

# A Short Course on Time Series Forecasting

by

Dr. Tanujit Chakraborty

Ph.D from Indian Statistical Institute, Kolkata, India  
IIIT Bangalore and Sorbonne University



## 2021 *This Is What Happens In An Internet Minute*



# The World is Data Rich

Astronomy



Social Networks



Healthcare



Banking



Genomics



Weather measurements



## Recent Trends and Buzzwords

- **Statistics** is the study of the collection, analysis, interpretation, presentation and organization of data. Statistics as a key Technology has its applications ranging from natural science, technology to social sciences and human welfare.
- **Data science** is the study of the generalizable extraction of knowledge from data, yet the key word is science.
- **Machine learning** gives computers the ability to learn without being explicitly programmed.
- **Artificial Intelligence** is the study of intelligent agents: any device that perceives its environment and takes actions that maximize its chance of success at some goal.
- **Forecasting** is estimating how the sequence of observations will continue into the future. Whether it is the rise/fall in exchange rates, the outcome of elections, or winners at the Oscars, there is sure to be something you want to know.

## Reputations can be made and lost by making Forecast

"I think there is a world market for maybe five computers."

(Chairman of IBM, 1943)

"There is no reason anyone would want a computer in their home."

(President, DEC, 1977)

"There's no chance that the iPhone is going to get any significant market share. No chance."

(Steve Ballmer, CEO Microsoft, April 2007)

"We're going to be opening relatively soon ... The virus ... will go away in April."

(Donald Trump, February 2020)

*"Prediction is very difficult, especially if it's about the future!"*

*- Niels Bohr, Danish Physicist & Nobel laureate in Physics.*

- Forecasting has fascinated people for thousands of years, sometimes being considered a sign of divine inspiration, and sometimes being seen as a criminal activity.
- The Jewish prophet Isaiah wrote in about 700 BC "Tell us what the future holds, so we may know that you are gods." (Isaiah 41:23)
- One hundred years later, in ancient Babylon, forecasters would foretell the future based on the distribution of maggots in a rotten sheep's liver.



Forecasting by maggots: Clay model of sheep's liver, stored in British Museum.

- Beginning in the 800 BC, a priestess known as the Pythia would answer questions about the future at the Temple of Apollo on Greece's Mount Parnassus.
- It is said that she, the Oracle of Delphi, dispensed her wisdom in a trance – caused, some believe, by the **hallucinogenic gases** that would seep up through natural vents in the rock.



Forecasting by hallucination

# Forecasters are to blame!

Forecasting. Prediction.  
Divination.

- Forecasters had a tougher time under the emperor Constantius, who issued a decree in AD357 forbidding anyone “to consult a soothsayer, a mathematician, or a forecaster – May curiosity to foretell the future be silenced forever.”
- A similar ban on forecasting occurred in England in 1736 when it became an offence to defraud by charging money for predictions. The British Vagrancy Act (1736) made it an offence to defraud by charging money for predictions. The punishment was three months' imprisonment with hard labour!



**Vagrant forecasters**



## Forecasters are to blame!

- **News report on 16 August 2006:**  
A Russian woman is suing weather forecasters for wrecking her holiday. A court in Uljanovsk heard that Alyona Gabitova had been promised 28 degrees and sunshine when she planned a camping trip to a local nature reserve, newspaper Nowyje Iswestija said.
- But it did nothing but pour with rain the whole time, leaving her with a cold. Gabitova has asked the court to order the weather service to pay the cost of her travel.

