If multiple factors are used to predict returns of securities, they are given equal or different weights to form a final "score" indicating whether the price will be likely to increase or decrease. However, if certain factors are strongly correlated with each other, the above weighting method will eventually lead to repeated exposure of portfolio to a certain style of factors. The performance of the entire combination will be heavily biased towards this factor, hence weakening the effects of other factors. More specifically, when the repeated exposed factor performs well, the portfolio achieves higher excess returns; when the factor performs poorly, the portfolio exhibits larger drawdown, which means higher risk. The goal of combining multiple factors is to reduce portfolio volatility and achieve more stable return. Therefore, we need to solve this problem by factor orthogonalization.

On every cross section, we can obtain the value of all securities in the market on every factor we are interested in. We represent the number of securities on the cross section as N, the number of factors as K. $f^k = [f_1^k, f_2^k, f_3^k, \dots, f_N^k]'$ is the values of all securities on the kth factor. For fair comparison between different factor, we assume every factor is z-score normalized, that is: $\overline{f^k} = 0, \|f^k\| = 1.$ $F_{N\times K} = [f^1, f^2, f^3, \dots, f^K]$ is a matrix containing K linearly independent factor vectors. These factors are assumed to be linear independent. We hope to perform linear transformation on $F_{N\times K}$ to get a new orthogonal matrix $\tilde{F}_{N\times K} = [\tilde{f}^1, \tilde{f}^2, \tilde{f}^3, \dots, \tilde{f}^K]$, whose column vectors are orthogonal to each other (i.e., $\forall i,j,\ i \neq j, (\tilde{f}^i)'(\tilde{f}^j) = 0$).

Two methods are presented below.

1. **Symmetric Orthogonalization**

We define a transition matrix $S_{K\times K}$ that rotates $F_{N\times K}$ to $\tilde{F}_{N\times K}$:

$$\tilde{F}_{N\times K} = F_{N\times K} S_{K\times K} \tag{1}$$

To obtain $S_{K\times K}$, we first calculate the covariance matrix of $F_{N\times K}$, $\Sigma_{K\times K}$, then the overlap matrix of $F_{N\times K}$,

$$M_{K\times K} = (N-1)\Sigma_{K\times K} \tag{2}$$

Rotated matrix $\tilde{F}_{N\times K}$ is an orthogonal matrix, therefore:

$$\begin{gathered}(\tilde{F}_{N\times K})' \tilde{F}_{N\times K} = (F_{N\times K} S_{K\times K})' F_{N\times K} S_{K\times K} = \\ S'_{K\times K} F'_{N\times K} F_{N\times K} S_{K\times K} = S'_{K\times K} M_{K\times K} S_{K\times K} = I_{K\times K}\end{gathered} \tag{3}$$

Hence,

$$S_{K\times K} S'_{K\times K} = M_{K\times K}^{-1} \tag{4}$$

Any matrix $S_{K\times K}$ that meets the above requirement of (4) is an eligible transition matrix.

$$S_{K\times K} = M_{K\times K}^{-1/2} C_{K\times K} \tag{5}$$

Where $C_{K\times K}$ is an arbitrary orthogonal matrix.

Now the problem is to solve $M_{K\times K}^{-1/2}$. Since $M_{K\times K}$ is symmetric, there exits a positive definite matrix $U_{K\times K}$ satisfying:

$$U'_{K\times K} M_{K\times K} U_{K\times K} = D_{K\times K} \tag{6}$$

Where

$$D_{K\times K}=\begin{bmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_K\end{bmatrix} \tag{7}$$

$U_{K\times K}$ and $D_{K\times K}$ are the eigenvector matrix & eigenvalues diagonal matrix of $M_{K\times K}$, respectively. Moreover, $U'_{K\times K}=U^{-1}_{K\times K}$, $\forall k,\lambda_k>0$. From (6) we can obtain:

$$M_{K\times K}=U_{K\times K}D_{K\times K}U'_{K\times K} \tag{8}$$

Thus,

$$M^{-1}_{K\times K}=U_{K\times K}D^{-1}_{K\times K}U'_{K\times K} \tag{9}$$

$$M^{-1/2}_{K\times K}M^{-1/2}_{K\times K}=U_{K\times K}D^{-1/2}_{K\times K}I_{K\times K}D^{-1/2}_{K\times K}U'_{K\times K} \tag{10}$$

