DATA UNIFICATION METHODS

Data harmonization and cleaning, what we call data unification, is of paramount importance before you start analyzing and looking for insights.

In particular, it would seem almost all companies and their analysts think that harmonization is about tagging, so that datasets can connect with each other.

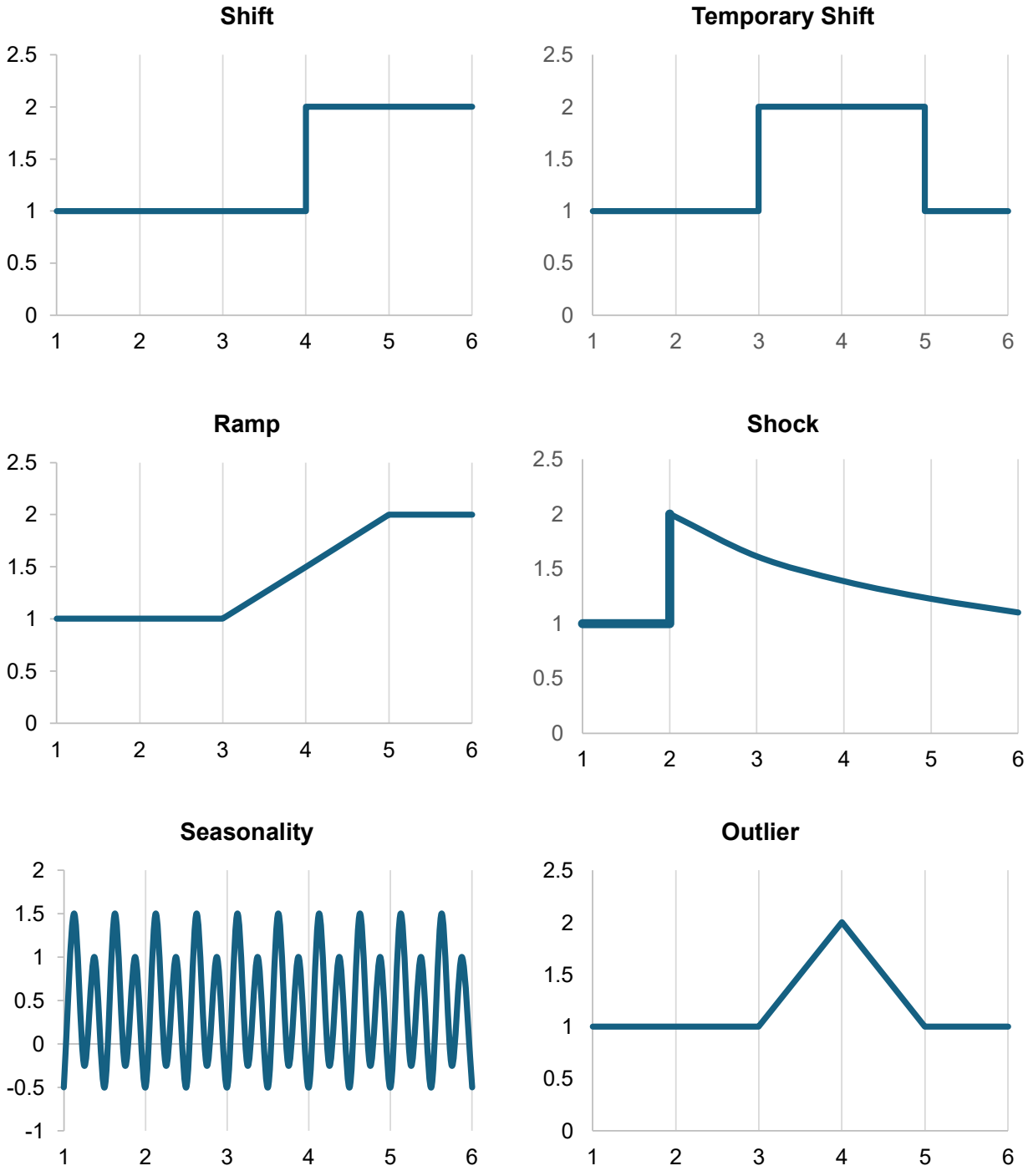
Important, but if you do only this, your company will be in deep trouble over time. Harmonization is so much more. Yet, in 21 years of dealing with these issues at the world's largest companies, we have yet to encounter a single company that does this well and systematically. A shocking state of affairs.

The good news that this is a large opportunity for the bottom line. This is because the precision in predictions will increase if harmonization is done well. The problem is that it is alien territory for executives who mostly think about traditional issues.

The graph on the next page shows the IMF's six archetypes of data adjustments (can also be found in most textbooks).

1. Shifts in time series is the most common adjustments to be made. It needs to be done in almost 100% of cases we have encountered, but we have never seen it. The IMF notes that it takes experience to do this right, but is not that difficult.
2. Temporary shifts are common. Covid led to a negative temporary shift in 2020, and a positive shift in 2021 in most countries. This can be handled systematically with well established methods, but what we have seen are mostly amateur patches.
3. Ramping is a variant of shift. It happens when a phenomenon id building gradually. Take GLP-1's gradual impact on junk food and junk beverages.
4. Shocks can for example be an increase in VAT. It has an immediate effect in the short term, then tends to peter. Shocks can have four shapes, not shown here.
5. Seasonality should always be adjusted for in short term models, but sometimes also in models with strategic horizons (5-10 years) even though they seem to have no seasonal element.
6. Finally, outliers will always exist. They sometimes need to be adjusted for, but not always. It depends on whether they influence the analysis or not.

DATA ADJUSTMENT METHODS



Source: IMF Quarterly National Accounts Manual; Tellusant summary

Tellusant, Inc. | 240 Elm Street | Somerville, MA 02144 | USA
https://tellusant.com | info@tellusant.com | +1-617-394-1800