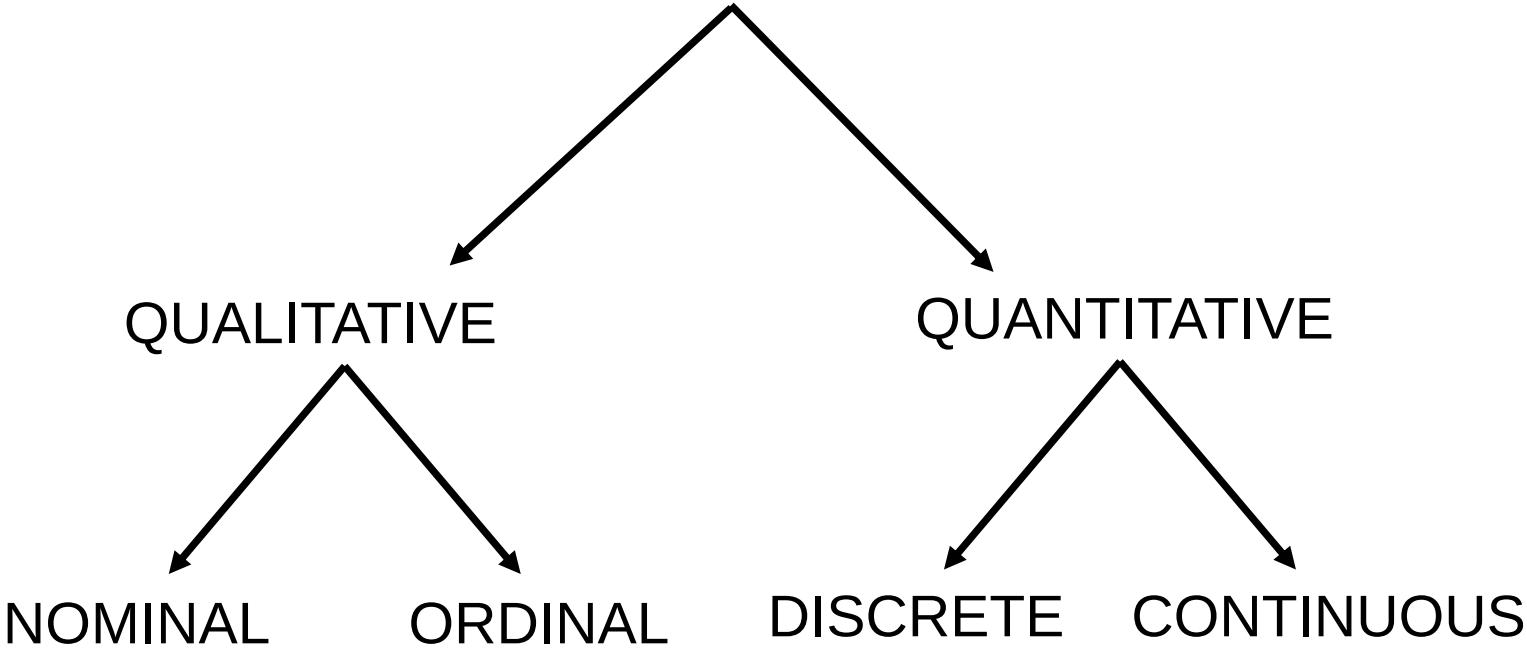
Definitions

- A **variable** is a characteristic or condition that can change or take on different values.
- **Datum** is one observation about the variable being measured.
- **Data** are a collection of observations.

TYPES OF VARIABLES



Qualitative variables

Nominal qualitative variable

Qualitative, nominal or categorical variable is data that comprises of categories that *cannot* be rank ordered – each category is just different.

What is your gender?
(please tick)

Male	
Female	

What is your favorite team?
(please tick)

Real Madrid	
Barcelona	
None	

Ordinal qualitative variable

It is a qualitative variable that comprises order.

How satisfied are you with the level of service you have received? *(please tick)*

Very satisfied	
Somewhat satisfied	
Neutral	
Somewhat dissatisfied	
Very dissatisfied	

Quantitative variables

Discrete quantitative variables

A quantitative variable with possible values of only specific points on a scale is called a discrete variable.

Number of children

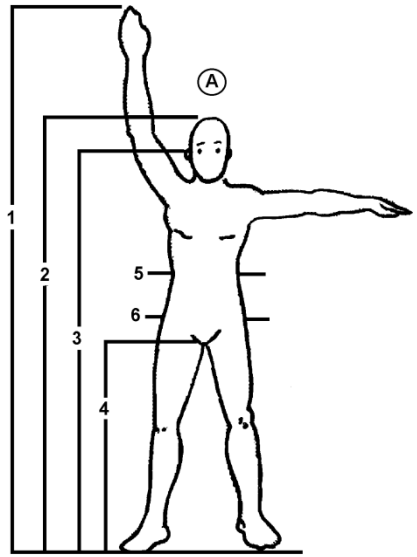
S={0, 1, 2, 3, …}

Number of head minus number of tails

S={-3, -1, 1, 3}

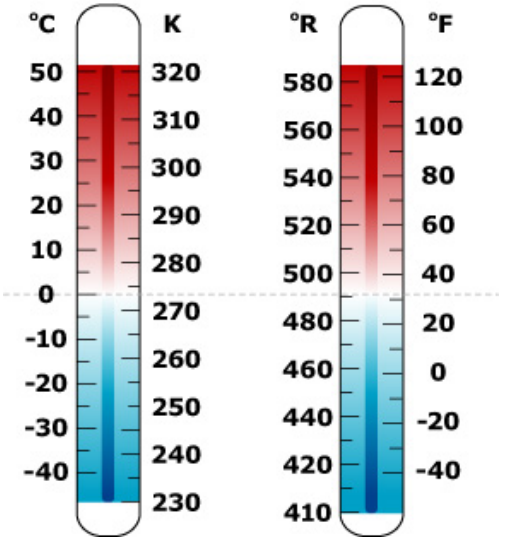
Continuous quantitative variables

This is a variable where the scale is continuous and not made up of discrete steps.