Since $M^{-1/2}_{K\times K}$ is symmetric, the above equation can be written as:

$$M^{-1/2}_{K\times K}M^{-1/2}_{K\times K}=U_{K\times K}D^{-1/2}_{K\times K}U'_{K\times K}U_{K\times K}D^{-1/2}_{K\times K}U'_{K\times K} \tag{11}$$

A particular solution of $M^{-1/2}_{K\times K}$ is:

$$M^{-1/2}_{K\times K}=U_{K\times K}D^{-1/2}_{K\times K}U'_{K\times K} \tag{12}$$

Where

$$D^{-1/2}_{K\times K}=\begin{bmatrix}1/\sqrt{\lambda_1} & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & 1/\sqrt{\lambda_K}\end{bmatrix} \tag{13}$$

Substitute into equation (5), the transition matrix can be obtained:

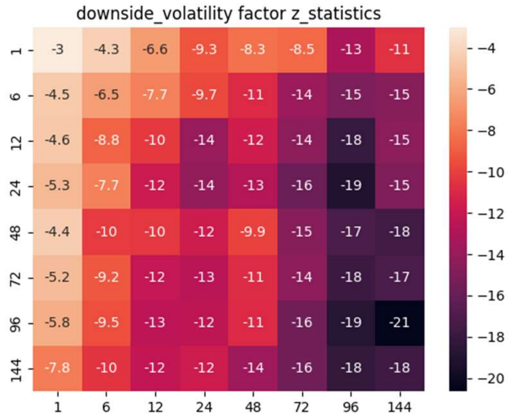
$$S_{K\times K}=U_{K\times K}D^{-1/2}_{K\times K}U'_{K\times K}C_{K\times K} \tag{14}$$

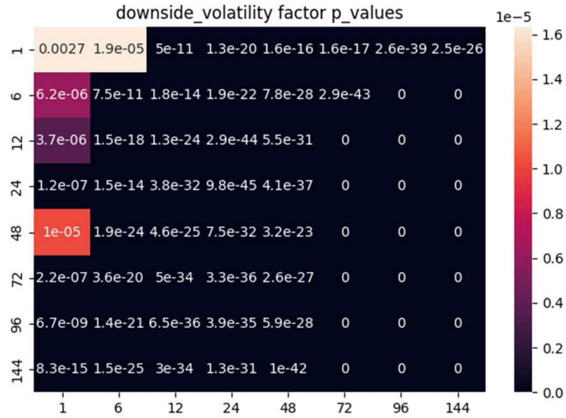
In symmetric orthogonalization, the arbitrary orthogonal matrix is chosen as: $C_{K\times K}=I_{K\times K}$.

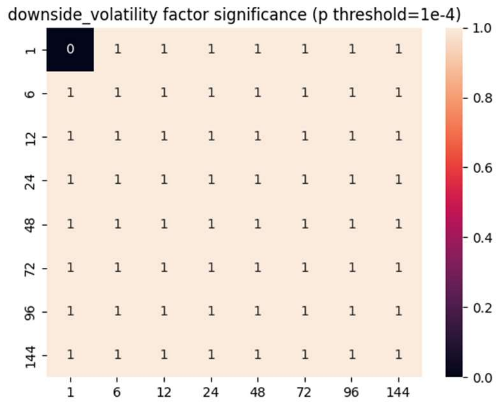
Now we are going to apply symmetric orthogonalization to the 5 factors we have taken as example in the previous report, which are respectively **downside volatility, skewness, trend strength, kurtosis, price-volume correlation**. All of them are calculated based on intraday 1-min data for maximized retention of original market information.

The orthogonalized factor performances are presented in the figures below. Note that in symmetric orthogonalization, the factors after orthogonalization have an explicit correspondence with the factors before orthogonalization.

