SONDAJE SEMINAR 13.12.2024

Compararea mediilor din 2 esantione identitice

Esantion N impartit in 2 esantioane dupa dif criterii: rural;urban, taxa/buget, varsta sub 35/varsta peste 35.

Pt fiecare subgrupa pot calcula; medie, disperie – acestia sunt estimatori pentru parametrii din populatie

In urma calculelor putem observa ca excista diferente intre cele 2 grupe. Trebuie sa vedem daca aceste diferente sub semnificative.

In ce situatii facem testarea?

Pas1:

Esantion trebuie sa fie independent, aleator, HOMOSCEDASCITATE(ADICA H0: dispera 1=disperia 2 ; H1: dispersia 1 NU e egal cu dispersia 2)

Daca se accepta H0 se aplica un test student(test t) conform relatiei:

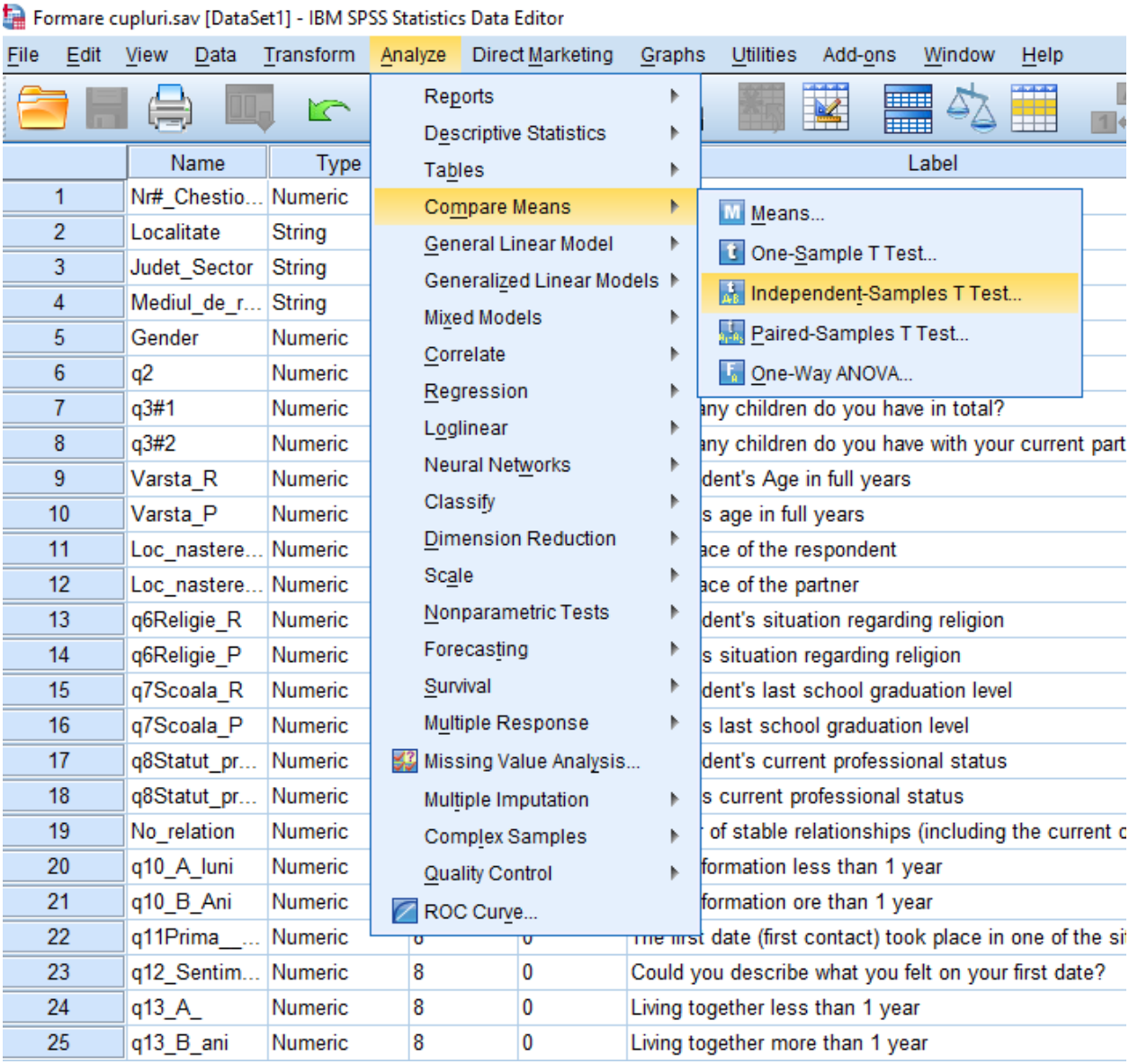
Daca se accepta H1(adica dispersiile nu sunt egale) se aplica

Pas2:

