

### Attribution ShareAlike (CC BY-SA)image

**Data Science for Public Managers -**

**Data Science Project Scoping**

**How do we scope actionable data-driven projects that solve public problems?**

This training program was developed by GobLab UAI, the public innovation lab of the School of Government of Adolfo Ibáñez University (Chile) and the Center for Data Science and Public Policy of the University of Chicago (USA). Part of the team from University of Chicago has now moved to Carnegie Mellon University and is continuing this work there.

It was delivered for the first time to 71 senior public managers in Chile between May and July 2019. The curriculum design and first edition were funded by a grant of the Network of Innovators of the Chilean government’s Innovation Lab (Laboratorio de Gobierno) the Chilean Civil Service (Servicio Civil) and Adolfo Ibáñez University.

Adolfo Ibáñez University and the University of Chicago make this curriculum available to the community so that it can be replicated and adapted, under a creative commons license. We ask you to inform us if you are using it and any changes you have made to the mail [goblab@uai.cl](mailto:goblab@uai.cl) as we would like to continue improving and learning on the best ways to teach data science for social good.

**General Goal**

* To identify and scope problems that achieve a government agency’s goals and can be addressed with data science
* To formulate a data science project proposal

**Specific Goals**

After taking this course, you should be able to:

* Determine the feasibility of using data science to address a problem facing government agencies
* Understand how to define the goals of the project and the actions that will be informed by this project
* Define a scope that can be turned into a project plan
* Understand the ethical challenges you must address during project scoping and execution.

**Participant Profile:**

For managers:

* At the central, regional and local level
* From the executive, legislative and judicial branch
* It is important to note that you **do not need** to know Data Science, Artificial Intelligence or Analytic techniques to take this program.

**Duration:** 24 hours (3 days of 8 hours each). The curriculum below is designed for classes scheduled once every other week or once a week. If you want to carry out this course in an intensive format, some adjustments are necessary. Please reach out to [goblab@uai.cl](mailto:goblab@uai.cl) if you plan to teach it in an intensive format.

**Methodology:**

The course is in workshop format which includes lectures, individual and group work. Participants will complete practical exercises to refine their initial proposal into a data science project definition.

There are two homework, due before the second and third day.

**Teachers**

This training program is designed to be taught by two teachers: a data scientist and a project or program manager.

**Videos available in** <https://gobierno.uai.cl/centros/goblab/proyecto-ciencia-de-datos-para-directivos-publicos/>

**Contents**

1. Pre-feasibility analysis
   * Real and significant problem
   * Ability to act on the problem
   * Priority and commitment from the institution
   * Data accessibility
   * Identification of risks
2. Methodology for data science project scoping:
   * Goals:
     + Characteristics of a goal
     + Constraints
     + Trade-offs
   * Actions:
     + Available actions
     + Effectiveness of the actions
     + Modification of actions
   * Data:
     + Types of data
     + Data dictionary
     + Data maturity matrix
   * Analysis:
     + Description
     + Detection
     + Prediction
     + Optimization
     + Behavior change
3. Ethical data use and management:
   * Privacy
   * Transparency
   * Accountability
   * Discrimination
   * Social license
   * Consent
4. How to get started with a project:
   * Team
   * Infrastructure
   * Project cycle and plan

**TRAINING PROGRAM STRUCTURE**

**Before the course starts**

Each participant must identify a problem that is important to her/his organization that could potentially be solved using data and particularly data science techniques. The participant must answer the following questions:

1. Name of the project
2. Project justification: Describe the problem your organization is facing. Describe which actors are impacted/affected by the problem? Why Is the problem important for your organization?
3. Current State: How is your organization currently solving this problem?
4. Goals: Which goals would be achieved if the problem is solved? Which is the expected deliverable?
5. Data: What data do you need to solve this problem? Does your organization own or have access to it? Which sources do you have access to? Describe them in detail: data source names, content and how many years of historical data are available.

**Day 1: Overview of project scoping and deep dive into goals**

|  |  |
| --- | --- |
| **8.30 - 9.00** | **Registration** |
| 9.00 - 10.00 | Introductions and overview of the training program |
| 10.00 - 11.00 | Case Studies: What is and what is not Data Science and what types of problems can it help solve |
| **11.00 - 11.30** | **Coffee break** |
| 11.30 - 12.15 | Deep dive into feasibility and initial criteria |
| 12.15 - 13.00 | Overview of project scoping using a participant’s project proposal   * One pre-selected participant presents his/her project proposal to the rest of participants (5') * Group work: Each group will create a list of questions by topic to evaluate if the project as proposed satisfy the initial criteria of feasibility (15’). * Q&A with the pre-selected participant (25’) |
| **13.00 - 14.00** | **Lunch** |
| 14.00 – 15.00 | Peer feedback of projects   * Work in pairs to provide written feedback on project proposals (40’) * Volunteers discuss the feedback received |
| 15.00 - 16.00 | A bird’s-eye view of the project scoping definition  A complete cycle of the scoping process is shown using a preselected participant’s proposal |
| **16.00 - 16.20** | **Coffee break** |
| 16.20 - 17.30 | Deep dive into goals |
| 17.30 - 18.00 | Assignment description  Quick feedback form |

**Day 1 Supplies**

* 2 printed copies of each student’s project (sent before the course)
* 1 pen for each student

**Assignment for students:**

* Based on what you learned today and the feedback from your peers, complete the Data Science Project Scoping Worksheet. For this assignment, focus on emphasis Section III: Problem Definition.

