# TEMPLATE FOR A DATA MANAGEMENT PLAN

The following **template** should be used to develop a Data Management Plan (DMP) to accompany a research proposal. The notes (*in italics*) provide further context and guidance for its completion. Where substantial data is generated from the research, the DMP will be more in depth and therefore likely to be 2 or 3 pages long [(3 pages maximum length - See MRC Je-S Help and Guidance for DMP)](https://je-s.rcuk.ac.uk/Handbook/pages/GuidanceonCompletingaStandardG/CaseforSupportandAttachments/MRCSpecificGuidance.htm#Data_Management_Plan__exactly_1__Mandatory_requirement___Maximum_of_1_DMP) for low impact studies generating small amounts of data, DMPs will be short ie less than half a page.

If you opt NOT to use the template the topics listed in the template MUST be addressed.

|  |  |
| --- | --- |
| **0. Proposal name** | |
| Compare the ocular behaviour of novices and visualization experts | |
| **1. Description of the data** | |
| **1.1 Type of study**  Study the ocular behaviour between novices and experts for financial documents, specifically income statements from financial reports.  **1.2 Types of data**  The primary data that will be handled in this study is quantitative data, specifically collected during the participant's examination of a data visual using eye tracking technology.  Prior to the study, participants will complete a survey to record their expertise in data visualisation and financial data. Additionally, a post-study survey will be conducted to assess the usefulness of data visualization for understanding income statements.  **1.3 Format and scale of the data**  The eye gaze data collected during the experiment will consist of x and y coordinates for each eye, recorded as tuples. This means that for every position the participant looks at, there will be four data points in total: two tuples for each eye. The Tobii X2-30 eye tracker used in this study has a sampling rate of 30 Hz, meaning that 30 data points will be collected every second. As the experiment duration is limited to 1 minute, a maximum of 1800 data points will be generated for each run. | |
| **2. Data collection / generation** | |
| **2.1 Methodologies for data collection / generation**  The data for this study will be collected via 2 means, survey collection from the participants and the eye tracking data. To control the Tobii eye tracker, a Python script running on Python version 3.8 will be used. The script will automatically terminate the eye tracking data collection once the designated time limit has been reached.  The data will be stored on a CSV and the raw data will be moved to an R environment which will then be processed using R.  With regards to the survey, no identifiable information will be asked as part of this survey.  **2.2 Data quality and standards**  The quality of data will be controlled by the same python script and eye tracker being used as part of this study. This will ensure the only variations will be from the physical environment such as lighting environment and participant. The eye tracker and the display will be set up using the same laptop to ensure the same device is used throughout. | |
| **3. Data management, documentation and curation** | |
| **3.1 Managing, storing and curating data.**  The eye gaze data will first be collected on the working laptop and then pushed and stored on a GitHub repository. When the data needs to be analysed, an R environment will be created and turned into a data frame meaning the raw data will remain untouched.  The survey data will be collected on the day of the experiment and will be saved in the repository. Once the data has been extracted from within, the surveys will be removed.  **3.2 Metadata standards and data documentation**  The methodology to gathering gaze data, code and hardware will be defined in the main report. This will allow for the procedure to be recreated. Data visualisations and tables will also be part of the report which will allow for comparisons in any future work.  **3.3 Data preservation strategy and standards**  Once the study has concluded, all raw data will be removed from the GitHub repository. | |
| **4. Data security and confidentiality of potentially disclosive information** | |
| **4.1 Formal information/data security standards**  *n/a*  **4.2 Main risks to data security**  Confidentiality is not a concern since no identifiable information will be captured or stored as part of the study. | |
| **5. Data sharing and access** | |
| Identify any data repository (-ies) that are, or will be, entrusted with storing, curating and/or sharing data from your study, where they exist for particular disciplinary domains or data types. [Information on repositories is available here.](http://www.wellcome.ac.uk/About-us/Policy/Spotlight-issues/Data-sharing/Guidance-for-researchers/WTX060360.htm)  **5.1 Suitability for sharing**  The data is suitable for sharing as it contains onto eye gaze data. Survey data captured as part of this study will not be released, only summarised at a high level.  **5.2 Discovery by potential users of the research data**  If the outcome of this study is successful, an outcome of this will be to provide guidelines on creating data visualisations for income statements. This will allows for future users to search by financial data visualisation which the report and code will be freely available on GitHub. The sharing of gaze data will not be shared although the code, visualisation of gaze data and high level analysis of the participants will be shared.  **5.3 Governance of access**  Only the main principal investigator will be able to make the decision to share the research data with new users. As mentioned the eye gaze data will be stored on a GitHub repository and removed once the study has come to a conclusion.  **5.4 The study team’s exclusive use of the data**  *MRC’s requirement is for timely data sharing, with the understanding that a limited, defined period of exclusive use of data for primary research is reasonable according to the nature and value of the data, and that this restriction on sharing should be based on simple, clear principles*. *What are the timescale/dependencies for when data will be accessible to others outside of your team? Summarize the principles of your current/intended policy.bg*  **5.5 Restrictions or delays to sharing, with planned actions to limit such restrictions**  *Restriction to data sharing may be due to participant confidentiality, consent agreements or IPR. Strategies to limit restrictions may include data being anonymised or aggregated; gaining participant consent for data sharing; gaining copyright permissions. For prospective studies, consent procedures should include provision for data sharing to maximise the value of the data for wider research use, while providing adequate safeguards for participants. As part of the consent process, proposed procedures for data sharing should be set out clearly and current and potential future risks associated with this explained to research participants.*  **5.6 Regulation of responsibilities of users**  *Indicate whether external users are (will be) bound by* [*data sharing agreements*](https://www.mrc.ac.uk/publications/browse/mrc-policy-and-guidance-on-sharing-of-research-data-from-population-and-patient-studies/)*, setting out their main responsibilities (please see page 13 section 7, titled* [*Data-sharing agreements*](https://www.mrc.ac.uk/publications/browse/mrc-policy-and-guidance-on-sharing-of-research-data-from-population-and-patient-studies/) *of the PDF file generated by selecting either of two links above).* | |
| **6. Responsibilities** | |
| *Apart from the PI, who is responsible at your organisation/within your consortia for:*   * *study-wide data management* * *metadata creation,* * *data security* * *quality assurance of data.* | |
| **7. Relevant institutional, departmental or study policies on data sharing and data security** | |
| *Please complete, where such policies are (i) relevant to your study, and (ii) are in the public domain, e.g. accessible through the internet.*  *Add any others that are relevant* | |
| **Policy** | **URL or Reference** |
| Data Management Policy & Procedures |  |
| Data Security Policy |  |
| Data Sharing Policy | *e.g. a* [*study policy of sharing research data*](https://www.mrc.ac.uk/publications/browse/mrc-policy-and-guidance-on-sharing-of-research-data-from-population-and-patient-studies/) |
| Institutional Information Policy |  |
| Other: |  |
| Other |  |
| **8. Author of this Data Management Plan (Name)** and, if different to that of the Principal Investigator, their **telephone & email contact details** | |
|  | |