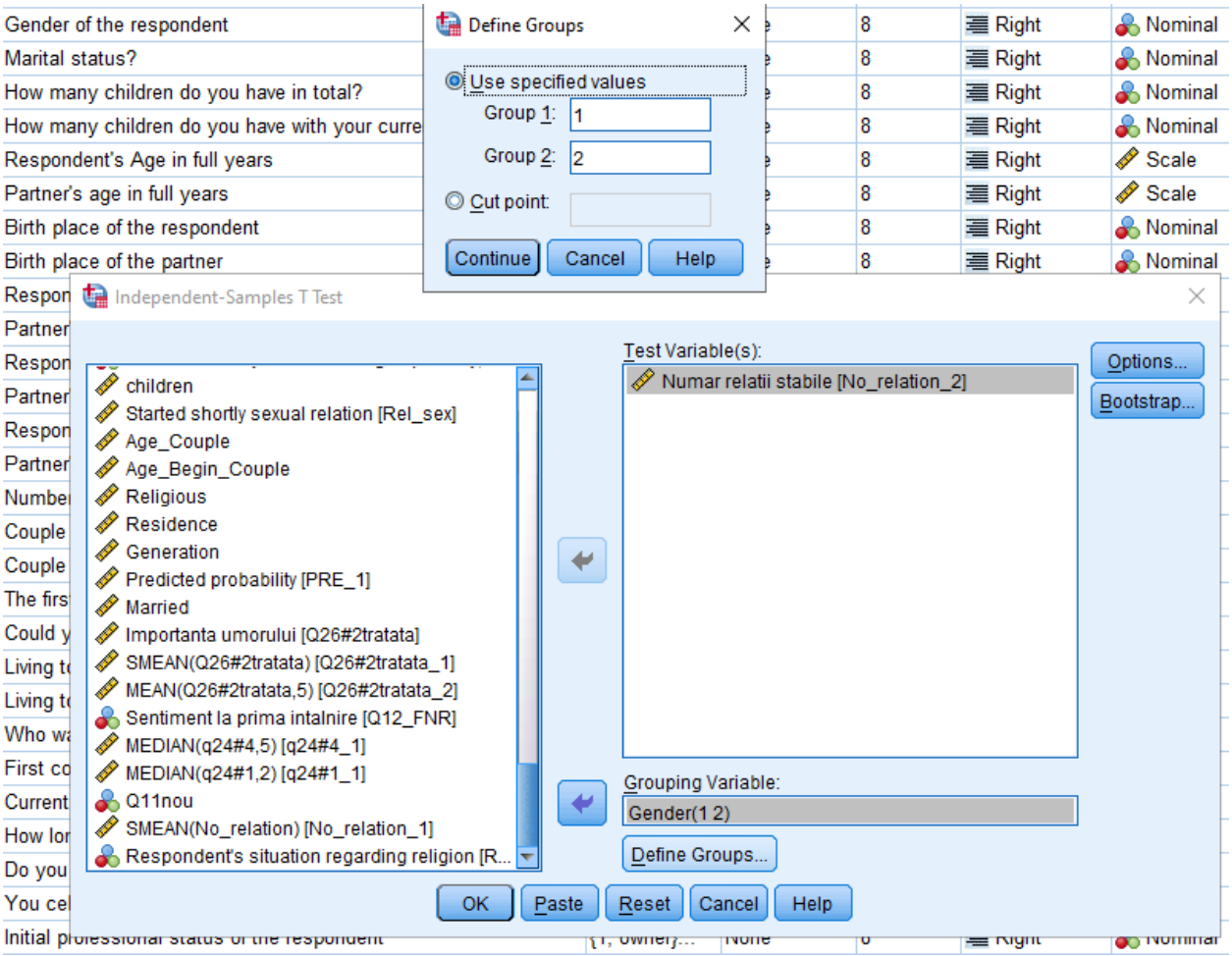
H0: media 1=media 2

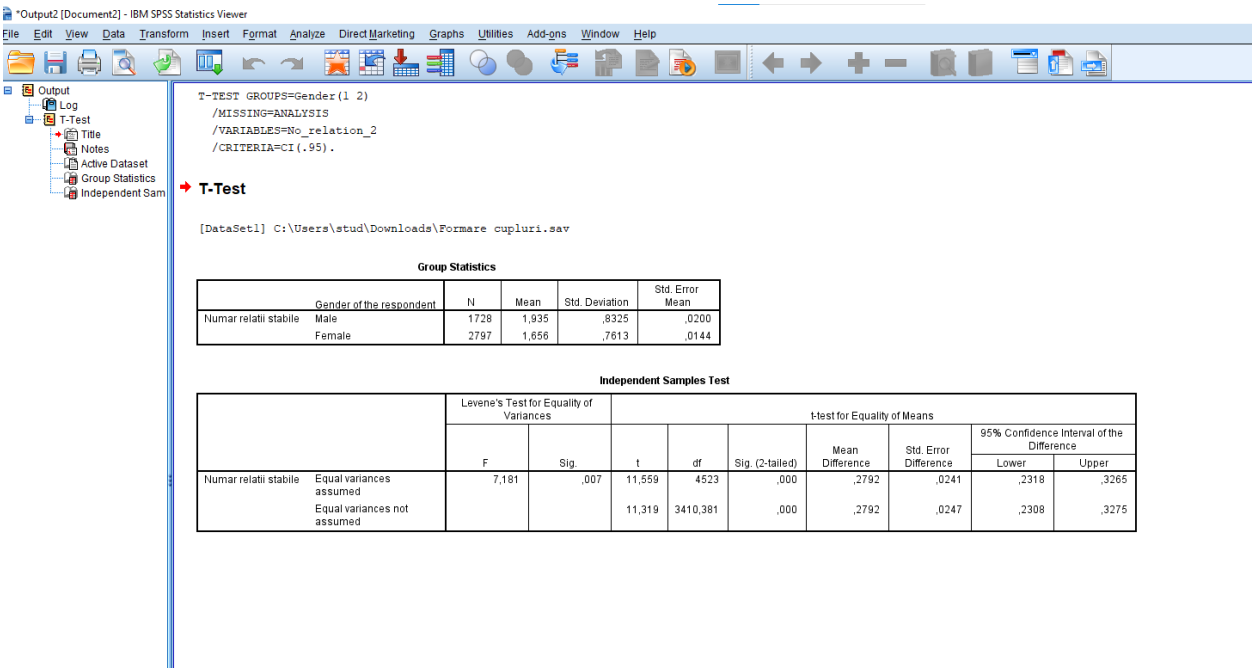
H1: media 1 NU e egal cu media 2

Pentru a exemplifica vrem sa vedem daca exista dif semnificative intre nr de relatii stabile si variabila gen.



Independent Samples T test





Group Statistics

	Gender of the respondent	N	Mean	Std. Deviation	Std. Error Mean
Numar relatii stabile	Male	1728	1,935	,8325	,0200
	Female	2797	1,656	,7613	,0144

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Numar relatii stabile	Equal variances assumed	7,181	,007	11,559	4523	,000	,2792	,0241	,2318	,3265
	Equal variances not assumed			11,319	3410,381	,000	,2792	,0247	,2308	,3275

Group Statistics

	Gender of the respondent	N	Mean	Std. Deviation	Std. Error Mean
Numar relatii stabile	Male	1728	1,935	,8325	,0200
	Female	2797	1,656	,7613	,0144

Deci ins ca barbatii au un nr de relatii stabile de 1.9

Deci x1=1.9 si x2=1,65

Deci am putea spune ca pers de gen masculin au mai multe relatii stabile. Verificam daca dif e semnificativa

In tabelul de mai jos avem rezultatele testului

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Numar relatii stabile	Equal variances assumed	7,181	,007	11,559	4523	,000	,2792	,0241	,2318	,3265
	Equal variances not assumed			11,319	3410,381	,000	,2792	,0247	,2308	,3275

F=7,2

Nivelul de semnificatie=0,007 < 0,05 deci se accepta ipoteza alternative(adica H1: dispersia 1 NU e egal cu dispersia 2)

Conditia de homoscen nu e indeplinita deci aplicam conditia cu ceva grade de libertate. Aceastya se gaseste pe linia 2 in tabelul 2(chinar cu rosu)

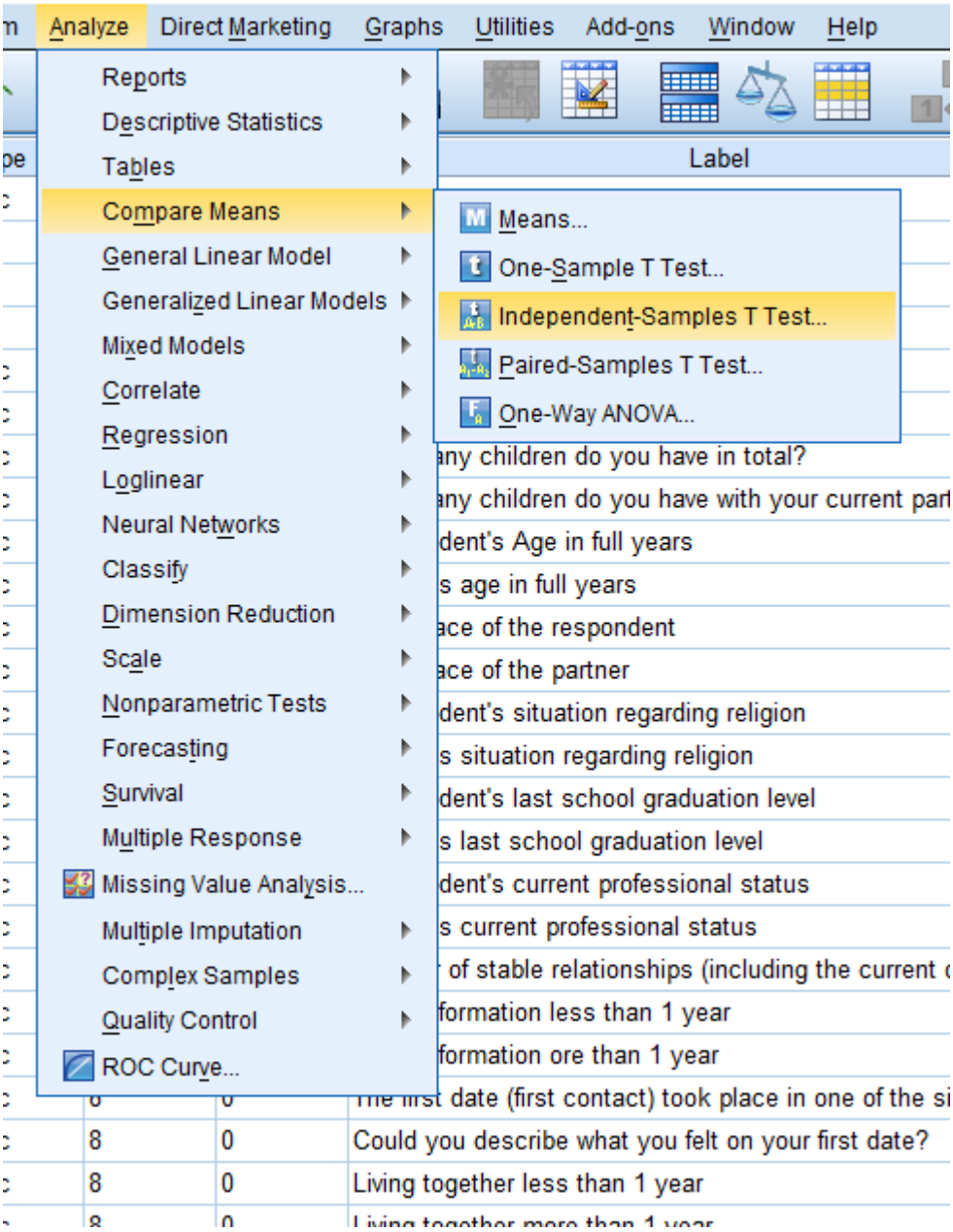
Deci ignoram testul de pe prima linie si ne interseaza doar cel de pe linia 2

