

# Medical Statistics, Informatics and Telemedicine

## Practice Lab #1

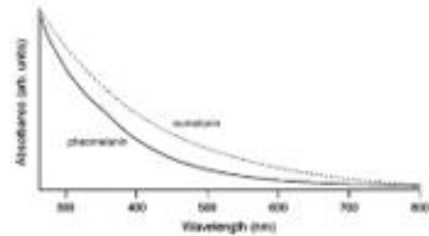
- Variable types
- Importing data to R Commander

repetition

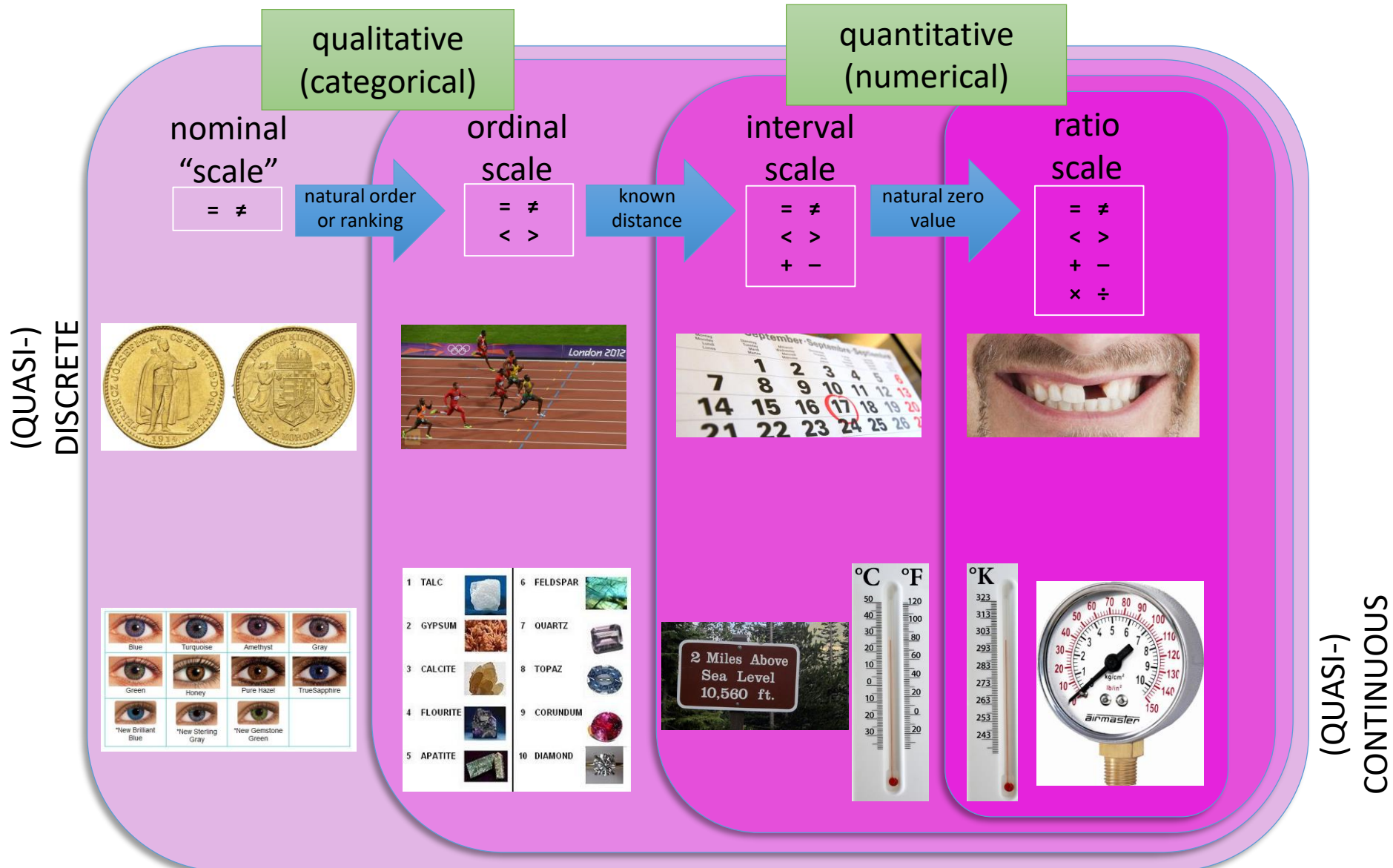
optional

# Variables

That we can measure or observe.