Interval variables

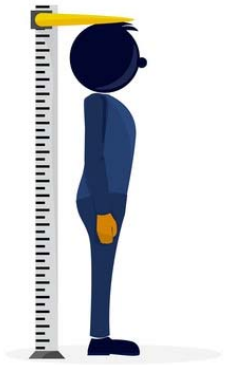
Interval variables measured on a *continuous* scale and has no true zero point. Examples:



Ratio variables

A *ratio* variable, has all the properties of an interval variable, and also has a clear definition of 0.0.

- Age
- Weight
- Height

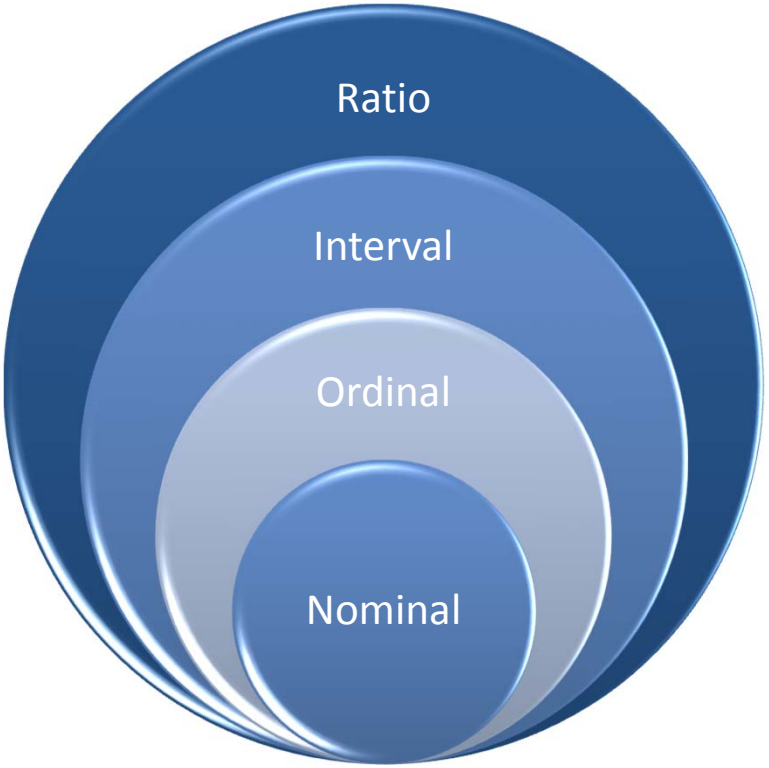


Are velocity and temperature continuous variables?

Depends on instrument

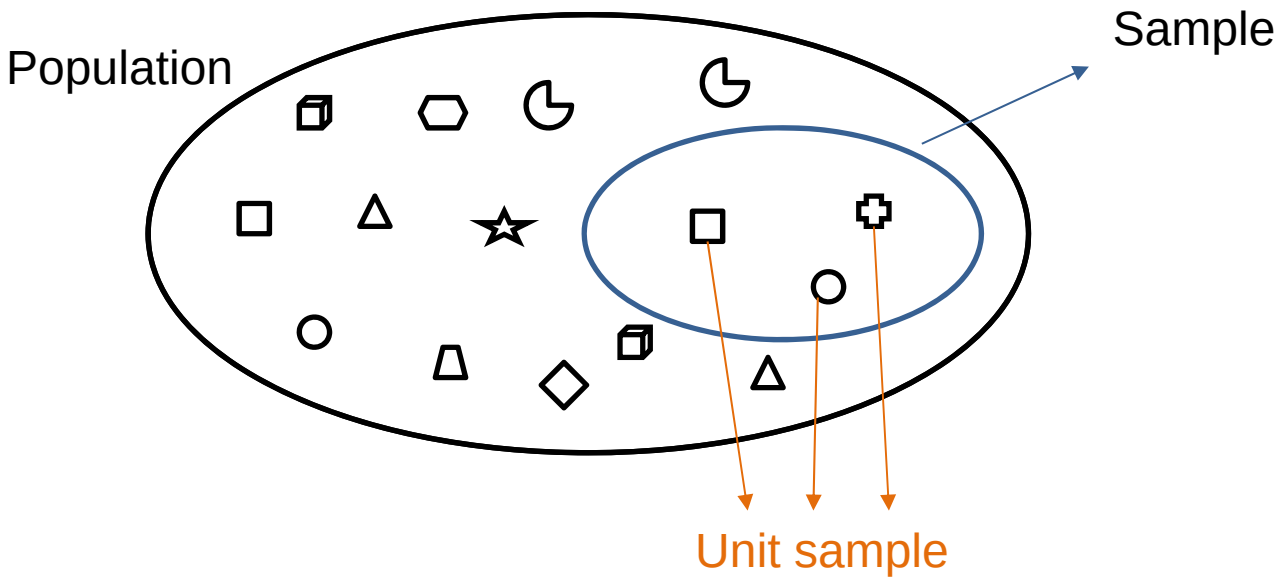
Hierarchical data order

These levels of measurement can be placed in hierarchical order.