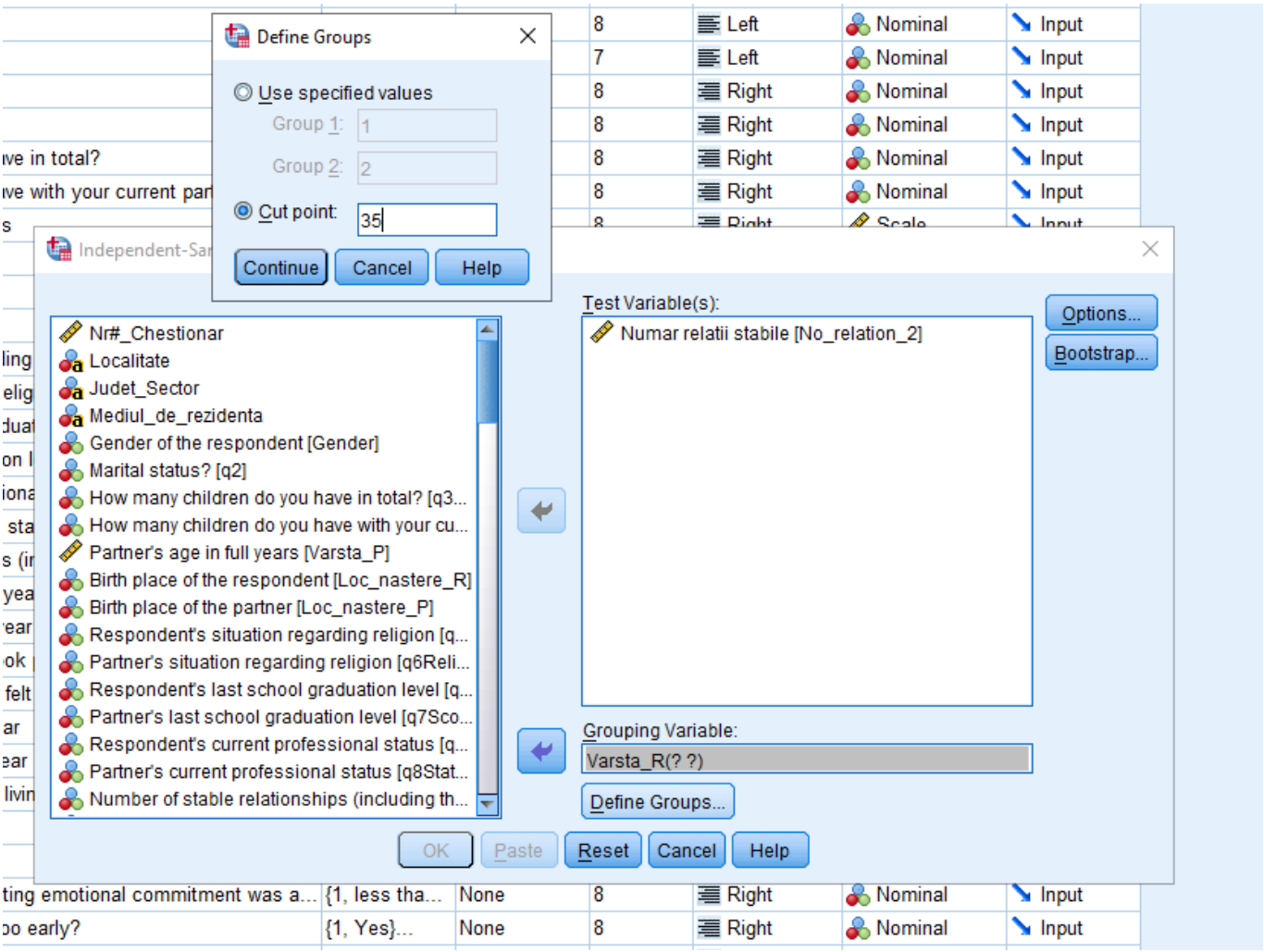
Sig (2-tailed) = 0,00 < 0,05 deci acceptam H1: media 1 difera de media 2

Garantam ca cele 2 medii sunt dif cu o prob de 99,9% (pt ca sig(2-tailed) = 0,00)

ALT EXEMPLU

Nr de realtii stabile in functie de varsta


Analyze
Direct Marketing
Graphs
Utilities
Add-ons
Window
Help
Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Quality Control
ROC Curve...
Means...
One-Sample T Test...
Independent-Samples T Test...
Paired-Samples T Test...
One-Way ANOVA...
Label
any children do you have in total?
any children do you have with your current part
dent's Age in full years
s age in full years
ace of the respondent
ace of the partner
dent's situation regarding religion
s situation regarding religion
dent's last school graduation level
s last school graduation level
dent's current professional status
s current professional status
of stable relationships (including the current
formation less than 1 year
formation ore than 1 year
8 0 Could you describe what you felt on your first date?
8 0 Living together less than 1 year



La varsta nu mai putem face cu 0,1 cu punem cut point , deci sub 35 si peste 35 de ani.

T-Test

[DataSet1] C:\Users\stud\Downloads\Formare cupluri.sav

Group Statistics

	Respondent's Age in full years	N	Mean	Std. Deviation	Std. Error Mean
Numar relatii stabile	>= 35	1496	1,589	,7796	,0202
	< 35	3029	1,848	,7973	,0145

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Numar relatii stabile	Equal variances assumed	,723	,395	-10,358	4523	,000	-,2591	,0250	-,3081	-,2100
	Equal variances not assumed			-10,437	3038,405	,000	-,2591	,0248	-,3077	-,2104

Group Statistics

	Respondent's Age in full years	N	Mean	Std. Deviation	Std. Error Mean
Numar relatii stabile	>= 35	1496	1,589	,7796	,0202
	< 35	3029	1,848	,7973	,0145

Cei pana in 35 de ani(care sunt si mai multi in esnation) au un nr de 1,84 nr de realtii stabile

Si cei peste 35 de ani au un nr de 1,58

**Sondajul a fost facut acum 15 ani, arata ca tinerii si-au schimbat comportamentul in timp

Este semnficativa diferenta?

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Numar relatii stabile	Equal variances assumed	,723	,395	-10,358	4523	,000	-,2591	,0250	-,3081	-,2100
	Equal variances not assumed			-10,437	3038,405	,000	-,2591	,0248	-,3077	-,2104

Sig = 0,395 > 0,05 deci dispersiile sunt egale9se accepta H0) si se respecta homo

Ne uitam la prima linie

Garantam cu sig(2 tailed) adica 99,9% ca tinerii pana in 35 de ani au mai multe realtii stabile.

Doua esantioane diferite

Ce ins esantion dependent? (se mai numeste si panel) – ACESTEA SE POT FOLOSI SI PENTRU OBSERVATII PERECHI

OBSERVATII PERECHI = acelasi esantion – variabila 1 cu variabila 2 sa fie comparabile(ex: nota 1-10 rel professor si nota 1-10 rel personal facultate/ varsta la care ar trebui sa se casatoreasca un barbat si varsta la care ar trebui sa se casatoreasca o femeie)

Am un esantion n iar la momentul de timp t1 aplic un chestionar

Acelassi esantion va primi acelasi chestionar la momentul t2

Pot sa aplic estimatorii media la moment t1 si media la moment t2

Se va calc diferenta de i(Di)

Di= x2i-x1i

D estimator pt diferenta D parametru in populatie

Ipoteze test:

H0: media 1=media 2 echivalent cu diferenta medie D =0

H1: media 1!=media 2 echivalent cu diferenta medie D !=0

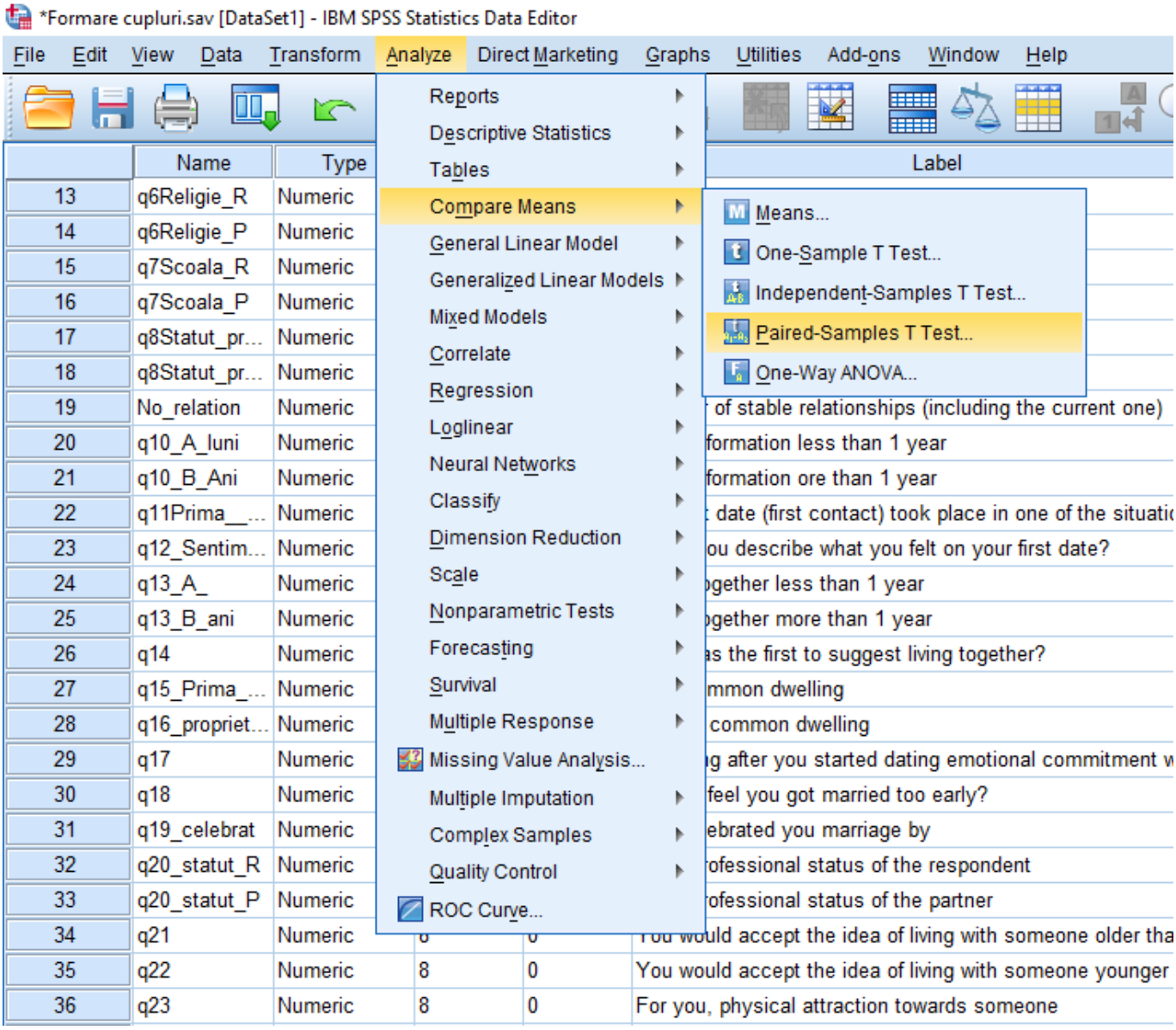
Ceva formula pt t=x2 mediu – x1 mediu / rad din(sd^2/ n-1)

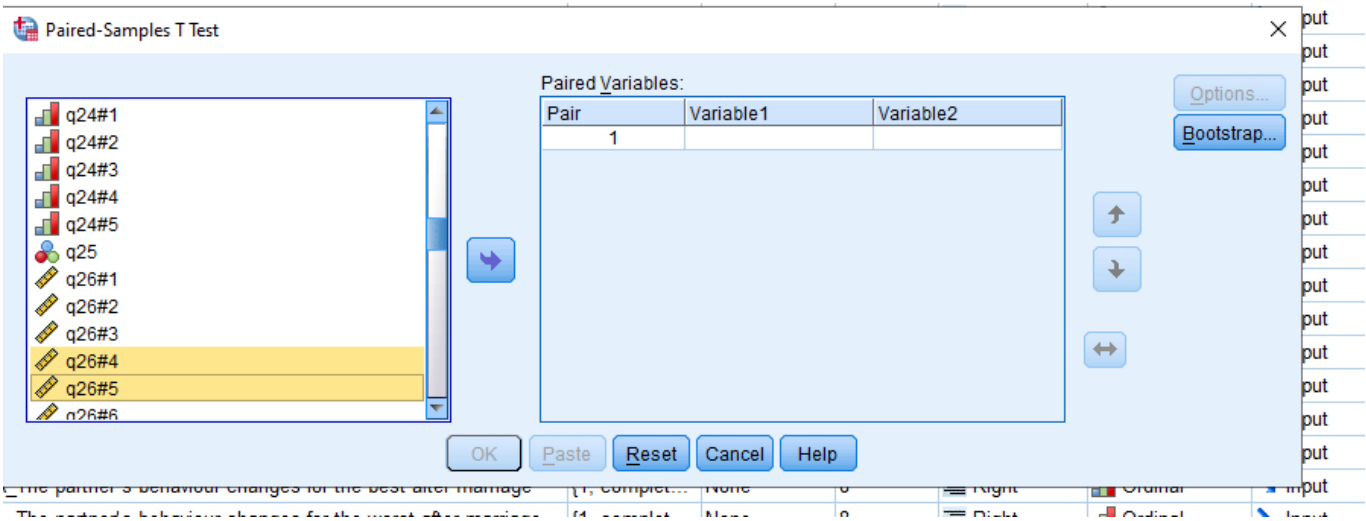
Df=n-1

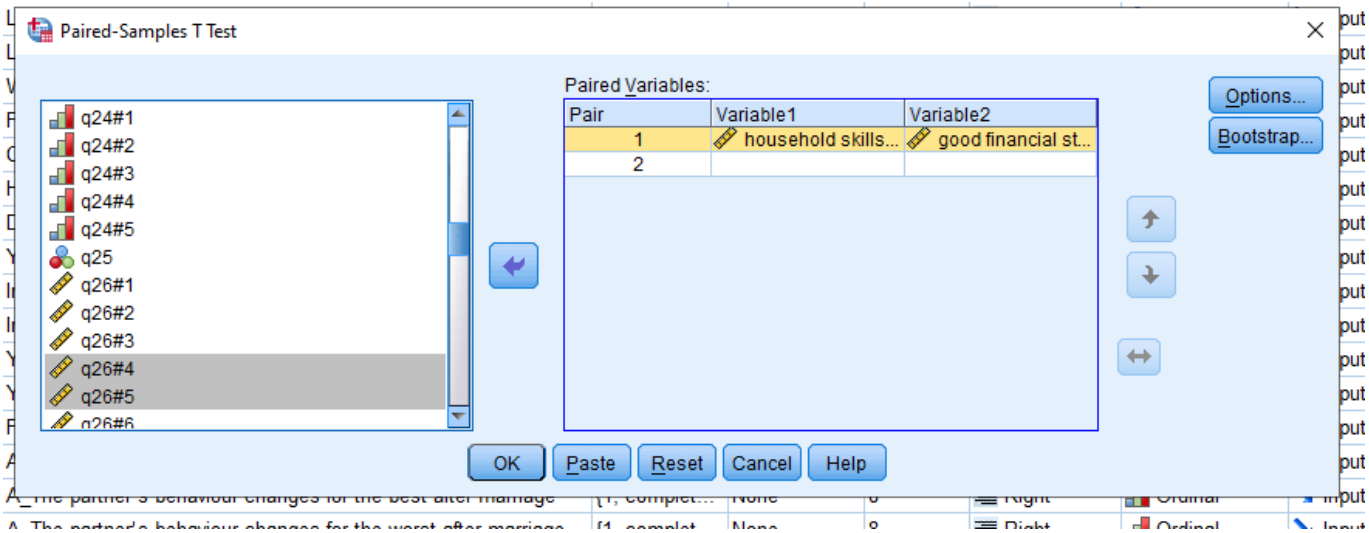
Aplicam si noi acum acest test:

La acestea s-au oferit note:

