Analysis of engagement in a maths support centre - effects, predictors and outlier groups

Piotr Kedziora

June 2020

Supervisor: Assistant Professor Anthony Cronin

University College Dublin, School of Mathematics and Statistics, Undergraduate Summer Research Project

Contents

1 Abstract

The maths support centre (MSC) in UCD has been consistently recording data of student visits for the past 11 semesters. The most interesting variables to look at are the attendance (which in this context, is measurable engagement) and looking at whatever grades exist. With this in mind, we first take a look at how attendance patterns throughout lectures and MSC visits impacted the grades of one outlier module. We then explore how an epidemiology model can be used to potentially predict attendance in the centre. Between all this, we include some miscellaneous material, such as the analysis of short sessions in the MSC, showing if they are mistakes or characteristic to one particular module.

2 Introduction

The main goal of this project was to explore many different avenues of data from a Maths Support Centre (MSC) in UCD. We take into consideration various literature outlining the benefits of a MSC and look to see if those benefits present themselves in UCD. We look into various attendance patterns of one particular module that's considered an outlier. We also investigate the variance of duration times between various demographic groups and report on any anomalies that present themselves in the data.

Finally, we attempt to predict the attendance patterns in the centre by using a model inspired by the SIR epidemiology model in collaboration with researchers from Swinburne.

This report contains research conducted in various areas of MSC data, and i hope it proves useful for research in the future.

A main topic of this research is **engagement**, but what exactly does that mean? **Engagement can mean..**

- Studying - using textbooks, looking up resources online, the learning process is all about engaging with the subject

- Interpersonal engagement - asking questions to lecturers and other students

- Attendance to lectures/Maths Support Centre

These are all types of engagement, but **attendance** is one that is easily quantifiable if consistently recorded.

... and so for the remainder of this report, attendance will be the primary independent variable.

3 Literature Review

"The term mathematics support centre is usually taken to mean a dedicated, physical space in which mathematics support is offered... While centres may offer a range of self-help learning materials and space for peer-peer learning, an almost universal feature is the availability of a tutor (or tutors) in the centre at specified times for one-to-one or small group advice."
— **Grove, Croft, Lawson & Petrie (2017)**

The positive impact of mathematics support centre engagement on students university grades has been studied extensively for years, mainly research coming from Ireland, UK and Australia.
Students who use Mathematics Support:

- are twice as likely to complete their courses [1]

- obtain results 8% higher than their peers [after controlling for an on-entry diagnostic test] [**Lee˙2008**]

- are more likely to be "at-risk" of failing their course (Mac an Bhaird & O'Shea, 2009), but also more likely to be a higher achieving student (Pell & Croft, 2008) [4]

- ..once are 1.63 times more likely to pass their module than for one who had never engaged with the service. [3]

Edwards and Carroll from 2018 [2], used epidemiology models to predict attendance patterns in their respective support centre. This is where the inspiration came from to pursue fitting a model to UCD's respective data.

4 Methodology

The goal of this research was to focus on obtaining quantitative results from MSC data. Python was the language of choice for this report, as it was something I had the most experience with in my undergrad. The raw data I received from my supervisor contained 6 types of data. "Student ID (encrypted for privacy), Topic (a brief few words describing the subject of the visit), Code (Module), Start (The date and time of the visit), and duration (how long a session lasted). Start and Duration were the main quantitative variables, so this is where most of the focus went.

5 Results

Initially, I decided to look into numerous distributions of average duration times (x) versus number of sessions (y) in the MSC between various demographic groups:

