

Emerging Markets indecies correlation to GDP

An exploration into four emerging markets. This paper will explore the correlation of FTSE indicies and GDP.

Even Oscar Harlert

2023-10-18

Introduction

Emerging markets can be great investment opportunity. In order to make good investment decisions and understand the risk and return for any investment you have to analyze historical data.

To measure countries development there are several metrics that can be useful. In this paper we will use keep things simple and only use one metric and that is gross domestic product (GDP onwards). We could use a lot more metrics but to limit the scope of this paper we are only using GDP. In a separate dataset we will have the population for the countries in question.

Data sets

There are two data sets acquired for this paper. The first one is downloaded from [refinitive/datastream](#). For more information regarding the data provider you can follow this [link](#). This will be refereed to as dataset one. The data in the file are time-series spanning from 19. September 2003 to 19. September 2023. The measurements are done with a monthly interval. All data points are represented in US dollar.

The second data set is downloaded from The World Bank. The data-set is actually divide into two(*might be more*) files but will be combined in RStudio as they each only contain one kind of data, GDP and population.

Project libraries

This project will make use of several libraries in RStudio. In order to use them we need to load them. To avoid any problems with replicating the code I've made a file to load all. This file also contains a comment with an install command if you are missing any of them on your local setup. To access the install command you can mark the area in the file and press "CTRL+Shift+C". This will remove the comment and allow you to press CTRL+ENTER to execute the code and install dependencies.

```
source("r_scripts/01libraries.R")
```

here() starts at D:/R_folder/Assignments/Emerging_markets

```
-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
v dplyr      1.1.3      v readr      2.1.4
v forcats    1.0.0      v stringr    1.5.0
v ggplot2    3.4.3      v tibble     3.2.1
v lubridate  1.9.3      v tidyr      1.3.0
v purrr      1.0.2
-- Conflicts ----- tidyverse_conflicts() --
x dplyr::filter() masks stats::filter()
x dplyr::lag()     masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become
```

Attaching package: 'psych'

The following objects are masked from 'package:ggplot2':

%,%, alpha

Attaching package: 'janitor'

The following objects are masked from 'package:stats':

chisq.test, fisher.test

To make the code cleaner, I've made a script folder and separated the required libraries in

that file. The command above simply executes the commands in that file. The following are the lines that has been executed.

```
library(here)
library(tidyverse)
library(readxl)
library(psych)
library(janitor)
library(readxl)
```

NB! We will add any library we add and check it at the end to assure that the project is not missing any dependencies.

Data wrangling

We start our project out by downloading data from refinitive. Since the data is actually only accessible from the terminals located on floor five in the tower building at NMBU you would have to access these in order to download the exact same data. I've provided a link to my set here (insert link to github). The data is stored as a excel file that has been stored in the project folder. To make it possible to work with in r we start by importing the file.

```
index_emerging_markets <- read_excel("datasets/FTSE 20.09.23 - EOH.xlsx", skip = 6)
```

The data in this file is currently in absolute values. In order to compare them to each other we need to normalize them

The data is downloaded in US dollars and the 20 year frame is from 19 september 2003 and 19. September 2023.

From this point we want to normalize the data so we have a fair comparison point with the base of 100. In order to do so we are using the following line:

```
normalized <- index_emerging_markets |>
  mutate(
    FTSEWorldN = FTSEWorld / (146.2 / 100),
    IndiaN = India / (332.5 / 100),
    EgyptN = Egypt / (59.04 / 100),
    ChinaN = China / (672.28 / 100),
    BrazilN = Brazil / (150.65 / 100),
  )
```

The code above adds 6 new columns where the data is normalized and starts at 100 on the 19. September 2003.

Here we now have a table of the number of the following:

- n = Number of observations
- mean = The average change
- sd = standard deviation, the average deviation from the mean
- skew = The tilt of the distribution curve
- kurtosis = Messurment of the tail of the distrubtion curve.

This data will allow for us to compare the different data sets.

Future:

- Check the $\text{Cov}()$ and $\text{Cor}()$

Sidenote

Access to the project files

The files and data for this project has been made public on GitHub and you can find it on this link. I personally always found it useful to look at other peoples work when learning and hopefully this can come to use for someone else. All links are relative and should be fully replaceable.