Lecture 1
An Overview of Time Series Analysis
Readings: CC08 Chapter 1; SS17 Chapter 1; BD16 Chapter 1.1 - 1.3

MATH 8090 Time Series Analysis
Week 1

Whitney Huang
Clemson University

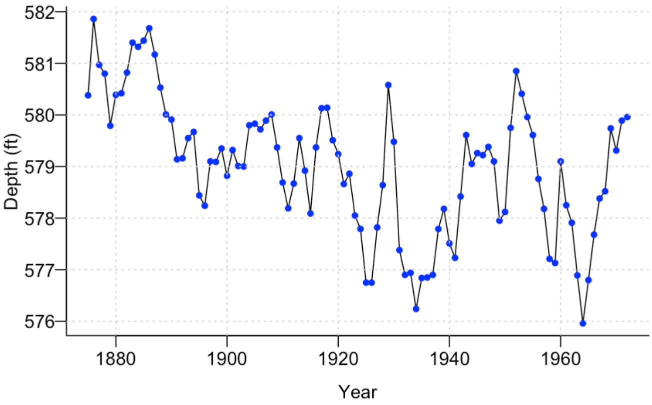
Agenda

1. **Time Series Data**
2. **Time Series Models**
3. **Objectives of Time Series Analysis**

Level of Lake Huron 1875–1972

Annual measurements of the level of Lake Huron in feet.
[Source: Brockwell & Davis, 1991]

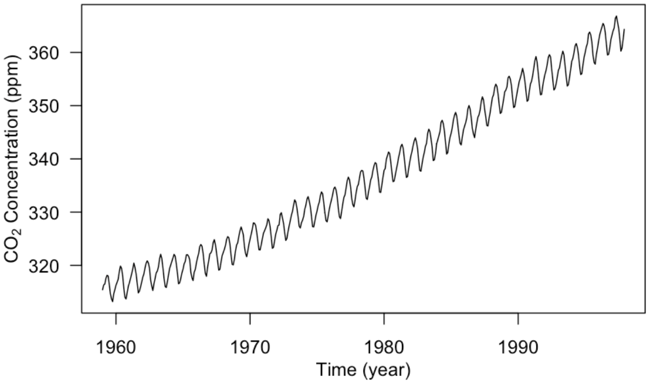
```{r}
par(mar = c(3.2, 3.2, 0.5, 0.5), mgp = c(2, 0.5, 0), bty = "L")
data(LakeHuron)
plot(LakeHuron, ylab = "Depth (ft)", xlab = "Year", las = 1)
points(LakeHuron, cex = 0.8, col = "blue", pch = 16)
grid()
```



Mauna Loa Atmospheric CO_2 Concentration

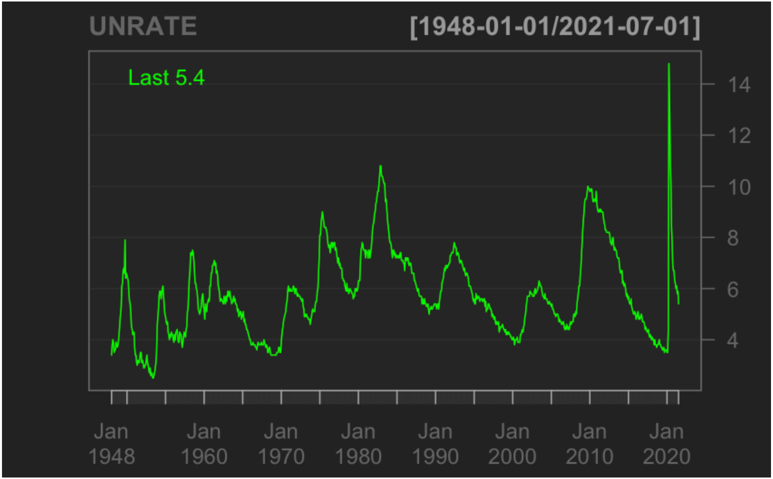
Monthly atmospheric concentrations of CO_2 at the Mauna Loa Observatory [Source: Keeling & Whorf, Scripps Institution of Oceanography]

```{r}
data(co2)
par(mar = c(3.8, 4, 0.8, 0.6))
plot(co2, las = 1, xlab = "", ylab = "")
mtext("Time (year)", side = 1, line = 2)
mtext(expression(paste("CO"[2], " Concentration (ppm)")), side = 2, line = 2.5)
```