# Type of variables, measurement scales



# Importing data I.

- Importing data table from Excel

Use the *frmgham2\_period3\_NA\_factor.xls* data table

- At first open in excel and investigate:
  - Is the first row contains variable names?
  - How missing data is labelled?
  - What kind (type) of variable we have?

# Importing data I.

- Importing data table from Excel

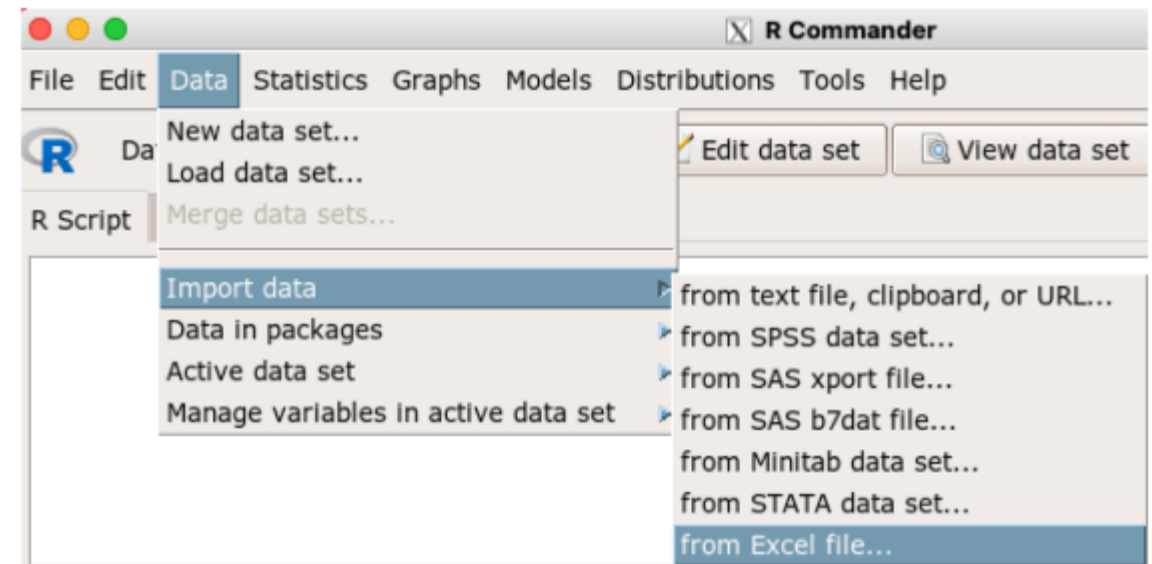
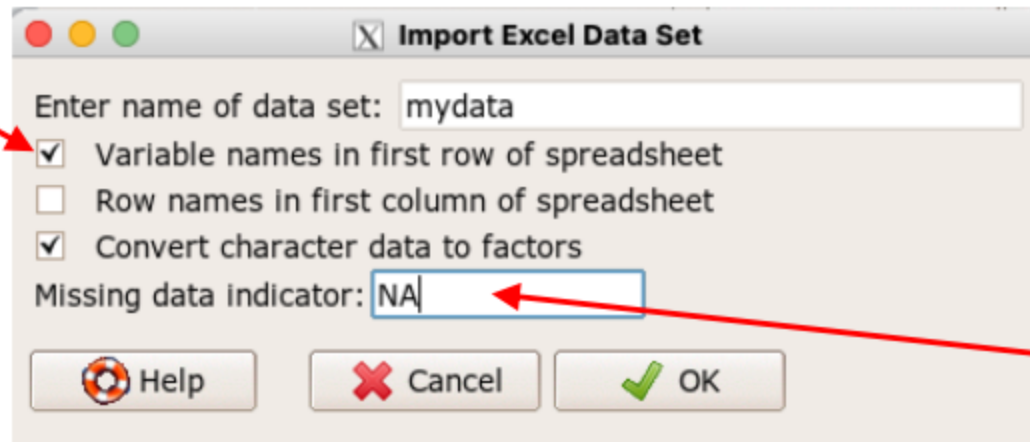
Use the *frmgham2\_period3\_NA\_factor.xls* data table

- At first open in excel and investigate:
  - Is the first row contains variable names?
  - How missing data is labelled?
  - What kind (type) of variable we have?

