Attendance Monitoring Visualisation

Interim Report

Daniel Bates - Candidate 234558
Supervised by Dr Paul Newbury

Contents

1 Introduction

1.1 Problem Area

Attendance and attainment in higher education has a proven link, and several studies have found a correlation between the two[1][2][3][4]. In the below figure, though the axis are somewhat unclear, a clear correlation between attendance and attainment can be seen.

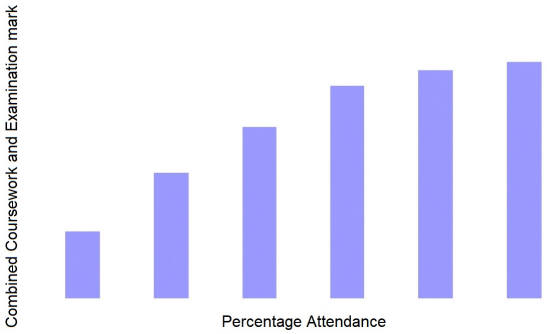


Figure 1.1: Class of combined coursework and examination mark versus percentage attendance at teaching sessions for those submitting work. (n=153); From a Computing module at the University of Central England, Birmingham[1].

There is a large task found by school administrators in monitoring attendance in order to provide support for students, as this is often for whole schools. As of 2020/21, the University of Sussex had 18,510 full-time students and 10 Schools[5]. With an approximate 18,510 full-time students, and assuming one staff member monitoring attendance per school, we can get a rough idea of the number of students monitored by an individual:

$$\frac{18510}{10} = 1851$$

This disparity is only going to increase: nationwide university uptake is only increasing in this and recent years[6], and Sussex experiences this as much as other universities[5].
This also ignores that in monitoring attendance, staff want to encourage students to improve it during the semester, not after. So regular updates or snapshots will be wanted in order to check in at several points in the semester.
Therefore, with multiple checks per semester, almost 2000 students each, and

with most of these staff taking on other roles like Student Experience, this problem area is ripe for automation and likely in need of useful visualisation for the sake of users (relevant staff), as well as to enable users to identify and support students who are not engaging, and therefore improve student welfare.

1.2 Objectives

This project aims to produce meaningful and actionable visualisations from sample data (though if put into production, will need to be suited for live data), as well as allowing for easily configurable filters on this data that allow staff to produce their own such visualisations.

1.2.1 Primary Objectives

1. Import attendance data into an effective and efficient database
2. Construct a well-structured and efficient database to store attendance data, and import sample attendance data
3. Create meaningful visualisations of attendance statistics for use by relevant staff
   (a) Per Student
   (b) Per Degree
   (c) Per Stage
   (d) Per Department & School
4. Get user feedback from relevant staff: in this case our Director of Student Experience, Dr Kate Howland, who due to her background in Interaction Design and as a user may be an asset in terms of specific feedback.
5. Apply this user feedback to improve visualisations and the system at large
6. Allow users to apply filters to data and create visualisations with this dataset

1.2.2 Extensions

1. Create a full-stack web application around the visualisations, featuring

   (a) A ReactJS front-end
   (b) A Python (Flask) back-end
   (c) PostgreSQL Database
2. Make the application user configurable, including different dashboards of visualisations
3. Investigate further ways to visualise attendance data e.g. Nightingale Rose Charts
4. Find meaningful statistics for student engagement beyond just quantity

1.3 Motivations

This project sets out to research and deliver useful attendance visualisations to aid staff (primarily at the University of Sussex, but this could be extended beyond) concerned with attendance.
As it is based around a web application, database, and data analysis, this project closely relates to my degree (Computer Science Integrated Masters) and has potential to inform decisions for my masters year, as well as my preferred career in Software Engineering, as that much like my project, would include plenty of project and time management opportunities, as well as design around users which is currently being informed by the Human-Computer Interaction module.
This project will allow me to test and improve my skills in Databases following the module in second year, revise and extend my statistics knowledge from Mathematical Concepts, and investigate the domain of Data Science.
In addition, pursuing this project has potential to benefit the School of Engineering and Informatics if visualisations and the final application prove genuinely useful in the long term for attendance monitoring.

