# ST237-10 Visualisation and Communication of Data

#### 23/24

Department Statistics Level Undergraduate Level 2 Module leader Martyn Plummer Credit value 10 Module duration 10 weeks Assessment 100% coursework Study location University of Warwick main campus, Coventry

## Description

## Introductory description

This module introduces students to contemporary practice of data visualisation. The module takes a hand-on approach where concepts are introduced using the grammar of graphics (ggplot2) framework implemented in R.

This module is offered as an optional module to Statistics students and as an unusual option to students from other departments, space permitting. Student taking this module from other departments **must** have taken ST120 Introduction to Probability and ST121 Statistics Laboratory.

Pre-registration is required. Please see the module page for details.

Module web page

#### **Module aims**

The module aims to develop knowledge of

- 1. Visualisation concepts, tasks, and workflows
- 2. Critical appraisal of visualisations produced by third parties.

- 3. Programming for visualisations in R
- 4. Presentation of visualisations to a general audience.

# **Outline syllabus**

This is an indicative module outline only to give an indication of the sort of topics that may be covered. Actual sessions held may differ.

- 1. Data visualisation in the pre-informatics age. Historical examples of good pen-and-ink graphics.
- 2. Tufte's principles of good graphical design: graphical excellence; graphical integrity; avoiding chart junk; data-ink maximisation.
- 3. File formats for graphics: vector and and bitmap graphics; portability; embedding graphics in documents.
- 4. Wilkinson's grammar of graphics. A pattern language for visualizations that allows us to build visualisations from reusable components.
- 5. Graphics in R using ggplot2, an implementation of Wilkinson's grammar of graphics in R.
- 6. Colour theory. Colour spaces. Classification of colour palettes. Accessibility for colour blind readers.
- 7. Interactive graphics using Shiny.

#### Learning outcomes

By the end of the module, students should be able to:

- · Critically appraise data visualisations created by third parties
- · Create their own improved versions of flawed visualisations in R
- Choose appropriate visualisations for different tasks
- Create novel visualisations in R from a given data set and a description of the required output

## Indicative reading list

Wickham, H (2009) Ggplot2: elegant graphics for data analysis

C.O. Wilke (2019) Fundamentals of data visualization: a primer on making informative and compelling essential

Nussbaumer Knaflic. C (2019) Storytelling with Data

Tufte, ER (2001) The visual display of quantitative information

Wilkinson, L (200) The grammar of graphics

Cleveland, WS (1993) Visualizing data

View reading list on Talis Aspire

# Interdisciplinary

This module requires students to develop programming and data-analytic skills for solving realworld visualisation problems from a wide variety of disciplines.

# Subject specific skills

- 1. Demonstrate facility with data processing and visualisation methods in R
- 2. Create readable, valid, reliable, and modular code for visualisation
- 3. Analyze problems, abstracting their essential information formulating them using appropriate programming

concepts to facilitate their solution.

4. Demonstrate programming skills and knowledge of fundamental programming concepts, both explicitly and by

applying them to the solution of real-world problems

# Transferable skills

- 1. Problem solving skills: The module requires students to solve problems presenting their conclusions as logical and coherent arguments.
- 2. Written communication skills: Students complete written assessments that require precise and unambiguous communication in the manner and style expected in mathematical sciences.
- 3. Verbal communication skills: Students are encouraged to discuss and debate formative assessment and lecture material within small-group tutorials sessions. Students can continually discuss specific aspects of the module with the module leader. This is facilitated by statistics staff office hours.
- 4. Team working and working effectively with others: Students are encouraged to discuss and debate formative assessment and lecture material within small-group tutorials sessions.
- 5. Professionalism: Students work autonomously by developing and sustaining effective approaches to learning, including time-management, organisation, flexibility, creativity, collaboratively and intellectual integrity.

# Study

# Study time

Туре

Lectures Tutorials Private study Total

#### Required

20 sessions of 1 hour (20%) 8 sessions of 1 hour (8%) 47 hours (47%) 100 hours

Туре		
Assessment		
Total		

#### Required 25 hours (25%) 100 hours

## Private study description

Weekly revision of lecture notes and materials, wider reading and practice/programming exercises, working on assessed coursework.

## Costs

No further costs have been identified for this module.

#### Assessment

You do not need to pass all assessment components to pass the module.

#### Assessment group A

Weighting	Study time	Eligible for self-	
		certification	

Assessment component

Assignment 1 20% 5 hours No

Write a report on a set of visualisations created by third parties. A collection of visualisations from contemporary media will be presented and students will be requested to write a critical appraisal of each one. Students should be able to assess the aesthetics, adequacy of the visualisation to effectively communicate the intended message, and criticise potentially misleading aspects. They will also be required to suggest ways that the visualisations will be improved but will not need to implement these suggestions in R code.

Reassessment component is the same

Assessment component

Assignment 2

30%

7 hours 30 minutes No

Write a report containing improved versions of a given set of visualisations and a narrative explaining the design choices made in improving them. As in assignment 1, a collection of visualisations from contemporary media will be presented and students will be requested to critically appraise them. They will also be required to produce an improved set of visualisations,

#### Weighting

Study time

Eligible for selfcertification

provide well-documented R code to reproduce the output, and explain their choices in a report.

Reassessment component is the same

Assessment component

Assignment 3 50% 12 hours 30 minutes No Create a set of novel visualisations given data sets and a briefing on the required message to be conveyed by the visualisations. Unlike previous exercises there will be no visual cues provided in the form of existing visualisations, only a text description. Students will create a written report containing the visualisations, an explanation of their design and aesthetic choices, and welldocumented R code to reproduce the output.

Reassessment component is the same

#### Feedback on assessment

Individual feedback will be provided on coursework by class tutors.

Cohort level feedback will provided after each assessment.

# Availability

#### Courses

This module is Option list A for:

- Year 2 of USTA-G302 Undergraduate Data Science
- Year 2 of USTA-GG14 Undergraduate Mathematics and Statistics (BSc)
- Year 2 of USTA-Y602 Undergraduate Mathematics, Operational Research, Statistics and Economics