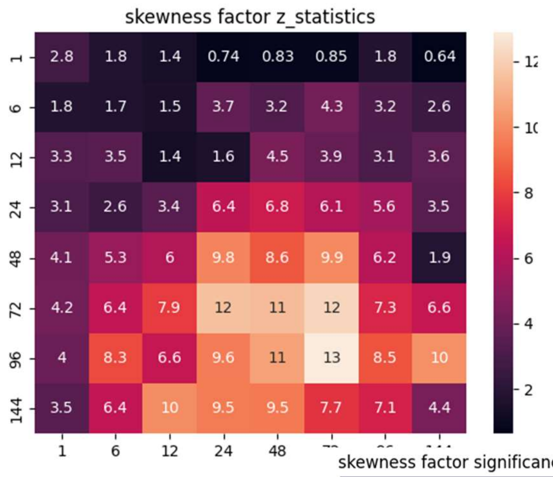
1) Downside volatility

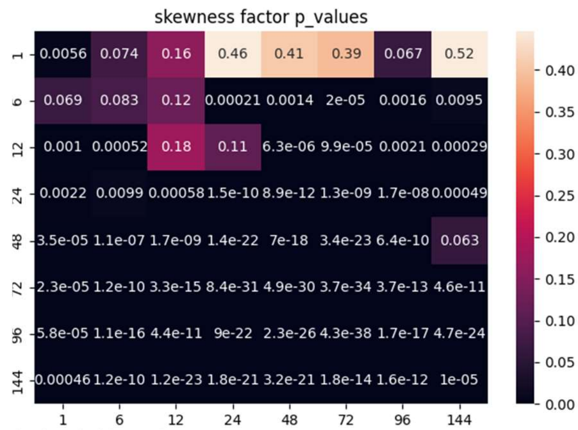


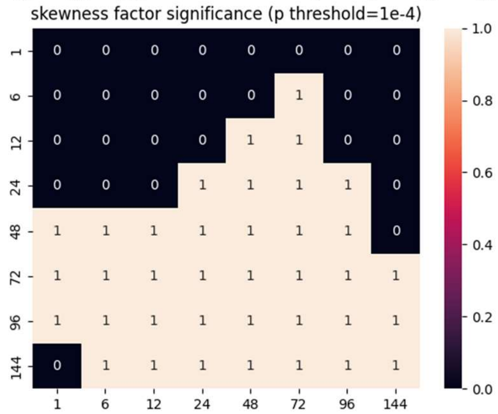




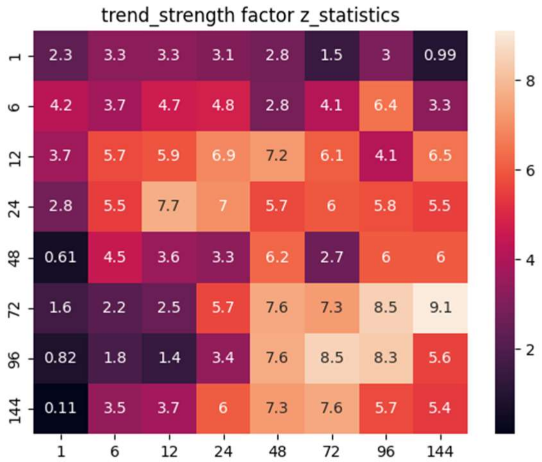
2) Skewness

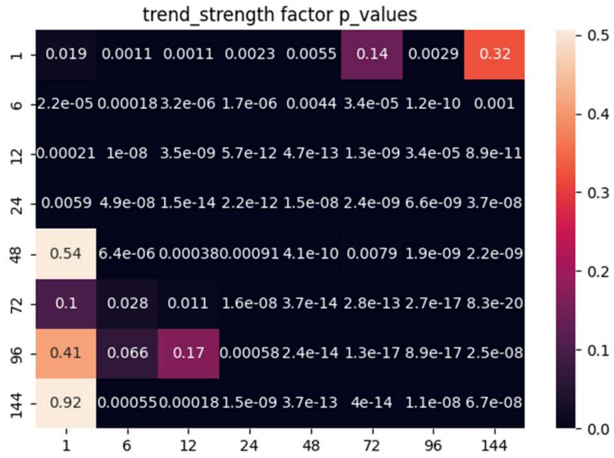


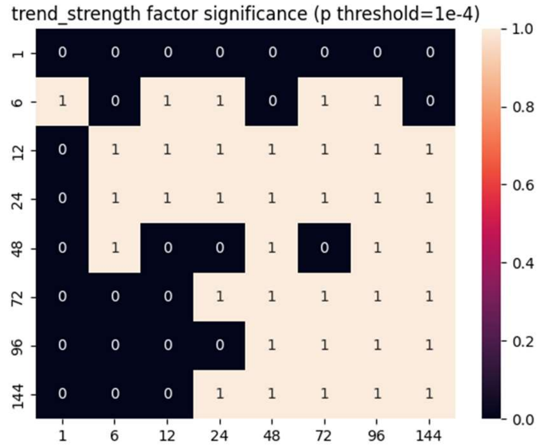




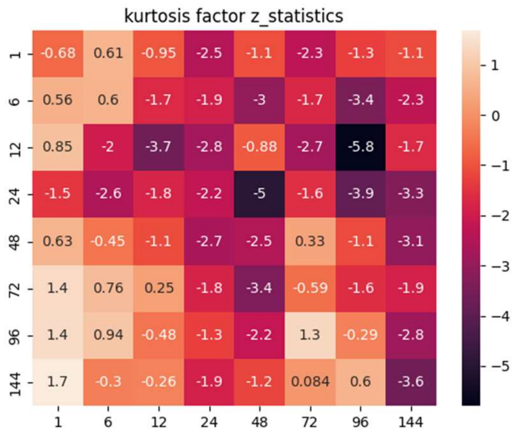
3) Trend strength

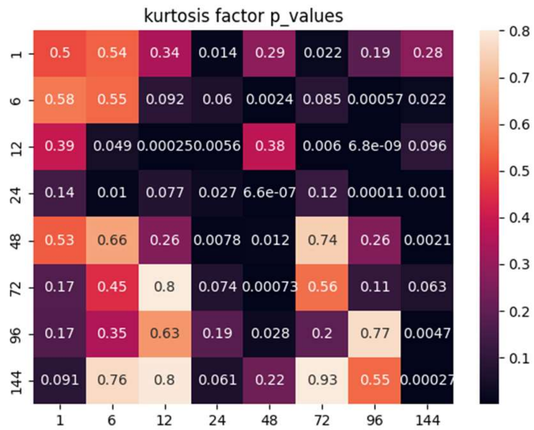


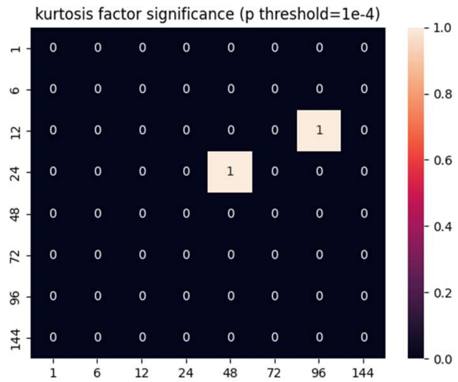




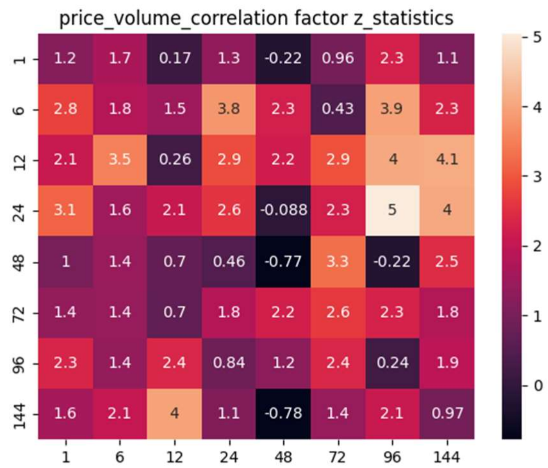
4) Kurtosis

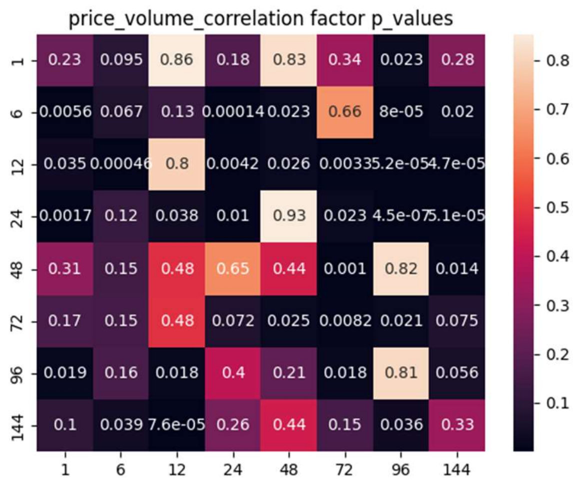


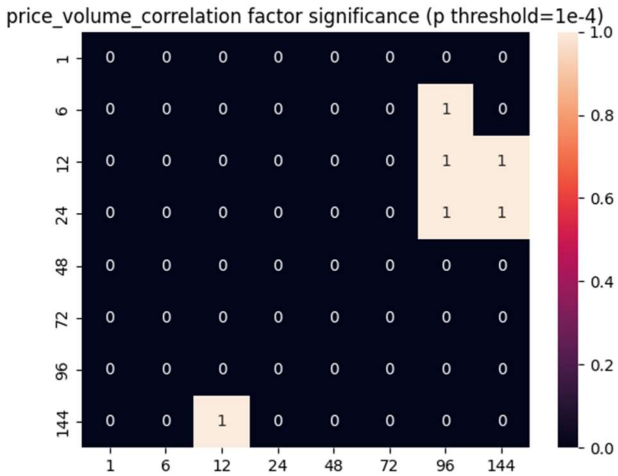




5) Price-volume correlation







We have observed that after excluding the influence from the other, **kurtosis** & **price-volume correlation** no longer contributes predictive power, which means they can be explained by some linear combination of the other 3 factors. **Trend strength**, **downside volatility** & **skewness**, which