

I. Observarea statistică (culegerea datelor)

1. Populațiile statistice cercetate (aferente fiecărui domeniu studiat în parte)

Pensiunile din România în luna iunie.

2. Stabilirea variabilelor studiate (unde este cazul se va elabora chestionar)

- Denumire pensiune - variabilă nominală
- Regiune- variabilă nominală
- Obiectiv turistic - variabilă nominală
- Stațiune - variabilă nominală
- Mare - variabilă nominală
- Oras - variabilă nominală
- Loc izolat - variabilă nominală
- Numar de stele- variabilă scale
- Scor - variabilă scale
- Personal - variabilă scale
- Servicii - variabilă scale
- Curatenie - variabilă scale
- Confort - variabilă scale
- Pret calitate - variabilă scale
- Localizare - variabilă scale
- Tarif - variabilă scale
- Număr locuri - variabilă scale
- Număr de camere - variabilă scale
- Oferă mic dejun - variabilă nominală
- Acces la bucatarie - variabilă nominală
- Grădina/curte - variabilă nominală
- Animale de companie - variabilă nominală
- ATV - variabilă nominală
- Șemineu - variabilă nominală
- Restaurant - variabilă nominală
- Teren de sport - variabilă nominală
- Ciubăr - variabilă nominală
- Sauna - variabilă nominală
- Închirieri de biciclete - variabilă nominală
- Tichete de vacanță - variabilă nominală
- Aer condiționat - variabilă nominală

- Scor_ordinal -variabilă ordinală
- Număr_locuri_ordinal- variabilă ordinală

3. **Tipul de sondaj aplicat**- sondaj aleator simplu fără revenire

4. **Formarea eşantioanelor** (unde este cazul se va administra chestionarul)

Am selectat pensiuni din județele: Brașov, Harghita, Mureș, Suceava.

5. **Culegerea datelor** (unde este cazul se va administra chestionarul)

Datele au fost culese din cadrul site-ului **turistinfo.ro**

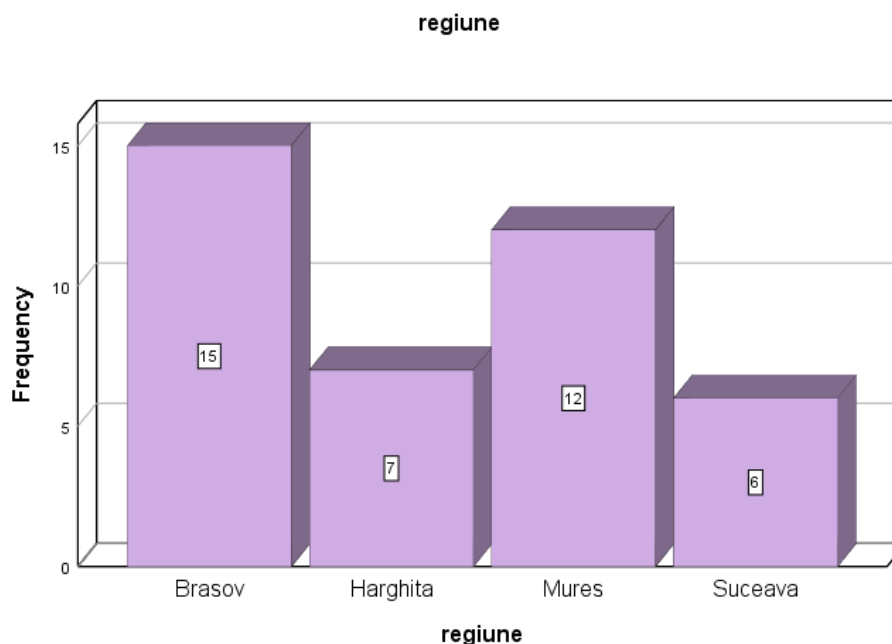
II. Prelucrarea datelor

(utilizând programul SPSS sau Excel)

Prelucrări univariate :

- Analiza unei variabile nominale

regiune					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Brasov	15	37.5	37.5	37.5
	Harghita	7	17.5	17.5	55.0
	Mures	12	30.0	30.0	85.0
	Suceava	6	15.0	15.0	100.0
	Total	40	100.0	100.0	



