Wellbore Data-Quality Narrative Visualization

(ZainalH2@illinois.edu)

1. **Messaging:** The "Wellbore Data Quality Narrative Visualization" is a comprehensive narrative visualization tool that provides readers with a detailed overview of the quality of wellbore data. It not only presents a high-level summary but also enables readers to delve deeper into the charts to effectively address and comprehend any concerns related to data quality. By offering a narrative component, this visualization further aids readers in identifying specific areas that may need improvement in order to elevate the overall quality of wellbore data.

2. **Narrative Structure:** I use a drop-down story structure to convey the narrative message effectively. The narration starts off with an easily accessible "***Overview***" chart, which offers readers a clear visualization and facilitates the identification of the wellbore data-quality history within the specified timeframe, spanning from 03-June-2023 to 29-Jul-2023.

   Within the visualization, readers could delve deeper into the "***Wellbores***" chart by drilling down into specific data points, thereby gaining a comprehensive understanding of the contributing factors behind any concerns related to data quality. This interactive feature enables readers to identify and analyze the wellbore that is responsible for the observed issues in data quality, facilitating a more detailed examination of the underlying causes.

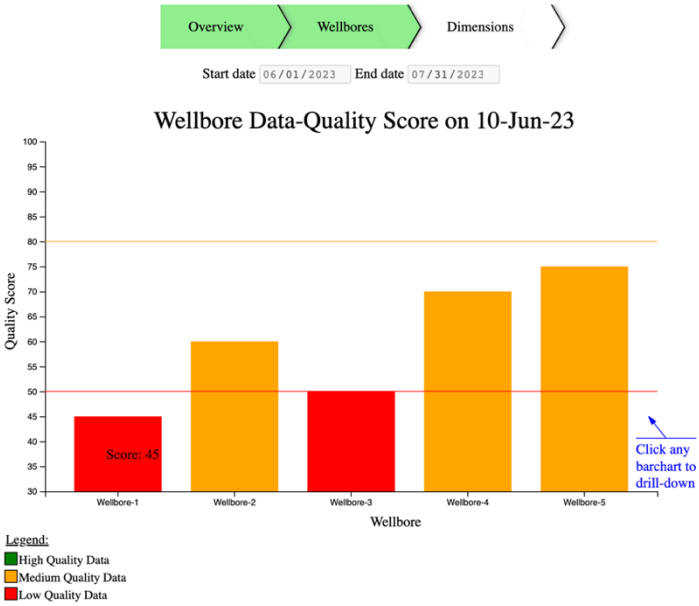
   Furthermore, readers can drill down into the "***Dimensions***" chart to understand which quality dimension needs to be addressed in order to improve the wellbore's data quality. This interactive functionality empowers readers to identify and distinguish the precise areas that require intervention, enabling targeted efforts to address and enhance the identified quality issues.

   At the top of the page, a user-friendly navigation bar is provided, allowing readers to effortlessly navigate to the previous chart and refer to previously explored information. However, the navigation bar cannot be used to drill down into the information forward; readers must click a data point on the chart to progress to the next chart.

   Here is a sample narrative illustrating how the reader drills down into the information to highlight data quality concerns:

   1. The reader analyzes the overview chart and identifies a low data-quality data point on the 10th of June 2023:

2. The reader drills down into the low-quality data point and discovers that *Wellbore-1* has the lowest data quality score:

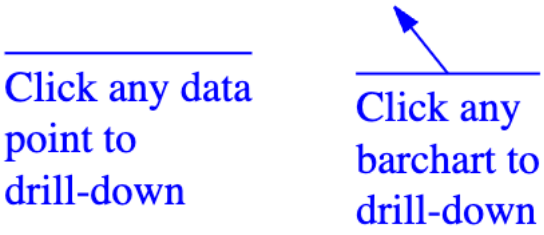


3. Furthermore, the reader drills down into the *Wellbore-1* data point and finds that the "Completeness" quality dimension has the lowest score. At this point, the reader concludes that enhancing the wellbore's data quality involves completing the missing data for *Wellbore-1*. Additionally, the "Validity" and "Consistency" dimensions have low scores and can be addressed for improvement:

3. **Visual Structure:** I utilize a ***user-directed visual structure*** to guide readers between scenes, ensuring their understanding of the data and facilitating navigation. The following elements are incorporated into the visual structure:

   - Navigation bar: It enables readers to quickly navigate to the previous scene, taking into account the data context of that scene. However, the navigation bar is not used for forward navigation. To drill down into the information, readers must click on data points on the charts.

   - Annotations: These assist readers in identifying how to drill down into the information presented on the charts.



   - Text narration: An essential element of the visual structure is the text narrative provided at the bottom of each chart helps readers comprehend the charts and their significance.

