

# Hackathon!

**Alistair Johnson**

11 Nov 2023



# Learn by doing

- Dive into the data early!
- Time is short, so be practical.
- Be prepared to pivot.



# Challenge focus: Acute kidney injury

- Analyse definitions
- Extract variables
- Build predictive models

# Tell us what you found (noon Monday)

- Think about your slide deck from the beginning
- We are (mostly!) not interested in performance measures



# Evaluation

- Defining the problem and putting it in context
- Identifying the elements of the study
- Methodology
- Presentation
- Reproducibility



# Enjoy yourself

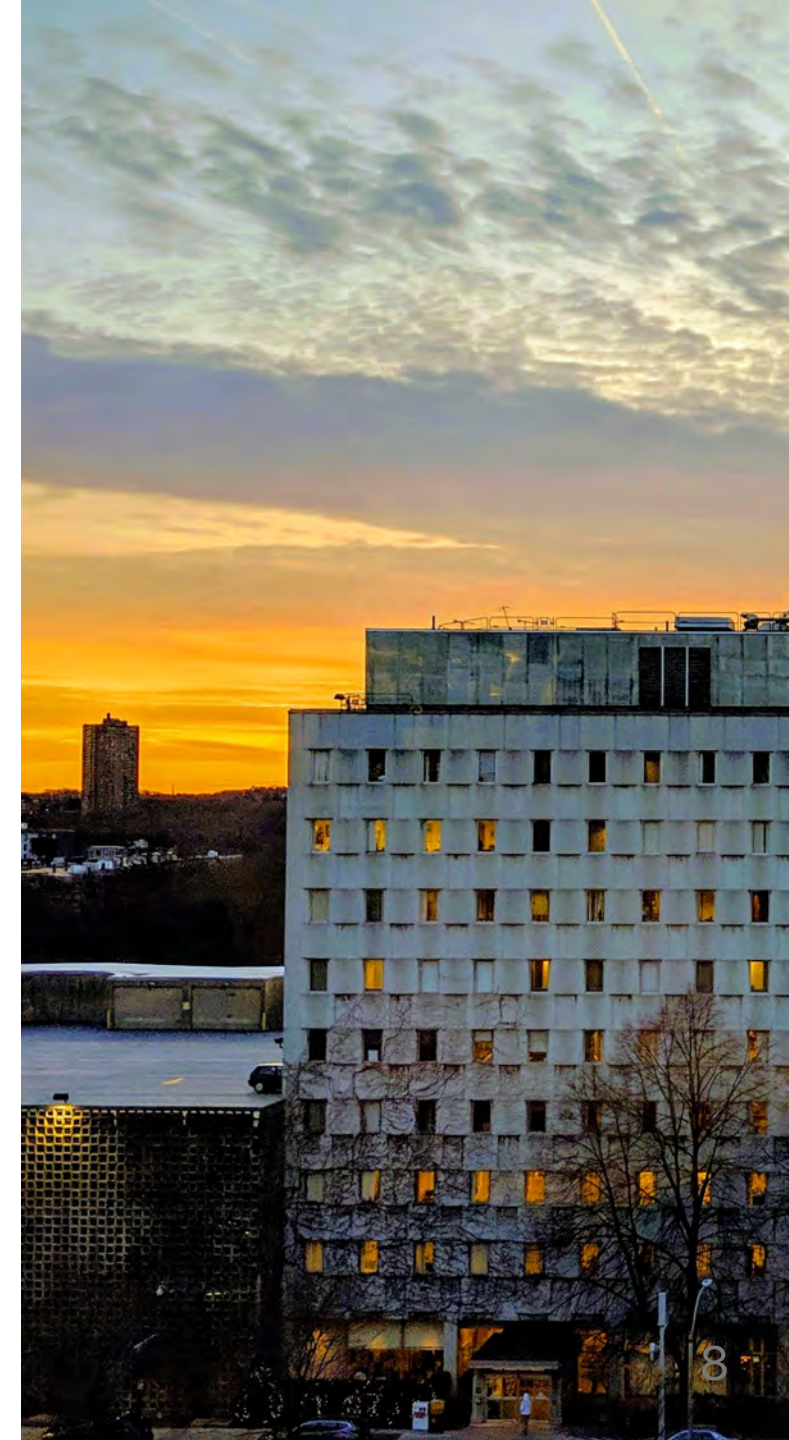
- Think beyond the datathon. How will you continue?



# Data

# MIMIC-IV

- Highly-detailed critical care database
- >40k patients
- Vital signs, medications, labs..
- Complies with US privacy laws (HIPAA), deidentified.
- Accessible to approved researchers.





The background of the slide is a dark, semi-transparent ECG (heart rate) signal. The signal is plotted on a grid with red horizontal and vertical lines. The waveform is black and shows several distinct peaks and troughs, characteristic of a heart rate signal. The overall tone is professional and scientific.

# PhysioNet

The Research Resource for Complex Physiologic Signals

Data

Software

Challenges

Tutorials

# MIMIC-IV

Alistair Johnson , Lucas Bulgarelli , Tom Pollard , Steven Horng , Leo Anthony Celi , Roger Mark 

Published: Jan. 6, 2023. Version: 2.2

**When using this resource, please cite:** [\(show more options\)](#)


Johnson, A., Bulgarelli, L., Pollard, T., Horng, S., Celi, L. A., & Mark, R. (2023). MIMIC-IV (version 2.2). *PhysioNet*. <https://doi.org/10.13026/6mm1-ek67>.

**Additionally, please cite the original publication:**

Johnson, A.E.W., Bulgarelli, L., Shen, L. et al. MIMIC-IV, a freely accessible electronic health record dataset. *Sci Data* 10, 1 (2023). <https://doi.org/10.1038/s41597-022-01899-x>

## Abstract

Retrospectively collected medical data has the opportunity to improve patient care through knowledge discovery and algorithm development. Broad reuse of medical data is desirable for the greatest public good, but data sharing must be done in a manner which protects patient privacy. The Medical Information Mart for Intensive Care (MIMIC)-III database provided critical care data for over 40,000 patients

Contents 

### Share



### Access

#### Access Policy:

Only credentialed users who sign the DUA can access the files.

#### License (for files):

[PhysioNet Credentialed Health](#)

**Sign up at:**

<https://physionet.org/events/4S3IZ9iPeqiZ/>

# Resources



# Hackathon materials

<https://github.com/xborrat/NEFRoHack>

# BigQuery

<https://console.cloud.google.com/bigquery>

# Colab

<https://colab.research.google.com/>

**Make use of the people  
around you!**