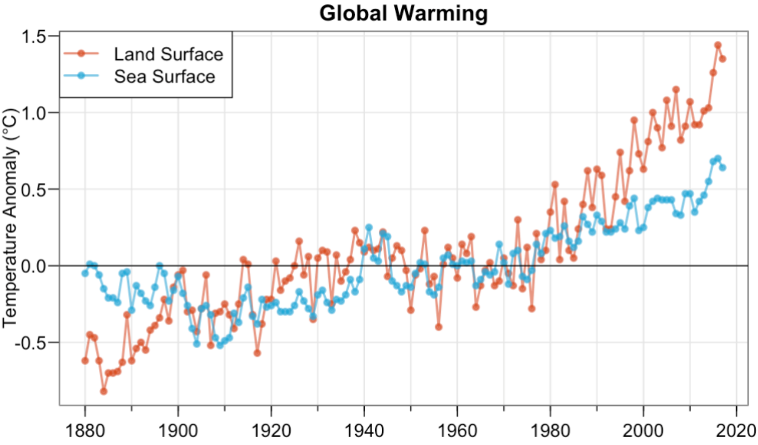
US Unemployment Rate 1948 Jan. – 2021 July

[Source: St. Louis Federal Reserve Bank's FRED system]



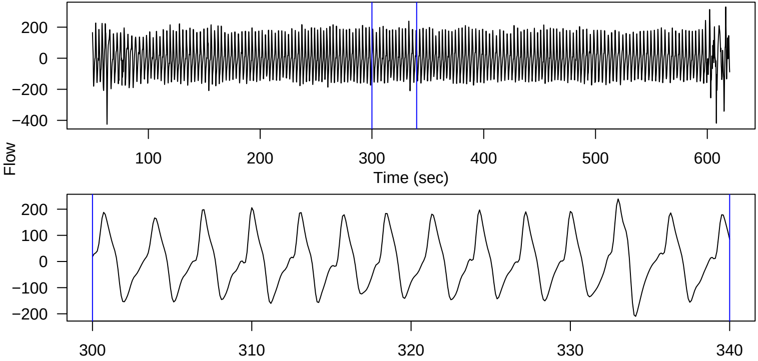
Global Annual Temperature Anomalies

[Source: NASA GISS Surface Temperature Analysis]



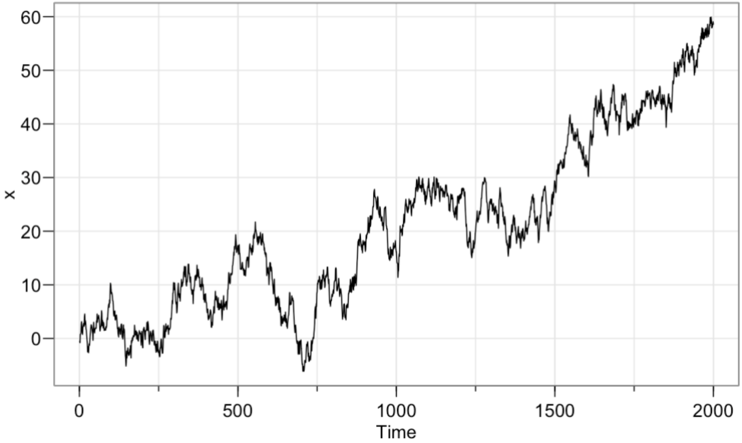
Sleep Airflow Signal

A “normal” patient’s 10 Hz sleep airflow signal [Source: H. et al. 2022]



A Simulated Time Series

```{r}
set.seed(123)
w <- rnorm(2000); x <- cumsum(w); tsplot(x, las = 1)
```



Time Series Data

- A time series is a collection of observations $\{y_t, t \in T\}$ taken sequentially in time (t) with the index set T
  - $T = \{0, 1, 2, \cdots, T\} \subset \mathbb{Z} \Rightarrow$ discrete-time time series
  - $T = [0, T] \subset \mathbb{R} \Rightarrow$ continuous-time time series
- A discrete-time time series might be intrinsically discrete or might arise from a underlying continuous-time time series via
  - sampling (e.g., instantaneous wind speed)
  - aggregation (e.g., daily accumulated precipitation amount)
  - extrema (e.g., daily maximum temperature)
- We will focus on dealing with discrete-time real-valued $(Y_t \in \mathbb{R})$ time series in this course

