# Tool: Data Qualification Checklist

**Bring the Data into the Decision**

**Instructions:**

Use this guide to ask key questions about the data and decision it is being used to make.

**Question:** Who are the stakeholders, and what does each have at stake?

List all key stakeholders in the table. For each, indicate how a good or bad outcome will affect that stakeholder.

|  |  |
| --- | --- |
| Stakeholder | What’s at Stake? |
|  |  |
|  |  |
|  |  |

**Question:** Do any of the Key Performance Indicators (KPIs) dominate others, or is there a priority order?

List the stakeholders along the top of the table.

In the first column, list any KPIs these stakeholders have an interest in. For each stakeholder, indicate whether each KPI is high, medium, or low priority.

|  |  |  |  |
| --- | --- | --- | --- |
| Stakeholders  KPIs | *(identify a stakeholder here)* | *(identify a stakeholder here)* | *(identify a stakeholder here)* |
| *(list a KPI here)* |  |  |  |
| *(list a KPI here)* |  |  |  |
| *(list a KPI here)* |  |  |  |

Use the table above to create a prioritized list of KPIs

1.

2.

3.

List the data that is or could be available

|  |  |  |
| --- | --- | --- |
| **Data available** | | |
| **Observational or Experimental** | **Limitations** | **Visualization(s)** |
|  |  |  |
|  |  |  |
|  |  |  |

|  |  |  |
| --- | --- | --- |
| **Data that could be collected** | | |
| **Observational or Experimental** | **Limitations** | **Visualization(s)** |
|  |  |  |
|  |  |  |
|  |  |  |

Questions to ask about each dataset

Data/Decision Alignment

* Are there alternative summary measures or visualizations for the data?
* Are there alternative KPIs I need to consider?
* Am I using the data correctly?
* Are there no acceptable outcomes that emerge from looking at the data?

Data Characteristics

* Do the KPIs move in expected and useful ways?
  + Are there unusual features (for example a bimodal distribution)?
  + Is the correct summary measure being used?
  + Would changing the granularity of the measure give a more accurate picture?
* Are the data of good quality for making inference?
  + Was the sample random?
  + Is the sample representative?
  + Are there potential sources of bias?
* Can the data be used to infer causality?
  + Is the data experimentally obtained?
  + Are there any potential confounds (alternative explanations of the outcome)?