

Indicații tehnice de realizare a laboratoarelor

CUPRINS

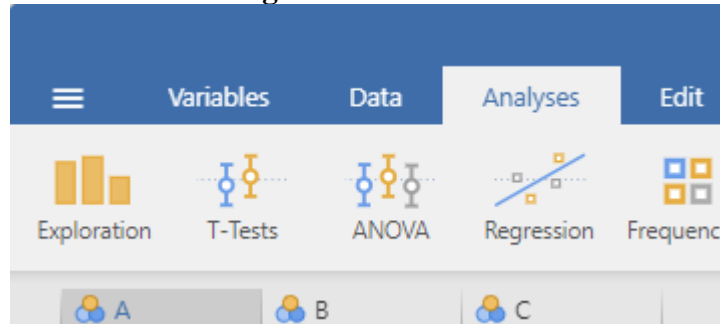
INSTRUCȚIUNI UTILIZARE PROGRAM JAMOVI	3
IMPORTAREA BAZEI DE DATE.....	3
VERIFICAREA TIPULUI DATELOR.....	3
REALIZARE ANALIZE ÎNTRE 2 VARIABILE CALITATIVE – TEST HI PĂTRAT, FISHER, TABEL CONTINGENTA, OR, RR, RA, GRAFIC COLOANE.....	4
<i>Tabel de contingență.....</i>	7
<i>Rezultatele testelor Hi pătrat și Fisher exact.....</i>	7
<i>Indicatori medicali ai forței asocierii</i>	8
<i>Grafic de tip coloane</i>	8
REALIZARE TEST STUDENT PENTRU DOUĂ GRUPURI INDEPENDENTE	9
<i>Evaluare condiții de aplicare test Student – normalitatea.....</i>	10
<i>Alegerea între testul Student cu varianțe egale și testul Welch – cu varianțe inegale.....</i>	11
<i>Statistica descriptivă.....</i>	12
Numeric	12
Grafic de medii	12
ANALIZA ROC	13
<i>Tabele cu valori limită și statistici asociate.....</i>	15
<i>Grafic de tip curbă ROC.....</i>	15
<i>Comparare curbe ROC prin teste statistice</i>	16
ANALIZA SUPRAVIEȚUIRII (GRAFICUL KAPLAN MEYER, TESTUL LOG-RANK, REGRESIA COX – HR).....	18
<i>Mediana timpului de supraviețuire</i>	19
<i>Probabilitatea de supraviețuire în momente diferite</i>	20
<i>Curba de supraviețuire Kaplan-Meier, cu testul log-rank.....</i>	20
META-ANALIZA	21
<i>Meta-analiza pentru compararea a două grupuri în ce privește date de tip dihotomial</i>	21
Estimatorul asocierii	23
Heterogenitatea statistică a studiilor.....	24
Graficul pădure (forest plot).....	24
Eroarea sistematică de publicare	25
Statistici	25
Grafic de tip pâlnie (funnel plot).....	26
INSTALARE MODUL DE ANALIZE SUPLIMENTARE (EX. PENTRU ANALIZA ROC).....	27
INSTRUCȚIUNI UTILIZARE EXCEL.....	28
DESCRIEREA DATELOR	28
<i>variabile calitative:</i>	28
Tabele de frecvență (două metode)	28
Grafice de tip sectorial (pie) (două metode).....	29
Tabele de contingență	30
Graficul asociat tabelului de contingență	31
<i>variabile cantitative:</i>	31
descrierea unei variabile.....	31
Media, mediana, deviația standard și intervalul de încredere 95% pentru medie.	31
Grafic boxplot.....	31
Tabele de frecvență.....	32
Histograme.....	33
descrierea relației între două variabile.....	34
diagramă nor de puncte (Scatter)	34
<i>variabile de supraviețuire:</i>	35
ACTIVAREA MODULULUI DATA ANALYSIS DIN MENIUL TOOLS ÎN EXCEL	35

ANALIZA DATELOR.....	36
<i>Realizarea testului Student (test t) în Excel</i>	36
<i>Realizarea regresiei lineare în Excel</i>	37
INSTRUCȚIUNI UTILIZARE EPI INFO 7.....	37
DESCHIDEREA PROGRAMULUI EPI INFO 7	37
DESCHIDEREA COMPONENTEI DE ANALIZĂ STATISTICĂ	37
IMPORTAREA UNUI FIȘIER EXCEL ÎN EPIINFO7.....	39
SELECTAREA FOLDERULUI UNDE VOR FI SALVATE REZULTATELE ANALIZEI STATISTICE	42
AFIȘAREA DATELOR IMPORTATE	43
REALIZAREA TABELELOR DE FRECVENȚĂ – CU INTERVALE DE ÎNCREDERE.....	44
TESTE STATISTICE, INDICATORI, REGRESII	45
<i>Realizarea testului X^2 (chi pătrat) în Epi Info și aflarea tabelului de contingență, a Riscului relativ(RR), ratei șansei(OR), diferenței de risc(RD)</i>	45
<i>Compararea datelor cantitative (Test student/ANOVA/ ...) în Epi Info</i>	45
<i>Realizarea testului log-rank pentru compararea supraviețuirii a două grupuri, în EpiInfo</i>	47
<i>Aflarea ratei hazardului prin efectuarea regresiei Cox (identificarea predictorilor în analiza datelor de supraviețuire), în EpiInfo</i>	48
REALIZARE GRAFICE ÎN EPIINFO 7.....	49
<i>Grafice pentru descrierea unei variabile calitative (Pie, Bar, Column)</i>	49
<i>Grafic pentru a arăta relația între două variabile calitative (Bar/Column)</i>	49
<i>Salvarea graficelor</i>	51
<i>Editarea graficelor (titlu, axe)</i>	52

Instrucțiuni utilizare program Jamovi

Importarea bazei de date

Apăsați pe icoana de meniu **hamburger**. 



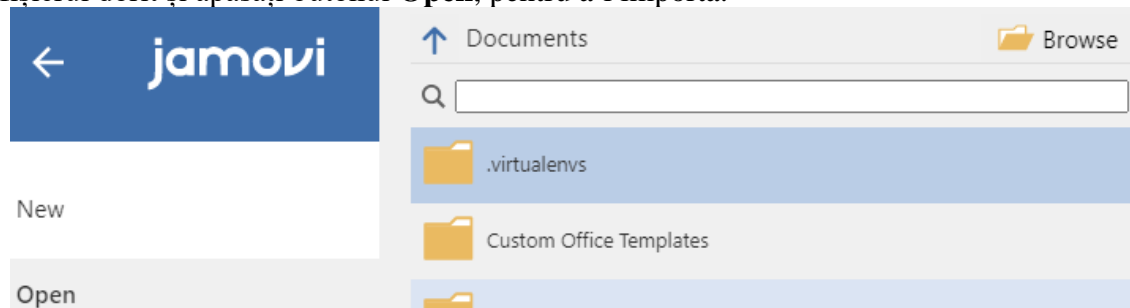
Apoi selectați opțiunea **Open** (deschidere)



New

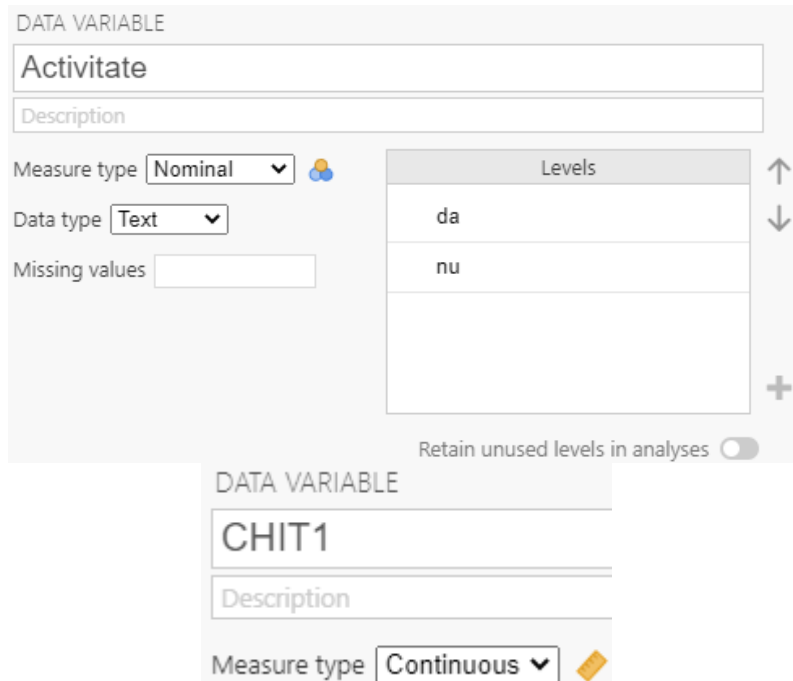
Open

Ulterior căutați dosarul cu baza de date apăsând pe butonul **Browse**, selectați fișierul dorit și apăsați butonul **Open**, pentru a-l importa.



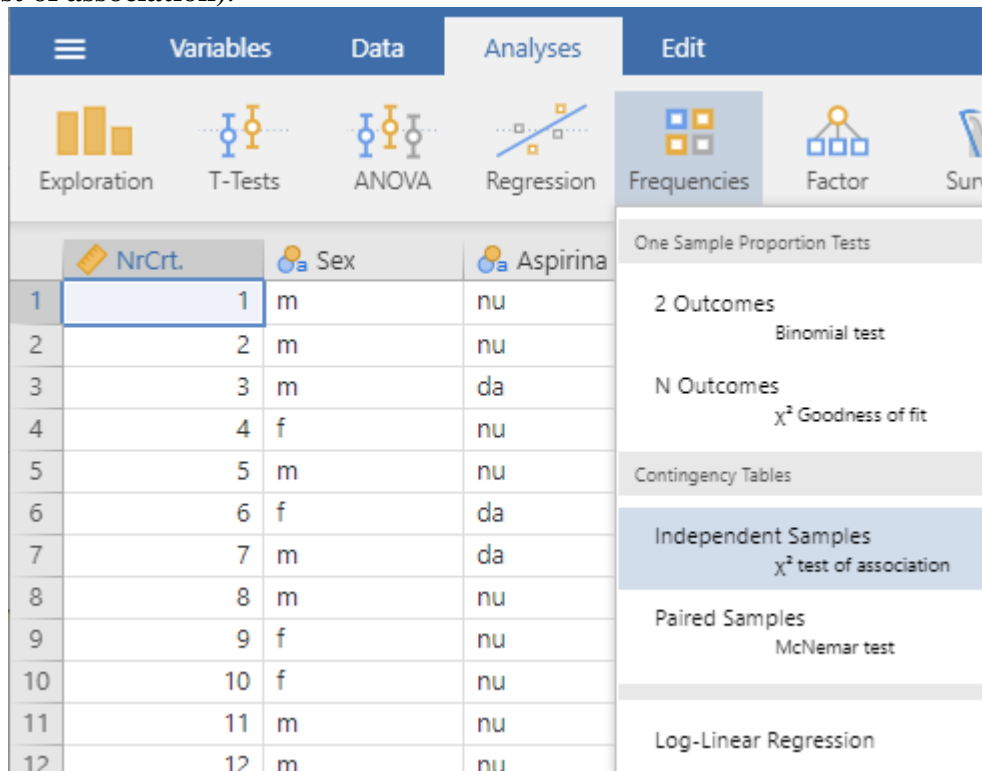
Verificarea tipului datelor

Înainte de începerea analizei, este necesar să verificăm dacă programul a clasificat corect variabilele din baza de date. Variabila Activitate este o variabilă calitativă, iar celelalte 4 variabile sunt cantitative. Pentru verificarea tipului unei variabile dăm **dublu clic** pe **titlul coloanei** (numele variabilei). În cazul variabilei Activitate ne asigurăm că în zona de selecție **Measure type** este selectată opțiunea **Nominal**. Pentru celelalte variabile cantitative ne asigurăm că în zona de selecție Measure type este selectată opțiunea **Continuous**.



Realizare analize între 2 variabile calitative – Test Hi pătrat, Fisher, tabel contingenta, OR, RR, RA, grafic coloane

În tabul Analyses, modulul Frequencies se alege opțiunea Independent Samples (χ^2 test of association).



Variabila de grupare (ex. factorul de risc, sau tratamentul), se deplasează apăsând pe săgeata din dreptul câmpului Rows.

Variabila care este de interes (ex. boala, sau complicația, sau rezultatul tratamentului), se deplasează apăsând săgeata din dreptul câmpului **Columns**.

Ca opțiuni recomandăm alegerile următoare: Tests: **X²**, **Fisher's exact test**; Comparative measures (2x2 only) (modificați în funcție de tipul de culegere al datelor în funcție de grupare): rația șanselor (**Odds ratio**), riscul relativ (**Relative risk**), riscul atribuabil/reducerea absolută de risc (**Difference in proportion**), intervale de încredere (**Confidence intervals**); frecvențe (Counts): frecvențe observate (**Observed counts**), frecvențe așteptate/teoretice (**Expected counts**); procente (Percentages): pe linii (**Row**); Grafice (Plots): coloană (**Bar Plot**), pe axa Y (**Y-Axis**): procente (**Percentages**), pe linii (**within rows**); tip de coloane (**Bar Type**), suprapuse (Stacked); pe axa X (**X-Axis**): linii (**Rows**).

Contingency Tables



- NrCr.
- Sex
- D1Ulceratie (mm)**
- D2Ulceratie (mm)
- D1-D2 (mm)

Rows: Aspirina

Columns: Vindecare

Counts (optional):

Layers:

Statistics

Tests

- χ^2
- χ^2 continuity correction
- Likelihood ratio
- Fisher's exact test
- z test for difference in 2 proportions

Hypothesis

- Group 1 \neq Group 2
- Group 1 $>$ Group 2
- Group 1 $<$ Group 2

Comparative Measures (2x2 only)

- Odds ratio
- Log odds ratio
- Relative risk
- Difference in proportions
- Confidence intervals

Interval 95 %

Compare rows

▼ | Cells

Counts

Observed counts

Expected counts

Percentages

Row

Column

Total

▼ | Plots

Plots

Bar Plot

Y-Axis

Counts

Percentages within rows ▼

Bar Type

Side by side

Stacked

X-Axis

Rows

Columns

Tabel de contingență

Contingency Tables

Aspirina		Vindecare		Total
		da	nu	
da	Observed	182	18	200
	Expected	162	38.0	200
	% within row	91.0 %	9.0 %	100.0 %
nu	Observed	142	58	200
	Expected	162	38.0	200
	% within row	71.0 %	29.0 %	100.0 %
Total	Observed	324	76	400
	Expected	324	76.0	400
	% within row	81.0 %	19.0 %	100.0 %

Rezultatele testelor Hi pătrat și Fisher exact

Valoarea lui p, corespunzătoare testelor se află în coloana p.

χ^2 Tests			
	Value	df	p
χ^2	26.0	1	< .001
Fisher's exact test			< .001
N	400		

Indicatori medicali ai forței asocierii

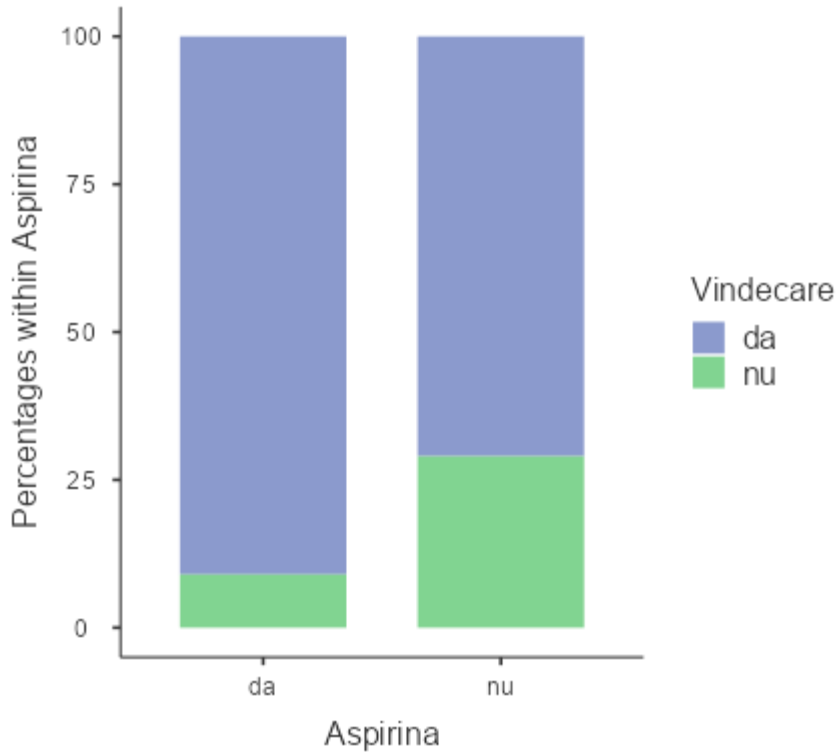
Rezultatele privind indicatorii medicali sunt prezentate în tabelul Comparative Measures, unde estimatorul punctual se află în coloana valori (**Value**), iar capetele intervalului de încredere 95% se află în coloanele **Lower, Upper (95% Confidence Intervals)**.

	Value	95% Confidence Intervals	
		Lower	Upper
Difference in 2 proportions	0.200*	0.126	0.274
Odds ratio	4.13	2.33	7.32
Relative risk	1.28*	1.16	1.41

* Rows compared

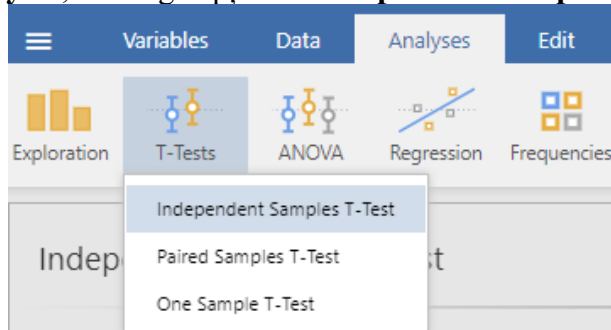
Grafic de tip coloane

La final avem reprezentarea grafică a asocierii dintre cele două variabile calitative.



Realizare test Student pentru două grupuri independente

În tabul **Analyses**, se alege opțiunea **Independent Samples T-Test**.



Variabila calitativă care **identifică grupurile** comparate se deplasează apăsând săgeata din dreptul câmpului **Grouping Variable**.

Variabila/ele cantitative care sunt de interes se deplasează apăsând săgeata din dreptul câmpului **Dependent Variable**.