Overview

The rest of this report details my Professional Considerations (2), Requirements Analysis and background research (3), Requirements Specification (4), and last, my Project Plan (5). My User Testing Compliance form is attached in the appendix (A).

2 Professional and Ethical Considerations

2.1 BCS Code of Conduct

Sections of the BCS Code of Conduct[7] relevant to this project are:

1.a. have due regard for public health, privacy, security and wellbeing of others and the environment.

All data used in this project will be anonymised and is already held by the University in accordance with data protection legislation. This project may in future be used to identify real at-risk students (as allowed by the University through agreements made with students during admission), however no current non-anonymised students will be identified in the course of this project. The project will include user testing by a relevant member of staff, as well as my supervisor, see 2.2.

1.b. have due regard for the legitimate rights of Third Parties.

This project does not engage third parties, but will make significant effort to avoid misuse of student data albeit anonymised. Any and all third-party code and libraries will be properly acknowledged.

1.c conduct your professional activities without discrimination on the grounds of sex, sexual orientation, marital status, nationality, colour, race, ethnic origin, religion, age or disability, or of any other condition or requirement.

No protected characteristics will be exposed in the sample data, as it has been anonymised. Therefore this project cannot and will not discriminate on any of these grounds.

1.d promote equal access to the benefits of IT and seek to promote the inclusion of all sectors in society wherever opportunities arise.

Effort will be made to accommodate screen-reading technologies in HTML to make the application accessible to those affected by blindness, however this will be limited to text such as tables as visualisations like graphs are unlikely to translate well. There is no auditory element planned for the application so it is unlikely to need adaptation for deaf users, and if the user interface is intuitively made, unlikely to need adaptations for users with motor disabilities either.

2.d ensure that you have the knowledge and understanding of Legislation and that you comply with such Legislation, in carrying

out your professional responsibilities.

I understand data protection regulations, and will not make any decisions that may impose upon an individuals right to data protection and security without due consultation with my supervisor and any other relevant persons.

2.e. respect and value alternative viewpoints and, seek, accept and offer honest criticisms of work.

User testing and supervisor feedback will be presented in full, and any alternative viewpoints will be considered.

2.g. reject and will not make any offer of bribery or unethical inducement.

There will be no incentives and no knowledge withheld from users during user feedback. See 2.2 for more.

3.d NOT disclose or authorise to be disclosed, or use for personal gain, or to benefit a third party, confidential information except with the permission of your Relevant Authority, or as required by Legislation.

All information given to me for use in this project will not be shared under any circumstances, except for supervision.

2.2 User Testing

It has been determined that this project would comply with all points on the User Testing Compliance Form[8], and the completed form can be found in Appendix A. The following section acknowledges the 12 points of the User Testing Compliance Form and gives reason why this project complies with each:

1. Participants were not exposed to any risks greater than those encountered in their normal working life.

Due to the nature of the project, what is shown to participants is certain to be a fairly simple interface on a web-page, which is not a risk greater than most people's normal working lives.

2. The study materials were paper-based, or comprised software running on standard hardware.

Materials will be software running on standard hardware (likely a networked computer in Chichester I, or my personal laptop).

3. All participants explicitly stated that they agreed to take part, and that their data could be used in the project.

I will require participants to sign an informative consent form before taking part, notifying them that their feedback will only be used to improve the project, and will not be shared or published.

4. No incentives were offered to the participants.

Participants will not be offered any incentives.

5. No information about the evaluation or materials was intentionally withheld from the participants.

The introductory script will inform the participants everything relevant to the evaluation of the project.

6. No participant was under the age of 18.

I will not engage participants under the age of 18.

7. No participant had a disability or impairment that may have limited their understanding or communication or capacity to consent.

I will not engage participants with a limited understanding or capability to consent to participating in user testing.

8. Neither I nor my supervisor are in a position of authority or influence over any of the participants.

Both I and my supervisor will only seek out participants we are on an equal hierarchical footing with (Student-Student, Staff-Staff only).

9. All participants were informed that they could withdraw at any time.

All participants will be informed of this in the introduction script, and consent form.

10. All participants have been informed of my contact details, and the contact details of my supervisor.

All participants will be informed of this in the introduction and debriefing scripts, consent form, as well as in any introductory email.