(a) 1st year modules

(b) 2nd year modules

(c) Modules inside school of maths
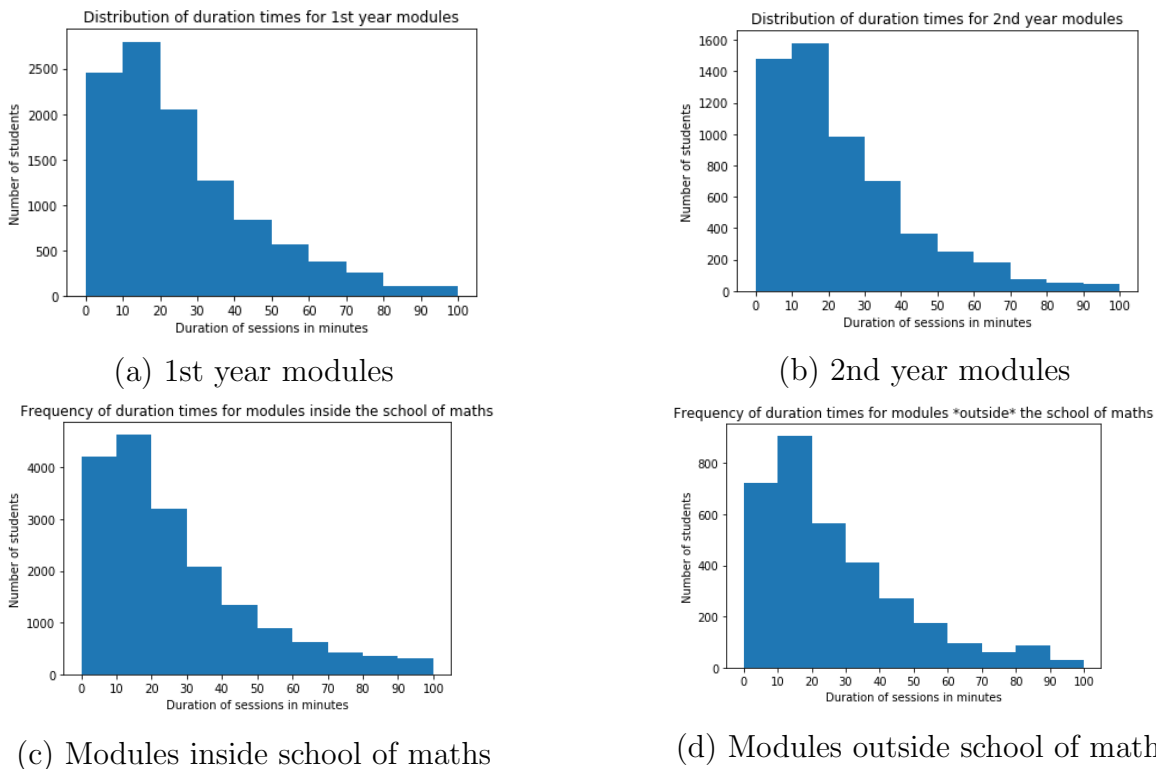

(d) Modules outside school of maths

Figure 1: Average duration times between demographic groups

It was found that between 1st year modules, 2nd year modules as well as modules taught inside and outside of school of maths all shared a similar distribution throughout. Furthermore, I decided to look into the top 10 most popular modules from the MSC and investigate what the average duration times of these were. What I found was that the access module MATH00030 seemed to be an outlier having a much longer average duration than other modules.

This result may have stemmed from the module's smaller class size, where students do not typically enter from the traditional Leaving cert route.

Average duration of the most popular modules in MSC over 5 years
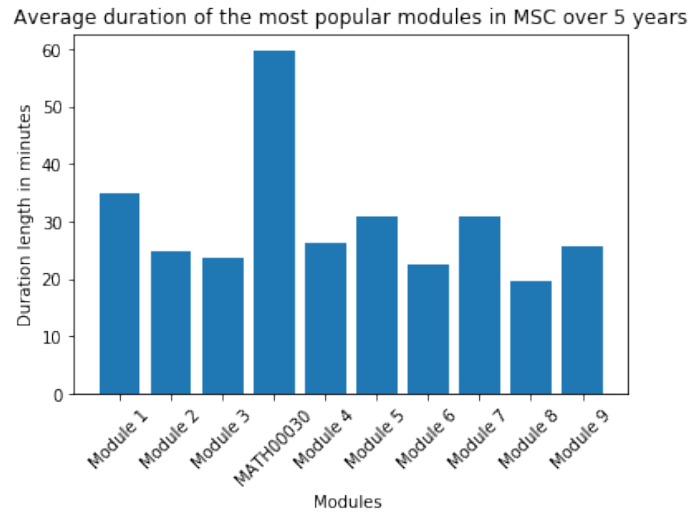
So within this module, I decided to carry out further investigation, focusing specifically on one particular year: 2014, which is where a lot more data was available.