Se bifează următoarele opțiuni:

- în secțiunea **Tests: Student's, Welch's**
- în secțiunea **Additional statistics: Mean difference, Confidence interval, Descriptives, Descriptive plots.**
- În secțiunea **Assumptions checks: Homogeneity test, Normality test, Q-Q plot.**

Independent Samples T-Test →

Variables

- NrCrt.
- Sex
- Vindecare
- D1Ulceratie (mm)
- D2Ulceratie (mm)

Dependent Variables

- D1-D2 (mm)

Grouping Variable

→ Aspirina 📊

Tests

Student's

Bayes factor

Prior

Welch's

Mann-Whitney U

Hypothesis

Group 1 ≠ Group 2

Group 1 > Group 2

Group 1 < Group 2

Missing values

Exclude cases analysis by analysis

Exclude cases listwise

Additional Statistics

Mean difference

Confidence interval %

Effect size

Confidence interval %

Descriptives

Descriptives plots

Assumption Checks

Homogeneity test

Normality test

Q-Q plot

Evaluare condiții de aplicare test Student – normalitatea

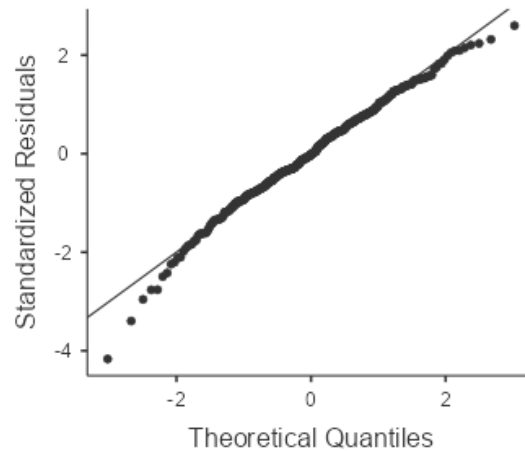
Testul Student poate fi aplicat doar **dacă datele** urmează o **distribuție normală**. O **variantă** de evaluare a normalității datelor este cu ajutorul **testelor statistice de evaluare a normalității**. Dacă eșantionul are **sub 50 de observații** ne uităm la rezultatul testului **Shapiro-Wilk**, dacă are **peste 50 de observații** ne uităm la rezultatul testului **Kolmogorov-Smirnov**. Rezultatul de interes este **valoarea p**, care se află în coloana p din tabelul **Tests of Normality**. Dacă valoarea **p** este **mai mica decât 0,05**, este o **sugestie** că datele **nu urmează o distribuție normală**, **altfel** este o **sugestie** că datele urmează o **distribuție normală**.

Tests of Normality

		statistic	p
D1-D2 (mm)	Shapiro-Wilk	0.990	0.009
	Kolmogorov-Smirnov	0.0298	0.870
	Anderson-Darling	0.502	0.206

Note. Additional results provided by moretests

O altă variantă de evaluare a normalității datelor este cu ajutorul **graficului Q-Q plot**. Dacă **tendința** observațiilor este ca acestea să fie dispuse **departe de linia continuă** (care reprezintă o distribuție normală), atunci este o **sugestie** că datele **nu urmează o distribuție normală, altfel**, dacă tendința e ca observațiile să fie apropiate de linia continuă, este o **sugestie** că datele urmează o **distribuție normală**.



Alegerea între testul Student cu varianțe egale și testul Welch – cu varianțe inegale

În cazul **testului Student pentru grupuri independente**, există **două variante**, pentru situația în care se presupun **varianțe egale**, sau **varianțe inegale**. Pentru a alege între cele două se pot folosi **teste pentru compararea varianțelor**. În tabelul **Homogeneity of Variances Tests**, privim **valoarea p** obținută cu **testul Levene**. Dacă valoarea **p** este **mai mica decât 0,05**, este o **sugestie** că grupurile **au varianțe inegale, altfel** este o **sugestie** că grupurile **au varianțe egale**.

Homogeneity of Variances Tests

		F	df	df2	p
D1-D2 (mm)	Levene's	1.35	1	398	0.246
	Variance ratio	0.785	199	199	0.088

Note. Additional results provided by moretests

Dacă **varianțele** sunt presupuse a fi **egale**, alegem rezultatele corespunzând pentru **testul Student**, altfel le alegem pe cele corespunzând pentru **testul Welch** din tabelul **Independent Samples T-Test**. În coloana **P** avem **valoarea p** a testului, în coloana **Mean difference**, avem **diferența între mediile** celor două grupuri, iar în **Lower, Upper (95% Confidence Interval)**, se află **intervalul de încredere 95%** asociat diferenței dintre mediile grupurilor.

Independent Samples T-Test

Independent Samples T-Test

		Statistic	df	p	Mean difference	SE difference	95% Confidence Interval	
							Lower	Upper
D1-D2 (mm)	Student's t	3.66	398	< .001	1.61	0.439	0.746	2.47
	Welch's t	3.66	392	< .001	1.61	0.439	0.746	2.47

Note. $H_0: \mu_{da} = \mu_{nu}$

Statistica descriptivă

Numeric

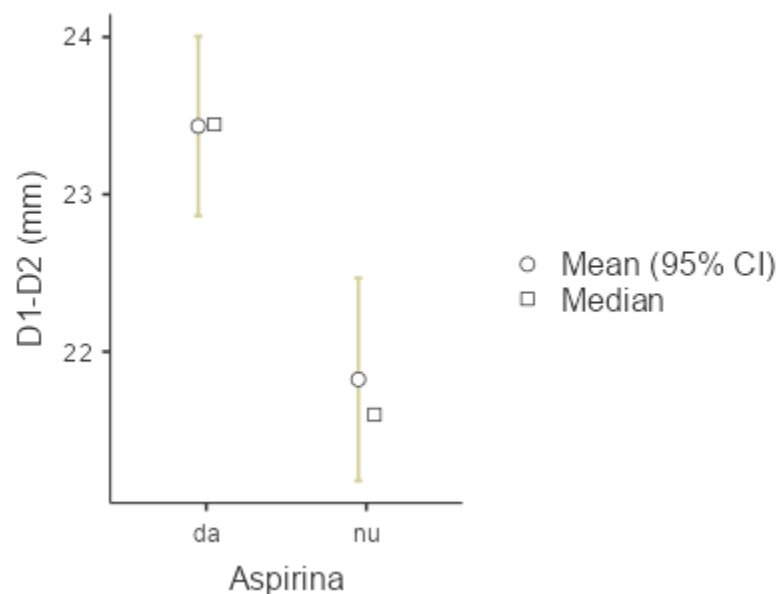
Statisticile descriptive se pot regăsi numeric în tabelul **Group descriptives**. **Numărul de subiecți** se află în coloana **N**, **media** în colana **Mean**, **deviația standard** în coloana **SD**.

Group Descriptives						
	Group	N	Mean	Median	SD	SE
D1-D2 (mm)	da	200	23.4	23.4	4.12	0.291
	nu	200	21.8	21.6	4.65	0.329

Grafic de medii

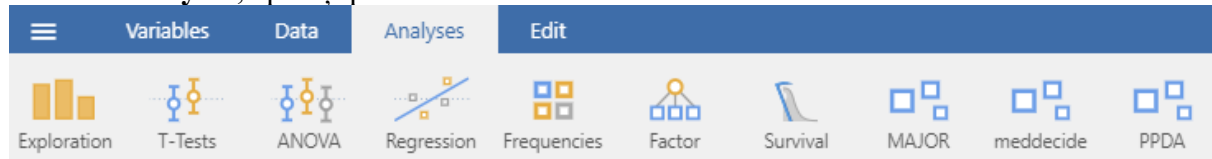
De asemenea este furnizat un **grafic de medii**, în care cercul reprezintă media, barele de eroare reprezintă intervalul de încredere 95% și cu pătrat este reprezentată mediana.

D1-D2 (mm)

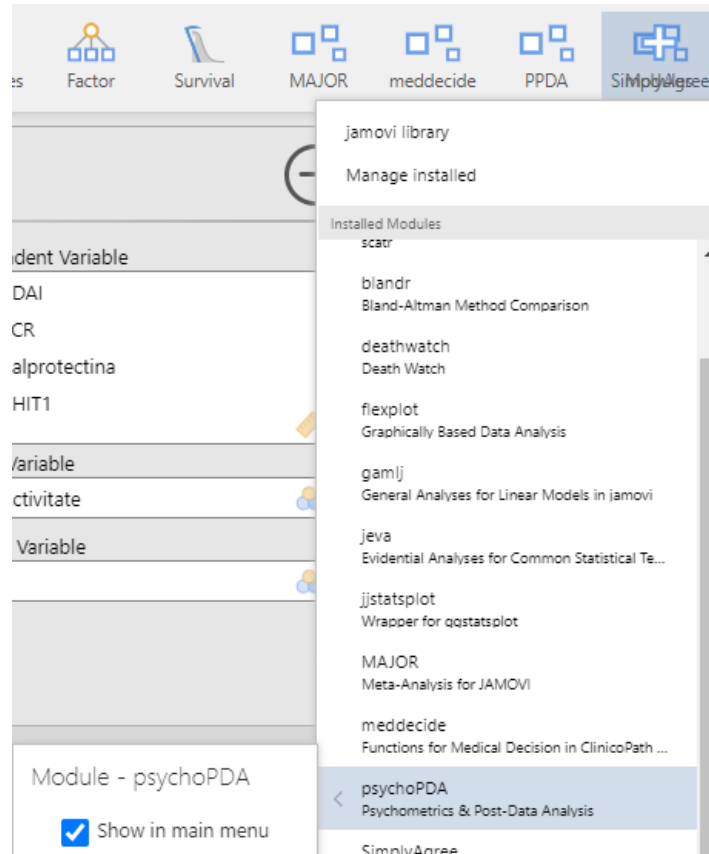


Analiza ROC

În tabul **Analyses**, apăsați pe butonul **PPDA**.



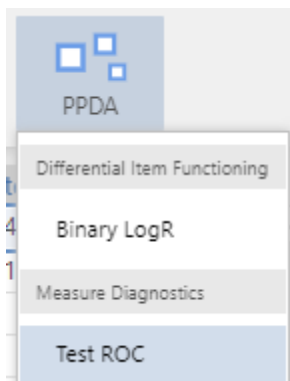
Dacă nu îl găsiți, este posibil să nu fie afișat. În tabul **Analyses**, în dreapta, apăsați pe + **Modules** și verificați dacă nu este prezent în lista de module instalate.



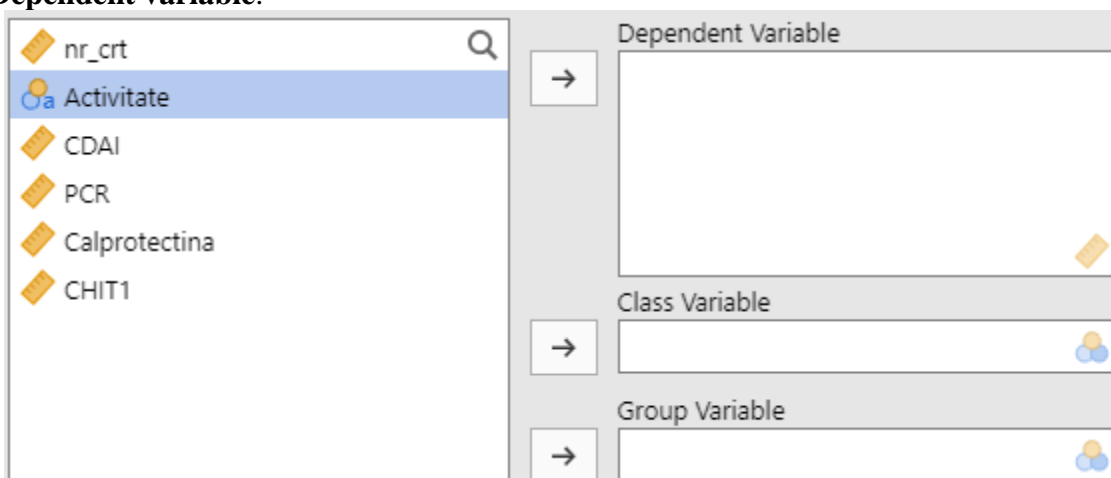
Puteți bifa ca acel modul să fie vizibil în tabul **Analyses**, dând clic pe opțiunea **Show in main menu**.

Dacă acesta nu este instalat urmați pașii indicați în capitolul **Instalare modul de analize suplimentare**.

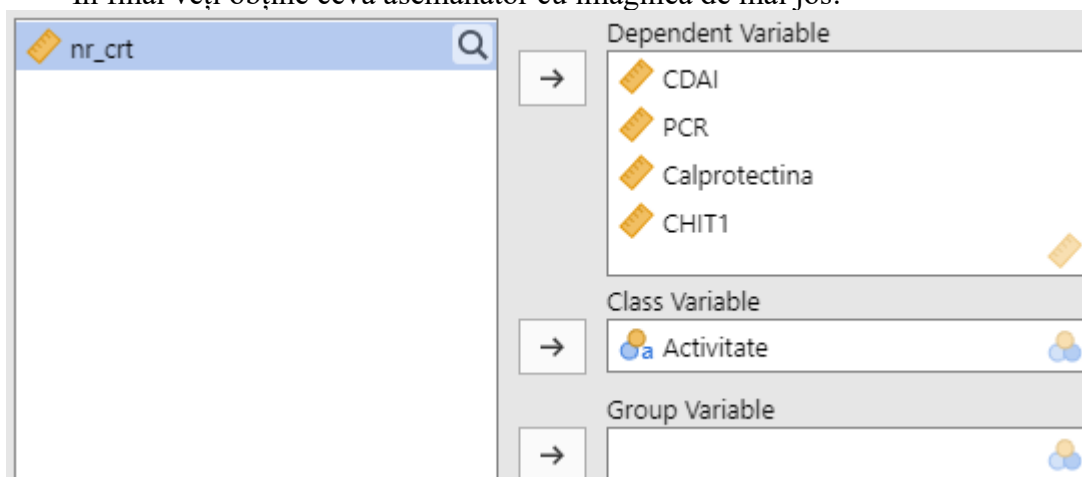
Apăsând pe modulul **PPDA**, selectați din meniu opțiunea **Test ROC**



Selecțai variabila care reprezintă **testul standard** (ex. Activitate), și apăsați butonul **săgeată** din dreptul câmpului **Class variable**. Selecțai **variabilele** cantitative care reprezintă **testele de interes** și apăsați butonul săgeată din dreptul câmpului **Dependent variable**.



În final veți obține ceva asemănător cu imaginea de mai jos:



În câmpul **Positive class**, se specifică pentru testul standard cum se numește categoria de interes (ex: da – pentru prezența activității sau bolii sau stadiu avansat. Avem **grijă** să scriem corect cu **majuscule** și **minuscule**, întrucât programul face diferențe).

CHIT1 (pg/mL)

Class Variable

Group Variable

All observed scores

Positive class

Tabele cu valori limită și statistici asociate

Programul a făcut deja analiza ROC și vă prezintă în partea dreaptă tabele pentru diferite valori limită (**Cutpoint**), cu sensibilitatea (**Se**), specificitatea (**Sp**), valoarea predictivă pozitivă (**PPV**) și negativă (**NPV**), indicele Youden (**Youden's index**), aria de sub curba ROC (**AUC**).

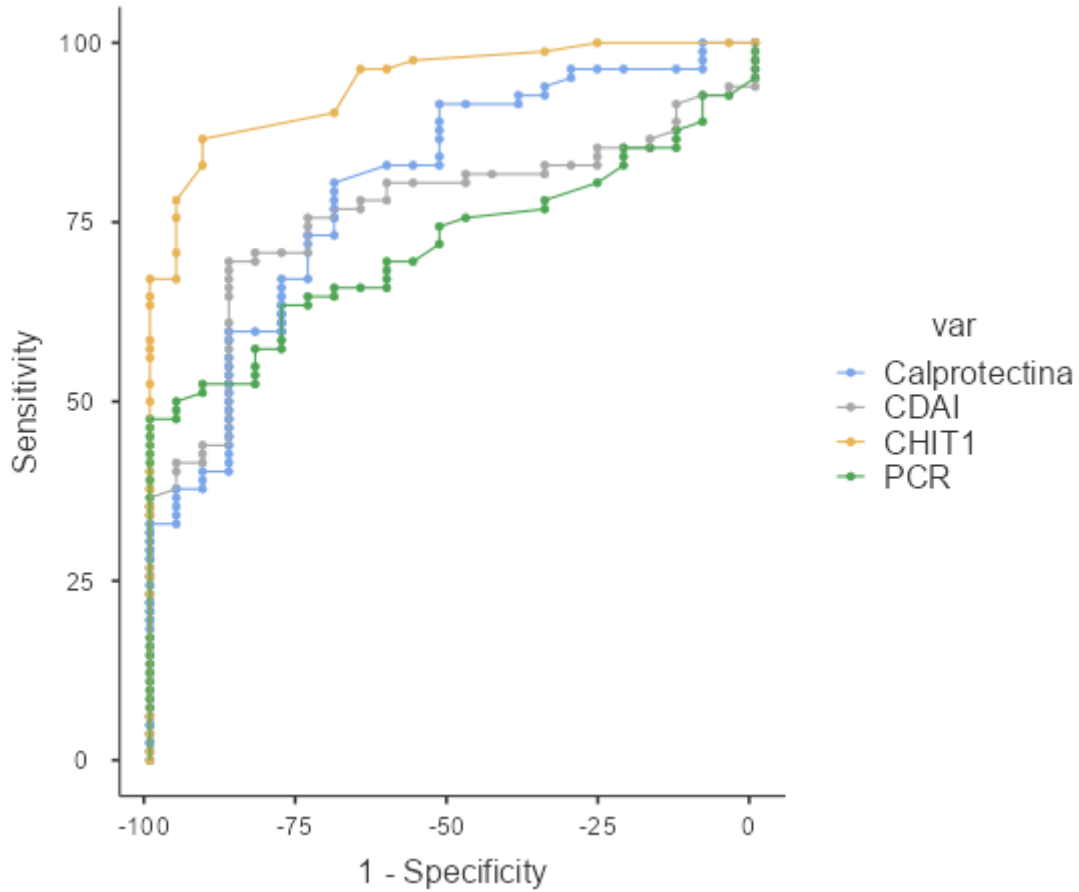
Astfel **cea mai bună valoare limită**, care are indicele Youden are valoarea cea mai mare, este 210, având sensibilitatea și specificitatea asociate de 80,49%, respectiv 69,57%.