are in nature respectively 1st, 2nd, & 3rd order of intraday returns. Hence it makes sense that they are not linearly correlated. **Kurtosis,** which is the 4th order of intraday price returns, dose not provide extra information. **Price-volume correlation** is supposed to carry extra information from trade volume, but in cryptocurrency market, trade volume seems not valuable. We are going to test it on other assets (stock, commodity, etc.) in the future research.

In the next section, the conclusion above will be verified using another orthogonalization method.

2. Gram-Schmidt Orthogonalization

Gram-Schmidt Orthogonalization is a common orthogonalization method. Its main idea is that given a set of vectors, orthogonalize and normalize each vector to all previous vectors in a given order. From a set of linearly independent factor column vectors $f^1, f^2, f^3, \dots, f^K$, we can construct a set of orthogonal vectors $\tilde{f}^1, \tilde{f}^2, \tilde{f}^3, \dots, \tilde{f}^K$

step by step. Orthogonalized vectors can be written as:

$$\tilde{f}^1 = f^1$$

$$\tilde{f}^2 = f^2 - \frac{\langle f^2, \tilde{f}^1 \rangle}{\langle \tilde{f}^1, \tilde{f}^1 \rangle} \tilde{f}^1$$

$$\tilde{f}^3 = f^3 - \frac{\langle f^3, \tilde{f}^1 \rangle}{\langle \tilde{f}^1, \tilde{f}^1 \rangle} \tilde{f}^1 - \frac{\langle f^3, \tilde{f}^2 \rangle}{\langle \tilde{f}^2, \tilde{f}^2 \rangle} \tilde{f}^2$$

$$\ldots = \ldots \qquad \ldots \qquad \ldots$$

$$\tilde{f}^k = f^k - \frac{\langle f^k, \tilde{f}^1 \rangle}{\langle \tilde{f}^1, \tilde{f}^1 \rangle} \tilde{f}^1 - \frac{\langle f^k, \tilde{f}^2 \rangle}{\langle \tilde{f}^2, \tilde{f}^2 \rangle} \tilde{f}^2 - \cdots - \frac{\langle f^k, \tilde{f}^{k-1} \rangle}{\langle \tilde{f}^{k-1}, \tilde{f}^{k-1} \rangle} \tilde{f}^{k-1}$$

After normalization:

