STATISTICAL ANALYSIS OF AVERAGE PRODUCTION AT WHEAT AND CORN IN 2016 IN THE COUNTIES OF ROMANIA

SABĂU ADRIAN[1], GABOR CRISTINA[1], BALAJ ALEXANDRU[1], RUJESCU CIPRIAN*[1], SALA FLORIN[2]

[1]*Banat University of Agricultural Sciences and Veterinary Medicine "King Michael I of Romania" from Timisoara, Romania, Faculty of Agricultural Management*
[2]*Banat University of Agricultural Sciences and Veterinary Medicine "King Michael I of Romania" from Timisoara, Romania, Faculty of Agriculture*
*Corresponding author's e-mail: rujescu@usab-tm.ro

Abstract: The average production for the surface unit for wheat and corn presents a high variety from one county to another in Romania in 2016. So, for wheat were recorded average values at hectare between 2800 kg/ha and 5200 kg/ha and at corn values between 3000 kg/ha and 5700 kg/ha. The aim of this paper is to realize a study based on statistically methods, which needs to facilitate the way we observe the variation at distinct geographical areas, respectively determination of some particularities of them.

Key words: analysis, cluster, average production, wheat, corn

INTRODUCTION

Statistical data processing about the average hectare production of cereals is a persistent preoccupation and also an actual one [6,8,9,10]. Necessities of those steps are sustained by the multiple changing values which appear from year to year in one area or another [13,16]. Further, the detailed knowledge of recent aspects on values about average production of cereals is necessary for understanding some actual aspects about the general condition of agriculture. Starting from statistically facts which INS Romania provides us [13], the aim of the study is to present in a large and actual way the values of average production at hectare both for wheat and corn, in the counties of Romania in 2016.

About the techniques of statistical analysis, literature gives us multiple ways to follow [2,7]. Cluster analysis is about grouping elements in a set, depending on the distance between the elements. This theoretical basis appeared early on XX century, starting from the necessity to realize some classification in psychology, anthropology, biology, afterwards evolving in parallel with algorithmic development and software applications [1,5]. Even if on first sight the notion of distance has a geometrical interpretation, however between the elements from a set, when those can be described quantitatively, it can be defined some expressions which can show the distance between them. Developing colouristic methods [3,4] have a real success showing the objectivity of classifications [11,14,15].

MATERIAL AND METHOD

The average values of wheat and corn production are given by the statistical data, for every county from Romania in 2016 (table 1). These are the ones given by INS through Tempo online [13].

The outliers values are those statistical data which are positioned at the extremes of a series of data and are not often representative of that series, their appearance is most often accidentally. However, there are situations in which these values describe very accurate some particular situation. These outliers values are determined directly using quartiles method, more precis, the values are those from the outside of the interval:

$$(Q_1 - 1.5\,IQR, Q_3 + 1.5\,IQR)$$

where IQR is interquartile deviation Q_3-Q_1.

There are the points $A_i(x_i, y_i)$, $i = 1,..,41$, if we write with x_i the value of average production at wheat, and with y_i the value of average production of corn, for every county aside, then between A_i points can be settled the distances using Euclidean distance formula:

$$d(A_i, A_j) = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}.$$

Classification of county values, A_i, using the cluster analysis it's realized depending the least values of points distances. The software application used for the clusters was Past 3. Also, the same application was used for the boxplot chart, but also for calculation of statistical indicators [12].

Heat map was built using the ACDSee PhotoManager application, following the distinctive color mapping of Romania's counties based on RGB color indexes. The coloring algorithm was realized based on matrix (R, G, B) of value (0,0,255) for the county with the lowest production and (255,0,0) for the county with the highest cereals production. The G value was zero everywhere, and the intermediate values are obtained proportionally with every county production.

Table 1.

