# M7
## Predictive Analytics

ENİS KAYIŞ

# Instructor

- ## Enis Kayış
  - Assistant Professor in Industrial Engineering
  - Email: enis.kayis@ozyegin.edu.tr
  - Phone: +90 (216) 594-9370

- ## Short Bio:
  - Education:
    - PhD in Management Science and Engineering, Stanford University, 2009
      - Essays on Procurement Contracting
    - MS in Statistics, Stanford University, 2007
    - BS in Industrial Engineering and Mathematics, Bogazici University, 2002
  - Work Experience:
    - Research Scientist at HP Labs, since 2009
      - Research projects in areas including demand estimation, product portfolio management and pricing, healthcare operations, and forecasting

# Module Description

- This course is designed to introduce predictive analytics methods for various parameters in businesses, such as demand estimation. Topics covered are quantitative estimation techniques, time series analysis, moving averages, exponential smoothing, linear trend, ARIMA methods and estimation error analysis. Real-world data and cases will be used to study applications of these techniques using statistical software.

# Suggested Readings

- Required Textbook:
  - Forecasting: principles and practice, by Rob J. Hyndman and George Athanasopoulos, http://otexts.com/fpp2/
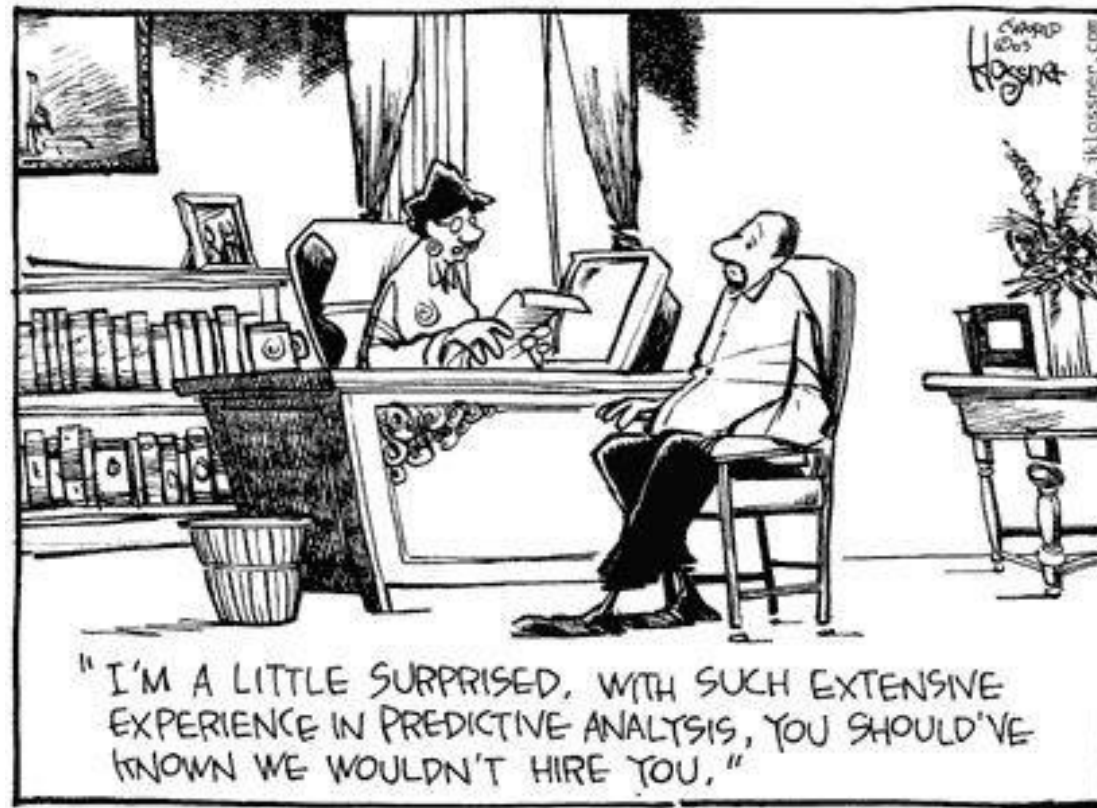
# Grading

- Homeworks (100%)
  - 3 homeworks

# Software Packages

▸ What are the available computer predictive analytics packages?