11. The evaluation was described in detail with all of the participants at the beginning of the session, and participants were fully debriefed at the end of the session. All participants were given the opportunity to ask questions at both the beginning and end of the session.

The evaluation will be described in detail in the introduction and debriefing scripts, and it will be made clear that questions are welcome both at the delivery of the script, and at the end of the session.

12. All the data collected from the participants is stored securely, and in an anonymous form.

User feedback will be anonymised and stored on the University of Sussex's OneDrive shortly after the session.

3 Requirements Analysis

3.1 Background Research

3.1.1 Similar Systems

Finding similar systems beyond similar projects of the same description is difficult: many different systems are in use and most at universities are bespoke implementations, whereas similar solutions in use at Secondary Schools are often whole packages and closed-source. Therefore, while finding some examples of similar systems is possible, they are only really useful for aesthetic design inspiration.

Panasonic AVS ('Attendance Visualization System')[9] is a solution from Panasonic System Networks, which features real time visualisation in table form, as well as exportable csv reports. This implementation is somewhat limited by only having one visualisation and designed for employers, possibly in some sort of production line.

While it is an older piece of software, it's a good point of comparison as it is so different - my users would have cohorts much larger than the number of active employees in AVS, and though clocking in/out isn't something a student does, their attendance in various sessions could be visualised similarly. A tabular view similar to the figure3.1 would be useful, but only when it contains data filtered by a particular parameter.

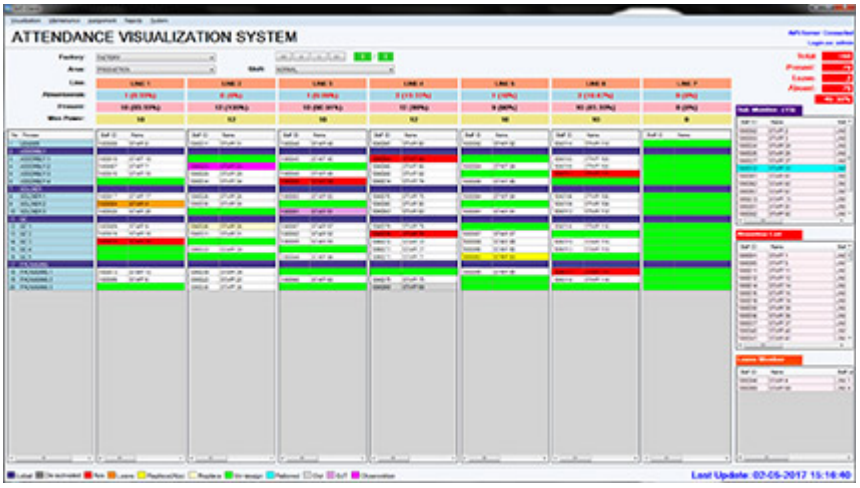
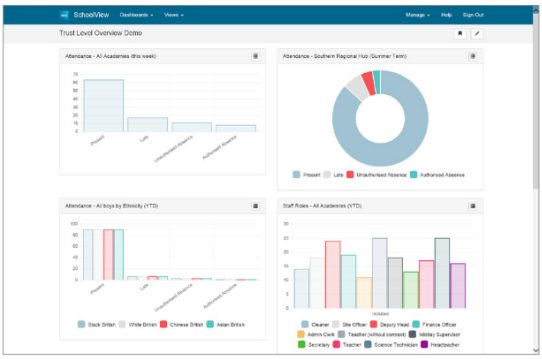


Figure 3.1: Panasonic's AVS, in a tabular view, from their promotional material[9]

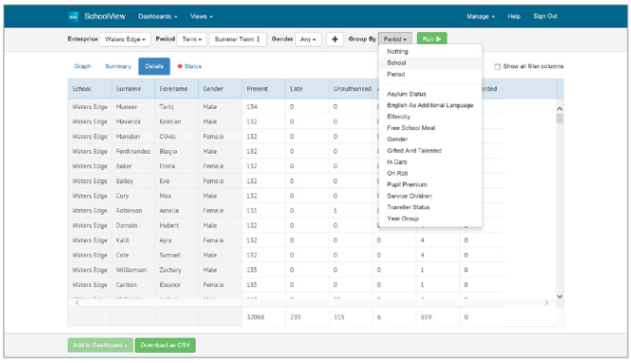
.