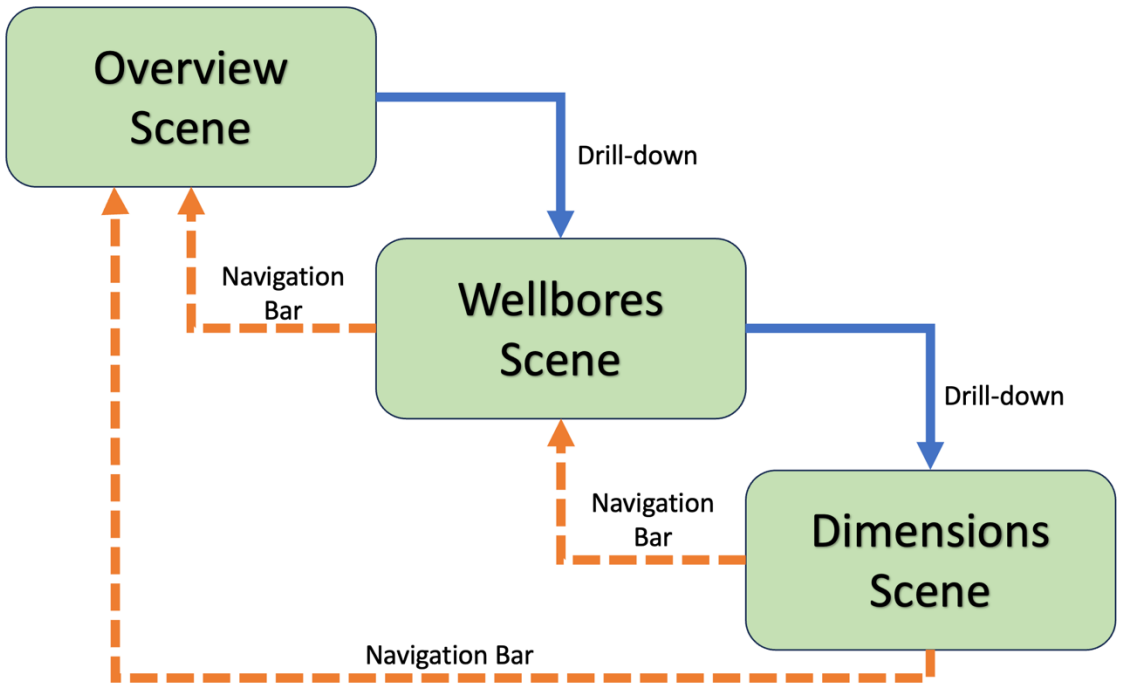
The narrative revolves around wellbore data quality. A wellbore refers to the hole or opening that is drilled into the ground during the process of constructing a well, providing access to underground resources such as oil, gas, water, or geothermal energy. The starting point of the drill-down narrative is the "Data-Quality Score Overview" line chart, spanning from 03-Jun-2023 to 29-Jul-2023. The parameters of the overview chart are the start and end dates. The line chart distinguishes low-quality data with a red circle, medium-quality data with an orange circle, and high-quality data with a green circle. Readers can quickly identify concerns related to low-quality data. By clicking on a circle data point, readers can drill down to the "Wellbores" bar chart to review the specific concern. This chart displays the "Low Quality Threshold Line" and the "Medium Quality Threshold Line." Data falling below (or on) the "Low Quality Threshold Line" is considered low-quality, while data falling between the "Low Quality Threshold Line" and the "Medium Quality Threshold Line" is considered medium-quality. Data above the "Medium Quality Threshold Line" is considered high-quality. A tooltip is provided when hovering over a data point, offering detailed information about that specific point.

- Consistently color-coded data points: The visual presentation of color-coded data points prompts readers to focus on data quality. Red indicates low data quality, orange represents medium data quality, and green signifies high data quality.

- Chart title and axis: These elements provide context for the data and aid readers in understanding the transition from one scene to the next.

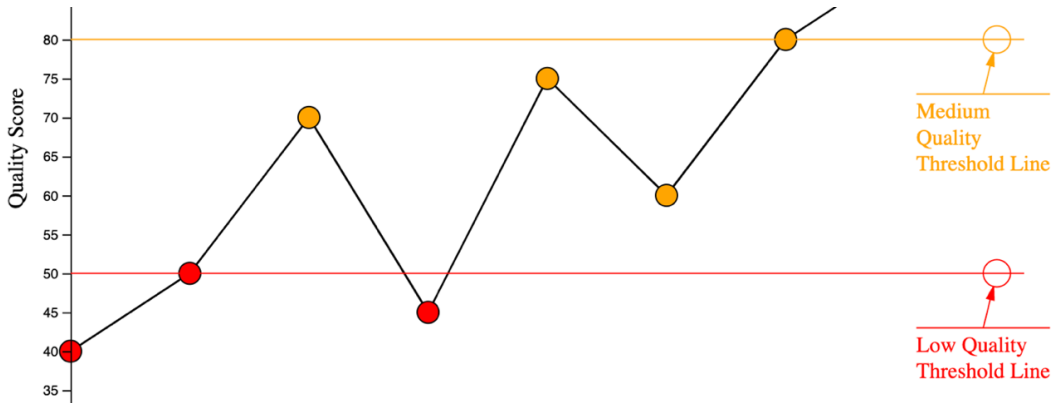
4. **Scenes**: There are three scenes presented in the following order:

