DATA 201: Time Series Analysis
Lecture 1: Introduction

Lulu Wang

Data Analytics
Dickinson College

1/21/2025

Course Overview and Logistics

Introduction

Examples

Problems

Course Logistics

- **Instructor: Dr. Lulu Wang**
- **Class Schedule:**
  - Tome 120, Tues/Fri 1:30 - 2:45 pm
- **Office Hours: Calendly link**
  - Fri 3:00 - 4:00 pm
  - Tues/Thurs 12:00 -1:00 pm
  - OR by appointment
- **Contact: wanglu@dickinson.edu**
- **Course Website: Moolde**
- **Textbook:**
  - No textbook is required; a comprehensive list of references is included in the syllabus. Lecture notes and readings will be provided.

The Plan

This course aims to introduce you to quantitative financial data analysis **through a practical, hands-on approach.**

- **Linear Regression Models:**
  - Widely used in finance to investigate relationships, e.g., measuring stock or portfolio "beta" with respect to the market.
- **Time Series Models:**
  - Focus on predicting future values of a variable based on its past (e.g., AR, MA, ARIMA models).
  - Essential in finance for modeling price and return dynamics, especially when economic indicators are imperfect or infrequent.
- **Volatility Models:**
  - Such as GARCH (Generalized Autoregressive Conditional Heteroskedasticity) and its variants. These models assume volatility changes over time and aim to forecast risk, a central concept in portfolio management and risk control.

The Plan

Every lecture note integrates the discussion of financial econometric techniques with Python implementations. You will learn:

- **Data Manipulation:**
  - Using pandas for reading, cleaning, and transforming datasets.
- **Visualization:**
  - Using matplotlib or seaborn to create informative plots.
- **Statistical Modeling:**
  - Using statsmodels and scikit-learn to fit regression or time-series models.
- **Efficient Workflows:**
  - Writing Python scripts/notebooks that clean data, perform analyses, generate tables, and create publication-ready figures.

Python and Programming

- Python will be our primary focus throughout the course.
- Use of Python and Jupyter Notebook will be required for this course.
- You may choose from various environments, such as Google Colab (cloud-based, requires a Google account) or an IDE like Visual Studio Code.

Prerequisites

- DATA 180, MATH 170, and MATH 121 (or DATA 180 and MATH 225).
- Prior exposure to basic time series concepts and the Python language is beneficial but not required.

Learning Outcomes

- Discuss the linear regression model, its assumptions, and its relevance in finance.
- Build time series models to forecast economic and financial variables.
- Measure financial risks under different distributional assumptions.
- Develop, implement, and present (both written and orally) a financial data analysis project.

Grading Rubric

- Participation (10%)
- Assignments (35%)
- Midterm (20%)
- Final project (35%)

Read the syllabus posted on Moolde carefully: all policies, especially academic integrity.

Assignments, Exams, Project Discussion

- **5-8 Assignments.**
- **Midterm:** One midterm
- **Final Project:**
  - Teams of 2-4 members.
  - Investigate a financial or economic problem via empirical analysis in Python.
  - Proposal due mid-semester.
  - You will produce a written report and a brief presentation.

Recommended Papers (1/6)

The list below provides some papers that you might find useful as a starting point for your project:

- W.N. Goetzmann, L. Peng, J. Yen (2009), *The Subprime Crisis and House Price Appreciation*, NBER working paper 15334.
- C. Himmelberg, C. Mayer, T. Sinai (2005), *Assessing High House Prices: Bubbles, Fundamentals, and Misperceptions*, Staff Report n. 218, NY Fed.
- C.S. Asness (2000), *Stocks vs Bonds: Explaining the Equity Risk Premium*, Financial Analysts Journal, 56(2), pp. 96-113.

Recommended Papers (2/6)

- C.J. Neely, D.E. Rapach, J. Tu, G. Zhou (2014), *Forecasting the Equity Risk Premium: The Role of Technical Indicators*, Management Science, 60(7), pp. 1772-1791.
- R. Cantor and F. Packer (1996), *Determinants and Impact of Sovereign Credit Ratings*, FRBNY Economic Policy Review, October.
- M. Scheicher (2008), *How Has CDO Market Pricing Changed During the Turmoil? Evidence from CDS Index Tranches*, ECB working paper, n. 910.

Recommended Papers (3/6)

- B. Barber, R. Lehavy, M. McNichols, B. Trueman (2003), *Reassessing the Returns to Analysts' Stock Recommendations*, Financial Analysts Journal, 59(2), pp. 88-96.
- C. Baumeister and L. Kilian (2013), *Forecasting the Real Price of Oil in a Changing World: A Forecast Combination Approach*, CFS working paper.
- C.B. Erb and C.R. Harvey (2006), *The Strategic and Tactical Value of Commodity Futures*, Financial Analysts Journal, 62(2), pp. 69-97.

