

广东省公共卫生研究院

粤公卫院函〔2025〕27号

广东省公共卫生研究院关于举办 R 语言 医学数据挖掘技术培训班的通知

各有关单位：

为提高卫生专业技术人员和相关科研工作者医学数据的分析与挖掘能力，我院将于 6 月底在东莞市举办“R 语言医学数据挖掘技术培训班”。本次培训班将基于 R 语言讲解医学数据的处理以及机器学习等多种分析方法，结合实际应用案例进行详细介绍和实践操作。培训班相关事项安排如下：

一、培训内容

- （一）R 语言数据处理基础及 AI 辅助编程；
- （二）回归分析方法；
- （三）机器学习分类方法；
- （四）机器学习聚类和降维方法；
- （五）Meta 分析方法；
- （六）ggplot2 绘图技术。

培训班日程安排详见附件。

二、培训对象

各相关单位的卫生健康工作者和研究人员，以及对 R 语言数据分析感兴趣的各界人士。

三、培训安排

（一）时间：6 月 24—28 日，24 日下午报到，28 日上午离会。

（二）培训方式：线下培训与线上培训同步进行。

（三）线下培训地点：东莞凯利酒店（南城市中心），广东省东莞市南城区元美路西侧。

线上培训：通过小鹅通平台直播。

（四）培训费用

1.线下培训：培训费 2800 元/人，含学费、场地费、资料费和午餐费。学员的早餐费、晚餐费、住宿费和交通费需自理。

2.线上培训：培训费 2200 元/人，包含学费、资料费。

3.收费方式：培训费请通过银行转账进行缴费，收款账户如下：

收款单位：广东省公共卫生研究院

账号：3602088709200021474（写全 19 位）

开户银行：工商银行广州怡乐支行

请务必备注：R 培训班线上/线下+姓名

（五）报名方式：学员请于 6 月 20 日前扫描下图二维码填写报名信息。如有疑问，请联系沈老师或容老师，联系电话：020—31051612；也可加入报名咨询 QQ 群，群号：801116971。



报名二维码

四、其他事项

（一）报名成功后，工作人员将组织学员加入培训学员 QQ 群。培训资料和小鹅通平台直播事宜将在 QQ 群中发布。

（二）培训内容包含软件实践操作，请学员准备 Windows 系统电脑，提前安装好 R 软件和 RStudio 软件，工作人员将在培训学员 QQ 群上传相关软件和安装指南。

（三）培训费发票统一以电子发票形式于培训结束后一周内发到学员邮箱。

（四）培训地点所在酒店可提供住宿，学员请自行拨打酒店前台电话（余经理 134 1229 5103，0769-22286888）预订，大床房/双床房团购价 328 元/间/天（含早餐）。房间有限，请提前预订。

附件：R 语言医学数据挖掘技术培训班日程表


广东省公共卫生研究院
2025 年 5 月 14 日

附件

R 语言医学数据挖掘技术培训班日程表

时间：6 月 24—28 日

时间		课程主题	课程内容
6 月 24 日下午		报到	
6 月 25 日	8:30-12:00	R 语言数据处理基础及 AI 辅助编程	1.R 语言基础（R 语言简介、基础语法、数据结构；R 包安装及加载、路径管理） 2.数据处理及应用（数据批量读写与合并、向量处理及常见函数，以水痘个案调查的数据清理为例） 3.AI 工具辅助 R 语言代码编写和代码解读示例
	14:00-17:00	ggplot2 绘图技术	1.ggplot2 绘图系统和要素介绍 2.数据可视化方法（基于疾病监测数据的箱型图、时序图、误差棒图、热力图、六边形地图、分面图、多图合并、交互式图等）
6 月 26 日	9:00-12:00	回归分析方法	1.广义线性模型（以臭氧暴露与人群死亡关系分析为例） 2.混合线性模型（以特定人群肺功能的影响因素分析为例）
	14:00-17:00	机器学习分类方法	1.Logistic 回归 2.支持向量机 3.随机森林 （基于 3 种机器学习算法，建立模型对乳腺癌进行分类和预测，计算召回率、精确度等指标，绘制 ROC 曲线，对模型预测效果进行比较）
6 月 27 日	9:00-12:00	机器学习聚类和降维方法	1.K-means 聚类（以某地社会卫生经济学指标聚类分析为例） 2.主成分分析（以某种药物化合物的成分质量评价指标为例，提取指标的主要成分） 3.因子分析（以健康体检指标数据的因子分析为例，降低相关特征变量维度）
	14:00-17:00	Meta 分析	1.分类型变量 Meta 分析 2.连续型变量 Meta 分析 （以多地区/多中心的癌症发生风险 RR 值、HR 值、 β 值等效应值的 Meta 分析为例；以 EV71 疫苗知晓率和保护率的 Meta 分析为例）
6 月 28 日上午		离会	