Scale: Calprotectina

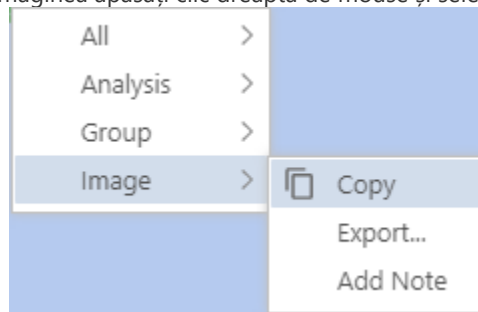
Cutpoint	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Youden's index	AUC	Metric Score
210	80.49%	69.57%	90.41%	50%	0.501	0.801	1.50
215	79.27%	69.57%	90.28%	48.48%	0.488	0.801	1.49
220	78.05%	69.57%	90.14%	47.06%	0.476	0.801	1.48
221.6	76.83%	69.57%	90%	45.71%	0.464	0.801	1.46
240	75.61%	69.57%	89.86%	44.44%	0.452	0.801	1.45
280	73.17%	73.91%	90.91%	43.59%	0.471	0.801	1.47
300	71.95%	73.91%	90.77%	42.5%	0.459	0.801	1.46
400	67.07%	78.26%	91.67%	40%	0.453	0.801	1.45
596	59.76%	86.96%	94.23%	37.74%	0.467	0.801	1.47
600	58.54%	86.96%	94.12%	37.04%	0.455	0.801	1.45

Grafic de tip curbă ROC

Mai jos este afișat graficul cu curba ROC (Receiver operating characteristic - caracteristica de funcționare a receptorului).

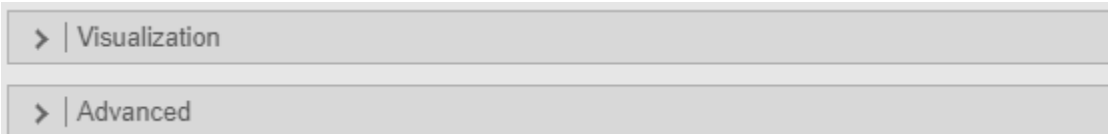


Pentru a copia imaginea apăsați clic dreapta de mouse și selectați **Image Copy**



Comparare curbe ROC prin teste statistice

Pentru realizarea de **teste statistice care să compare curbele ROC** între ele, apăsați pe **butonul >** din dreptul **Advanced**, și selectați opțiunea **De Long's test**. Pentru identificarea valorii limită se selectează la **Method Maximize metric**, iar la **Metric Youden-Index**.



Advanced

DeLong's test

Method: Maximize metric

Metric: Youden-Index

Tolerance metric: 0.05

Rezultatele arată în genul celor de mai jos:

DeLong Test of Difference between AUCs

Estimated AUC's:

	AUC	SD(Hanley)	P(H0: AUC=0.5)	SD(DeLong)	P(H0: AUC=0.5)
1	0.762	0.050	0.000	0.048	0.000
2	0.717	0.055	0.000	0.049	0.000
3	0.801	0.045	0.000	0.050	0.000
4	0.942	0.022	0.000	0.022	0.000

Pairwise comparisons:

	AUC Difference	CI(lower)	CI(upper)	P.Value	Correlation
1 vs. 2	0.045	-0.059	0.149	0.400	0.400
1 vs. 3	-0.039	-0.145	0.066	0.465	0.409
1 vs. 4	-0.179	-0.266	-0.092	0.000	0.406
2 vs. 3	-0.084	-0.187	0.019	0.110	0.437
2 vs. 4	-0.224	-0.318	-0.131	0.000	0.268
3 vs. 4	-0.140	-0.228	-0.052	0.002	0.455

Overall test:

p-value = 1.28e-06

În primul tabel sunt trecute estimările **suprafețelor de sub curba ROC** pentru fiecare test diagnostic, deviația standard, precum și un **test statistic de semnificație pentru o curbă ROC**, pentru fiecare variabilă selectată în ordinea selectării lor. Astfel pentru scorul de activitate clinică CDAI, suprafața de sub curba ROC este 0,717, iar rezultatul este statistic semnificativ P (H0: AUC=0.5) fiind mai mic de 0,05 (valoare afișată 0,001). **Ipoteza nulă** a testului pentru **o curbă ROC** este: aria de sub curba ROC pentru CDAI comparat cu examenul histopatologic, nu este statistic semnificativ diferită de 0,5. **Ipoteza alternativă** a testului este: aria de sub curba ROC pentru CDAI comparat cu examenul histopatologic, este statistic semnificativ diferită de 0,5.

În al doilea tabel sunt **comparații între curbele ROC** luate **două câte două**, prezentându-se diferența dintre suprafețele curbelor ROC (AUC Difference), cu interval de încredere 95%, valoare test statistic (**P.Value**). Spre exemplu, comparând scorul de activitate clinică CDAI (1) cu proteina C reactivă (2), rezultatul nu este statistic semnificativ (p=0,400). **Ipoteza nulă a testului de comparație a două curbe ROC**: Nu există diferență semnificativă statistic între acuratețea diagnostică a scorului de activitate clinică CDAI și a proteinei C reactive, măsurată prin AUC, având ca standard examenul histopatologic. **Ipoteza alternativă a testului de comparație a două curbe ROC**: Există diferență semnificativă statistic între acuratețea diagnostică a scorului de activitate clinică CDAI și a proteinei C reactive, măsurată prin AUC, având ca standard examenul histopatologic.

În final este trecut rezultatul unui test statistic global care compară toate curbele ROC între ele.

Analiza supraviețuirii (graficul Kaplan Meyer, testul log-rank, regresia Cox – HR)

Asigurați-vă că **jsurvival - Modulul de supraviețuire al ClinicoPath** pentru Jamovi 0.0.2 este instalat. Consultați capitolul **Instalarea unui modul de analize suplimentar**.

Verificați dacă tipul de variabile este corect ales și dacă ordinea categoriilor pentru fiecare variabilă este cea preferată.

Mutați variabila care indică timpul de **supraviețuire** făcând clic pe săgeata de lângă câmpul **Timp scurs (Time Elapsed)**. Mutați variabila care indică **starea/evenimentul** cenzurii făcând clic pe săgeata de lângă câmpul **Rezultat**. Asigurați-vă că este selectată **categoria** care indică **evenimentul** în lista **Nivel eveniment (Event Level)** (de exemplu, 1). Mutați variabila de grupare făcând clic pe săgeata de lângă câmpul **Variabilă explicativă (Explanatory Variable)** (de exemplu, Tip, Tratament, Factor de risc).

În secțiunea **Parcele**, verificați opțiunea **Complot de supraviețuire**.

Survival Analysis

ID
Traitement

Time Elapsed
Temp_mois

Outcome
Censure
Event Level 1

Explanatory Variable
Type_histologique

> | Advanced Elapsed Time Options

> | Analysis with Multiple Outcomes

> | Single Arm Survival

> | Pairwise Comparisons

∨ | Plots

Plots

Survival Plot

Puteți modifica opțiunile graficului: pentru a afișa valoarea **p**, tabelul cu numărul de subiecți expuși riscului (**risktable**).

Plot Arguments

Plot End Time

Time Interval

Start y-axis

End y-axis

95% CI

risktable

censored

p-value

Survival Tables

Survival table

Cutpoints

Pentru a afla probabilitatea de supraviețuire în momente diferite, puteți specifica aceste perioade în secțiunea **Tabel de supraviețuire**, câmpul **Puncte de tăiere**.

Mediana timpului de supraviețuire

În secțiunea **Rezumatul și tabelul privind supraviețuirea mediană (Median Survival Summary and Table)**, supraviețuirea mediană cu intervale de încredere de 95% pot fi găsite în textul dinaintea tabelului. De asemenea, în tabel, în coloana **Mediană (Median)**, găsim mediana pe grupe, iar intervalul de încredere 95% în ultimele două coloane. De exemplu, pentru grupul histologie cu celule mici, supraviețuirea mediană este de 2 luni.

Median Survival Summary and Table - Type_histologique

When Type_histologique is Petites cellules, median survival is 2 [1 - 2, 95% CI] months.
 When Type_histologique is Squameux, median survival is 4 [3 - 10, 95% CI] months.

Median Survival Table: Levels for Type_histologique

Levels	Records	Events	rmean	se_rmean	Median	95% Confidence Interval	
						Lower	Upper
Type_histologique=Petites cellules	48	45	2.72	0.497	2.00	1.00	2.00
Type_histologique=Squameux	35	31	7.76	1.601	4.00	3.00	10.00

Regresie Cox – hazard ratio

În tabelul cu rezumatul și tabelul **regresiei Cox (Cox Table)**, puteți găsi hazard ratio cu intervalul de încredere de 95%, în coloana **HR (univariable)**. Pentru a afla care grup este **grupul de referință**, în coloana **HR (Univariable)** căutați grupul care afișează **semnul -**. În exemplul de mai jos, putem vedea că grupul de referință este tipul histopatologic de celule mici. Deci, HR se obține prin împărțirea hazardului participanților cu cancer cu histologie scuamoasă, la hazardul participanților cu cancer cu histologie cu celule mici.

Cox Regression Summary and Table - Type_histologique

When Type_histologique is Squameux, there is 0.43 (0.26-0.72, p=0.001) times risk than when Type_histologique is Petites cellules.

Cox Table- Type_histologique

Explanatory	Levels	all	HR (Univariable)
Type_histologique	Petites cellules	48 (57.8)	-
	Squameux	35 (42.2)	0.43 (0.26-0.72, p=0.001)

Probabilitatea de supraviețuire în momente diferite

Probabilitatea de supraviețuire la diferite momente de timp (de exemplu, în coloana Timp, 12 luni) poate fi găsită în **Rezumatul și tabelul Supraviețuire (1,3,5-yr Survival Summary and Table)**, în coloana **Supraviețuire (Survival)**, iar intervalul de încredere 95% în ultimele două coloane, dar și în format text, înainte de tabel. De exemplu, pentru grupul histologie cu celule mici, probabilitatea de supraviețuire la 12 luni este de 5,9%.

1, 3, 5-yr Survival Summary and Table - Type_histologique

When Type_histologique=Petites cellules, 12 month survival is 6% [1.6%-22%, 95% CI].

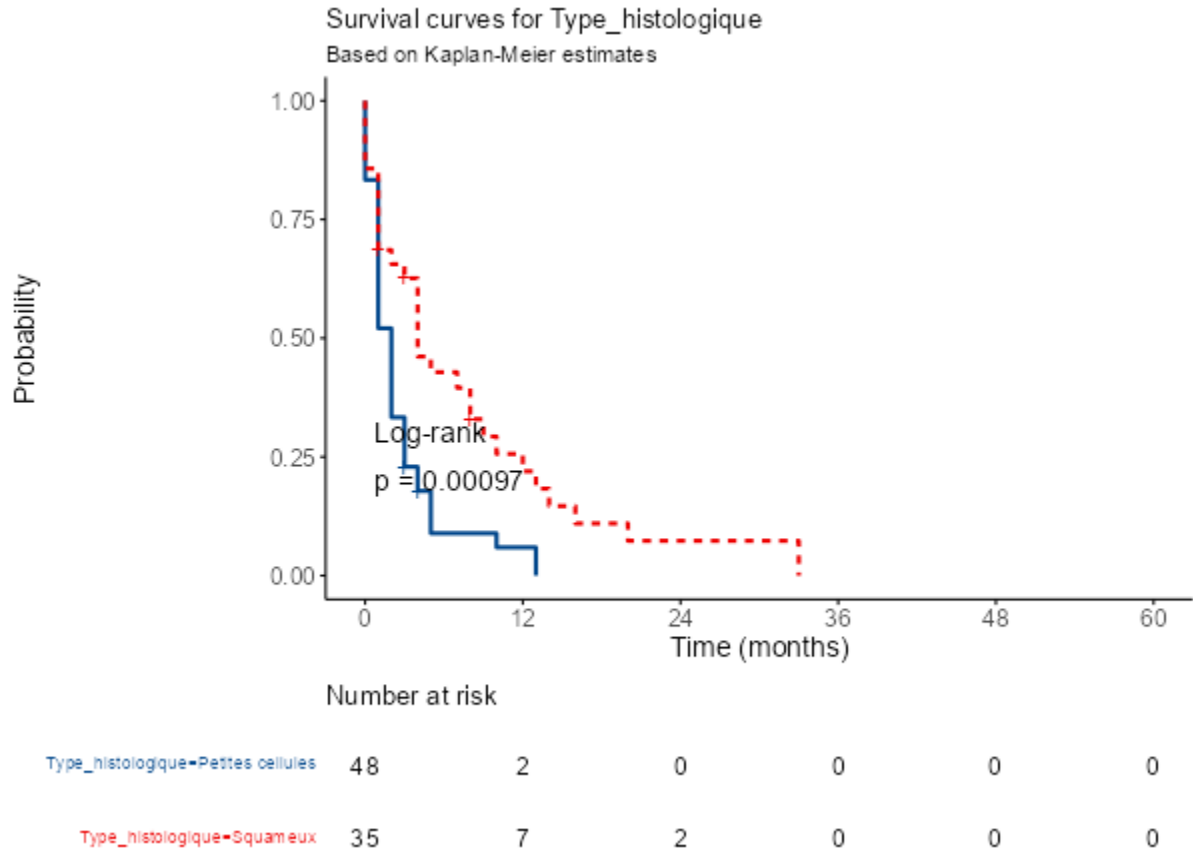
When Type_histologique=Squameux, 12 month survival is 22% [11.1%-43%, 95% CI].

1, 3, 5 year Survival - Type_histologique

Levels	time	Number at Risk	Number of Events	Survival	95% Confidence Interval	
					Lower	Upper
Type_histologique=Petites cellules	12	2	43	5.9 %	1.6 %	21.6 %
Type_histologique=Squameux	12	7	25	22.0 %	11.1 %	43.3 %

Curba de supraviețuire Kaplan-Meier, cu testul log-rank

Curba de supraviețuire Kaplan-Meier, cu valoarea p a testului log-rank și tabelul cu numărul de subiecți la risc, pot fi găsite în secțiunea **Survival Plot**. În tabelul cu numărul subiecților la risc, putem observa la începutul studiului că există 48 de subiecți în viață (cu risc de a face evenimentul – deces) în grupul cu histologie cu celule mici și 35 de subiecți în grupul cu histologie scuamoasă. La 12 luni, au mai rămas doar 2 subiecți în viață (cu risc de a face evenimentul – deces) în grupul cu histologie cu celule mici și 7 subiecți în grup cu histologie scuamoasă.



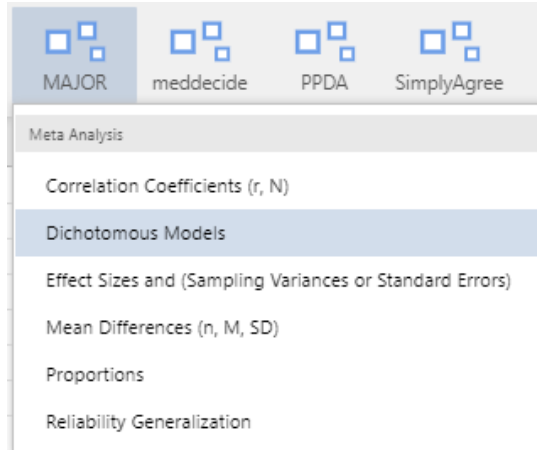
Meta-analiza

Meta-analiza pentru compararea a două grupuri în ce privește date de tip dihotomial

Ne asigurăm că **modulul MAJOR** este **instalat**. Dacă nu este instalat urmați instrucțiunile din capitolul Instalare modul de analize suplimentare.

Ne asigurăm că **tipul variabilelor** care conțin numărul de evenimente, respectiv numărul total de subiecți este de tip **Continuous**.

Din **modulul MAJOR**, alegem opțiunea **Dichotomous Models**.



Deplasați variabila care indică numărul de evenimente în grupul experimental/de intervenție/expus în câmpul **Number of Incidents in Experimental Group** (număr de incidente în grupul experimental); variabila care indică numărul de evenimente în grupul de control/neexpus în câmpul **Number of Incidents in Control Group** (număr de incidente în grupul de control); variabila care indică numărul total de participanți în grupul experimental/de intervenție/expus în câmpul **Total Sample Size for Experimental Group** (număr de subiecți pentru grupul experimental); variabila care indică numărul total de participanți în grupul de control/neexpus în câmpul **Total Sample Size for Control Group** (număr de subiecți pentru grupul de control); variabila care indică denumirea studiului în câmpul **Study Label** (eticheta studiului).

În secțiunea de opțiuni ale modelului (**Model Options**), alegeți la estimatorul modelului (**Model Estimator**), **Restricted Maximum-Likelihood**; la estimatorul de interes al asocierii (**Model Measures**), Log odds ratio (pentru studii caz martor), log risk ratio sau risk difference (pentru studii de cohortă); bifați **Back-Transform Log Odds Ratio to Odds Ratio** (transformare logaritm din odds ratio în odds ratio).

Restul opțiunilor pot fi lăsate așa cum sunt. Dacă se dorește se poate alege în caz de heterogenitate redusă la Fail-Safe N method Rosenberg.

Dichotomous Models ➔

Tip studiu 🔍

➔ Number of Incidents in Experimental Group
Evenimente Expusi

➔ Number of Incidents in Control Group
Evenimente Control

➔ Total Sample Size for Experimental Group
Total Expusi

➔ Total Sample Size for Control Group
Total Control

➔ Study Label
Tip studiu

➔ Moderator (optional)

▼ | Model Options

Model estimator Restricted Maximum-Likelihood ▼

Model measures Log odds ratio ▼

Moderator type No Moderator ▼

Back-Transform Log Odds Ratio to Odds Ratio

Estimatorul asocierii

Dacă s-a selectat ca estimator al asocierii Log odds ratio, precum și dacă s-a bifat **Back-Transform Log Odds Ratio to Odds Ratio** (transformare logaritm din odds ratio în odds ratio), atunci în tabelul denumit la fel, se găsește valoarea estimatorului punctual al OR, precum și intervalul de încredere 95%.

Back-Transform Log Odds Ratio to Odds Ratio		
Odds Ratio	CI Lower Bound	CI Upper Bound
1.542	1.135	2.096

În **tabelul Random-Effects Model** (modelul cu efecte aleatorii), este trecută valoarea estimatorului asocierii log odds ratio/log risk ratio/risk difference, în coloana estimate, cu intervalul de încredere în ultimele două coloane. Valoarea p pentru semnificația statistică a acestuia se află în coloana p.