  ▸ There are a lot of computer packages with forecasting tools included are available such as: SPSS, Minitab, SAS, Eviews, Excel, Oracle, … etc

▸ We are going to actively use Python as part of this class

▸ When necessary, we will also use rpy2 package of Python to leverage some functions of R.

# Predictive Analytics

"I'M A LITTLE SURPRISED. WITH SUCH EXTENSIVE EXPERIENCE IN PREDICTIVE ANALYSIS, YOU SHOULD'VE KNOWN WE WOULDN'T HIRE YOU."
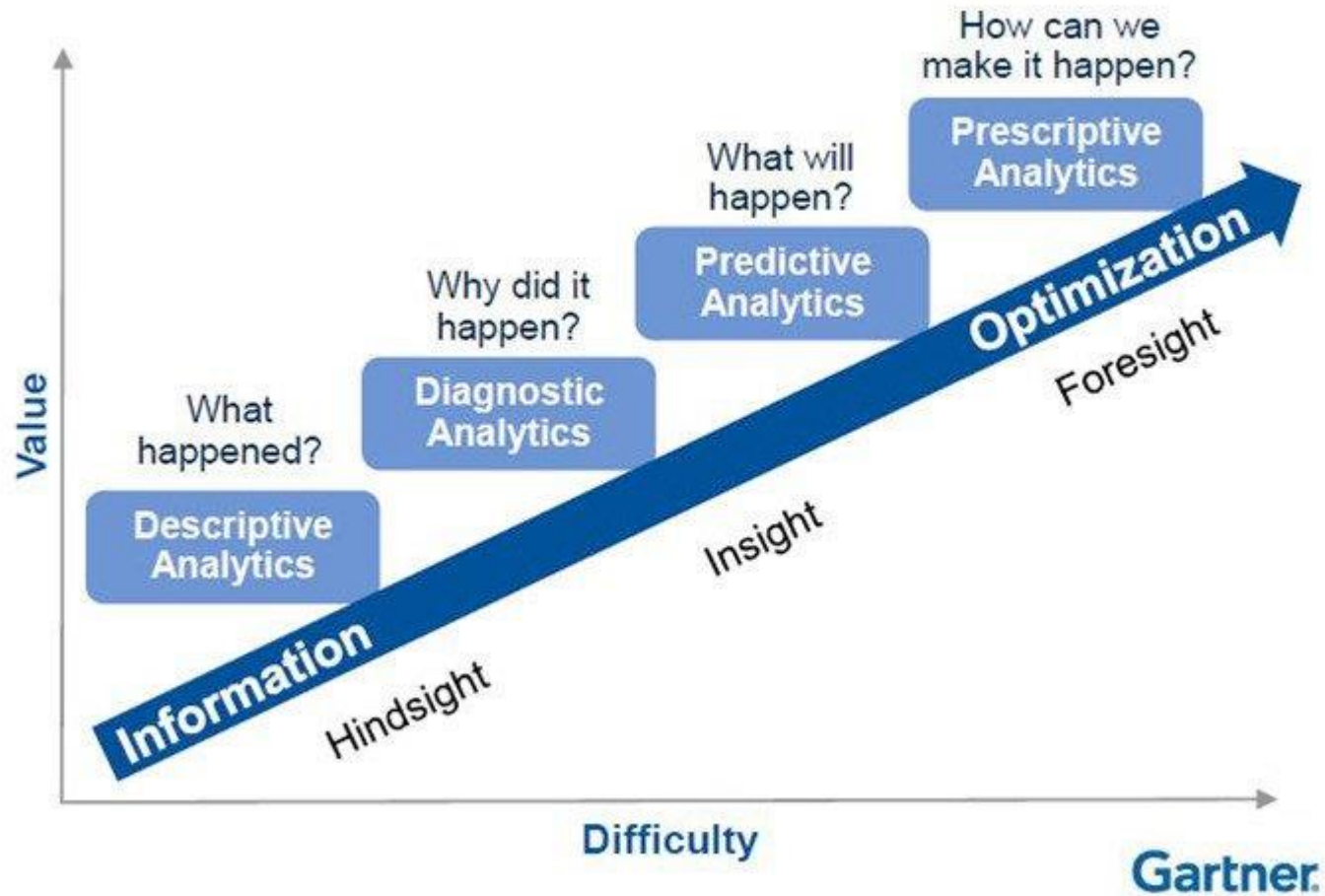
# What is Predictive Analytics?