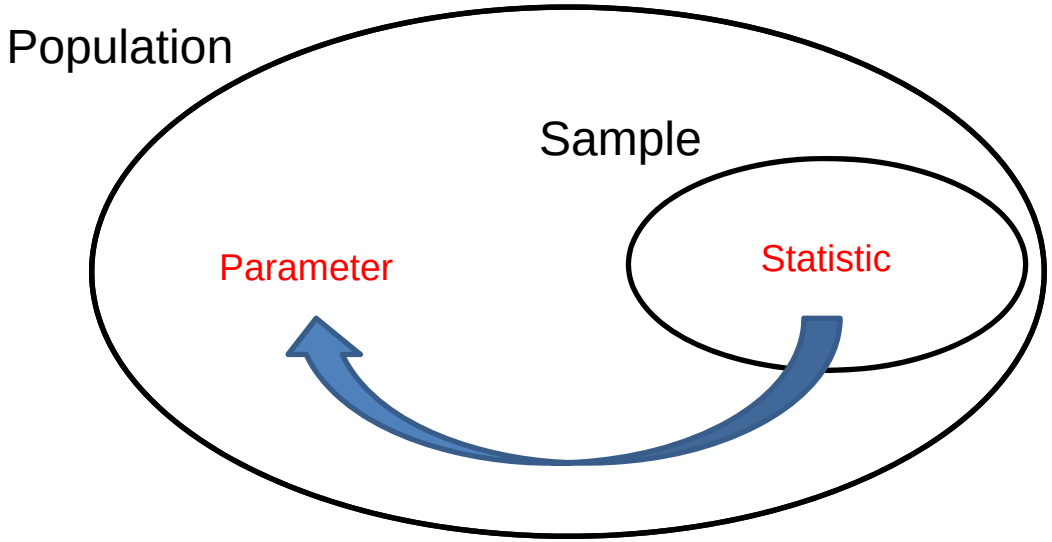
Population and Sample

A sample is a group of units selected from a population.



Parameter versus Statistic

A descriptive value for a population is called a **parameter** and a descriptive value for a sample is called a **statistic**.



How to organize the data?

Design matrix

Sex	Age	Smoke	Country	Married
Female	23	Yes	USA	Yes
Male	43	Yes	Colombia	Yes
Male	19	Not	Brazil	Yes
Male	23	Yes	Brazil	Not
Female	NA	Not	Canada	Yes
Female	78	Yes	USA	Yes
Male	54	Not	Spain	Not
Male	76	Yes	Colombia	Not
Female	43	Not	Peru	Yes

9 Individuals

Dimension 9 x 5

5 Variables

Statistic tools

Tables

Tables
- Absolute frequency tables
- Relative frequency tables

Tables

One way frequency table

Number of children	Absolute frequency	Relative frequency
0	2	0,02 → = 2/93
1	23	0,25 → = 23/93
2	41	0,44
3	18	0,19
5	8	0,09
7	1	0,01
Total	**93**	**1**

For nominal, ordinal and discrete variables.

Tables

Two way frequency table

Gender\ Hobby	Dance	Sports	TV	**Total**
Male	2	10	8	**20**
Female	16	6	8	**30**
Total	**18**	**16**	**16**	**50**

For nominal, ordinal and discrete variables.

Tables

Two way relative frequency table

Gender\ Hobby	Dance	Sports	TV	**Total**
Male	0,04	0,20	0,16	**0,4**
Female	0,32	0,12	0,16	**0,6**
Total	**0,36**	**0,32**	**0,32**	**1**

For nominal, ordinal and discrete variables.

Tables

Two way relative frequency table

by row

Gender\ Hobby	Dance	Sports	TV	**Total**
Male	0,10	0,50	0,40	**1**
Female	0,53	0,20	0,27	**1**

For nominal, ordinal and discrete variables.

Tables

Two way relative frequency table

by column

Gender\ Hobby	Dance	Sports	TV
Male	0,11	0,63	0,50
Female	0,89	0,38	0,50
Total	**1**	**1**	**1**

For nominal, ordinal and discrete variables.

Tables

Frequency table

Age	Absolute frequency	Relative frequency
10-14	2	0,050
15-19	16	0,400
20-24	18	0,450
25-29	3	0,075
30-34	1	0,025
Total	**40**	**1**