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
1	RANDID	SEX	TOTCHOL	AGE	SYSBP	DIABP	CURSMOKE	CIGPDAY	BMI	DIABETES	BPMEDS	HEARTRTE	GLUCOSE	EDUC	
2	2448	1	209	52	121	66 n		0	NA		0	69	92	4	
3	6238	2	237	58	108	66 n		0		28.5	0	0	80	71	2
4	11252	2	NA	58	155	90 y		30	24.61	0	0	74	NA		3
5	11263	2	220	55	180	106 n		0	31.17	1	1	86	81	2	3
6	12806	2	320	57	110	46 y		30	22.02	0	0	75	87	2	2
7	14367	1	280	64	168	100 n		0	25.72	0	0	92	82	1	1
8	16365	1	211	55	173	123 n		0	29.11	0	1	75	85	1	1
9	16799	2	291	62	120	70 n		0	21.98	0	NA	62	83	1	1
10	22727	2	150	62	124	78 n		0	26.62	0	0	68	125	2	2

# Importing data I.

Use the *frmgham2\_period3\_NA\_factor.xls* and import to R Commander



missing data:  
NA (all caps,  
without slash!)

# Importing data II.

- Import data from .csv file

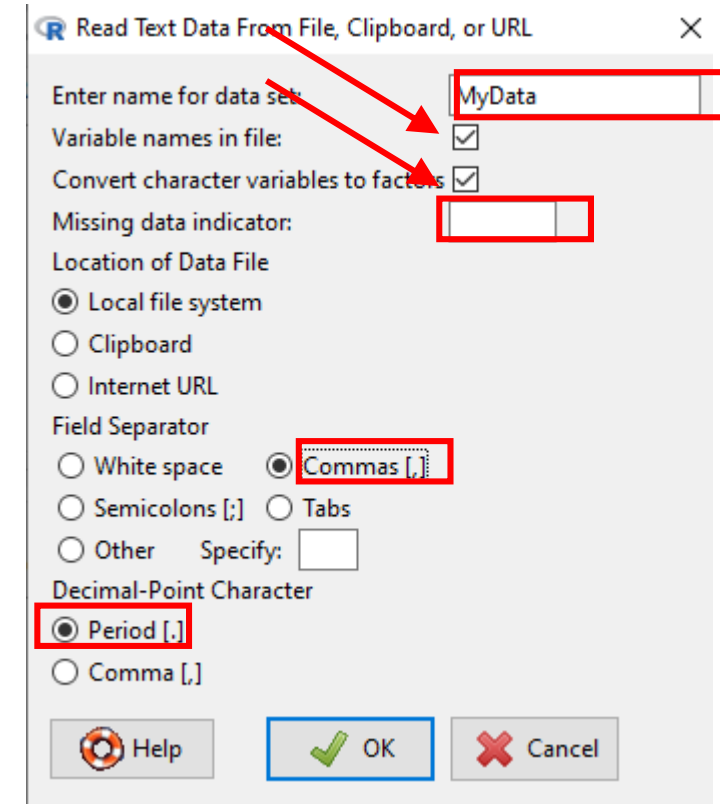
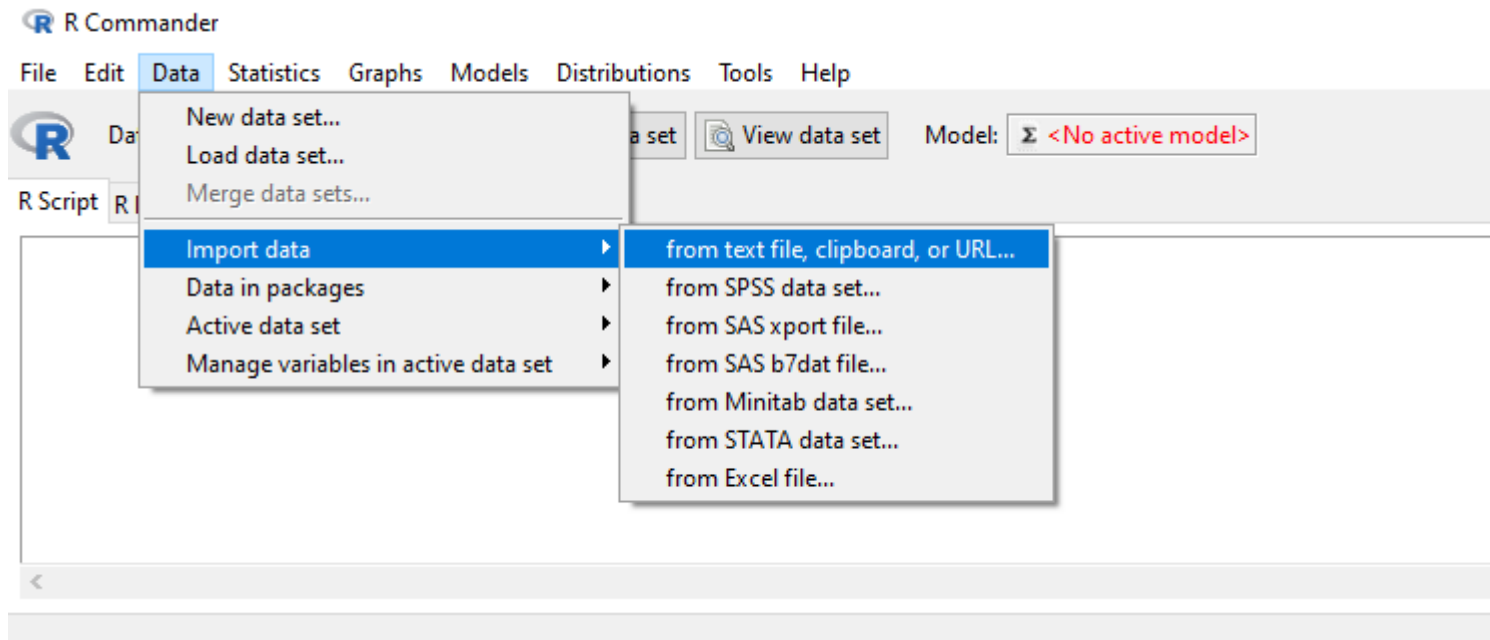
Use the *frmgham2.csv* data table

- We should know:
  - Is the first row contains variable names?
  - How missing data is labelled?