The average production at hectares in the counties of Romania in 2016, at wheat and corn

Nr. crt.	County	Wheat	Corn	Nr. crt.	County	Wheat	Corn
1	Alba	4150	5935	21	Harghita	3222	3507
2	Arad	4931	5403	22	Hunedoara	3749	2409
3	Arges	2871	3017	23	Ialomita	5210	5717
4	Bacau	3586	4184	24	Iasi	3611	3634
5	Bihor	2935	4217	25	Ilfov	4157	4314
6	Bistrita-Nasaud	3142	3061	26	Maramures	3274	3527
7	Botosani	3021	3128	27	Mehedinti	3661	3519
8	Braila	4972	5760	28	Mures	4132	4565
9	Brasov	3857	3842	29	Neamt	4524	5355
10	Buzau	4408	4090	30	Olt	3170	2904
11	Calarasi	5049	5193	31	Prahova	4091	3852
12	Caras-Severin	3976	4235	32	Salaj	3310	3734
13	Cluj	3911	4667	33	Satu Mare	4200	4824
14	Constanta	4008	3191	34	Sibiu	3561	4091
15	Covasna	3399	4129	35	Suceava	3230	3156
16	Dambovita	3071	2889	36	Teleorman	3606	3616
17	Dolj	3357	4025	37	Timis	4625	6500
18	Galati	4203	3261	38	Tulcea	4065	3232
19	Giurgiu	3466	3681	39	Valcea	3114	3716
20	Gorj	3176	3061	40	Vaslui	3668	2554
				41	Vrancea	3440	3385

**Source:[13]*

RESEARCH RESULTS

Statistical indicators

Speaking about wheat production, the lowest values was 2871 kg/ha, and the maximum was 5210 kg/ha, that means a 1.8 higher value comparative. The value of average production at wheat is 3783 kg/ha which places our country on an inferior place regarding a ranking similar with the one of European Union [16]. The first quartile from our series is 3252 kg/ha, so 25% from the values of average production of counties are below this value. The median is 3661 kg/ha, so this value is delimiting the series of average production in two equal parts, and the third quartile is 4153 kg/ha.

So only a quarter of series values are superior to this value. The coefficient of variation of this series is 16.25% so we have a relatively homogeneous series. There wasn't observed outliers values.

Speaking about corn production, 25% of average production values recorded at the level of countiees from Romania in 2016 were inferior than 3211 kg/ha, that means the value of first quartile. 50% of values are lower than 3734 kg/ha, and 25% are bigger than 4439 kg/ha. The coefficient of variation in this case has a higher value, 24,39% because of standard deviation which are superior wheat values. This fact indicates a relative heterogeneity of series, this being characterized through big jumps from a area to another. At corn culture we observed an outlier value, of 6500 kg/ha. Even if sometimes those outliers values can be interpreted as being aberrant values, in this case, considering the similar results from the past years, we consider it grounded and based on a particularly efficient agro-technical support that is practiced in some areas of the country.

The most important statistical indicators, reprocess based on calculation that have been made with the Past 3 application, we have them in table 2.

Table 2.

Statistical indicators referring to average production at hectares in counties of Romania in 2016, at wheat and corn

| Wheat		Corn	
N	41	N	41
Min	2871	Min	2409
Max	5210	Max	6500
Sum	155109	Sum	163080
Mean	3783.146	Mean	3977.561
Std. error	96.018	Std. error	151.550
Variance	377998.3	Variance	941672.1
Stand. dev	614.8157	Stand. dev	970.397
Median	3661	Median	3734
Q1=25 prcntil	3252	Q1=25 prcntil	3211.5
Q3=75 prcntil	4153.5	Q3=75 prcntil	4439.5
IQR=Q3- Q1	901.5	IQR=Q3- Q1	1228
Q1-1.5IQR	1899.75	Q1-1.5IQR	1369.5
Q3+1.5IQR	5505.75	Q3+1.5IQR	6281.5
Outliers	-	Outliers	6500
Skewness	0.647	Skewness	0.811
Kurtosis	-0.315	Kurtosis	0.122
Geom. mean	3736.534	Geom. mean	3869.916
Coeff. var	16.25%	Coeff. var	24.39%
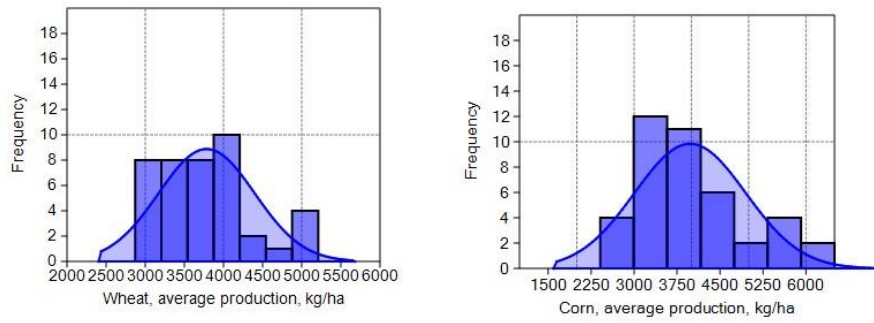
Figure 1. Histograms of average productions per hectare for wheat and corn in Romanian counties

