Do the means and distributions align with the SJR values of SCImago's economic subject category lists?

Imre Dobos[1] – Erzsébet Kovács[2] – Péter Sasvári[3]

ABSTRACT: Only a few sections of the Hungarian Academy of Sciences (HAS) maintain ranked journal lists, one of which is the IX Section of Economics and Law. The doctoral committees within this section assess candidates for the title Doctor of the Academy of Sciences based on eight lists of ranked international journals. These lists remain unchanged for a fixed period (three to five years). This study examines the extent to which the average SCImago Subject Ranking (SJR) values differ across subject categories within the economic subject areas of the SCImago database. To analyze these differences, we apply various models of analysis of variance (ANOVA), including Welch's ANOVA and the Kruskal-Wallis ANOVA.

KEYWORDS: scientometrics, list of journals, ANOVA
JEL-CODES: A20, A23, O31, O35, O38
DOI: https://doi.org/10.35551/PFQ_2025_2_1

Introduction

In the doctoral procedures of the Hungarian Academy of Sciences (HAS), a key aspect of evaluating candidates' scientific performance is the publication of articles in peer-reviewed journals that appear on specific, curated lists reflecting scientific judgment. Within the IX Section of Economics and Law (GJO) of HAS, doctoral and scientific committees compile these journal lists according to predefined procedures

1 Corresponding author, DSc, Professor, Budapest University of Technology and Economics, Faculty of Economic and Social Sciences, Budapest, dobos.imre@gtk.bme.hu
2 CSc, Professor Emerita, Corvinus University of Budapest, Budapest, erzsebet.kovacs@uni-corvinus.hu
3 PhD, Habilitated Associate Professor, University of Public Service, Faculty of Public Governance and International Studies, Budapest, and Associate Professor, University of Miskolc, Faculty of Mechanical Engineering and Informatics, Miskolc, sasvari.peter@uni-nke.hu

to determine whether a candidate's scientific and professional achievements meet the requirements for obtaining a doctoral title. The principles and methods used for reviewing international journal lists within the IX Section were introduced by Dobos et al. (2023). This study also addresses the statistical challenges encountered during the compilation of journal lists, particularly in estimating the SCImago Journal Ranking (SJR) indicator.

SJR serves as a measure of the prestige of academic journals, taking into account both the number of citations a journal receives and the prestige of the citing journals. The study describes the algorithm used to construct the journal list—which, in statistical terms, represents the database. Additionally, it presents a statistical method for assigning SJR values to journals that are not included in the SCImago database. Finally, it compares the current and newly compiled journal lists.

This study examines whether the distributional characteristics of journal lists derived from the SCImago database are identical across subfields or whether significant differences exist. This question is particularly important because the compilation of the journal list was based on the assumption that the average values of different subfields do not differ. Put more simply: does a researcher specializing in marketing have the same chance of publishing in a high-prestige journal as a researcher specializing in economics?

The ANOVA (Analysis of Variance) method, widely used in statistical analysis, is employed to determine whether the means of different populations or datasets are equal or nearly equal. ANOVA is frequently applied in various areas of economic research, such as analyzing digital indicators (Tarjáni et al., 2022), investigating innovation activities of food manufacturing companies (Erdei et al., 2021), assessing tourism security (Tokodi, 2023), and modeling corporate insolvency (Ágoston, 2022). These applications highlight the need for a deeper theoretical examination of the variance analysis methodology.

The next section of this study describes the dataset compilation process, followed by a presentation of ANOVA methods and applied post hoc analytical procedures. The results of the post hoc analyses performed on the dataset are then discussed. The study concludes with a summary of findings.

Compilation of the Database in the Field of Economic Sciences

The Doctoral Committee of Economic Sciences (GMDB) compiled the journal list using the freely accessible Scopus/SCImago database (SCImago, n.d.).