SIMS SchoolView [10] is a bit closer to what this project seeks to achieve. SIMS is a modular school management system used primarily by Primary and Secondary schools. While the needs of these users are different to that of a University School, it's still very useful to see some examples of features I will want to incorporate into this project. It exposes both student-level data, filterable by multiple different attributes, as well as useful pre-configured dashboards for student and staff attendance 3.2.

For this project I would use similar dashboards but ideally they would be somewhat user-configurable as user needs are likely to change over time. I would also like to incorporate filterable table data similar to the Pupil and Staff view.



(a) Overview Dashboard



(b) Pupil and Staff Level data view

Figure 3.2: Two examples of SIMS SchoolView[10], from their promotional material.

Sutjarittham et al: 'Tool to Access and Visualize Classroom Attendance Data from a Smart Campus' [11] is a tool that was demonstrated at the

2018 17th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN). It was part of a research effort to investigate optimisation of classroom usage at the University of New South Wales (Sydney)[12].

While it was aimed toward optimising use of teaching spaces rather than student welfare, it's still very useful as an academic reference and for its good use of visualisations, as well as the web-based interface they have used.

3.1.2 Technologies

My chosen technology stack is a combination of ReactJS[13], Python for both a Flask[14] web server (API) and to allow for easy manipulation of data, as well as a PostgreSQL[15] database.

ReactJS is a JavaScript (node.js) library best known for the ease it applies to building user interfaces, as well as its multi-platform support via React Native, and plenty of packages that extend its capabilities. For these reasons it is a good choice for the user interface, and visualisations especially.

This is why I intend to use React for my front-end presentation and logic. However it cannot interact with a database on its own, which leads us onto the next level of my technology stack...

Flask is a Python Web Framework that is simple and easy to extend, which makes it perfect for my purposes as a middleman for React and PostgreSQL and there are plenty of existing implementations interfacing Flask with each.

Writing the web-server/API in Python also allows me more flexibility when it comes to data manipulation using libraries like pandas, as well as the possibility of training a predictive model on the data.

PostgreSQL is a object-relational database software/server, known for its feature set over the competition and wide use in web applications. In particular, I've chosen it over say MySQL for it's ability to handle multiple encrypted sessions at once, which could prove useful in the eventuality my API (flask server) is overrun with requests. I've also selected Postgres (as it's commonly called) for its good community support as this will enable me to more easily deal with issues as the crop up.

All three technologies are well known in the field (so finding any issues should be fairly straightforward), and known inter-operate.

4 Requirements Specification

4.1 Mandatory Requirements

R1 Construct a well-structured database to store attendance data, and import sample attendance data into it. The fields corresponding roughly to the current excel document headers (see Appendix B):

- User becoming the Primary Key for the 'students' table which contains from User to Course Code
- Level of Study being flattened to a Boolean 'isUG', True meaning Undergraduate, False meaning Postgraduate Taught
- Course Title can be moved to a separate table 'course' that maps it to the Course Code (Primary Key).
- Everything after Course Code will repeat each snapshot, so a table 'snapshots' with composite primary key User and 'Snapshot Date' (a Date variable corresponding to the date of insertion) will be used to differentiate them.
- All percentage fields can be dropped as they can be easily re-calculated from their associated values.
- Further steps may need to be taken to normalise the database as the project goes on.

R2 Write code that presents 2D Line graphs/histograms of session attendance (as a percentage of sessions that have been attended over sessions that could have been attended) over time for the following groups:

1. Each Individual
2. Each Degree and Stage (e.g. Computer Science - Year 3)
3. Each Degree
4. Each Stage
5. Each Department (group degrees to achieve this)
6. The School as a whole

R3 Create a tabular view that requires at least one filter set with 'at risk students' preset, using a suitable threshold percentage of missed sessions (e.g. below 40% attendance).

R4 Perform User Testing with my supervisor, as well as the school's Director of Student Experience, if possible.

