

Lista 6 - Introdução a Análise de Dados

Raspagem de dados

Guilherme Masuko

May 2023

Vamos estudar como três variáveis macroeconômicas brasileiras se relacionam.

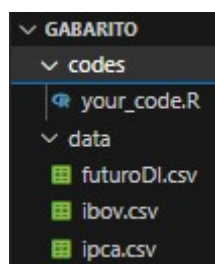
- Inflação
- Taxa de Juros
- Ibovespa

Para cada um desses indicadores, utilizaremos uma base de dados na forma de série temporal.

Questão 1

Faça a importação dos dados. Certifique-se de manter seu diretório organizado. Exemplo:

Figure 1: Diretório de Exemplo



- a) Faça a importação do arquivo 'futuroDI.csv' em um dataframe. Utilizaremos o futuro DI¹ como proxy da taxa de juros brasileira (Selic). Essa base de dados é composta de 22 colunas contendo dados de Junho de 2012 à Outubro de 2022. Abaixo descrevemos os dados.

¹<<https://www.bradesco corretora.com.br/SiteBradescoCorretora/Produtos/Mercados-Futuros/Produtos-Oferecidos/DI-Futuro>>

- X: Data

- BRPRE___.BMF: é a taxa do futuro DI onde os __ são preenchidos pelo horizonte à frente referente à expectativa de taxa de juros, ex: 1M representa o futuro DI um mês à frente. Fonte: Reuters.

b) Faça a importação do arquivo 'ipca.csv' em um dataframe. Essa base de dados é composta de 2 colunas contendo dados de Junho de 2012 à Abril de 2023. Abaixo descrevemos os dados.

- X: Data

- valor: Variação da inflação do mês anterior em relação ao mês atual.

c) Faça a importação do arquivo 'ibov.csv' em um dataframe. Essa base de dados é composta de 2 colunas contendo dados de Junho de 2012 à Abril de 2023. Abaixo descrevemos os dados.

- X: Data

- IBOV: Pontos da bolsa brasileira Ibovespa.

Questão 2

Para cada um dos dataframes, transforme a coluna X em índice (nome das linhas).

Questão 3

Para cada dataframe, faça as alterações abaixo:

a) Para o dataframe `futuroDI`, mantenha somente a coluna referente ao futuro DI para um mês.

b) Para o dataframe `ipca`, mantenha somente a `valor`.

c) Para o dataframe `ibov`, mantenha somente a `IBOV`.

Questão 4

Renomeie as colunas remanescentes para os dataframes `futuroDI`, `ipca` e `ibov`, de `BRPRE1M.BMF`, `valor` e `IBOV`, para `Futuro_DI`, `IPCA` e `Ibovespa`.

Questão 5

Precisamos fazer manipulações no dataframe `ipca` para que cada linha tome o valor acumulado da inflação dos últimos 12 meses (assim como o Banco Central mede em <<https://www.bcb.gov.br/>>). Obtemos essa medida calculando a seguinte fórmula.

$$\begin{aligned}\pi_{12t} &= \prod_{j=0}^{11} (1 + \pi_{t-j}) - 1 \\ &= (1 + \pi_t) \cdot (1 + \pi_{t-1}) \cdot \dots \cdot (1 + \pi_{t-11}) - 1\end{aligned}$$

onde π é a inflação.

Para isso, siga os passos:

- Crie colunas com os valores defasados. Seu dataframe deve ficar da seguinte maneira.

Figure 2: Dataframe do IPCA

	IPCA	IPCA_1	IPCA_2	IPCA_3	IPCA_4	IPCA_5	IPCA_6	IPCA_7	IPCA_8	IPCA_9	IPCA_10	IPCA_11
2012-06-30	0.0008	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2012-07-31	0.0043	0.0008	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2012-08-31	0.0041	0.0043	0.0008	NA	NA	NA	NA	NA	NA	NA	NA	NA
2012-09-30	0.0057	0.0041	0.0043	0.0008	NA	NA	NA	NA	NA	NA	NA	NA
2012-10-31	0.0059	0.0057	0.0041	0.0043	0.0008	NA	NA	NA	NA	NA	NA	NA
2012-11-30	0.0060	0.0059	0.0057	0.0041	0.0043	0.0008	NA	NA	NA	NA	NA	NA
2012-12-31	0.0079	0.0060	0.0059	0.0057	0.0041	0.0043	0.0008	NA	NA	NA	NA	NA
2013-01-31	0.0086	0.0079	0.0060	0.0059	0.0057	0.0041	0.0043	0.0008	NA	NA	NA	NA
2013-02-28	0.0060	0.0086	0.0079	0.0060	0.0059	0.0057	0.0041	0.0043	0.0008	NA	NA	NA
2013-03-31	0.0047	0.0060	0.0086	0.0079	0.0060	0.0059	0.0057	0.0041	0.0043	0.0008	NA	NA
2013-04-30	0.0055	0.0047	0.0060	0.0086	0.0079	0.0060	0.0059	0.0057	0.0041	0.0043	0.0008	NA
2013-05-31	0.0037	0.0055	0.0047	0.0060	0.0086	0.0079	0.0060	0.0059	0.0057	0.0041	0.0043	0.0008

- Compute π_{12t} . Chame essa coluna de IPCA_Acumulado.
- Mantenha apenas a colunas IPCA_Acumulado no dataframe ipca.

Questão 6

- Una os três dataframes: futuroDI, ipca e ibov.
- Drope todas linhas que tenham NA em alguma das colunas.

O dataframe final deve parecer como:

Figure 3: Dataframe Final

	IPCA_Acumulado	Ibovespa	Futuro_DI
2013-05-31	0.06503960	53506	0.07719
2013-06-30	0.06695514	47457	0.08104
2013-07-31	0.06270559	48234	0.08265
2013-08-31	0.06090637	50008	0.08728
2013-09-30	0.05858560	52338	0.09118
2013-10-31	0.05837513	54256	0.09366
2013-11-30	0.05774389	52482	0.09778
2013-12-31	0.05910818	51507	0.09957
2014-01-31	0.05585294	47639	0.10339
2014-02-28	0.05679754	47094	0.10567
2014-03-31	0.06153088	50415	0.10743

Questão 7

Calcule e interprete as correlações entre as variáveis:

- `IPCA_Acumulado` e `Futuro_DI`.
- `Ibovespa` e `Futuro_DI`.