Chapter Summaries:

**Chapter 1**: We provide a brief introduction to market microstructure and trading from practitioners’ point of view. The terms used in this chapter, all can be traced back to academic literature; but the discussion is kept simple and direct. The data, which is central to all the analyses and inferences, is then introduced. The complexity of using data that can arise at irregular intervals can be better understood with Level III data, as illustrated here. Finally, the last part of this chapter contains a brief academic review of market Microstructure—a topic about the mechanics of trading and how the trading can be influenced by market designs. This evolving field is of much interest to all: regulators, practitioners and academics.

**Chapter 2:** We provide a review of some basic and some advanced statistical methodologies that are useful for developing trading algorithms. We begin with time series models for univariate data and provide a broad discussion of autoregressive integrated moving average (ARIMA) models—from model identification, estimation, inference to model diagnostics. The stock price and return data exhibit some unique features and so we identify certain stylized facts regarding their behavior that have been empirically confirmed; this work will greatly help to discern any anomalies as and when they arise, as these anomalies generally indicate deviation from efficient market hypothesis. Although for modeling price and return data, only lower order (ARIMA) models are needed, increasingly other trading features such as volume, volatility—that are being used in developing trading strategies—require higher order models. In particular, predicting future volume flow is useful to determine when to enter the market. Wherever possible we illustrate the methodology with examples and make the data accessible to the reader. We also introduce some novel methodologies that have potential for developing competing trading models.

**Chapter 3:** Methodologies for analyzing multiple time series data are discussed. These have applications from pairs trading to portfolio optimization to trading in multiple markets. A key feature of this chapter is the discussion of multivariate dimension-reduction techniques that play an important role in large-scale data modeling. Because the data is chronological, the commonly used techniques such as principal component analysis may not be valid; we discuss the concept of reduced-rank regression that has become widely used in practice. Modeling of multiple non-stationary series such as stock prices is discussed. The concept of co-integration, commonality, co-movement are covered in detail with illustrative examples. Recent research on Multivariate GARCH models is also discussed with applications in foreign exchange market.

**Chapter 4**: This chapter contains a discussion on topics such as State-Space models, Regime Switching and Theory of Point Processes as trading takes place on a continual basis. Established stylized facts in the High Frequency Trading context including models for multiple assets are discussed. The use of volume via its relationship to volatility is summarized. As most of the trading outfits still use price-bar data, analysis based on time-aggregated data is mentioned. Finally, this chapter also contains a very brief treatment of other modern topics in machine learning: Neural Networks, Reinforcement Learning and Boosting methods.

**Chapter 5:** We present trading algorithms based on statistical analyses of market data. These analyses are also guided by established principles in financial economics. The market is assumed to consist of informed and noise traders. It is also postulated that the market is efficient, and the informed traders generally gain at the expense of noise traders. Any information about a stock is quickly impounded in price by the Efficient Market Hypothesis (EMH). If there are anomalies, statistical arbitrage techniques can be employed to successfully detect and exploit them for profitable opportunities. We present some commonly used strategies based on the statistical methods described in Chapters 2–4. Although the literature is replete with numerous strategies, most of them are some variations or combinations of basic strategies described here. As price-based strategies have been more or less fully exploited, it is important to consider strategies that use other information such as volume, market volatility, related stock performance, etc. Also presented are improved strategies based on machine learning methods.

**Chapter 6**: We discuss mean-variance portfolio theory—fundamental to diversification of risk over various financial assets. Known and latent multifactor models and tests related to CAPM and APT are discussed with implications for portfolio investing and with an illustrative example. Regularization methods in the context of large-scale portfolio allocation are mentioned, as well as how the trading strategies are handled. As portfolio tracking and rebalancing is quite common in practice and because rebalancing requires trading, this chapter provides a comprehensive view of portfolio trading strategies that take into consideration transaction costs.

**Chapter 7:** This chapter deals with the use of automated analysis of news and market sentiment data for trading. This is essential because market consists of noise traders who may trade on sentiment and the short-term price movements are not easily explained by the relevant indicators that may be slow to vary; the commonly studied stock related indicators such as earnings announcements occur only on periodic basis. The sentiment analysis captures an understanding of how noise traders are likely to trade and this topic can be considered as part of the emerging area of Behavioral Finance. This chapter presents an analysis of sentiments captured in stock related tweets and refined by Natural Language Processing techniques with an application to trading strategies.

**Chapter 8:** Execution strategies have come to play an important role in Algorithmic Trading. To develop successful execution strategies, it is necessary to have a good understanding of the limit order book dynamics. In this chapter, we cover from the practitioner’s perspective, the normalizing analytics in capturing microstructure signals. The construction of the limit order book and its key descriptive indicators are covered. Recent work on modeling the dynamics of order book via Hawkes process is discussed and some empirical results on hidden liquidity are discussed based on Level III data.

**Chapter 9:**  In this chapter, we focus on one of the fundamental subjects of execution: Market Impact (MI). The impact of current trade on the future trades is of interest to traders to develop pre and post trade analytics. We bring in the intuition on the inner works of market impact and build the functional form of standard MI models from first principles. We then go into more details through a review of select literature to guide the reader who may be interested in more advanced studies. The empirical estimation of transaction costs is discussed through a unique data set, which confirms the square root law that is commonly followed in industry.

**Chapter 10:** This chapter begins with the standard execution benchmarks from the practitioner’s point of view. Then the evolution of the execution strategies across the three fundamental layers: Scheduling, Order Placement and Order Routing are reviewed. Formal description of select execution models is also provided. The smart routing algorithm which is essential in the context of fragmented markets is discussed. Finally, execution algorithms for multiple assets and extension of the algorithms to other asset classes beyond equities are covered.

**Chapter 11:** Algorithmic Trading is primarily a technology endeavor, involving the orchestration and connectivity of numerous different systems. In this chapter we describe the full technology stack necessary to operate a trading business, from a standard Broker-Dealer infrastructure, to specificities of an HFT infrastructure, to requirements and flexibilities of an ATS infrastructure. A better understanding of the inner workflows and connections between different components will equip the reader with hands-on experience to deploy a trading platform that is stable and scalable.

**Chapter 12:** Designing and creating high performing algorithms is no simple matter and requires significant research work on the modeling and calibration of the strategy itself and the underlying execution. This chapter provides a high-level view of the necessary research stack and what is entailed in running a large trading operation at scale, from the data infrastructure, to the automated calibration and deployment of models, and to simulation environment to be used as a testing ground. Finally, we address the need to develop flexible and powerful transaction cost analysis (TCA) environment in order to measure, compare and improve performance of strategies in various settings.