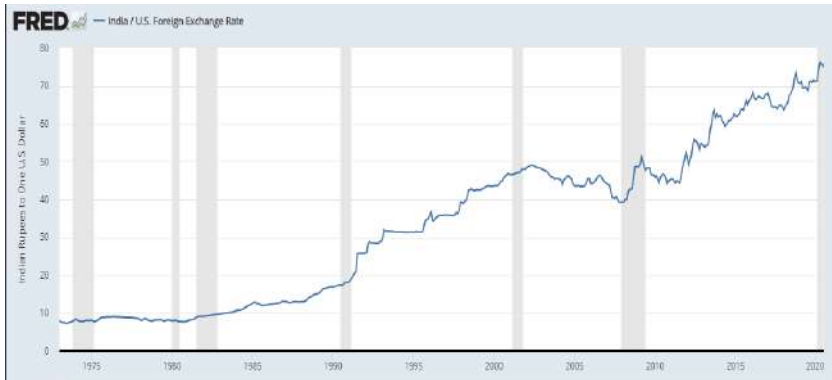


**Weather Rock**

- “There are four ways economists can lose their reputation. Gambling is the quickest, sex is the most pleasurable and drink the slowest. But forecasting is the surest.” (Max Walsh, Australian economic and political commentator, The Age, 1993)
- “Precautions should be taken against running into unforeseen occurrences or events.” (Horoscope, New York Times)

# A Sample Forecasting Problem: India / U.S. Foreign Exchange Rate

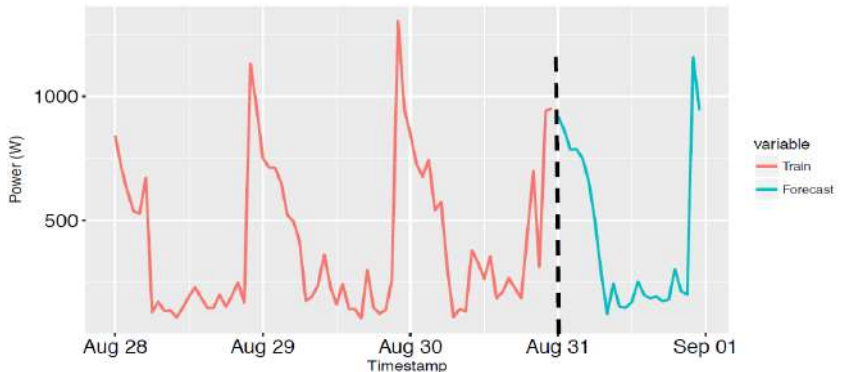
- Source : **FRED ECONOMICS DATA** (Shaded-areas indicate US recessions)
- Units : **Indian Rupees to One U.S. Dollar, Not Seasonally Adjusted**
- Frequency : **Monthly (Averages of daily figures)**



- **Time series** is a set of observations, each one being recorded at a specific time. (e.g., Forecasting of major economic variables like Annual GDP of a country, Unemployment, Inflation, Exchange rates, Production and Consumption).
- **Discrete time series** is one in which the set of time points at which observations are made is a discrete set. (e.g., All above including irregularly spaced data)
- **Continuous time series** are obtained when observations are made continuously over some time intervals. (e.g., ECG graph)
- **Forecast** is an estimate of the probability distribution of a variable to be observed in the future.

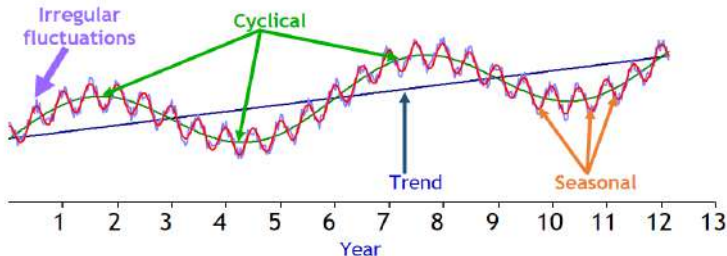
# Statistical Forecasting : Assumptions

- Time series Forecasting : Data collected at regular intervals of time (e.g., Weather and Electricity Forecasting).
- Assumptions: (a) Historical Information is available; (b) Past patterns will continue in the future.
- Statistical models, though, only work in the short term, are not very good for long-term forecasting because these assumption.