The journal list is based on the SCImago database, which is updated annually in May; for this study, we considered the May 2022 version. New journals may be added to the list, while others may be removed if they fail to meet strict quality standards, leading to annual fluctuations in the number of listed journals.

Within this database, economic science journals are classified into two main subject areas, which are further divided into specific subject categories.

The study covers the following two subject areas:

- Economics, Econometrics, and Finance (EEF)
- Business, Management, and Accounting (BMA)

The two main subject areas contain 1,111 and 1,427 journals, meaning that the SCImago database includes 1,111 journals related to economics and 1,427 journals related to business and management. Naturally, some journals are linked to both subject areas, accepting articles in both economics and business disciplines. The number of such overlapping journals is 410, resulting in a total of 2,341 journals available in the economic sciences field at the time of data collection.

The two subject areas are further divided into thirteen subject categories, which are presented in Table 1, along with the number of journals in each category. It is important to note that the journal counts in the table cannot be summed directly, as some journals belong to multiple subject categories and, in some cases, even to multiple subject areas.

1. Table: Number of Scopus/SCImago business journals per subject

Subject Area	Nr.	Subject Category	Cleaned Journals (count)	Total Journals (count)	Cleaned/ Total
Business, Management and Accounting	1	Accounting	124	159	77,99
	2	Business and International Management	403	423	
	3	Business, Management and Accounting (miscellaneous)	269	343	95,27
	4	Industrial Relations	48	59	
	5	Management Information Systems	93	116	78,43
	6	Management of Technology and Innovation	134	262	81,36
	7	Marketing	91	199	80,17
	8	Organizational Behavior and Human Resource Management	109	207	

Subject Area	Nr.	Subject Category	Cleaned Journals (count)	Total Journals (count)	Cleaned/ Total
Business, Management and Accounting	9	Strategy and Management	143	458	51,15
	10	Tourism, Leisure and Hospitality Management	94	126	45,73
Economics, Econometrics and Finance	11	Economics and Econometrics	449	694	
	12	Economics and Econometrics (miscellaneous)	318	407	52,66
	13	Finance	66	298	31,22

Source: Own compilation based on SCImago data.

The journals associated with each subject category were filtered to ensure that each journal was assigned to only one category. This step was necessary because the applied methodology requires that the thirteen subject categories form disjoint sets, meaning that no journal could be included in multiple categories.

A key question in this process was determining the principles for assigning journals to specific subject categories. To address this, we applied the rule that each journal was assigned to the subject category in which it had the highest quartile ranking. If a journal's assignment could not be clearly determined, the categorization was performed randomly. This step was essential because, in all types of ANOVA analysis, data independence and the formation of disjoint sets are fundamental requirements.

It is important to note that the SCImago database used in this study represents the full population, as all journals within the field of economic sciences were considered in the analysis.

Types of Analysis of Variance (ANOVA) and Application Conditions

The term Analysis of Variance (ANOVA) does not fully reflect the primary objective of the method. Its main function is not merely to analyze variances but rather to compare the means of different datasets or populations. The various ANOVA methods and their statistical characteristics are summarized in Table 2.

2. Table: Summary of ANOVA methods

	Non-Parametric ANOVA (Ordinal Scale)	Parametric ANOVA (Interval and/or Ratio Scale)		
	Kruskal-Wallis ANOVA	Fisher's ANOVA (Traditional, Homoscedastic)	Welch's ANOVA (Heteroscedastic)	Generalized ANOVA
Independence	Yes	Yes	Yes	Yes
Normality	No	Yes	Yes	No
Homogeneity	No	Yes	No	No

Source: Own compilation.

