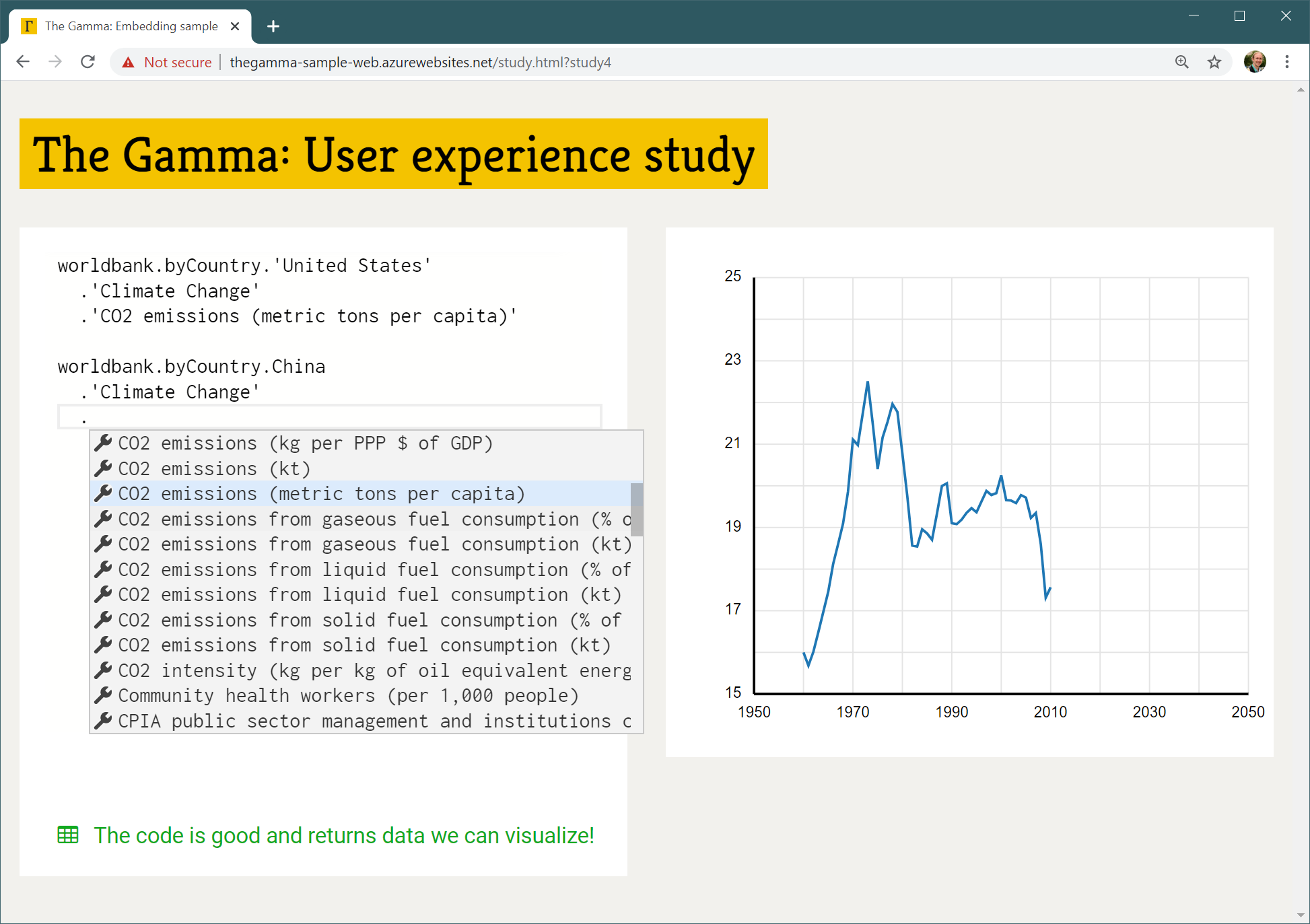
TheGamma user study (4)

The Gamma is a simple programming environment for data exploration that aims to be accessible to non-experts such as journalists. TheGamma is text-based like typical programming languages. In the study, we’ll be using TheGamma in a simple split-screen environment. You see your code on the left and a data visualization automatically generated for the results on the right.



Most of programming in TheGamma is done through a single interaction principle – choosing a member from a list of members offered by the environment. A program is just a path that starts with a data source (such as “worldbank”) and follows through the data to find e.g. “CO2 emissions (metric tons per capita)” of a country such as “United States” as in the above example. You can construct programs by typing the data source name (such as “worldbank”) and then repeatedly typing “.” and choosing one of the offered items – either by clicking or via keyboard arrow keys and “Enter”.

In this study:

1. You will first watch a very brief presentation explaining TheGamma
2. You will be asked to try modify existing sample program to achieve a slightly different goal
3. We will discuss your experience, difficulties and things you learned

The problem:

The “olympics” data set contains information about Olympic medalists throughout the entire history of summer Olympic games. The following script groups the data by Team (that is, the country). For each country, it then sums the number of Gold medals that the country has won over the entire history. We then sort the data by the number of Gold medals, take top 5 countries and get a data series that contains the number of Gold medals for each Team. This is visualized as a bar chart.

olympics

.'group data'.'by Team'.'sum Gold'.then

.'sort data'.'by Gold descending'.then

.paging.take(5)

.'get series'.'with key Team'.'and value Gold'

We want to write a similar script that calculates the individual athletes with the greatest number of gold medals from the London 2012 Olympic games. To do this, you will first need to filter the data (to get only medals from London 2012 games). You will then need to group the data by Athlete (rather than by Team). The rest of the script will stay very similar.

To get started, open <http://bit.ly/turing-study-4> in your web browser!