Frequently Asked Questions

What is The Atlas of Economic Complexity?

The Atlas is a research and data visualization tool that allows people to learn more about the economic structure of their country, including the growth opportunities that exist in the latent productive capabilities a country has. The Atlas puts the capabilities and know-how of a country at the heart of its growth prospects, where the diversity and complexity of existing capabilities heavily influence how growth happens.

Where can I learn more about the ideas and methodology behind the Atlas?

More detail on our ideas and methodology can be found in the companion book, **The Atlas of Economic Complexity: Mapping Paths to Prosperity**. The first part contains context, motivating questions, and then a detailed description of the methodology.

You can also check out our research publications on the product space and our theory of economic growth based on these ideas:

The Product Space Conditions the Development of Nations

The Building Blocks of Economic Complexity

The Network Structure of Economic Output

The **Glossary** also provides a detailed definition of many terms.

Where do you get your data?

We get the raw data from the **United Nations Comtrade** database, which is publicly available. We use both trade classification types, Harmonized System (HS) and Standard International Trade Classification (SITC) data, across key dimensions of exporter, importer, product, and year. Due to limited, delayed, or inaccurate reporting of trade data to UN Comtrade, our research team - led by Sebastian Bustos, with key inputs from Muhammed Yildirim and Ricardo Hausmann - developed a unique method to clean the data to account for inconsistent reporting practices and thereby generate estimates of trade flows between countries. This data cleaning is known as the *Bustos-Yildirim Method* in the literature.

How do I access your data?

The Atlas 2016 dataset can be**found here**.

Why doesn't your data match Comtrade exactly? Has your data been modified?

We perform our own data cleaning on the dataset. We exploit the fact that trade flows are in theory recorded twice: as exports and imports. In practice, however, trade values are recorded imperfectly and reported on an untimely basis, and sometimes not reported at all.

Our data cleaning process can be summarized in the following three steps:

We first correct import values (which are reported including the costs of freight and insurance – CIF) in order to compare to the same flows reported by exporters (which are reported free on board – FOB);

We then estimate an index of reliability of countries when reporting trade flows, based on the consistency of trade totals reported by all exporter and importer combinations over time;

Our final step is to generate our estimate of the trade values using the data reported by exporters and importers, by taking into account how reliable each country is.

This is essentially an accounting process: we cross-reference the reported exports and imports of countries against each other.

Why should I use your website rather than Comtrade itself, WITS, or some other export data source?

The Atlas features a unique cleaning of trade data to account for inconsistent reporting standards and presents a specific point of view of how trade data predicts economic growth and new economic opportunities. The Atlas studies the relatedness of products, as a measure of the capabilities and know-how of countries, rather than merely the export volumes of specific products. By mapping global production and diversification patterns in the product space, we are able to use what a country currently produces to predict what new product sectors are likely to develop in the coming years—and to make recommendations for which sectors offer the greatest complexity gains, paving the way for higher incomes.

After a decade of research on The Atlas, we have generated a set of stylized facts on trade and production, including that poor countries are defined by their production of few products that nearly every country can produce, while rich countries produce many products, and products that few countries can produce. Beyond basic visualizations of existing trade, we offer the Product Space and Product Feasibility scatterplots to offer a positive perspective of the products that are more likely to

succeed. We find these methods predict product appearance and overall economic growth better than any other single measure of the economy.

What is the difference between the HS and SITC datasets?

The difference is a tradeoff between time length and detail. The SITC data go back further in time, from 1962 to now, but struggles to dissect new products that did not exist in 1962 (e.g. cell phones). The HS data is only available back to 1995, but provides a much more detailed and recent product classification. This is especially important for product introduced in the last two decades, such as high-tech products and electronic goods. Using one or the other therefore depends on the nature of your analysis, but we generally find they lead to highly similar conclusions.

Why do you use product classification at the four-digit level of specificity? What are the high-level categories you use?

The data reported to Comtrade are not universally available at the six-digit level, but we do have such data for most countries. The reason for using four-digit data (specifying around 900 products) is that it represents a good compromise between level of detail, accuracy of the estimated similarities and complexities and differentiation between truly separate product types. That is, although we think it is still helpful to reflect on the differences in capabilities required to make t-shirts or men's suits, there is less of a case for differentiating between men's linen suits of less-than-1cm thickness and men's cotton suits of greater-than-1cm thickness.

Moreover, although product data exists at the eight-digit level and beyond and we have used this level of analysis to study specific sectors within a country, we generally find that using the data at the six-digit level suggests unwarranted levels of precision and, at over 5,000 products, muddies the visualizations.

In aggregating the product data, we use the standard "Chapter" classification of the HS classification, with additional aggregation across a few Chapters of related and miscellaneous product categories.

Why doesn't your data contain HS codes for services? (98xx / 99xx)

Services data are not reported by every country, as the Customs Agency responsible for reporting trade to Comtrade often does not collect services data, as service trade does not go through customs. This biases our calculations towards service-reporting countries, so we disregard services information entirely to avoid that bias.

Why are certain countries named and classified the way they are?

We use **the official names**and data provided to UN Comtrade.

Why are some countries missing from the complexity ranking?

We currently do not provide complexity rankings for countries for which we are not certain of the reliability of the data, often involving countries that are non-reporters (where inferring data from importing countries has weaknesses), heavy re-exporters (not allowing certainty of differentiating true value-added of the country), or smaller countries where discrepancies in reported data lead to misleading conclusions.

Why doesn't your website contain data or visualizations for X?

We have a specific focus on studying growth and development from a capability and know-how perspective of economic complexity, and the content of our site reflects that purpose.

Within this context, if you have a suggestion for additional content, feel free to**contact us** with your ideas.

OK, so this tells me which countries export which products, but do we also know something about in which regions within these countries these products are made?

We have currently produced two online Atlases that dig deeper into the geography of production within specific countries: one for **Mexico** and one for **Colombia**. In fact, these atlases also allow you to explore the industry space of a location, the network of industries that are similar in terms of the capabilities they require. Because these atlases are based on employment data and not just trade data, we can here also depict the position of services.