HPDM097 'Making a difference with health data'
GROUP ASSIGNMENT 2

Problem background

An NHS trust has asked you as a data scientist to help them plan the future capacity of their **Acute Stroke Unit (ASU)**. The ASU currently admits three types of patient:

- Patients who have suffered an acute stroke
- Patients who have suffered a transient ischaemic attack (TIA)
- Patients who have complex neurological conditions.

The ASU currently has 9 inpatient beds available. When a patient is admitted they undergo a period of treatment and then are discharged elsewhere. The NHS data confirms that each type of patient has its own length of stay distribution. This high level process is illustrated in Figure 1. **A technical appendix is provided at the end of the assignment briefing.**

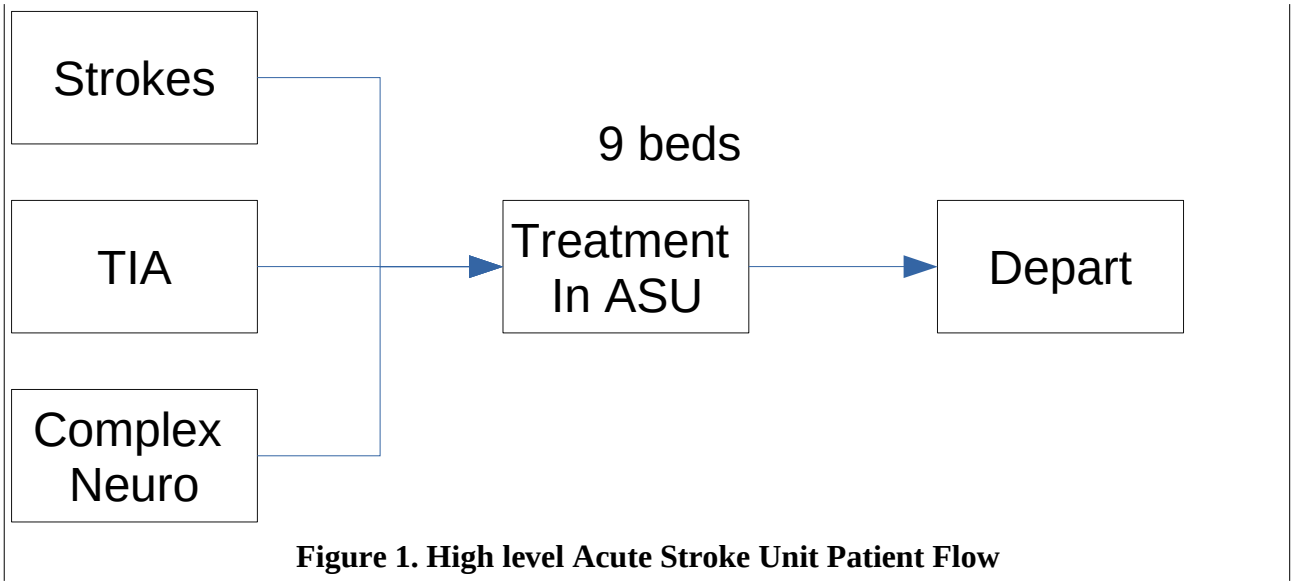

Figure 1. High level Acute Stroke Unit Patient Flow

Primary (essential) research questions:

- The ASU managers would like to know what bed capacity is needed to achieve an annual target of **90% of patients being admitted to the stroke unit within 4 hours of their arrival at the hospital.**
- The managers would also like to know if they can achieve 90% performance and maintain a **bed utilisation** in excess of 70%
- The period of interest for each performance measure is 1 year.

Secondary (desirable) research questions:

- By 2030 the ASU has forecast a 10% increase in patients requiring an admission. They would like to understand the impact on **the percentage of patients admitted to the stoke unit within 4 hours of arrival, bed occupancy and if more bed capacity is needed**.

Your Task

- Design and build a **discrete-event simulation** model of the ASU using **SimPy**.
- **At a minimum** the model should be able to answer the NHS's **essential** research questions.
- The model should take account of each type of patient, their arrival rate and length of stay.

- The analysis of model performance should use an appropriate run length and make use of high quality output analysis techniques.
- Results must be clearly presented and justified.
- Your modelling, results and recommendations must be submitted in a lab report.

Lab report details:
The purpose of the lab report is to succinctly document and communicate your model and analyses to a team of NHS analysts. You can assume the readers will have knowledge of the ASU system and so you do not need to re-explain it. The report should include a title, executive summary, aims, planned analyses, findings and conclusions. The markdown in lab report **should be no more than 1000 words in length**. The report should be neatly presented and include numbered sections and subsections where appropriate. The report should be in clear English and be **spell checked**. Lab reports must be provided in **Jupyter Notebooks**. The notebook should include runnable python code to reproduce your findings. Outputs may include tables, and charts that *clearly* convey your results. Please note that code listings do not count towards the word count. Please prefix the lab report with your group number e.g. 'group_1_report.ipynb'. Note that the simulation model can be contained in separate modules or a package and imported into the notebook to be run. The notebook should describe the workings of the model to the NHS analysts.

Coding standards:
Students should undertake the assignment using Python 3. As covered in the course, students are expected to produce code of the highest standard and where possible follow PEP8 guidelines for Python. All code should be well structured, documented and make use of functions, appropriate data structures and data science packages.

Group Assignment Submission.
Assignments must be handed in ***on 3rd*** April 2023 by 12pm. This deadline is strict.