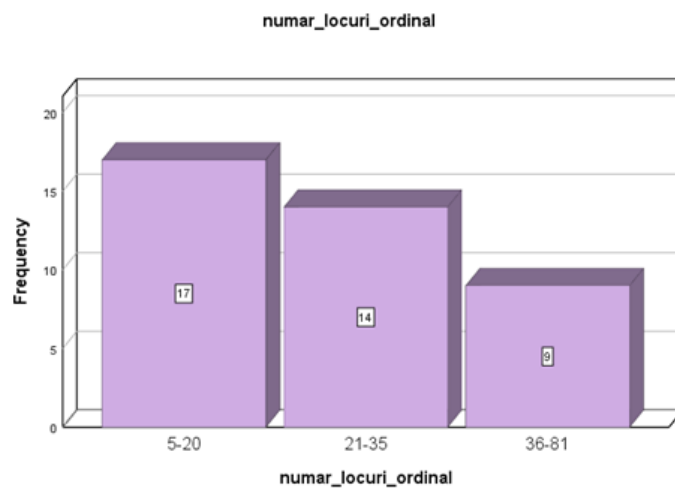
Numarul total de pensuni selectate este 40. Dintre acestea, 15 (37.5%) sunt din judetul Brasov, 7 (17.5%) din Harghita, 12 (30%) din Mures si 6 (15%) din Suceava.

- Analiza unei variabile ordinale

Statistics		
numar_locuri_ordinal		
N	Valid	40
	Missing	0
Median		2.0000
Mode		1.00
Percentiles	25	1.0000
	50	2.0000
	75	2.0000

Ținând cont de modul de codificare a variabilei observăm că jumătate din pensuni au cel mult 35 de locuri, iar ceilalți 50% au între 35 și 81 de locuri. Cele mai multe pensuni au între 5 și 20 de locuri. Folosind valorile quartile, distribuția eșantionului în raport cu variabila ”număr locuri” se prezintă astfel: 25% dintre pensuni au până la 20 de locuri, iar 75% au până în 35 de locuri, sau se poate spune că 25% dintre pensuni au mai mult de 35 de locuri .

numar_locuri_ordinal					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	5-20	17	42.5	42.5	42.5
	21-35	14	35.0	35.0	77.5
	36-81	9	22.5	22.5	100.0
	Total	40	100.0	100.0	

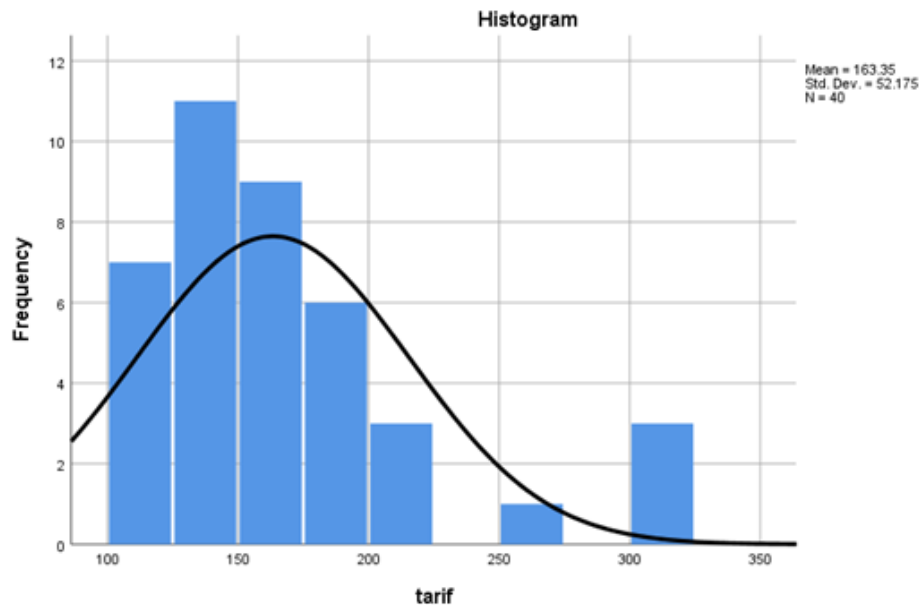


- Analiza unei variabile scale

Statistics		
tarif		
N	Valid	40
	Missing	0
Mean		163.35
Median		150.00
Mode		150
Std. Deviation		52.175
Variance		2722.233
Skewness		1.435
Std. Error of Skewness		.374
Kurtosis		1.601
Std. Error of Kurtosis		.733
Minimum		100
Maximum		300
Percentiles	25	130.00
	50	150.00
	75	180.00

tarif					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	100	2	5.0	5.0	5.0
	110	1	2.5	2.5	7.5
	120	4	10.0	10.0	17.5
	125	1	2.5	2.5	20.0
	130	6	15.0	15.0	35.0
	135	2	5.0	5.0	40.0
	140	2	5.0	5.0	45.0
	150	7	17.5	17.5	62.5
	160	1	2.5	2.5	65.0
	170	1	2.5	2.5	67.5
	180	4	10.0	10.0	77.5

