



Trainer: Bondor Cosmina-loana, PhD, MsC

# Descriptive statistics

Variables type

- A** ALWAYS
- S** SEEK
- K** KNOWLEDGE

# Objectives for descriptive statistics

**Scales of measurement**

**Variables type**

**Frequencies**

Central tendency

Dispersion

Relationships between two measurement

Ways to assess the reliability and the validity of measurements

Misleading charts examples

Exercises



Descriptive statistics 1

# There is a “common sense” statistics

You use "common sense" statistics when you answer to the question:

What would be appropriate to wear today?

- I have to plan a party:

How many beers should I buy? How much food?

What type of music should I prepare?

- Will I pass this exam?

“common sense” statistics is different than medical evidence

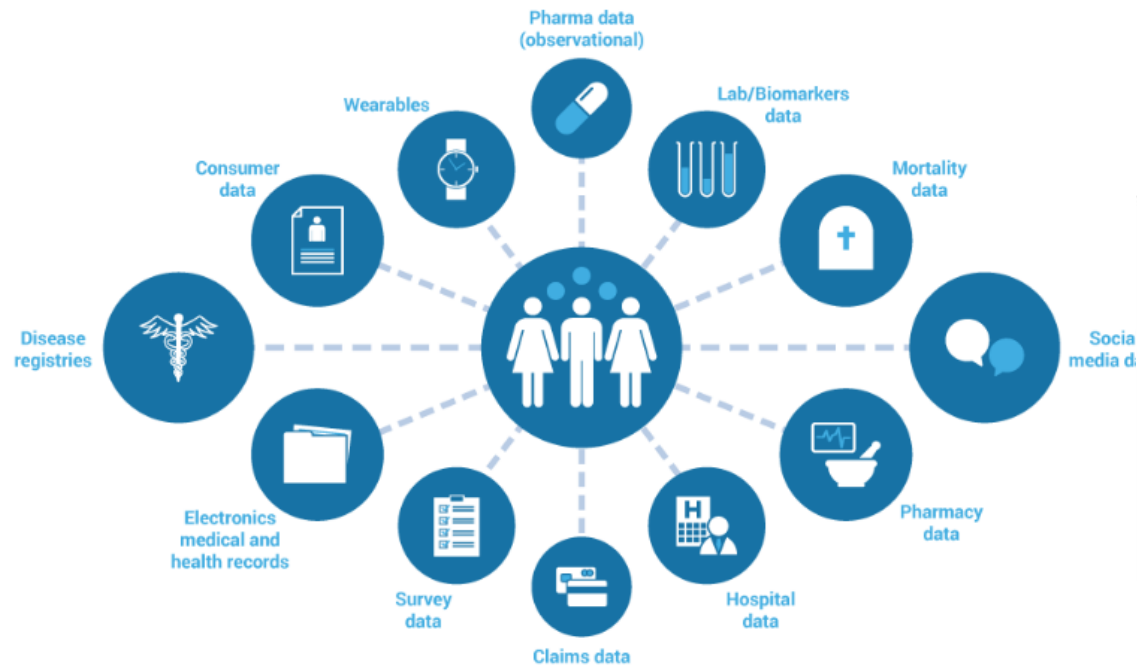
Medical research



Questions/  
Hypothesis



Design and  
Analytics



Evidence based medicine



Data can be found in many different environment



# Glycemia

- blood glucose concentration

healthy subjects

tested in the morning before food or water – glycemia a Jeun (fasting glycemia)

subjects with diabetes

tested at any time during the day (e.g. postprandial if it's taken after a meal)

– possible values in an interval

... 20-650 ... mg/dl



# What data can be memorized about - Glycemia

– Ex. 3 different possibilities:

- **measuring – record the measured value**

- possible values interval
  - ... 20-650 ... mg/dl

- **measuring, divide into categories**

- record the category

- < 70 mg/dl - hypoglycemia
- 70 – 120 mg/dl normal glycemia
- 120 – 180 mg/dl prediabetes
- $\geq 180$  mg/dl diabetes

- measuring, than divide in only 2 categories

- record Diabetes (yes/no)

- if glycemia is tested twice  $\geq 180$  mg/dl

	A	B	C	D
			Glycemia classification (hypoglycemia, normal glycemia, prediabetes, diabetes)	Diabetes (Yes/no)
1	Id_patient	<b>Glycemia</b>		
2		<b>1</b>	<b>70</b> normal glycemia	No

# What data can be memorized?

	A	B	C	D
1	Id_patient	<b>Glycemia</b>	Glycemia classification (hypoglycemia, normal glycemia, prediabetes, diabetes)	Diabetes (Yes/no)
2	1	70	normal glycemia	No
3	2	90	normal glycemia	No

# What data can be memorized?

The value of glycemia

The classification

The diagnostic

	A	B	C	D
1	Id_patient	Glycemia	Glycemia classification (hypoglycemia, normal glycemia, prediabetes, diabetes)	Diabetes (Yes/no)
2	1	70	normal glycemia	No
3	2	90	normal glycemia	No
4	3	95	normal glycemia	No
5	4	105	normal glycemia	No
6	5	120	prediabetes	No
7	6	200	diabetes	Yes
8	7	250	diabetes	Yes
9	8	150	prediabetes	No
10	9	100	normal glycemia	No
11	10	90	normal glycemia	No

# Variables and data

Variable = A characteristic of an individual – an attribute

It may take different values (**data**) depending on the individual from a **set of possible values**.

The opposite of a variable is a constant - a fixed value.

	A	B	C	D
1	Id_patient	Glycemia	Glycemia classification (hypoglycemia, normal glycemia, prediabetes, diabetes)	Diabetes (Yes/no)
2	1	70	normal glycemia	No
3	2	90	normal glycemia	No
4	3	95	normal glycemia	No
5	4	105	normal glycemia	No
6	5	120	prediabetes	No
7	6	200	diabetes	Yes
8	7	250	diabetes	Yes
9	8	150	prediabetes	No
10	9	100	normal glycemia	No
11	10	90	normal glycemia	No
12	11	70	normal glycemia	No
13	12	80	normal glycemia	No
14	13	90	normal glycemia	No
15	14	120	prediabetes	No
16	15	150	prediabetes	No
17	16	205	diabetes	Yes
18	17	95	normal glycemia	No
19	18	105	normal glycemia	No
20	19	75	normal glycemia	No
21	20	50	hypoglycemia	No

# When work with patient data: Ensure Ethical and Legal Framework

**Obtain informed consent** from patients or participants.

Adhere to regulations such as:

**GDPR** (General Data Protection Regulation - EU)

**HIPAA** (Health Insurance Portability and Accountability Act - US)

**Institutional Review Board (IRB)** approvals for research

Respect patient autonomy and confidentiality

# Is there a difference between these variables/data?

B	C
	Glycemia classification (hypoglycemia, normal glycemia, prediabetes, diabetes)
Glycemia	
70	normal glycemia
90	normal glycemia
95	normal glycemia
105	normal glycemia
120	prediabetes
200	diabetes
250	diabetes
150	prediabetes
100	normal glycemia
90	normal glycemia
70	normal glycemia
80	normal glycemia
90	normal glycemia
120	prediabetes
150	prediabetes
205	diabetes
95	normal glycemia
105	normal glycemia
75	normal glycemia
50	hypoglycemia

Data can be

- as a number → numerical, quantitative
- as a quality → qualitative

B	C
	Glycemia classification (hypoglycemia, normal glycemia, prediabetes, diabetes)
Glycemia	
70	normal glycemia
90	normal glycemia
95	normal glycemia
105	normal glycemia
120	prediabetes
200	diabetes
250	diabetes
150	prediabetes
100	normal glycemia
90	normal glycemia
70	normal glycemia
80	normal glycemia
90	normal glycemia
120	prediabetes
150	prediabetes
205	diabetes
95	normal glycemia
105	normal glycemia
75	normal glycemia
50	hypoglycemia

# Numerical scale of measurement

Individuals' data are evaluated on **scales of measurement**

## Interval scale

- quantitative characteristics, measurable in real numbers in a range

## Ratio scale

- quantitative characteristics, measurable in real numbers, which can be divided

## Discrete scale

- quantitative characteristics, measurable in integer numbers, decimals are meaningless

Ex: Interval: glycemia      Ratio: weight, age      Discrete: number of children in a family

Usually - have a unit of measurement

Answer the question: How much?

The difference between any two units is the same.

If we make a measurement,  
then the data are measured on a numerical scale,  
we measure a

quatitative variable

Measurement



numerical scale



quantitative variable

# Quantitative variable (numerical)

continuous

discrete

## Continuous quantitative variable

- It can be measured with decimals
  - ex. Weight, Age, Total cholesterol

## Discrete quantitative variable

- decimals are meaningless
- ex. Number of children, Number of pregnancies

G
Tooth number
12
13
12
14
12
13
14
13
12
14
9
12
10
10
12
14
10
15
14
14

B	J
Glycemia	Cranian perimeter (cm)
70	34
90	74
95	24
105	66
120	38
200	74
250	45
150	41
100	20
90	38
70	56
80	34
90	20
120	20
150	38
205	66
95	20
105	49
75	34
50	59
	31
	41

# Example - Body height

– continuous quantitative variable

variable name: Body height

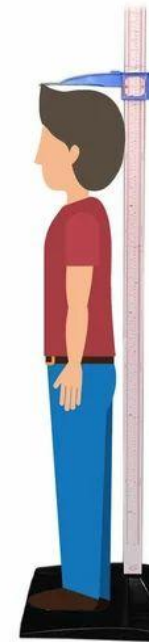
range of possible values: ...45-210... cm

We can measure it with a device

has unit of measure: cm

Decimals has meaning

Answer the question: How much?