We investigate 3 main types of attendance within this module

- **Lecture attendance**: Attendance to class every week, recorded by lecturer.

- **MSC workshop attendance**: Weekly dedicated 90-min sessions for access students

- **Independent MSC attendance**: How often these access students came and used the MSC by themselves

What we hoped to see is a fairly consistent positive correlation between attendance of all three types and the *Final Grade (%)*.
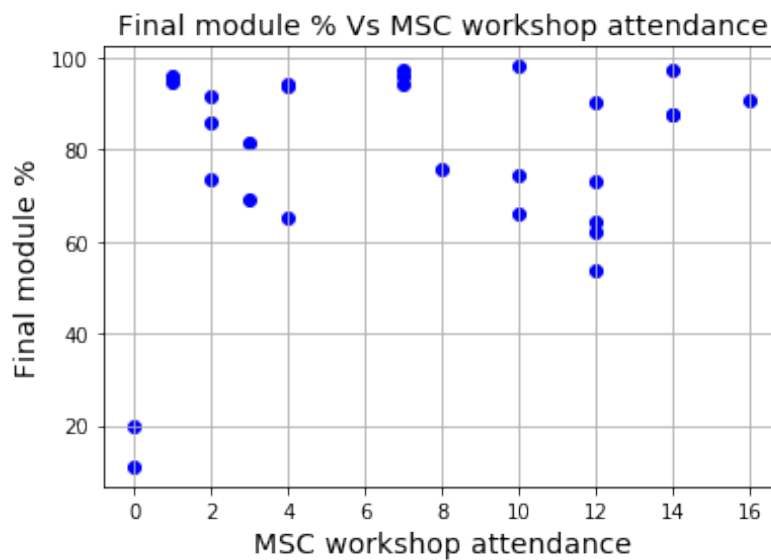
Access module vs attendance patterns

Lecture attendance
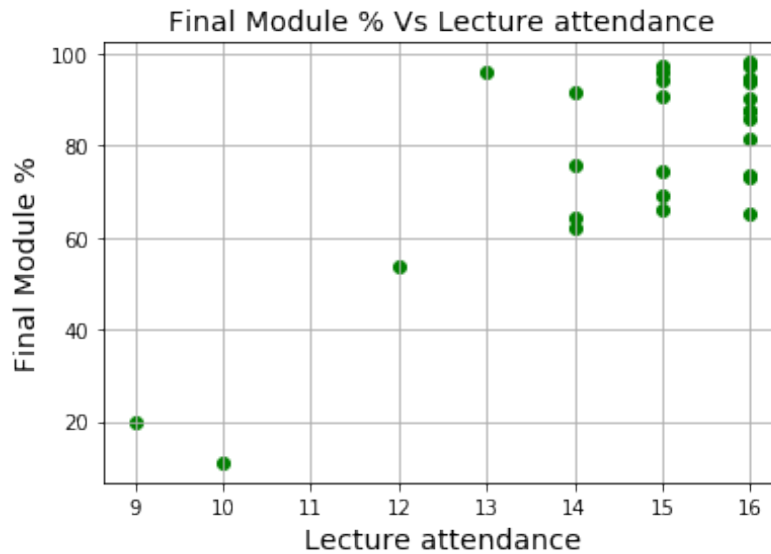
- As expected, the higher the lecture attendance, the higher the performance.

- Only 2 people failed, and they attended 10 or less lectures.

