

**Q1: Once you have cleaned your data, upload the cleaned data file to a github repository & include the link here:**

<https://github.com/KoolaidDaBeast/ArduinoDocumnets/blob/master/PreparingData.pdf>

**Q2: Submit brief details on which inconsistencies you have found, what was your approach for fixing them and discuss the Pros. and Cons of your approach, using the given table below:**

Incontinences found	Approach for fixing	Pros and Cons of my approach
Humidity values were in encased in double quotations marks	Removed double quotation marks	Pros – The humidity column are all now numbers.
Gaps	Removed the gaps	Pros – There are no gaps in the data Cons – Not easy to read if there are no gaps.
Misleading column names	Rename the columns so they explicitly say what each column is about	Pros <ul style="list-style-type: none"> <li>- The humidity column has units.</li> <li>- The column name are easier to understand</li> </ul> Cons <ul style="list-style-type: none"> <li>- If someone doesn't read the humidity column they will not know what the numbers in the column mean</li> <li>- Column names are bulky</li> </ul>
Null Values	Simply replaced the null with a – to show that there is no reading for this time.	Pros <ul style="list-style-type: none"> <li>- You know something happened to the sensor at that point in time</li> <li>- Tells the reader that no value is recorded</li> </ul> Cons

		<ul style="list-style-type: none"> <li>- No value for that reading of course</li> </ul>
Row Id	Simply used a formula to add 1000 from the previous row to the row below it so all the rows are incrementing by a thousand.	<p>Pros</p> <ul style="list-style-type: none"> <li>- Row Id's are consistent</li> </ul> <p>Cons</p> <ul style="list-style-type: none"> <li>- RowID becomes large</li> </ul>