**Data Quality Form**

**👉 Business Question:  What roller coaster manufacturer has the most roller coasters still operating?**

For each item below, give a score ranging from 1 to 5.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **1** | **2** | **3** | **4** | **5** |
| Poor | Below Average | Average | Good | Excellent |

1. **Relevance**
   * **DESCRIPTION**: How relevant is this data to the business question you are trying to answer? Are all of the fields useful as a part of the analysis?
   * **SCORE**:  5 – Each manufacturer is listed in the CSV file, the parks are listed as well as each roller coaster.
2. **Timeliness**
   * **DESCRIPTION**: How up to date is the data? Is it recent enough and covering a long enough range of time to be useful?
   * **SCORE**: 3 – The last time it was up to date was back in October 2022. As I was looking at the data, I noticed that the data still showed the Vortex at Kings Island was still operating when it had been torn down in 2020.
3. **Historical Usage**
   * **DESCRIPTION**:  Has this data been utilized in other analyses?  If there is not much usage of the data source, is there a reason for this?  Could other users be avoiding this data set for a particular reason?
   * **SCORE**: 4 – I’m not sure how to compare this to others, I just know the download to view ratio was .18%. This is dependent on what the viewers were looking for and I cannot make any assumption as to why. The usability is a 10.
4. **Cleanliness**
   * **DESCRIPTION**: How clean is the data set? Are there a large amount of null values, duplicated rows, or inconsistent formatting of the data? Will it take a lot of work to get it in usable form?
   * **SCORE**: 3 – There’s some null values that need to be cleaned, plus I would like to update the Vortex one to removed/closed. There are also the remove/closed items in the operating status, I would like to make this one in the same.
5. **Credibility**
   * **DESCRIPTION**: What source is this data set coming from?  Is this considered a trustworthy source?  Are there any potential biases that this data source may hold that could skew their presentation of the data?
   * **SCORE**: 4 – The source is coming from Wikipedia. I think this is a credible source. The Vortex roller coaster being removed has been updated and the date it closed is listed. The data was last updated in October 2022 and is updated annually. If this particular roller coaster was not updated, it’s hard telling what else needs updated.

**TOTAL SCORE:  19/25**