Following the histograms of the average productions we observe a high average production rate of 4000 kg / ha for wheat and 3000-4000 kg / ha for maize.
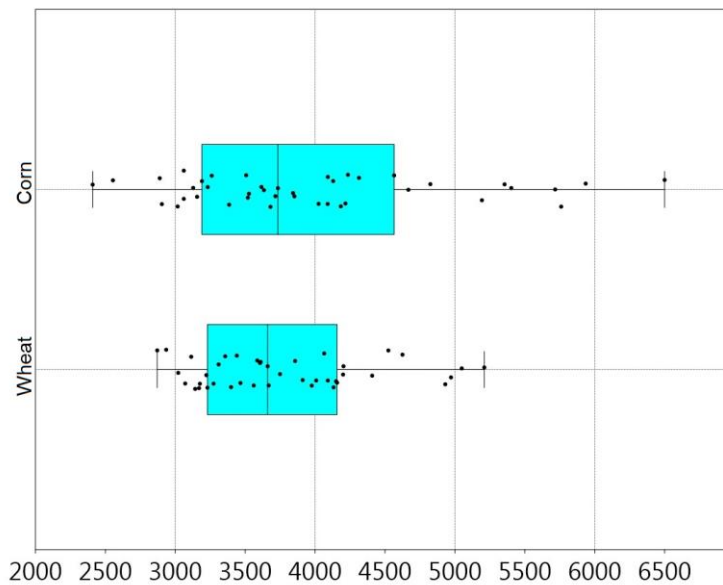

Figure 2. Boxplot chart with jitter, regarding the average production of wheat and corn in the counties of Romania

Distribution of average production values can be oserved easily analyzing the Box and jitter chart [12,15]. The values form inside of rectangle are the ones between quartiles 1 and 3. The exterior segments, through their extremities indicate the minimum values and also the maximum values of the series. To notice, is the fact that the outlier values is outside of this marking.

The classification of counties regarding the average production values at wheat and corn, obtained on Euclidean distance, are standing in clusters chart. Basically, the distance between the counties can be observed watching the vertical axis. These values were calculated after this way:

$$d_{ij} = \sqrt{\left(\text{average production wheat}_i - \text{average production wheat}_j\right)^2 + \left(\text{average production corn}_i - \text{average production corn}_j\right)^2}$$
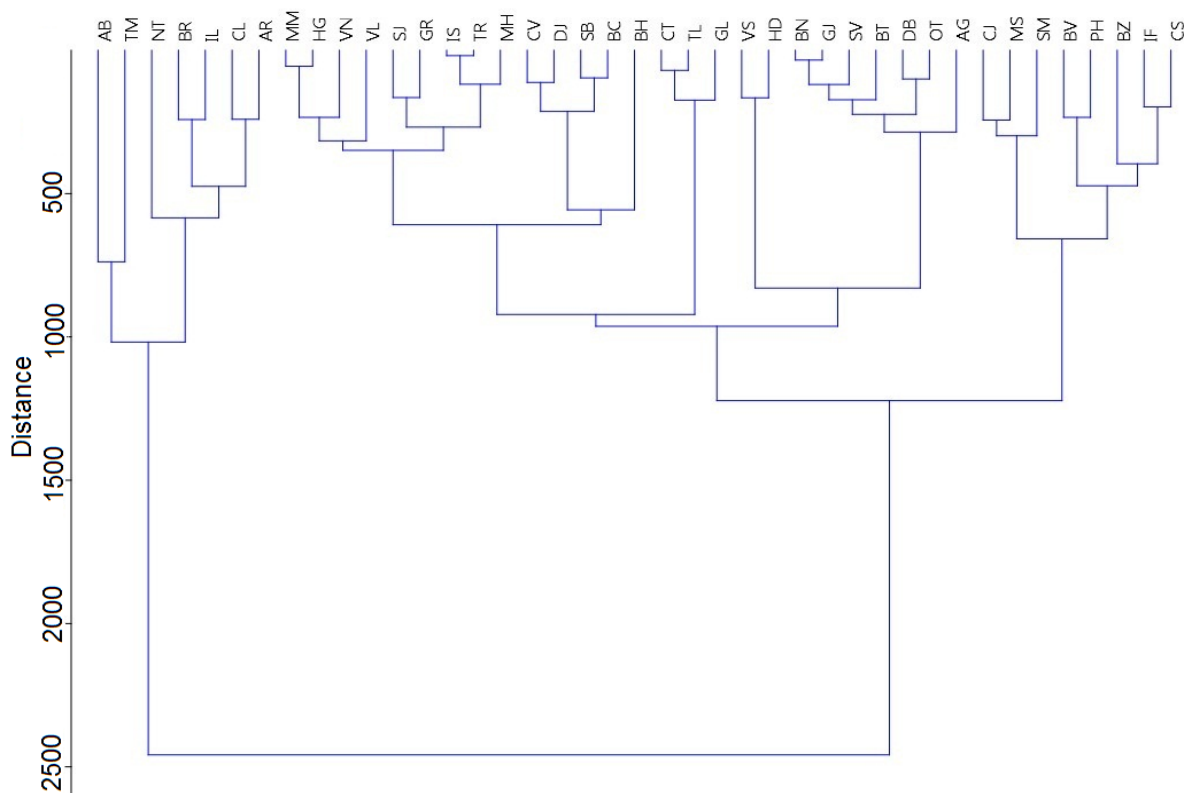
Figure 3. Clusters for the average production of wheat and corn in 2016 at county level in Romania

For proper coloring of conty maps having in sight the value of production, the related values of R,G,B can be determined as a image of function f(p)=ap+b, where p is production of a any county. The determination of coefficients a and b can be realized this way:

$$\begin{cases} f(p_{min}) = 0 \\ f(p_{max}) = 255 \end{cases}$$

$$\begin{cases} ap_{min} + b = 0 \\ ap_{max} + b = 255 \end{cases}$$

$$\Delta = \begin{vmatrix} p_{min} & 1 \\ p_{max} & 1 \end{vmatrix} = p_{min} - p_{max}, \Delta_a = \begin{vmatrix} 0 & 1 \\ 255 & 1 \end{vmatrix} = -255, \Delta_b = \begin{vmatrix} p_{min} & 0 \\ p_{max} & 255 \end{vmatrix} = 255p_{min}$$

$$\begin{cases} a = \dfrac{255}{p_{max} - p_{min}} \\ b = \dfrac{-255p_{min}}{p_{max} - p_{min}} \end{cases}$$

So R values for production, are given by:

$$R = f(p) = \frac{255}{p_{max} - p_{min}} \cdot p + \frac{-255p_{min}}{p_{max} - p_{min}}.$$

The values of G can be determined sorting in reverse order R values, and G values were considered null. So, in this way was determined the exact RGB values for every individual county in side, the data is shown in the table below.

Table 3.