R5 Create a three dimensional graph including attended sessions, time and stage, to visualise the 'fabric' of attendance for the school

R6 Apply group filters from R2 to other fields where there is enough data to meaningfully show a trend, such as assessments turned in on time (of assessments turned in), and academic advising sessions attended of sessions scheduled.

R7 Construct a well-structured back end (API for the database) to facilitate the other requirements data needs, with modularity and efficiency as focuses.

4.2 Desirable Requirements

E1 Create a script to enable easy bulk data insertion from an excel document

E2 Investigate applying additional visualisations, including a Nightingale Rose (as it's good for changing magnitude over time), and a Bubble Chart (change in a value over time, with a magnitude included in the size of the bubbles).

E3 Create a Dashboard, housing useful visualisations for use by Attendance Monitoring staff

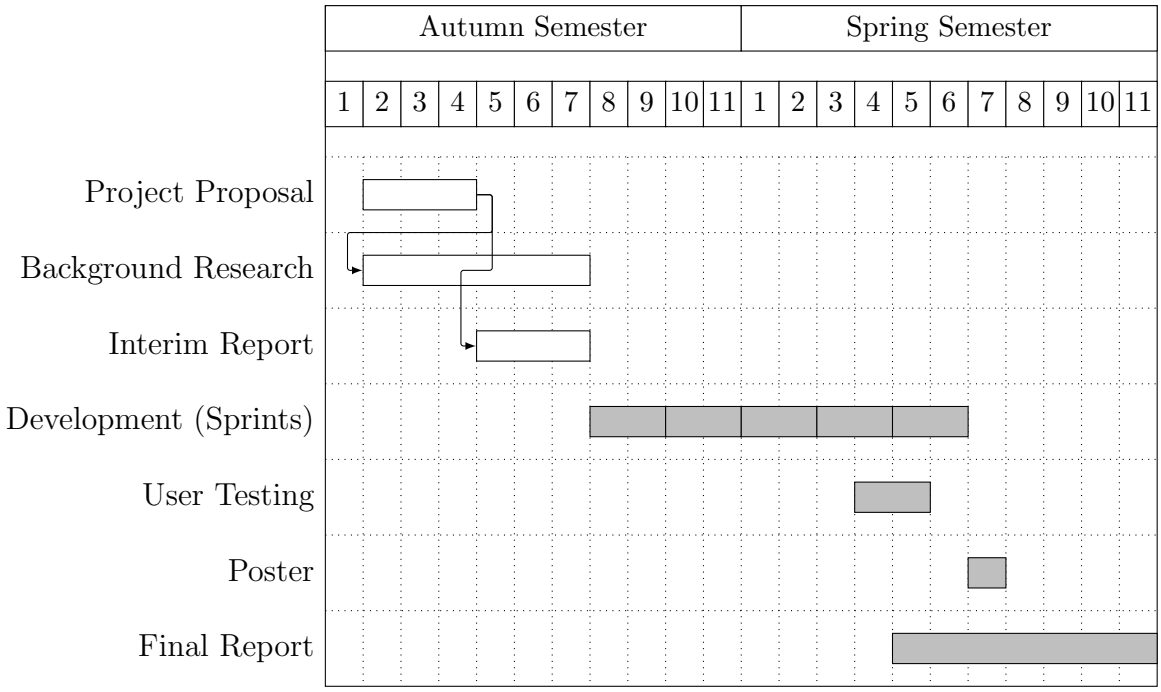
E4 When additional data is added to the system, and new at-risk students are found, alert the user

E5 Investigate additional filters beyond R2

E6 Investigate if a machine learning model (from SciKit-Learn or similar libraries) can successfully predict later attendance based on earlier attendance, or likeliness of submission of assessments.

5 Project Plan

I will employ the use of sprints when developing the software side of my project, which will coincide with User Testing so that any feedback can be considered and acted upon.



5.1 Tasks

- **Project Proposal**
- Background Research - Background reading performed mainly for the Interim Report.
- **Interim Report**
- Development - All programming task e.g. experimenting with visualisations, implementing frontend and backend systems, and a database.
- User Testing - A period to facilitate user testing of the system in a mostly finished state, and gain additional constructive user feedback.
- **Poster**
- **Final Report**

(Deliverables in **Bold**)

References

[1] John Colby. Attendance and Attainment - a comparative study. *Innovation in Teaching and Learning in Information and Computer Sciences*, 4(2):1–13, May 2005. Publisher: Routledge _eprint: https://doi.org/10.11120/ital.2005.04020002.

