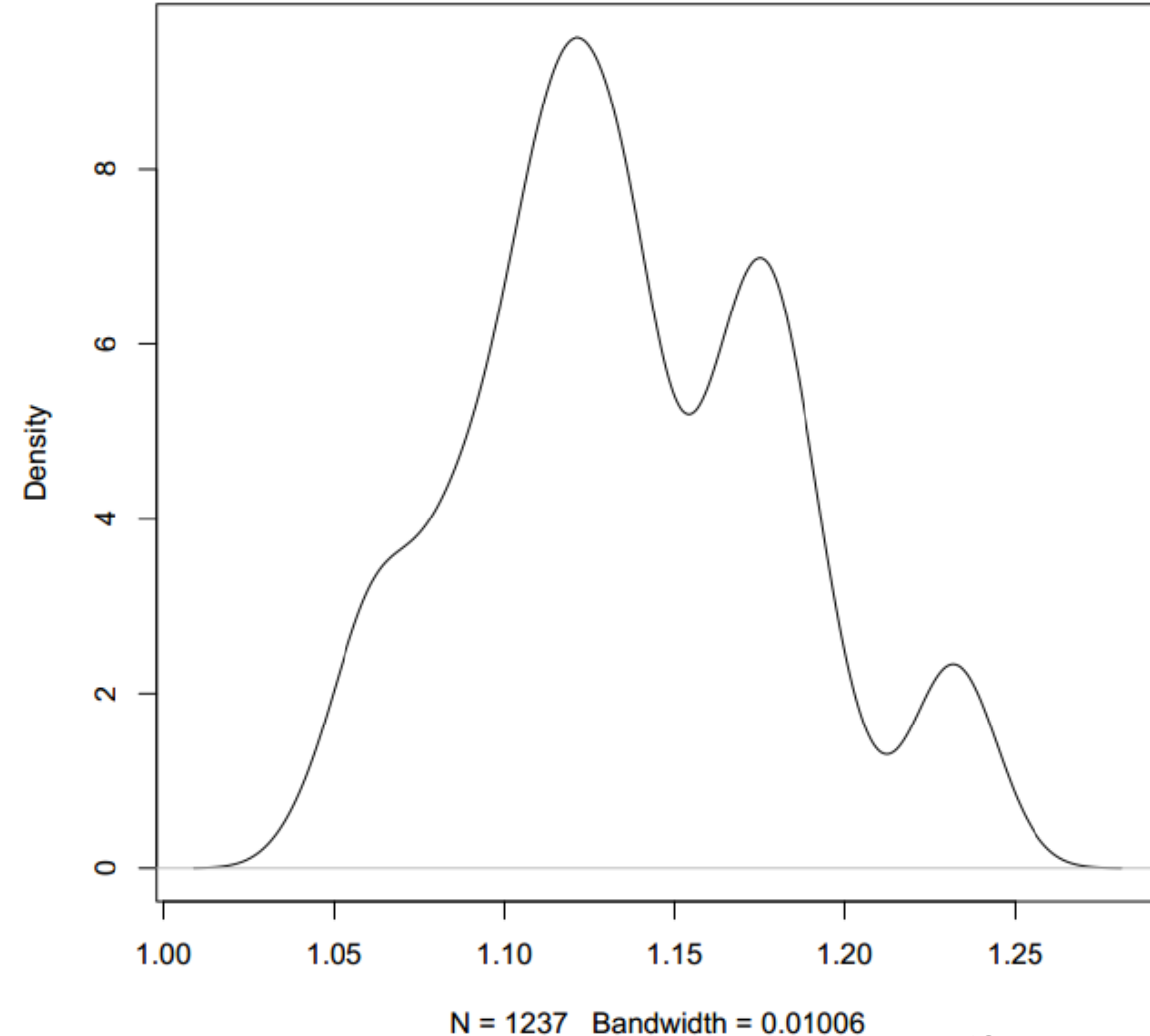
```
boxplot(data$Open)
```



graphe de densité

Il utilise une courbe lisse pour montrer où les valeurs se situent dans l'ensemble de données, combien de données sont regroupées à certaines valeurs et si il y a des pics ou des vallées dans les données

```
plot(density(data$Open))
```



