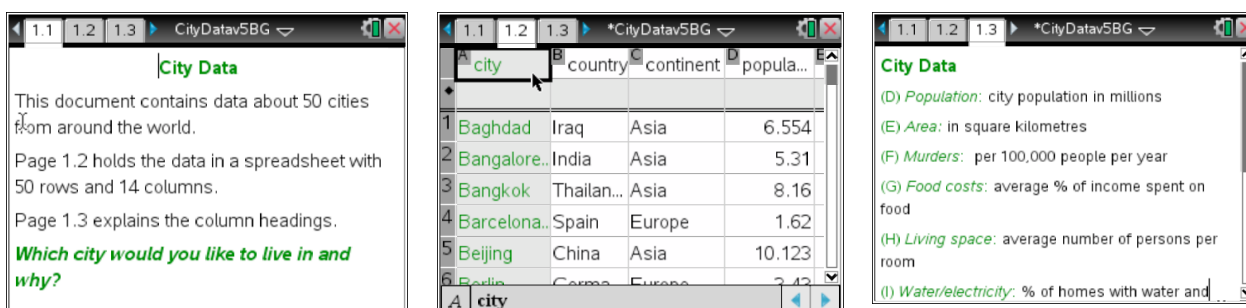# City Data
Teacher Notes

Introduction

This activity allows students to explore data about 50 world cities using the various applications of TI-Nspire.

Accompanying these teacher notes there are handouts for students (2 pages) and a tns file entitled CityData.tns.

This activity has been developed by Barrie Galpin from an initial idea by Andy Kemp of Taunton School, Somerset.

Words of warning about these data

The origin of much of this data seems lost in the mists of time. An original data file of 50 world cities was compiled some years ago (we would be absolutely delighted to acknowledge the original compilers!) What is clear is that a lot of the original data are now well out of date. We have updated, as far as possible, the population data and added a column of city areas but we have not been able to update or check much of the remaining numerical data.

A more fundamental problem when comparing data about cities is that what comprises a city is far from well defined. Take London for example. The City of London has an area of just over one square mile and a population of about 8000, but the entity generally known as London is said to have a population of 7-9 million, depending on where and when you draw the boundary. Most cities in the world are similarly ill-defined. To that must be added the fact that the populations of most cities in the developing world are expanding dramatically as people move away from rural areas. All this makes it almost impossible to compare like with like and even to answer apparently simple questions such as what is the biggest city in the world?

Very often students are given data in mathematics that have been pre-checked and sanitised. This document provides them with some dirty data and encourages them to criticise, correct or reject some of it before moving on to analyse it. We believe that the data provide a wealth of worthwhile activities for students, raising issues that have relevance well beyond the confines of the mathematics curriculum.

Assumed skills

Students will need only basic familiarity with TI-Nspire. They will need to be able to:
    select and open the document CityData.tns;
    move between pages of the document;
    use the cursor to move around the cells of the spreadsheet.

TI-*nspire*™

City Data
Student Handout

Getting started

On your handheld open the document called CityData.tns.

Have a look at the data in the spreadsheet on page 1.2.

The numbers are explained on page 1.3.

Let's be critical

Before you use data like these you should always think critically about them.
What do the numbers mean? Look carefully at page 1.3 and think carefully about how these statistics might have been calculated.
Can you believe these numbers? Are there any mistakes?
Can it be corrected or is it better ignored?

Is the data "dirty"?

Scroll across and down through the tables of numbers - do any of them stand out as a bit odd?
Are any seriously suspect? Can you correct them or delete them?
Are there any gaps? Why could this be? Can you fill the gaps or can you ignore them?

Where are these cities?

Why do you think these 50 cities have been chosen?
Are they evenly spread across the continents?
One way of checking this is to plot the continents on a Data & Spreadsheets page: press ctrl doc▾ 5
Now click on the vertical axis to add a variable and choose **continent**.

Secondary schools

Look down column K, which is said to show the percentage of children in secondary school.
Do you believe this?
Can you explain these data?

Telephones

Look down column J which is supposed to show the number of telephones per 100 people.
What do you think "telephones" means here?
Could you check the likely validity of this statistic for the UK cities?

How many people?

Can you get an idea of the spread of populations of the 50 cities? Which is the largest and which the smallest?

You could sort the data by population but an easier way is to draw a Box Plot.

On the Data & Statistics page remove the variable from the vertical axis by pressing ⬚menu ⬚2 ⬚A.