[2] Loretta Newman-Ford, Karen Fitzgibbon, Stephen Lloyd, and Stephen Thomas. A large-scale investigation into the relationship between attendance and attainment: a study using an innovative, electronic attendance monitoring system. *Studies in Higher Education*, 33(6):699–717, December 2008. Publisher: Routledge _eprint: https://doi.org/10.1080/03075070802457066.

[3] Bernard Meulenbroek and Maartje van den Bogaard. Attendance and attainment in a Calculus course. *European Journal of Engineering Education*, 38(5):532–542, October 2013. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/03043797.2013.811478.

[4] Gabrielle E. Kelly. Lecture attendance rates at university and related factors. *Journal of Further and Higher Education*, 36(1):17–40, February 2012. Publisher: Routledge _eprint: https://doi.org/10.1080/0309877X.2011.596196.

[5] Facts and figures : Rankings and figures : About us : University of Sussex. `https://www.sussex.ac.uk/about/facts/facts-figures`. [Online; accessed 31-10-2022].

[6] Record numbers of students take up university places. `https://www.gov.uk/government/news/record-numbers-of-students-take-up-university-places`. [Online; accessed 31-10-2022].

[7] BCS Code of Conduct for members - Ethics for IT professionals | BCS. [Online; accessed 05-11-2022].

[8] Research Governance : Research : School of Engineering and Informatics : Schools and services : University of Sussex. `https://www.sussex.ac.uk/ei/internal/research/researchgov`. [Online; accessed 26-10-2022].

[9] Attendance Visualization System - Panasonic System Networks Malaysia Sdn. Bhd. (42154-T) [PSNM]. `https://snm.my.panasonic.com/attendance-visualization-system`. [Online; accessed 01-11-2022].

[10] Sims schoolview. https://www.ess-sims.co.uk/sites/default/files/2021-06/SIMS%20SchoolView%20-%20June%202021.pdf.

[11] Thanchanok Sutjarittham, Hassan Habibi Gharakheili, Salil Kanhere, and Vijay Sivaraman. Demo Abstract: A Tool to Access and Visualize Classroom Attendance Data from a Smart Campus. In *Conference: 2018 17th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pages 140–141, April 2018.

[12] Thanchanok Sutjarittham, Hassan Habibi Gharakheili, Salil S. Kanhere, and Vijay Sivaraman. Data-driven monitoring and optimization of classroom usage in a smart campus. In *Proceedings of the 17th ACM/IEEE International Conference on Information Processing in Sensor Networks*, IPSN '18, pages 224–229, Porto, Portugal, April 2018. IEEE Press.

[13] React – A JavaScript library for building user interfaces.

[14] Welcome to Flask — Flask Documentation (2.2.x).

[15] PostgreSQL Global Development Group. PostgreSQL, November 2022.

Appendix A User Testing Compliance Form

User Testing Compliance Form for UG and PGT Projects*
School of Engineering and Informatics
University of Sussex

This form should be used in conjunction with the document entitled "Research Ethics Guidance for UG and PGT Projects".

Prior to conducting your project, you and your supervisor will have discussed the ethical implications of your research. If it was determined that your proposed project would comply with **all** of the points in this form, then both you and your supervisor should complete and sign the form on page 3, and submit the signed copy with your final project report/dissertation.

If this is not the case, you should refer back to the "Research Ethics Guidance for UG and PGT Projects" document for further guidance.

---

1. Participants were not exposed to any risks greater than those encountered in their normal working life.

   Investigators have a responsibility to protect participants from physical, mental and emotional harm during the investigation. The risk of harm must be no greater than in ordinary life. Areas of potential risk that require ethical approval include, but are not limited to, investigations that require participant mobility (e.g. walking, running, use of public transport), unusual or repetitive activity or movement, physical hazards or discomfort, emotional distress, use of sensory deprivation (e.g. ear plugs or blindfolds), sensitive topics (e.g. sexual activity, drug use, political behaviour, ethnicity) or those which might induce discomfort, stress or anxiety (e.g. violent video games), bright or flashing lights, loud or disorienting noises, smell, taste, vibration, or force feedback.