The difference between any two units is the same.

e.g. between 50 and 51 cm the difference is the same as between 100 and 101 cm

K
Height (cm)
54
55
60
55
45
60
59
51
60
62
58
58
57
61
62
63
61
58
59
55

# Nominal scale

**Nominal scale** - qualitative characteristics, qualities of an individual

## Ex. Diabetes, Profession

- not measurable
- observable/classified/transform
- the data cannot be ordered

Answer the questions:

- How is it?
- what is it
- ...

e.g. viruses or bacteria in a cell culture

D	C
Diabetes (Yes/no)	<b>Profession</b>
No	pupil
No	studenta
No	computer sci
No	economist
No	unemployed
Yes	banker
Yes	unemployed
No	worker
No	worker
No	salesman
No	student
No	teacher
No	weiter
No	salesman
No	worker
Yes	unemployed
No	computer sci
No	ospatar
No	weiter
No	student

# Ordinal scale

## Ordinal scale

- qualitative characteristics, qualities of an individual
- Nominal data **with an order between data categories**

Ex. Stages I-IV of cancer, rheumatoid arthritis class I to IV

- The difference between the categories **is not the same**
- e.g. Mark at informatics 0-10

Measured on the ordinal scale

- scores
- percentages
- proportions

Apgar  
Score

10

10

9

10

10

8

8

9

9

10

10

10

10

6

10

10

10

7

7

8

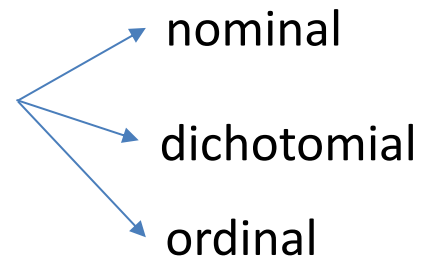
If we have observed/measured and classified (divided) into categories, then the values belong to one

## Qualitative variable

C	C	D
Glycemia classification (hypoglycemia, normal glycemia, prediabetes, diabetes)	<b>Profession</b>	Diabetes (Yes/no)
normal glycemia	pupil	
normal glycemia	studenta	No
normal glycemia	computer sci	No
normal glycemia	economist	No
normal glycemia	unemployed	No
prediabetes	banker	No
diabetes	unemployed	Yes
diabetes	worker	Yes
prediabetes	worker	No
normal glycemia	salesman	No
normal glycemia	student	No
normal glycemia	teacher	No
normal glycemia	weiter	No
normal glycemia	salesman	No
normal glycemia	worker	No
prediabetes	unemployed	No
prediabetes	computer sci	No
diabetes	ospatar	Yes
normal glycemia	weiter	No
normal glycemia	student	No
normal glycemia	dr	No
normal glycemia	salesman	No
hypoglycemia		No

Observed / divide in categories  nominal/ordinal scale  qualitative variable

# Qualitative variables



## Nominal variable

qualitative characteristics, qualities of an individual  
with **more than two** categories

## Dichotomial variable

qualitative characteristics, qualities of an individual  
with **two categories**

## Ordinal variable

qualitative characteristics, qualities of an individual  
has categories that **can be ordered**

C	H	Apgar Score
		10
		10
<b>Profession</b>	<b>Surgery</b>	9
pupil	No	10
studenta	No	10
computer sci	No	10
economist	No	8
unemployed	No	8
banker	No	9
unemployed	No	9
worker	No	9
worker	No	10
salesman	No	10
student	No	10
teacher	No	10
weiter	No	10
salesman	No	6
worker	No	10
unemployed	No	10
computer sci	No	10
ospatar	No	10
weiter	Yes	7
student	Yes	7
dr	Yes	8
salesman		

# Example - Maxillary

- Dichotomial qualitative variable

variable name: "Maxillary"

possible values (data): Superior/Inferior (two categories)

can't be measure (it has no unit of measure)

can be observe

The possible values cannot be ordered

D
Maxillary
superior
inferior
superior
inferior
superior
inferior
superior
inferior
inferior
superior
superior
inferior
inferior
inferior
inferior
inferior
inferior
superior
inferior
superior
superior

# Example - Dentition

## - Nominal qualitative variable

name of the variable “Dentition”

possible values (categories)

permanent, temporary, mixt, etc. (more than two categories)

can't be measure (it has no unit of measure)

can be observe

The possible values cannot be ordered

F
<b>Dentition</b>
mixt
permanent
permanent
permanent
mixt
temporary
permanent
permanent
mixt
permanent
mixt
permanent
temporary
temporary
mixt
mixt
temporary
permanent
permanent
permanent

# Ordinal qualitative variable

Ex. Any score

- depression score etc.
- Any grade
- grade in chemistry, physics, etc.

C	Apgar Score
Glycemia classification (hypoglycemia, normal glycemia, prediabetes, diabetes)	10
normal glycemia	10
normal glycemia	9
normal glycemia	10
normal glycemia	10
normal glycemia	8
prediabetes	8
diabetes	9
diabetes	9
prediabetes	9
normal glycemia	10
normal glycemia	10
normal glycemia	10
normal glycemia	10
normal glycemia	10
normal glycemia	6
prediabetes	10
prediabetes	10
diabetes	10
normal glycemia	10
normal glycemia	7
normal glycemia	7
hypoglycemia	8

# Example - Grade in Romanian language

– **Ordinal qualitative** variable

the name of the variable: Grade in the Romanian language

possible values (categories) 4,5,6,7,8,9,10

**there is an order** between the values:

10 reflects more acquired knowledge than 9, but not twice as much as 5

The difference between categories is not the same (grades are not quantities)