41	q24#5	Numeric	8	0	A_To get married means to enter a too long term commitment	{1, Acord to...	None	8	Right	Ordinal	Input
42	q25	Numeric	8	0	If you had a daughter or a son at the age of marriage, would yo...	{1, yes, esp...	None	8	Right	Nominal	Input
43	q26#1	Numeric	8	0	appearance	{0, Cel mai ...	None	8	Right	Scale	Input
44	q26#2	Numeric	8	0	humour	{0, Cel mai ...	None	8	Right	Scale	Input
45	q26#3	Numeric	8	0	intelligence	{0, Cel mai ...	None	8	Right	Scale	Input
46	q26#4	Numeric	8	0	household skills	{0, Cel mai ...	None	8	Right	Scale	Input
47	q26#5	Numeric	8	0	good financial status	{0, Cel mai ...	None	8	Right	Scale	Input
48	q26#6	Numeric	8	0	fidelity	{0, Cel mai ...	None	8	Right	Scale	Input
49	q27	Numeric	8	0	If you decide to get married, you intend to have	{1, Numai c...	None	8	Right	Nominal	Input
50	q28#1	Numeric	8	0	the birth of a child	{0, Cel mai ...	None	8	Right	Nominal	Input


*Formare cupluri.sav [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help
Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Quality Control
ROC Curve...
Means...
One-Sample T Test...
Independent-Samples T Test...
Paired-Samples T Test...
One-Way ANOVA...
Name Type Label
13 q6Religie_R Numeric
14 q6Religie_P Numeric
15 q7Scoala_R Numeric
16 q7Scoala_P Numeric
17 q8Statut_pr... Numeric
18 q8Statut_pr... Numeric
19 No_relation Numeric of stable relationships (including the current one)
20 q10_A_luni Numeric formation less than 1 year
21 q10_B_Ani Numeric formation ore than 1 year
22 q11Prima__... Numeric date (first contact) took place in one of the situati
23 q12_Sentim... Numeric ou describe what you felt on your first date?
24 q13_A_ Numeric gether less than 1 year
25 q13_B_ani Numeric gether more than 1 year
26 q14 Numeric s the first to suggest living together?
27 q15_Prima_... Numeric mmon dwelling
28 q16_propriet... Numeric common dwelling
29 q17 Numeric g after you started dating emotional commitment w
30 q18 Numeric feel you got married too early?
31 q19_celebrat Numeric ebrated you marriage by
32 q20_statut_R Numeric ofessional status of the respondent
33 q20_statut_P Numeric ofessional status of the partner
34 q21 Numeric 0 0 You would accept the idea of living with someone older tha
35 q22 Numeric 8 0 You would accept the idea of living with someone younger
36 q23 Numeric 8 0 For you, physical attraction towards someone


Paired-Samples T Test
q24#1
q24#2
q24#3
q24#4
q24#5
q25
q26#1
q26#2
q26#3
q26#4
q26#5
Paired Variables:
Pair Variable1 Variable2
1
Options...
Bootstrap...
OK Paste Reset Cancel Help



T-Test

```
[DataSet1] C:\Users\stud\Downloads\Formare cupluri.sav
```

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	household skills	7,78	4511	2,087	,031
	good financial status	6,57	4511	2,563	,038

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	household skills & good financial status	4511	,239	,000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	household skills - good financial status	1,205	2,893	,043	1,121	1,290	27,983	4510	,000

Pt indeletnicitiri gospo media e 7,78

Pt relatie buna financiara e 6,57

Distanta medie este = 1,2 = d = chinar galben

Chinar portocaliu = testul t

Sig(2 tailed) = 99,9% = 0,00 <0,06 deci accept H1

Importanta acordata treb gospo e semnificativ mai mare dect imp acordata statului financiar(am ajuns la comcluzia asta comparand mediile

Compararea mediilor din 3 sau mai multe esantione dependete

Avem nev de o var dependent care sa fie numerica(neaparat)

Si o variabila factor care sa imparta in 3 sau mai multe grupuri

Premisele testarii:

Pas1:

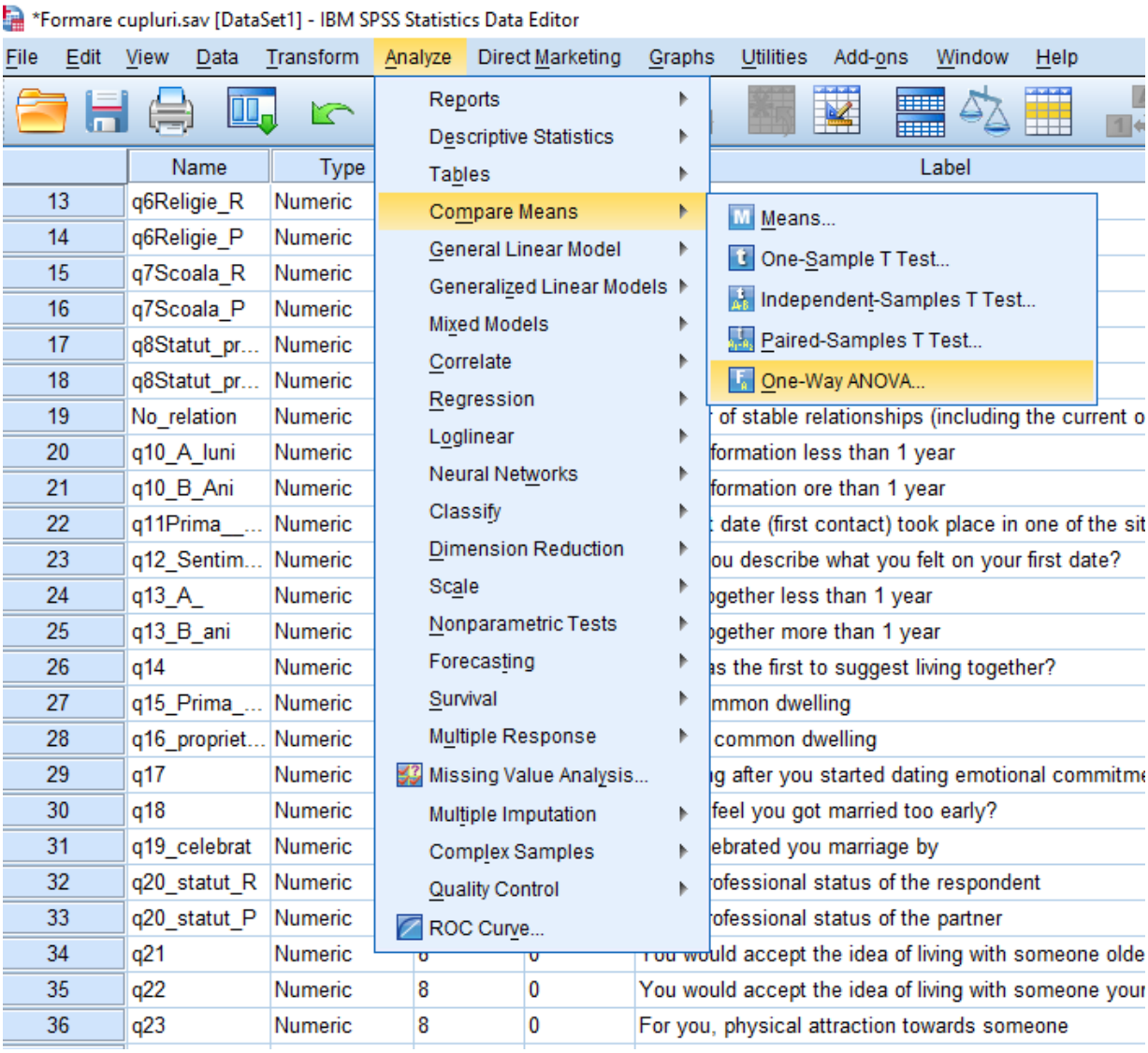
Esantion trebuie sa fie independent, aleator, HOMOSCEDASCITATE(ADICA H0: media i =media j oricare I si j ; H1: media i !=media j oricare I si j)