	189	1	2.5	2.5	80.0
	190	1	2.5	2.5	82.5
	220	3	7.5	7.5	90.0
	250	1	2.5	2.5	92.5
	300	3	7.5	7.5	100.0
	Total	40	100.0	100.0	



Avem un eșantion de 40 de pensiuni din Harghita, Mureș, Suceava și Brașov. Tarifele pentru o noapte de weekend într-o cameră matrimonială variază între 100 și 300 de lei. Prețul mediu este 163.35 de lei, cu o abatere de 52.175 lei. Jumătate dintre pensiuni au prețuri cuprinse între 100 și 150 de lei, iar cealaltă jumătate au prețuri cuprinse între 150 și 300 de lei. Cele mai multe pensiuni au tarife în jur de 150 de lei/noapte.

În ceea ce privește parametrii formei, se observă că seria este asimetrică pozitiv ($\text{skewness} > 0$), iar în ceea ce privește boltirea seriei, avem o serie leptokurtică, adică boltirea corespunzătoare curbei respective este mai înaltă și mai ascuțită decât curba normală ($\text{kurtosis} > 0$).

- Teste de semnificatie

→ Testarea semnificatiei unei medii (Testul T)

Formulăm ipotezele statistice ale testului T:

H_0 : Valoarea medie a variabilei „Număr de camere” analizate nu diferă semnificativ de valoarea 12.

H_1 : Valoarea medie a variabilei „Număr de camere” analizate diferă semnificativ de valoarea 12.

One-Sample Statistics				
	<i>N</i>	<i>Mean</i>	<i>Std. Deviation</i>	<i>Std. Error Mean</i>
<i>nr_de_camere</i>	40	10.70	5.464	.864

Din acest tabel se poate observa faptul că valoarea medie în eșantion este de 10.7, iar abaterea standard este 5,464.

One-Sample Test						
	<i>Test Value = 12</i>					
	<i>t</i>	<i>df</i>	<i>Sig. (2-tailed)</i>	<i>Mean Difference</i>	<i>95% Confidence Interval of the Difference</i>	
					<i>Lower</i>	<i>Upper</i>
<i>nr_de_camere</i>	-1.505	39	.140	-1.300	-3.05	.45

Iar din tabelul de mai sus se poate observa valoarea lui $t_{\text{calc}} = -1.505$, avem 39 de grade de libertate și un $\text{sig} = 0.140$.

Fixăm pragul de semnificație 0.05.

$\text{Sig} = 0.140 > 0.05 \Rightarrow$ ipoteza nulă se acceptă, adică valoarea medie a variabilei analizate nu diferă semnificativ de 12

→ Testarea semnificației unei proporții (Testul Binomial)

- ★ Variabilă binară

Formulăm ipotezele statistice ale testului Binomial:

H_0 = Proporția pensiunilor care nu oferă mic dejun nu diferă semnificativ de 45%.

H_1 = Proporția pensiunilor care nu oferă mic dejun diferă semnificativ de 45%.

Binomial Test						
		Category	N	Observed Prop.	Test Prop.	Exact Sig. (1-tailed)
ofera_mic_dejun	Group 1	Nu	19	.48	.45	.435
	Group 2	Da	21	.53		
	Total		40	1.00		

Observăm că la nivelul eșantionului de 40 de pensiuni, 19 nu oferă mic dejun, adică o proporție de 48%. Dacă fixăm un prag de semnificație de 5%, rezultă: $\text{sig}=0.435>0.05$, deci acceptăm ipoteza nulă potrivit căreia procentul din eșantion a pensiunilor care nu oferă mic dejun nu diferă semnificativ de 45%. Putem afirma că procentul observat din eșantion este mai mare, 48%.