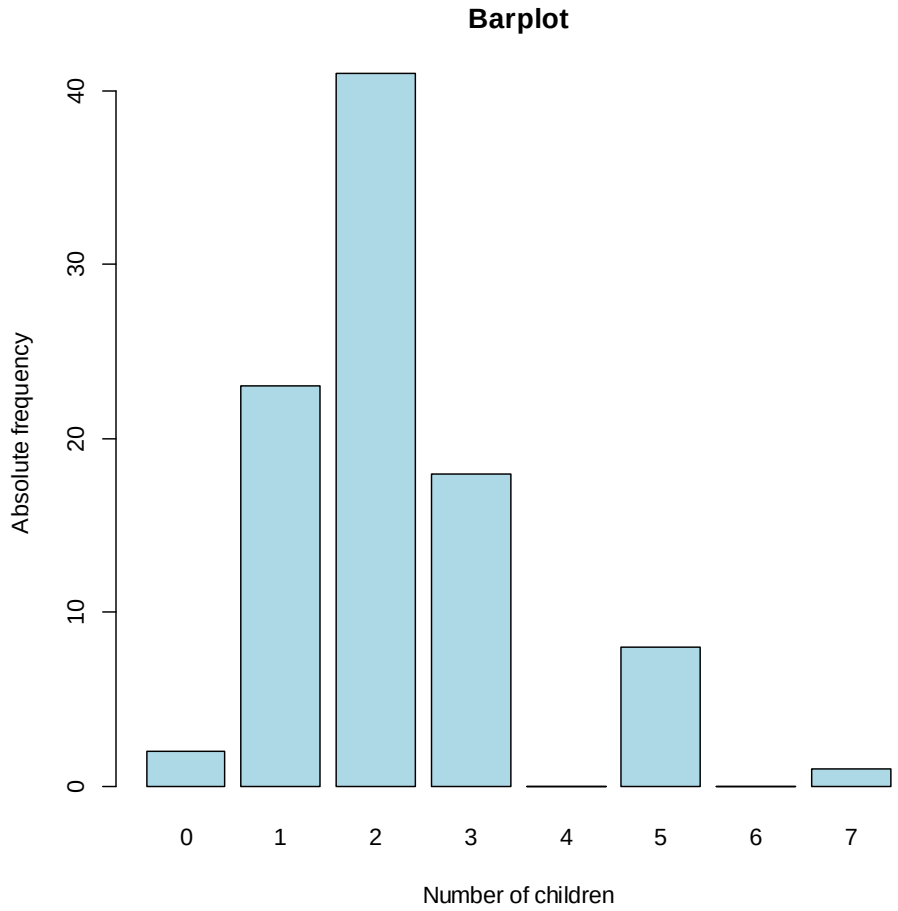
For quantitive variables.

Homework: check Sturges rule.

Graphs

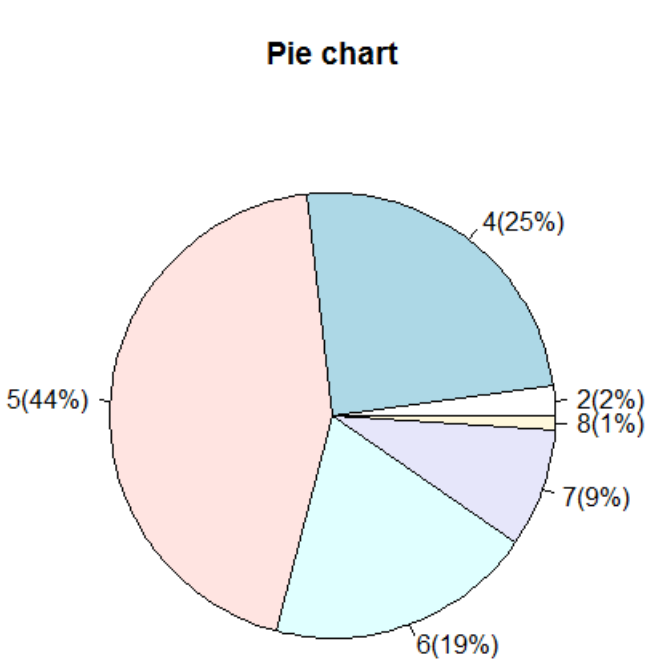
- Bar chart
- Pie chart
- Pictograms
- Histogram
- Density plot
- Scatter plot
- Time series plot
- Boxplot

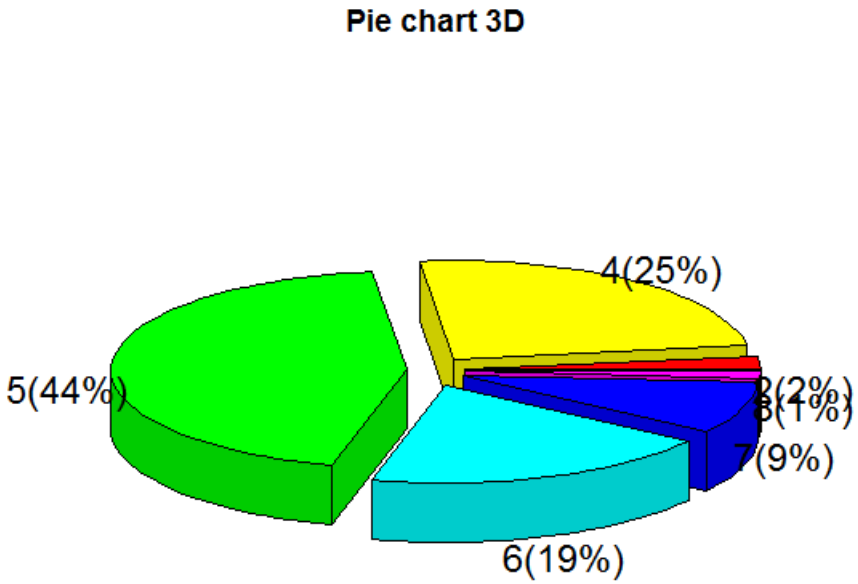
Graphs



For nominal, ordinal and discrete variables.