the difference between 4 and 5 means that did you pass the exam or not

the difference between 9 and 10 means that you had some more good answers

# Summary - Variables

**Qualitative variables type are:**

**Dichotomial**

**Ordinal**

**Nominal**

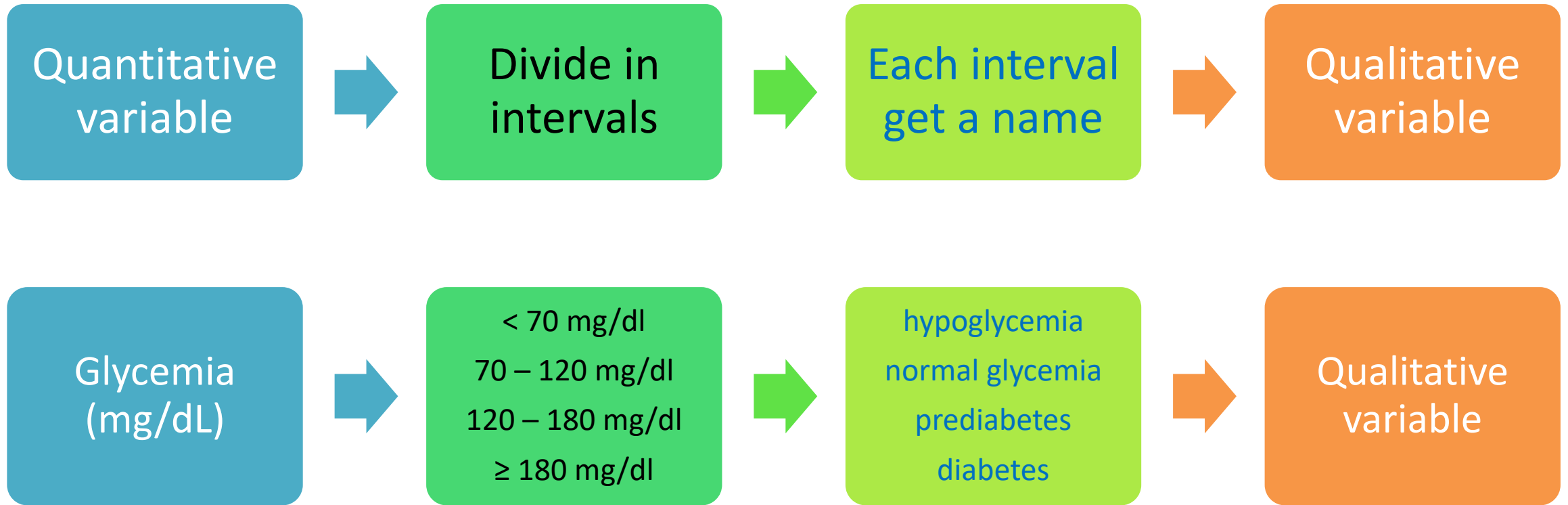
**Numerical/Quantitative:**

**Continuous**

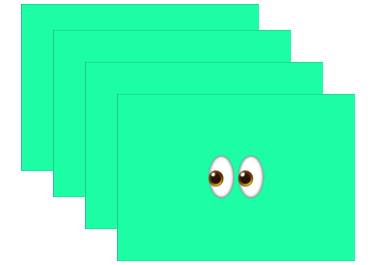
**Discrete**



# Transforming a variable



# Variables type



<https://www.wooclap.com/XZLYSC>

## Qualitative

Dichotomial

Nominal

Ordinal

## Numerical/Quantitative

Continuos

Discrete



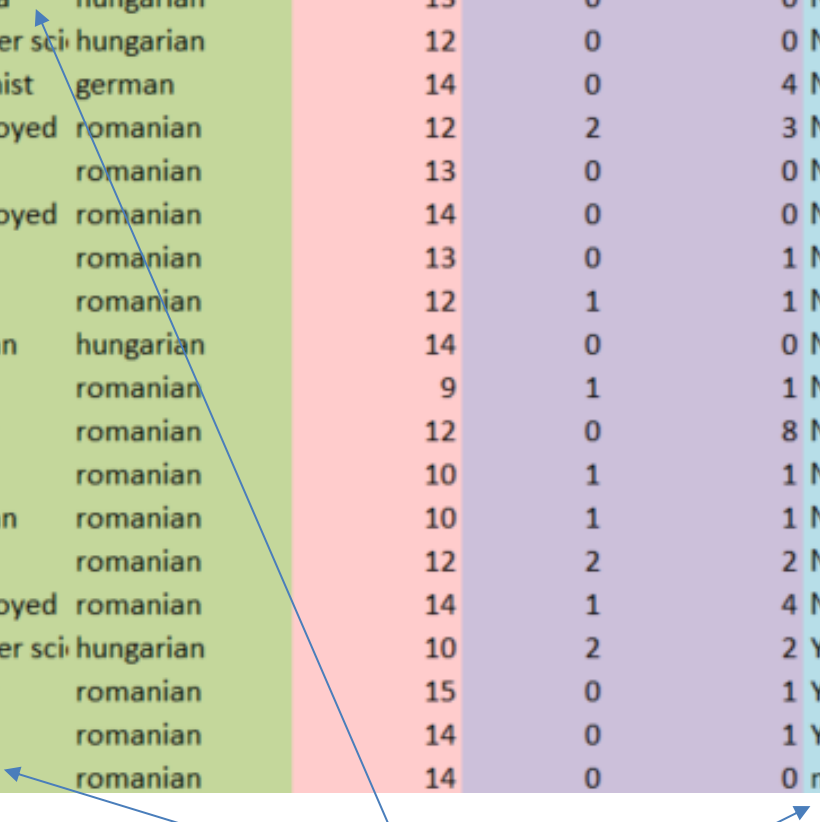
# Best Practices in Medical Data Collection Prepared for Statistical Analysis

- Plan
  - The variables
  - The type of the variables
  - How missing data will be handled - Collect the reasons
    - e.g. "patient refused", "not applicable"
- Define
  - the variables. Use dictionaries
  - the allowable data
    - use uniform coding
      - ex. use "Yes/No" or "Y/N" for the same variable. Not both.
- Use table data forms
  - e.g. rows = observations, columns = variables
  - Avoid free-text fields
- collect sufficient metadata: when/how/device/method
  - e.g. "blood pressure measured seated after 5 min rest"
- maintain logs of data cleaning or transformations
- collect all potential factors
  - e.g. age, sex, other diseases, etc.
- collect basic data, not the transformed/calculated data
  - e.g. we prefer glycemia in mg/dl
    - we can always change it to diabetes/no diabetes
    - the reverse is not valid
  - e.g. collect weight and height, instead of body mass index (BMI)
    - $BMI = \text{weight (kg)} / \text{height (m}^2\text{)}$

# Example

	A	B	C	D	E	F	G	H	I	J	K	L
1	Group	Mother age	Profession	Ethnicity	Number of pregnancy weeks at birth	Number of births	Number of pregnancies	Surgery	Sex at birth	Cranian perimeter (cm)	Height (cm)	Apgar Score
2	1	16	pupil	romanian	12	0	0	No	M	34	54	10
3	1	22	studenta	hungarian	13	0	0	No	M	74	55	10
4	1	26	computer sci	hungarian	12	0	0	No	F	24	60	9
5	1	34	economist	german	14	0	4	No	F	66	55	10
6	1	30	unemployed	romanian	12	2	3	No	F	38	45	10
7	1	26	banker	romanian	13	0	0	No	F	74	60	8
8	1	29	unemployed	romanian	14	0	0	No	M	45	59	8
9	1	28	worker	romanian	13	0	1	No	F	41	51	9
10	1	25	worker	romanian	12	1	1	No	M	20	60	9
11	1	24	salesman	hungarian	14	0	0	No	M	38	62	10
12	1	22	student	romanian	9	1	1	No	F	56	58	10
13	1	36	teacher	romanian	12	0	8	No	F	34	58	10
14	1	29	weiter	romanian	10	1	1	No	F	20	57	10
15	1	33	salesman	romanian	10	1	1	No	F	20	61	6
16	1	34	worker	romanian	12	2	2	No	F	38	62	10
17	1	31	unemployed	romanian	14	1	4	No	M	66	63	10
18	1	30	computer sci	hungarian	10	2	2	Yes	F	20	61	10
19	1	20	ospatar	romanian	15	0	1	Yes	M	49	58	7
20	0	25	weiter	romanian	14	0	1	Yes	F	34	59	7
21	0	21	student	romanian	14	0	0	nu	F	59	55	8

mistakes



# Descriptive statistics for qualitative variables

# Objectives

Summarizing qualitative variables

Frequencies

Tables

Charts

Gender	Dentition	Apgar Score
M	mixt	
M	permanent	10
F	permanent	10
F	permanent	9
F	mixt	10
F	temporary	10
M	permanent	8
F	permanent	8
M	mixt	9
M	permanent	9
F	mixt	10
F	permanent	10
F	temporary	10
F	temporary	10
F	mixt	6
M	mixt	10
F	temporary	10
M	permanent	10
F	permanent	7
F	permanent	7
F	permanent	8

# Represent qualitative variables

- dichotomial variables
- nominal variables
- ordinal variables with few categories

?

Please describe these data

What can you say?

Gender

M  
M  
F  
F  
F  
F  
F  
M  
F  
M  
M  
F  
F  
F  
F  
F  
F  
M  
F  
M  
F  
F

# Frequency table: Dichotomical variable

Gender have two possible values in our data (two categories)

Gender		
F		
M		

we get 20 patients:  
these are the data  
about their gender in  
the file

Gender  
M  
M  
F  
F  
F  
F  
F  
M  
F  
M  
M  
F  
F  
F  
F  
F  
F  
F  
M  
F  
M  
F  
F

20  
subjects

# Frequency table: Dichotomical variable

Gender	Number = Absolute frequency	
F	7	
M	13	
Total	20	

How many female?

How many male?

Verify:  $7+13=20$  – total number of subjects

Gender

M  
M  
F  
F  
F  
F  
F  
M  
F  
M  
M  
F  
F  
F  
F  
F  
F  
M  
F  
M  
F  
F

# Frequency table: Dichotomical variable

Female percentage =  $7/20 * 100 = 35\%$

Gender	Number = Absolute frequency	Percentage (%) = Relative frequency
F	7	35
M	13	65
Total	20	100

Male percentage =  $13/20 * 100 = 65\%$

Verify:  $35 + 65 = 100\%$

Gender

M  
M  
F  
F  
F  
F  
M  
F  
M  
M  
F  
F  
F  
F  
F  
M  
F  
M  
F  
F

# Chart – Pie chart

Gender	Number = Absolute frequency	Percentage (%) = Relative frequency
F	7	35
M	13	65
Total	20	100

Gender

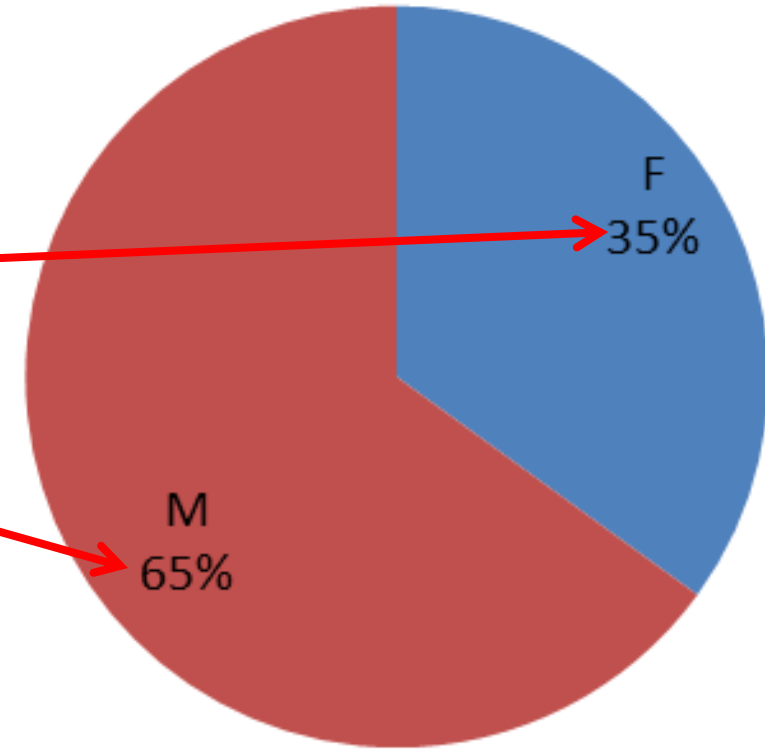


Figure. Gender percentages distribution

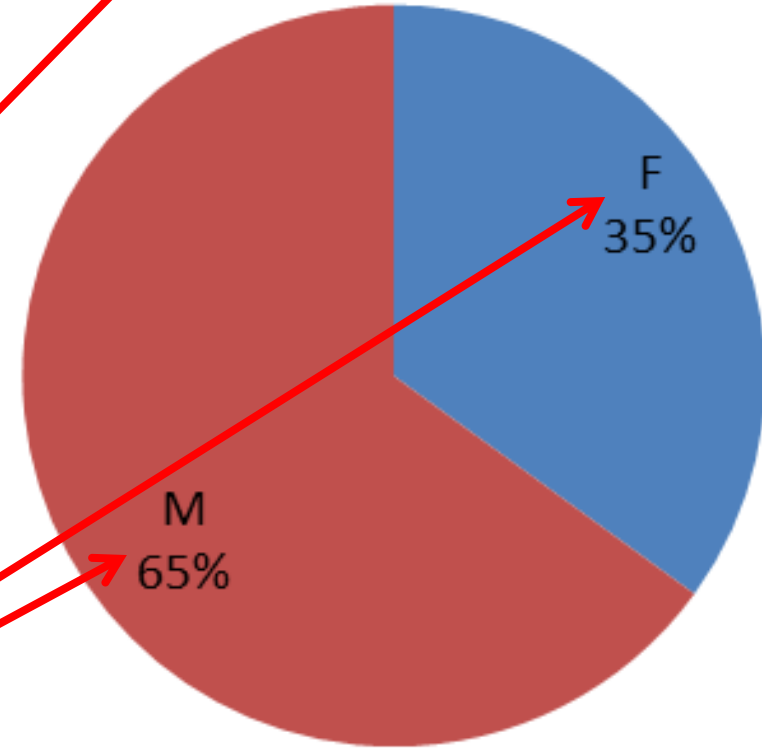
Gender

M  
M  
F  
F  
F  
F  
M  
F  
M  
M  
F  
F  
F  
F  
F  
M  
F  
M  
F  
F

# Chart – Pie chart

Gender	Number = Absolute frequency	Percentage (%) = Relative frequency
F	7	35
M	13	65
Total	20	100