**Teacher’s instructions:**

* Assignment’s due date should be before the second day, preferably with enough time to do the following tasks:
  + Provide written feedback to each project
* Identify the most promising/feasible project, in order to distribute at least one feasible project in each group for the group work for Day 2.

**Day 2: Deep dives into actions/interventions and data**

|  |  |
| --- | --- |
| 9.00 - 9.30 | Recap of Day 1 |
| 9.30 - 11.00 | Group work: Problem Definition Feedback  Instructions:  Each participant will be part of a group of 4 persons.  Note: the following numbers were calculated with the assumption of 40 participants in the training program. You may adjust the times depending on the size of the class. We recommend that each student has a printed copy of the following instructions.  *Instructions for Students:*  *Part 1- In each team (30’):*   * *One person will be the timekeeper* * *One person will explain the problem in 3 minutes (section 3 of the worksheet). Do not use other sections of the worksheet!* * *Each remaining participant will give feedback for one minute* * *Each participation (explanation + feedback) should be 7 minutes tops!* * *Repeat this for the four group members*   *Part 2- Individual work (5’):*   * *Prepare to present your learnings i.e. “My project definition is correct/incorrect in the following points …”; “My improvements were … / The improvements I need to do are …”*   *Part 3 - Discussion (50’)*  *Each participant will present in 45 seconds:*   * *My project definition is correct/incorrect in the following points …* * *My improvements were … / The improvements I need to do are …*   Use a timer on the screen! |
| **11.00 - 11.20** | **Coffee break** |
| 11.20 - 13.00 | Deep dive into Actions |
| **13.00 - 14.00** | **Lunch** |
| 14.00 - 15.00 | Deep dive into Types of Data and Data Maturity Framework |
| 15.00 - 16.00 | Data Maturity Matrix exercise  Each student will work with a peer and should have a copy of their project worksheet and the data maturity framework. Each student will have 30 minutes to complete the exercise.  *Instructions for students:*   * *For each project select* ***one*** *data source, discuss it with your peer and fill the data maturity matrix* |
| **16.00 - 16.20** | **Coffee break** |
| 16.20 - 17.30 | Groups work on refining actions and data   * Use the previous established groups of four * Choose **one** project per group and discuss how to improve the actions and data   *Instructions for the students:*  *Part I - 20 minutes*   * *Improve the parts V and VI of the chosen project worksheet* * *Summarize the improved sections in a large paper sheet (like the ones in a flip chart).*   *Part II - 15 minutes*   * *3 volunteers share the findings with the class* |
| 17.30 - 18.00 | Assignment description  Quick feedback form |

**Day 2 Supplies**

* large paper sheets
* markers
* masking tape
* 1 printed copy of each homework

**Assignment:** Talk to people responsible for collecting and storing data and for determing/allocating interventions/actions in your organizations and update the project scoping worksheet with:

1. New data sources and actions that you learned about
2. Refine existing data sources and actions in your worksheet with more detail.

**Teacher’s instructions:**

* Assignment’s due date should be before the third day, preferably with enough time to do the following tasks:
  + Provide written feedback to each project
* Identify the most promising/feasible project in order to select those to assign to each group for the final iteration (activity on day three after lunch)

**Day 3: Deep dive into analysis and ethics - Final iteration of selected projects**

|  |  |
| --- | --- |
| 9.00 - 9.30 | Recap of day 2 |
| 9.30 - 11.00 | Deep dive into Analysis |
| **11.00 - 11.00** | **Coffee break** |
| 11.30 - 13.00 | Deep dive into Ethics |
| **13.00 - 14.00** | **Lunch** |
| 14.00 - 14.30 | Analysis Q&A |
| 14.30 - 15.00 | Final project iteration  *Instructions for the students*   * *30’* * *Choosing one project, the group will do a final iteration on the project worksheet to improve it.* * *Summarize each of the following in a large paper sheet (like the ones in a flip chart):*   + *Problem description*   + *Goals*   + *Actions*   + *Data*   + *Analysis*   + *Validation*   + *Ethics* |
| 15.00 - 16.00 | Final feedback for projects   * Each group will present their selected project in 3 minutes * Each student will give feedback to each project in a post-it * At the end of the session, students stick their feedback post-its to each project |
| **16.00 - 16.20** | **Coffee break** |
| 16.20 - 17.30 | How to get a project started |
| 17.30 - 18.00 | Wrap up and evaluation |
| 18.00 - 19.00 | Celebration |

**Day 3 Supplies:**

* large paper sheets
* markers
* post-its
* masking tape