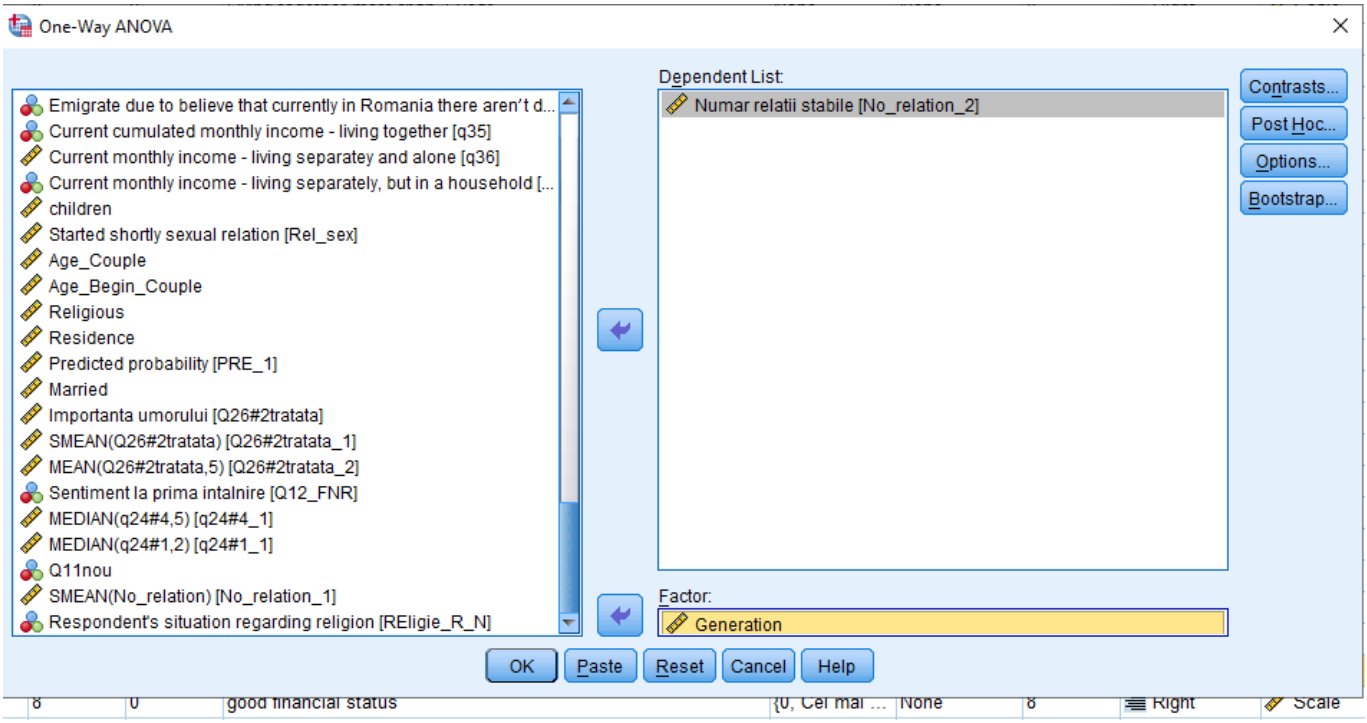
Daca se accepta H0 se aplica un test ANOVA

Daca se accepta H1 testul ANOVA NU MAI POATE FI INTERPRETAT(tot o s o avem in outrput) aplicam corectia Welch cu F*, df1, df* 9df=grade de libertate penalizate)

Var depen: nr de relatii stabile

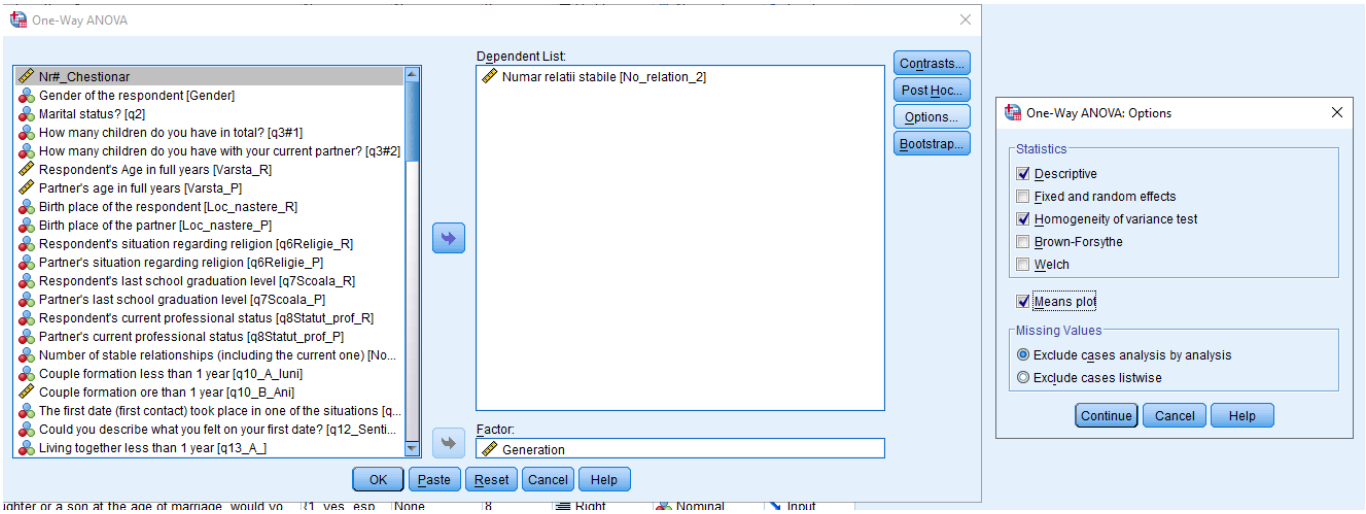
Var factor: gerupe de varsta


*Formare cupluri.sav [DataSet1] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help
Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Quality Control
ROC Curve...
Means...
One-Sample T Test...
Independent-Samples T Test...
Paired-Samples T Test...
One-Way ANOVA...
Name Type Label
13 q6Religie_R Numeric
14 q6Religie_P Numeric
15 q7Scoala_R Numeric
16 q7Scoala_P Numeric
17 q8Statut_pr... Numeric
18 q8Statut_pr... Numeric
19 No_relation Numeric of stable relationships (including the current o
20 q10_A_luni Numeric formation less than 1 year
21 q10_B_Ani Numeric formation ore than 1 year
22 q11Prima__... Numeric date (first contact) took place in one of the sit
23 q12_Sentim... Numeric ou describe what you felt on your first date?
24 q13_A_ Numeric gether less than 1 year
25 q13_B_ani Numeric gether more than 1 year
26 q14 Numeric s the first to suggest living together?
27 q15_Prima_... Numeric mmon dwelling
28 q16_propriet... Numeric common dwelling
29 q17 Numeric g after you started dating emotional commitme
30 q18 Numeric feel you got married too early?
31 q19_celebrat Numeric ebrated you marriage by
32 q20_statut_R Numeric ofessional status of the respondent
33 q20_statut_P Numeric ofessional status of the partner
34 q21 Numeric ould accept the idea of living with someone olde
35 q22 Numeric 8 0 You would accept the idea of living with someone your
36 q23 Numeric 8 0 For you, physical attraction towards someone



Folosim ca factor var generatii(pe la final o gasim)

Alegem din OPTIONS ASA:



Oneway

[DataSet1] C:\Users\stud\Downloads\Formare cupluri.sav

Descriptives

Numar relatii stabile

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
<25 ani	1560	1,771	,7760	,0196	1,733	1,810	1,0	3,0
25-34 ani	1469	1,930	,8117	,0212	1,888	1,971	1,0	3,0
>=35 ani	1496	1,589	,7796	,0202	1,550	1,629	1,0	3,0
Total	4525	1,762	,8008	,0119	1,739	1,786	1,0	3,0

Test of Homogeneity of Variances

Numar relatii stabile

Levene Statistic	df1	df2	Sig.
,902	2	4522	,406

ANOVA

Numar relatii stabile

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	86,271	2	43,135	69,304	,000
Within Groups	2814,524	4522	,622		
Total	2900,795	4524			