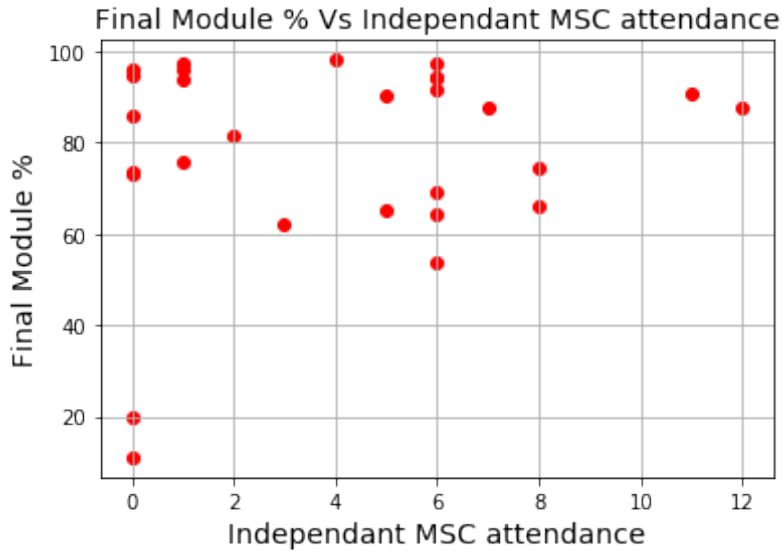- The correlation coefficient is: $r = 0.7884$

MSC workshop attendance

- Seems quite inconsistent and non-linear.

- Doesn't seem to correlate with final grade.

- Correlation coefficient : $r = 0.2382$

Independent MSC attendance

- Similar to MSC workshop attendance, non-linear, low correlation.

- Correlation coefficient: $r = 0.1884$

Summary statistics - Ordinary Least Squares regression

```
                         OLS Regression Results
==============================================================================
Dep. Variable:        Final Grade (%)   R-squared:                       0.624
Model:                           OLS    Adj. R-squared:                  0.576
Method:                Least Squares    F-statistic:                     13.25
Date:               Thu, 30 Jul 2020    Prob (F-statistic):           2.64e-05
Time:                       12:49:18    Log-Likelihood:                 -111.93
No. Observations:                 28    AIC:                             231.9
Df Residuals:                     24    BIC:                             237.2
Df Model:                          3
Covariance Type:           nonrobust
==============================================================================
                             coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const                    -61.5552     22.491     -2.737      0.011    -107.975     -15.135
Independant MSC attendance -0.1710      1.018     -0.168      0.868      -2.273       1.930
MSC workshop attendance    0.2514      0.720      0.349      0.730      -1.234       1.737
Lecture attendance         9.3841      1.564      6.000      0.000       6.156      12.612
==============================================================================
Omnibus:                       1.218   Durbin-Watson:                   1.872
Prob(Omnibus):                 0.544   Jarque-Bera (JB):                0.696
Skew:                          0.386   Prob(JB):                        0.706
Kurtosis:                      3.004   Cond. No.                        145.
==============================================================================
```

Analysis

- p-value for both MSC attendance data is too high, indicates that evidence is not strong enough to support a hypothesis that claims there is a relationship between MSC attendance and grades for this module.

- Do strong students not attend MSC voluntarily or perhaps the opening times don't suit them as these are part-time adult students with jobs, caring responsibilities etc.

- People seem to prioritise lectures rather than the MSC

Comparing to other years

Analysis of Access Modules and attendance vs grades

It can be worth it now to compare this module's grade performance vs lecture attendance over a period of a few years. Below are graphs showing grades vs attendance of MATH00030 and its semester 2's continuation MATH00040.