Random-Effects Model (k = 10)

	Estimate	se	Z	p	CI Lower Bound	CI Upper Bound
Intercept	0.433	0.157	2.77	0.006	0.126	0.740

Note. Tau² Estimator: Restricted Maximum-Likelihood

Heterogenitatea statistică a studiilor

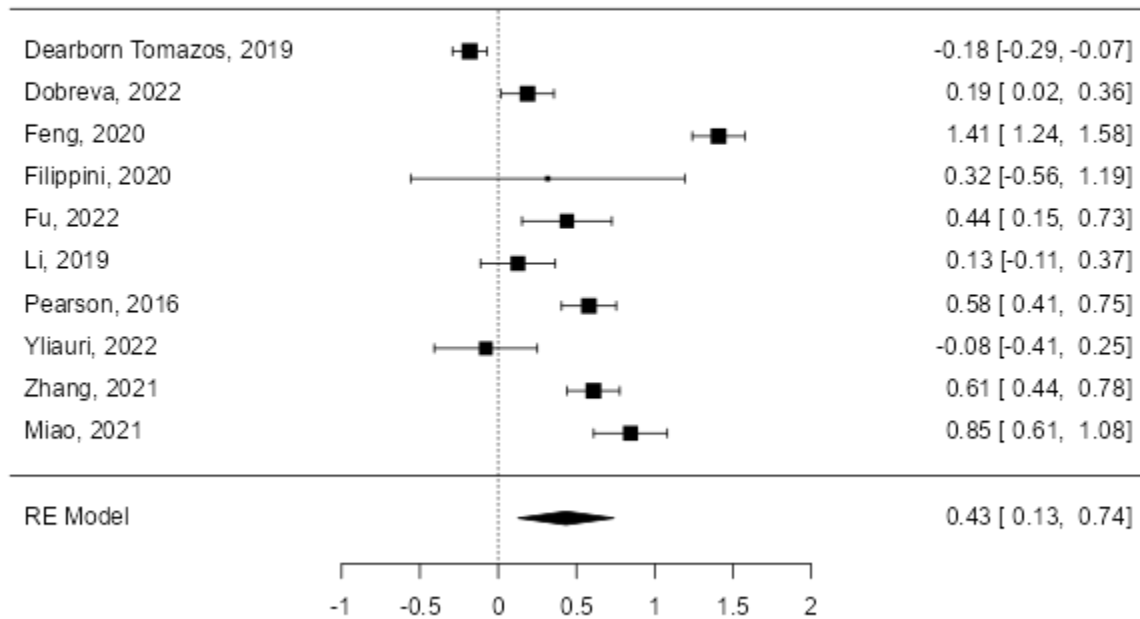
În tabelul Heterogeneity Statistics (statistici cu privire la heterogeneitatea), în coloana **I²** se află indicele de inconsistență, iar în coloana **p**, valoarea p a testului asociat.

Heterogeneity Statistics

Tau	Tau ²	I ²	H ²	R ²	df	Q	p
0.471	0.222 (SE= 0.1147)	95.71%	23.304	.	9.000	288.885	< .001

Graficul pădure (forest plot)

În secțiunea grafic pădure (**Forest plot**), aveți graficul omonim. Din păcate, pentru estimatorii asocierii log odds ratio, log risk ratio, în grafic sunt trecute valorile logaritmice ale OR sau RR. Avantajul este simetria intervalelor de încredere. Dezavantajul este că este mai dificil de interpretat. Valoarea lui 0 este echivalentă cu un OR sau un RR de 1. Valorile de pe grafic mai mari decât 0, pentru OR/RR sunt mai mari decât 1, iar valorile mai mici decât 0 sunt mai mici decât 1.



Eroarea sistematică de publicare

Statistici

Pentru evaluarea erorii sistematice de publicare se poate folosi rezultatul testului statistic al regresiei pentru asimetria graficului pâlnie (**Regression Test for Funnel Plot Asymetry**), valoarea p din tabel, din secțiunea **Publication Bias Assessment** (evaluarea erorii sistematice de publicare).

Publication Bias Assessment

Fail-Safe N Analysis (File Drawer Analysis)

Fail-safe N	p
593.000	< .001

Note. Fail-safe N Calculation Using the Rosenthal Approach

Rank Correlation Test for Funnel Plot Asymmetry

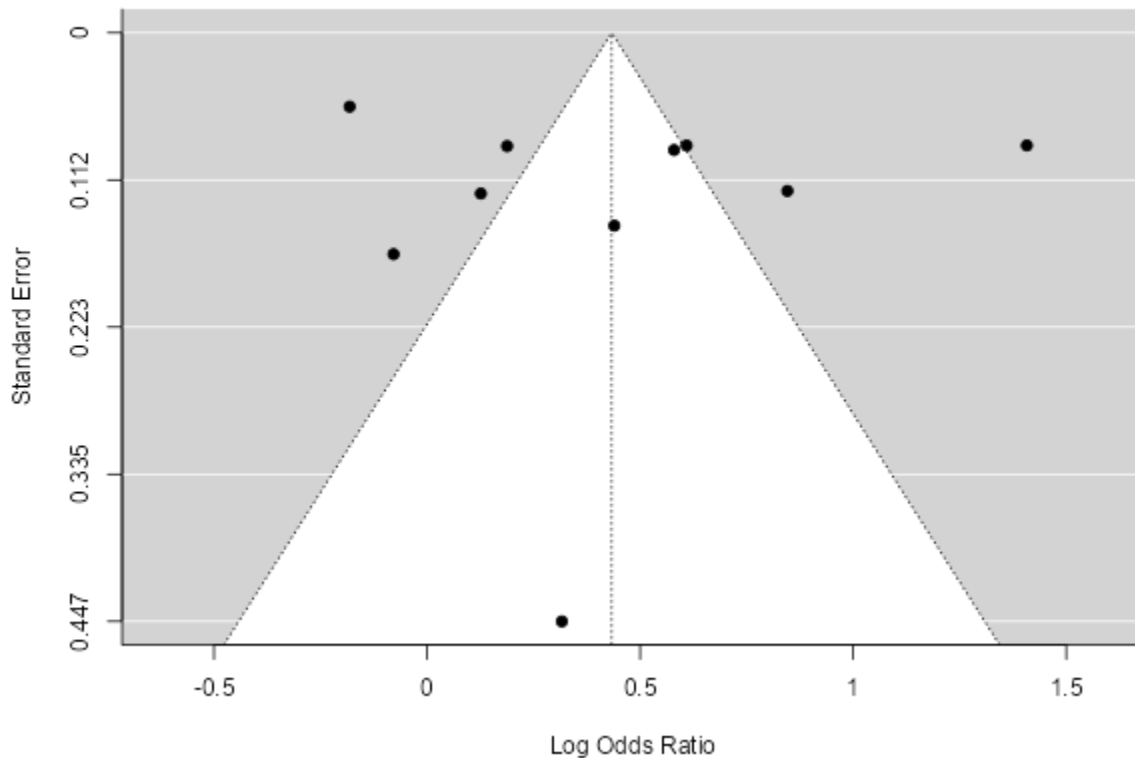
Kendall's Tau	p
-0.156	0.601

Regression Test for Funnel Plot Asymmetry

Z	p
-0.306	0.760

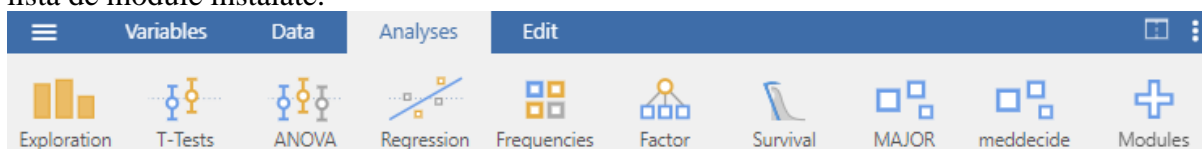
Grafic de tip pâlnie (funnel plot)

Graficul de tip pâlnie se găsește în secțiunea **Funnel plot**.

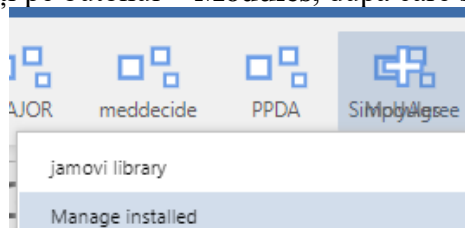


Instalare modul de analize suplimentare (ex. pentru analiza ROC)

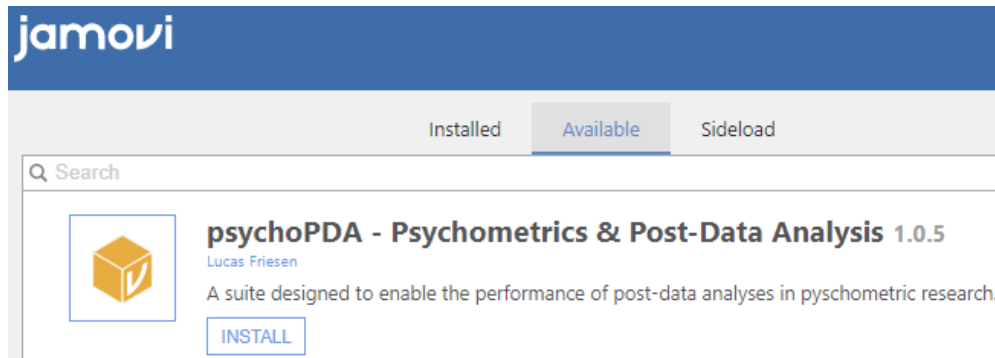
În tabul Analyses, în dreapta, apăsați pe + Modules și verificați dacă nu este prezent în lista de module instalate.



Pentru a instala module de analiză suplimentare celor implicite, selectați tabul **Analyses**, iar în partea dreaptă apăsați pe butonul + **Modules**, după care selectați **Manage installed**.



Selectați tabul **Available**, derulați până găsiți modulul dorit (ex. psychoPDA) și apăsați butonul **INSTALL**.



Reveniți la fereastra de analiză apăsând butonul săgeată:



Instrucțiuni utilizare Excel

Descrierea datelor

variabile **calitative**:

Tabele de frecvență (două metode)

- În **Excel** utilizați funcția **COUNTIF** pentru a număra de câte ori apare fiecare valoare în baza de date (frecvența absolută a valorii). *Ex. pentru a afla câte persoane de sex feminin (codate în fișier cu F), există în eșantion, se creează o celulă care să conțină ceva de genul =COUNTIF(A2:A58,"F"), întrucât datele se aflau în celulele de la A2 până la A58. Un tabel de frecvență arată în felul următor (Linia cu Total a fost creată manual după numărarea subiecților cu COUNTIF):*

Sex	Număr subiecți
Feminin	600
Masculin	500
Total	1100

- În **Excel**, selectați oricare celulă care conține date. Apoi selectați opțiunea **PivotTable and PivotChart Report...** din meniul **DATA**. Apare o fereastră în care apăsați pe butonul **Next**. În fereastra nou apărută Excel a selecționat pentru dumneavoastră toate datele din foaia de calcul. Dacă nu este așa, trebuie selectate manual, inclusiv titlurile coloanelor. Apăsați pe butonul **Next**. În fereastra nou apărută vă asigurați că este selecționat ca rezultatele să apară pe o nouă foaie de calcul (opțiunea **New worksheet**). Apăsați pe

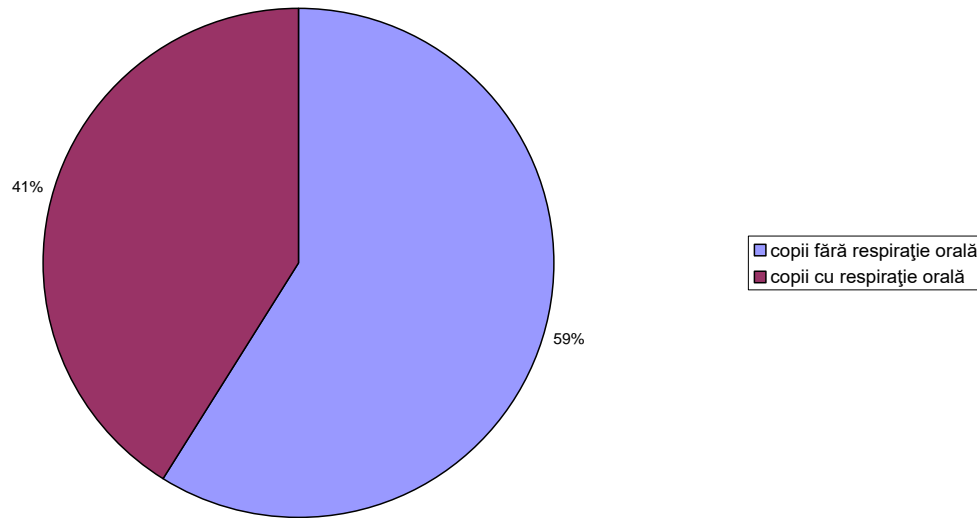
butonul **Finish**. Apare o nouă foaie de calcul cu un tabel gol și o fereastră **Pivot Table Field List**. Din această fereastră trageți variabila dorită (ex. *sex*) în tabelul gol în zona **Drop Row Fields Here**, pentru ca Excel să identifice categoriile. Puteți redenumi la nevoie aceste categorii pentru a fi explicite (ex. *dacă sexul este codat cu f pentru feminin va fi redenumit în feminin*). Din aceeași fereastră **Pivot Table Field List**, trageți aceeași variabilă în tabelul gol în zona **Drop Data Fields Here**. Excel realizează tabelul de frecvență inclusiv cu totalul. Redenumiți Grand Total în Total, pentru ca tabelul să fie complet în limba română. Dacă efectuați din nou un **PivotTable**, înainte de fereastra de selectare a datelor vă întreabă dacă nu doriți să folosiți selecția anterioară pentru a economisi memorie. Selectați **OK**, iar apoi selectați selecția anterioară din lista de selecții. Apoi urmați pașii ca și mai sus.

Grafice de tip sectorial (pie) (două metode)

- Dacă ați creat tabelul de frecvență cu ajutorul opțiunii **Pivot Table**, trebuie să selectați din meniul **Insert** opțiunea **Chart**. Pentru schimbarea tipului de grafic dați click dreapta în zona albă a graficului spre colțul stânga sus. Selectați **Chart Type** și alegeți **Pie**. Pentru mascarea butoanelor de pe grafic dați click dreapta pe butonul **Count of** și selectați **Hide PivotChart Field Buttons**. Pentru afișarea frecvențelor sau a procentajelor dați click dreapta în zona albă a graficului spre colțul stânga sus. Selectați **Chart Options** și din fereastra apărută în tab-ul **Data Labels** bifați opțiunea **Percentage** sau **Value**.

sau

- În **Excel** urmați indicațiile de mai sus pentru a realiza tabelul de frecvențe cu ajutorul funcției **COUNTIF**. Selectați tabelul ce conține denumirile categoriilor și frecvențele corespunzătoare (fără total). Dați apoi **INSERT GRAPH** tip **pie**. Apăsăți pe butonul **Next**. În fereastra nou apărută (**Chart Options**) dați click pe tab-ul **Data Labels** și bifați opțiunea **Percentage** sau **Value**. Apăsăți pe butonul **Next**. Apăsăți pe butonul **Finish**. *Exemplu de grafic de tip sectorial:*



Tabele de contingență

- În **Excel**, selectați oricare celulă care conține date. Apoi selectați opțiunea **PivotTable and PivotChart Report...** din meniul **DATA**. Apare o fereastră în care apăsați pe butonul **Next**. În fereastra nou apărută Excel a selecționat pentru dumneavoastră toate datele din foaia de calcul. Dacă nu este așa, trebuie selectate manual, inclusiv titlurile coloanelor. Apăsați pe butonul **Next**. În fereastra nou apărută vă asigurați că este selecționat ca rezultatele să apară pe o nouă foaie de calcul (opțiunea **New worksheet**). Apăsați pe butonul **Finish**. Apare o nouă foaie de calcul cu un tabel gol și o fereastră **Pivot Table Field List**. Din această fereastră trageți de obicei factorul pronostic (*ex. pt. studii prognostice sau de factori de risc se trage factorul pronostic/de risc; pt. studii diagnostice se trage noul studiu diagnostic; pt. studii terapeutice se trage variabila reprezentând tratamentul*) în tabelul gol în zona **Drop Row Fields Here**, pentru ca Excel să identifice categoriile. Din această cu câmpurile trageți variabila rezultat urmărit (*ex. pt. studii prognostice sau de factori de risc se trage variabila reprezentând boala; pt. studii diagnostice se trage variabila reprezentând studiu diagnostic de referință [standard de aur]; pt. studii terapeutice se trage variabila reprezentând rezultatul urmărit [răspuns/vindecare/ameliorare]*) în tabelul gol în zona **Drop Column Fields Here**, pentru ca Excel să identifice categoriile. Puteți redenumi la nevoie aceste categorii pentru a fi explicite (*ex.*

dacă sexul este codat cu f pentru feminin va fi redenumit în feminin). Din aceeași fereastră **Pivot Table Field List**, trageți oricare variabilă din cele două folosite în tabelul gol în zona **Drop Data Fields Here**. Excel realizează tabelul de contingență inclusiv cu totalul. Redenumiți Grand Total în Total, pentru ca tabelul să fie complet în limba română.

Graficul asociat tabelului de contingență

- După ce ați creat tabelul de frecvență cu ajutorul opțiunii **Pivot Table**, trebuie să selectați din meniul **Insert** opțiunea **Chart**. Pentru mascarea butoanelor de pe grafic dați click dreapta pe butonul **Count of** și selectați **Hide PivotChart Field Buttons**. Pentru afișarea frecvențelor dați click dreapta în zona albă a graficului spre colțul stânga sus. Selectați **Chart Options** și din fereastra apărută în tab-ul **Data Labels** bifați opțiunea **Percentage** sau **Value**, iar în tabul **Titles** introduceți dacă doriți titlul graficului și neapărat introduceți definițiile axelor cu unitățile de măsură, fără a folosi prescurtări.

variabile **cantitative**:

descrierea unei variabile

Media, mediana, deviația standard și intervalul de încredere 95% pentru medie.

- În **Excel** folosiți **Descriptive Statistics** din **Data Analysis** din meniul **Tools**. În fereastra **Descriptive Statistics** selectați opțiunile **Summary Statistics, Confidence Level for Mean**. Apare o nouă foaie de calcul, cu datele dorite: mean (media - folosită când datele au distribuție normală (histograma are formă de clopot)), median (mediana – folosită când datele nu au distribuție normală (histograma nu are formă de clopot)), standard deviation (deviația standard). Pentru aflarea limitei inferioare a intervalului de încredere al mediei calculați media minus Confidence Level (95,0%). Pentru aflarea limitei superioare a intervalului de încredere al mediei calculați media plus Confidence Level (95,0%).

