Uma imagem com texto

Descrição gerada automaticamente

**Previsão do Preço diário do Ouro**

**Docente:**

Profª. Doutora Isabel Pereira

**Discentes:**

Ana Rita Cheganças Nº 104633

Nuno Pedrosa Nº 94471

Índice

[Introdução 3](#_Toc107098963)

[Dados e análise exploratória 3](#_Toc107098964)

[*Dataset* 3](#_Toc107098965)

[Exploração dos dados 3](#_Toc107098966)

[Modelos ARIMA/SARIMA 8](#_Toc107098967)

[Previsões 8](#_Toc107098968)

[Discussão de Resultados/Conclusões 8](#_Toc107098969)

[Bibliografia 9](#_Toc107098970)

Índice de Figuras

[Figura 1 - QQPlot para Preço dário do Ouro 4](#_Toc107098947)

[Figura 2 - Histograma para Preço diário do Ouro 4](#_Toc107098948)

[Figura 3 - Gráfico Preço diário do Ouro 5](#_Toc107098949)

[Figura 4 - Comparação dados originais com 1ª's diferenças, log, e 1ª's diferenças do log 5](#_Toc107098950)

[Figura 5 - Modelo aditivo para os dados log 6](#_Toc107098951)

[Figura 6 - Modelo multiplicativo para os dados log 7](#_Toc107098952)

[Figura 7 - Modelo aditivo para os dados 1ª's diferenças de log 7](#_Toc107098953)

[Figura 8 - Modelo multiplicativo para os dados 1ª's diferenças de log 7](#_Toc107098954)

# Introdução

O presente relatório foi realizado no âmbito da Unidade Curricular de Séries Temporais no decurso do Mestrado de Ciência de Dados. Este trabalho tem como principal objetivo aplicação dos modelos SARIMA, *Bootstrap*, *forecasting* e *Exponential Smoothing* lecionados em aula. O tema adotado pelo nosso grupo tem por base a previsão dos preços diários do ouro. Este relatório encontra-se dividido em quatro partes distintas. Numa primeira parte do relatório, encontra-se a explicação do *dataset* utilizado e respetiva análise exploratória realizada ao mesmo. Numa segunda parte, apresenta-se a demonstração e explicação dos modelos aplicados. De seguida, apresenta-se as previsões realizadas. Por fim, na quarta parte, poderá encontrar-se a discussão dos resultados obtidos e respetivas conclusões.

O tema escolhido para este trabalho foi o preço do ouro. Este é um material escasso, que não consegue ser gerado pelo ser humano. Assim, o seu valor é completamente alicerçado em torno da sua durabilidade e qualidade, que se mantém com o passar do tempo, desde a antiguidade. Devido a isto, em termos históricos, o ouro foi muito utilizado para moeda de troca. Posteriormente, quando o dinheiro em papel foi introduzido, o ouro passou a ser considerado principalmente como uma reserva de valor. Algo muito importante, pois permite controlar o poder econômico diante de alterações sistêmicas sobre a economia, tanto a nível individual, como a nível dos cofres de governos e de instituições bancárias. Logo, é muito importante ter uma previsão do preço do ouro no futuro para que as reservas deste sejam geridas adequadamente.

# Dados e análise exploratória

## *Dataset*

A escolha do *dataset* para realização deste trabalho, recaiu sobre o preço diário do ouro. Este dataset foi retirado apartir do website [Kaggle](https://www.kaggle.com/datasets/psycon/daily-gold-price-historical-data). Para além disso, tem na sua constituição uma coluna referente à data, outra com o preço de abertura, uma com o preço de fecho, com o preço mais alto, preço mais Elevado, e por fim uma com a moeda de transição. Para o objetivo do nosso trabalho, foram selecionadas apenas duas colunas para serem trabalhadas, nomeadamente a dos preços de fecho e a da data correspondente.

## Exploração dos dados

De forma a percebermos melhor os dados, foi realizada uma análise exploratória aos mesmos. Com isto, conseguimos perceber que os dados apresentavam um preço médio de fecho e uma variância de 1044.9 e 269110.4, respetivamente. O preço mais baixo que o ouro atingiu foi cerca de 255.1, e o valor máximo atingido foi de 2054.6.

A análise do *QQplot* e do Histograma, permitem aferir a normalidade das variáveis a nível de representação gráfica, ou seja, permitem comparar a distribuição dos nossos dados com uma distribuição normal. Desta forma, conseguimos concluir que os nossos dados não contêm uma distribuição normal.

