1. **Notebook mini project**

**The general question you will address is:** What mathematical models best fit an empirical dataset?

### Workflow ###

Clean data 🡪 fit curve 🡪 model selection using AIC, BIC or R^2 🡪 plotting 🡪 run\_miniproject script 🡪 Report(LaTeX, 11pt, article document class, 1.5-spaced……)

\*Silwood Masters Student Guidebook

### report structure ###

## Samraat’s post

An Introduction section with bullet points listing what you will cover in it

A Methods section with bullet points listing what you will be covering there (basically the main topic headings, such as "data cleaning", "model fitting", etc)

A Results section with at least 2 display items:

One Figure showing an example of all the models you have managed to fit: plot the model curves on a single dataset. Please include a Caption, and if possible, a Legend stating what the fIgure is showing - the "take-home message".

Another Figure or Table showing a summary of model performance (to convey a clear image/idea of which one wins). Here again, include a Caption and if possible, a legend.

In the Results, you can also include some (bulleted) points for what additional figures/tables you might like to include.

No Discussion section needed yet, but you can put in a placeholder section if you like, and some bullet points on what you would like to cover there.

###

**Title**

Up to 10-15 words are reasonable\

**Abstract:**

The sigmoidal curve has extensive application in the real world. With some complex system lacking a specific model, it can be used[1]. In ecology, specifically, population growth is a typical application. Most cases, the curve will have successive lag, growth, and asymptotic phases. After it gets the asymptote phase which population size has gotten the K value, some data’s population size will drop, which circumstance will not be discussed in the article. In this project, I used computational based methods analysed three models:polynomial(cubic) model, logistic model and Gompertz model.After fitting 285 datasets, the result shows that the polynomial model performs better than the other two models. +(Succinct discussion).

**Introduction (“expectations” for the readers)**

This article analysed the population (log)size against time. Addressed the basic steps choosing the Cubic, Logistic and Gompertz models[2], basically by comparing AIC and R^2.

**Methods**

{states briefly how each of the scripting languages (bash, R, Python) was used and what packages within them were used and a justification of why.}

1. Computational based model fitting project: git, R, python…
2. Choose the polynomial, logistic and Gompertz model to fit the data, and comparing the model by comparing AIC and R^2
3. To capture the lag phase, more complicated growth models, the Gompertz model[2], has been chosen in this project, which is asymmetrical compared with the logistic model.

**Results**

(plot histogram)

**Discussion** (reminds the reader about what the original goals of the study were, states out key findings succinctly, and then discusses their implications in the wider context and then finished off with some caveats and a conclusion that delivers the final takeaway messages.)

\*Avoid sub-sectioning (with headers) the Introduction and Discussion sections as it breaks the flow of your “narrative”. But subset the methods and results

\* checklist

1. check spell

2.use vector graphic

3.take home messages: tables, figures, legends, text caption

# Reference\

####### mid night thoughts #######

Maybe I can use the datasets have just a few data point to test the prediction effectiveness of the models. Cause like the data has some dataset has just several points and some have a lot of points

Q: some data has mess point should I just throw it away? (add the pic)

Cause technically it can be fitted in the code with generating parameters. While my code still has some problems with plotting, I cannot see the effect of those models intuitively, so I don’t know how those models work with those data.