Recommended Papers (4/6)

- A. Petajisto (2013), *Active Share and Mutual Fund Performance*, Financial Analysts Journal, 69(4), pp. 73-93.
- J. Hasanhodzic and A. Lo (2007), *Can Hedge Fund Returns Be Replicated? The Linear Case*, Journal of Investment Management, 5(2), pp. 5-45.
- X. Lou and R. Sadka (2011), *Liquidity Level or Liquidity Risk? Evidence from the Financial Crisis*, Financial Analysts Journal, 67(3), pp. 51-62.

Recommended Papers (5/6)

- L.K.C. Chan, N. Jegadeesh, J. Lakonishok (1999), *The Profitability of Momentum Strategies*, Financial Analysts Journal, 55(6), pp. 80-90.
- M.M. Copeland and T.E. Copeland (1999), *Market Timing: Style and Size Rotation Using the VIX*, Financial Analysts Journal, 55(2), pp. 73-81.
- M. Haug and M. Hirschey (2006), *The January Effect*, Financial Analysts Journal, 62(5), pp. 78-88.
- S.C. Andrade, V. Chhaochharia, M.E. Fuerst (2013), *Sell in May and Go Away... Just Won't Go Away*, Financial Analysts Journal, 69(4), pp. 94-105.
- M. Pojarliev and R.M. Levich (2008), *Do Professional Currency Managers Beat the Benchmark?*, Financial Analysts Journal, 64, pp. 18-32.

Recommended Papers (5/6)

- L. Menkhoff, L. Sarno, M. Schmeling, A. Schrimpf (2012), *Currency Momentum Strategies*, Journal of Financial Economics, 106(3), pp. 660-684.
- C.J. Neely and P.A. Weller (2011), *Technical Analysis in the Foreign Exchange Market*, St. Louis Fed, working paper.
- V. Bhansali (2008), *Tail Risk Management*, Journal of Portfolio Management, 34, pp. 68-75.

Main References

- Tsay, *Analysis of Financial Time Series*, Wiley.
- Tsay, *An Introduction to Analysis of Financial Data with R*, Wiley.
- Hilpisch, *Python for Finance: Mastering Data-Driven Finance*, O'Reilly.
- Christoffersen, *Elements of Financial Risk Management*, Academic Press.
- Enders, *Applied Econometric Time Series*, Wiley.
- Hamilton, *Time Series Analysis*, Princeton University Press.
- Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, O'Reilly.

Academic Integrity

- While discussing assignments with your peers is encouraged, all work you submit must be entirely your own. You are NOT allowed to show your solution to a classmate, ask to see another student's solution.
- You may not use code from the internet (e.g., ChatGPT, StackOverflow) or post your code and answers where others can access them.
- **As a general rule, you are expected to write and debug your own code without copying others' work.** If you are found to have cheated or plagiarized, you will receive a zero for the assignment on your first offense. A second incident will lead to a failing grade for the course.

What is Financial Time Series Analysis (Financial Econometrics)?

- **Definition:** Application of statistical/econometric methods to financial data.
- This field has become increasingly important as financial markets continue to generate massive amounts of data every day.

  Importance:
  - Understand market behavior
  - Forecast future returns, volatility, risks
  - Make informed investment decisions

  Scope:
  - Modeling asset prices & returns
  - Risk management (Value-at-Risk, volatility forecasts)
  - High-frequency trading analytics

Example 1: Annual Data (Valuation Ratios)

Valuation Ratios: Price-to-Earnings (PE) & Price-to-Dividends (PD)

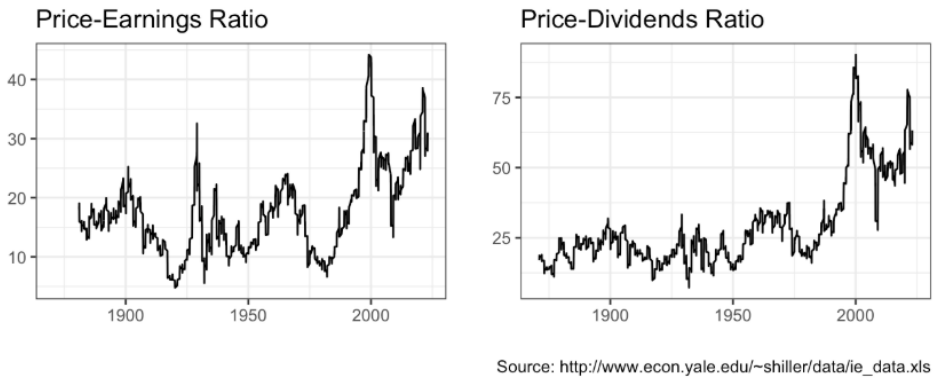

Figure 1.1: Annual Cyclically Adjusted Price-to-Earnings (CAPE) and Price-to-Dividends ratio for the Standard and Poors 500 Index starting in 1871.