2. The study materials were paper-based, or comprised software running on standard hardware.

   Participants should not be exposed to any risks associated with the use of non-standard equipment: anything other than pen-and-paper, standard PCs, mobile phones, and tablet computers is considered non-standard.

3. All participants explicitly stated that they agreed to take part, and that their data could be used in the project.

   Participants cannot take part in the study without their knowledge or consent (i.e. no covert observation). Covert observation, deception or withholding information are deemed to be high risk and require ethical approval through the relevant C-REC.

---

*This checklist was originally developed by Professor Steven Brewster at the University of Glasgow, and modified by Dr Judith Good for use at the University of Sussex with his permission.

If the results of the evaluation are likely to be used beyond the term of the project (for example, the software is to be deployed, the data is to be published or there are future secondary uses of the data), then it will be necessary to obtain signed consent from each participant. Otherwise, verbal consent is sufficient, and should be explicitly requested in the introductory script (see Appendix 1).

4. No incentives were offered to the participants.
   The payment of participants must not be used to induce them to risk harm beyond that which they risk without payment in their normal lifestyle. People volunteering to participate in research may be compensated financially e.g. for reasonable travel expenses. Payments made to individuals must not be so large as to induce individuals to risk harm beyond that which they would usually undertake.

5. No information about the evaluation or materials was intentionally withheld from the participants.
   Withholding information from participants or misleading them is unacceptable without justifiable reasons for doing so. Any projects requiring deception (for example, only telling participants of the true purpose of the study afterwards so as not to influence their behaviour) are deemed high risk and require approval from the relevant C-REC.

6. No participant was under the age of 18.
   Any studies involving children or young people are deemed to be high risk and require ethical approval through the relevant C-REC.

7. No participant had a disability or impairment that may have limited their understanding or communication or capacity to consent.
   Projects involving participants with disabilities are deemed to be high risk and require ethical approval from the relevant C-REC.

8. Neither I nor my supervisor are in a position of authority or influence over any of the participants.
   A position of authority or influence over any participant must not be allowed to pressurise participants to take part in, or remain in, any study.

9. All participants were informed that they could withdraw at any time.
   All participants have the right to withdraw at any time during the investigation. They should be told this in the introductory script (see Appendix 1).

10. All participants have been informed of my contact details, and the contact details of my supervisor.
    All participants must be able to contact the investigator and/or the supervisor after the investigation. They should be given contact details for both student and supervisor as part of the debriefing.

11. The evaluation was described in detail with all of the participants at the beginning of the session, and participants were fully debriefed at the end of the session. All participants were given the opportunity to ask questions at both the beginning and end of the session.

    Participants must be provided with sufficient information prior to starting the session, and in the debriefing, to enable them to understand the nature of the investigation.

12. All the data collected from the participants is stored securely, and in an anonymous form.

    All participant data (hard-copy and soft-copy) should be stored securely (i.e. locked filing cabinets for hard copy, password protected computer for electronic data), and in an anonymised form.

Project title: Attendance Monitoring Visualisation

Student's Name: Daniel Bates

Student's Registration Number: 22004544

Student's Signature: _____

Date: _____

Supervisor's Name: Dr Paul Newbury

Supervisor's Signature: _____

Date: _____

Appendix B Existing Excel Columns

Fields currently found in the sample excel document.

- User
- Level of Study
- Year of Course
- Registration Status
- Course Title
- Course Code
- Teaching Sessions
- Attended
- Explained Absences
- Non Attendance
- % Attendance
- % Attendance (Unexcused)
- Last Attendance
- Assessments Submitted
- Explained Non-Submission
- Non Submission
- Within Late Period Flag
- % Submitted
- Last Submitted
- Academic Advising Sessions
- Attended (AA)
- Explained Non Attendances (AA)
- Non Attendances (AA)
- Attendance Not Recorded (AA)
- % Attendance (AA)

- Last Attended (AA)