



Beware of outliers: Student worksheet

<http://topdrawer.aamt.edu.au/Statistics/Misunderstandings/Misunderstandings-of-averages/Outliers>

The table below gives the salt content of 30 products in Australian supermarkets. The salt content is measured as the number of milligrams of salt per 100 grams of the product.

Product	Sodium content
Tomato sauce (Brand A)	871
Tomato sauce (Brand B)	1250
Soy sauce	6458
Tinned tomatoes	270
Four bean mix	250
Corn kernels	205
Mushrooms (tinned)	340
Tomato paste	843
Chunky pasta sauce	482
Chicken stock	521
Beef stock	450
Peanuts	360
Beef stroganoff mix	780
Dry roasted macadamias	340
Cashews	510
Butter	460
Sunflower spread	380
Baked beans	290
Peanut butter	458
French onion soup	335
Chicken noodle soup	355
Margarine original	590
Pink salmon	600
Gravy mix supreme	595
Vegetable stock	547
Precooked meat	1036
Spaghetti	410
Mixed nuts with green pistachios	190
Red salmon	530



1. Plot the sodium content, either by hand or by entering the data into a software package, using a stacked dot plot.
2. Describe the shape of the plot and note the scale on the horizontal axis.
3. What is the mean of the data set?
What is the median of the data set?
Explain why these values are so different.
4. Replot the data set without the value for soy sauce.
5. Describe the shape of the plot and note the scale on the horizontal axis.
6. What is the mean of this reduced data set?
What is the median of this reduced data set?
What is the relationship of these values now?
Which one (mean or median) is different for the two plots?
Which one (mean or median) is the same for the two plots?
Why did the change occur?
7. Did you use the same or a different scale for the two plots?
How did this affect the shape of the data distribution?
8. If you were writing an article for your local newspaper on the salt content of food, which graph and which measure of centre would you use? Explain.
9. Write a headline for your article.
10. (Optional) Write an article about salt content in food.