Long-Term View:

- Data from 1871 to present (Robert Shiller's dataset)
- Historically, PE $\approx$ 5–25, soared above 40 in late 1990s
- PD ratio has also varied widely

Questions:

- What drives market valuations over the long run?
- Are current valuations too high or justified by fundamentals?

Example 2: Daily Data (S&P 500)

Daily Price & Returns:

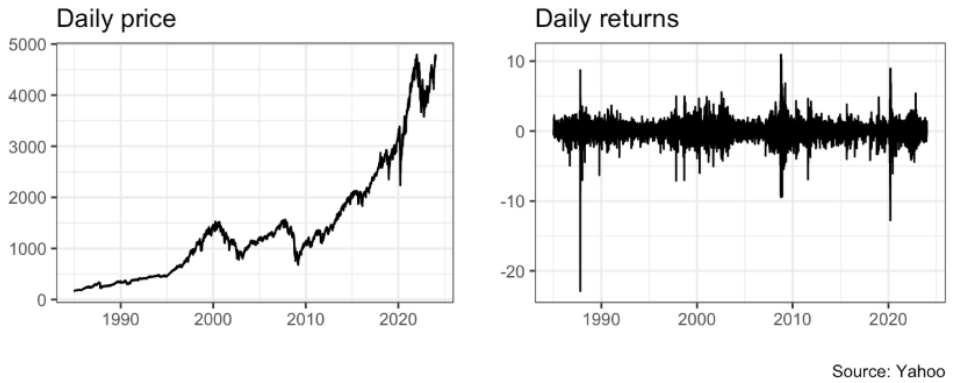

Figure 1.2: Daily prices and returns for the Standard and Poor 500 Index.

- S&P 500 from 1985 to 2024
- Notable crashes (e.g., 1987, 2008, 2020 pandemic)

Volatility Clustering:

- Calm vs. turbulent periods
- Large daily swings in crisis times

Questions:

- Why do these boom-bust cycles occur?
- Are returns or volatility predictable?

Example 3: High-Frequency Data (Intraday FX)

Intraday Quotes: USD/JPY at 1-minute intervals

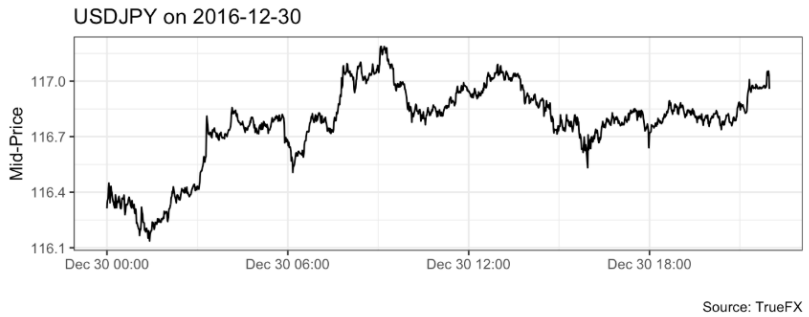

Figure 1.3: Intra-day mid-point between bid and ask price for the USD to JPY exchange rate sampled at the 1 minute frequency.

Key Observations:

- Thousands of quotes per day per asset
- Bid-ask spreads vary with liquidity and market events

Challenges:

- Massive data (millions of rows per month)
- Need for specialized data handling

Potential Research Topics:

