# **Data Science Project Scoping Worksheet**

This worksheet is designed for social good organizations (government agencies, non-profits, social enterprises, and others) to scope actionable data science projects.

**1. Project Name:**

U.S. Medical Insurance Costs

**2. Problem Description:**

What factors lead to high insurance debt.

**2.1 What is the problem you are facing?**

There are many factors that can lead to building up large insurance debt so the issue is finding out if there is a main factor that causes this or a combination of separate factors.

**2.2 Who/what is affected by this problem?** (people of certain type, organizations, neighborhoods, environment)

U.S. residents from different regions.

**2.4 Why is solving this problem a priority for your organization?**

Finding the cause of large insurance debt accumulation can demonstrate how Americans can change their lifestyle to prevent or ease the debt they accumulate.

**3. Goals: What are the business/policy goals that will be accomplished by solving this problem and what constraints do you have? (in order of priority)**

|  |  |  |
| --- | --- | --- |
|  | **Goal** | **Constraints** |
| 1 | Find what is the main cause of insurance debt. |  |
| 2 | Find a combination of factors that can lead to insurance debt. |  |
| 3 | Do children have a big impact on other type of life style choices. |  |

**4. Analysis**

* Typical data science projects include a combination of analysis, typically including description, detection, prediction, optimization, and/or behavior change.
* Again, the analysis is not the goal of the project - the **analysis** helps you use the **data** you have to inform the **actions** you have access to in order to achieve your **goals.**
* Choose the right set of analysis for each problem
* You must validate the analysis and use a validation process that matches how your analysis will be used in practice

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Analysis 1:** | **Analysis 2:** | **Analysis 3:** |
| **Analysis type**  *e.g. Description, Prediction, Detection, Behavior Change* |  |  |  |
| **Purpose of the analysis** *eg. understand historical behavior of individuals, estimate risk of disease of patient, identify which actions will diminish overfishing in the region* |  |  |  |
| **Which action will this analysis inform?**  *eg. inspections of compliance regarding fishing quotas* |  |  |  |
| **How will you validate this analysis using existing data?**  **What methodology and what metrics will you use?**  *eg. using historical data, running an RCT* |  |  |  |