Data exercise

Bella Private Markets

Situation

You are approached by a client representing a large institutional investor in the US. They ask for advice on the best types of venture capital (VC) funds to invest in. In particular, they are curious as to which strategies and industries would generate the best returns.

Data

You have at your disposal the attached anonymised datasets. The file fund_cashflows.csv provides cashflows associated with venture capital funds, and the file fund_cashflows.csv provides various characteristics for each of the funds.

Key variable descriptions

The following are some key variables that you may (or may not) require to complete the analysis:

- FUND ID: This is the identifier of the VC fund in question.
- VINTAGE / INCEPTION YEAR: The year the fund began operations.
- STRATEGY: The strategy of the fund, corresponding to the types of companies they invest in (e.g. Early Stage, Start-up, etc.)
- CORE INDUSTRIES: The industries that the VC fund focusses its investments in.
- TRANSACTION TYPE: One of - Capital Call, Distribution, Value
- TRANSACTION DATE: The date of the transaction
- TRANSACTION AMOUNT: The value of the transaction.
- PRIMARY REGION FOCUS: Core geographic area of investments.

Task

You have roughly 2 hours to go through the following problems. You may use whatever resources you would like online, as well as any tool you would like, but you must provide:

1. Reproducible code which generates your results - you may use any language, such as R or Python.
2. A presentation of your findings

These two may be woven together in a format such as a Jupyter Notebook or an R Markdown, or provided separately. If you present your results separately, you may do so using any appropriate format, e.g.: PDF, LaTeX, Markdown, Word Document.

Problem 1: Calculate performance for each fund

You are required to calculate the performance metric, Total Value to Paid In (TVPI) **for each fund** in the dataset. The TVPI metric is calculated using three intermediate values:

1. Sum of all Distributions, D.[1]
2. The absolute value of the sum of all Capital Calls, C.[2]
3. The last (by transaction date) Value, V.[3]

To calculate these values, you will require the variables FUND ID, TRANSACTION TYPE, TRANSACTION DATE, and TRANSACTION AMOUNT.

The equation for TVPI is $TVPI = (D + V)/C$

Problem 2: What types of funds have the best performance profile?

Using the STRATEGY and CORE INDUSTRIES variables, along with the measure of performance you have just calculated, calculate statistics which would allow you to answer which strategies and industries (and potentially, combinations of the two) would provide the most attractive returns for investors.

Problem 3: Present your findings

In whatever format you would like, present your findings. This may include charts and commentary on implications of your results. If in report format, limit your answer to no more than 1-2 pages, including charts and graphics.

Submit

Submit your materials by replying to the email you received with the instructions. If you do not manage to complete each part of the exercise, submit what you can - we're looking to get a sense of your skills and thought process, and not everyone is expected to provide a full answer. If you have any questions, email rdewan@bella-pm.com.

[1] Distributions may be thought of as capital returned to the investor.
[2] Capital calls represent capital drawn from investors.
[3] Valuations represent the remaining value held in fund investments which have yet to be liquidated and distributed.