RGB color indices for wheat and corn production

County/Wheat	R	G	B	County/Corn	R	G	B
Alba	139	0	43	Alba	219	0	9
Arad	224	0	21	Arad	186	0	37
Arges	0	0	255	Arges	37	0	186
Bacau	77	0	107	Bacau	110	0	68
Bihor	6	0	237	Bihor	112	0	60
Bistrita-Nasaud	29	0	180	Bistrita-Nasaud	40	0	183
Botosani	16	0	229	Botosani	44	0	150
Braila	229	0	16	Braila	208	0	29
Brasov	107	0	77	Brasov	89	0	81
Buzau	167	0	32	Buzau	104	0	75
Calarasi	237	0	6	Calarasi	173	0	40
Caras-Severin	120	0	64	Caras-Severin	113	0	53
Cluj	113	0	75	Cluj	140	0	46
Constanta	123	0	62	Constanta	48	0	134
Covasna	57	0	130	Covasna	107	0	69
Dambovita	21	0	224	Dambovita	29	0	208
Dolj	52	0	133	Dolj	100	0	76
Galati	145	0	33	Galati	53	0	113
Giurgiu	64	0	120	Giurgiu	79	0	89
Gorj	33	0	145	Gorj	40	0	173
Harghita	38	0	144	Harghita	68	0	110
Hunedoara	95	0	80	Hunedoara	0	0	255
Ialomita	255	0	0	Ialomita	206	0	30
Iasi	80	0	86	Iasi	76	0	100
Ilfov	140	0	39	Ilfov	118	0	51
Maramures	43		139	Maramures	69	0	104
Mehedinti	86	0	86	Mehedinti	69	0	107
Mures	137	0	47	Mures	134	0	48
Neamt	180	0	29	Neamt	183	0	40
Olt	32	0	167	Olt	30	0	206
Prahova	133	0	52	Prahova	89	0	79
Salaj	47	0	137	Salaj	82	0	82
Satu Mare	144	0	38	Satu Mare	150	0	44
Sibiu	75	0	113	Sibiu	104	0	69
Suceava	39	0	140	Suceava	46	0	140
Teleorman	80	0	95	Teleorman	75	0	104
Timis	191	0	26	Timis	255	0	0
Tulcea	130	0	57	Tulcea	51	0	118
Valcea	26	0	191	Valcea	81	0	89
Vaslui	86	0	80	Vaslui	9	0	219
Vrancea	62	0	123	Vrancea	60	0	112

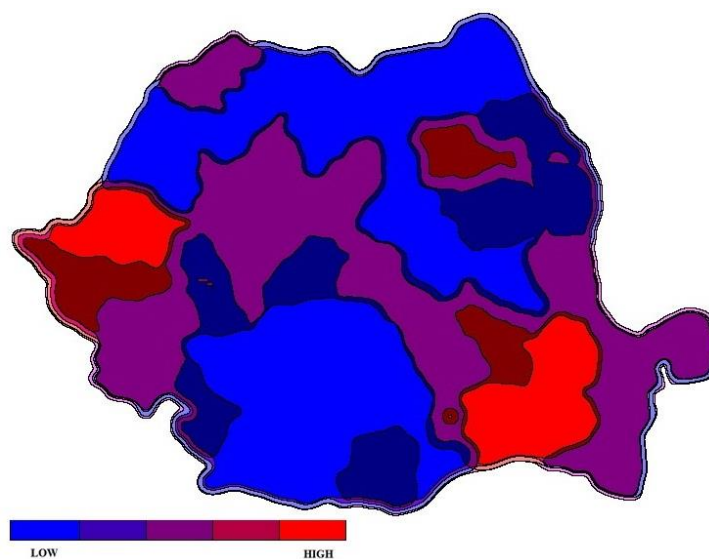*Source: our calculation*

Heat maps are given below.