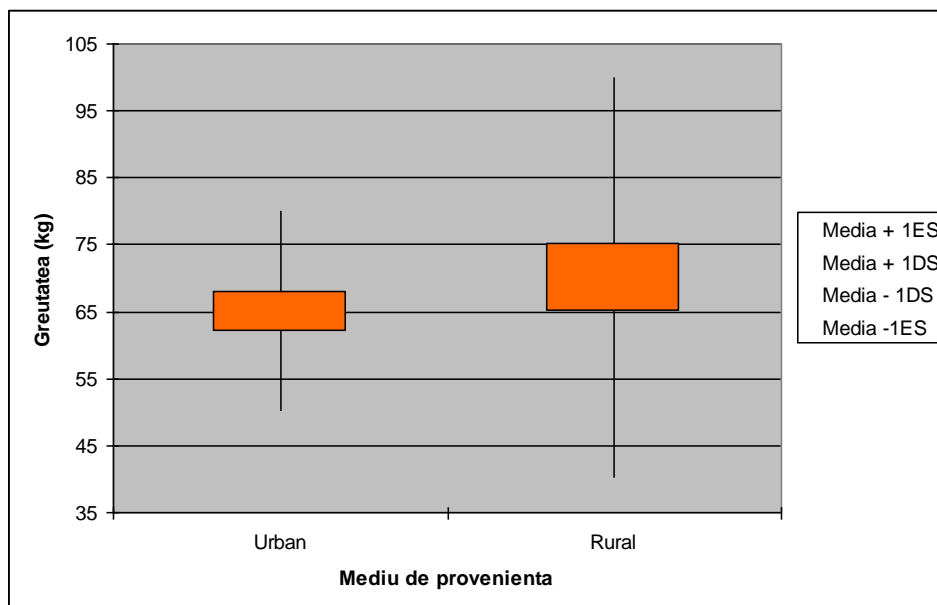
Grafic boxplot

- În Excel aflați cu ajutorul Descriptive Statistics media, deviația standard (DS) și eroarea standard (ES) pentru variabila(ele) care vă

interesează. Creați în celulele unei foi de calcul Excel un mic tabel de forma (e important să aranjați exact cum vedeți în tabel prima coloană. Nu e necesar să aveți mai multe categorii.):

	Urban	Rural
Media + 1ES	68	75
Media + 1DS	80	100
Media - 1DS	50	40
Media -1ES	62	65

- Selectați apoi acest tabel. Apoi selectați din meniu **Insert Chart**. Din fereastra apărută selectați **Stock** apăsați butonul **Next**. În fereastra nou apărută în tabul **Data Range** selectați **Rows**. Apăsați butonul **Next**. În noua fereastră denumită **Chart Options** în tabul **Titles** introduceți dacă doriți titlul graficului și neapărat introduceți definițiile axelor cu unitățile de măsură, fără a folosi prescurtări. În tab-ul **Gridlines** puteți adăuga sau elimina linii care să ajute orientarea în grafic. În tab-ul **Data Labels** asigurați-vă că e debifat tot. Apăsați butonul **Next**. În fereastra denumită **Chart Locations** selectați dacă doriți ca graficul să fie salvat pe foaia curentă, sau într-o nouă foaie de calcul pentru care trebuie introdusă o denumire sugestivă pentru regăsirea facilă ulterioară. Apăsați butonul **Finish**. Cu click dreapta puteți modifica culorile cutiei (cutiilor)



Tabele de frecvență

- În **Excel** identificați valoarea minimă și maximă pe care o ia variabila de interes, și intervalul dintre acestea (utilizând **Descriptive Statistics** din **Data Analysis – minimum, maximum, range**). Împărțiți intervalul dintre minim și maxim la 7 pentru a afla dimensiunea intervalului de frecvențe. În foaia de calcul cu variabila de interes creați o coloană denumită **Intervale nume variabilă de interes (unitate de măsură)** care să cuprindă 6 numere, distanțate la dimensiunea anterior calculată, situate între minim plus dimensiunea intervalului și maxim minus dimensiunea intervalului, reprezentând intervalele de frecvență dorite. Folosiți **Histogram** din **Data Analysis** din meniul **Tools**. La **Input Range** selectați variabila de interes inclusiv denumirea acesteia din primul rând. La **Bin Range** selectați coloana nou creată (denumirea acesteia și datele reprezentând intervalele de frecvență). Selectați **Labels**. La **New Workseet Ply** scrieți denumirea foii de calcul în care va apărea tabelul de frecvență. Selectați **Chart Output** dacă doriți să afișați histograma. *Exemplu de tabel de frecvență:*

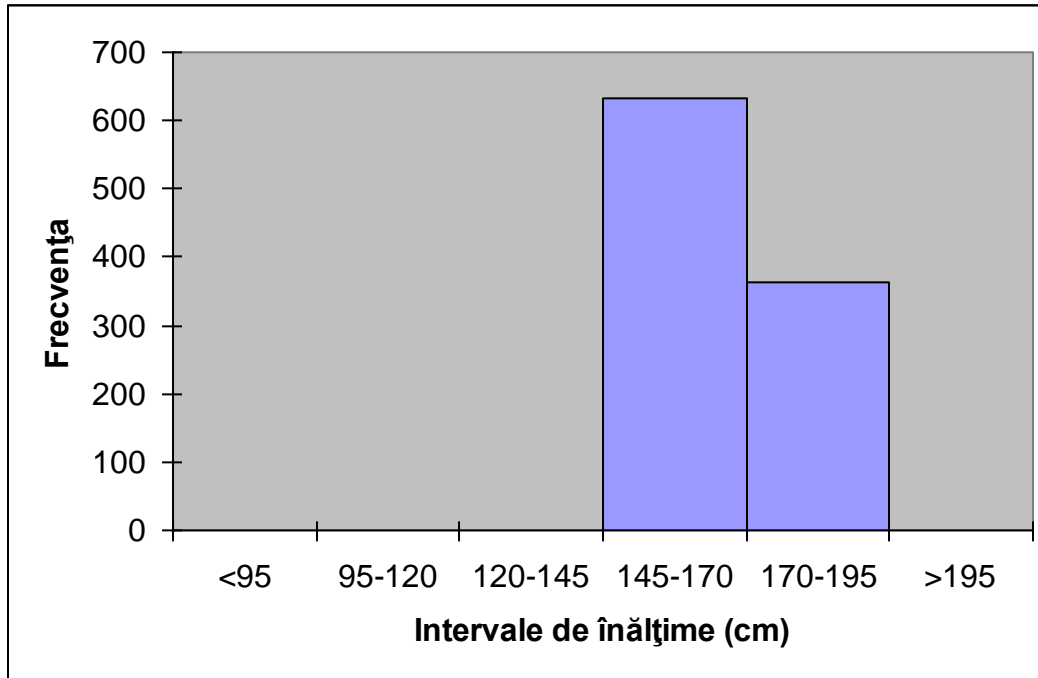
Intervale de înălțime (cm)	Frecvența
<95	1
95-120	0
120-145	1
145-170	634
170-195	364
>195	0

Histograme

- În **Excel** conform explicației de la tabele de frecvență, dar în fereastra **Histogram** e necesară selectarea **Chart Output**.
- **Pentru unirea coloanelor histogramei**. Selectați oricare coloană a acesteia și apăsați click dreapta. Selectați **Format data series...** . În fereastra apărută în tabul **Options**, câmpul **Gap width** tastați 0. Apăsați butonul **OK**.
- **Nu uitați să definiți corect și în aceeași limbă** (română, după dorință altă limbă, în funcție de textul în care va fi inserat graficul) **axele și să eliminați legenda**, utilizând fereastra **Chart Options**, apelabilă printr-

un click dreapta în zona albă din colțul dreapta sus a graficului și selectarea **Chart Options...** . În tab-ul **Legend** debifați legenda, în tabul **Titles** introduceți dacă doriți titlul graficului și neapărat introduceți definițiile axelor cu unitățile de măsură, fără a folosi prescurtări.

- Exemplu histogramă:



descrierea relației între **două variabile**

diagramă nor de puncte (Scatter)

- În **Excel** selectați coloanele corespunzătoare celor două variabile inclusiv titlul acestor coloane. Apoi selectați din meniul **Insert Chart**. Din fereastra apărută selectați **XY (Scatter)** apăsați butonul **Next**. În fereastra nou apărută în tabul **Data Range** este selectat **Columns**. Se dă click pe tabul **Series** și se verifică ce variabile se află pe axa X și pe axa Y. În mod normal pe axa X se află prima coloană din stânga a selecției, iar pe Y cealaltă coloană. Dacă doriți ca variabilele să fie altfel afișate selectați în câmpurile **X values** și **Y values** variabilele dorite. Apăsați butonul **Next**. În noua fereastră denumită **Chart Options** în tabul **Titles** introduceți dacă doriți titlul graficului și neapărat introduceți definițiile axelor cu unitățile de măsură, fără a

folosi prescurtări. În tab-ul **Gridlines** puteți adăuga sau elimina linii care să ajute orientarea în grafic. În tab-ul **Legend** debifați legenda. În tab-ul **Data Labels** asigurați-vă că e debifat tot. Apăsați butonul **Next**. În fereastra denumită **Chart Locations** selectați dacă doriți ca graficul să fie salvat pe foaia curentă, sau într-o nouă foaie de calcul pentru care trebuie introdusă o denumire sugestivă pentru regăsirea facilă ulterioară. Apăsați butonul **Finish**.

- **Adăugarea dreptei de regresie și a coeficientului de determinare pe grafic.** Selectați oricare punct al graficului sectorial. Apăsați click dreapta, selectați **Add trendline...** . În fereastra apărută în tabul **Type**, selectați **Linear**. În tabul **Options** bifați **Display equation on chart** și **Display R-squared value on chart**, după dorință. Apăsați butonul **OK**.
- **Modificarea încadrării punctelor astfel încât să ocupe cea mai mare suprafață a graficului.** Selectați axa X, sau Y, dacă punctele nu ocupă majoritatea suprafeței pentru axa dorită, fiind înghesuite într-o parte. Apăsați click dreapta, selectați **Format axis...** . În fereastra apărută în tabul **Scale**. În câmpurile minimum, maximum alegeți valori pentru a încadra corespunzător punctele din graficul de nor de puncte. Apăsați butonul **OK**.
- **Modificarea caracteristicilor oricărui element al graficului.** Selectați elementul dorit apoi apăsați click dreapta. Din opțiunile afișate selectați formatările dorite. Apăsați butonul **OK**.

variabile de **supraviețuire**:

- **Mediana** timpului de supraviețuire
 - În **Excel** folosiți **Descriptive Statistics** din **Data Analysis** din meniul **Tools**. În fereastra **Descriptive Statistics** selectați **Summary Statistics**.
- **Apoi urmați instrucțiunile din EpiInfo – comanda Kaplan-Meyer Survival**

Activarea modului Data Analysis din meniul Tools în Excel

- În meniul **Tools** se selectează **Add Ins**. În fereastra apărută se bifează în pătratele corespunzătoare modulele **Analysis ToolPak** și **Analysis ToolPak -**

VBA, iar apoi se dă OK. Apoi se selectează o celulă goală în foaia de calcul iar în meniul **Tools** va apărea modulul **Data Analysis**.

- Dacă **Data Analysis** nu apare deși e bifat corespunzător în **Add Ins**, atunci debifați **Analysis ToolPak** și **Analysis ToolPak** în **Add Ins**, iar apoi urmați din nou procedura de bifare prezentată mai sus.

Analiza datelor

Realizarea testului Student (test t) în Excel

Primul pas este să grupați datele în funcție de ce sunteți interesați. Gruparea se realizează prin sortare. (ex. vă interesează să comparați media colesterolului unui grup de subiecți din mediu urban cu subiecți din mediul rural. Atunci sortați după coloana (variabila) de mediu).

- **Sortarea datelor.** Se selectează toate datele !!! din foaia de calcul inclusiv primul rând !!! care conține denumirea variabilelor. Din meniul **Data** se alege opțiunea **Sort**. Din fereastra de sortare există un câmp (**Sort by**) cu meniu derulant din care se alege coloana (numele variabilei) pe care doriți să o sortați (ex. mediu, dacă doriți să comparați media colesterolului în mediul urban comparativ cu cel rural). Se apasă butonul **OK**.
- **Testarea diferenței**
Pentru situația eșantioanelor dependente (perechi) – *t:Test: Paired Two Sample for Means*, sau independente (*t:Test: Two-Sample Assuming Equal Variances* sau *t:Test: Two-Sample Assuming Equal Variances*)

Din meniul **Tools** se selectează **Data Analysis**. De aici se alege unul din testele menționate în paragraful anterior. În fereastra testului se alege în primul câmp (**Variable 1 range**) datele numerice ale primului grup (fără a selecta numele variabilei), iar în al doilea rând (**Variable 2 range**) datele numerice ale celui de-al doilea grup (fără a selecta numele variabilei) (ex. pentru a media colesterolului în mediul urban comparativ cu cel rural, selectați celulele cu informația din coloana de colesterol corespunzătoare subiecților din mediul rural în primul câmp și selectați celulele cu informația din coloana de colesterol corespunzătoare subiecților din mediul urban în al doilea câmp). **Atenție nu selectați celulele din coloana cu variabila în funcție de care ați făcut sortarea !!!** (ex. de mediu). În câmpul **Hypothesized mean difference** introduceți valoarea 0. În câmpul **New worksheet ply** introduceți o denumire sugestivă și scurtă pentru testul făcut. Se apasă butonul **OK**. În foaia de calcul apărută, găsiți media celor două grupuri comparate (**Mean**), varianța acestora (**Variance**), numărul de subiecți din grup (**Observations**), și valoare lui p bilateral/bidirecțional (**P (T<=t) two-tail**). Mediile se folosesc pentru a vedea **efectul clinic** observat (se face **diferența mediilor** pentru a se **aprecia importanța CLINICĂ a diferenței**). Dacă era vorba de un test pentru grupuri independente, și varianța celor două grupuri e asemănătoare ar fi trebuit folosit testul *t:Test: Two-Sample Assuming Equal*

Variances, iar dacă varianța celor două grupuri e mult diferită ar fi trebuit să folosiți testul *t:Test: Two-Sample Assuming Unequal Variances*. Dacă ați folosit testul incorect, refaceți cu testul corect. Valoarea lui p pe bilateral vă dă **semnificația STATISTICĂ a diferenței dintre mediile grupurilor comparate**. Dacă valoarea lui p bidirecțional este de genul 3,22342E-6, înseamnă că p este egal cu $3,22342 * 10^{-6}$, adică p este egal cu 0,0000032234 (p poate avea valori doar între 0 și 1, fiind o probabilitate).

Realizarea regresiei lineare în Excel

Dacă regresia lineară e multiplă (mai mult de 2 variabile dependente) trebuie să aveți în foaia de calcul o coloană cu variabila dependentă, și un bloc compact de variabile independente. Dacă coloanele cu variabilele independente nu sunt una lângă cealaltă regresia nu funcționează. De aceea prin proceduri ce copiere de coloane, creare de coloane noi (meniu **Insert** comanda **Columns**) trebuie să vă asigurați că variabilele independente formează un bloc compact.

Din meniul **Tools** se selectează **Data Analysis**. În fereastra apărută se selectează **Regression**. În fereastra regresiei în câmpul **Input Y range** se selectează datele corespunzătoare variabilei dependente (coloana inclusiv cu numele variabilei), iar în câmpul **Input X range** se selectează datele corespunzătoare variabilei (variabilelor) dependente (coloana inclusiv cu numele variabilei, dacă sunt mai multe variabile dependente se selectează compact toate acestea (nu se utilizează tasta CTRL pentru unirea selecțiilor)). Se bifează **Labels** și **Confidence Level**, iar în câmpul **New worksheet ply** se scrie o denumire scurtă și sugestivă pentru foaia de calcul cu regresia. Se apasă butonul **OK**.

