**In no more than 250 words, write a response to the following questions:**

**Does data in spreadsheets ever need to be worked with "at scale"? Is spreadsheet data ever "dirty"? Under what circumstances might you want to use tools like this set from eBay for working with csv data in your current or a previous job? Can you imagine any uses of a spreadsheet-to-JSON data converter?**

**Do not just quote what you read. Reference at least 3 sources for information about additional tools and discussions of data cleansing and format conversion processes and so on.**

In my experience, spreadsheet data is nearly always “dirty”, but that is not necessarily a function of the spreadsheet specifically. Spreadsheets are simply the most approachable form of data collection/aggregation for organizations because nearly everyone is familiar with them. Very often, however, there is not a single point person to maintain the spreadsheet, so it can quickly become a headache to work with (speaking from experience) – one person may end up manually entering a calculated value, whereas another types the equation directly in the cell. The inconsistencies make working with the sheet as a whole very challenging. This (somewhat outdated, but still relevant) blog gives a novel example of a single point person controlling a common spreadsheet, and then exporting it as JSON for easier management (<http://blog.pamelafox.org/2013/06/exporting-google-spreadsheet-as-json.html>).

On the other hand, JSON allows for semi-structured data, which may serve a specific use-case better. The spreadsheet-to-JSON converter could work wonderfully for a spreadsheet with missing or optional entries (or a sparse spreadsheet if one exists). Key-value pairs can be created but the user is no longer managing large quantities of missing data. This also generally makes it easier for humans to interpret logically, as pointed out here (<https://ezeelive.com/json-advantages-disadvantages/>).

One of the most useful tools I use at my job is the Python pandas package, specifically pandas.read\_csv and pandas.read\_excel (<https://pandas.pydata.org/pandas-docs/stable/index.html>). I regularly export data from a building monitoring system into a .txt file which can be read in as CSV. From there, I can do any type of data analysis and manipulation I need to.