

## Data Analysis Assessment: Covid Restrictions

### Instructions

To answer this and the following questions, please download the attached dataset, that relates to Covid 19 restrictions. Using these data, your task is to answer a series of questions regarding the effect of these restrictions on financial markets.

Also make sure to read the accompanying data dictionary. Some of the data have been altered for the purpose of this assessment and you may not necessarily have the same data as other candidates.

The first set of questions concern the quality of the data. At one point you will be asked to identify a stock ticker with an obvious data quality issue and to remove it from further analysis. You should try to complete this part before you move on to the second and main part of the assessment, i.e. from Question 6 onwards.

You are free to use any software available to you to complete the assignment. The scoring of the test is set up to account for small variation in results that can arise from different programs, i.e. full marks will still be given as long as you are within a reasonable tolerance.

Many questions will ask you to calculate average returns for a given period of time or country. Please read the instructions below to make sure that you answer in the correct format:

- Please provide your answer in percentages and round to two-decimal precision: for example, a return of 4.3163% should be entered as "4.32".
- Use negative numbers for negative returns: for example, if you are asked to compute the market return and the market dropped by 6.152% you should enter "-6.15".
- Some of the questions ask about *annualised* returns. You can turn daily returns into annualised returns using the following formula:  $r_{annual} = (1 + r_{daily})^{252} - 1$ . For example, a daily return of 0.05% (0.0005) corresponds to an annualised return of approximately 13.42% (0.1342).
- Some of the questions ask about *cumulative* returns over several days. You can calculate the cumulative returns using the following formula:  $r_{cum} = \prod_{t=1}^T (1 + r_t) - 1$ .
- Some of the questions ask you to calculate a *simple average* of returns, while others ask for *weighted average* returns. It is often good practise to calculate weighted averages to account for the fact that bigger companies are more important to the economy. For this purpose, we included a file "wgt.csv" with weights derived from a  $\log_{10}$  of the market capitalisation. When computing a weighted average, note that the weights must sum to 1, so you may need to renormalise the weights when computing a weighted average for a subset of stocks.

## **Data overview**

The folder contains the following 6 csv files:

File	Description
companies.csv	Table with information on 500 companies (see details below).
restrictions.csv	Table with information on Covid restrictions (see details below).
vaccines.csv	Table with information on Covid vaccines (see details below).
ret.csv	Table with daily returns of companies, calculated as $\frac{price_t}{price_{t-1}} - 1$ . The first column contains the dates, and the subsequent entries in the top row show the tickers.
cap.csv	Matrix with market capitalisation of companies in USD. The first column contains the dates, and the subsequent entries in the top row show the tickers.
wgt.csv	Matrix with weights, derived from $\log_{10}(cap)$ . The first column contains the dates, and the subsequent entries in the top row show the tickers.

The companies.csv file contains the following columns:

Column Name	Description
ticker	The stock ticker of the company. For the purpose of this exercise, tickers are defined as numbers from 1 to 500 that are unique to individual companies.
company	The name of the company.
country	The country where the company is listed.
industry	The industry classification of the company.

The restrictions.csv file contains the following columns:

Column Name	Description
country	Country that introduces the restriction.
measure	The name of the restriction that was taken.
definition	A brief description of the measure.
date_start	The first date that the restriction was in place.
date_end	The last date that the restriction was in place (if applicable).

The vaccines.csv file contains (amongst others) the following columns:

Column Name	Description
name	The name of the vaccine candidate.
resultsDate	The date that the first interim Results from Phase III trials were published.