Table 2 illustrates the application conditions of the traditional ANOVA model, which can be categorized as follows:

1. Independence: Populations must be independent and disjoint, following the same distribution. This means that each observation or data point belongs to only one population, ensuring that elements in the database are independent of each other.
2. Normality: The distribution of the dependent variables must be (1) normal, (2) approximately normal, or (3) at least symmetric in all examined populations.
   Note: If the sample size exceeds 25, normality testing is generally not required.
3. Homogeneity: The variance of independent variables must be equal across all populations.
   - If the population sizes are approximately equal, variance homogeneity testing can be omitted (see Field, 2024).
   - Variance homogeneity testing is also unnecessary if the population standard deviations do not correlate with the means.

This condition, namely the lack of correlation between population standard deviations and means, implies that the covariance between the two variables is close to zero. Consequently, this also suggests that the means and/or standard deviations are approximately centered around zero.

Based on the points discussed so far, it is clear that normality assumptions can be disregarded if certain distributional properties hold or if the sample sizes exceed a certain threshold. Similarly, one of the key assumptions of the classical ANOVA model, variance homogeneity (i.e., equal standard deviations across populations), can also be relaxed if the sample sizes are approximately equal or if there is no correlation between means and standard deviations.

These considerations indicate that the assumptions of the traditional ANOVA model can be widely relaxed, significantly expanding the applicability of the ANOVA test and the F-test. As a result, these methods can be more flexibly used to compare population means.

For assessing normality, various methods are available, such as the Kolmogorov–Smirnov statistic, although other techniques can also be employed. The SPSS28 software package includes multiple tests for examining variance equality. If any of these assumptions are violated, alternative statistical methods can be applied.

If variance homogeneity is not met, an alternative to the traditional Fisher's ANOVA is the Welch ANOVA, which is specifically designed for heteroscedastic conditions (i.e., cases where population variances differ). Although the concepts of homoscedasticity and heteroscedasticity are primarily used in the context of regression model residuals, they are also relevant in ANOVA applications. In such cases, SPSS replaces the classical unweighted F-test with the Levene-weighted F-test to account for unequal variances.

If neither normality nor variance homogeneity is satisfied, two possible solutions exist. One approach is to use simulation methods to approximate the distribution of the F-test, which can then be applied to determine whether population means are equal. In the literature, this is often referred to as a generalized F-test, where the distribution is derived from simulation techniques (for more details, see Desi, 2022; Islam & Abbas, 2022; Lantz, 2013).

If the assumptions of traditional ANOVA are violated, a non-parametric test can be employed. In such cases, the comparison is conducted on a lower, ordinal scale, such as with the Kruskal–Wallis ANOVA, which is based on the H-test. This method compares the entire dataset by evaluating the median rank values of populations rather than their means.

These approaches allow us to determine whether population means differ from the overall dataset mean. However, these tests do not identify which population means are specifically equal to each other—a question that will be addressed in the next section.

Post Hoc Tests in ANOVA

After performing ANOVA, there are two main types of post hoc analyses. The first method involves pairwise comparisons of population means using the t-test or its modified versions, conducting as many pairwise comparisons as the number of populations included in the analysis. This approach is referred to in the literature as multiple comparisons, and it is also supported by SPSS28. However, a major drawback of this method is that pairwise comparisons are not transitive, meaning that the results are not always consistent, an issue that will be discussed in detail later.

The second method focuses on identifying homogeneous subsets, grouping populations whose means are statistically indistinguishable. This approach results in the formation of three main subsets: populations with lower, equal, or higher means. While this method helps identify statistically similar groups, it does not provide information on the magnitude of differences between population means.

To determine homogeneous subsets, an iterative algorithm can be applied. The first step involves sorting the population means in ascending order and conducting an ANOVA test to check whether the means are statistically equivalent. If ANOVA does not indicate significant differences, the algorithm terminates, and the population means are considered identical. If ANOVA rejects the equality of means, the next step involves dividing the sample into two groups by removing the smallest and then the largest mean, followed by another ANOVA test on the remaining populations. If the ANOVA test does not reject mean equality for any of the newly formed groups, the procedure ends. However, if significant differences persist, the process continues by further eliminating populations from the dataset until statistically homogeneous groups are identified.