As expected, there is some positive correlation: the more classes a student attends, the more likely they are to score high, or at least pass.
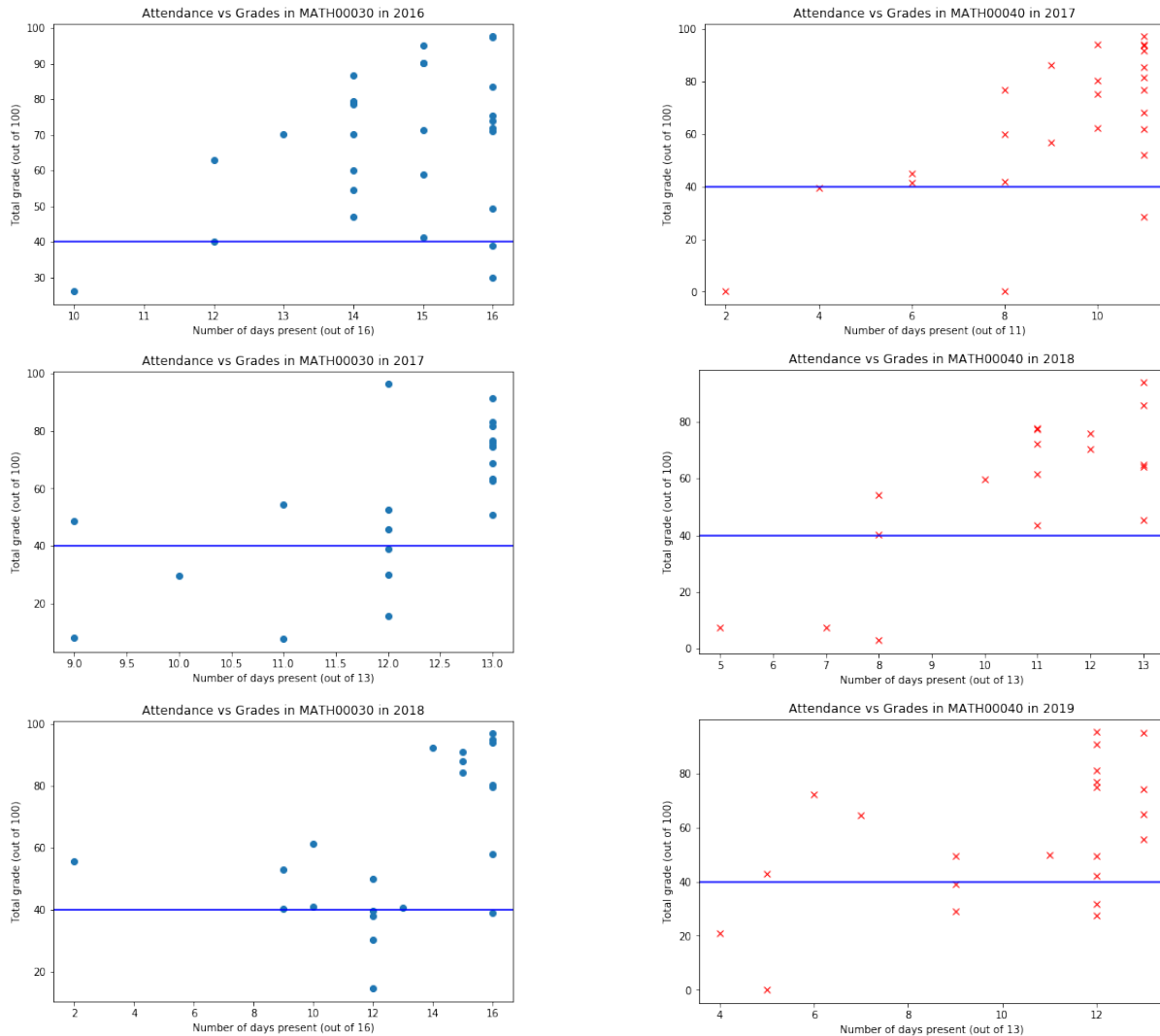
Figure 2: Put your caption here

Summary statistics to include in this section: Correlation coefficient (r), mean of each data set, pass rate,

The correlation coefficients (in order):