- ▶ Predictive Analytics
  - ▶ is the art and science of using data to make better informed decisions.
  - ▶ helps you predict what may happen in the future
  - ▶ helps you uncover hidden patterns and relationships in your data
  - ▶ provides you with valuable, actionable insights.

# Analytics Spectrum

# Predictive Analytics Process

- *A predictive model is made up of a number of predictors, variable factors that are likely to influence or predict future behavior.*

- *The end result is both a set of factors that predict, to a relatively high degree, the outcome of an event, as well as what that outcome will be.*

  - *A customer's gender, age and purchase history might predict the likelihood of a future sale.*

- *To create a predictive model, data is collected for the relevant factors, a statistical model is formulated, predictions are made and the model is validated.*

# Predictions versus Forecasts

▸ Forecasting would be a subset of prediction.

▸ Any time you predict into the future it is a forecast.

▸ All forecasts are predictions, but not all predictions are forecasts:

   ▸ e.g.: Use regression to explain the relationship between two variables (sales and marketing budget).

# Forecasting

▸ Most cultures have been concerned with forecasting.

- ▸ Sometimes the forecaster was held in high regard, as was the oracle at Delphi.
- ▸ Often, however, forecasting is regarded as a necessary evil and is frowned upon.
  - ▸ According to a current sage (Drucker, 1973, p. 124), " . . . forecasting is not a respectable human activity and not worthwhile beyond the shortest of periods."
- ▸ Sometimes it has been illegal.
  - ▸ For example, in Rome in 357 A.D. Emperor Constantius II issued an edict forbidden anyone " to consult a soothsayer, a mathematician, or a forecaster... May curiosity to foretell the future be silenced forever,"
- ▸ In recent years, however, forecasting have become a respectable activity.

# Course Material

| Week | Subject |
|------|---------|
| 1 | Introduction to predictive analytics and forecasting |
| 2 | Time series decomposition |
| 3 | Naïve and Moving average methods |
| 4 | Exponential smoothing |
| 5 | ARIMA models |
| 6 | Seasonal ARIMA models |
| 7 | Multiple Regression and Forecasting |

# Introduction to Forecasting

- **Why Forecasting is Necessary?**
  - All organizations operate in an atmosphere of uncertainty so that decisions must be made today to affect the future organization.
  - Educated guess versus pure guess.
  - The most effective forecasting is to combine both quantitative and qualitative techniques.

- **Who needs Forecasting and why?**
  - Almost every organization needs explicit or implicit forecasting, because every organization must plan to meet the conditions of future for which it has imperfect knowledge.

# Forecasting is difficult

## A Timeline of
## Very Bad Future Predictions

**1800**

Rail travel at high speed is not possible, because passengers, unable to breathe, would die of asphyxia.

Dr. Dionysys Larder, Professor of Natural Philosophy & Astronomy, University College London

**1859**

Drill for oil? You mean drill into the ground to try and find oil? You're crazy!

Associates of Edwin L. Drake refusing his suggestion to drill for oil in 1859 (Later that year, Drake succeeded in drilling the first oil well.)

**1876**

This telephone has too many shortcomings to be seriously considered as a means of communication.

Western Union internal memo

**1880**

Everyone acquainted with the subject will recognize it as a conspicuous failure.

Henry Morton, president of the Stevens Institute of Technology, on Edison's light bulb

**1902**

Flight by machines heavier than air is unpractical and insignificant, if not utterly impossible.

Simon Newcomb, Canadian-American astronomer and mathematician, 18 months before the Wright Brothers' flight at Kittyhawk

**1903**

The horse is here to stay, but the automobile is only a novelty, a fad.