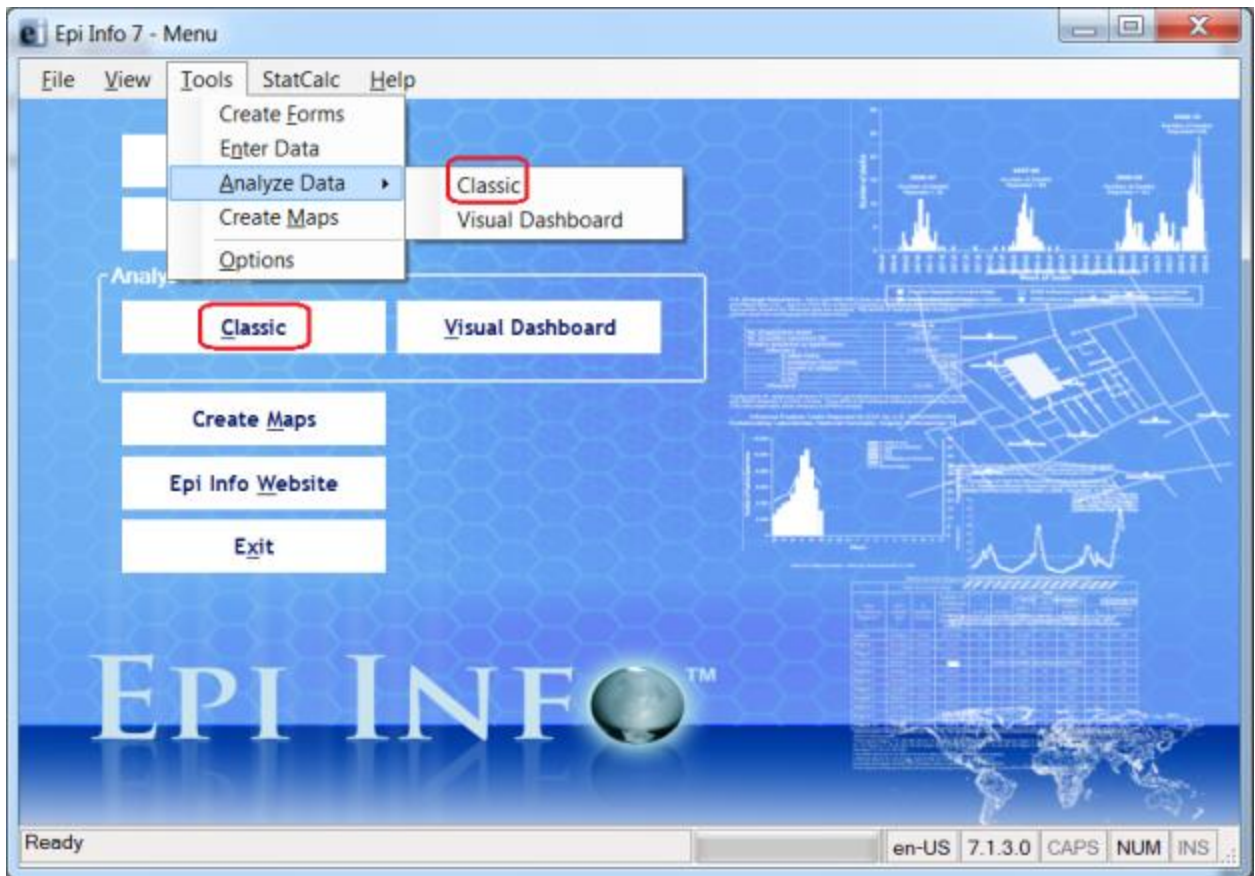
Instrucțiuni utilizare Epi Info 7

Deschiderea programului Epi Info 7

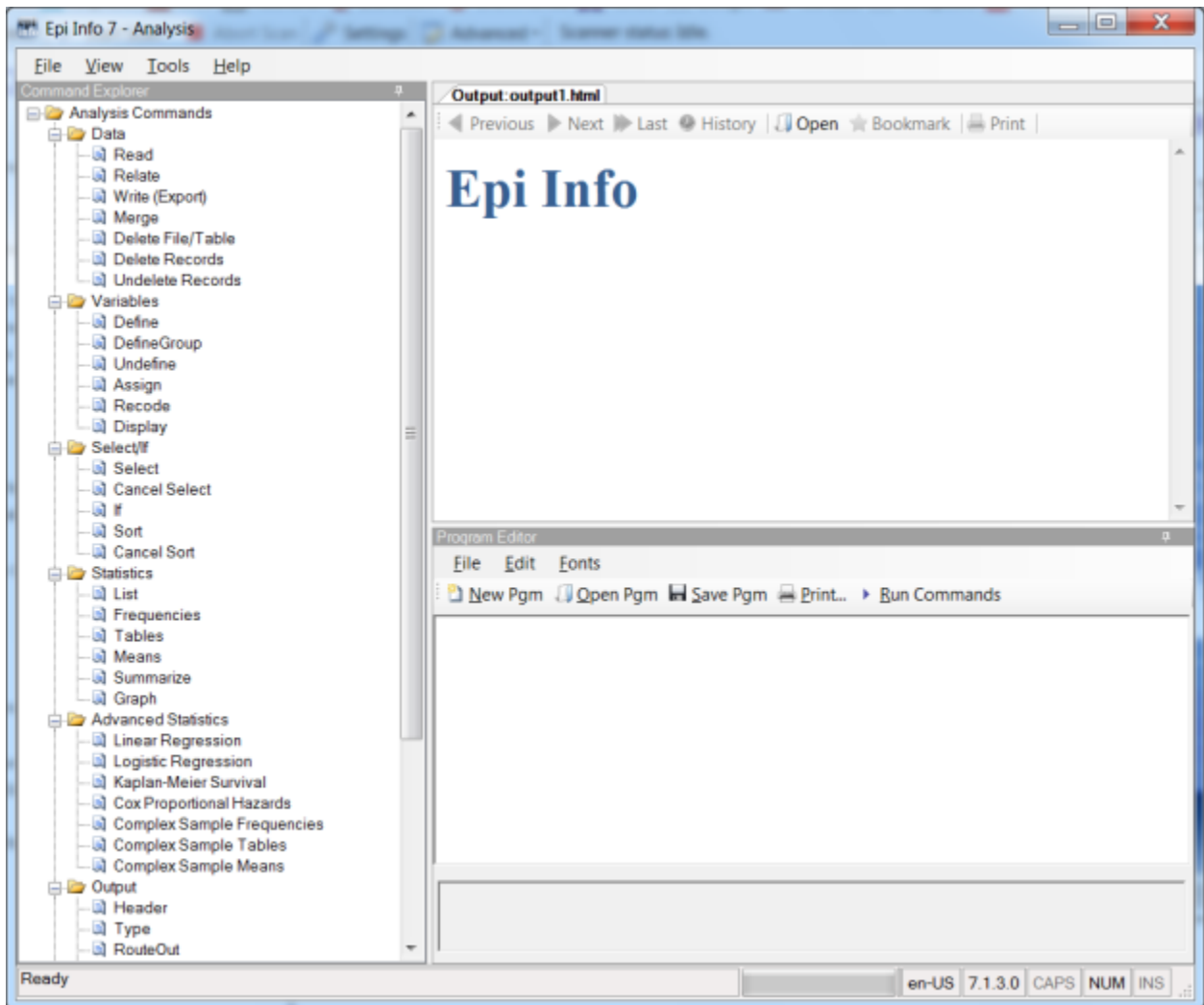
Butonul **Start**, link **Epi Info™ 7** sau **Start, All Programs, CDC, Epi Info 7, Epi Info™ 7**

Deschiderea componentei de analiză statistică

În fereastra programului, apăsați pe butonul **Classic**, în secțiunea **Analyze Data**, sau **intrați în meniul Tools, Analyze Data, Clasic**.

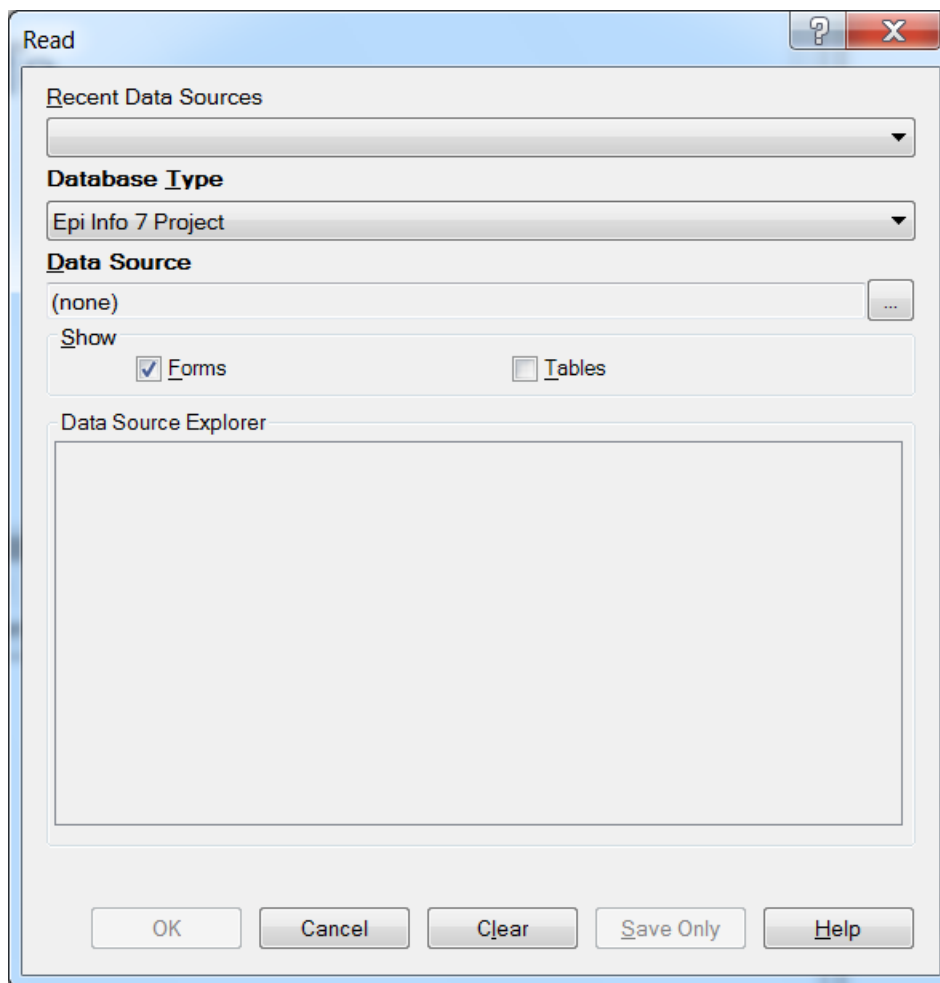


Fereastra pentru analiza datelor:

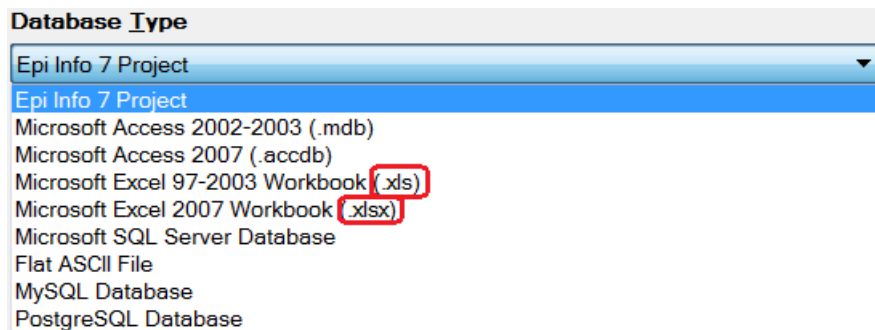




Importarea unui fișier Excel în EpiInfo7

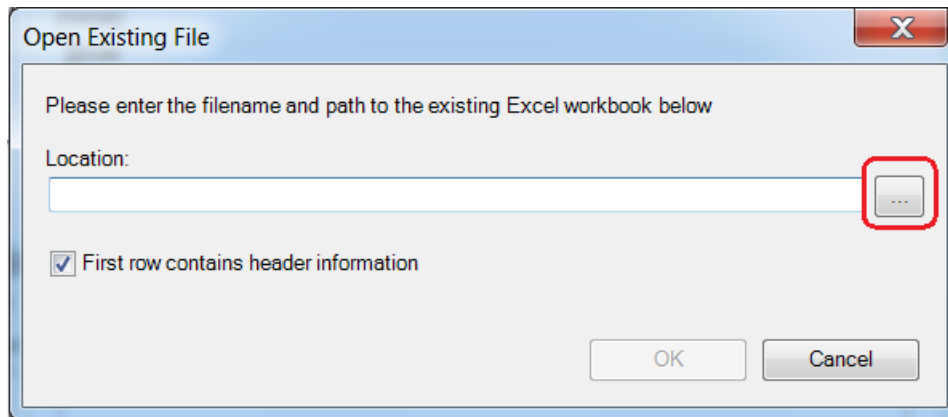
În secțiunea de comenzi (**Command Explorer** – partea stângă a ferestrei) alegem în **Analysis Commands**, secțiunea de date : **Data** - comanda **Read**. Se deschide fereastra de importare unde se alege la Database type tipul fișierului de importat precum e explicat mai jos.



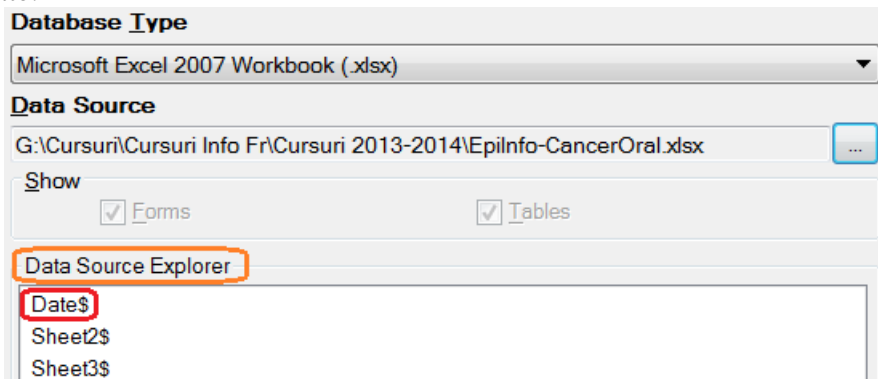
Dacă este un fișier **Excel** cu extensia **.xlsx**, se alege **Microsoft Excel 2007 Workbook (.xlsx)**, dacă este un fișier cu extensia **.xls**, se alege **Microsoft Excel 97-2003 Workbook (.xls)**.



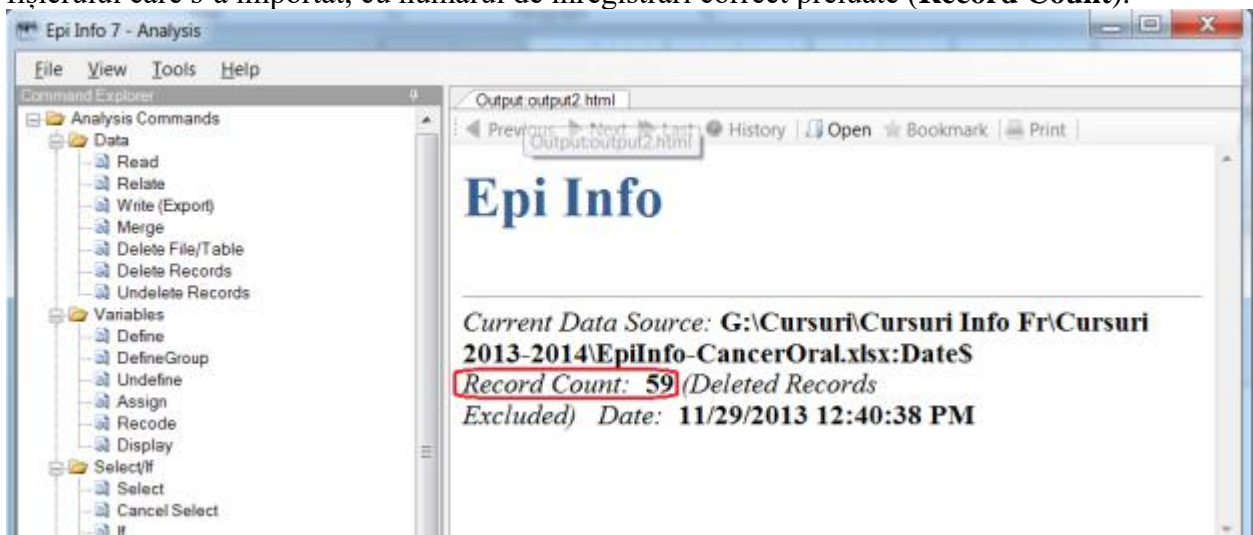
Ulterior se alege fișierul dorit pentru importare apăsând butonul  din dreptul **Data Source**. În fereastra deschisă se apasă butonul  corespunzător **Location**, urmând să se caute fișierul dorit. După selectarea lui se poate apăsa butonul **OK**, din această fereastră lăsând opțiunea **First row contains header information** bifată.



În urma acestor operațiuni Epi Info vă afișează foile de calcul (worksheets) în zona **Data Source Explorer**. Selectați foaia de calcul care conține datele pentru analiză și apăsați apoi butonul **OK** pentru a finaliza importarea. În acest exemplu se alege foaia de date Date.

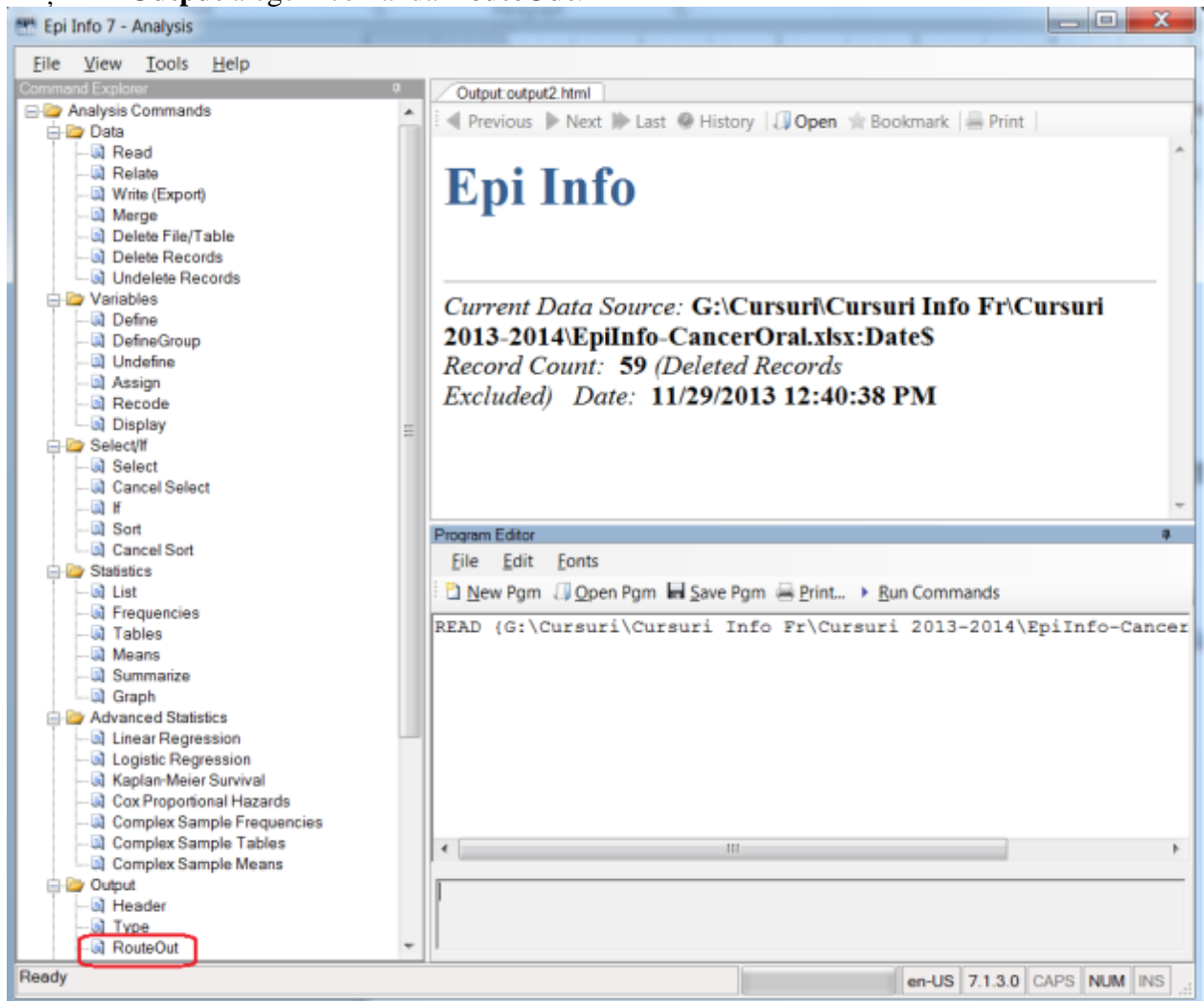



În urma importării, Epi Info afișează în zona de rezultate (**Output**), numele fișierului care s-a importat, cu numărul de înregistrări corect preluate (**Record Count**).

