

University of Glasgow  
School of Mathematics and Statistics  
Data Analytics MSc (ODL)

---

## **Project**

### **Data Summary**

Carlos Yanez Santibanez

---

15 October 2022

## 1 Project Plan

The project outline has been organised in “sprints” to approximately match the meeting schedule, with the aim to make the most of the supervisor’s input. Each sprint will aim to :

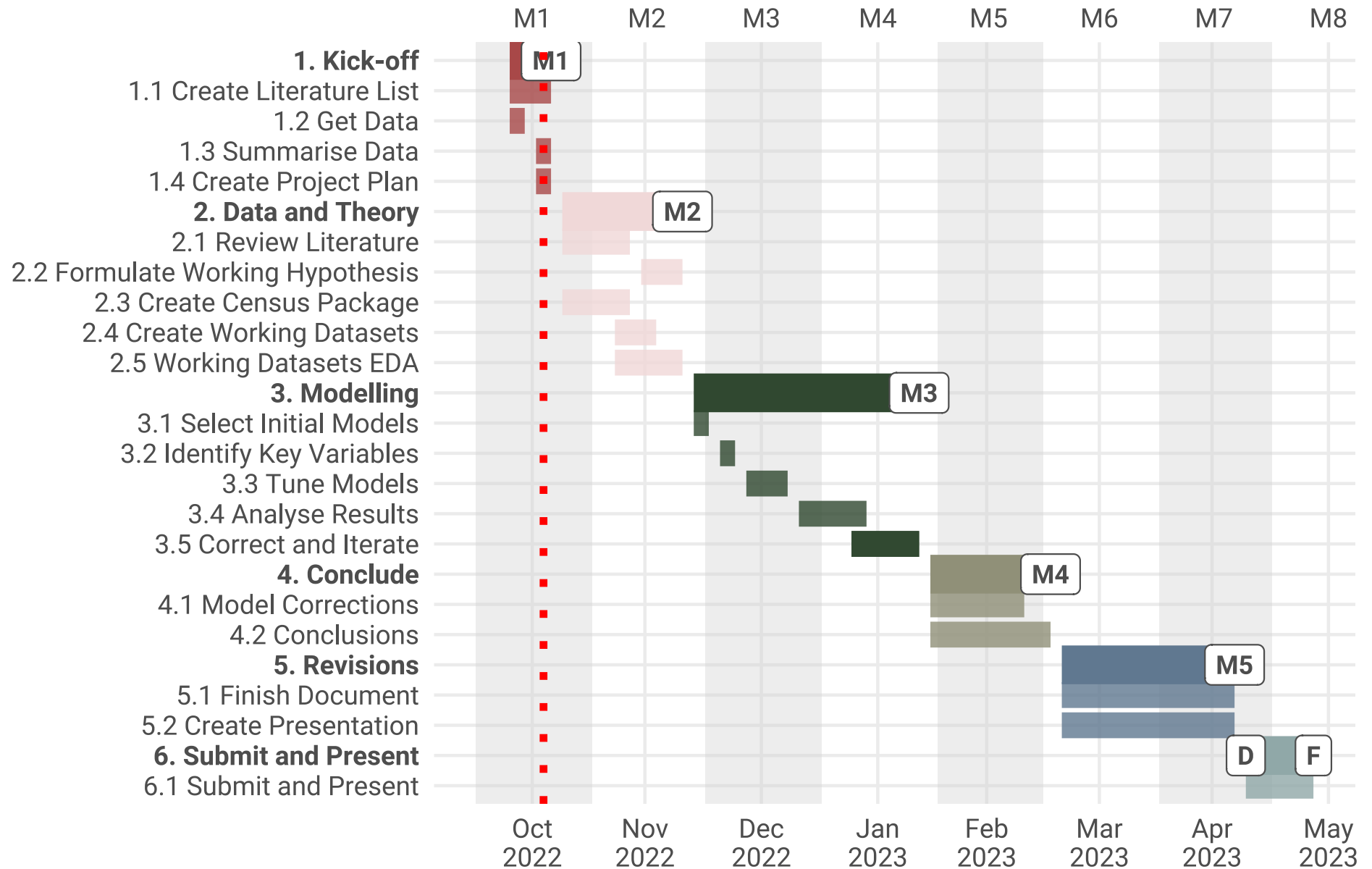
- **Phase 1 - Kick-off** (*this meeting*): Present and discuss project plan, approach, existing data and literature review plan.
- **Phase 2 - Data and Theory**: Conduct detailed data review, creating code/packages to retrieve tidy data where needed (.e.g. Census). Conduct detailed literature review, formulate initial hypotheses (which data are important, possible modelling approach).
- **Phase 3 - Modelling** : Conduct core analytical work. Validate/ tune initial hypotheses: compare models and variables of significance. Analyse results against literature. Iterate over initial assumptions.
- **Phase 4 - Conclude** : Implement any model corrections and complete draft for revision.
- **Phase 5 - Revision** : Finish documentation and code, prepare presentation.
- **Phase 6 - Submit and Present**: Submit work, present.

This plan has been made under the below assumptions:

- This plan is targeted to meet the April 2023 submission deadline. To achieve this objective, sufficient analytical progress needs to be made by the end of the year (halfway phase 3). If that’s not the case the plan needs to be adjusted to reflect a June submission.
- Taken the above into consideration, phases 4,5 and 6 have a general plan at this stage, which needs to be updated by EOY.
- This plan assumes that all documentation/deliverables will be continuously generated - thus there is no explicitly separate allocation of time to this.



## 1.1 Gantt Chart



## 1.2 Tasks

Table 1: Milestones

Group	Activity	Start	End
1. Kick-off	1.1 Create Literature List	10 Oct	21 Oct
1. Kick-off	1.2 Get Data	10 Oct	14 Oct
1. Kick-off	1.3 Summarise Data	17 Oct	21 Oct
1. Kick-off	1.4 Create Project Plan	17 Oct	21 Oct
2. Data and Theory	2.1 Review Literature	24 Oct	11 Nov
2. Data and Theory	2.2 Formulate Working Hypothesis	14 Nov	25 Nov
2. Data and Theory	2.3 Create Census Package	24 Oct	11 Nov
2. Data and Theory	2.4 Create Working Datasets	7 Nov	18 Nov
2. Data and Theory	2.5 Working Datasets EDA	7 Nov	25 Nov
3. Modelling	3.1 Select Initial Models	28 Nov	2 Dec
3. Modelling	3.2 Identify Key Variables	5 Dec	9 Dec
3. Modelling	3.3 Tune Models	12 Dec	23 Dec
3. Modelling	3.4 Analyse Results	26 Dec	13 Jan
3. Modelling	3.5 Correct and Iterate	9 Jan	27 Jan
3. Modelling	3.5 Correct and Iterate	9 Jan	27 Jan
3. Modelling	3.5 Correct and Iterate	9 Jan	27 Jan
4. Conclude	4.1 Model Corrections	30 Jan	24 Feb
4. Conclude	4.2 Conclusions	30 Jan	3 Mar
5. Revisions	5.1 Finish Document	6 Mar	21 Apr
5. Revisions	5.2 Create Presentation	6 Mar	21 Apr
6. Submit and Present	6.1 Submit and Present	24 Apr	12 May