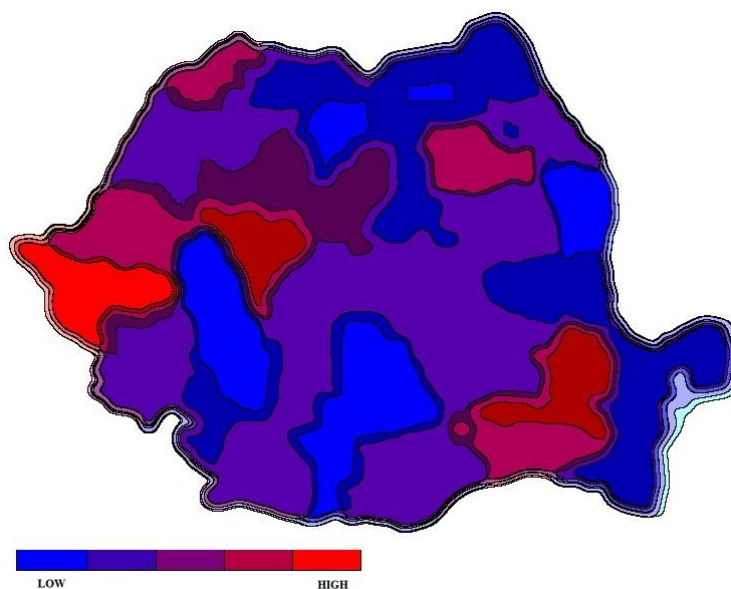**Figure 4. Heat map for average wheat production**

Figure 5. Heat map for average corn production

CONCLUSIONS

The cereal production is an important variable at the level of counties from Romania. More exactly for wheat there are average values from 2800 kg/ha to 5200 kg/ha and for corn there are values starting from 2400 kg/ha to 6500 kg/ha. Therefore, the steps to improve the production process of the cereals must be carried out in a differentiated manner, the measures to be applied are being necessary to be made or applied in a particular way to each area.

REFERENCES

[1]. **BAILEY K**. 1994, Numerical Taxonomy and Cluster Analysis Typologies and Taxonomies
[2]. **DODGE, Y.** (Ed.), 2006, The Oxford dictionary of statistical terms. Oxford University Press on Demand.

311

[3]. HIRSCH R., 2004, Exploring Colour Photography: A Complete Guide, Laurence King Publishing

[4]. HUNT R. W. G., 2004, The Reproduction of Colour (6th ed.). Chichester UK: Wiley–IS&T Series in Imaging Science and Technology

[5]. ROTARIU T. et al. 1999, Metode statistice aplicate în ştiinţele sociale, Editura Polirom 1999

[6]. RUJESCU C., FEHER ANDREA, RAIKOV MIROSLAV, SALA F., 2017, Comparative study on the evolution of average cereal production in the European Union, Lucrări Ştiinţifice Management Agricol, vol 19, no 1:101-106

[7]. RUJESCU CIPRIAN, Statistică matematică, Editura Artpress 2015

[8]. SALA F., RUJESCU C., FEHER ANDREA, 2016, Evolution scenario for the average wheat yield in the west of Romania Lucrări Ştiinţifice – vol. 59/2016, seria Agronomie, USAMV Iaşi, Facultatea de Agriculură

[9]. SALA, F., RUJESCU, C., CONSTANTINESCU, C., 2016, Causes and solutions for the remediation of the poor allocation of P and K to wheat crops in Romania. AgroLife Scientific Journal 5(1): 184-193.

[10]. VLAD, I. V., 2015, Strategia de dezvoltare a României în următorii 20 de ani, vol. I, Editura Academiei Române.

[11]. WILKINSON L., FRIENDLY M., 2009, The History of the Cluster Heat Map, The American Statistician, Vol. 63, No. 2

[12]. WILLIAMSON, D. F., PARKER, R. A., & KENDRICK, J. S., 1989. The box plot: a simple visual method to interpret data. Annals of internal medicine, 110(11), 916-921.

[13]. http://statistici.insse.ro/shop/ accessed at 05.09.2017

[14]. http://www.gretchenpeterson.com/blog/archives/2694

[15]. https://folk.uio.no/ohammer/past/, accessed 13.10.2017

[16]. http://ec.europa.eu/eurostat accessed at 05.09.2017