- Predicting short-term price movements
- Estimating volatility in real-time

Example 4: Cross-Sectional Data

Scatter Plot: Log(Market Cap) vs. Next-Month Returns

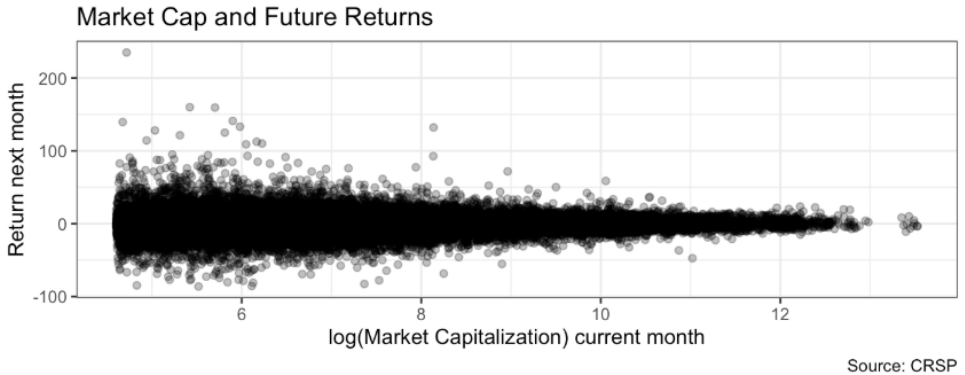

Figure 1.4: Scatter plot of the logarithm of the market capitalization for a stock listed in the NYSE, AMEX, or NASDAQ against the percentage return in the following month. Each point represents a stock-month pair in 2015. Stocks with market cap below 100 million dollars are dropped from the sample.

- Example: All NYSE/AMEX/NASDAQ stocks in 2015

Questions:

- Do small-cap stocks outperform large-cap stocks?
- Is size a predictor of future returns?
- Are small caps riskier or just higher returning on average?

Data Structure:

- Could be a single snapshot or panel data (time + cross-sections)

Problem 1: Stock Return Risk Analysis

- **Objective**: Explain excess return of a stock (risk premium) by decomposing it into different risk components (factors).
- **Two type of risk**
  - *Systematic Risk*: Market-wide influences (e.g., macroeconomic conditions).
  - *Idiosyncratic Risk*: Firm-specific factors (e.g., company news, management changes).

Solution 1.1: Capital Asset Pricing Model (CAPM)

- **Core Idea**
  - The *excess return* of a stock is determined by its exposure to market-wide risk (β).
  - Other firm-specific risks are assumed to be diversifiable in an efficient market.
- **Excess Return Formula**

$$R_i - R_f = \beta_i \Big[R_m - R_f \Big] + \epsilon$$

  - R_i: Return of the stock i
  - R_f: Risk-free rate
  - β_i: Sensitivity of stock i to the market
  - R_m: Return of the market
- **Pros & Cons**
  - *Pros*: Simple, widely used, foundational model.
  - *Cons*: Single-factor approach; ignores other known return drivers (e.g., size, value).

Solution 1.2: Fama–French Factor Models

- **Motivation**
  - Empirical evidence shows that factors beyond market beta (e.g., size, value) explain stock returns.
  - Fama–French 3-Factor Model adds SMB (Size) and HML (Value).
- **3-Factor Model Equation**

$$R_i - R_f = \alpha + \beta_{\text{MKT}}(R_{\text{MKT}} - R_f) + \beta_{\text{SMB}}\,\text{SMB} + \beta_{\text{HML}}\,\text{HML} + \epsilon$$

  - SMB: Return spread of Small vs. Big market-cap stocks.
  - HML: Return spread of High vs. Low book-to-market (Value vs. Growth).
- **Extensions**
  - 5-Factor Model adds Profitability (RMW) & Investment (CMA).
  - Multiple factors capture different dimensions of systematic risk.

Solution 1.3: Barra's Multi-Factor Models

Barra-Type Factor

- **Purpose**
  - Risk decomposition and portfolio construction at a granular level.
  - Factor-based approach to identify risk exposures (style, industry, macro, etc.).
- **Key Features**
  - *Statistical vs. Fundamental Factors*:
    - Statistical: Derived from covariance matrices using principal components.
    - Fundamental: Based on observable firm attributes (growth, liquidity, etc.).
  - Modular design for different markets (e.g., Barra U.S. Equity Model).
- **Usage**
  - Helps portfolio managers monitor and control factor exposures.
  - Provides risk forecasts and aids in performance attribution.

Problem 2: Factor Investing

- **What is Factor Investing?**
  - An investment strategy aiming to systematically capture return drivers (factors) like value, size, momentum, etc.
  - tilts portfolios toward specific characteristics associated with long-term outperformance.
- **Factor Model Rationale**
  - Certain factor exposures (e.g., small-cap, high-value) have historically offered additional return premia.
  - Provides a structured, rules-based way to select stocks.

Identifying Relevant Factors

- **Common Equity Factors**
  1. Value: Stocks deemed "cheap" (e.g., high book-to-market).
  2. Size: Small-cap vs. large-cap.
  3. Momentum: Recent winners tend to keep winning (and vice versa).
  4. Quality: Strong balance sheets, stable earnings, low debt.
  5. Low Volatility: Stocks with lower return variance.
- **Selecting Factors**
  - *Economic Intuition*: Factors should have a risk-based or behavioral rationale.
  - *Empirical Evidence*: Validate with historical performance.
- **Diversification Benefits**
  - Combining multiple factors can reduce overlaps and stabilize returns.