This iteration continues until:

- the null hypothesis is rejected, or
- we obtain sequences where the iteration cannot proceed further because the null hypothesis cannot be rejected, or
- no more populations remain for analysis.

The advantage of this method is that it can identify populations with equal means. However, its drawback is that it often results in multiple homogeneous subsets, making it more challenging to select the most appropriate solution.

Table 3 presents a comparison of the possible post hoc tests. It is important to note that most of the tests listed in the table are available in SPSS28, except for the Ryan-Einot-Gabriel-Welch (R-E-G-W) Q, Waller-Duncan, and Dunnett tests.

3. Table: Comparison of ANOVA Post Hoc tests

Post Hoc Test	Comparison s	Homogeneity	Population Size
Duncan	Ordered Means	Yes	Equal
Student-Newman-Keuls (SNK)	Ordered Means	Yes	Equal
Tukey HSD	Ordered Means	Yes	Equal
Gabriel	Ordered Means	Yes	
Hochberg	Ordered Means	Yes	
Ryan-Einot-Gabriel- Welch (R-E-G-W) F	Ordered Means	Yes	
Tukey b	Ordered Means	Yes	
Bonferroni Test for Minimum Significant Difference	Mean Pairs	Yes	Different
Scheffé	Mean Pairs	Yes	Different
Least Significant Difference (LSD)	Mean Pairs	No	Different
Šidák	Mean Pairs	No	
Tamhane T_2	Mean Pairs	No	

Post Hoc Test	Comparison s	Homogeneity	Population Size
Games-Howell	Mean Pairs	No	
Dunnett's T_3	Mean Pairs	No	For Smaller Populations
Dunnett's *C*	Mean Pairs	No	For Larger Populations

Source: Own compilation based on Wikipedia (n.d.).

The third and fourth columns of the table represent variance homogeneity. As emphasized in the ANOVA methodology discussion, the assumption of equal variances across populations and/or nearly equal sample sizes ensures that the traditional ANOVA method can be applied alongside the F-test.

In Table 3, when using SPSS28, it is assumed that variance homogeneity holds for the first eleven tests, while for the four tests highlighted in gray, this assumption is not made. The homogeneity of variances can be tested using the Levene test, and based on its results, either a post hoc test assuming equal variances can be applied, or alternatively, a post hoc test that does not assume variance homogeneity should be used.

Comparison of Mean SJR Values Across Subject Category Groups Using Post Hoc Analysis

When presenting the dataset, we already indicated that the database was made independent by ensuring that each journal was assigned to only one population. As a result, two main conditions required for the application of ANOVA—normality and homoscedasticity—needed to be examined.

To check for normality, we applied the Kolmogorov–Smirnov statistic. Using SPSS28, we calculated whether the deviation from normality was statistically significant for each economic subject category and for the entire dataset. Although the results indicated that the distributions did not strictly follow normality, the ANOVA and F-test remained applicable since, in every subject category, the number of journals exceeded 25. Generally, this is sufficient to disregard the normality assumption.

The homogeneity of variances was examined using the Levene test, with its results summarized in Table 4. The test revealed that the assumption of homoscedasticity was not met, meaning that the population variances differed significantly. Consequently, we also examined whether the population sizes were approximately equal. Given the significant variation in population sizes, we concluded that the assumption of variance homogeneity could be disregarded. This was further supported by a strong correlation (0.810) between the means and standard deviations of the 13 subject categories, indicating a high degree of association between these two metrics.

4. Table: The result of Levene's test

Variance Homogeneity Test					
		Levene Statistic	df1	df2	Significance (p-value)
SJR	Mean-Based	6,777	12	2328	< 0,001
	Median-Based	3,336	12	2328	< 0,001
	Median and Adjusted Degrees of Freedom-Based	3,336	12	1263,1301	< 0,001
	Trimmed Mean-Based	4,208	12	2328	< 0,001

Source: Own compilation based on SPSS28.

