The paper “Are Transformers Effective for Time Series Forecasting?”, published in 2022, presents a critical examination of the use of Transformer-based models in the domain of long-term time series forecasting (LTSF). In recent years, Transformers, known for their effectiveness in handling long sequences in fields like natural language processing (NLP) and computer vision, have been increasingly applied to time series analysis. However, the authors of this paper challenge the appropriateness of Transformers for LTSF tasks. Transformers, originally designed for NLP tasks, are built upon the concept of self-attention, which allows them to process each element of an input sequence in the context of all other elements. This design is particularly effective in capturing the semantic correlations among the elements in a sequence. For example, in language processing, Transformers can understand the relationship between words in a sentence regardless of their positional distance from each other. This capability is largely due to the architecture's ability to weigh the influence of different parts of the input data without the constraints of sequential processing, as seen in traditional models like Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks. In contrast to Transformers, LSTM networks are a type of RNN specifically designed to remember long-term dependencies in sequential data. LSTMs achieve this through a complex system of gates that regulate the flow of information, allowing them to retain or forget information over long sequences. This makes LSTMs particularly suited for tasks where the order of data points is crucial, such as time series forecasting. The authors of the paper argue that the core strength of Transformers in semantic analysis may not translate effectively to time series forecasting. The main concern raised is the permutation-invariant nature of the self-attention mechanism in Transformers, which could lead to a loss of temporal information crucial in time series data. In time series analysis, the order of data points is of paramount importance as it directly represents the progression of time and the associated changes. Although Transformers employ positional encoding to retain some order information, the authors suggest that this may not be sufficient for accurately capturing temporal dynamics inherent in time series data. To substantiate their claims, the authors introduce LTSF-Linear, a simple linear model, and demonstrate that it surprisingly outperforms complex Transformer-based models in LTSF tasks across various real-life datasets. This finding is significant as it suggests that simpler models, which are inherently more adept at handling the ordered nature of time series data, may be more suitable for LTSF tasks than the currently popular Transformer-based models. Furthermore, the paper delves into a comprehensive empirical study exploring the impact of various design elements of LTSF models, such as the ability to model long inputs, sensitivity to the order of time series, the effect of positional encoding, and the efficiency of different models. The results indicate that the forecasting errors of Transformer-based models do not decrease (and sometimes even increase) with the expansion of look-back window sizes, questioning their ability to extract and utilize long-term temporal relationships effectively. In conclusion, the paper urges the scientific community to reassess the efficacy of Transformer-based models for time series forecasting, particularly in the context of LTSF. It opens new avenues for research in developing more appropriate models for time series analysis, emphasizing the need for methods that can more accurately capture and leverage the temporal dynamics of data. The authors also advocate for future studies to revisit the use of Transformer-based solutions in other time series analysis tasks beyond LTSF.