  - How the data („cells”) are separated?
  - What is the decimal separator?

# Importing data II.

- Importing data from .csv file

Use the *frmgham2.csv* data table





# View Data Table

R Commander

File Edit Data Statistics Graphs Models Distributions Tools Help



Data set:

MyData



Edit data set



View data set

Model:



<No active model>

R Script

R Markdown

MyData

	RANDID	SEX	TOTCHOL	AGE	SYSBP	DIABP	CURSMOKE	CIGPDAY	BMI	DIABETES	BPMEDS	HEART
1	2448	1	195	39	106.0	70.0	0	0	26.97	0	0	
2	2448	1	209	52	121.0	66.0	0	0	NA	0	0	
3	6238	2	250	46	121.0	81.0	0	0	28.73	0	0	
4	6238	2	260	52	105.0	69.5	0	0	29.43	0	0	
5	6238	2	237	58	108.0	66.0	0	0	28.50	0	0	
6	9428	1	245	48	127.5	80.0	1	20	25.34	0	0	
7	9428	1	283	54	141.0	89.0	1	30	25.34	0	0	
8	10552	2	225	61	150.0	95.0	1	30	28.58	0	0	
9	10552	2	232	67	183.0	109.0	1	20	30.18	0	0	
10	11252	2	285	46	130.0	84.0	1	23	23.10	0	0	
11	11252	2	343	51	109.0	77.0	1	30	23.48	0	0	
12	11252	2	NA	58	155.0	90.0	1	30	24.61	0	0	