Régression Linéaire

La régression multiple est une technique statistique pour prédire une variable dépendante continue à partir de plusieurs variables indépendantes. Elle repose sur l'hypothèse d'une relation linéaire entre la variable dépendante et les variables indépendantes

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon$$

```
x <- data[,c('Open', 'High', 'Low', 'J.1', 'J.2', 'J.3', 'J.4', 'J.5', 'J.6', 'Moy2W', 'MoyMonth')]
y <- data$'J (Close)'
```

Open	High	Low	J (Close)	J.1	J.2	J.3	J.4	J.5	J.6	Moy2W	MoyMonth
1.0886	1.09464	1.08744	1.09095	1.08861	1.08729	1.088	1.08818	1.09293	1.09744	1.09552	1.10908
1.09095	1.09422	1.08826	1.0897	1.09095	1.08861	1.08729	1.088	1.08818	1.09293	1.09437	1.10804
1.08971	1.09917	1.0895	1.09824	1.0897	1.09095	1.08861	1.08729	1.088	1.08818	1.09417	1.10724
1.09896	1.09912	1.09756	1.0982	1.09824	1.0897	1.09095	1.08861	1.08729	1.088	1.09370	1.10639

- Y représente la variable cible J(close) .
- X représente les variables indépendantes utilisées pour prédire la variable dépendante Y .

Principe de la méthode :

Le Modèle de Régression Linéaire essaie de trouver une relation linéaire entre les variables indépendantes **X** et la variable dépendante **Y** (la variable cible), Afin de pouvoir prédire les valeurs de **Y** à partir des nouvelles valeurs entrées de **X** .

Modèle de Régression Linéaire

Ensembles d'entraînement et de test

```
set.seed(42)
trainIndex <- createDataPartition(y, p = .8, list = FALSE)
X_train <- X[trainIndex,]
y_train <- y[trainIndex]
X_test <- X[-trainIndex,]
y_test <- y[-trainIndex]
```

Creation du modèle

```
model <- lm(y_train ~ ., data = data.frame(X_train, y_train))
y_pred <- predict(model, newdata = X_test)
```

Affichage des coefficients du modèle :

```
coefficients <- summary(model)$coefficients  
print(coefficients)
```

Résultat :

	Variable	Coefficient
0	Intercept	0.000902
1	Open	-0.587725
2	High	0.733139
3	Low	0.857363
4	J.1	-0.030418
5	J.2	0.040958
6	J.3	-0.028453
7	J.4	-0.013037
8	J.5	-0.007774
9	J.6	-0.008958
10	Moy2W	0.060289
11	MoyMonth	-0.015822

→ On stocke Les coefficients du modèle, pour les variables indépendantes et pour l'intercept dans un DataFrame appelé coefficients .

→ On affiche les coefficients du modèle.

Pour voir leur importance et leur participation dans la prédiction de la variable cible Y.

le coefficient R^2

```
r2 <- summary(model)$r.squared  
print(paste("R²:", round(r2, 4)))
```

On calcule le coefficient de détermination R^2 pour évaluer la performance du modèle .

Affichage de la valeur de R^2 .

```
"R²: 0.9983"
```

Comparaison des valeurs réelles et prédites



Time	Open	High	Low	J (Close)	J.1	J.2	J.3	J.4	J.5	J.6	Moy2W	MoyMonth
10/8/2020	1.17657	1.17816	1.17326	1.17668	1.17657	1.17371	1.17891	1.17195	1.17157	1.17439	1.17122	1.17685

```
variables_1 <- data.frame('Open' = 1.1737, 'High' = 1.17818, 'Low' = 1.17249,
                          'J.1' = 1.17657, 'J.2' = 1.17371, 'J.3' = 1.17891, 'J.4' = 1.17195, 'J.5' = 1.17157, 'J.6' = 1.17439,
                          'Moy2W' = 1.17122, 'MoyMonth' = 1.17685)
prediction1 <- predict(model, newdata = variables_1)
print(paste("Le Resultat de Prédiction est :", prediction1))
```

Résultat :

[1.17607]

Pour la journée 11/8/2020 notre modèle prédite
que le taux de change sera : **1.17607**

Conclusion

Grâce à notre modèle de prédiction, nous sommes en mesure d'anticiper les mouvements futurs de ce taux de change. Cela peut aider les investisseurs et les entreprises à prendre des décisions éclairées, que ce soit pour acheter ou vendre des devises.

