

Data Science for Healthcare Using R

Microcertificate - Weekly Schedule

Please note that the schedule may change once the course begins; however, all listed content will be covered during the microcertificate.

Important Dates

Milestone	Date
In-person kickoff session at St. Paul's Hospital (also via live stream)	October 16, 2026
Holiday Break	December 14, 2026 – January 7, 2027
Course resumes — in-person lecture at St. Paul's Hospital (also via live stream)	January 8, 2027
Capstone Project	January 25 – February 5, 2027

Week 1 — Friday, October 16 – Sunday, October 25 (In-Person/Livestream Kickoff, Oct 16 — St. Paul's Hospital)

R Foundations, Data Import/Cleaning, Tidyverse Intro

- Write and execute basic R commands
 - Vectors, matrices, lists, and data frames
 - Indexing, subsetting, and filtering techniques
 - The concept of vectorized functions
 - Import and export data from various formats
 - Identify and handle missing values and outliers
 - dplyr functions (part 1): pipes, arrange, filter, pull, select, relocate, mutate, transmute, rename
 - Concept of tidy vs. non-tidy data
-

Week 2 — Monday, October 26 – Sunday, November 1

Conditionals & Loops + File System / Git Setup

- if / else if / else conditional logic
- Loops for efficient computation (for loops, apply-family loops)
- Avoiding loops with vectorization (reinforcement)
- Working with the file system: reading directory contents; copying, moving, renaming, deleting files; extracting file attributes (timestamps); creating files inside a loop
- Git for version control

Week 3 — Monday, November 2 – Sunday, November 8

Regex Deep Dive + Tidyverse, Part 2

- Regex with dplyr and the stringr package
 - Regex application to data cleansing
 - Tidy: wide and long data, pivoting, unite, separate.
 - Handling NAs.
 - Piping and joining.
-

Week 4 — Monday, November 9 – Sunday, November 15

Exploratory Data Analysis (EDA), Part 1

- Summarizing data using simple descriptive statistics (mean, median, quantiles, min, max, qqplot)
 - group_by and summarize for category-level insights
 - Publication-quality figures using base R graphics
-

Week 5 — Monday, November 16 – Sunday, November 22

Exploratory Data Analysis (EDA), Part 2

- Visualizations using ggplot2
 - aes, geom, stats, scales, facets and themes
 - dot plots, line plots, bar charts, box plots, violin plots and visual customization
 - Completing the EDA module (4-hour block split across Weeks 4–5)
-

Week 6 — Monday, November 23 – Sunday, November 29

Basic Statistical Analysis in R

- Common statistical tests (t-tests, chi-square, correlation)
 - Generating p-values and confidence intervals
 - Remainder of week: applied practice exercises
-

Week 7 — Monday, November 30 – Sunday, December 6

Reproducible Research with R

- R Markdown for PDF/HTML output
 - Reproducible research workflows
 - Automating an R analysis
 - kableExtra for polished tables
-

Week 8 — Monday, December 7 – Sunday, December 13

Effective Use of LLMs for Code Generation

- What ChatGPT/LLMs are useful for — and where they fall short
 - Effective prompting to generate code blocks efficiently
 - Spotting and avoiding poorly written LLM-generated code
 - Remainder of week: open project/practice time
-

Holiday Break — December 14, 2026 – Thursday January 7, 2027 (No new content)

Week 9 — Friday, January 8 – Sunday, January 17 (In-Person/Livestream Session, Jan 8 — St. Paul's Hospital)

Dashboard Production, Part 1

- Building dashboards with flexdashboard
-

Week 10 — Week of January 18 – Sunday, January 24

Dashboard Production, Part 2

- Interactive dashboards with Shiny
 - Completes the module set
 - **Capstone Project briefing** — dataset options, expectations
-

Capstone Project - Monday, January 25 – Friday, February 5th

- Students will have two weeks to submit their capstone project as their final assessment
-

DRAFT