The president of the Michigan Savings Bank, advising Henry Ford's lawyer not to invest in the Ford Motor Company

**1916**

The idea that cavalry will be replaced by these iron coaches is absurd. It is little short of treasonous.

Comment of Aide-de-camp to Field Marshal Haig, at tank demonstration

**1916**

The cinema is little more than a fad. It's canned drama. What audiences really want to see is flesh and blood on the stage.

Charlie Chaplin, actor, producer, director, and studio founder

**1921**

The wireless music box has no imaginable commercial value. Who would pay for a message sent to no one in particular?

Associates of commercial radio and television pioneer, David Sarnoff, responding to his call for investment in the radio

**1946**

Television won't last because people will soon get tired of staring at a plywood box every night.

Darryl Zanuck, movie producer, 20th Century Fox

**1977**

There is no reason for any individual to have a computer in his home.

Ken Olson, president, chairman and founder of Digital Equipment Corporation
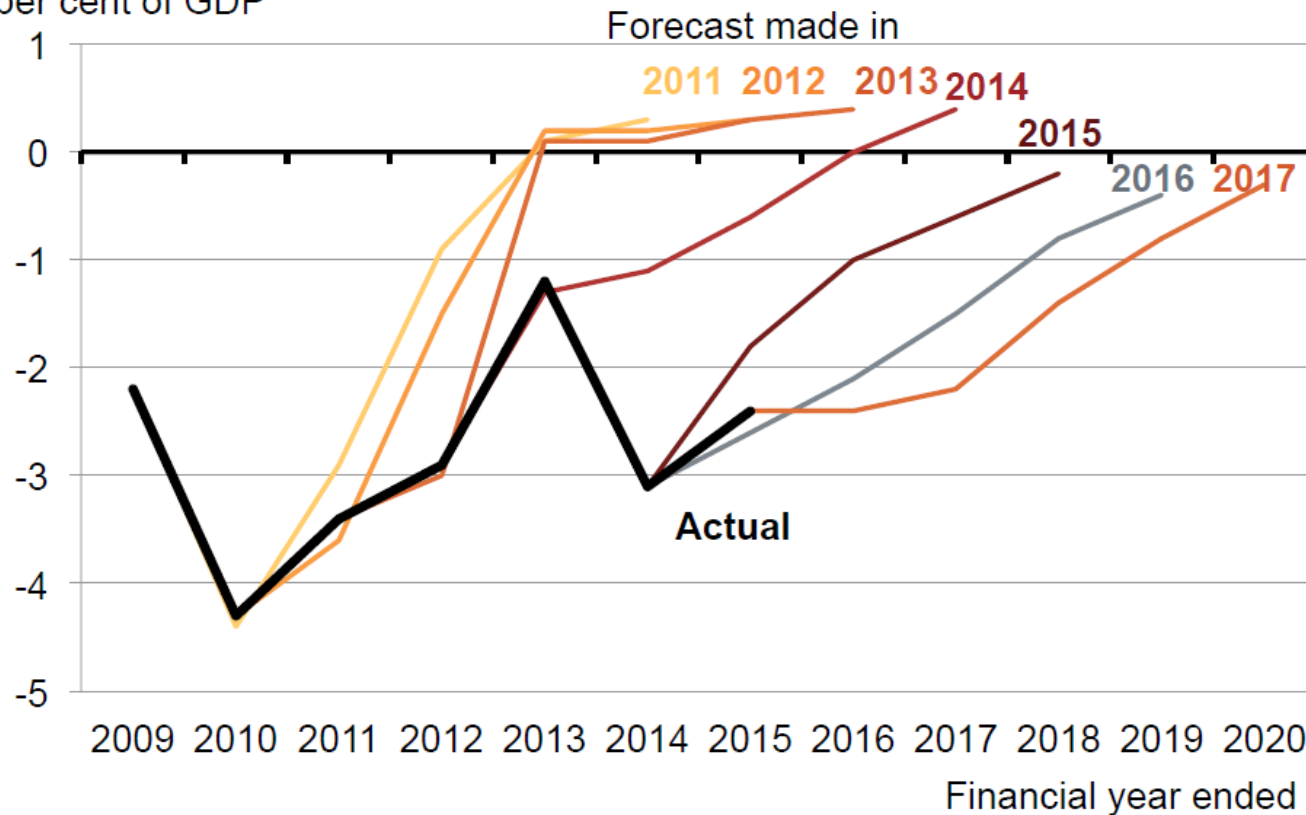
**1995**

The truth is no online database will replace your daily newspaper...