# Time Series Components

- **Trend ( $T_t$ )** : pattern exists when there is a long-term increase or decrease in the data.
- **Seasonal ( $S_t$ )** : pattern exists when a series is influenced by seasonal factors (e.g., the quarter of the year, the month, or day of the week).
- **Cyclic ( $C_t$ )** : pattern exists when data exhibit rises and falls that are not of fixed period (duration usually of at least 2 years).
- **Irregular component ( $I_t$ )**: random and unpredictable portion of the series.

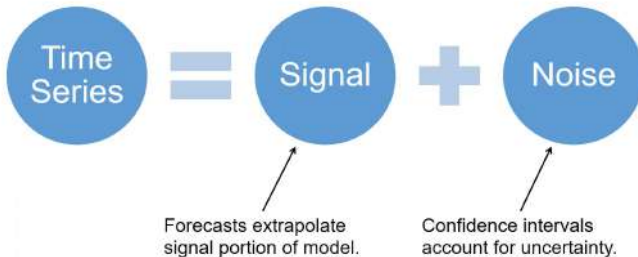


- **Additive model** is more appropriate if magnitude of seasonal fluctuations does not vary with level.
- **Multiplicative model** is more appropriate if seasonal fluctuations are proportional to level of series.
- **In Economic Series**, multiplicative decomposition is more prevalent.
- **Alternative:** Use a Box-Cox transformation and then use additive decomposition.
- **Logs turn multiplicative relationship into an additive relationship :**  
$$Y_t = T_t \times S_t \times C_t \times I_t \implies \log Y_t = \log T_t + \log S_t + \log C_t + \log I_t$$
- **A stationary series is:** roughly horizontal, constant variance and no patterns predictable in the long-term.

# Statistical Forecasting

The science of forecasting got going properly in the 1980s when people realised that if you took all of the ideas that people had developed in different fields, and you thought of it as a collection of techniques and overlaid that with analytical and scientific thinking, then forecasting itself could be considered a scientific discipline.

## Statistical Forecasting





- Exponential smoothing (Brown, 1959).
- ARIMA models (Box, Jenkins, 1970).
- VAR models (Sims, Granger, 1980).
- Non-linear models (Granger, Tong, Hamilton, Teräsvirta, 1980s).
- ARCH/GARCH (Engle, Bollerslev, 1982).
- State space models (Harvey, West, Harrison, 1989).
- Neural Networks and Recurrent Neural network (1993 - 1995).
- Long short-term memory (LSTM) (1997).

# Advantages of Statistical models

- ① Based on empirical data
- ② Objective measure of uncertainty.
- ③ Computable, Replicable, Testable.
- ④ Able to compute prediction intervals.

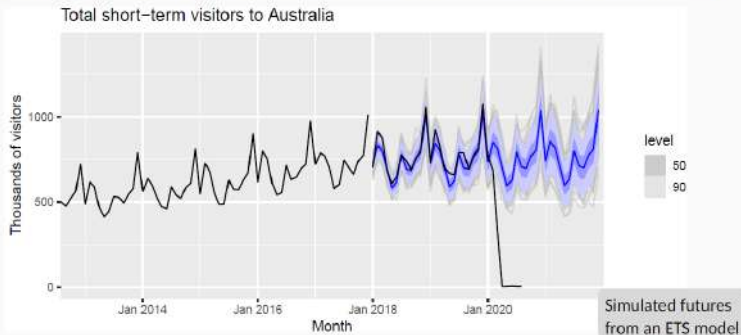
# Advantages of Machine Learning models

- ① Smart algorithms, few assumptions and applied to huge data sets.
- ② Solve problems which traditional statistical methods can't handle (largely due to size of data sets).
- ③ Strong emphasis on out-of-sample predictive performance (the test data).
- ④ Recognition that many problems are about prediction not p-values.