Based on the above results, the classical Fisher's ANOVA cannot be applied, as the assumption of variance homogeneity is not met. Therefore, four possible alternatives remain.

However, one of these, generalized ANOVA, is not a viable option, as its applicability is still a subject of academic debate, and SPSS does not support its direct implementation. As a result, the further analyses in this study will be conducted using one of the three remaining methods.

Application of the Welch ANOVA Model

Among the three possible alternatives, the first is Welch's ANOVA. In this case, the assumption of normality, or an equivalent condition—a large population size—is met, while homoscedasticity is not. Since Welch's ANOVA does not require variance homogeneity, this method is suitable for our analysis.

Table 5 presents the Welch ANOVA results based on the analysis conducted in SPSS. The findings indicate that the mean SJR values of the subject category journal lists differ significantly, meaning that they cannot be considered statistically identical.

5. Table: Results of Welch's ANOVA

Robust Test for Equality of Means				
SJR				
	Statistic[a]	df1	df2	Significance (p-value)
Welch	4,691	12	571,240	< 0,001

a. Asymptotically F-distributed.

Source: Own compilation based on SPSS28.

Since the means are not equal, the next question is which subject category lists have statistically similar mean values. To determine this, we apply the previously introduced post hoc analysis.

SPSS's one-way ANOVA analysis provides four possible post hoc tests, which were outlined in Table 3. However, these methods only allow for pairwise comparisons and are not suitable for identifying homogeneous subsets.

Among the four methods, we first examine the pairwise comparison matrix of the Games–Howell test, presented in Table 6. We selected this test because it is the most commonly recommended method in both academic literature and practical applications when variance homogeneity is not assumed.

The Games–Howell matrix is symmetric, and subject categories with significantly different means are highlighted in gray in the table. Subject categories not highlighted can be considered statistically similar in terms of mean SJR values.

6. Table: Games-Howell pairwise comparison's significance levels

Subject Categories	12	13	3	4	10	5	6	2	9	7	8	11	1
12		1,000	0,998	1,000	0,674	0,569	0,326	0,089	0,217	0,064	0,076	**0,042**	**0,001**
13	1,000		0,999	1,000	0,694	0,596	0,342	0,089	0,228	0,066	0,082	**0,044**	**0,001**
3	0,998	0,999		1,000	0,980	0,888	0,724	0,363	0,575	0,249	0,171	0,146	**0,007**
4	1,000	1,000	1,000		1,000	0,996	0,991	0,976	0,977	0,906	0,417	0,678	0,246
10	0,674	0,694	0,980	1,000		1,000	1,000	0,997	0,997	0,959	0,514	0,751	0,201
5	0,569	0,596	0,888	0,996	1,000		1,000	1,000	1,000	1,000	0,894	0,998	0,890
6	0,326	0,342	0,724	0,991	1,000	1,000		1,000	1,000	1,000	0,870	0,996	0,828
2	0,089	0,089	0,363	0,976	0,997	1,000	1,000		1,000	1,000	0,857	0,995	0,769
9	0,217	0,228	0,575	0,977	0,997	1,000	1,000	1,000		1,000	0,912	0,999	0,900
7	0,064	0,066	0,249	0,906	0,959	1,000	1,000	1,000	1,000		0,960	1,000	0,968
8	0,076	0,082	0,171	0,417	0,514	0,894	0,870	0,857	0,912	0,960		0,999	1,000

Subject Categories	12	13	3	4	10	5	6	2	9	7	8	11	1
11	**0,042**	**0,044**	0,146	0,678	0,751	0,998	0,996	0,995	0,999	1,000	0,999		1,000
1	**0,001**	**0,001**	**0,007**	0,246	0,201	0,890	0,828	0,769	0,900	0,968	1,000	1,000	

Source: Own compilation based on SPSS28.