Graphs

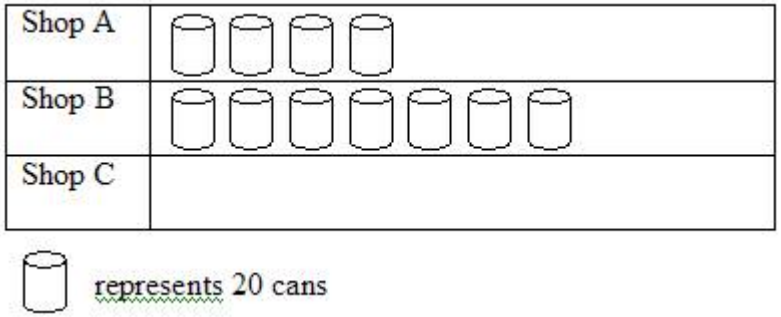




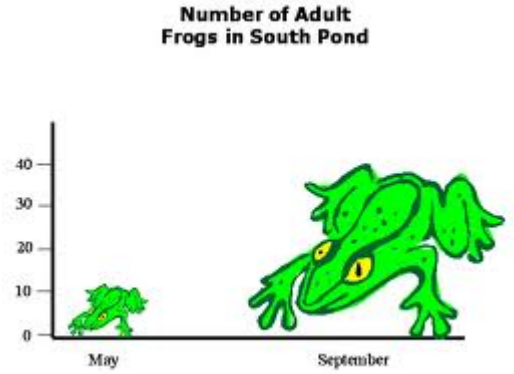
For nominal, ordinal and discrete variables.