★ Variabilă scale

Formulăm ipotezele statistice ale testului Binomial:

H_0 =Proporția pensiunilor care au cel mult 3 stele nu diferă semnificativ de 79,99%.

H_1 = Proporția pensiunilor care au cel mult 3 stele diferă semnificativ de 79,99%

Binomial Test						
		Category	N	Observed Prop.	Test Prop.	Exact Sig. (1-tailed)
nr_de_stele	Group 1	≤ 3	34	.850000	.799900	.285
	Group 2	> 3	6	.150000		
	Total		40	1.000000		

Din rezultatele obținute constatăm că proporția observată în eșantion este de 85%. Dacă fixăm un prag de semnificație de 5%, rezultă $\text{sig}=0.285>0.05$, se acceptă ipoteza nulă, respectiv proporția pensiunilor care au cel mult 3 stele nu diferă semnificativ de 79,99%.

★ Variabila ordinală

Formulăm ipotezele statistice ale testului Binomial:

H_0 = Proporția pensiunilor care au cel mult 35 de locuri nu diferă semnificativ de 66.66%

H_1 = Proporția pensiunilor care au cel mult 35 de locuri diferă semnificativ de 66.66%

Binomial Test						
		Category	N	Observed Prop.	Test Prop.	Exact Sig. (1-tailed)
numar_locuri_ordinal	Group 1	≤ 2	31	.775000	.666600	.096
	Group 2	> 2	9	.225000		
	Total		40	1.000000		

Din rezultatele obținute constatăm că proporția observată în eșantion a pensiunilor care au cel mult 35 de locuri este de 77.5%. Dacă fixăm un prag de semnificație de 5%, rezultă $\text{sig}=0.096>0.05$, se acceptă ipoteza nulă, respectiv proporția pensiunilor care au cel mult 35 de locuri nu diferă semnificativ de 66.66%.

Prelucrari bivariate:

→ Analiza asocierii dintre doua variabile nominale:

În vederea realizării analizei, alegem variabilele nominale: **Restaurant, Acces bucătărie.**

Pentru măsurarea asocierii vom folosi **coeficientul Phi.**

restaurant * acces_bucatarie Crosstabulation				
Count				
		acces_bucatarie		Total
		Nu	Da	
restaurant	Nu	4	22	26
	Da	11	3	14
Total		15	25	40

Chi-Square Tests					
	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	15.502 ^a	1	.000		
Continuity Correction ^b	12.923	1	.000		
Likelihood Ratio	16.052	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	15.114	1	.000		
N of Valid Cases	40				
a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 5.25.					
b. Computed only for a 2x2 table					

Symmetric Measures			
		Value	Approximate Significance
Nominal by Nominal	Phi	-.623	.000
	Cramer's V	.623	.000
	Contingency Coefficient	.528	.000
N of Valid Cases		40	

Astfel formulăm ipoteza nulă:

H0 : cele doua variabile sunt independente (nu există asociere)

H1 : cele doua variabile sunt dependente (există asociere)

Avem **Phi= -0,623**

Sig= 0.000082 <0,05 (nivelul de semnificație fixat) rezulta ca ipoteza nula este respinsa, este acceptata ipoteza alternativa, respectiv exista asociere între existenta restaurantului si accesul în bucătărie.

Valoarea **coeficientului** fiind negativa rezulta ca sunt mai multe cazuri de **da** cu **nu** sau invers decat **da** cu **da** si **nu** cu **nu**.)

→ Analiza legăturii dintre variabilele cantitative:

În vederea realizării analizei, alegem variabilele scale: **Confort, Curatenie**.

Pentru analiza legăturii vom folosi **Testul lui Pearson**.

Formulam ipotezele:

H0: nu exista asociere între gradul de mulțumire al clienților privind confortul si gradul de mulțumire privind curățenia.

H1: exista asociere între gradul de mulțumire al clienților privind confortul si gradul de mulțumire privind curățenia.

Correlations			
		confort	curatenie
confort	Pearson Correlation	1	.769**
	Sig. (2-tailed)		.000
	N	40	40
curatenie	Pearson Correlation	.769**	1
	Sig. (2-tailed)	.000	
	N	40	40

** . Correlation is significant at the 0.01 level (2-tailed).

Coeficientul lui **Pearson= 0.769**>0 -> exista o legătură directă între gradul de mulțumire al clienților privind confort si gradul de mulțumire privind curățenia din cadrul pensiunilor.