- **Models for annual data:** Exponential Smoothing State Space (ETS), ARIMA, Dynamic regression, ARNN and LSTM model.
- **Models for quarterly, monthly data:** ETS, ARIMA/SARIMA, Dynamic regression, Seasonal and Trend decomposition using Loess (STL), ARNN and LSTM model.
- **Models for weekly, daily, hourly data:** ARIMA/SARIMA, Dynamic regression, STL.

# Random futures

A forecast is an estimate of the probability distribution of a variable to be observed in the future.



Mathematical/Statistical/ML models are simplifications of reality – and life is sometimes too complex to model accurately.

## Exercise: Which is easiest to forecast?

- Exchange rate of USD/INR next week.
- Daily electricity demand in 3 days time.
- Time of sunrise this day next year.
- Maximum temperature tomorrow.
- Google stock price tomorrow.

## Which is easiest to forecast? (Easy to Tough)

- ① Time of sunrise this day next year.
- ② Maximum temperature tomorrow.
- ③ Daily electricity demand in 3 days time.
- ④ Google stock price tomorrow.
- ⑤ Exchange rate of USD/INR next week.

**How do we measure “easiest”?**

**What makes something easy/difficult to forecast?**

Something is **easier to forecast** if:

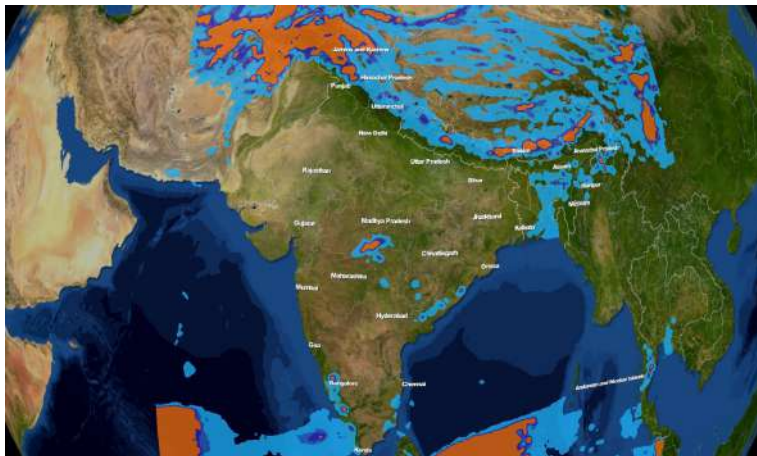
- We have a good understanding of the **factors** that contribute to it, and can measure them (for stock price and exchange rates causes are mostly unknown).
- There is lots of **data available**.
- The future is somewhat **similar to the past**.
- **The forecasts cannot affect the thing we are trying to forecast** (say, Warren Buffett, CEO of Berkshire Hathaway, make some comment that stock price may change!).
- **When should we give up?** When there is insufficient data? When the models give implausible forecasts?



# What can we forecast? - Weather

Publicly available Weather statistics Data of Norwegian Centre for Climate Services

<https://seklima.met.no/>



# What can we forecast? - Stock Price

Publicly available stock price data: <https://finance.yahoo.com/>



## What can we forecast? - Sales

Time series models are extremely useful in pharmaceutical sales forecasting and in predicting demand for items.