$r = 0.3587$

$r = 0.6707$

$r = 0.6480$

$r = 0.7923$

$r = 0.4922$

$r = 0.5254$

Investigating short sessions

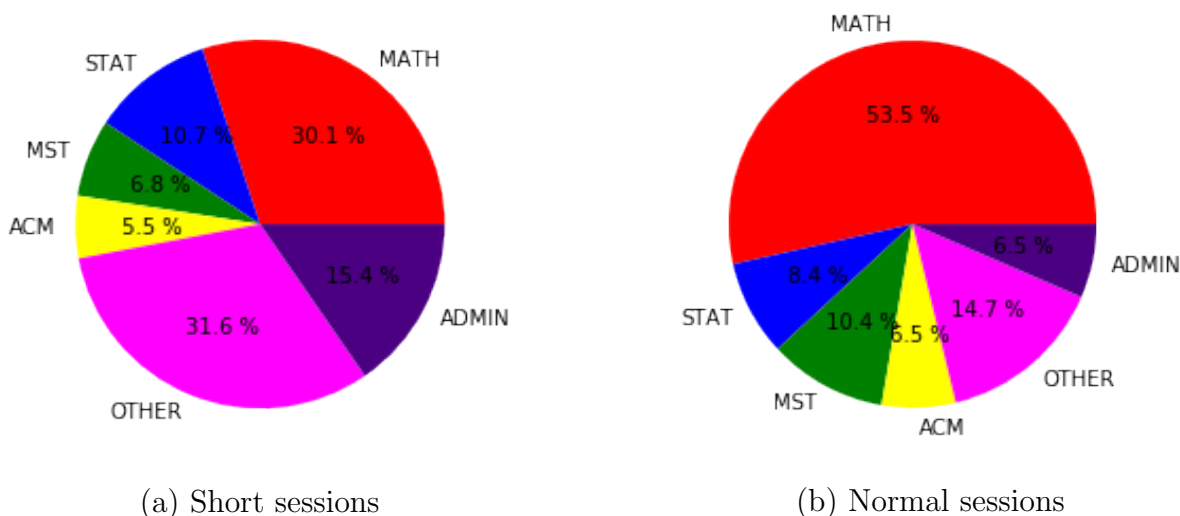(a) Short sessions   (b) Normal sessions

Figure 3: Breakdown by modules

This data includes September 2018 to present. This is when the MSC introduced the 'ADMIN' module for tutors to use. This is to indicate where the student left without being seen.

Apart from the ADMIN module, I looked at the other modules that were frequently listed as 0 or 1 minute sessions. It can be assumed that sometimes these are a mistake. A tutor simply forgets to log a student into the system at the appropriate time, and

Here we see an over representation in the 'Other' modules, that is, modules outside the school of maths. The general claim is that these are students from a biology, chemistry or arts course and come in about a very quick query that usually takes a short amount of time.

We can see that the stats and other modules are over represented in the ultra short sessions. Regarding the 'other' modules, that is, these are the modules that are not in the school of maths - students might just need a small bit of advice on their assignment that naturally just takes a very short amount of time; reading a graph, or checking a small calculation etc.

However, the MSC is known to have an under representation of statistics tutors, leading to students arriving to seek help only to find out there's no one qualified enough to help them with their query.

6 Attendance modelling with SIR

Now that we know certain effects of an MSC, and we've studied a particularly involved group within the centre, a good question to ask is: **can we predict attendance?**

The SIR model stands for Susceptible, Infected, Recovered and it is primarily used to predict outcomes of a disease spreading throughout a population.

$$\frac{dS}{dt} = -\beta SI$$

$$\frac{dI}{dt} = \beta SI - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$

However, it can also be used to model information spread or rumours, and thus it can be hypothesised that it can be used to predict information spread. (Daley Kendall, 1964)

The model

Researchers in Melbourne University and Swinburne University of Technology (also in Melbourne) have constructed a model inspired by the SIR to model information spread through word of mouth

$$\frac{dI}{dt} = \alpha(t)S + \beta(t)S\frac{I}{N}$$

- $I(t) =$ Cumulative number of students at day t

- $S(t) =$ Potential MSC users who are unaware of the service

- $N = S + I$

- $\alpha(t) = ae^{-bt}$ or sometimes constant

- β constant

Here are two diagrams modelling two consecutive semesters in UCD. Here, the $\alpha(t)$ parameter is an exponential function. As we see, the model is a good fit!

This also follows the behaviour of Swinburne's MSC very closely...