Test of Homogeneity of Variances

Numar relatii stabile

Levene Statistic	df1	df2	Sig.
,902	2	4522	,406

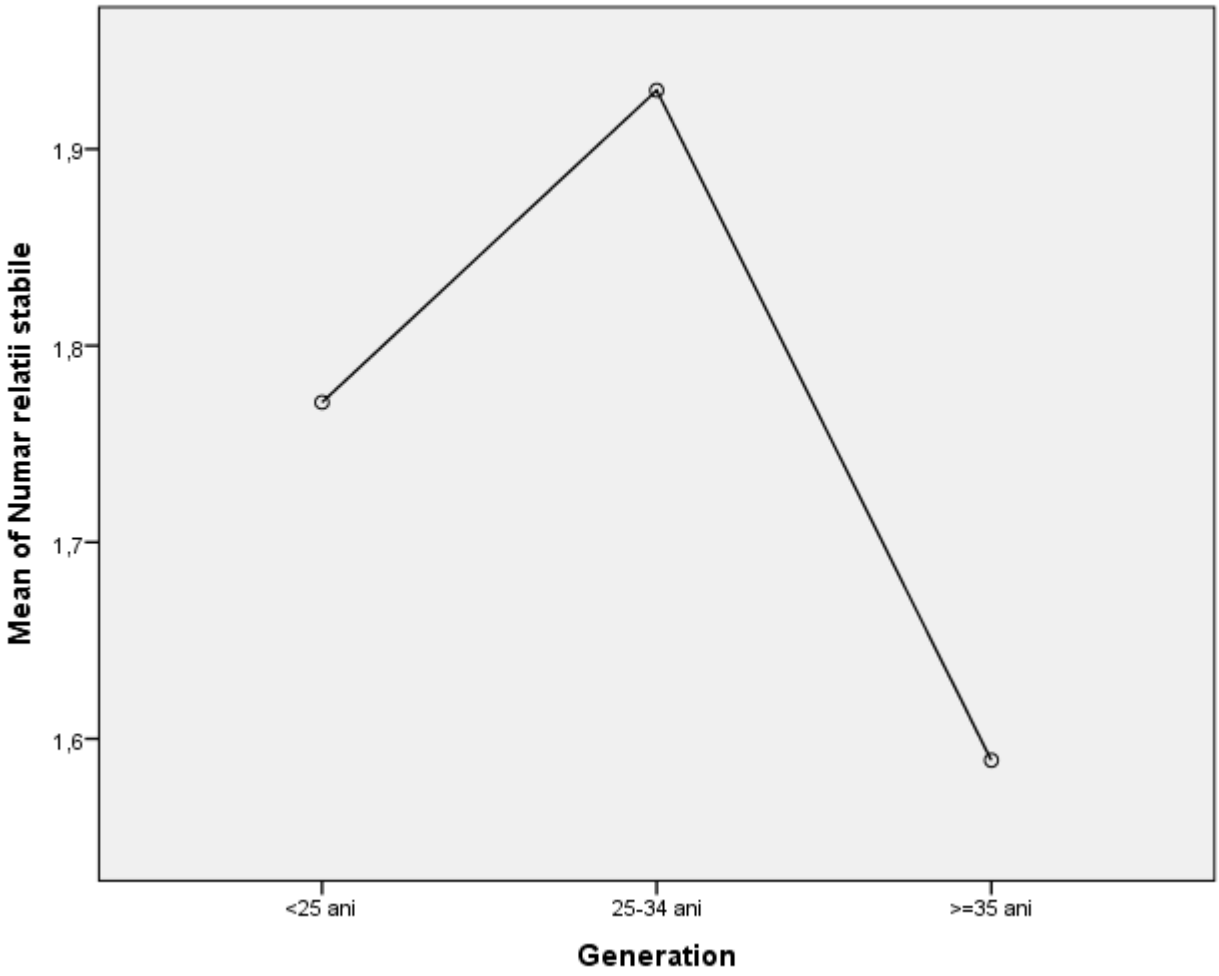
ANOVA

Sig= 0,406> 0,05 deci H0; disp i=disp j

Chenar albastru = 0,00< 0,05 deci acceptam H1= mediile nu sunt egale macar la 2 dintre chestiile comparate, nu ne zice care dintre ele

Acum graficul:

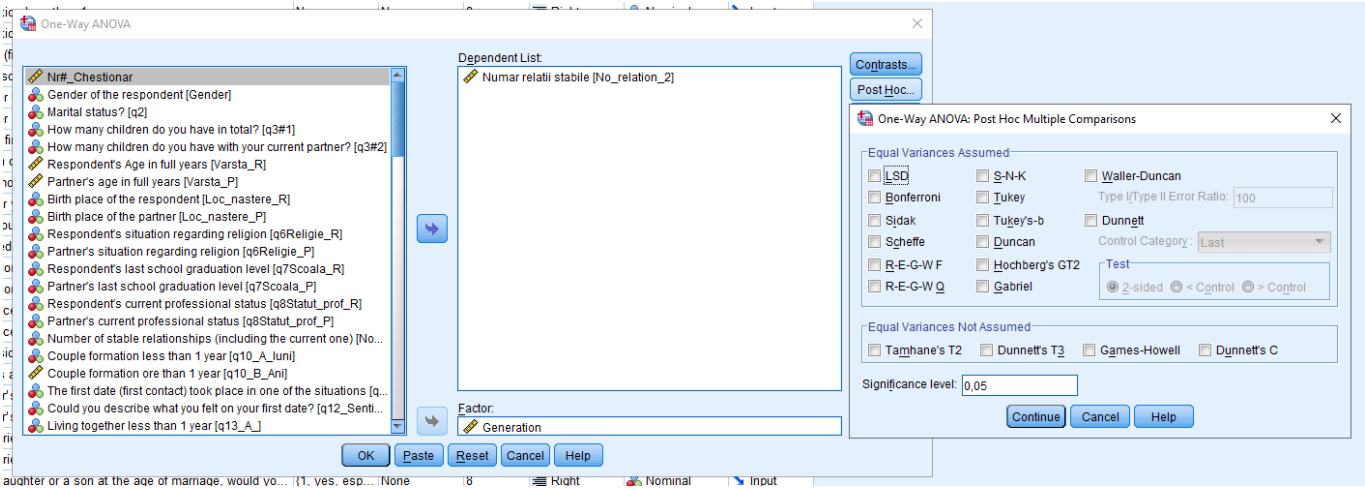
Means Plots



Acest lucru datorita schimbarii comp: cei peste 35 de ani au avut mai putine relatii per total(communist)

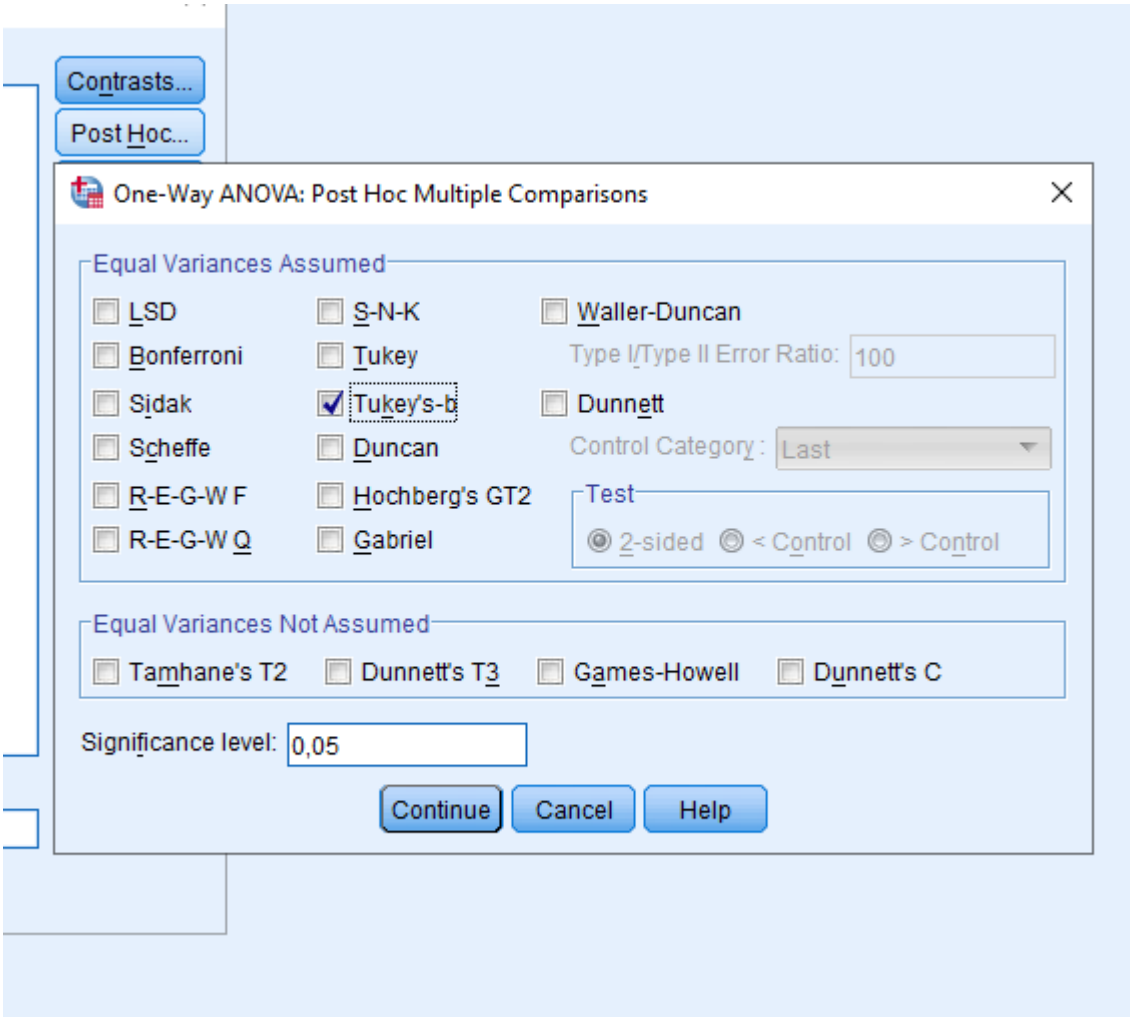
Cel putin una din medii e diferita – vrem sa verificam care e relatia dintre toate mediile SI PENTRU ASTA MAI OPTAM PT CEVA