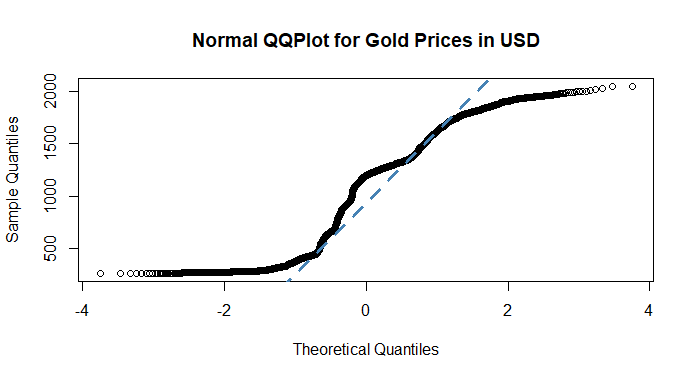


Figura - QQPlot para Preço dário do Ouro

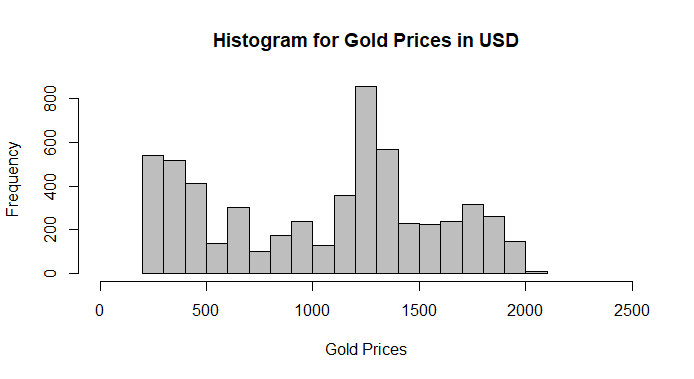


Figura - Histograma para Preço diário do Ouro

Recorremos à visualização gráfica do conjunto completo de dados, de forma a termos uma ideia do tipo de série temporal presente. Assim sendo, a figura 3 mostra-nos que podemos estar perante uma *random walk*.

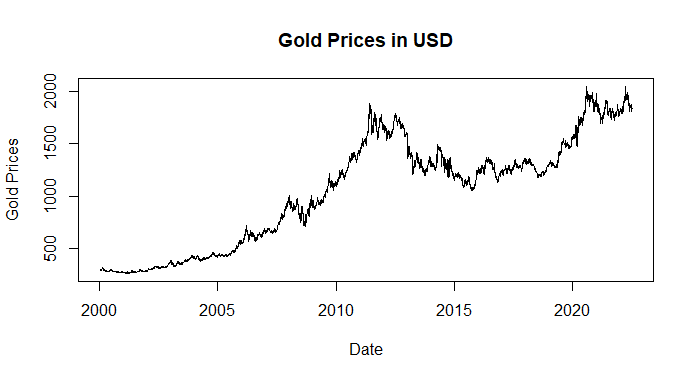


Figura - Gráfico Preço diário do Ouro

Realizámos o cálculo das primeiras diferenças, do logaritmo, e das primeiras diferenças do logaritmo, de forma a melhorar a disposição dos dados. A figura 4 apresenta os resultados obtidos, comparando assim com os dados originais.

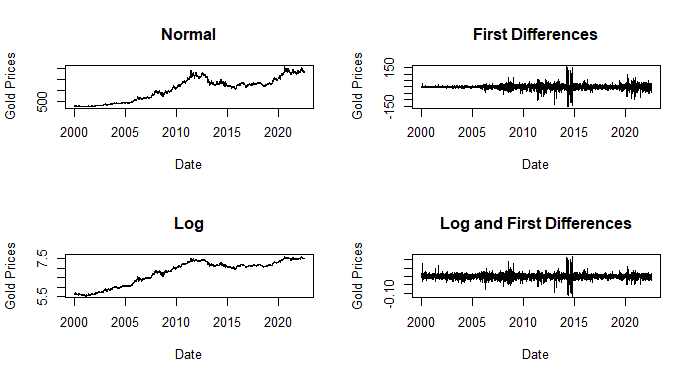


Figura - Comparação dados originais com 1ª's diferenças, log, e 1ª's diferenças do log

Tendo em conta os resultados alcançados, começamos a utilizar o *dataset* com os valores calculados através do logaritmo e o *dataset* das 1ª’s diferenças do logaritmo.

Posteriormente, fomos verificar a estacionariedade de cada *dataset*. Através do teste adf, conseguimos concluir que o *dataset* com os dados do logaritmo, apresentam um *p-value* superior a 5%, o que significa que não se rejeita H0, pelo que não apresentam estacionariedade, o que é um índicio da existência de uma R*andom Walk*. Por outro lado, o *dataset* com os dados das 1ª’s diferenças do logaritmo, apresentam um *p-value* inferior a 5%, o que significa que se rejeita H0, que tem estacionariedade, e que é índicio da existência de uma *White Noise*.