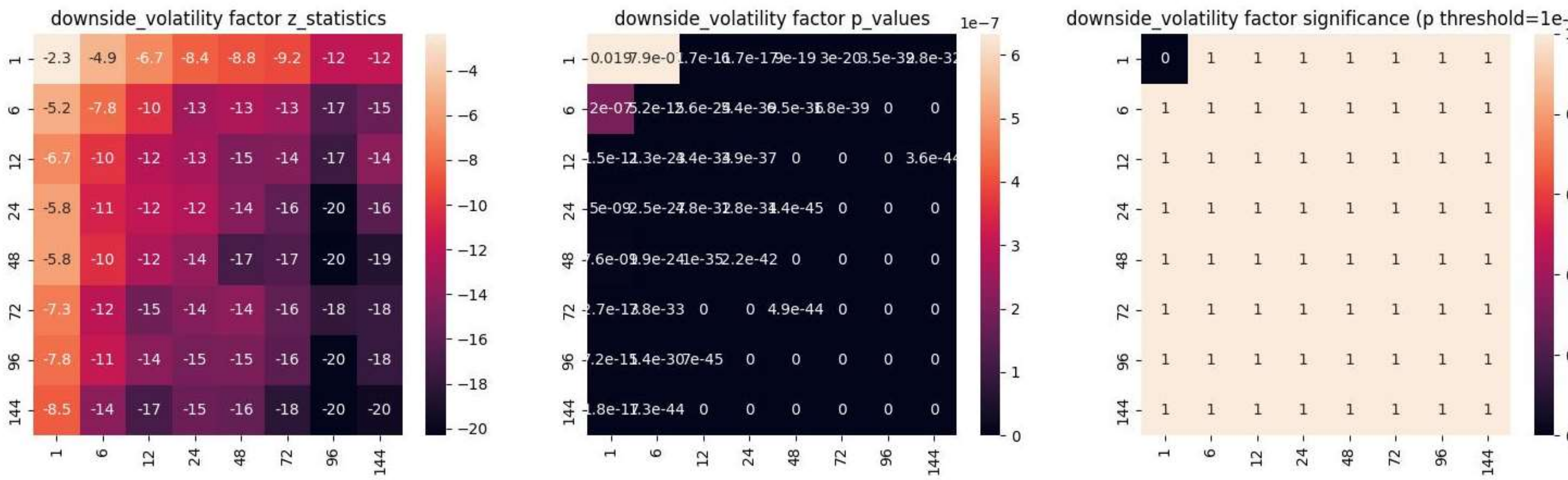
$$e^k = \frac{\tilde{f}^k}{\|\tilde{f}^k\|} (k = 1, 2, \ldots, K)$$

While using Gram-Schmidt Orthogonalization, the order of factor orthogonalization must be predefined. It can be a fixed sequence, or be dynamically determined according to some criteria. In this research, we use a fixed sequence: ['downside_volatility_1', 'skewness_1', 'trend_strength_1', 'kurtosis_1', 'price_volume_correlation_1'], intuitively following the rank of their significance get from previous tests.

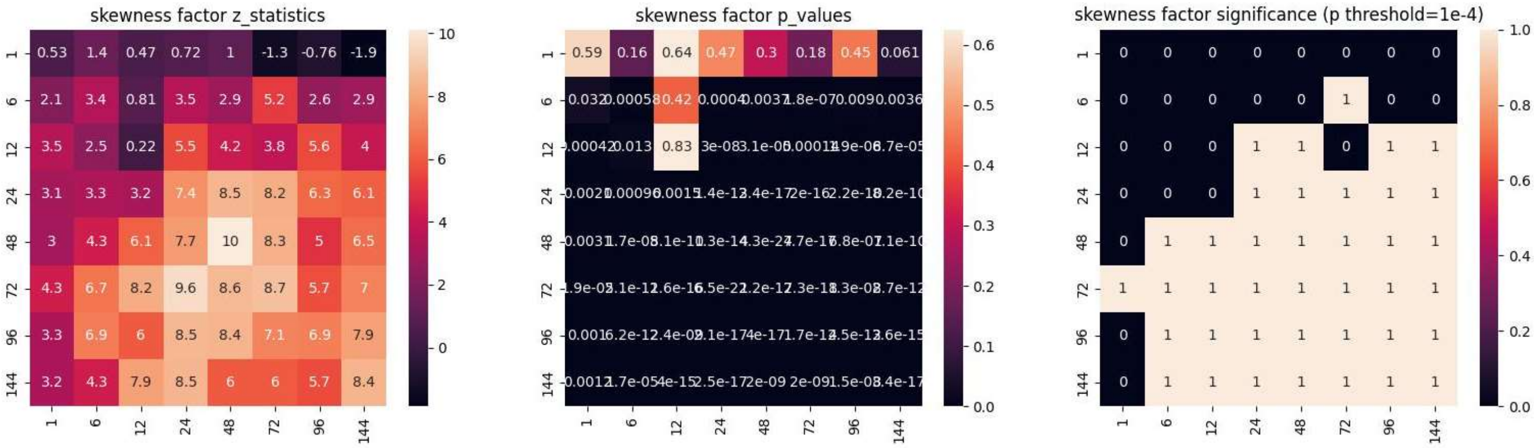
In nature, Gram-Schmidt Orthogonalization is a series of repetitive linear regressions. One new variable will be taken as explained variable while the existing variables are explanatory variable. After linear regression, the new variable will be transformed to the residuals of this regression, then added to existing variables. Hence, this method is suitable to test whether a newly discovered factor provides extra information given existing factors.