Gender



Title of the chart

Percentages for all categories

Categories titles

Figure. Gender percentages distribution

Gender

M  
M  
F  
F  
F  
F  
M  
F  
M  
M  
F  
F  
F  
F  
F  
M  
F  
M  
F  
F

# Chart – Column chart

Gender	Number = Absolute frequency	Percentage (%) = Relative frequency
F	7	35
M	13	65
Total	20	100

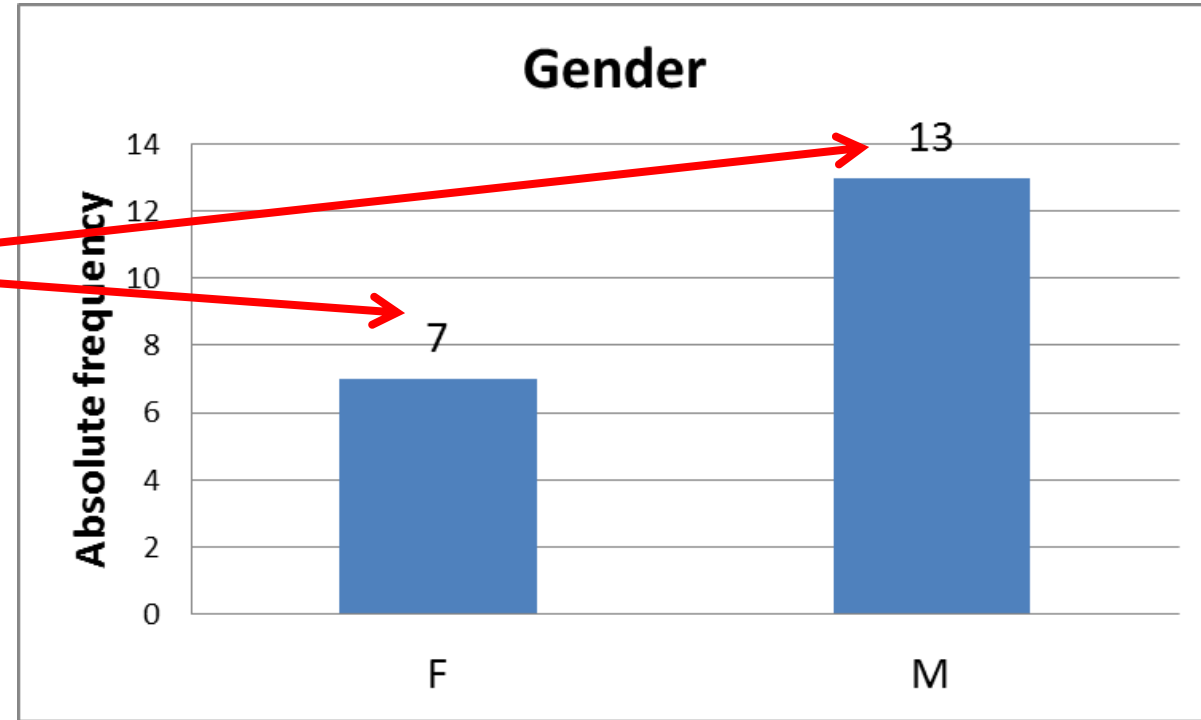


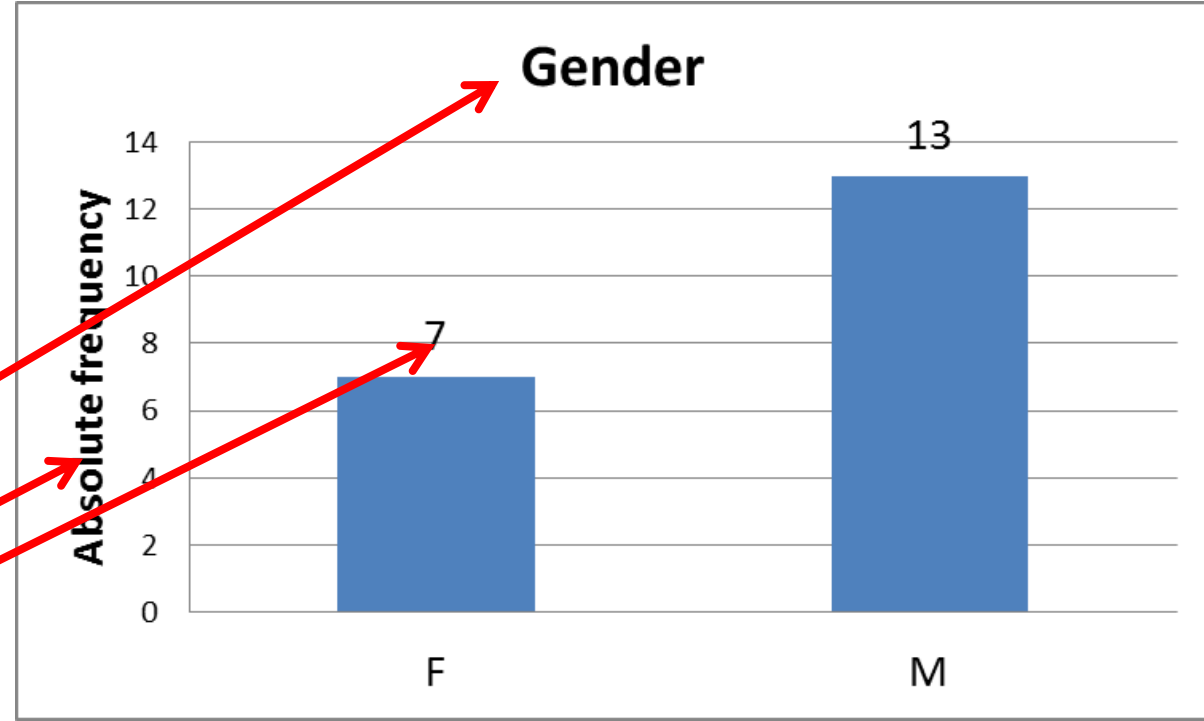
Figure. Gender distribution

Gender

M  
M  
F  
F  
F  
F  
M  
F  
M  
M  
F  
F  
F  
F  
F  
M  
F  
M  
F  
F

# Chart – Column chart

Gender	Number = Absolute frequency	Percentage (%) = Relative frequency
F	7	35
M	13	65
Total	20	100



Title of the chart

Axis title

Values

Figure. Gender distribution

Dentition  
mixt  
permanent  
permanent  
permanent  
mixt  
temporary  
permanent  
permanent  
mixt  
permanent  
mixt  
permanent  
temporary  
temporary  
mixt  
mixt  
temporary  
permanent  
permanent  
permanent

# Frequency table: Nominal variable

How many people?

Dentition	Absolute frequency	Relative frequency
temporary	4	20
permanent	10	50
mixt	6	30
Total	20	100