The results clearly show that the mean SJR values of subject categories 2, 4, 5, 6, 7, 8, 9, and 10 are statistically similar based on pairwise comparisons. Moreover, these eight subject categories have mean SJR values that are statistically identical to all other subject categories, as indicated by the shaded areas in the matrix.

If pairwise comparisons formed a transitive relationship, then the means of all subject categories would be equal. However, the results reveal that there are five pairwise comparisons where the means differ significantly. This contradiction suggests that pairwise comparisons alone are not suitable for clearly distinguishing homogeneous subsets.

An alternative approach is to construct homogeneous subsets directly. Since SPSS does not generate homogeneous subsets in heteroscedastic cases, we performed the classification manually by sequentially eliminating subject categories. To achieve this, we first arranged the mean SJR values in ascending order, as presented in Table 7.

7. Table: Means and confidence intervals in ascending order

				Mean SJR Values and 95% Confidence Intervals by Subject Category		
Nr.	**Subject Category**	**Mean**	**Standard Deviation**	**Lower Bound**	**Upper Bound**	**Count**
12	Economics and Econometrics (miscellaneous)	0,498	1,036	0,384	0,612	318
13	Finance	0,508	0,424	0,404	0,613	66
3	Business, Management and Accounting (miscellaneous)	0,573	0,742	0,484	0,662	269
4	Industrial Relations	0,593	0,844	0,348	0,839	48
10	Tourism, Leisure and Hospitality Management	0,681	0,635	0,550	0,811	94
5	Management Information Systems	0,777	1,069	0,557	0,997	93
6	Management of Technology and Innovation	0,779	1,067	0,597	0,962	134
2	Business and International Management	0,787	1,446	0,646	0,929	403

				Mean SJR Values and 95% Confidence Intervals by Subject Category		
Nr.	**Subject Category**	**Mean**	**Standard Deviation**	**Lower Bound**	**Upper Bound**	**Count**
9	Strategy and Management	0,802	1,105	0,619	0,985	143
7	Marketing	0,842	0,837	0,668	1,017	91
8	Organizational Behavior and Human Resource Management	0,944	1,218	0,713	1,175	109
11	Economics and Econometrics	1,048	2,399	0,826	1,271	449
1	Accounting	1,162	2,202	0,770	1,553	124
	Összesen	**0,787**	**1,510**	**0,725**	**0,848**	**2341**

Source: Own compilation based on SPSS28.

To determine the homogeneous subsets, we applied the algorithm presented in Table 8. The numbers in the table represent the subject category numbers in the order listed in Table 7.

In the initial step (Step 0), the mean values of the subject categories were arranged in ascending order, and a Welch ANOVA test was conducted to determine whether the means were statistically identical. The results indicated that the means could not be considered equal, meaning that at least one subject category had a mean significantly different from the others.

In Table 8, this result is highlighted in a light color, indicating that a statistically significant difference exists between the means.

In the first step, we removed one subject category from both the beginning and the end of the sequence, reducing the dataset to 11 subject categories, and then performed the Welch test on this revised set.

The rationale behind this approach is that the subject categories with the lowest and highest means are most likely to differ significantly from the others.

The results of the new test indicated that at least one subject category still had a mean that was significantly different from the rest.

8. Table: Algorithm for determining homogeneous sets

	SJR Means in Ascending Order												
Step	0,498	0,508	0,573	0,593	0,681	0,777	0,779	0,787	0,802	0,842	0,944	1,048	1,162
0.	12	13	3	4	10	5	6	2	9	7	8	11	1
1.	12	13	3	4	10	5	6	2	9	7	8	11	
		13	3	4	10	5	6	2	9	7	8	11	1