The procedure for submitting group work to eBART is as follows:

- **Only one student per group should submit the code and lab report**. Nominate someone in your group to do this and make sure you arrange a time, before the deadline, to confirm that this has been done. If possible do this together (or via a shared screen in Teams) so you can all check that the submission is correct, it has been uploaded to the correct link, and that the file hasn't been corrupted.
- The remaining members of your group must submit a PDF with the names of their group members on it.

Your submission must consist of a single Zip file (group_[group number].zip) containing:
- A Jupyter Notebook (.ipynb) containing your lab report and runnable Python code.
- Any additional modules or python packages you have created.

See next page for technical appendix.

Technical appendix:

Current performance of the ASU

The NHS do not have hard data on performance, but they **believe** that with current demand and bed availability that they are seeing less than 80% of patients admitted to the ASU within 4 hours.

Data for the model.

A collaborating data scientist has analysed the available data for this project and provided the following inputs for your simulation modelling

Arrivals of patients who require an admission

The following parameters specify the **average inter-arrival time** of the different patient groups. All arrivals are random and follow an **Exponential** distribution.

Table T1: Patient inter-arrival times.	
Patient Type	**Mean IAT (days)**
Acute strokes	1.2
Transient Ischaemic Attack (TIA)	9.5
Complex Neurological	3.5

Acute length of stay parameters

Initial input modelling found that the length of stay data approximately follows a **log normal** distribution.

Table T2: Acute length of stay parameters		
Patient type	**Mean (days)**	**Stdev (days)**
Acute strokes	7.4	8.5
Transient Ischaemic Attacks	1.8	2.3
Complex Neurological	2.0	2.5

Modelling simplifications and assumptions:

- You can assume first come first served as a queuing discipline i.e. patients are admitted to the ASU unit in the order they arrived.
- For the basic modelling you can simplify the operations of the real system and assume that patients will wait indefinitely to be admitted to the ASU.
  - Occasionally you may find that the modelled ASU is **overloaded** i.e. patients are not admitted for many days due to a full ward and a build up in the queue for admissions.
  - In reality, patients may spend their acute stay as an ASU outlier where the stroke physicians conduct a *safari round* and care for patients in a separate ward.
  - Advanced projects may optionally wish to challenge this simplification and consider alternative more accurate ways of modelling the problem.
- The destination of discharge is not part of this project and you do not need to model the destination explicitly. The NHS trust does not have reliable data on where patients end up. In reality you should expect this to be a mix of patient mortality, transfer to rehabilitation, nursing homes and transfer home with support.

Marking scheme:

Mark range	Quality of assignment
Fail 0-49	• Absent or mostly incorrect simulation model and analysis with low adherence to a sound simulation study process. • Absent or poorly presented lab report, possibly in the incorrect format, with absent or limited use of tables, figures and several grammatical and typographical errors. The report exceeds the word limit. • Absent or highly limited recommendations to the NHS. • Poorly documented and formatted code that does not attempt to follow PEP8 guidelines; absent or highly limited use of good practice coding practices for code organisation in simulation studies. • Code does not reproduce results and contains multiple run time errors. It cannot be executed.
Pass 50-59	• Limited analysis of the simulation problem. The analysis and modelling may fail to follow best practices or be unclear. • Borderline passable report where reporting of the results would benefit from the use of tables, figures and sections to succinctly convey results. The report contains multiple grammatical and typographical errors, but is within the word limit. • Conclusions and recommendations are provided, but are poorly justified or are not fully supported by the analyses. • Demonstration of limited understanding of how to organise code and best practice in simulation studies. The code does not followed PEP8 guidelines. The readability and maintainability of the code could be improved by adopting best practices. • Code mostly reproduces results, but contains some minor runtime errors and minor bugs.
Merit 60-69	• Limited but mostly correct analyses and modelling that provides clear evidence of understanding simulation methods applied to stochastic health systems but falls short of the highest standards. • Generally well presented lab report with limited use of tables, figures, and sections. Several grammatical and typographical errors. The report is within the word limit. • Mostly clear recommendations to the NHS with some improvement possible in clarity and justification and support from analyses. • Generally well presented code that attempts to follow PEP8 guidelines. Code has limited documentation and demonstrates some evidence of following best practice for organisation and reuse in simulation studies. The readability and maintainability of the code could be further improved. • Code reproduces reported results and contains no run time errors, but may contain minor bugs.
Distinction 70–85	• Generally correct and well justified simulation modelling and analyses. Work demonstrates good understanding of simulation concepts, methods and limitations. Clear evidence of following a sound simulation study process. • Mostly well-presented lab report with good use of tables, figures, and sections. Limited grammatical and typographical errors. The report is within the word limit. • Clear and accurate recommendations for the NHS that are mostly justified and are supported by the analyses. Conclusions critique the analyses to identify strengths and limitations. • High standard code with minimal breaches of PEP8 guidelines. The code is mostly readable. Code makes use of best practices for organisation and reuse. • Code reproduces results, contains no runtime errors and minimal bugs.
Distinction++ 86-100	• Comprehensive publication quality analyses and results that make use of appropriate methods and may make use of advanced methodology. All analyses are justified, thoughtfully designed and employ a robust high quality approach to presentation and modelling. • Publication quality, succinct and thoughtful lab report. Excellent English, use of tables, figures, sections and very limited, if any, grammatical and typographical errors. • Clear, concise and accurate recommendations for the NHS that are fully justified by the included analyses. Conclusions include a critique and accurately and fully caveat the limitations. • Flawless publication quality code that closely follows PEP8 guidelines for style and documentation. The code is easy to read and maintain. The code make use of advanced concepts from Python where appropriate. • The code reproduces the results exactly with no runtime errors and no bugs

Students will be provided with individual feedback and a grade breakdown as above.