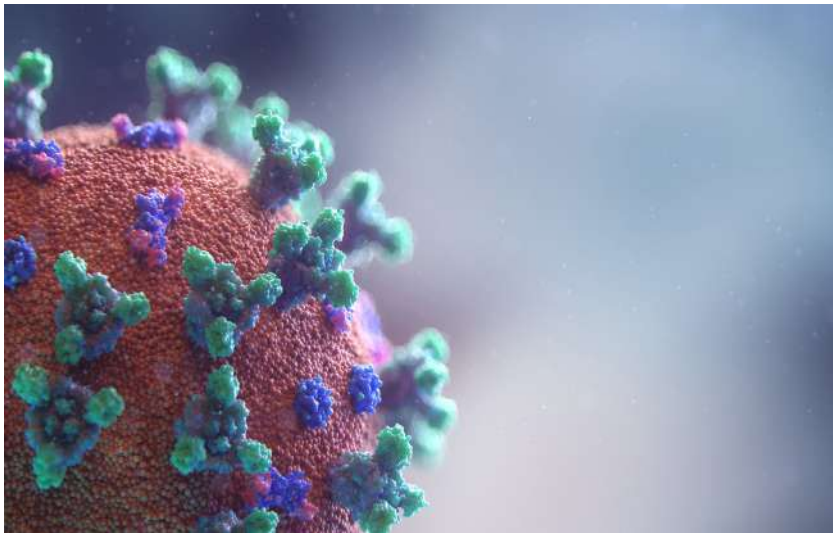
# What can we forecast? - Electricity

Models for electricity demand were developed so that they could plan generation capacity.



# What can we forecast? - COVID-19

Publicly available data at <https://ourworldindata.org/>



# Business Forecasting — Case Study - 1

**Client:** Car fleet companies around the world.



PROBLEM: HOW TO FORECAST RESALE VALUE OF VEHICLES?  
HOW SHOULD THIS AFFECT LEASING AND SALES POLICIES?

### Additional information:

- They can provide a large amount of data on previous vehicles and their eventual resale values.
- The resale values are currently estimated by a group of specialists. Statistical/ML-based forecasters can help the Car fleet companies.

## Business Forecasting — Case Study - 2

Client: Airline Company.





PROBLEM: HOW TO FORECAST PASSENGER TRAFFIC ON MAJOR ROUTES (SAY, ECONOMY CLASS PASSENGERS)?

Additional information:

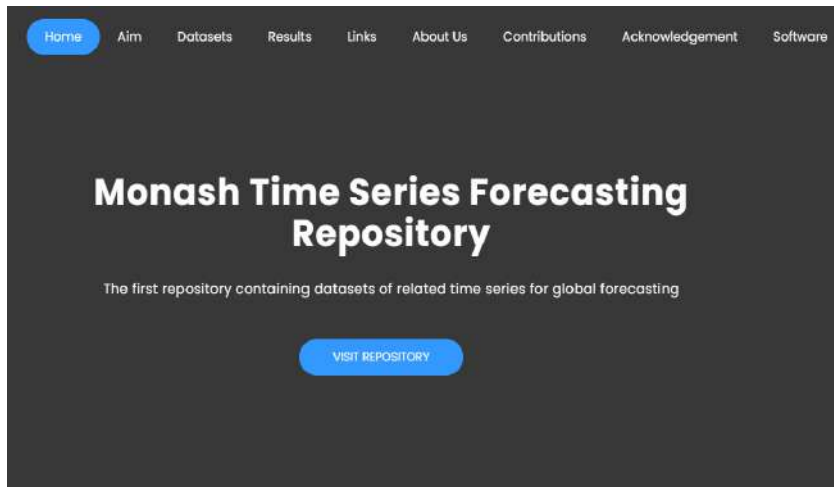
- They can provide a large amount of data on previous routes.
- Traffic is affected by school holidays, special events such as the FIFA World Cup, advertising campaigns, competition behaviour, etc.
- They have a highly capable team of people who are able to do most of the computing.

# Forecasting for Social Good

Our Epidemic Forecasting Library: <https://github.com/mad-stat/epicasting>.



Monash Time Series Data Repo: <https://forecastingdata.org/>.