    1. **Overview** scene: This scene visualizes the aggregated data quality scores of wellbores over a specific date range.
    2. **Wellbores** scene: In this scene, the quality scores of wellbores are visualized for a selected date.
    3. **Dimensions** scene: This scene visualizes the quality scores of dimensions for a chosen wellbore and selected date.

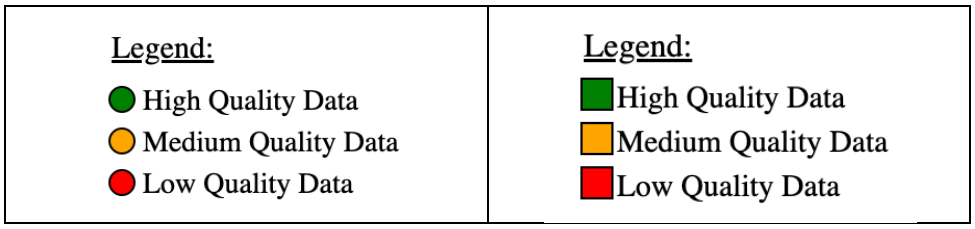


5. **Annotations**: Annotations are used to highlight data quality and guide readers to drill down into the information. These annotations are strategically positioned and remain fixed in their designated locations throughout each scene, ensuring consistent and reliable reference points for readers.

- Low and Medium Quality Threshold Lines: The purpose of these lines is to emphasize and group together the data points based on their respective data quality levels, namely low, medium, and high.

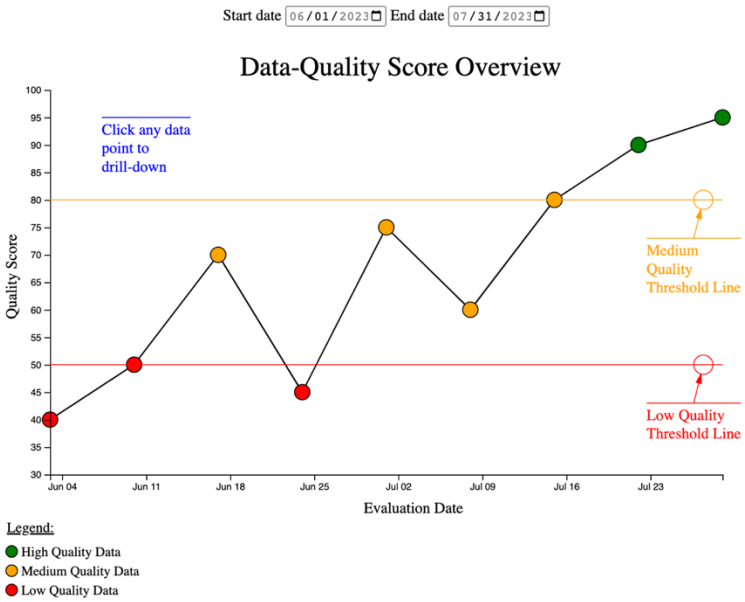


- Chart legends are provided to help readers understand the charted data quality. The legend symbol is adjusted according to the chart types. A circle-shaped legend is provided for the scatter-line chart, while a square-shaped legend is provided for the bar chart.