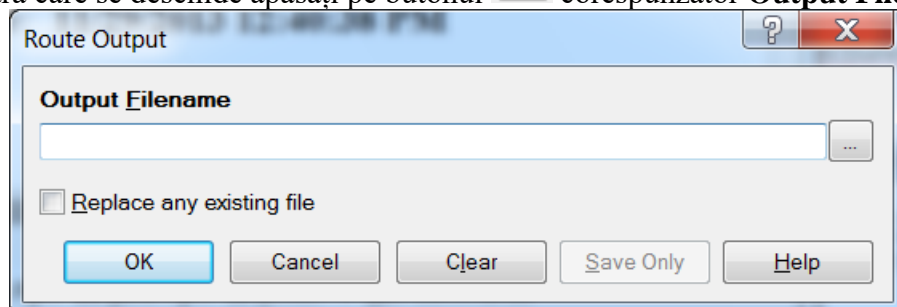


Selectarea folderului unde vor fi salvate rezultatele analizei statistice

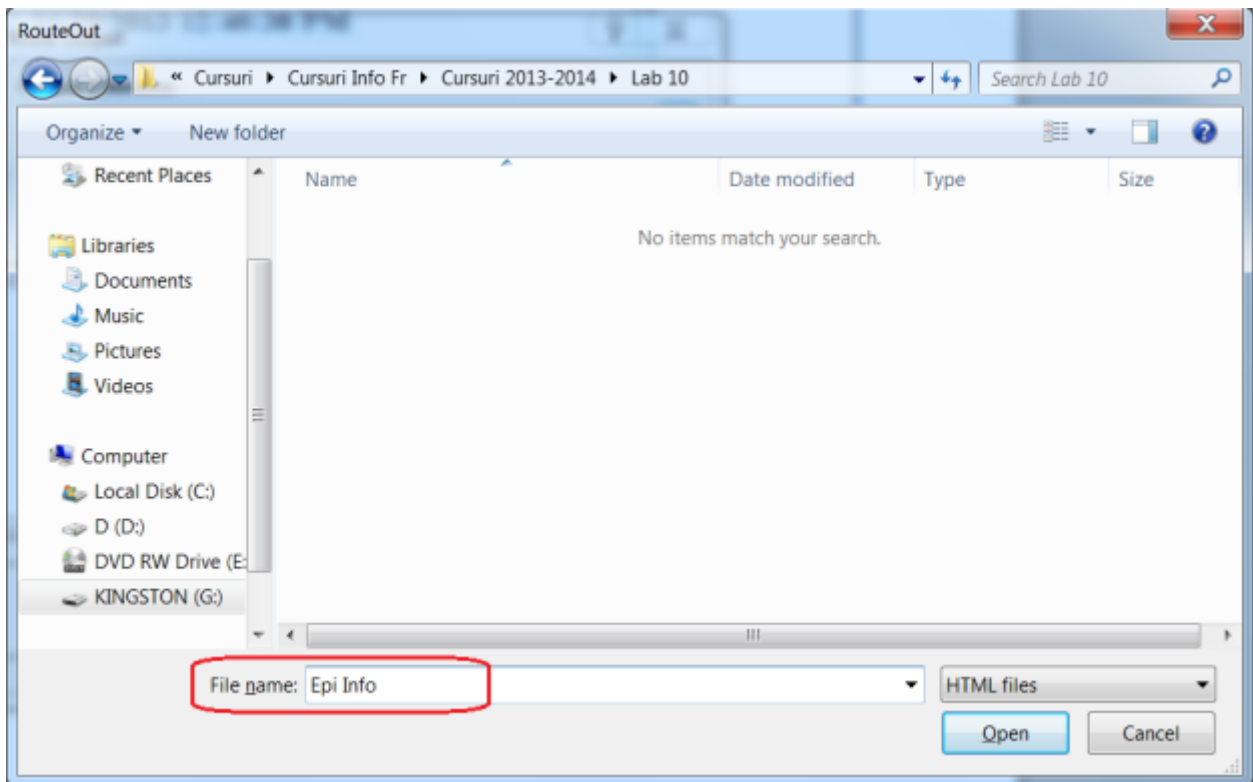
În zona de comenzi (**Command Explorer**) a ferestrei de analiză a Epi Info, secțiunea **Output** alegem comanda **RouteOut**.



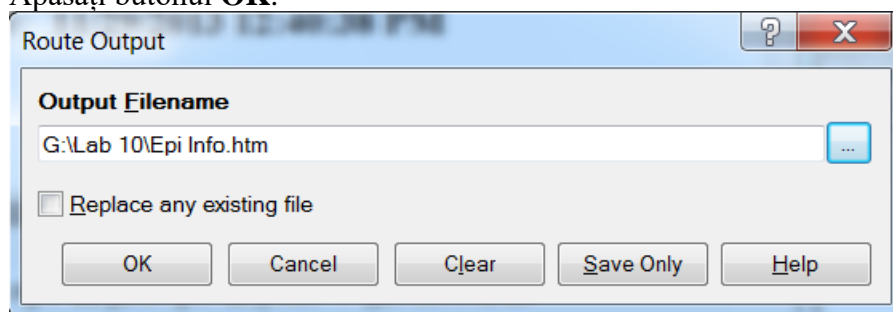
În fereastră care se deschide apăsați pe butonul  corespunzător **Output Filename**



Se deschide o nouă fereastră, unde veți căuta folderul (ex. Lab 10) unde doriți să se salveze fișierele rezultate din analiza statistică. Intrați în acel dosar iar în fereastră la **File name** tastați numele fișierului de analiză statistică (ex. **Epi Info**). Apăsați apoi butonul **Open**.



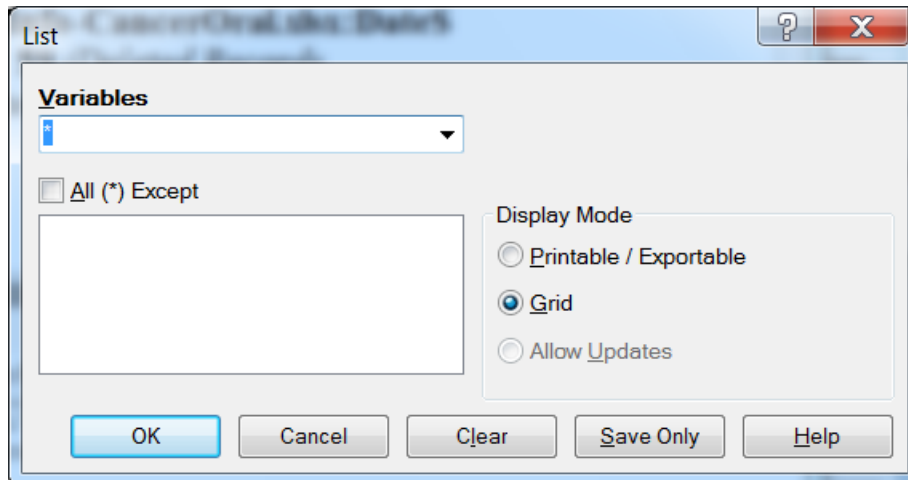
Fereastra **Route Output** va afișa calea și numele fișierului ce va conține analiza statistică. Apăsați butonul **OK**.



Afișarea datelor importate

Pentru a vedea conținutul fișierului importat, în zona de comenzi (**Command Explorer**) a ferestrei de analiză a Epi Info, secțiunea **Statistics** alegem comanda **List**.

În fereastra deschisă se lasă opțiunile implicite (de afișare a tuturor variabilelor) și se apasă butonul **OK**.



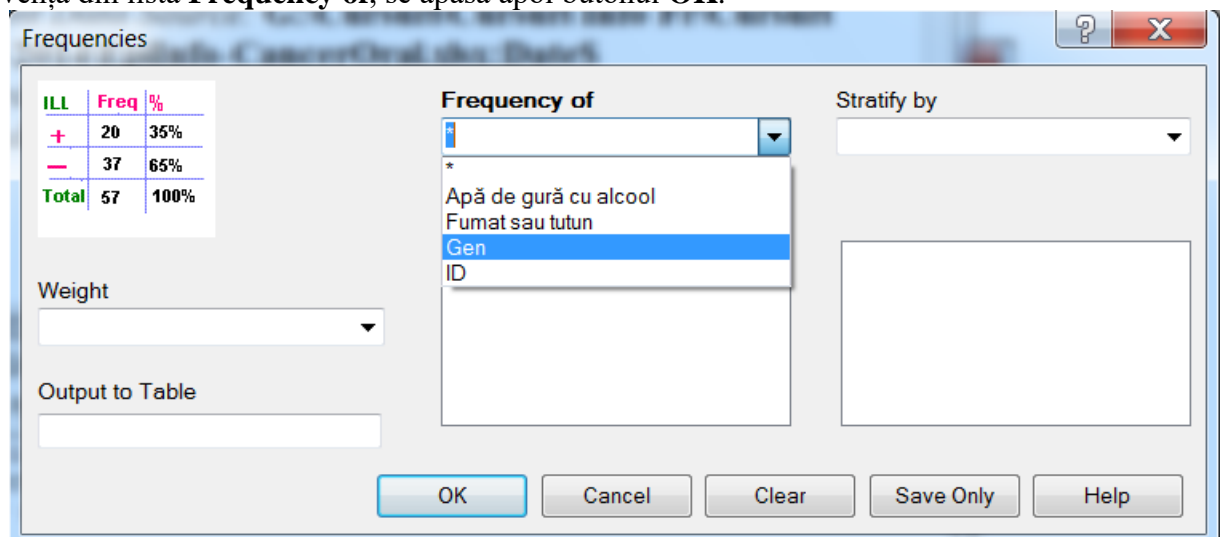
Se verifică dacă datele au fost corect importate. Apoi se poate închide această fereastră.

ID	Gen	Apă de gură cu alcool	Fumat sau tutun
1	M	Da	Da
2	M	Da	Nu
3	M	Da	Da
4	F	Nu	Nu
5	M	Da	Da
6	M	Nu	Da

Realizarea tabelului de frecvență – cu intervale de încredere




În zona de comenzi (**Command Explorer**) a ferestrei de analiză a Epi Info, secțiunea **Statistics** alegem comanda **Frequencies**.

În fereastra deschisă se alege variabila pentru care se dorește realizarea tabelului de frecvență din lista **Frequency of**, se apasă apoi butonul **OK**.



În fereastra de rezultate (**Output**), sunt afișate tabelele de frecvență precum și intervalele de încredere pentru fiecare categorie a variabilelor calitative.

FREQ Gen

GEN	Frequency	Percent	Cum. Percent	
F	14	23.73%	23.73%	
M	45	76.27%	100.00%	
Total	59	100.00%	100.00%	

95% Conf Limits

F 13.62% 36.59%

M 63.41% 86.38%

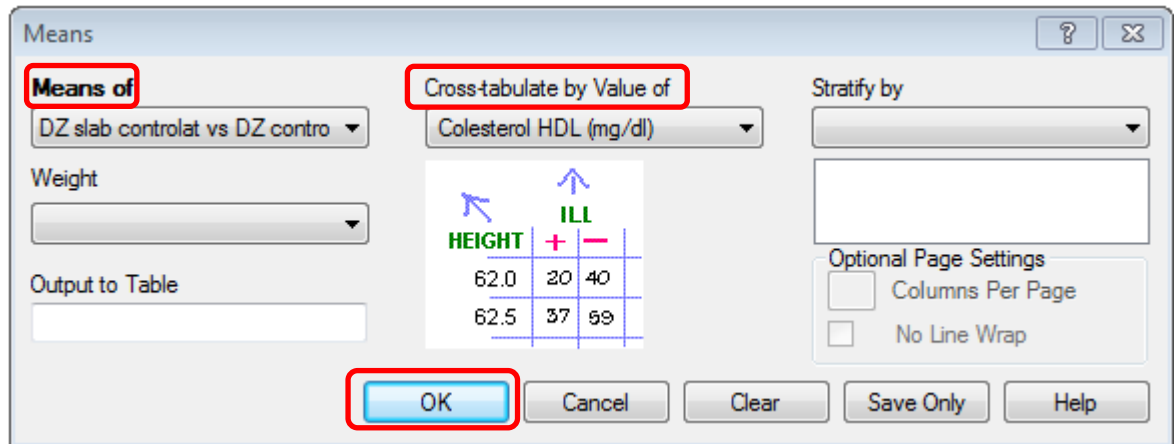
Teste statistice, indicatori, regresii

Realizarea testului X² (chi pătrat) în Epi Info și aflarea tabelului de contingență, a Riscului relativ(RR), ratei șansei(OR), diferenței de risc(RD)

Utilizați comanda **Tables** din modulul **Analysis** din **EpiInfo**. În fereastră se selectează la **exposure variable** variabila de expunere/factorul pronostic (*ex. consum de alcool*)/terapia (*ex. tratament chirurgical*)/una dintre variabile (*ex. mediu de proveniență*), respectiv la **outcome variable** se selectează variabila urmărită/boala (*ex. diabet*)/rezultatul urmărit (*ex. vindecare/ameliorare*)/cealaltă variabilă (*ex. genul*). Apoi se apasă butonul **OK**. În funcție de tipul de culegere a datelor utilizați indicatorii adecvați. Copiați atât estimatorul punctual cât și intervalul de încredere 95%. Tabelul de contingență, dacă îl copiați în alt program trebuie adaptat pentru ca informațiile pe linii și coloane să fie în pozițiile corecte (*ex. factor de expunere prezent deasupra liniei cu factorul de expunere absent, ...*). Pentru valoarea lui p alegeți ori testul **chi squared uncorrected** dacă măcar în 80% din celulele tabelului de contingență există cifre peste 5, sau testul **Fisher** dacă sub 80% din celulele tabelului de contingență există cifre peste 5.

Compararea datelor cantitative (Test student/ANOVA/ ...) în Epi Info

în **Command Explorer**, secțiunea **Statistics** alegeți comanda **Means**.



În fereastra deschisă se alege variabila cantitativă din lista de variabile din **Means of** și variabila de grupare din lista **Cross-tabulate by Value of**, apoi apăsați butonul OK.

În fereastra cu rezultate (output) se afișează:

1. **Statistica descriptivă** pe grupuri:

Descriptive Statistics for Each Value of Crosstab Variable						
	Obs	Total	Mean	Variance	Std Dev	
DZ controlat	45.0000	2502.0000	55.6000	112.4727	10.6053	
DZ slab controlat	37.0000	1932.0000	52.2162	77.8408	8.8227	
	Minimum	25%	Median	75%	Maximum	Mode
DZ controlat	39.0000	48.0000	55.0000	62.0000	80.0000	54.0000
DZ slab controlat	36.0000	46.0000	52.0000	58.5000	76.0000	41.0000

1. **Rezultatele testului t** (Student) ce compară mediile pentru eșantioane independente – se afișează două variante – cu varianțe egale (**Pooled**) sau inegale (**Unequal**). Se afișează diferența între mediile comparate cu interval de încredere și valoarea lui p.

T-Test

	Method	Mean	95% CL Mean	Std Dev
Diff (Group 1 - Group 2)	Pooled	3.3838	-0.9633 7.7309	9.8432
Diff (Group 1 - Group 2)	Satterthwaite	3.3838	-0.8867 7.6543	

Method	Variances	DF	t Value	Pr > t
Pooled	Equal	80	1.55	0.1253
Satterthwaite	Unequal	79.98	1.58	0.1188

2. **Rezultatele testului Anova** pentru compararea mediilor a două sau mai multe grupuri independente cu varianțe egale:

ANOVA, a Parametric Test for Inequality of Population Means

(For normally distributed data only)

Variation	SS	df	MS	F statistic
Between	232.49071	1	232.49071	2.39957
Within	7751.07027	80	96.88838	
Total	7983.56098	81		

P-value = 0.12532

3. Rezultatele testului Bartlett ce compară varianțele a două grupuri independente:

Bartlett's Test for Inequality of Population Variances

Bartlett's chi square= 1.30103 df=1 P value=0.25403

A small p-value (e.g., less than 0.05) suggests that the variances are not homogeneous and that the ANOVA may not be appropriate.

4. Rezultatele testelor neparametrice pentru compararea a două eșantioane independente (**Mann-Whitney test / Wilcoxon Two-Sample**) sau mai mult de două eșantioane independente (**Kruskal-Wallis test**):

Mann-Whitney/Wilcoxon Two-Sample Test (Kruskal-Wallis test for two groups)

Kruskal-Wallis H (equivalent to Chi square) = 1.4703

Degrees of freedom = 1

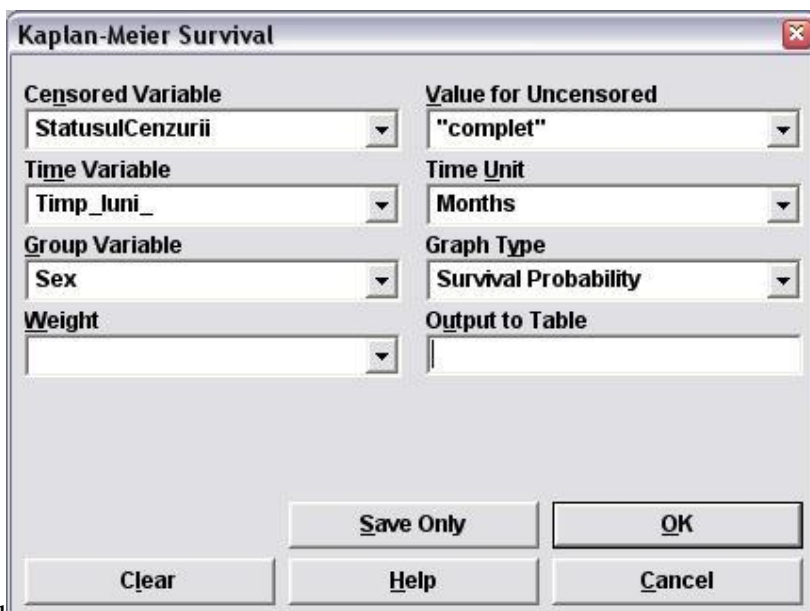
P value = 0.2253

Realizarea testului log-rank pentru compararea supraviețuirii a două grupuri, în EpiInfo

Operațiunea e identică cu cea a descrierii supraviețuirii pentru două grupuri. În **Epi-Info** modulul **Analysis** secțiunea **Advanced Statistics**, se selectează **Kaplan-Meyer Survival**. În fereastra apărută se completează conform imaginii:

Se vor obține curbele de supraviețuire într-un grafic, iar dedesubt vor fi afișate două teste. Unul din ele este testul log-rank.

