# How will you tell your data story?

Remember to give the context for your data **(the 5 ws):**

* What’s in the data, what is it about?
* Who is the data about? And who collected the data?
* Where is the data from? E.g. What region does it cover?
* When does the data cover? Is the data recent or old? Is it annual, or more frequent?
* **Why** was the data collected, and **why** should we care about it?

Have you selected your key facts, numbers, and charts? Remember, you shouldn’t just tell us what the numbers are, you should explain what the numbers mean; what does your analysis show?

## Writing with, and about, numbers: Examples from AEI

 “Any data that you are referencing **needs to be scrutinized**. Taking this data set at face value you could believe that sub-Saharan Africa is consistently the biggest consumer of coal. Only by interrogating the data you can find that this is brought about by only one country and there is no data available for most countries.”

**This is a great point, and data *does* need to be scrutinized. You could add information on *why* there’s no data for other countries, and what might be different about S.Africa, perhaps drawing on research or other sources.**

“From the histogram, most countries around the world consume little amount of coal per capita. Through the geo chart as well as filtering the outliers from the chart, high income and upper middle income countries consumed the most amount of coal per capita in 2009.”

**Also some interesting points, think about pulling out the key points in the histogram/geochart. Where should the reader look, what will they see (maybe give some specific numbers), and what is the significance of this?**

“The median coal usage globally is 0.2806. This data produced no lower outliers, but an astonishing 11 upper outliers. The impact of these outliers is felt when the mean coal usage per region is examined; only 2 regions averaged below the median, while all others scored above it to varying degrees. What this means is that policy and alternative energy solutions enacted in these outlying countries will have a disproportionately positive effect on total coal usage globally. These countries (particularly Australia) are urged to reexamine their energy infrastructure to accommodate cleaner energy sources.”

**Good to see some numbers being referred to. There’s some emotive language (“astonishing”), but it isn’t clear how we should interpret this (why is it unusual to have lots of upper outliers?). But, great to see some wider implications being drawn out; these could be backed up with research/external sources too!**

“Whilst leading economies like China, the USA and India lead the pack in terms of raw coal consumption, there is much to be discovered when normalizing this data. Calculating coal consumption per Capita, it is in fact less economically dominant countries with lower populations that have the highest coal consumption per capita. Calculating consumption this way changes the highest coal consumers over all income levels. There are a few questions one might ask regarding the data presented. Firstly, only raw coal consumption is documented, not the amount of energy extracted from said coal, which can vary. ....”

**Nice explanation based on the data, but could use specific numbers and charts to support the claims.**

 “The regional area of East Asia and Pacific's average coal consumption per capita was 0.8375 MTOE in 2011, greater than the global average of 0.5723 MTOE. Australia, Taiwan then China were the top per capita coal consumers in the region. The smallest coal consumers per capita were the Philippines, then Thailand and Vietnam. Whilst in 2011 Australia consumed 1.9988 MTOE coal per capita, the Philippines consumed a comparatively very small 0.0877 MTOE.”

**Great to use the numbers, and pull out some key features. But what do they mean? Combining this para with the previous one might bring the best of both of them out!**

 “Based on the box plot generated from the data above it is apparent that the east Asia and pacific region has a large inter-quartile range which highlights a large difference in coal consumption between the countries in the region. From this data it could be argued that this large disparity is a product of the greatly differing income levels of the countries within the region. ….”

**Again good to see the data being explained, and drawing on the boxplot to do this (maybe give the numbers too to help contextualise), and flag wider implications drawing on research to do so.**