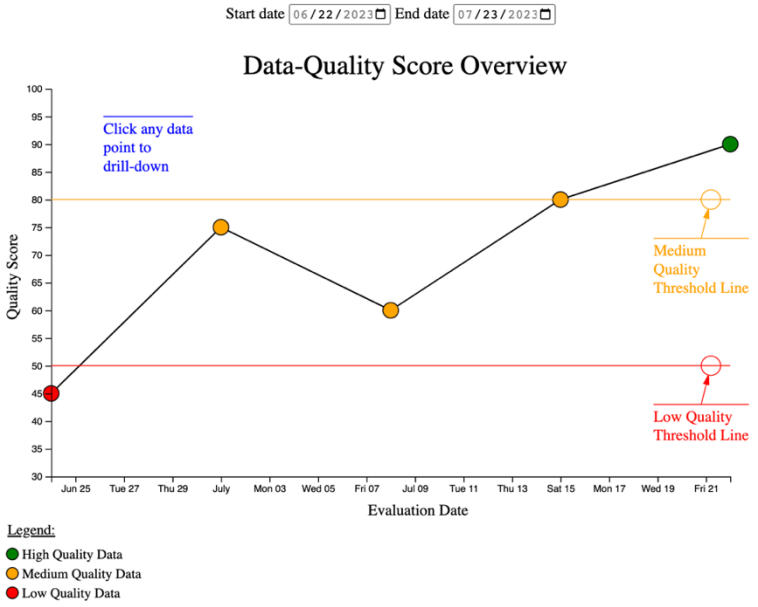


6. **Parameters:** These parameters are used to filter the data on the "Overview" scene. There are two parameters: the data quality history's start date and end date. These parameters are enabled for the "Overview" scene, but they are read-only information for the other scenes.

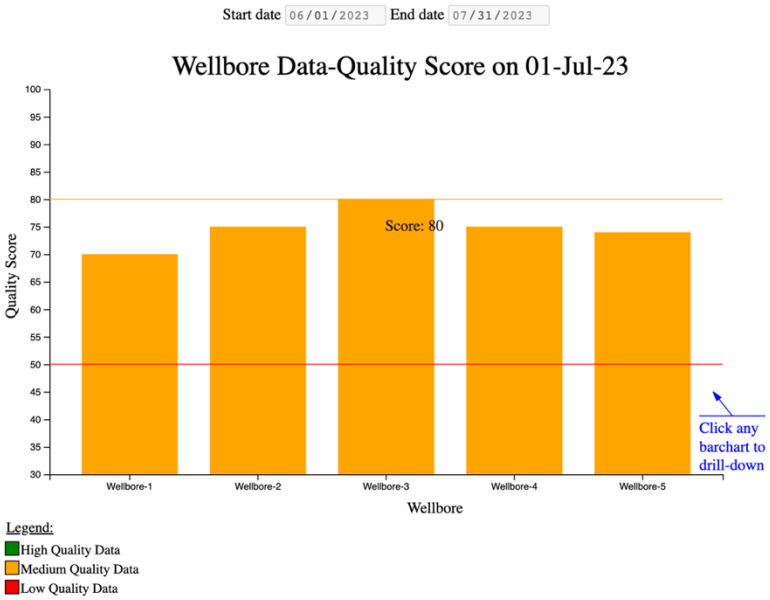
   Below is a sample overview chart with the start date set to 01-Jun-2023 and the end date set to 31-Jul-2023:



Below is another sample overview chart with the start date set to 22-Jun-2023 and the end date set to 23-Jul-2023:



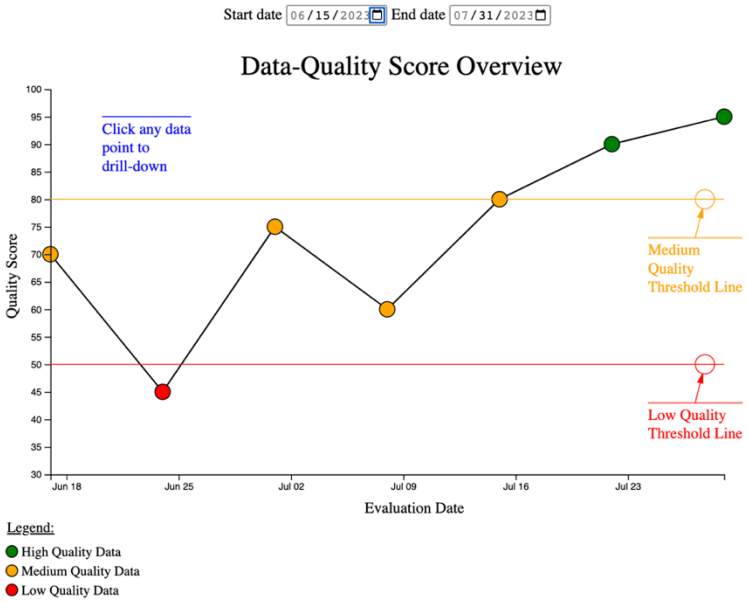
Please note that the start date and end date parameters are disabled for the other scenes, as indicated by the `grayed-out dates`:



7. **Triggers:** The overview chart is ***automatically*** refreshed and triggered to display the intended date range whenever the start date or end date is updated or changed, ensuring that it accurately reflects the intended date range as specified by the changes.

   Below is a sample case when readers try to update the start date from 01-Jun-2023 to 15-Jun-2023:

After the start date is updated, the overview chart is ***automatically*** refreshed and displayed as follows:



8. The source code is available here: https://github.com/zen030/zen030.github.io