Percentages

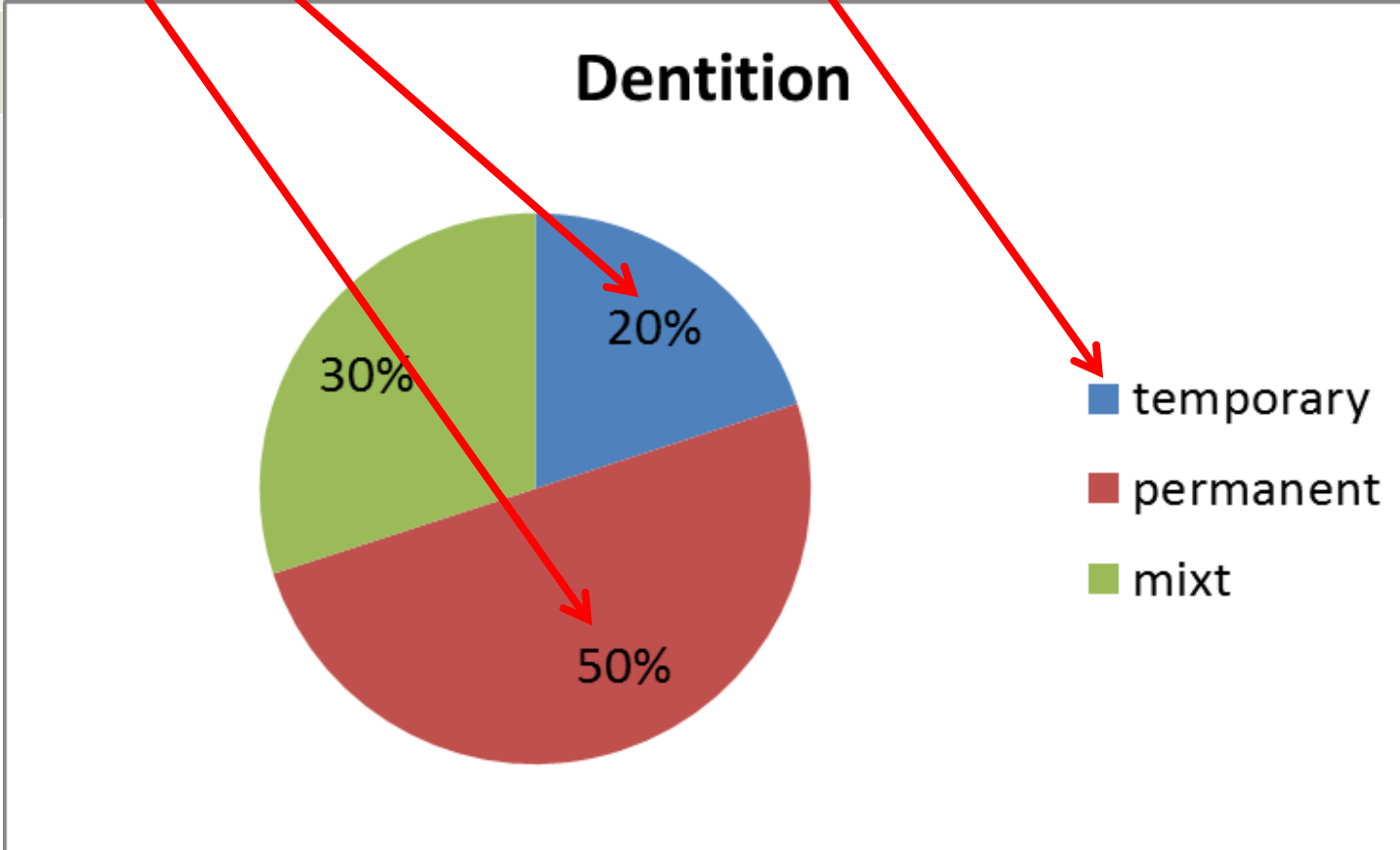
Dentition – **3 possible categories**: temporary, permanent, mixt

**Dentition**  
 mixt  
 permanent  
 permanent  
 permanent  
 mixt  
 temporary  
 permanent  
 permanent  
 mixt  
 permanent  
 mixt  
 permanent  
 temporary  
 temporary  
 mixt  
 mixt  
 temporary  
 permanent  
 permanent  
 permanent

Dentition	Absolute frequency	Relative frequency
temporary	4	20
permanent	10	50
mixt	6	
Total	20	

# Chart – Pie chart

Legend



**Dentition**  
 mixt  
 permanent  
 permanent  
 permanent  
 mixt  
 temporary  
 permanent  
 permanent  
 mixt  
 permanent  
 mixt  
 permanent  
 temporary  
 temporary  
 mixt  
 mixt  
 temporary  
 permanent  
 permanent  
 permanent

Dentition	Absolute frequency	Relative frequency
temporary	4	20
permanent	10	50
mixt	6	30
Total	20	100

# Chart – Bar chart

- ✓ Title of the chart
- ✓ Axis title
- ✓ Values



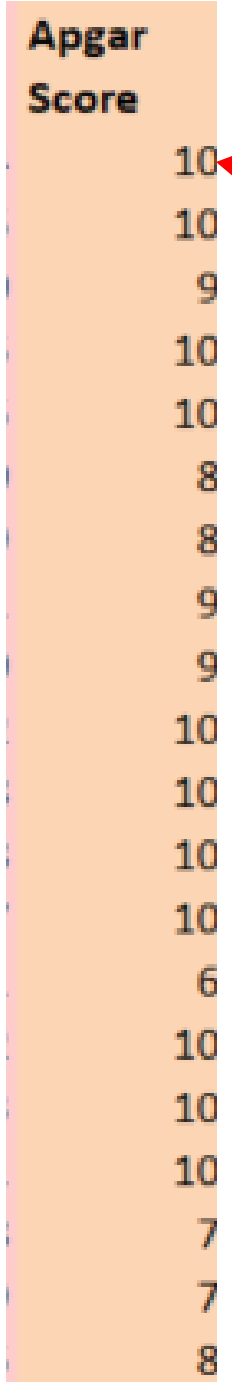
# Frequency table: ordinal variable

Apgar score	Absolute frequency	Relative frequency (%)
10	11	55
9	3	15
8	3	15
7	2	10
6	1	5
Total	20	100

Apgar score have 5 possible values in the table

How many 10 are in the Apgar score column?

How many 6?



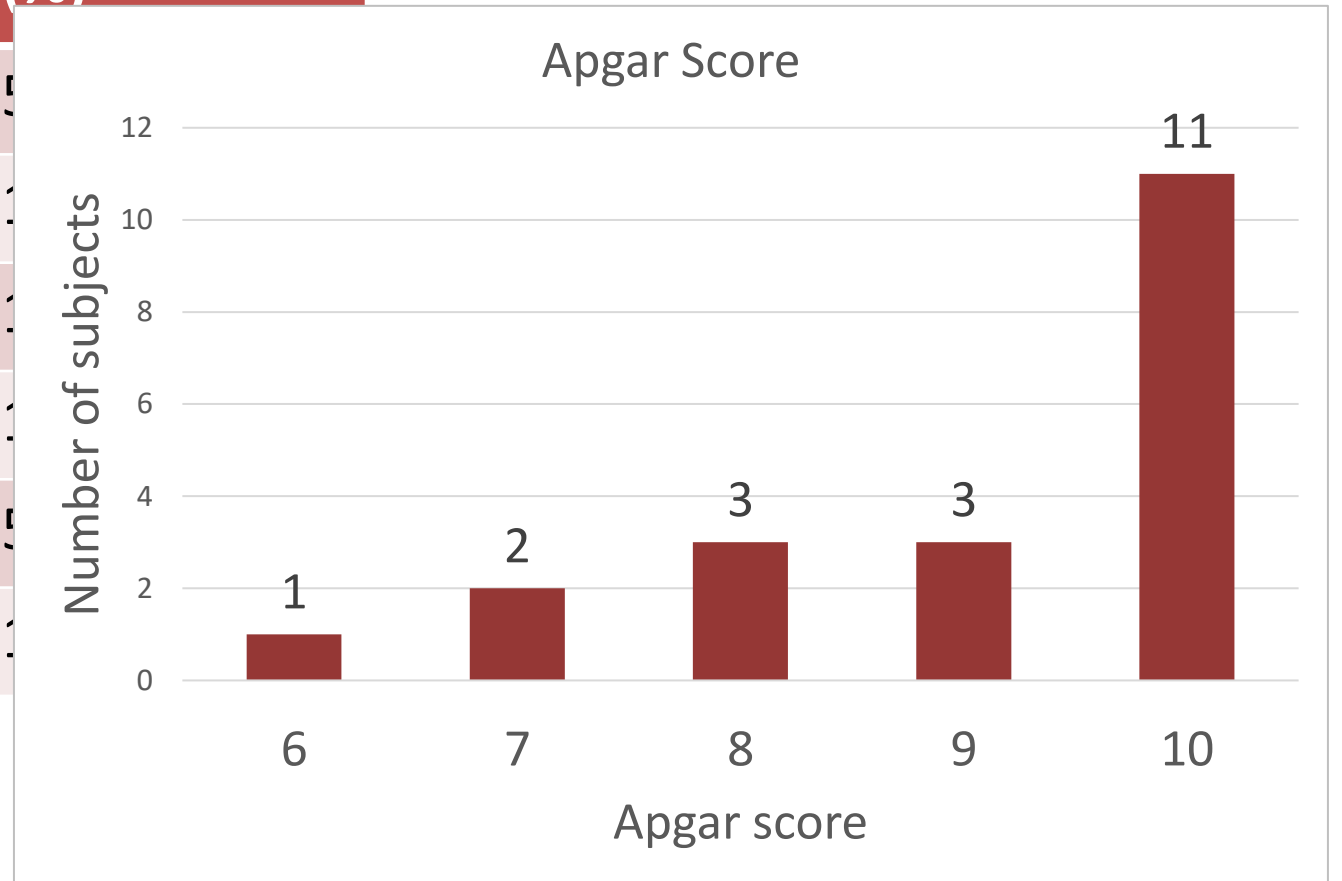
# Chart – Column/Bar Chart

✓ Title of the chart

✓ Axis title

✓ Values

Apgar score	Absolute frequency	Relative frequency (%)
10	11	55
9	3	15
8	3	15
7	2	10
6	1	5
Total	20	100



# Summarizing nominal & ordinal data

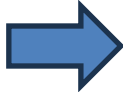
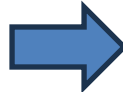
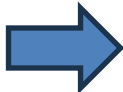
**Absolute frequency of category “A”** =  $a$  = number of apparition of category “A”

**Relative frequency in percentages of category “A”** =  $a/n * 100$   
(where  $n$  is the total number of subjects)

**Relative frequency of category “A”** =  $a/n$

**Table of frequencies** – a structured table that displays frequencies for all categories of a qualitative variable

# Summary – charts type

Dichotomial		<b>Pie chart or Column chart or Bar chart</b>
Nominal		<b>Pie chart or Column chart or Bar chart</b>
Ordinal		<b>Column chart or Bar chart</b>



# Calculate frequencies in Excel

with CountIf function

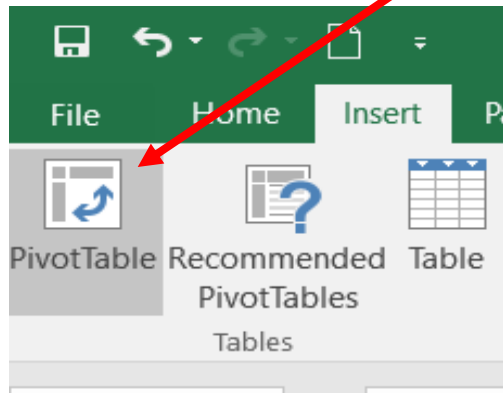
=CountIf(range, criteria)

The image shows an Excel spreadsheet with a pivot table. The pivot table is located in columns H through M. The data source is in column H, with rows 47 through 56. The pivot table has two columns: 'Symptoms at discharge' and 'Absolute frequency'. The 'Absolute frequency' column contains the following formulas: =CountIf(H2:H151,"Present"), =COUNTIF(range, criteria), and =SUM(...). A blue arrow points from the text '=CountIf(range, criteria)' to the formula in the 'Absolute frequency' column.