## A popular R Package: <https://github.com/robjhyndman/forecast>.

`forecast`: Forecasting Functions for Time Series and Linear Models

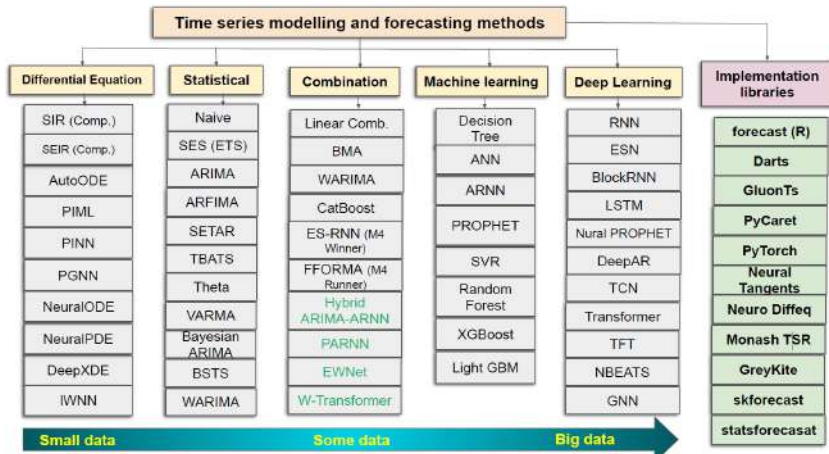
Methods and tools for displaying and analysing univariate time series forecasts including exponential smoothing via state space models and automatic ARIMA modelling.

Version: 8.17.0  
Depends: R (≥ 3.0.2)  
Imports: [colorspace](#), [fractdiff](#), [generics](#) (≥ 0.1.2), [ggplot2](#) (≥ 2.2.1), [graphics](#), [lmtest](#), [magrittr](#), [nnet](#), [parallel](#), [Rcpp](#) (≥ 0.11.0), [stats](#), [timeDate](#), [tseries](#), [urca](#), [zoo](#)  
LinkingTo: [Rcpp](#) (≥ 0.11.0), [RcppArmadillo](#) (≥ 0.2.35)  
Suggests: [forecTheta](#), [knitr](#), [methods](#), [markdown](#), [rtables](#), [seasonal](#), [testthat](#), [uroot](#)  
Published: 2022-07-25  
Author: Rob Hyndman  [aut, cre, cph], George Athanasopoulos  [aut], Christoph Bergmeir  [aut], Gabriel Caceres  [aut], Leanne Chhay [aut], Kirill Kuroptev [aut], Mitchell O'Hara-Wild  [aut], Fotios Petropoulos  [aut], Slava Razbash [aut], Earo Wang  [aut], Farah Yasmeen  [aut], Daniel Reid [ctb], David Shaub [ctb], Federico Garza [ctb], R Core Team [ctb, cph], Ross Ihaka [ctb, cph], Xiaoqian Wang [ctb], Yuan Tang  [ctb], Zhenyu Zhou [ctb]  
Maintainer: Rob Hyndman <Rob.Hyndman@monash.edu>  
BugReports: <https://github.com/robjhyndman/forecast/issues>  
License: [GPL-3](#)  
URL: <https://pkg.robjhyndman.com/forecast/>, <https://github.com/robjhyndman/forecast>

## Recent Forecasting Toolkit (Mostly in Python)

- Prophet by Meta (2018, American Statistician): [Link](#)
- Darts by Unit8 (2018): [Link](#)
- GluonTs by awslabs (2019, JMLR): [Link](#)
- tslearn Toolkit (2020, JMLR): [Link](#)
- Pytorch Forecasting by ADIPytorch (2018): [Link](#)
- STUMPY (2016, ICDM): [Link](#)
- Merlion by Salesforce (2021): [Link](#)
- NBEATS by Element AI (2020 ICLR): [Link](#)
- Kats by Meta (2021): [Link](#)
- GrayKite by LinkedIn (2021): [Link](#)

# List of Various Forecasting Models



The most Applied Forecasting Book for Students: <https://otexts.com/fpp3/>.

- What about interpretable neural networks and deep learning?
- Will we ever be able to forecast “black swans”?
- Does more data mean better forecasts?
- Are there many jobs for forecasters?
- Urgent need for a course on “Applied Forecasting” for all the students, if not any.
- Various resources are available at [www.forecasters.org](http://www.forecasters.org)

# HAPPY FORECASTING

“A good forecaster is not smarter than everyone else, he merely has his ignorance better organised.”

