# **Data Science Project Sc­oping Worksheet**

1. Project Name:

2. Organization Name:

3. Project Description:

4. Who are the agencies/departments that will need to be involved?

5. Who are the individuals in these organizations that are stakeholders? What are their role?

**6. Goals (in order of priority)**

What are you maximizing or minimizing?

Are there any constraints (budget, resources, etc.)?

|  |  |  |
| --- | --- | --- |
| Goal 1: | Goal 2: | Goal 3: |
| Constraints: | Constraints: | Constraints: |

**7. Actions**

What is the action?

Who is taking the action?

What/Who is it being taken on?

How often?

|  |  |  |
| --- | --- | --- |
| Action: | Action: | Action: |
| Questions | Questions | Questions |
| Who is taking the action? | Who is taking the action? | Who is taking the action? |
| What/Who is it being taken on? | What/Who is it being taken on? | What/Who is it being taken on? |
| How often is the decision to take this action made? | How often is the decision to take this action made? | How often is the decision to take this action made? |
| What channels are/can be used to take this action (in person, digital, etc.) | What channels are/can be used to take this action (in person, digital, etc.) | What channels are/can be used to take this action (in person, digital, etc.) |
| Other questions about the action | Other questions about the action | Other questions about the action |

**8. Data**

**A. What Data do you have internally?**

|  |  |  |
| --- | --- | --- |
| **Data Source** | **Data Source** | **Data Source** |
| What does it contain? | What does it contain? | What does it contain? |
| What level of granularity? | What level of granularity? | What level of granularity? |
| How frequently is it collected/updated after it’s captured? | How frequently is it collected/updated? | How frequently is it collected/updated? |
| Does it have unique identifiers that can be linked to other data sources? | Does it have unique identifiers that can be linked to other data sources? | Does it have unique identifiers that can be linked to other data sources? |
| Who’s the internal owner of the data? | Who’s the internal owner of the data? | Who’s the internal owner of the data? |
| How is it stored? | How is it stored? | How is it stored? |
| Other | Other | Other |

**B. What data can you get externally and /or from public sources?**

|  |  |  |
| --- | --- | --- |
| **Data Source** | **Data Source** | **Data Source** |
| What does it contain? | What does it contain? | What does it contain? |
| What level of granularity? | What level of granularity? | What level of granularity? |
| How frequently is it collected/updated after it’s captured? | How frequently is it collected/updated? | How frequently is it collected/updated? |
| Does it have unique identifiers that can be linked to other data sources? | Does it have unique identifiers that can be linked to other data sources? | Does it have unique identifiers that can be linked to other data sources? |
| Who’s the internal owner of the data? | Who’s the internal owner of the data? | Who’s the internal owner of the data? |
| How is it stored? | How is it stored? | How is it stored? |
| Other | Other | Other |

**C. What data would you need in addition to the ones above?**

Data Source:

Data Source:

Data Source:

**9. Analysis**

What analysis needs to be done?

What type of analysis is it? Description? Prediction? Detection? Behavior Change?

How will you validate the analysis using existing data?

What field trial or randomized controlled trial can you design to validate it in the field?

|  |  |  |
| --- | --- | --- |
| Analysis 1: | Analysis 2: | Analysis 3: |
| Analysis type: | Analysis type: | Analysis type: |
| Which action will this analysis inform? | Which action will this analysis inform? | Which action will this analysis inform? |
| How will you validate this analysis? | How will you validate this analysis? | How will you validate this analysis? |

|  |
| --- |
| The Center for Data Science and Public Policy at the University of Chicago is a joint effort between the Harris School of Public Policy and the Computation Institute with the goal of increasing the use of data-driven and computational methods in public policy research and practice. The Center conducts educational programs, does data science research and applied policy projects with governments, non-profits, and corporations, and create reusable software and data products across policy areas such as health, economic development, education, public safety, criminal justice, and environment. For more information about our programs and work, please visit **http://dsapp.uchicago.edu** or email us at [**dsapp@uchicago.edu**](mailto:dsapp@uchicago.edu) |