	H	I	J	K	L	M
	Symptoms at discharge (present/absent)			Symptoms at discharge	Absolute frequency	
47	Present			Present	=CountIf(H2:H151,"Present")	
31	Absent			Absent	=COUNTIF(range, criteria)	
56	Absent			Total	=SUM(...)	
38	Present					
32	Present					
30	Present					

# Calculate frequencies in Excel

with Pivot Table



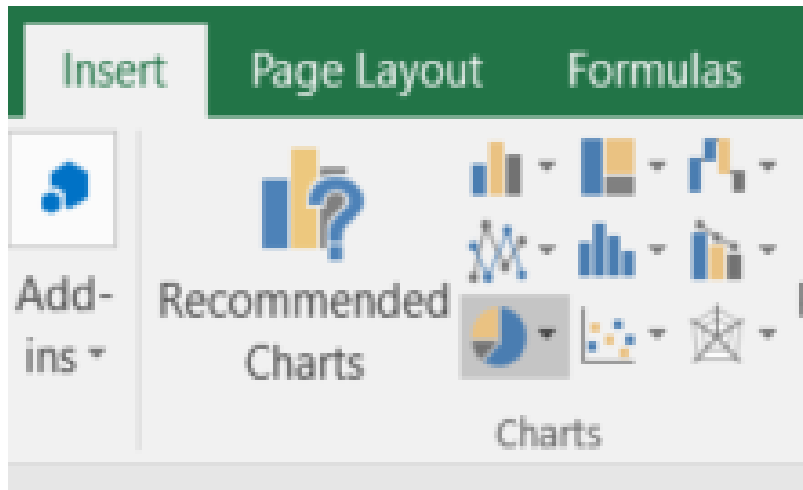
The screenshot shows the Excel interface with the 'Create PivotTable' dialog box open. The dialog box has three main sections: 'Choose the data that you want to analyze', 'Choose where you want the PivotTable report to be placed', and 'Choose whether you want to analyze multiple tables'. The 'Table/Range' field is set to 'Sheet1!\$A\$1:\$H\$151'. The 'New Worksheet' radio button is selected. The 'PivotTable Fields' task pane is also visible on the right, showing the 'Symptoms at discharge (present/absent)' field selected and placed in the 'Rows' and 'Σ Values' areas. A red arrow points from the 'Table/Range' field in the dialog box to the 'Symptoms at discharge (present/absent)' field in the task pane.

Row Labels	Count of Symptoms at discharge (present/absent)
Absent	98
Present	52
<b>Grand Total</b>	<b>150</b>

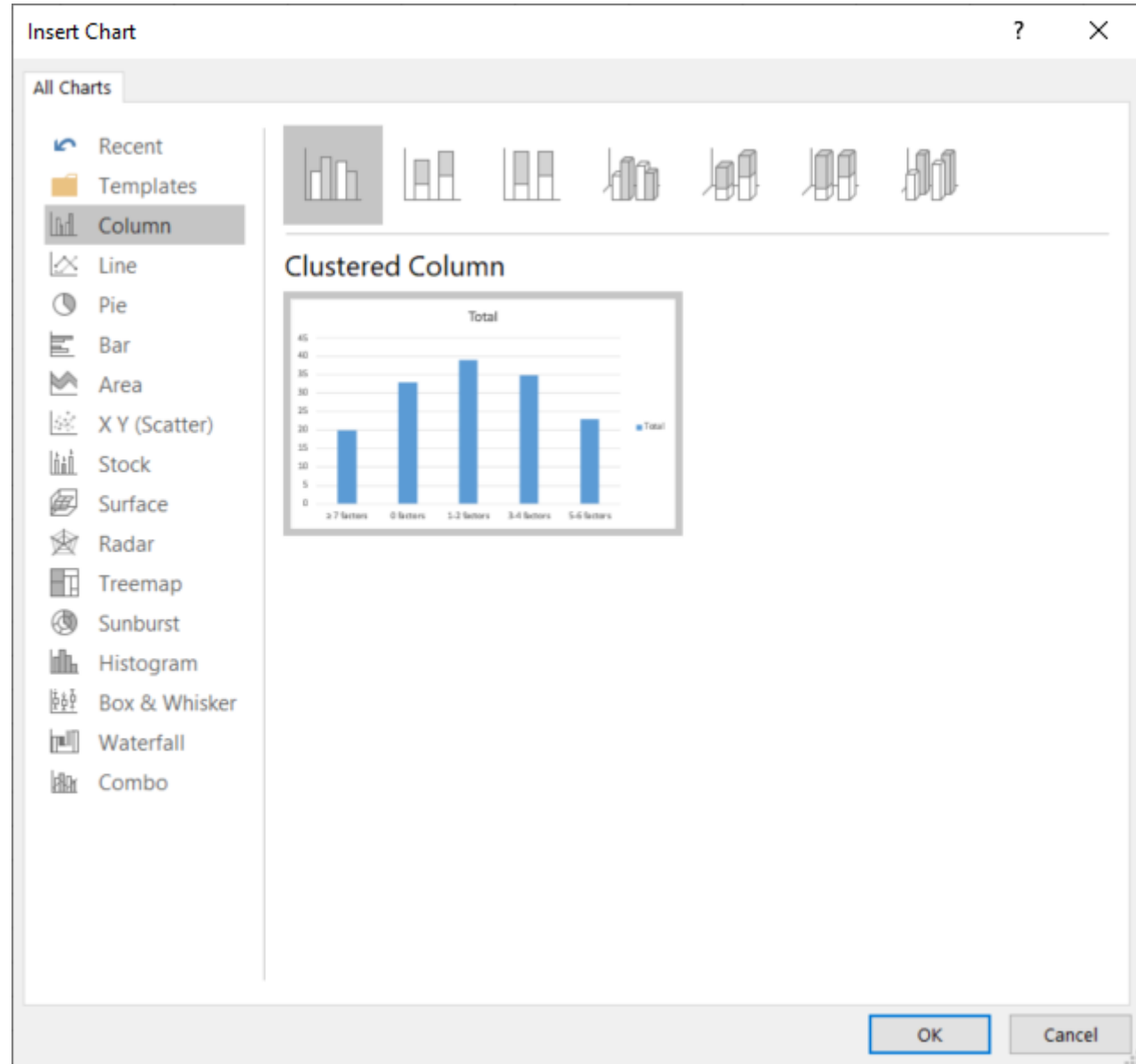
Select the data table

Drag and drop the variable on Rows and on  $\Sigma$  Values

# Insert a Chart in Excel



!To make a chart in Excel you should make a frequency table first!



# Without frequency table in Excel

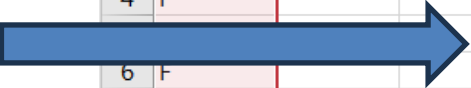
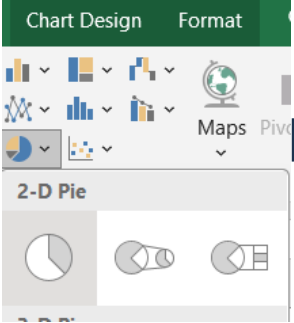
	A
1	Gender
2	M
3	M
4	F
5	F
6	F
7	F
8	M
9	F
10	M
11	M
12	F
13	F
14	F
15	F
16	F
17	M
18	F
19	M
20	F
21	F
22	

Select data



	A
1	Gender
2	M
3	M
4	F
5	F
6	F
7	F
8	M
9	F
10	M
11	M
12	F
13	F
14	F
15	F
16	F
17	M
18	F
19	M
20	F
21	F
22	

Insert Pie chart



	A	B	C	D	E	F	G	H	I	J
1	Gender									
2	M									
3	M									
4	F									
5										
6	F									
7	F									
8	M									
9	F									
10	M									
11	M									
12	F									
13	F									
14	F									
15	F									
16	F									
17	M									
18	F									
19	M									
20	F									
21	F									
22										