SJR Means in Ascending Order													
Step	0,498	0,508	0,573	0,593	0,681	0,777	0,779	0,787	0,802	0,842	0,944	1,048	1,162
2.	12	13	3	4	10	5	6	2	9	7	8		
		13	3	4	10	5	6	2	9	7	8	11	
			3	4	10	5	6	2	9	7	8	11	1
3.	12	13	3	4	10	5	6	2	9	7			
		13	3	4	10	5	6	2	9	7	8		
			3	4	10	5	6	2	9	7	8	11	
				4	**10**	**5**	**6**	**2**	**9**	**7**	**8**	**11**	**1**
4.	12	13	3	4	10	5	6	2	9				
		13	3	4	10	5	6	2	9	7			
			3	4	10	5	6	2	9	7	8		
5.	12	13	3	4	10	5	6	2					
		13	3	4	10	5	6	2	9				
			3	4	10	5	6	2	9	7			
6.	12	13	3	4	10	5	6						
		13	3	4	10	5	6	2					
			3	**4**	**10**	**5**	**6**	**2**	**9**				
7.	**12**	**13**	**3**	**4**	**10**	**5**							
		13	**3**	**4**	**10**	**5**	**6**						

Source: Own compilation based on SPSS28.

In the second step, we removed two elements from the sequence in different ways. First, we excluded the two largest elements, then one element from both the beginning and the end, and finally, we removed the two smallest elements. The three newly formed sequences were then subjected to the Welch ANOVA test again. The results indicated that the means were still not equal, prompting us to continue with the algorithm.

In the third step, we excluded three elements from the sequence in various combinations. The results showed that for three out of the four sequences examined, the means remained significantly different. However, in the fourth sequence, there was no significant difference, suggesting that the assumption of mean equality could be accepted.

Before proceeding to the next step, we must consider that if we remove even one more element from either the beginning or the end of this ten-element sequence, the remaining sequence will still show equal means. This means that the alternative hypothesis is not rejected, confirming the equality of means within this subset.

As a result, it is sufficient to examine only those mean sequences where the first element starts from one of the first three means. This simplification reduces the number of necessary combinations, making further analysis more efficient.

9. Table: The four homogeneous subsets with means

		Homogeneous Subsets			
Nr.	**Subject Category**	**1.**	**2.**	**3.**	**4.**
12	Economics and Econometrics (miscellaneous)	0,498			
13	Finance	0,508	0,508		
3	Business, Management and Accounting (miscellaneous)	0,573	0,573	0,573	
4	Industrial Relations	0,593	0,593	0,593	0,593
10	Tourism, Leisure and Hospitality Management	0,681	0,681	0,681	0,681
5	Management Information Systems	0,777	0,777	0,777	0,777
6	Management of Technology and Innovation		0,779	0,779	0,779
2	Business and International Management			0,787	0,787
9	Strategy and Management			0,802	0,802
7	Marketing			0,842	0,842
8	Organizational Behavior and Human Resource Management				0,944
11	Economics and Econometrics				1,048
1	Accounting				1,162
	Welch Significance Levels	**0,105**	**0,053**	**0,069**	**0,070**
	Fisher Significance Levels	0,102	0,079	0,230	0,153

Source: Own compilation based on SPSS28.

Based on the above results, in the fourth and fifth steps, we examined only three sequences at a time, each of which had significantly different means.

In the sixth step, after removing a total of six means from the beginning and the end of the sequence, the results indicated that in two cases, the means still differed significantly, while in the third case, mean equality could be accepted. This suggests that from the remaining subset, we could remove at most two additional subject categories without causing a statistically significant difference in means.

In the seventh and final step, for both remaining sequences, the null hypothesis of mean equality was not rejected, meaning that the algorithm was successfully concluded. The final homogeneous subsets are presented in Table 9.