Exploratory Time Series Analysis

- Start with a time series plot, i.e., to plot y_t versus t
- Look at the following:
  - Are there abrupt changes?
  - Are there “outliers”?
  - Is there a need to transform the data?
- Examine the trend, seasonal components, and the “noise” term

Features of Times Series

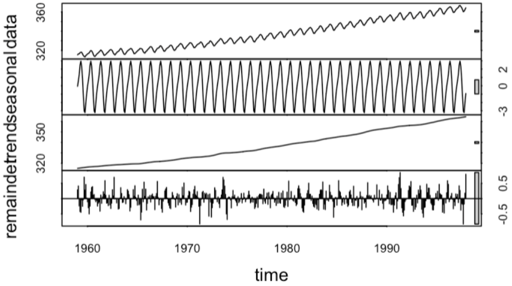
- Trends
  - One can think of trend, μ_t, as continuous changes, usually in the mean, over longer time scales ⇒ *"the essential idea of trend is that it shall be smooth"* - [Kendall, 1973]
  - Usually the form of the trend is unknown and needs to be estimated. When the trend is removed, we obtain a detrended series
- Seasonal or periodic components
  - A seasonal component s_t constantly repeats itself in time, i.e., $s_t = s_{t+kd}$
  - We need to estimate the form and/or the period d of the seasonal component to deseasonalize the series
- The "noise" process
  - The noise process, η_t, is the component that is neither trend nor seasonality
  - We will focus on finding plausible (typically stationary) statistical models for this process

Combining Trend, Seasonality, and Noise Together

There are two commonly used approaches

- Additive model:

$$y_t = \mu_t + s_t + \eta_t, \quad t = 1, \cdots, T$$



- Multiplicative model:

$$y_t = \mu_t s_t \eta_t, \quad t = 1, \cdots, T$$

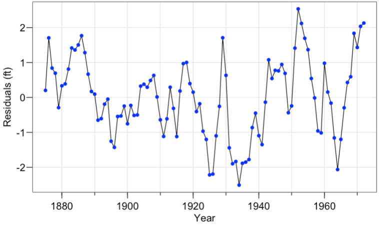
If all $\{y_t\}$ are positive then we obtain the additive model by taking logarithms:

$$\log y_t = \log \mu_t + \log s_t + \log \eta_t, \quad t = 1, \cdots, T$$

Time Series Models

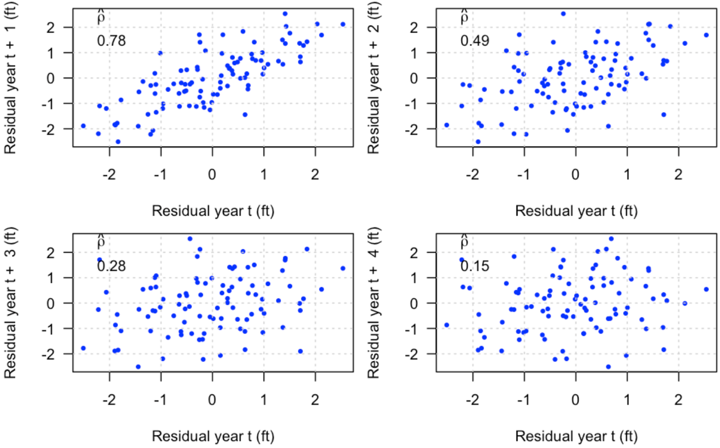
Lake Huron Time Series

- Time series analysis is the area of statistics which deals with the analysis of dependency between different observations (typically $\{\eta_t\}$)
- Some key features of the Lake Huron time series:
  - decreasing trend
  - some "random" fluctuations around the decreasing trend
- We extract the "noise" component by assuming a linear trend



