

### Countries of Asia

ID: **XXXX**

Name \_\_\_\_\_

Class \_\_\_\_\_

How do we make sense of large amounts of information? This is the essential task of statistics, and two main methods present themselves:

To use statistical calculations to reduce large amounts of data to small amounts of data, OR to view the information graphically.

In this activity, we use the statistical tools available to organise, to represent, and to seek to draw some conclusions from information about 25 countries in the Asia-Pacific region.

The screenshot shows a TI-nspire document window with a title bar containing tabs 1.1, 1.2, 1.3, and 1.4, and a menu bar with options RAD, AUTO, and REAL. The main content area has a title 'Countries of Asia' and the following text:

**Keywords:** Outlier, log transformation, regression.

**Source:** World Factbook (1995), Central Intelligence Agency, USA

Open the file *Stat\_ActXX\_Asian\_Countries\_EN.tns* on your handheld and follow along with your teacher to work through the activity. Use this document as a reference and to record your answers.

**Reference: OzDASL:** <http://www.statsci.org/data/oz/asia.html>

#### The Data

Country	Area	Population	Life	GDP	GDP/caput
Australia	7686850	18322231	77.78	374.6	20720
Bangladesh	144000	128094948	55.46	130.1	1040
Bhutan	47000	1780638	51.03	1.2	700
Brunei	5770	292266	71.24	4.4	16000
Burma(Myanmar)	678500	45103809	60.47	41.4	930
Cambodia	181040	10561373	49.46	6.4	630
China	9596960	1203097268	68.08	2978.8	2500
Hong.Kong	1040	5542869	80.18	136.1	24530
India	3287590	936545814	59.04	1253.9	1360
Indonesia	1919440	203583886	61.22	619.4	3090
Japan	377835	125506492	79.44	2527.4	20200
Laos	236800	4837237	52.20	4.0	850
Macau	16	490901	79.86	4.8	10000
Malaysia	329750	19723587	69.48	166.8	8650
Maldives	300	261310	65.49	0.36	1500
Nepal	140800	21560869	53.09	22.4	1060
North.Korea	120540	23486550	70.05	21.3	920
Pakistan	803940	131541920	57.86	248.5	1930
Philippines	300000	73265584	65.65	161.4	2310
Singapore	633	2890468	76.16	57.0	19940
South.Korea	98480	45553882	70.89	508.3	11270
Sri.Lanka	65610	18342660	72.14	57.6	3190
Taiwan	35980	21500583	75.47	257.0	12070
Thailand	514000	60271300	68.42	355.2	5970
Vietnam	329560	74393324	65.72	83.5	1140

**EXERCISES**

1. It is possible to put two properties such as **area** and **population** on the same chart - can you explain why this is not particularly useful in this situation?
2. Carefully describe your observations regarding area and population for these Asian countries.
3. Three countries in particular stand out: name these and explain why they are of interest in this context.
4. Could you think of any benefits that might follow from applying a logarithmic transformation to our data sets?
5. Carefully describe any advantages you see from using the logarithmic data sets. Are there any disadvantages?

**Extension:** Use these techniques now to present a report on the relationship between any other two data sets and what might be inferred in relation to Australia, China and India.

**SUGGESTED ANSWERS**

1. The very different scales make it difficult to see anything about area since it is swamped by population.
2. Most Asian countries are very similar in having very large populations in relatively small areas, i.e. the population density is high.
3. The exceptions are Australia (large area, small population), and China and India (large populations in large areas).
4. The log transformation keeps the order of data the same while reducing the spread. In this context, it could allow us to better place both data sets on a common scale.
5. The raw data clearly exposes our three focus nations as outliers, but does not allow us to see anything about the other nations - they are too closely clustered. The log data draws these others out and allows them to be studied. A linear regression of the log data is far more accurate than for the raw data, implying an exponential fit would be better.

**EXTENSION:** Some interesting relationships appear to exist between population and GDP (but less so with GDP per capita), between population and life expectancy, and between life expectancy and GDP per capita.

