

Course

Data-analysis and statistics

Date:	7 October, 21 October, 11 November 2015
Time:	10.00 – 17.00 h.
Location:	Utrecht, Academieggebouw, Belle van Zuylenzaal, Domplein 29
Course leader:	Dr. Maarten Kroesen and Dr. Eric Molin
Days:	3
ECTS:	1.5 (class) - 1.5 (assignment)
TUD GS credits:	3 (class) - 4 (assignment)
Course fee:	Free for TRAIL/Beta/OML members, others please contact the TRAIL office
Registration:	www.rstrail.nl

Objectives

After this course attendees are able to:

1. Recognize the probability density functions of various discrete (Bernoulli, Binomial, Geometric, Poisson) and continuous random variables (Uniform, Exponential, Normal).
2. Explain the basic principles behind statistical modelling (law of large numbers, Central Limit Theorem).
3. Choose appropriate bivariate data-analysis techniques* (given a particular research question), correctly apply these techniques (by formulating statistical hypotheses, checking the statistical assumptions and deriving the test statistic) and interpret their results in meaningful ways.
4. Estimate a multivariate regression model, check its assumptions (normality, linearity, homoscedasticity) and interpret its outcomes.

* The following data-analysis techniques will be treated: descriptive data analysis (mean, median, variance, standard deviation), univariate (one sample t-test, proportion test) and bivariate parametric tests (paired/independent samples t-test, ANOVA, Pearson correlation) and a non-parametric test (chi-square).

Course description

In this course attendees will actively work on solving concrete statistical problems in the domains of transport, infrastructure and logistics using various bivariate and multivariate data-analysis techniques. The course will extensively treat the basic principles behind statistical modelling so that attendees really understand what the results of statistical tests mean.

Assignment

Attendees have to apply several data-analysis techniques to their own data (or a given dataset) and report the results in a brief research report.

Program

- Day 1 Probability density functions, basic principles behind statistical modelling, descriptive statistics and bivariate data-analysis techniques (Maarten Kroesen)
- Day 2 Continuation of bivariate data-analysis techniques (Maarten Kroesen)
- Day 3 Multiple regression (Eric Molin)

Course material

Slides and online materials

Prerequisite

None