

# Statistical Methods for Data Science

## Semester 1 2021

### DATA7202

Slava Vaisman

The University of Queensland

*[r.vaisman@uq.edu.au](mailto:r.vaisman@uq.edu.au)*

# Overview (1)

## Course objectives

- ① Develop understanding of and practical proficiency with a variety of modern statistical techniques.
- ② Develop a variety of skills with statistical software (R, Python, and Matlab).
- ③ Develop ability to apply dynamic simulation to real-life problems.
- ④ Develop ability to clearly communicate the results from statistical analyses of data and the decisions made along the way.

# Overview (2)

## Assumed Background

- ➊ MATH7501 and MATH7502 or equivalent: multivariate calculus, matrix and vector algebra, basic optimization and approximations.
- ➋ STAT7203 or equivalent: probability to multivariate transformations, basic statistics, hypothesis testing, confidence intervals and regression
- ➌ Programming ability (e.g. CSSE7030) and familiarity with R and Python (e.g. DATA7001).

## Recommended resource

- D.P. Kroese, Z.I. Botev, T. Taimre, and R. Vaisman. Data Science and Machine Learning: Mathematical and Statistical Methods, Chapman and Hall/CRC, 2019  
<https://people.smp.uq.edu.au/DirkKroese/DSML/>
- and many more (please see the ECP).

# Overview (3)

**It is very important to follow the lectures and attend the practicals!**

## Assessment items

- ① Assignment 1 (22 Mar 21 17:00)
- ② Assignment 2 (19 Apr 21 17:00)
- ③ Assignment 3 (17 May 21 17:00)
- ④ Assignment 4 (14 Jun 21 17:00)

# Overview (4)

The following topics will be covered.

- Understanding statistical inference.
- Stochastic modeling.
- Statistical/ML models.
- Dynamic simulation.