Swinburne data
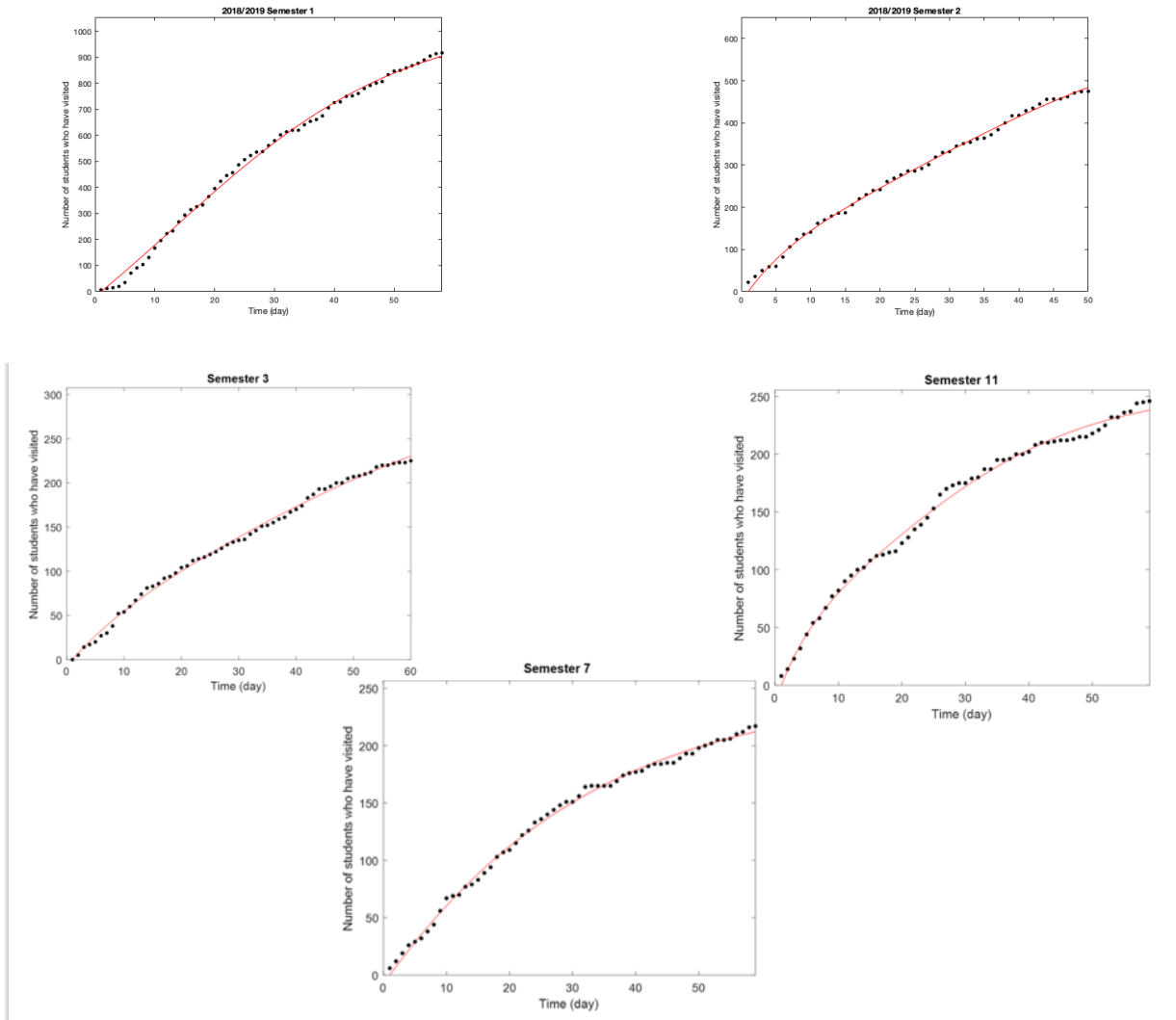
7 Conclusions

Conclusions

- While it is unclear from one sample that the MSC has a positive effect on grades, it is encouraging to see that many students do partake in the service, meaning that engagement is present.

- Now that we know how to predict attendance, a good next step would be to find a way to increase the effectiveness of advertising, to reach the failing students.

- With the model we have, we now have a proxy for estimating how many numbers we *should* be expecting throughout a semester.

8   References

References

[1] MacGillivray† Cuthbert. "Investigation of completion rates of Engineering students". In: *Delta 2007* (2007), p. 8.

[2] Ant Edwards and David Carroll. "Predictors of demand for mathematics support". In: *PME 42 Proceedings* (2018).

[3] Maria Jacob and Eabhnat Nı Fhloinn. "A quantitative, longitudinal analysis of the impact of mathematics support in an Irish university". In: *Teaching Mathematics and its Applications: An International Journal of the IMA* 38.4 (2019), pp. 216–229.

[4] Godfrey Pell and Tony Croft. "Mathematics support—support for all?" In: *Teaching mathematics and its applications* 27.4 (2008), pp. 167–173.