Por último, realizamos uma análise à sazonalidade de cada *dataset*. Existem dois modelos subjacentes a esta, nomeadamente, o modelo aditivo e o multiplicativo. Ambos foram testados para o *dataset* com os dados do logaritmo, e para o *dataset* com os dados das 1ª’s diferenças do logaritmo. Apesar de ambos os modelos terem sido testados, o que melhor se aplica no nosso caso, seria o multiplicativo. Assim sendo, a figura 6, demonstra o modelo multiplicativo para os dados do logaritmo, e conclui-se que não apresentam sazonalidade. Já a figura 8, demonstra o modelo multiplicativo para os dados das 1ª’s diferenças do logaritmo, e conclui-se que não apresentam sazonalidade.

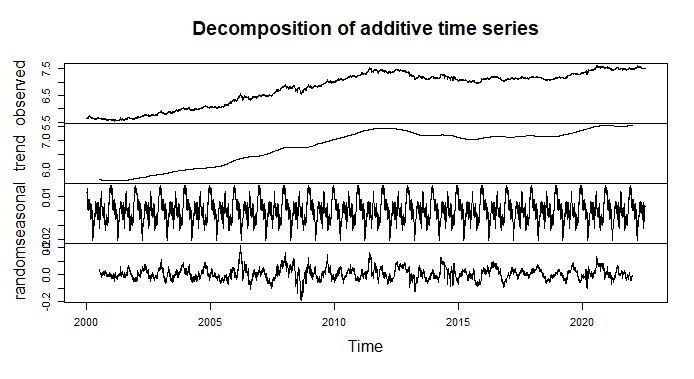


Figura - Modelo aditivo para os dados log

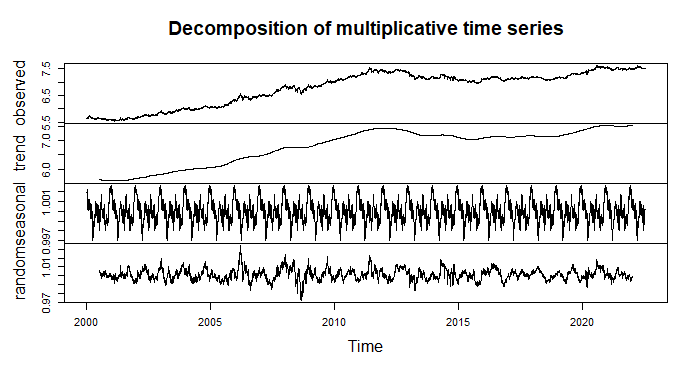


Figura - Modelo multiplicativo para os dados log

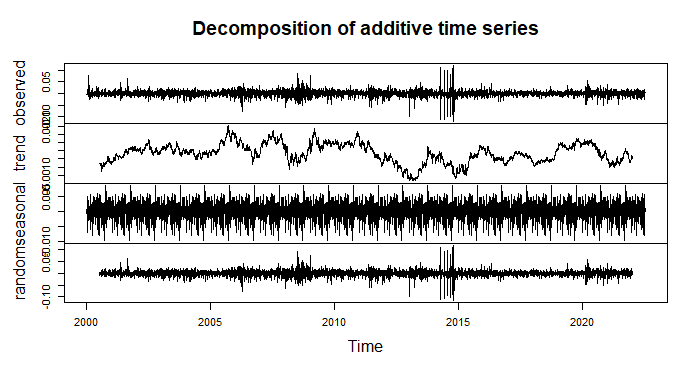


Figura - Modelo aditivo para os dados 1ª's diferenças de log

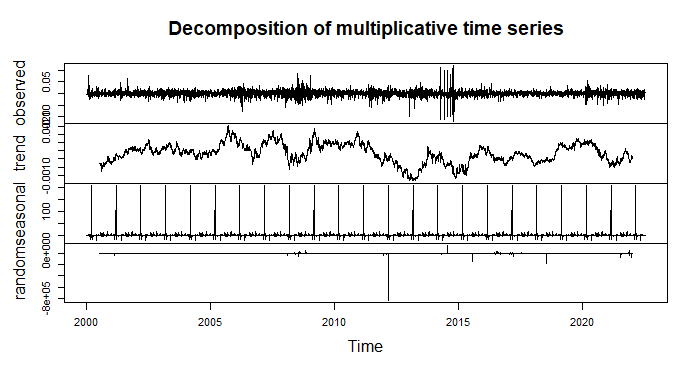


Figura - Modelo multiplicativo para os dados 1ª's diferenças de log

# Modelos ARIMA/SARIMA

# Previsões

# Discussão de Resultados/Conclusões

# Bibliografia

Fazer bibliografia automática do word. Colocar todos as referencias para aqui e depois fazer isso no final

Colocar slides da prof