1921 UN SECOLO DI STORIA DAVANTI A NOI —— 2021

UNIVERSITÀ CATTOLICA del Sacro Cuore

Time Series Analisys and Forecasting

Professor Stefano Norcia

MGO962

Contact details

Lecturer: Professor Stefano Norcia

stefano.norcia@unicatt.it

stefanonorcia.com

@stefanonorcia

@stefanonorcia-3ab0b61a

Bio

- I’m a Senior Software Engineer with almost two decades of experience in the field, I work with Web and Mobile technologies and I’m an open source enthusiast. I also have good skills in System Administration, Security and Microservices. I’m also interested in Data Science, R and Python.
- I have almost one decade experience in teaching Information Technology. I teach training courses for different companies about web and mobile software development, security, system administration and DevOps.

Unit objectives

1. To obtain an understanding of common statistical methods used in business and economic forecasting.
2. To develop the computer skills required to forecast business and economic time series data;
3. To gain insights into the problems of implementing and operating large scale forecasting systems for use in business.

Teaching

- One 120 minute lecture on tuesday at 14:30 each week for 10 weeks.
- One 60 minute lecture on thursday at 13:30 each week for 10 weeks.
- One 60 minute computer lab session on thursday at 14:30 each week for 10 weeks.

R

R Language:

`https://cran.csiro.au/`

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS.

Widely used in education, not much in the industry because of performance issues

Main R packages

Main R packages

```
# Data manipulation and plotting functions
library(tidyverse)
# Time series manipulation
library(tsibble)
# Tidy time series data
library(tsibbledata)
# Time series graphics and statistics
library(feasts)
# Forecasting functions
library(fable)
```

Main R packages

```
# Data manipulation and plotting functions
library(tidyverse)
# Time series manipulation
library(tsibble)
# Tidy time series data
library(tsibbledata)
# Time series graphics and statistics
library(feasts)
# Forecasting functions
library(fable)
```

```
# All of the above
library(fpp3)
```

Install required packages

```
install.packages(c(
  "tidyverse",
  "fpp3"
))
```

Python

Python Language: `https://www.python.org/`

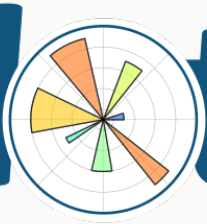
Python is an interpreted high-level general-purpose programming language. Its design philosophy emphasizes code readability with its use of significant indentation. Widely used in the industry for Data Science.

Main Python packages

pandas

statsmodels

matplotlib



NumPy

SciPy

Main Python packages

```
# Data manipulation
from pandas import read_csv, DataFrame
```

```
# Data plotting
from matplotlib import pyplot
```

```
# Forecasting models
from statsmodels.tsa.arima.model import ARIMA
```

Tools for R/Python

Overview of the most used tools for R/Python:

- RStudio
- PyCharm
- Jupyter Notebook
- Jupyter Lab
- Anaconda

RStudio

`https://www.rstudio.com`

RStudio is an Integrated Development Environment (IDE) for R, a programming language for statistical computing and graphics.

PyCharm

`https://www.jetbrains.com/pycharm/`

PyCharm is an integrated development environment (IDE) used in computer programming, specifically for the Python language.

Jupyter Web Environment

Jupyter Notebook/Lab Web Development Environment: `https://www.jupyter.org`

Jupyter Notebook (formerly IPython Notebooks) is a web-based interactive computational environment for creating notebook documents. JupyterHub is a multi-user server for Jupyter Notebooks

`https://www.jupyter.org`

Anaconda

Anaconda is a distribution of the Python and R programming languages for scientific computing, that aims to simplify package management and deployment.

- Free and Open Source
- Contains R-Studio, Jupyter Notebook, PyCharm and more

Our Development Environment : Languages

- I strongly advice to install a development environment on your personal computer
- R language
  - `https://cran.r-project.org/bin/windows/base/`
  - R-4.1.1-win.exe
- Python language
  - `https://www.python.org/downloads/release/python-397/`
  - python-3.9.7-amd64.exe

Our Development Environment : IDE

- I also advice to use an advanced IDE in order to take advantage of tools like the Debugger and Code Completion
- PyCharm Community Edition is free and supports R by adding plugins
  - `https://www.jetbrains.com/pycharm/download/#section=windows`
  - pycharm-community-2021.2.2.exe

Our Web Development Environment

- All participants will receive access to a web development environment

Jupyter Hub

jupyter.stefanonorcia.com

- You will receive credentials to access the development environment via Mail
- This is an Jupyter Hub installation on my VPS where you can practice with Python

Key reference

Hyndman, R. J. & Athanasopoulos, G. (2021)
***Forecasting: principles and practice*, 3rd edition**

OTexts.org/fpp3/

- Free and online
- Data sets in associated R packages
- R code for examples

Outline

	Topic	Chapter
1	Introduction to forecasting and R	1
2	Time series graphics	2
3	Time series decomposition	3
4	The forecaster’s toolbox	5
5	Exponential smoothing	8
6	Forecasting with ARIMA models	9
7	Multiple regression and forecasting	7
8	Dynamic regression	10

Assessment

- 3 short assignments: worth from 0 to 2 points each
- Final Exam (2 hours): 24 points

Assessment

- 3 short assignments: worth from 0 to 2 points each
- Final Exam (2 hours): 24 points

Task	Due Date	Value
Assignment 1	Sun 11:59pm week 6	2
Assignment 2	Sun 11:59pm week 8	2
Assignment 3	Sun 11:59pm week 10	2
Final exam	Official exam period	24

Blackboard site

- Includes all course materials
- Assignment submissions
- Forum for asking questions, etc.