Then enter **population** as the variable on the horizontal axis.

At first you see a Dot Plot but you can change this using the Plot Type menu: press ⬚menu ⬚1 ⬚2.

As you move the cursor over the plot you will see the extreme values picked out as well as the median and quartiles.

Which city has the largest population?
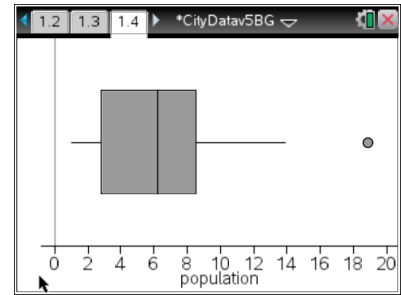
Now change the variable on the horizontal axis to **area**.

Which city in this data set has the largest area?

How packed are the cities?

You can calculate the average number of people per square kilometre, or population density, by calculating **population** times 1000000 divided by **area**.

At the top of column O of the spreadsheet type "density".

In the shaded cell below it enter the formula shown using these key presses:

> ⬚= ⬚var (choose **population**)
> ⬚× 1000000
> ⬚÷ ⬚var (choose **area**)

What are the extreme values of population density?

Where in the world do you think the greatest population densities occur? Plot **continent** on the vertical axis and **density** on the horizontal to see if you are right.

Space to live

How many people live in each room in cities in different parts of the world? Can you produce the dot plot shown here.

Now display this as a histogram. Does this type of plot have any advantages?
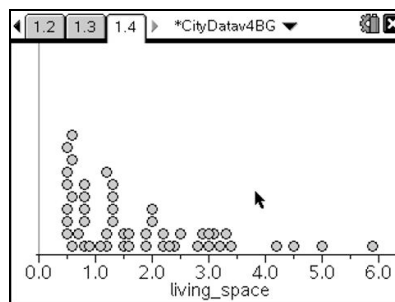
Can you produce a plot showing this information separated into continents?

Comparing variables

Do you think there might be a connection between

> **density** (number of people per square kilometre) and
> **living space** (average number of people per room)?

Draw a scatterplot of these two variables. Is this what you would expect?

Infant mortality

Investigate the way infant mortality rates vary across the cities in the various continents.

Do you think it might be linked to any of the other variables such as living space or water/electricity? Draw scatterplots to see if you can find a connection.

Choose another variable that interests you and investigate it in similar ways.

Notes about the student handout

Getting started

Here is the full list of explanations of the variables in the column headings

(D)	**population**:	city population in millions.
(E)	**area**:	in square kilometres.
(F)	**murders**:	per 100,000 people per year.
(G)	**food**:	average % of income spent on food.
(H)	**living space**:	average number of persons per room.
(I)	**water/electricity**:	% of homes with water and electricity.
(J)	**telephones**:	number per 100 people.
(K)	**secondary school**:	% of children in secondary school.
(L)	**infant deaths**:	deaths of infants (0-1 years) per 1000 live births.
(M)	**noise**:	level of background noise. (Scale of 1-10: 1 = low, 10 = high).
(N)	**traffic flow**:	average speed (m.p.h) in rush hour.

Is the data "dirty"?

There is a (deliberate) mistake in the population column: Hong Kong's population (a published estimate) should be entered not as 7,000,000 but as 7. This can be easily corrected. Students may come across other suspect pieces of data and this may provoke useful about whether to edit them or not.

Several of the columns do have gaps indicating, presumably, that no figure was available This will not affect the TI-Nspire's ability to draw plots – it will simply be ignored.

Where are these cities?

The 50 cities do include many of the largest in the world but they were probably originally chosen because the socio-economic data was available for them.

The "Dot Chart" indicates that cities are certainly not evenly spread across the continents. (e.g. there are none in Australasia).

Notice that the Plot Type menu offers alternatives. A Bar Chart and Pie Chart can be drawn and their suitability discussed.
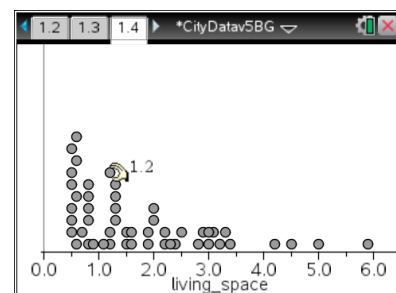
Secondary schools

This provides an opportunity to question the meaning and provenance of the data more carefully. In this column we see that 66% of childen in Birmingham are in secondary schools – perhaps indicating that 34% are in primary school. But Moscow is listed as 100% - are we to believe that all Muscovite children are in secondary school?