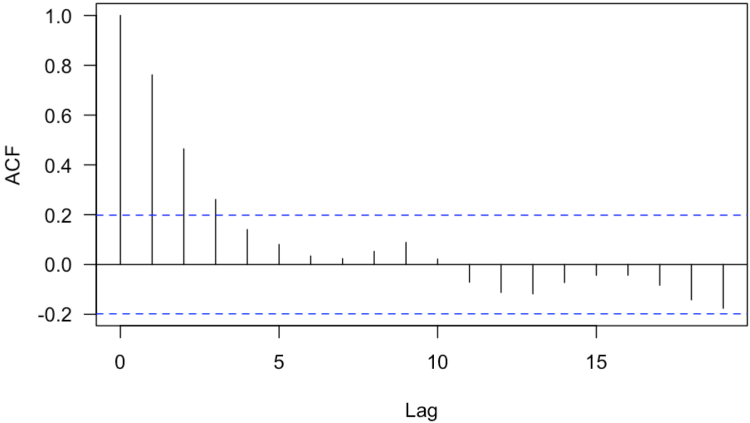
Exploring the Temporal Dependence Structure of $\{\eta_t\}$

$\{\eta_t\}$ exhibit a temporal dependence structure, meaning that the nearby (in time) values tend to be more alike than those that are far part. To observe this, let's create a few time lag plots



Further Exploration of the Temporal Dependence Structure

Let's plot the correlation as a function of the time lag



In a few weeks we will learn how to use this information to suggest an appropriate model

Time Series Models

- A time series model is a probabilistic model that describes ways that the series data $\{y_t\}$ could have been generated
- More specifically, a time series model is usually a probability model for $\{Y_t : t \in T\}$, a collection of random variables indexed in time
- We will try to keep our models for $\{Y_t\}$ as simple as possible by assuming stationarity $\Rightarrow$ some characteristic of the distribution of $\{Y_t\}$ does not depend on the time points, only on the "time lag"
- While most time series are not stationary, one either remove or model the non-stationary parts (e.g., de-trend or de-seasonalization) so that we are only left with a stationary component $\{\eta_t\}$. We typically further assume that the process is second order stationary $\Rightarrow$
  $\mathrm{E}[\eta_t] = 0, \quad \forall t \in T$ and
  $\mathrm{Cov}(\eta_t, \eta_{t'}) = \gamma(t' - t) = \mathrm{Cov}(\eta_{t+s}, \eta_{t'+s})$

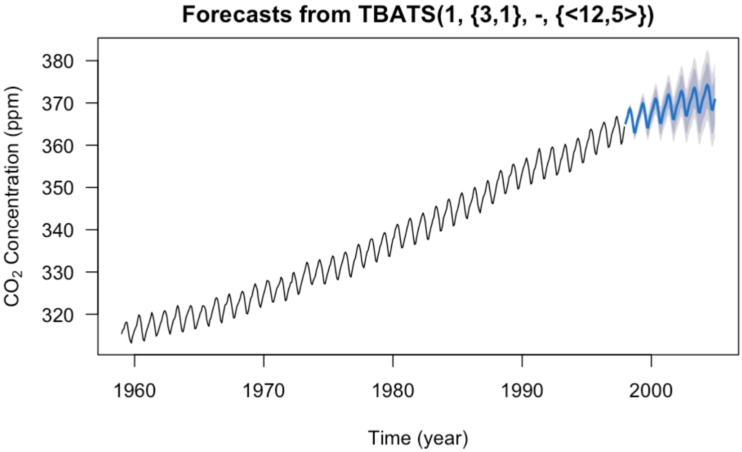
Objectives of Time Series Analysis

Some Objectives of Time Series Analysis

- **Modeling**: Find a statistical model that adequately explains the observed time series
- For example, identify a model which can account for the fact that the depths of Lake Huron are correlated with differ years and with a decreasing long-term trend
- The fitted model can be used for further statistical inference, for instant, to answer the question like: Is there evidence of decreasing trend in the Lake Huron depths?

Some Objectives of Time Series Analysis, Cont'd

Forecasting is perhaps the most common objective. One observe a time series of given length and wish to **predict** or **forecast** future values of the time series based on those already observed.



Some Objectives of Time Series Analysis, Cont'd

- **Adjustment**: an example would be seasonal adjustment, where the seasonal component is estimated and then removed to better understand the underlying trend
- **Simulation**: use a time series model (which adequately describes a physical process) as a surrogate to *simulate repeatedly in order to approximate how the physical process behaves*
- **Control**: adjust various input (control) parameters to make the time series fit more closely to a given standard (many examples from statistical quality control)