Clifford Stoll, Newsweek article entitled *The Internet? Bah!*

# Forecasting is difficult



Commonwealth plans to drift back to surplus show the triumph of experience over hope

# Introduction to Forecasting

▸ What is the purpose of forecasting?

 ▸ Predictions of forecasting are rarely precise, we can only minimize inevitable errors.

 ▸ So the purpose of forecasting is to reduce the range of uncertainty within which management judgments must be made.

▸ This purpose suggests two primary rules to which the forecasting process must adhere:

 ▸ Rule 1: The forecast must be technically correct and produce forecasts accurate enough to meet the firm's need.

 ▸ Rule 2: The forecasting procedure and its results must be effectively presented to management so that the forecasts are used in the decision making process to the firm's advantage; results must also be justified on a cost-benefit basis.

 ▸ Remark: Forecasters often pay particular attention to the first rule and expend less effort on the second.

# Introduction to Forecasting

▸ What Types of Forecasts are available?

1. Long-term or short-term forecasts.
   ▸ In this case forecasts can be classified in terms of their timescale on long-term or short-term forecasts.
   ▸ Long-term forecasts are used by top management are necessary to set the general course of an organization for the long.
   ▸ Short-term forecasts which are used by middle and first-line management to design immediate strategies to meet the needs of the immediate future.
   ▸ Forecasting method that is used to forecast sales next month would probably be NOT an inappropriate method to forecast sales in five years.

| Term | Timescale | Type of decision | Example |
|------|-----------|------------------|---------|
| Short | Up to 3-6 months | Operating | Inventory control, Production planning, and distribution |
| Medium | 3-6 months to 2 yrs | Tactical | Leasing of plant or equipment |
| Long | Above 2 yrs | Strategic | R & D, acquisitions, mergers, and product changes |

# Introduction to Forecasting

II. **A micro or macro forecasts**

- In this case forecasts can be classified in term of their position on micro-macro continuum.

- Different levels of management in an organization tend to focus on different levels of micro-macro continuum.

- The basic reason for the above classification is that different forecasting methods apply in each situation.

- A forecasting method that is used to forecast number of workers needed for the next several months for a plant manager (micro) would probably be an inappropriate method to forecast the total number of people employed in the entire country for the government (macro).

# Introduction to Forecasting

III. Quantitative or qualitative forecasts.

- In this case forecasts can be classified according to whether forecasts are from no formal mathematical model (because of lack of data) versus forecasts are of formal mathematical model.

- A forecasting method based on forecaster judgment as a result of the mental manipulation of past historical data is a qualitative forecast.

- A forecasting method using a mechanical procedures based on the manipulation of numerical data such as business data is a quantitative forecast.

IV. Nature of Output

- Point Forecast: A single number best guess for the future value

- Interval Forecast: A range of numbers within which the future value will fall

- Density Forecast: Entire probability distribution for the future value

# Predictive Analytics Steps

1. Problem formulation and data collection

   ‣ It is a process of understanding the problem to determine the appropriate collection of data and considering the related methodology (qualitative or quantitative) in view of the availability and the correctness of the existed data.

2. Data manipulation and cleaning

   ‣ It is the process of checking the data whether it is small or BIG and cleaning data from missing observations using appropriate estimation techniques.

3. Model building and evaluation

   ‣ It is the process of fitting the collected data into an appropriate prediction model with minimum errors to be used for predicting the future.

# Predictive Analytics Steps

4. Model implementation (actual prediction)

- It is the process of generating the actual model forecasts using the appropriate prediction model that have been chosen.

5. Prediction evaluation

- It is the process of evaluating the accuracy of the model predictions and the adequacy of the fitted model using different methods defined on the prediction errors.