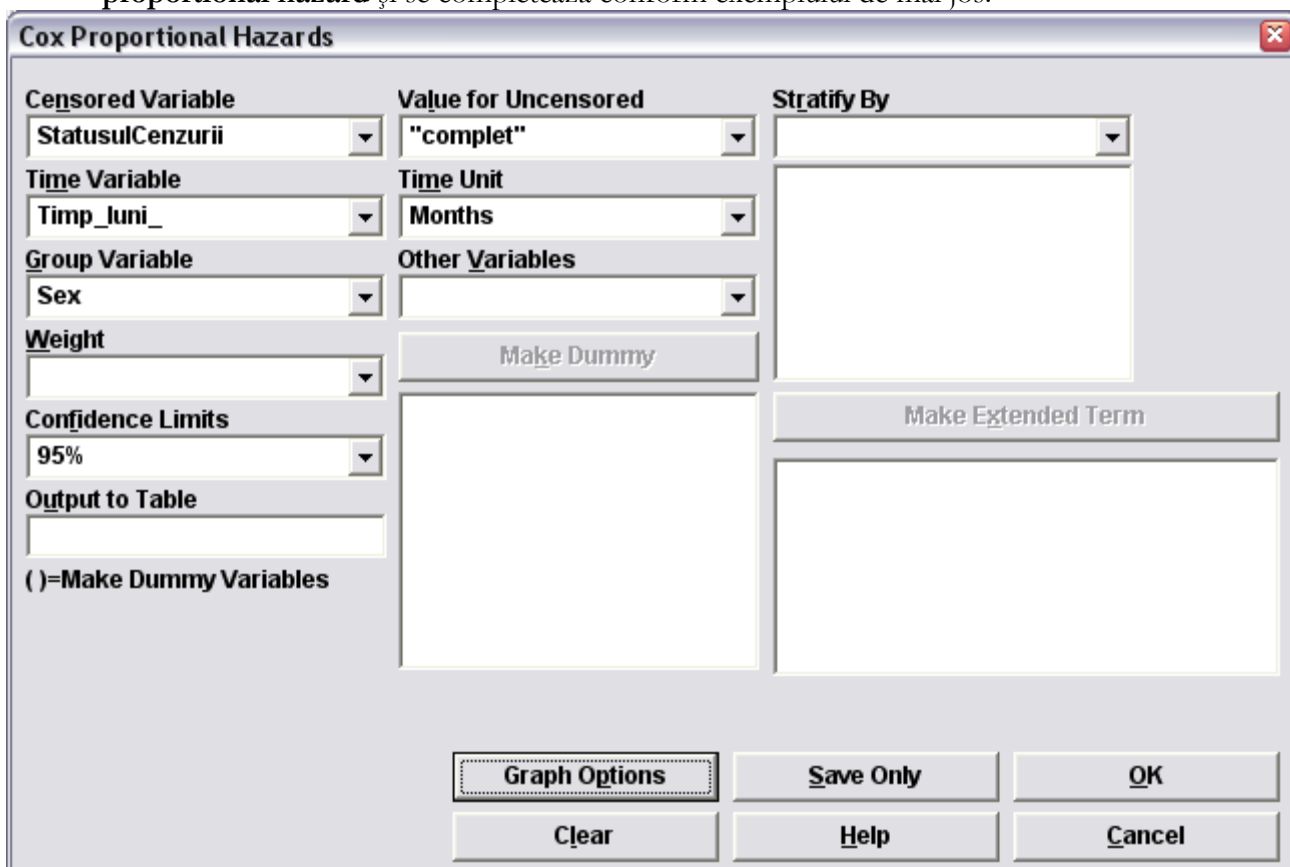
Dacă se dorește să se compare în funcție de o altă variabilă la Group Variable se selectează variabila de interes



(ex. Mediu, Intervenție, Stadiu ...).

Aflarea ratei hazardului prin efectuarea regresiei Cox (identificarea predictorilor în analiza datelor de supraviețuire), în EpiInfo

În Epi-Info modulul **Analysis** secțiunea **Advanced Statistics**, se selectează **Cox proportional hazard** și se completează conform exemplului de mai jos:



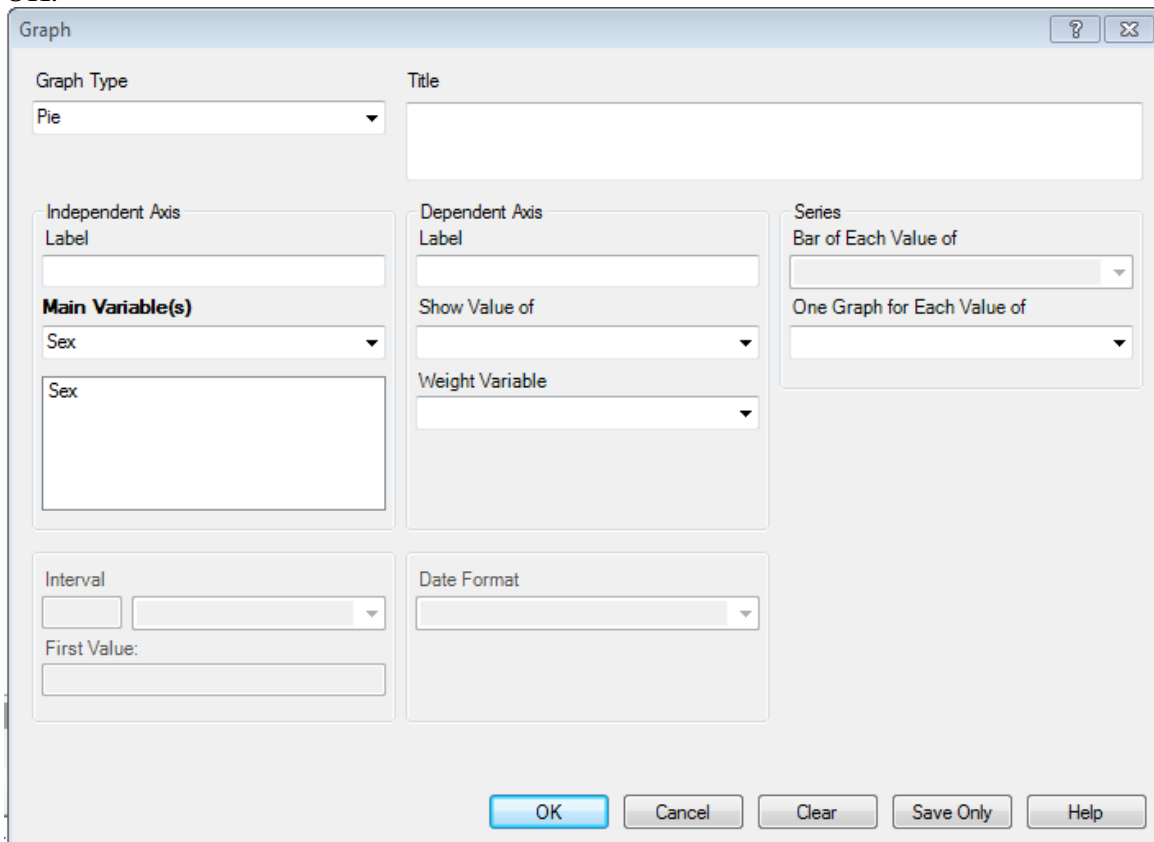
Rata hazardului este afișată dedesubtul graficului ajustat al supraviețuirii împreună cu **intervalul de încredere** asociat, precum și cu **semnificația modelului Cox** creat.

Dacă se dorește să se rata hazardului pentru o altă variabilă, la **Group Variable** se selectează variabila de interes (ex. Mediu, Terapie, Stadiu, Histologie ...).

Realizare grafice în Epilinfo 7

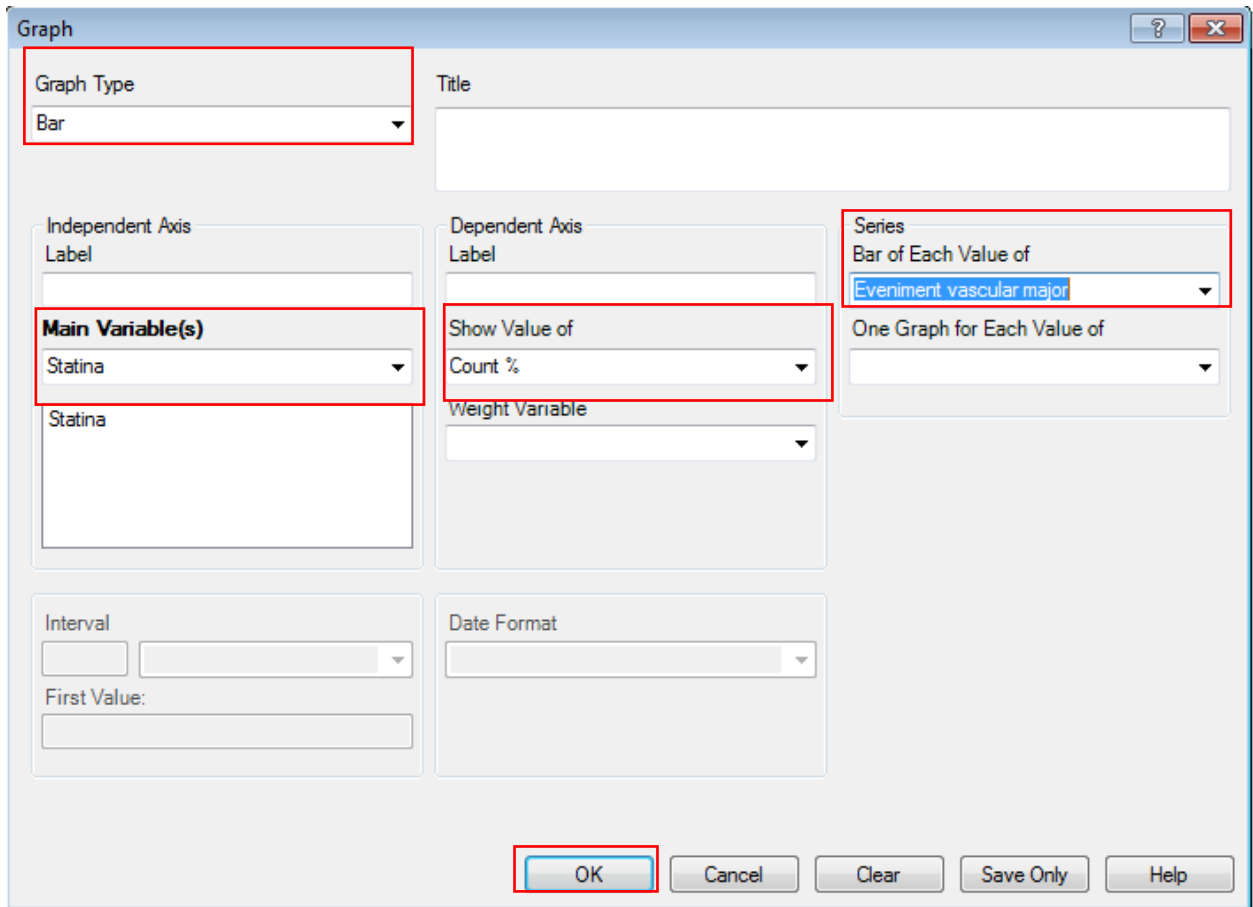
Grafice pentru descrierea unei variabile calitative (Pie, Bar, Column)

În secțiunea **Statistics** comanda **Graph**. La **Graph type** se alege tipul graficului (Pie, bar, column). La **Main Variable(s)** se alege variabila pentru care se dorește și se apasă OK:

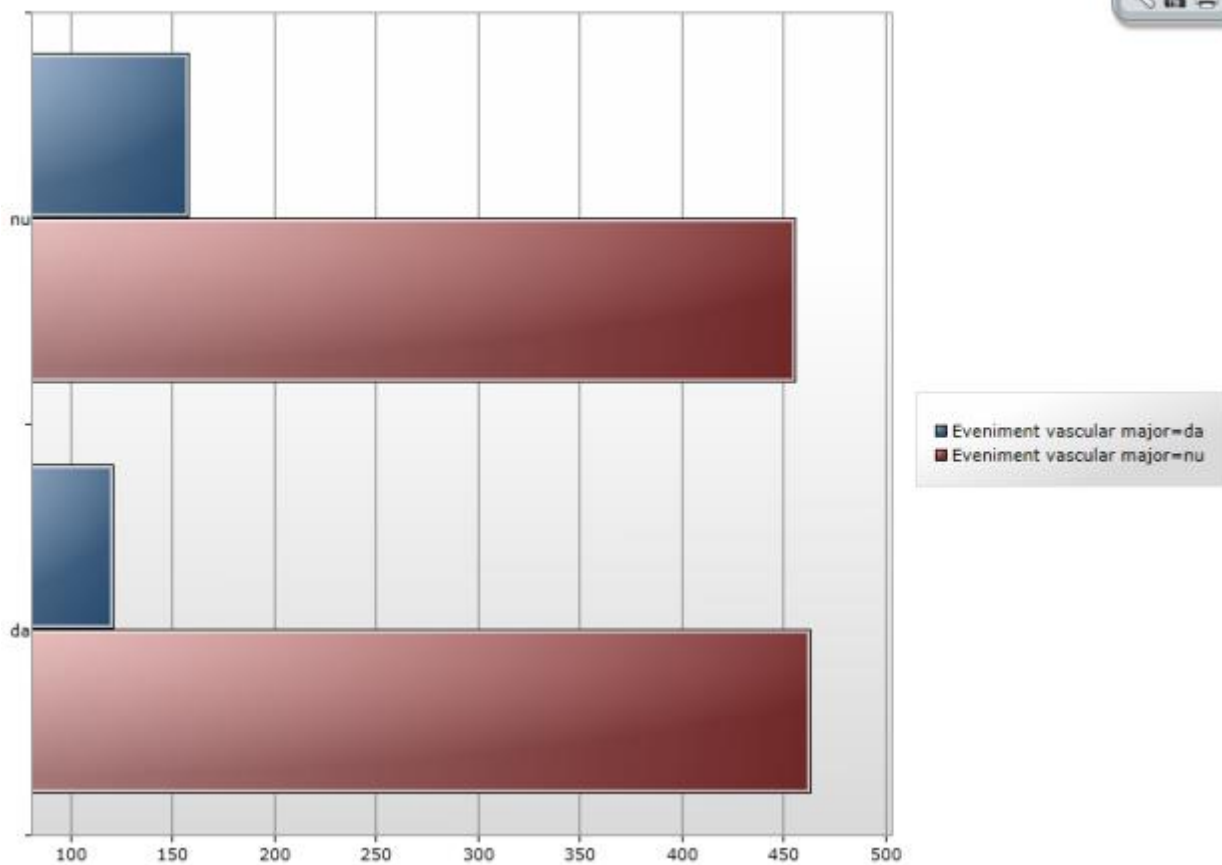


Grafic pentru a arăta relația între două variabile calitative (Bar/Column)

În secțiunea **Statistics** comanda **Graph**. La **Graph type** se alege tipul graficului (bar/ column). La **Main Variable(s)** se alege variabila de expunere (pentru studii de factori de risc), sau factorul de grupare, la **Bar of Each Value of** se alege boala (pentru studiile de tip factori de risc), sau obiectivul urmărit (complicație, vindecare, deces, ...), iar la **Show Value of** se alege **Count %** - pentru un grafic care arată procentele sau **Count** pentru un grafic care arată numărul de subiecți și se apasă OK. Pe unele stații în rețea nu merg graficele realizate cu Count %, atunci selectați Count.

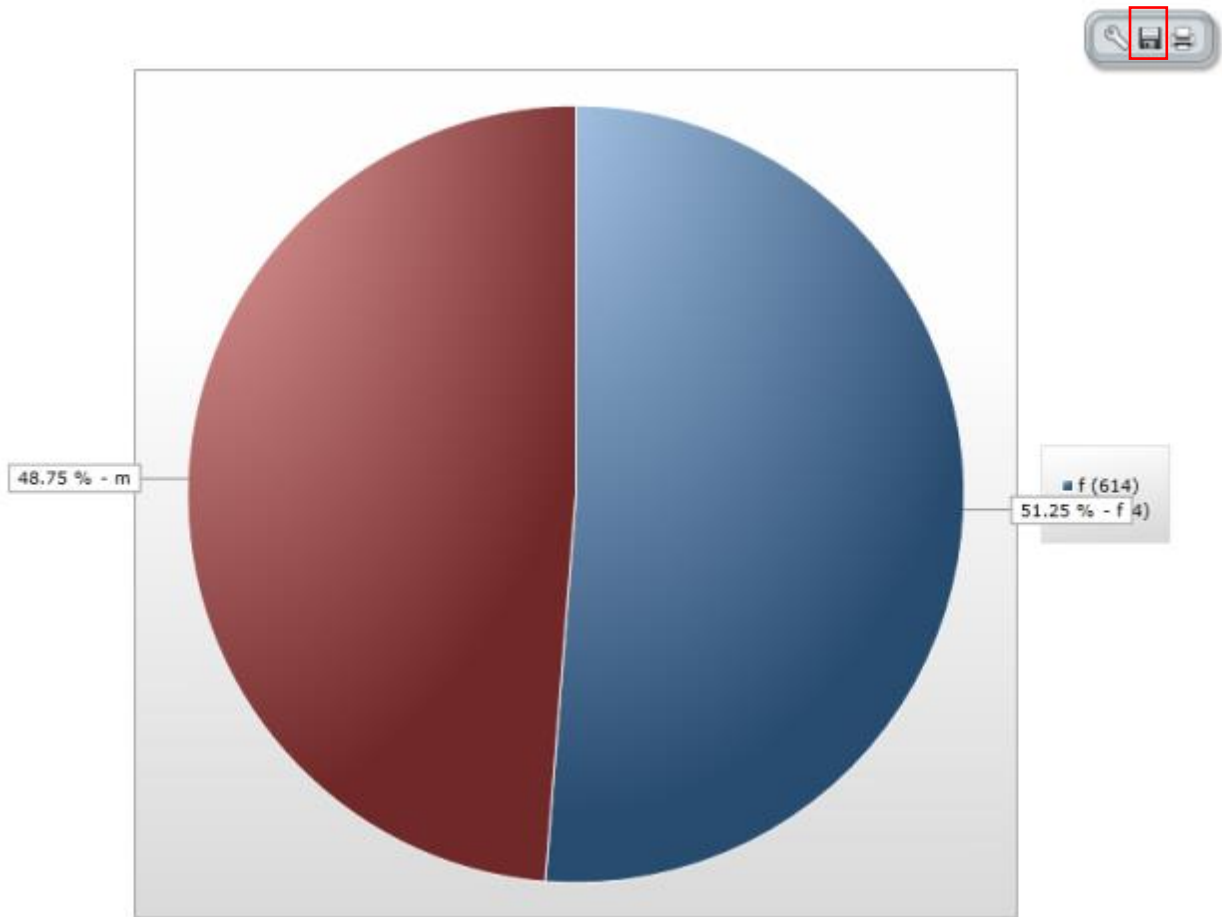


După aceea trebuie editat graficul pentru a trece definițiile axelor. (vezi capitolul de editare)



Salvarea graficelor

După ce ați făcut un grafic, întrucât e realizat în Silverlight nu se poate copia direct. În dreapta sus este o bară cu 3 butoane. Se apasă butonul cu imaginea unei diskete. Se scrie numele fișierului și se alege folderul unde se dorește a fi salvat graficul.



Editarea graficelor (titlu, axe)



În dreapta sus este o bară cu 3 butoane.   Se apasă butonul cu imaginea unei chei. Apare o fereastră în care se pot trece definițiile axelor X și Y. Titlul de obicei se trece în Word/Powerpoint nu direct pe grafic.

Chart Configuration ✕

Chart Title:

Legend Title:

X-Axis Label:

Y-Axis Label:

Copyright Daniel-Corneliu Leucuța