Chapter 1
Introduction to Forecasting

1 Introduction

These documents are meant to serve as a complement to the notes provided in class. These documents may go into more detail, and I will try to cover different examples, so that concepts are illustrated clearly throughout the course. That being said, it is best to think of these documents as "lecture notes" rather than a formal book.

2 Why Forecast?

Why are we interested in forecasting? In general, humans dislike uncertainty — we would rather be prepared to deal with 6 inches of snow than to have it fall out of the sky unexpectedly. Of course, the topic of this course is business and economic forecasting, but the point is generally true. For example, it is much easier to come up with a plan for production when we can anticipate orders (or a lack thereof) happening. Unfortunately, compared to weather forecasts, it can be harder to forecast in business and economics, since decision makers in the economy react to changing events and even to forecasts themselves.

2.1 Forecasting in the Private and Public Sectors

There are many examples of the importance of forecasts in both private and public sectors. In the private sector, an accurate sales forecast can be of vital importance. A sales forecast can be used when deciding on the marketing budget, when assessing realistic goals for sales growth, and when deciding on how much material to order and how many personel to have on staff at any given point in time. Some additonal private sector forecasts include:

- Forecasts of costs & revenue for tax planning purposes (Accounting).
- Forecasts of sales or sales growth for employee recruitment (Human Resources).

- Forecasts of product demand for planning raw material inventories (Operations Management).

Forecasts are also useful for decision making in the public sector and nonprofits. Some examples include:

- Forecasting demand for transit so cities know how many buses to buy and drivers to hire.
- Forecasting crime in specific areas at specific times of day to know how many officers to have on patrol.
- Forecasting health emergencies for staffing at emergency rooms.
- Forecasting tax collections so local and state governments know how big their budgets can be over the next year.

3 Different Types of Forecasts and Data

3.1 Qualitative vs. Quantitative Forecasts

To get a better handle on foreasting, we can break down different forecasting processes into different categories. One common distinction between forecasting processes is that of qualitative and quantitative forecasts.

Qualitative or "subjective" forecasts are formed without using formal statistical equations. These forecasts are typically based on market research, intuition and experience, or gut instinct. It is best to use qualitative forecasts when there is little available data, or when we expect the behavior of the variable we are forecasting to change dramatically from previous historical patterns. For example, a company forecasting sales of a new product, or forecasting the demand for a pharmaceutical drug during a major disease outbreak would likely be best forecasted by subject matter experts using their past experience as a guide.

Qualitative forecasts can be broken down further into two more categories: implicit and explicit. **Implicit qualitative** forecasts are forecasts that are not formed using any formal procedure. An example may be studying for an exam. When you study for an exam, you probably try to "guess" what will be on the test. However, you probably don't follow the same formal procedure every time you study. On the contrary, you were probably not even aware that what you were doing was *forecasting*!

On the other hand, **explicit qualitative** forecasts are formed using a formal process. These forecasts may even use numbers. For example, baseball scouts typically rate players on a 1-8 scale, and try to evaluate each player objectively and in the same manner. However, these scouts do not use any formal statistical procedures, so they are still considered qualitative forecasts. Another example could be a sales forecast for a new product in which analysts formed forecasts using sales data on similar products and their own intuition. Again, if they do not follow a formal statistical procedure, this is still a qualitative forecast.

Produced under the right circumstances, by the right people, qualitative forecasts can be highly accurate. In his book "Superforecasting: The Art and Science of Prediction", Philip Tetlock explains that the type of people that produce the best forecasts of world events tend to have the following in common:

- Intelligent and intellectually curious, but not "geniuses"
- Follow a formal forecasting process.
- Open minded - for example, they don't let political ideology seep in to their political or economic forecasts.
- Break down hard forecasting questions into easier parts, answer those parts, and aggregate.
- Start with the "outside view" and then move on to the "inside view." For example, if the question was: "What is the probability that the Los Angeles Lakers repeat as NBA champions this year?" they would start by looking at the recent probability that any team has repeated to win the NBA championship. Only then will they either increase or decrease this percentage based on things like roster construction, injuries, etc.