Gender M M F F F M F M M F F F M F M F

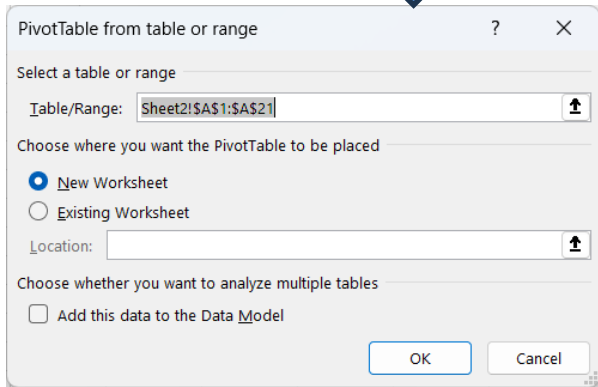
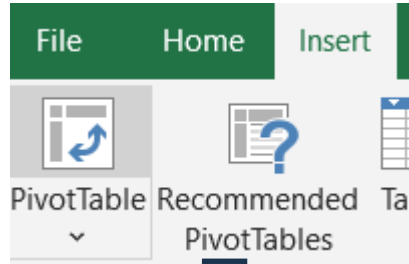
**Incorrect**

# With frequency table in Excel

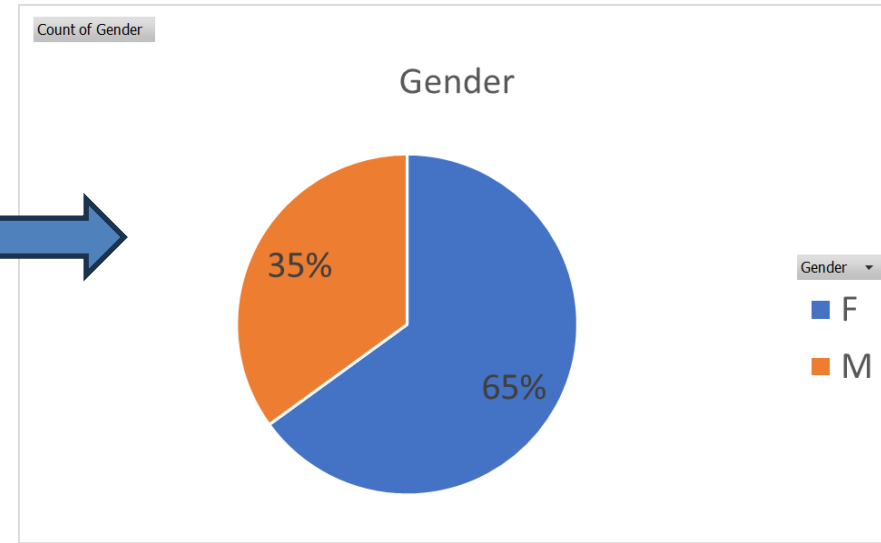
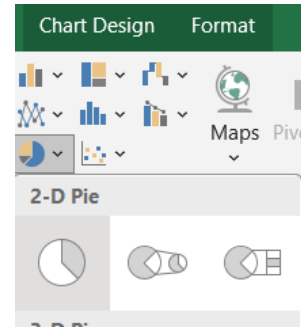
Select the title

	A
1	Gender
2	M
3	M
4	F
5	F
6	F
7	F
8	M
9	F
10	M
11	M
12	F
13	F
14	F
15	F
16	F
17	M
18	F
19	M
20	F
21	F

Insert Pivot table



Insert a Pie chart



Make the frequency table

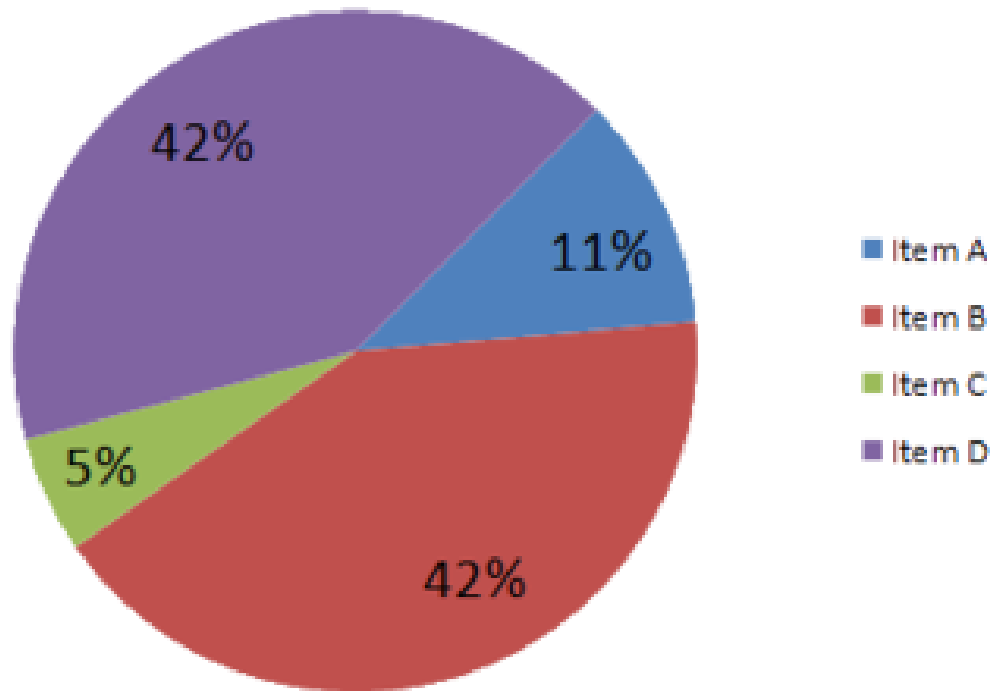
	Row Labels	Count of Gender
4	F	13
5	M	7
6	<b>Grand Total</b>	<b>20</b>

Correct

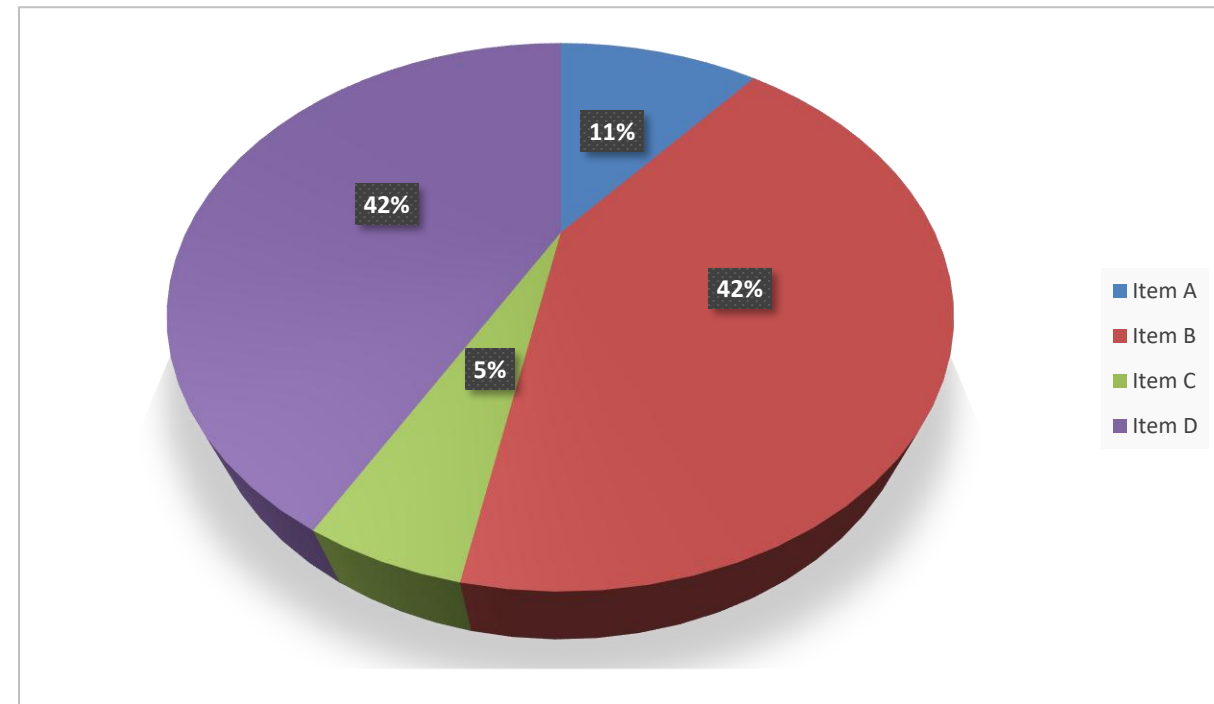
Obs. Frequency table can be made also with Countif function