Graphs

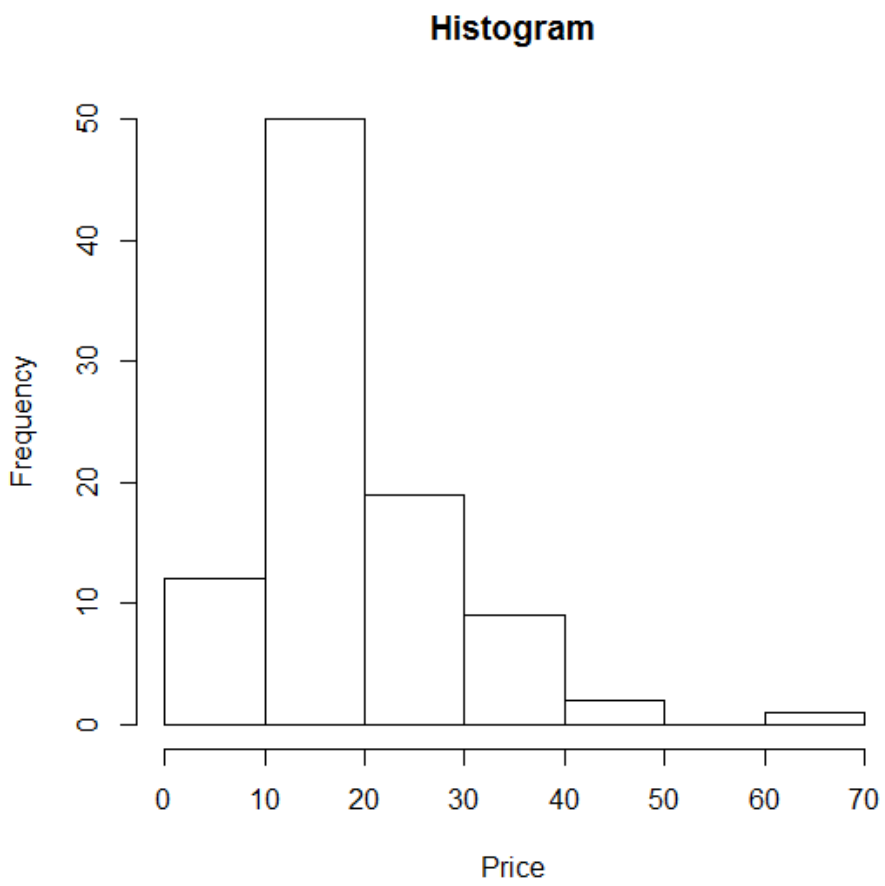
Statistic pictograms

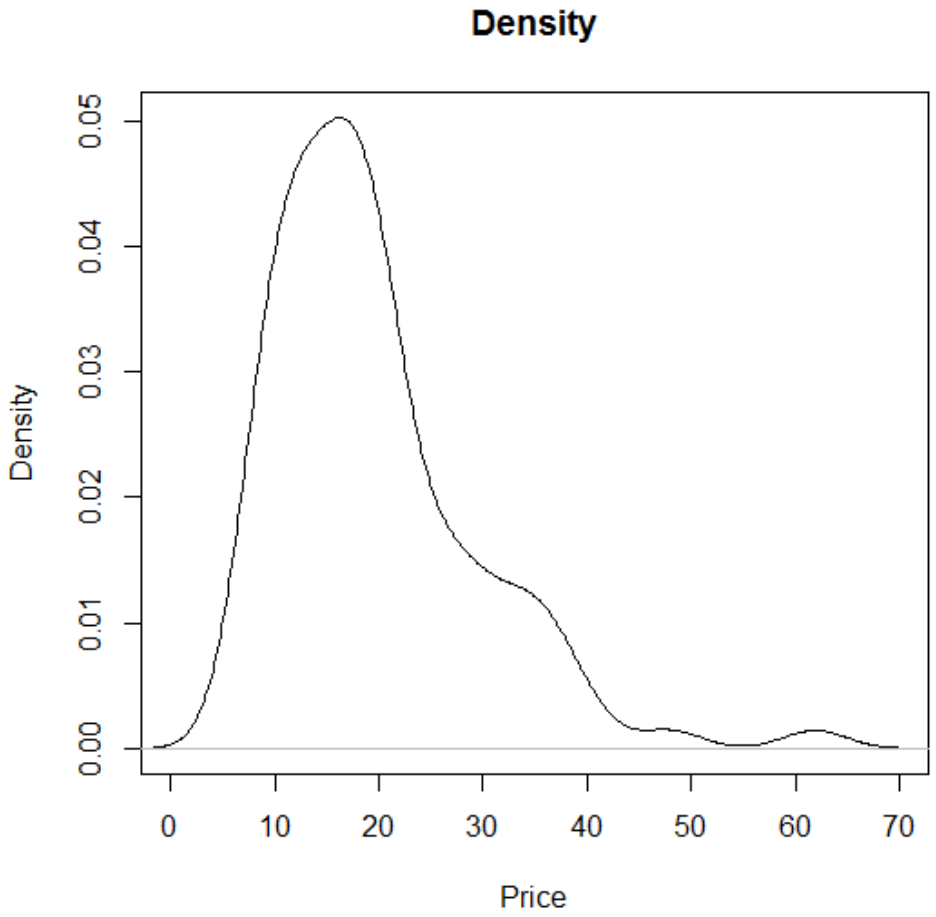


Do not recommended



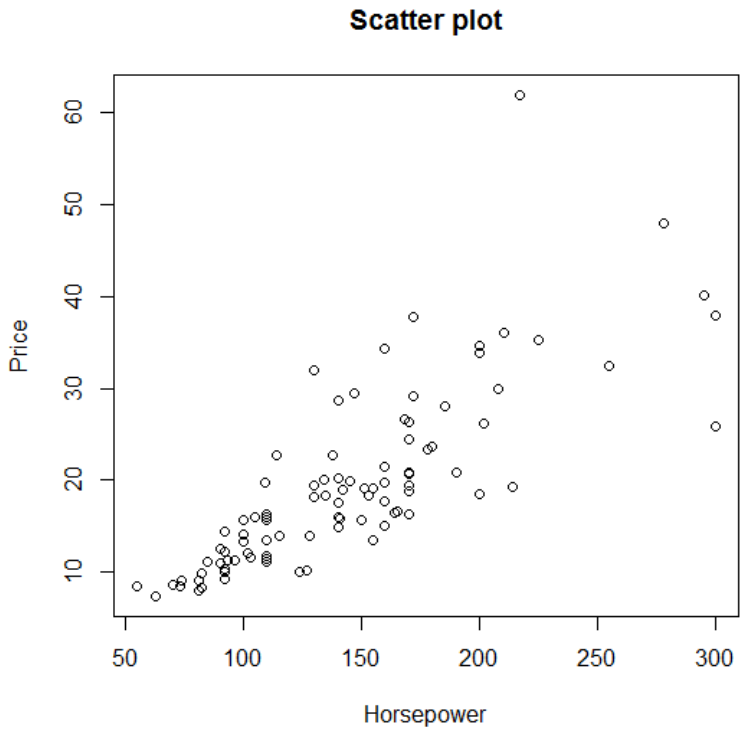
Graphs

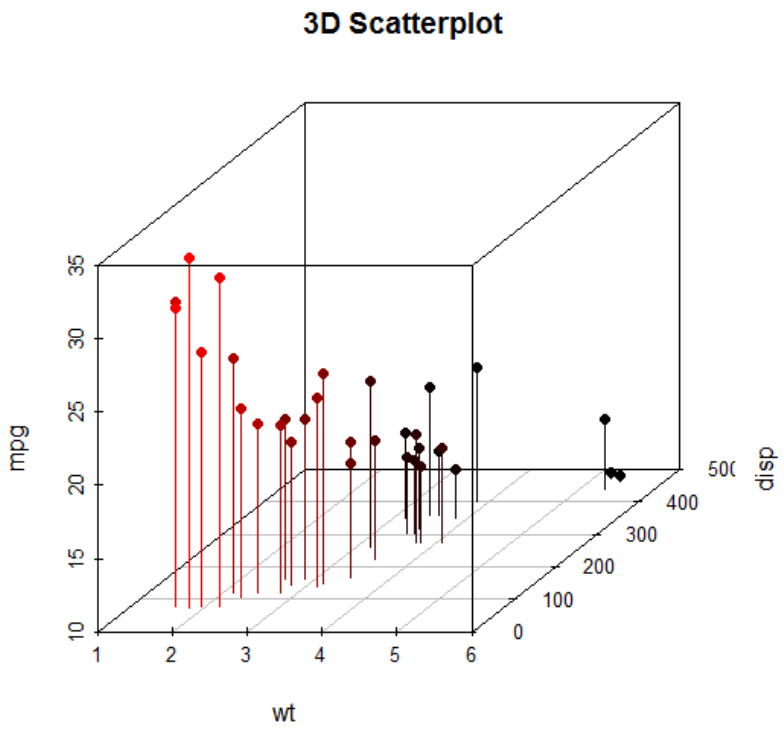




Only for numerical variables

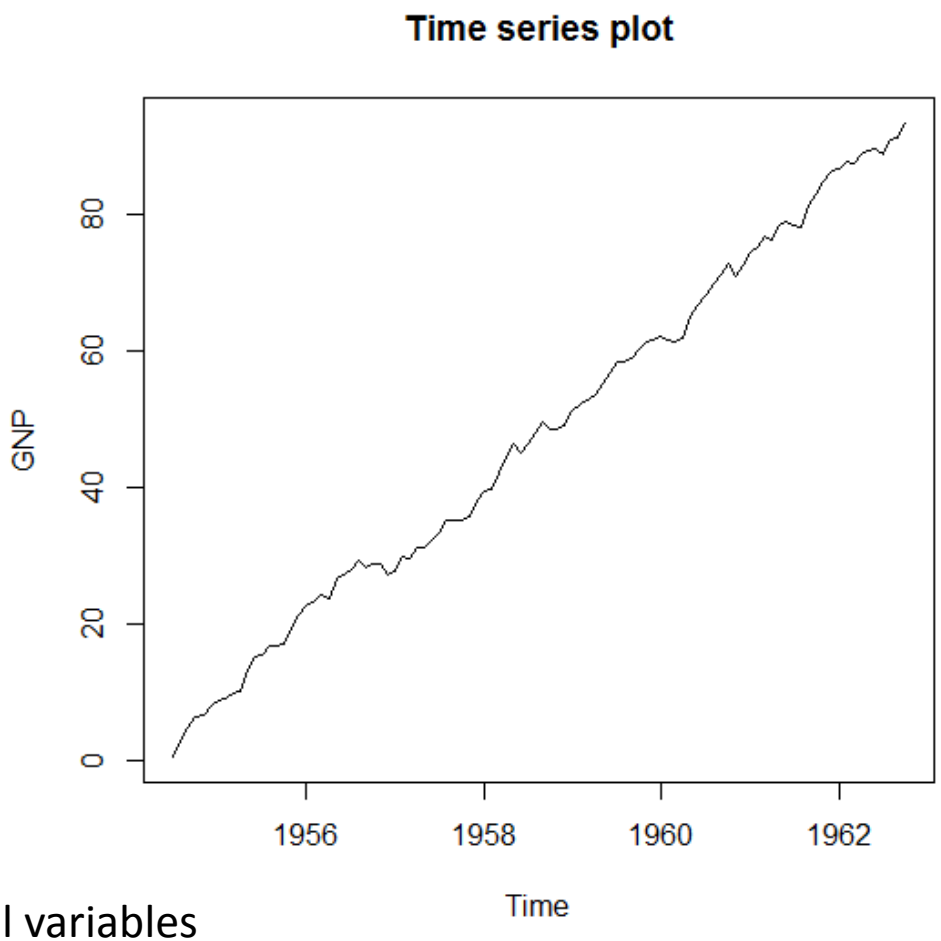
Graphs





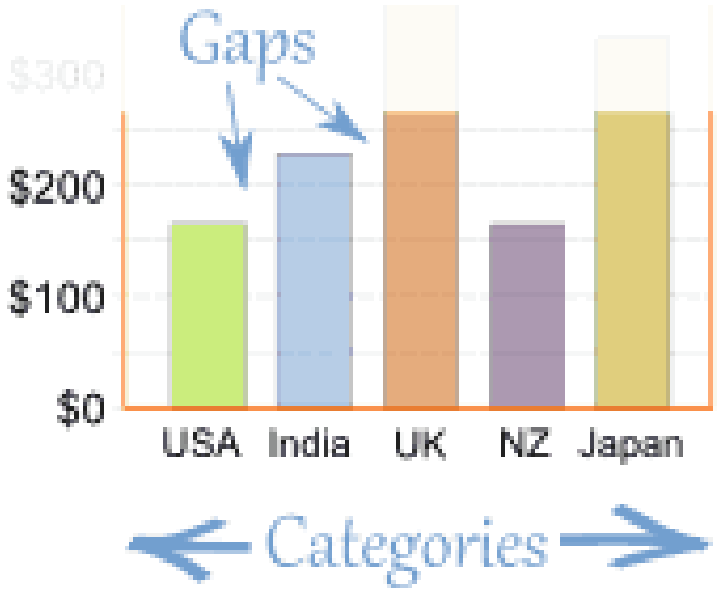
Only for numerical variables