Facem din nou Anova si alegem un test POST HOC



Disperiile sunt egale; alegem unu din testele din primul chinar

Alegem Tukeys-b



Post Hoc Tests

Homogeneous Subsets

Numar relatii stabile

Tukey B[a,b]

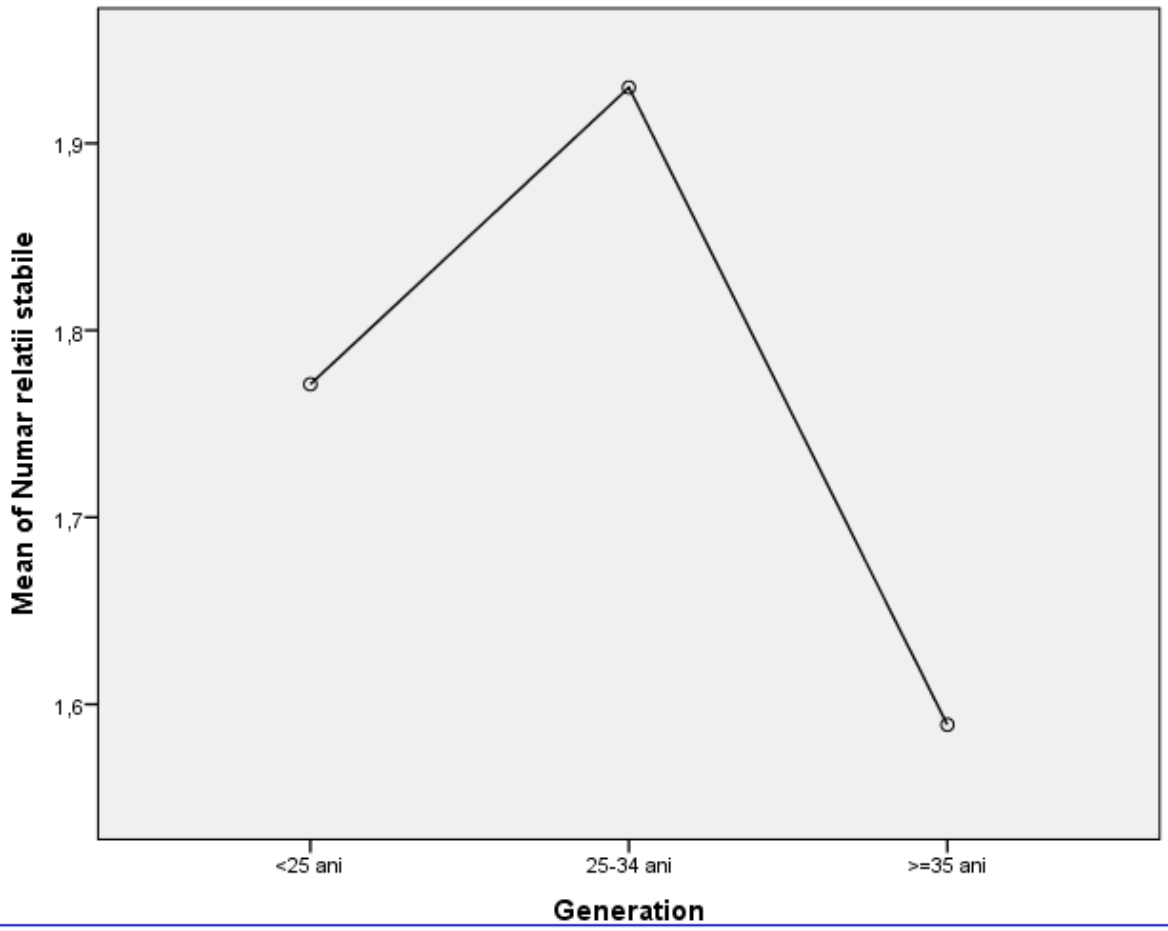
Generation	N	Subset for alpha = 0.05		
		1	2	3
>=35 ani	1496	1,589		
<25 ani	1560		1,771	
25-34 ani	1469			1,930

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 1507,379.

b. The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

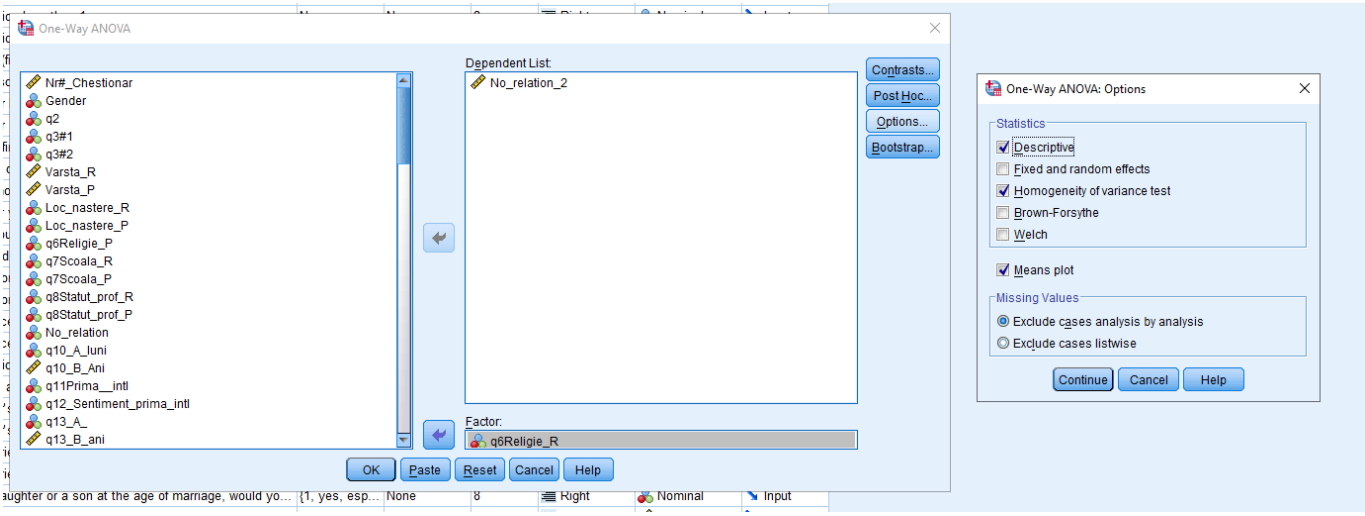
Means Plots



Toate cele 3 difera semnificativ; vezi nr in cascada cum ar fi pe trepte

ALT TEST

Aceeasi var(nr relatii)



Oneway

```
[DataSet1] C:\Users\stud\Downloads\Formare cupluri.sav
```

Descriptives

Numar relatii stabile

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Practica frecventa (regulata)	625	1,624	,7819	,0313	1,563	1,686	1,0	3,0
Practica ocazionala	2507	1,715	,7826	,0156	1,685	1,746	1,0	3,0
Nu practic, dar am un sentiment de apartenenta	1069	1,890	,8193	,0251	1,841	1,939	1,0	3,0
Nici practica, nici apartenenta	239	1,985	,8432	,0545	1,878	2,093	1,0	3,0
Nu stiu	85	1,932	,7651	,0830	1,767	2,097	1,0	3,0
Total	4525	1,762	,8008	,0119	1,739	1,786	1,0	3,0

Test of Homogeneity of Variances

Numar relatii stabile

Levene Statistic	df1	df2	Sig.
1,436	4	4520	,219

ANOVA

Numar relatii stabile

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	49,182	4	12,295	19,489	,000
Within Groups	2851,613	4520	,631		
Total	2900,795	4524			

E la fel…