This may be a case where the definition of terms is not uniform across the world and we suggest that this column of data be treated with scepticism.

Telephones

It is likely that this data excludes mobile phones, referring only to fixed telephone lines. Wikepedia has an interesting chart showing the number of fixed telephone lines per 100 inhabitants between 1997 and 2007. (http://en.wikipedia.org/wiki/Telephone#Usage)

The "Developing World" hovers just above 50%, while the "Developing world" grew from 6 to 16% during that time. The data in the spreadsheet conform broadly to these percentages.

TI-*nspire*

How many people?

If you wish to sort the data, select all 50 rows (cursor left in row 1 until the row is shaded then hold ⇧shift and cursor down) and then choose Sort from the Actions menu.

Drawing a plot is probably more satisfactory. However, if any students get the plot shown here it is because the Hong Kong population is still entered as 7000000!
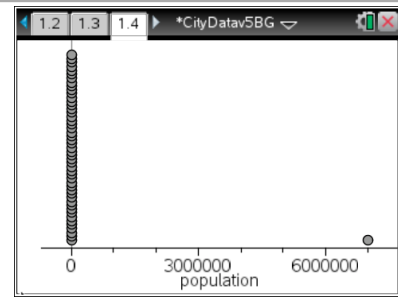
The largest population, 18.88, is for Shanghai. It is also recorded as having the largest area, 7037 km^2. You may wish to discuss with students the question of what constitutes a city. It is also instructive to look at discussion of population levels on Shanghai's official website: www.shanghai.gov.cn.

How packed are the cities?

Notice that options 2 and 3 in the Plot Properties menu allow you either to display a Box Plot's outliers or extend its whiskers.
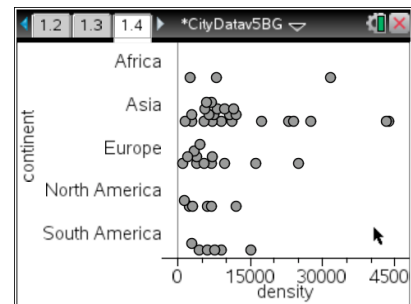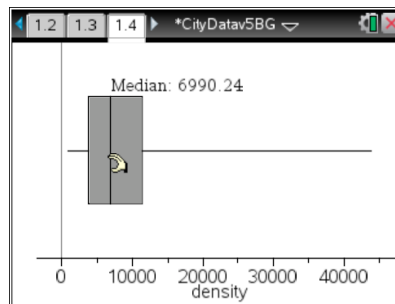
Space to live
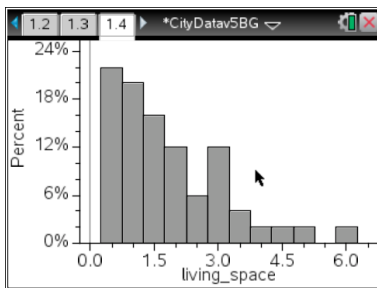
There are various histogram properties available in the Plot Properties menu, enabling a version such as that shown on the left to be produced. However, for the data separated into continents a Dot Plot (right) is probably preferable.
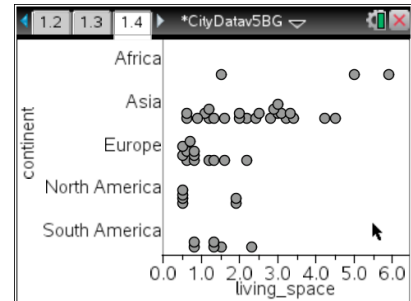
Comparing variables

The scatterplot is shown here.

Various regression facilities are available through the Analyse menu.
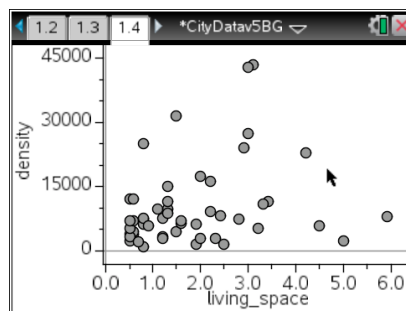
Infant mortality

If correlation is found between two variables it is important to stress that this does not necessarily imply that there is a causal relationship.