In contrast to qualitative forecasts, **quantitative** forecasts are formed using statistical equations. These are best implemented when we have high quality historical data, and when we expect past patterns in the data to continue into the future. In this course, we will focus most of our attention on quantitative forecasting. Note that *all* quantitative forecasts are explicit forecasts, since we have to actively use equations and follow a formal statistical procedure when we use them. Quantitative methods do not always outperform qualitative methods, but they often serve as a good start — a way to form an "outside view."

Often, the best forecasts don't fit neatly into either category, as they are a combination of quantitative and qualitative forecasts. While a quantitative forecast is often a good starting

Table 1: Strengths and Weaknesses of **Qualitative** Forecasting

Strengths	Weaknesses
No math background needed.	Can take many years to become "good"
Works well when combined with quantitative	Subject to behavioral biases

Table 2: Strengths and Weaknesses of **Quantitative** Forecasting

Strengths	Weaknesses
Statistically Unbiased	Requires training
Assumptions are explicit	Requires high quality data

point, having background knowledge and experience with the variable you are forecasting, the industry in which your are forecasting, or the general topic, will allow you to make adjustments and improvements to the quantitative forecasts. Since this course will be relatively broad, focusing on forecasting many different variables, we will not focus on how to make these qualitative adjustments. Instead, we will focus on how to best create a quantitative forecast for any given data set.

3.2 Cross-Sectional, Time Series, and Panel Data

Before forming a forecast, it is important to recognize the type of data you are working with, since the techniques you will use to form quantitaive forecasts will differ depending on what type of data you have.

Cross-Sectional data contains observations on different entities within the same time period. For example, if I were to give your class a quiz, each student would have a quiz score. Since you all took the quiz during the same time period, I would have data on several entities (students) measured at the same time. This example is cross-sectional data with just one variable (quiz score), but we could also collect additional variables. For example, instead of just collecting your quiz score, I could also give a short survey that asked your age, class year, gender identity, high school GPA, and SAT or ACT score. Since all of this information is collected at the same time, it is still considered cross-sectional data. However, now we have many variables for each entity instead of only one variable for each entity.

Time Series data is data that contains repeated observations on the same variable(s) over time for one entity. For example, the US unemployment rate has been measured every month since January, 1948. This data set is time series data, since it is for one entity (the US) and

the observations are repeated over time. Note that the unemployment rate has always been measured at the same *frequency* (monthly), but it would still be considered time series data if it used to be measured quarterly, and then switched to being measured monthly (i.e. if it had irregular frequency). In this class, we will only work with time series data that has regular frequency.

Like cross-sectional data, time-series data can contain more then one variable. For example, we might have a data set that contains both the US unemployment rate and an index of US industrial production, both measured monthly since January 1948. Since this data is only for one entity (the US), it is still considered time series data.

Finally, **panel** data (sometimes called "longitundinal" data) contains repeated observations on different entitites over time. Panel data is essentially a hybrid of cross-sectional and time series data. For example, if I gave one quiz each week in this class, and I tracked the weekly quiz score for each student, this would be considred panel data. This is because this data set contains observations on different entities (students) repeated over time (weekly). Another example could be the monthly unemployment rate and industrial production in both the U.S. and Canada from January, 1948 through the current month. There are multiple entities (U.S. and Canada) and the measurements are repeated over time.

4 Exercises

1. Name of two possible uses of forecasting in the private sector and two possible uses of forecasting in the public sector that were not mentioned in section 2.1.

2. What is the distinction between a qualitative and a quantitative forecast?

3. Describe a scenario in which a qualitative forecast might perform more accurately then a quantitative forecast and vice-versa.

4. What is the difference between time series, cross-sectional, and panel data?