

12 NPHC - PRE-CONFERENCE WORKSHOP 1																		
Title	:	Practical Epidemiological Analysis Using ChatGPT and R																
Introduction	:	ChatGPT is an advanced AI chatbot that supports epidemiological tasks and analyses. This workshop demonstrates practical data analysis techniques using real-world datasets, focusing on data from the MOH GitHub repository. Participants will gain hands-on experience in R for data cleaning, visualisation, and statistical testing to enhance analytical efficiency.																
Learning Outcomes	:	By the end of the workshop, participants will able to: <ol style="list-style-type: none"> 1. Use ChatGPT to guide epidemiological tasks. 2. Perform data cleaning and exploration using R. 3. Conduct epidemiological analyses such as incidence and trend analysis. 4. Apply ethical AI practices in public health research. 																
Target participants	:	Public health professionals, including specialists, epidemiologists, and officers involved in data analysis or research.																
Number of participants (min – max)	:	Min 10 pax Max 30 pax																
Pre-requisite for participants (if applicable)	:	Participants must have: <ol style="list-style-type: none"> 1. R and RStudio installed on their laptops. 2. Pre-downloaded datasets from the MOH GitHub repository or access to the repository. 3. Basic R programming knowledge is helpful but not mandatory. 																
Instructors:	:	<ol style="list-style-type: none"> i. Prof Dr Jamalludin Ab Rahman (Kulliyyah of Medicine, IIUM) ii. Dr Muhammad Adil Zainal Abidin, (Kulliyyah of Medicine, IIUM) 																
Tentative schedule:	:	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left;">Time</th> <th style="text-align: left;">Activity</th> <th style="text-align: left;">Description</th> </tr> </thead> <tbody> <tr> <td>8.30am</td> <td>Registration</td> <td>Sign-in and distribution of workshop materials.</td> </tr> <tr> <td>9.00am</td> <td>Introduction</td> <td>Overview of workshop objectives, tools (ChatGPT and R), and the dataset.</td> </tr> <tr> <td>9.30am</td> <td>Data Setup and Exploration</td> <td>Importing the dataset, understanding variables, and exploring data structure.</td> </tr> <tr> <td>10.15am</td> <td>Break</td> <td></td> </tr> </tbody> </table>		Time	Activity	Description	8.30am	Registration	Sign-in and distribution of workshop materials.	9.00am	Introduction	Overview of workshop objectives, tools (ChatGPT and R), and the dataset.	9.30am	Data Setup and Exploration	Importing the dataset, understanding variables, and exploring data structure.	10.15am	Break	
Time	Activity	Description																
8.30am	Registration	Sign-in and distribution of workshop materials.																
9.00am	Introduction	Overview of workshop objectives, tools (ChatGPT and R), and the dataset.																
9.30am	Data Setup and Exploration	Importing the dataset, understanding variables, and exploring data structure.																
10.15am	Break																	

	10.30am	Data Cleaning	Preparing data for analysis, handling missing values, and adding new variables.
	11.00am	Descriptive Analysis	Summarising data, mortality by age and gender, and visualisation techniques.
	11.45am	Mortality Rates and Stratification	Calculating mortality rates and performing stratified analysis by age and vaccination.
	12.15pm	Trend Analysis	Analysing and plotting mortality trends over time.
	12.45pm	Advanced Techniques	Hypothesis testing, regression analysis, and ChatGPT for R coding.
	1.15pm	Ethics and reporting	Ethical consideration and reporting guideline in statistical analysis plan.
	1.30pm	Q&A	Open discussion on workshop content and additional queries.
	1.45pm	Workshop Ends	

OUTLINE OF THE WORKSHOP

Outline:	:	<p>1. Introduction</p> <ul style="list-style-type: none"> • Objective: Introduce the workshop objectives, tools, and dataset. • Key Points: <ul style="list-style-type: none"> • Role of ChatGPT in epidemiological analysis. • Overview of R as a statistical tool. • Dataset introduction: COVID-19 Linelist Deaths from the Ministry of Health Malaysia. <p>2. Data Setup and Exploration</p> <ul style="list-style-type: none"> • Objective: Set up the environment and understand the dataset. • Activities: <ul style="list-style-type: none"> • Installing necessary R packages (tidyverse, lubridate, etc.).
-----------------	---	--

- Importing the dataset from the GitHub repository.
- Initial exploration: structure, summary statistics, and missing values.

3. Data Cleaning

- Objective: Prepare the dataset for analysis.
- Activities:
 - Handling missing data.
 - Creating derived variables, e.g., age groups (Children, Adults, Elderly).
 - Ensuring date formats are correct.

4. Descriptive Analysis

- Objective: Summarise key data features and visualise trends.
- Activities:
 - Mortality distribution by age, gender, and other demographics.
 - Creating visualisations (e.g., bar plots for age groups).

5. Mortality Rates and Stratification

- Objective: Calculate and analyse mortality rates.
- Activities:
 - Mortality rate computation using population data.
 - Stratified analysis by vaccination status and age group.

6. Trend Analysis

- Objective: Explore mortality trends over time.
- Activities:
 - Analysing daily mortality counts.
 - Visualising trends using line charts.

7. Advanced Techniques

- Objective: Apply advanced epidemiological analysis methods.
- Activities:

- Hypothesis testing (e.g., Chi-square tests for categorical variables).
- Logistic regression analysis to predict outcomes based on variables like age and vaccination status.
- Using ChatGPT for debugging, generating R scripts, and exploring alternative approaches.

8. Ethical Use and reporting guideline of AI

- Objective: Highlight the responsible use of AI in public health and transparency reporting based on current guideline.
- Key Points:
 - Benefits and limitations of AI tools like ChatGPT.
 - Ensuring data privacy and avoiding bias in analyses.
 - The importance of human oversight in AI-generated outputs.
 - Transparency reporting of using AI.

9. Q&A Session