Results are presented below:

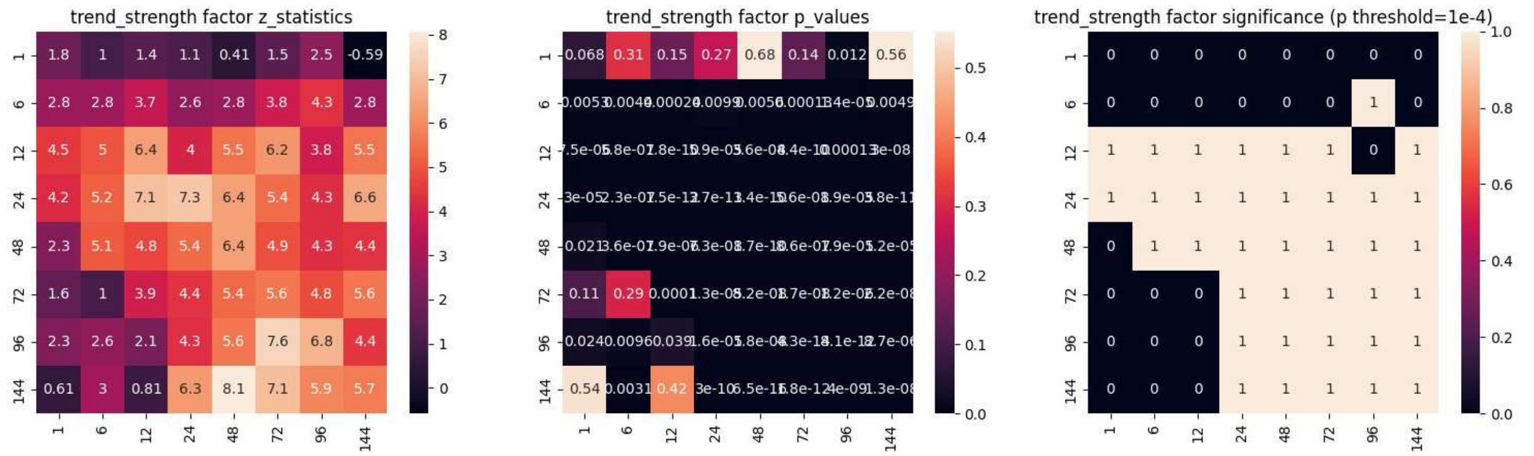
1) Downside volatility



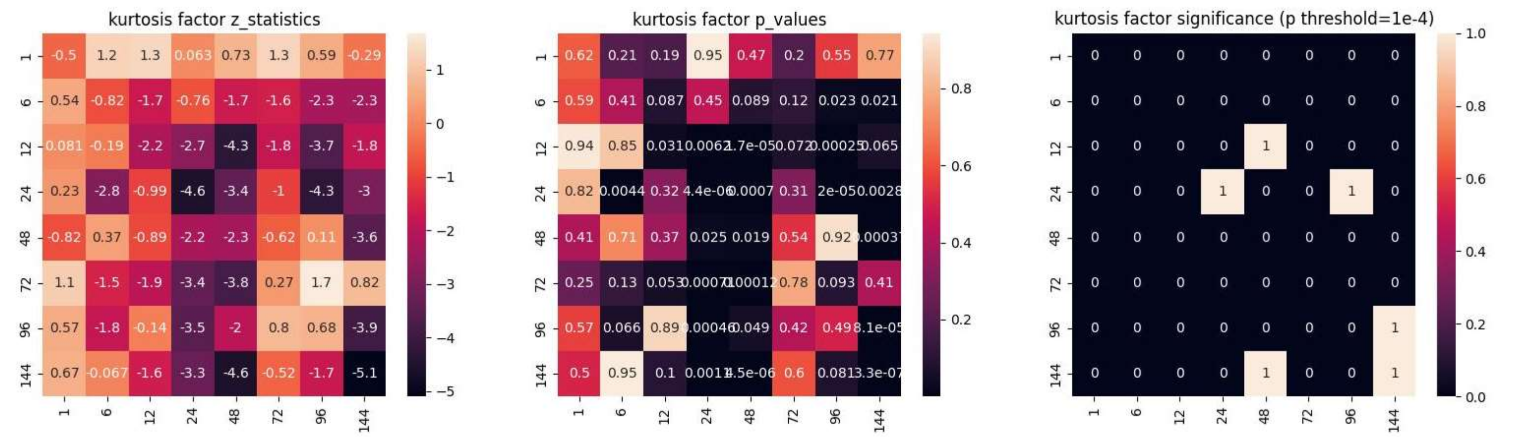
2) Skewness



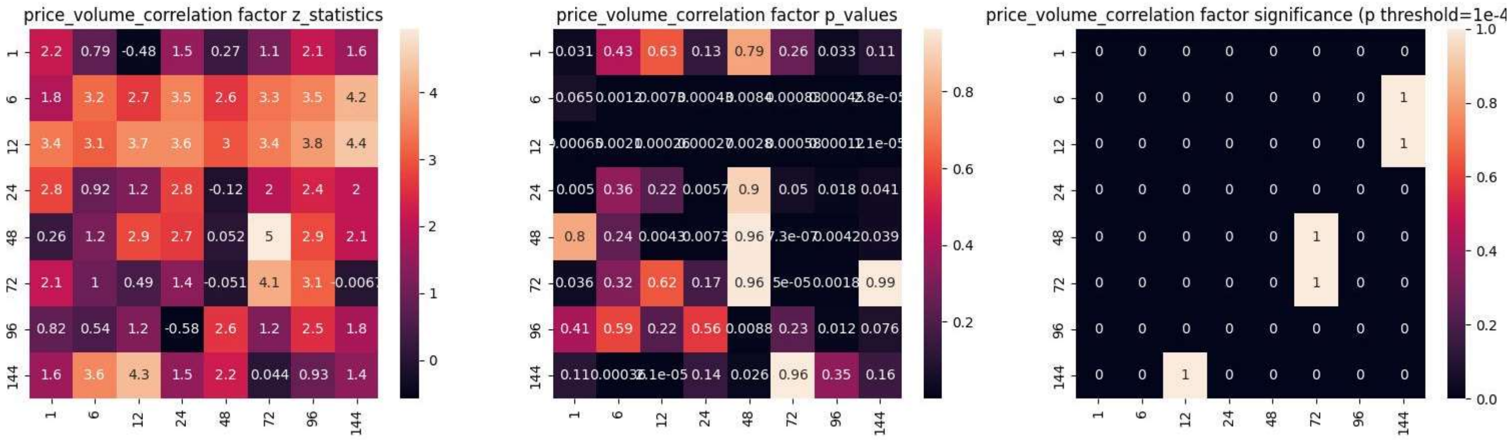
3) Trend strength



4) Kurtosis



5) Price-volume correlation



We have observed similar results as symmetric orthogonalization in section 1.

So far, we have constructed a small but complete pipeline for factor research. It will be easy to add in new factor, either calculated from fundamental data or dug out using deep learning method. After factor selection, the next topic is to assign weights each of them for better portfolio selection. It can be either simple as equal weights or complicated as recurrent neural network.