Pe de altă parte. valoarea **Sig-ului = 0,00** < 0,05, deci probabilitatea de acceptare a ipotezei nule este foarte redusă (aproape zero), deci ipoteza nulă se respinge și se acceptă ipoteza alternativă (*adică faptul că există legătură între gradul de mulțumire al clienților privind curățenia si gradul de mulțumire privind confortul*).

→ Analiza legăturii dintre variabilele ordinale:

Am construit două variabile ordinale **Scor_ordinal** și **Număr_locuri**

Folosim testul de corelație **Kendall**

Se formulează ipotezele statistice:

- H0: Nu există legătură (asociere) între variabilele analizate
- H1: Există legătură (asociere)

Correlations				
			scor_ordin	numar_locuri_ordinal
Kendall's tau_b	scor_ordin	Correlation Coefficient	1.000	-.152
		Sig. (2-tailed)	.	.298
		N	40	40
	numar_locuri_ordinal	Correlation Coefficient	-.152	1.000
		Sig. (2-tailed)	.298	.
		N	40	40

Obținem o matrice de coeficienți de corelație care este simetrică față de diagonala principală.

În tabelul de mai sus găsim valoarea coeficientului Knedall (K= -0.152)

În plus, avem Sig=0.298>0,05 (pragul fixat), deci ipoteza nulă se acceptă, deci nu există legătură între scorul obținut și numărul de locuri disponibile în pensiuni.

→ Analiza varianței (anova)

Alegem:

- o variabilă scale- dependentă (X)- TARIF
- o variabila nominală- factor (Y) - RESTAURANT

Formulăm ipotezele:

H0: Nu există diferențe între tarifele medii definite de prezenta restaurantului în pensiuni

H1: Există cel puțin două pensiuni al căror tarif diferă semnificativ, în funcție de prezența sau absența restaurantului (Există diferențe între tarifele medii definite de prezența restaurantului în pensiuni

Descriptives								
tarif								
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Nu	26	151.54	50.592	9.922	131.10	171.97	100	300
Da	14	185.29	49.470	13.221	156.72	213.85	120	300
Total	40	163.35	52.175	8.250	146.66	180.04	100	300

ANOVA					
tarif					
	Sum of Squares	df	Mean Square	F	Sig.
<i>Between Groups</i>	10363.781	1	10363.781	4.111	.04966
<i>Within Groups</i>	95803.319	38	2521.140		
Total	106167.100	39			

Din tabel rezultă:

- Variația din interiorul grupurilor (within groups)= 95803.319
- Variația dintre grupuri (between groups)= 10363.781
- F= 4.111
- Sig= 0.049667

Fixăm un nivel (prag) de semnificație de 0,05.

Cum $Sig < 0.05$, rezultă că ipoteza nulă este respinsă, respectiv este acceptată ipoteza alternativă, respectiv există diferențe între mediile variabilei dependente calculate pe grupe, adică *Tariful pensiunilor se diferențiază în raport cu prezența unui restaurant în pensiune.*

Alegem:

- o variabilă scale- dependentă (X)- TARIF
- o variabilă ordinală- factor (Y) - SCOR_ORD

Formulăm ipotezele:

H0: Nu există diferențe între tarifele medii definite de scorurile obținute de pensiuni

H1: Există cel puțin două pensiuni al căror tarif diferă semnificativ, în funcție de scorul obținut

Descriptives								
tarif								
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
7-8.99	8	176.13	34.971	12.364	146.89	205.36	120	220
9-9.25	6	155.83	45.433	18.548	108.15	203.51	100	220
9.26-10	26	161.15	58.486	11.470	137.53	184.78	100	300
Total	40	163.35	52.175	8.250	146.66	180.04	100	300

ANOVA					
tarif	Sum of Squares	df	Mean Square	F	Sig.
<i>Between Groups</i>	1770.007	2	885.004	.314	.733
<i>Within Groups</i>	104397.093	37	2821.543		
Total	106167.100	39			

Din tabel rezultă:

- Variația din interiorul grupurilor (within groups)= 104397.093
- Variația dintre grupuri (between groups)= 1770.007
- F= 0.314
- Sig= 0.733

Fixăm un nivel (prag) de semnificație de 0,05.