Graphs

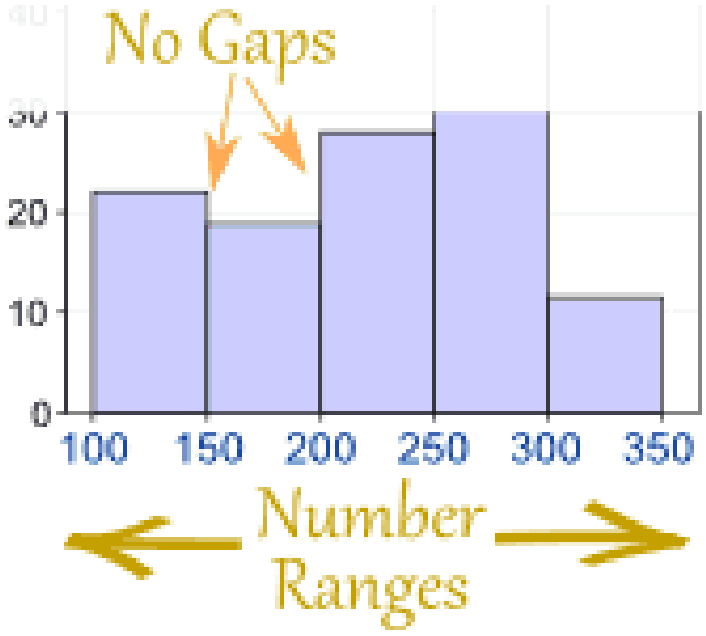


Only for numerical variables

Barplot vs histogram

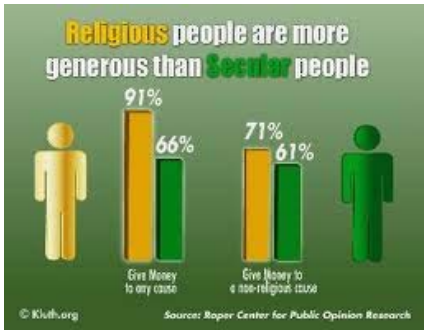


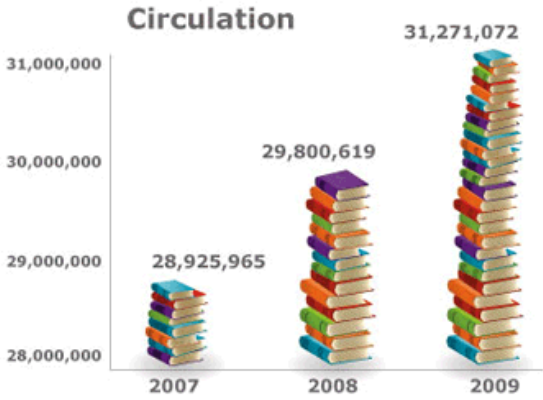
Bar Graph

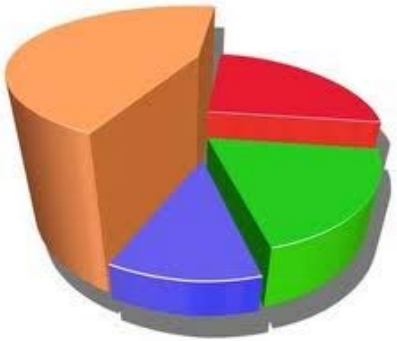


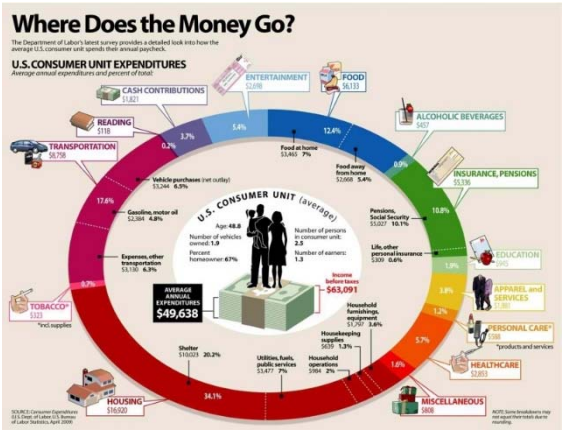
Histogram

Graphs examples on web

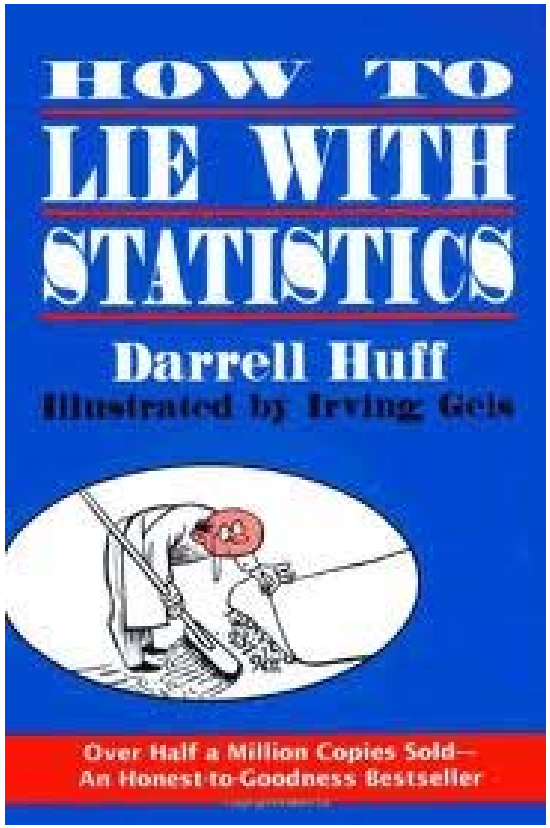








Recommended book



ENLACE

A cartoon


I DIDN'T HAVE ANYTHING USEFUL TO SAY SO I MADE THIS PIE CHART.
OOOH!
OOOH!
IT MUST BE TRUE BECAUSE IT'S PIE.
THAT WORKED TOO WELL.
I PLEDGE MY LIFE AND MY FORTUNE TO THE PIE!
www.dilbert.com

Recommended videos

http://www.youtube.com/watch?v=nUJNstRFvvo

http://www.youtube.com/watch?v=ETbc8GIhfHo

How are your data?