# Not recommended 3D charts

5% near the same as 11%



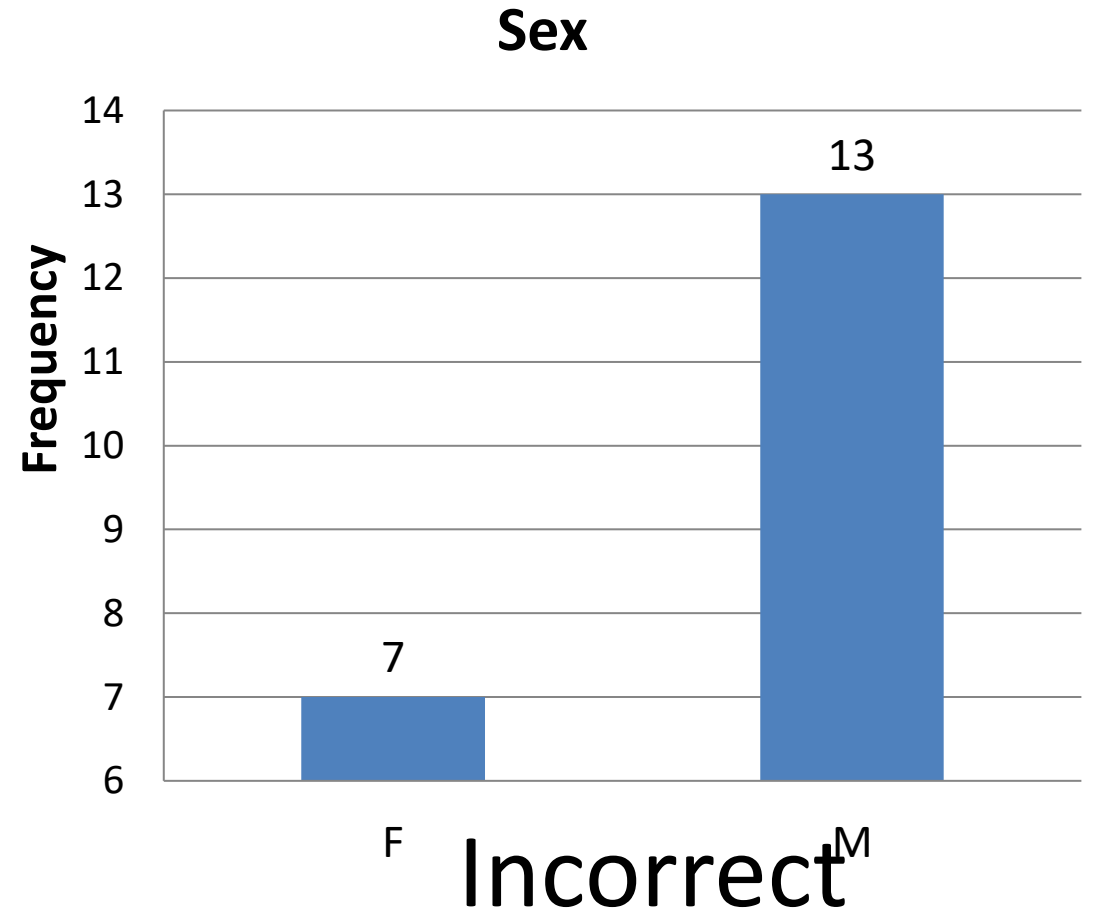
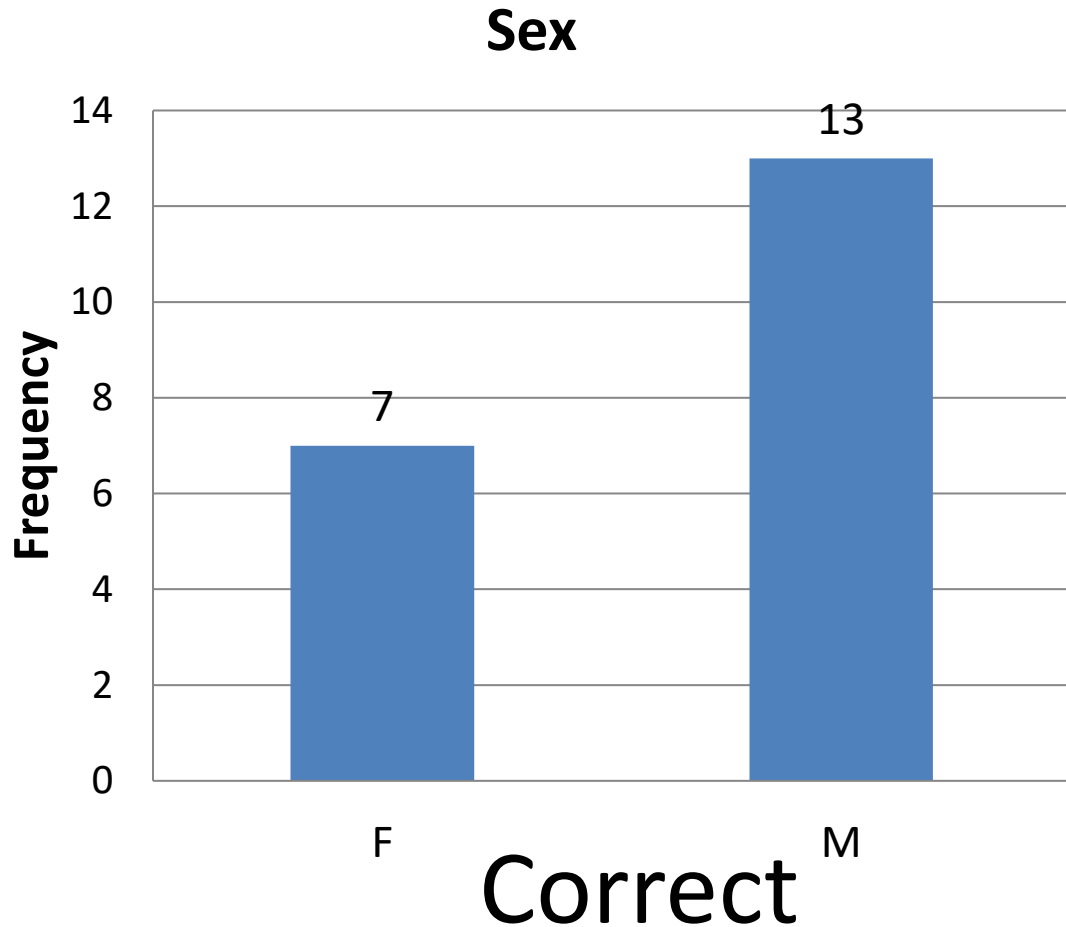
Correct



Incorrect

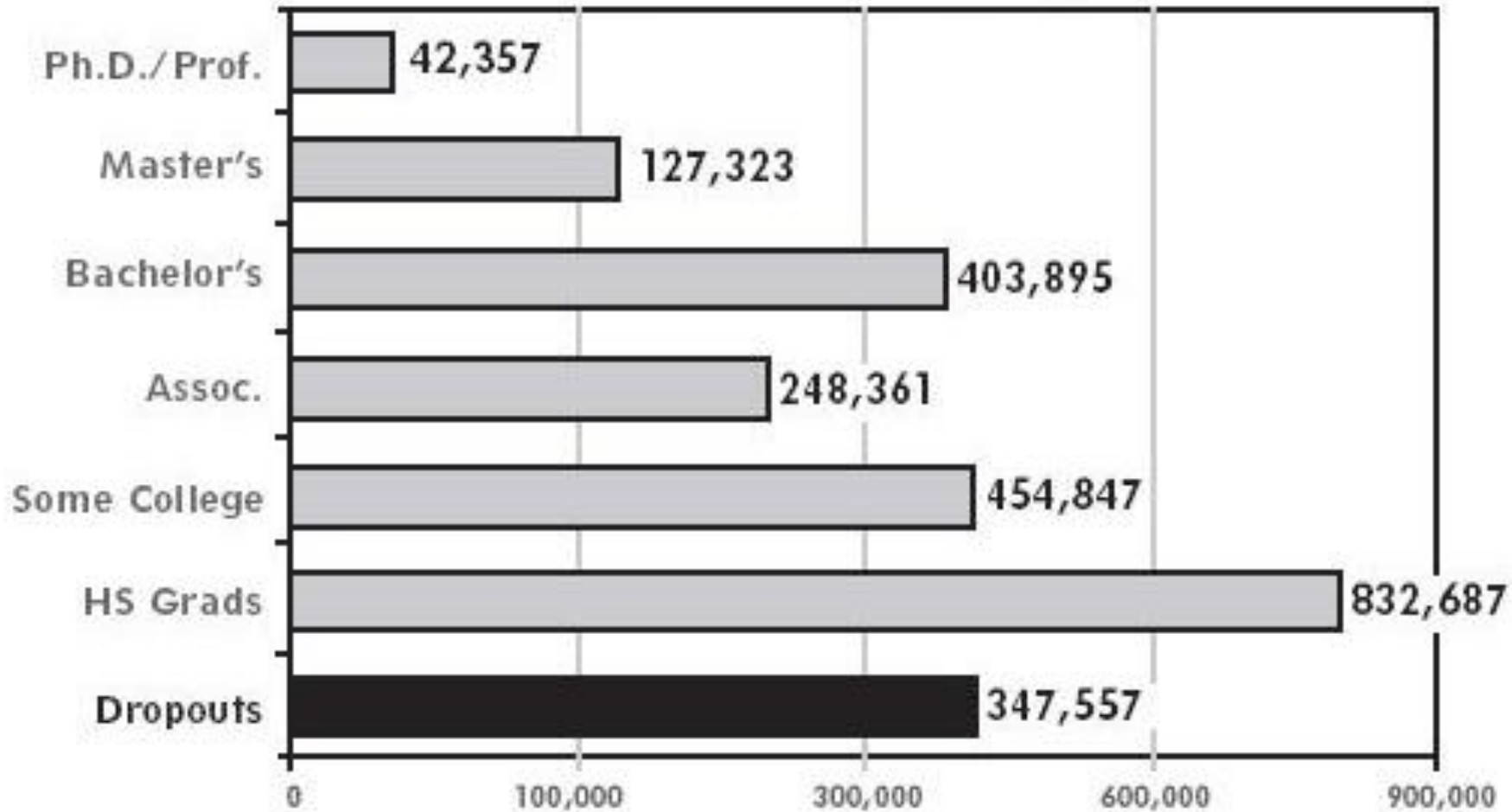
# Not recommended

to start the axis at another value than 0 (increasing the differences between the frequencies)



# Incorrect

South Carolina Residents Age 20-64



Source: PolEcon analysis of U.S. Census Bureau's Current Population Survey March 2004-06 Supplement data for South Carolina

Thank you!