Cum Sig>0.06, rezultă că ipoteza nulă se accepta, adică *Tariful pensiunilor* nu se diferențiază în raport cu *scorul obținut*.

→ Teste de comparare

★ Testarea egalității a două medii, cazul eșantioanelor independente

Verificăm dacă variabila TARIF se diferențiază în funcție de acceptarea TICHETELOR DE VACANȚĂ.

Avem ipotezele:

H0: Tariful pensiunilor care accepta tichete de vacanță nu diferă semnificativ de tariful pensiunilor care nu acceptă tichete de vacanță.

H1: Tariful pensiunilor care acceptă tichete de vacanță diferă semnificativ de tariful pensiunilor care nu acceptă tichete de vacanță.

Group Statistics					
	<i>tichete_vacanta</i>	<i>N</i>	<i>Mean</i>	<i>Std. Deviation</i>	<i>Std. Error Mean</i>
<i>tarif</i>	<i>Nu</i>	10	163.90	60.175	19.029
	<i>Da</i>	30	163.17	50.369	9.196

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
tarif	Equal variances assumed	.596	.445	.038	38	.970	.733	19.300	-38.338	39.805
	Equal variances not assumed			.035	13.467	.973	.733	21.134	-44.765	46.231

Deoarece testul T se poate aplica într-una din următoarele situații>

1. σ_1^2, σ_2^2 -nu se cunosc dar sunt egale $\sigma_1^2 = \sigma_2^2 = \sigma^2$
2. σ_1^2, σ_2^2 -nu se cunosc și sunt diferite

★ Testăm ipoteza privind egalitatea varianțelor.

Avem ipotezele:

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_1^2 \neq \sigma_2^2$$

În tabelul de mai sus (al lui *Levene*), găsim $Sig=0,445 > 0,05$, deci acceptăm ipoteza nulă. Prin urmare $\sigma_1^2 = \sigma_2^2$ este validă și suntem în cazul 1.

Dacă acceptăm această ipoteză, atunci, mai departe, citim tabelul t-test for Equality of Means. Și alegem din acest tabel linia unu corespunzătoare cazului în care egalitatea variantelor este asumată.

$$T_{\text{calculat}} = 0.038$$

$$\text{Grade libertate} = 38$$

$$Sig = 0,970$$

Dacă fixăm un prag de semnificație de 5%, rezulta: $Sig = 0,970 > 0,05$ deci, în testul de egalitate al mediilor, acceptăm ipoteza nulă, adică *tarifele pensiunilor care acceptă tichete de vacanță nu diferă în medie de tarifele pensiunilor care nu acceptă tichete de vacanță*.

Raport Final

-cele mai importante concluzii la care am ajuns sunt:

- Din totalul de 40 de pensiuni 15 (37.5%) sunt din județul Brașov, 7 (17.5%) din Harghita, 12 (30%) din Mureș și 6 (15%) din Suceava.

În urma aplicării testelor de semnificație, am observat că:

Aștept parere/sugestii/opinii/reclamatii

- Valoarea medie a variabilei "Număr de camere" nu diferă semnificativ de 12
- Procentul din eșantion a pensiunilor care nu oferă mic dejun nu diferă semnificativ de 45%. Putem afirma că procentul observat din eșantion este mai mare, 48%.
- Proportia pensiunilor care au cel mult 3 stele nu diferă semnificativ de 79,99%.
- Proportia pensiunilor care au cel mult 35 de locuri nu diferă semnificativ de 66.66%.

În urma analizelor bivariante, am ajuns la următoarele concluzii:

- Există asociere între existența restaurantului și accesul în bucătărie.
- Există legătură între gradul de mulțumire al clienților privind curățenia și gradul de mulțumire privind confortul.
- Nu există legătură între scorul obținut și numărul de locuri disponibile în pensiuni.

În urma analizei ANOVA, am obținut:

- Există diferențe între mediile variabilei dependente calculate pe grupe, adică *Tariful pensiunilor* se diferențiază în raport cu *prezența unui restaurant în pensiune*.
- *Tariful pensiunilor* nu se diferențiază în raport cu *scorul obținut*.

În urma aplicării testului de comparare, am obținut:

- Tarifele pensiunilor care acceptă tichete de vacanță nu diferă în medie de tarifele pensiunilor care nu acceptă tichete de vacanță.