Chapter 1

Statistical variables

Exercise 1 *Answer* :

1. The scale is ordinal. There is an inherent ordering in that a Major is higher than a Captain, which is higher than a Lieutenant.
2. Since clothes are categorized and have no inherent order, the scale is nominal.
3. The scale is interval because there are equal intervals between temperatures but no true zero point.
4. It is ordinal because higher scores are better than lower scores. However, there is no guarantee that the difference between, say, a 2 and a 3 represents the same difference in knowledge as the difference between a 4 and a 5.
5. The city that someone was born in has no inherent order, thus can only be a nominal scale.
6. Most statisticians agree that it is valid to compute means of ordinal data, although some vehemently disagree.

Exercise 2 *Answer* : For each observation, the following eight variables are available through the command :

```
> str(Forbes2000)
'data.frame':  2000 obs.  of 8 variables:
$ rank :  int 1 2 3 4 5 6 7 8 9 10 ...
$ name :  chr "Citigroup" "General Electric" "American Intl Group" "ExxonMobil" ...
$ country :  Factor w\ 61 levels "Africa","Australia",..:  60 60 60 60 56 60 56 28 60 60 ...
$ category :  Factor w\ 27 levels "Aerospace & defense",..:  2 6 16 19 19 2 2 8 9 20 ...
$ sales :  num 94.7 134.2 76.7 222.9 232.6 ...
$ profits :  num 17.85 15.59 6.46 20.96 10.27 ...
$ assets :  num 1264 627 648 167 178 ...
$ marketvalue:  num 255 329 195 277 174 ...
```

Let us mention that

- *rank* denotes the ranking of the company (ordinal),
- *name*: the name of the company (nominal),
- *country*: the country the company is situated in (nominal),
- *category*: a category describing the products the company produces (nominal),

- *sales*: the amount of sales of the company in billion US dollars (interval),
- *profits*: the profit of the company in billion US dollars (interval),
- *assets*: the assets of the company in billion US dollars (interval),
- *marketvalue*: the market value of the company in billion US dollars (interval).

We next turn to describe phenomenologically the distributional properties of a single 1-D statistical variable X in a specific statistical sample S of size n, drawn in the context of a survey from some population of study objects of size N.