The Hong Kong Polytechnic University

Subject Description Form

Subject Code	AMA2203
Subject Title	Introduction to R
Credit Value	1
Level	2
Pre-requisite/ Co-requisite/ Exclusion	Nil
Objectives	This subject aims to equip students with fundamental skills in using R for data manipulation, visualization, and basic statistical analysis. It is designed for beginners and focuses on efficient, modern workflows using the tidyverse ecosystem (e.g., dplyr, ggplot2). By the end of the subject, students will be able to transform raw data into reproducible outputs and reports, providing a foundation for further studies in statistics and data science, and machine learning.
Intended Learning Outcomes	Upon completion of the subject, students will be able to: (a) set up and navigate the RStudio Integrated Development Environment (IDE) and manage R projects efficiently. (b) import data from various sources (CSV, Excel, webpage) and clean, filter, and transform datasets using dplyr commands. (c) create professional-quality statistical graphics (scatter plots, histograms, boxplots) using the grammar of graphics (ggplot2). (d) apply basic programming concepts including vector operations, logical indexing, and writing simple custom functions. (e) generate reproducible data analysis reports using RMarkdown, integrating code, output, and narrative text.
Subject Synopsis/ Indicative Syllabus	The course is conducted as a weekly hands-on workshop (1 hour per week). 1. Getting Started: Introduction to R, RStudio, and the "Project" workflow. Basic syntax and arithmetic. 2. Vectors & Data Types: Understanding vectors, lists, data frames, and factors. 3. Data Import: Reading data using readr and readxl. Handling missing values (NA). 4. Data Transformation (Part 1): Introduction to the tidyverse pipe (%>% or |>). Using select() and filter(). 5. Data Transformation (Part 2): Creating new variables with mutate() and sorting with arrange(). 6. Data Aggregation: Grouped summaries using group_by() and summarize().

<table>
<tr><td></td><td>
7. Data Visualization (Basics): The Grammar of Graphics. ggplot2 layers: aesthetics (aes) and geometries (geom_point, geom_line).
8. Data Visualization (Distribution): Visualizing distributions using histograms, density plots, and boxplots.
9. Data Visualization (Customization): Faceting, themes, labels, and saving plots.
10. Basic Statistics in R: Performing T-tests, correlation matrices, and simple linear regression models (lm).
11. Control Flow & Functions: Writing simple if-else statements, for loops, and defining custom functions.
12. Reproducible Research: Introduction to RMarkdown/Quarto for knitting documents to HTML/PDF.
13. Review & Application: End-to-end case study: transforming a raw dataset into a final visual report.
</td></tr>
<tr><td>Teaching/Learning Methodology</td><td>The subject is delivered through weekly interactive computer workshops in a computer laboratory. The pedagogical approach focuses on “code-along” instructions, where students actively type and execute code simultaneously with the instructor to build immediate practical competence. Some micro-challenges may be included in this course for students to learn how to active debugging.</td></tr>
<tr><td>Assessment Methods in Alignment with Intended Learning Outcomes</td><td>
<table>
<tr><td rowspan="2">Specific assessment methods/tasks</td><td rowspan="2">% weighting</td><td colspan="6">Intended subject learning outcomes to be assessed (Please tick as appropriate)</td></tr>
<tr><td>a</td><td>b</td><td>c</td><td>d</td><td>e</td><td></td></tr>
<tr><td>1. Quizzes</td><td>30</td><td>✓</td><td>✓</td><td>✓</td><td>✓</td><td></td><td></td></tr>
<tr><td>2. Lab participation</td><td>20</td><td>✓</td><td>✓</td><td>✓</td><td>✓</td><td></td><td></td></tr>
<tr><td>3. Project</td><td>50</td><td>✓</td><td>✓</td><td>✓</td><td>✓</td><td>✓</td><td></td></tr>
<tr><td>Total</td><td>100 %</td><td colspan="6"></td></tr>
</table>

Explanation of the appropriateness of the assessment methods in assessing the intended learning outcomes:

Quizzes and lab participations force students to review regularly. This ensures that students identify and rectify conceptual gaps along the semester. The project serves as an authentic assessment by simulating a real-world analyst’s workflow: importing raw data, cleaning, analysing and reporting findings.
</td></tr>
<tr><td>Student Study Effort Expected</td><td>
<table>
<tr><td>Class contact:</td><td></td></tr>
<tr><td>▪ Workshops</td><td>13 Hrs.</td></tr>
<tr><td>Other student study effort:</td><td></td></tr>
<tr><td>▪ Project</td><td>12 Hrs.</td></tr>
</table>
</td></tr>
</table>

	▪ Self-study	15 Hrs.
	Total student study effort	40 Hrs.
Reading List and References	de Vries, A. & Meys, J. (2015). R for Dummies. Wiley. Wickham, H., & Crolemund, G. (2023). R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. O'Reilly.	