Interpretation of the Results:

- First Homogeneous Subset: The first six subject categories have statistically identical SJR values, whereas the subsequent categories have significantly higher mean SJR values compared to these first six.
- Fourth Homogeneous Subset: The ten subject categories in this group exhibit statistically similar mean values, but the first three categories in the dataset have significantly lower mean SJR values than the remaining ten categories.
- Second and Third Homogeneous Subsets: These groups contain both lower and higher mean journal categories, meaning that they cannot be clearly separated into distinct clusters.

We compared the results of the classic Fisher ANOVA and the Welch ANOVA. The comparison revealed that Fisher's ANOVA tends to overestimate the significance of mean equality, leading to a higher likelihood of Type I errors when variance homogeneity is not met.

However, both ANOVA methods ultimately produced similar overall results, reinforcing the validity of using Welch's ANOVA in cases where the assumption of variance homogeneity is violated.

Visual Methods for Post Hoc Analysis

The t-test can also be used to determine homogeneous subsets, as it is based on the calculation and comparison of confidence intervals of the means. To apply this method, we need to know the mean values of all 13 subject categories, as well as their 95% confidence intervals, which are provided in Table 7 based on SPSS standard calculations.

One approach to defining homogeneous subsets is to plot the lower and upper confidence interval bounds of the sorted SJR means. We then look for values between these boundaries that overlap with the highest number of subject category confidence intervals, meaning they remain within a common range for the longest time across different categories.

Although it is difficult to establish a precise algorithm for this procedure, a possible approach is to first consider the smallest upper bound and examine how long this constant remains within the lower and upper confidence limits. According to Table 10, this condition holds for the first six subject categories, which means they form a homogeneous subset.

This result matches the homogeneous subset obtained using Welch ANOVA, confirming that both methods lead to the same conclusion regarding the classification of subject categories.

In the second step, we considered the largest lower bound and examined how long this value remained within the confidence interval range. According to Table 10, in this case, eight subject categories remained within the examined range. However, this result differs from the Post hoc analysis of the Welch ANOVA.

The discrepancy arises because the upper bound for the Tourism, Leisure and Hospitality Management subject category was 0.811, which is lower than the determined largest lower bound of 0.826. At the same time, the fourth-highest mean subject category had an upper limit that would have allowed it to remain within the range.

As a result, Welch ANOVA produced a longer homogeneous subset than the t-test-based method, but the difference between the two methods remains minimal. This suggests that while both approaches generally lead to consistent conclusions, slight variations may arise due to the different statistical assumptions underlying each method.

Finally, to determine the final homogeneous subsets, we needed to identify the constant value that falls between the lowest upper bound and the highest lower bound, ensuring that this value remains within these two limits throughout. The two extreme values were 0.612 and 0.826.

Within this range, the interval between 0.662 and 0.811 covered the largest number of subject categories. This interval spanned from the third and fifth highest subject categories to the ninth and eleventh highest, forming a homogeneous subset of seven subject categories.

As a final conclusion, the analysis confirmed that no additional homogeneous subsets could be identified, thus concluding the procedure.

10. Table: Three homogeneous subsets of visual analysis

				Final Homogeneous Subsets				
Nr.	**Subject Category**	**Mean**	**Lower Bound**	**1.**	**2.**	**3.**	**4.**	**Upper Bound**
12	Economics and Econometrics (miscellaneous)	0,498	0,384	0,612				0,612
13	Finance	0,508	0,404	0,612				0,613
3	Business, Management and Accounting (miscellaneous)	0,573	0,484	0,612	0,662			0,662
4	Industrial Relations	0,593	0,348	0,612	0,662			0,839
10	Tourism, Leisure and Hospitality Management	0,681	0,550	0,612	0,662	0,811		0,811
5	Management Information Systems	0,777	0,557	0,612	0,662	0,811	0,826	0,997
6	Management of Technology and Innovation	0,779	0,597	0,612	0,662	0,811	0,826	0,962
2	Business and International Management	0,787	0,646		0,662	0,811	0,826	0,929
9	Strategy and Management	0,802	0,619		0,662	0,811	0,826	0,985

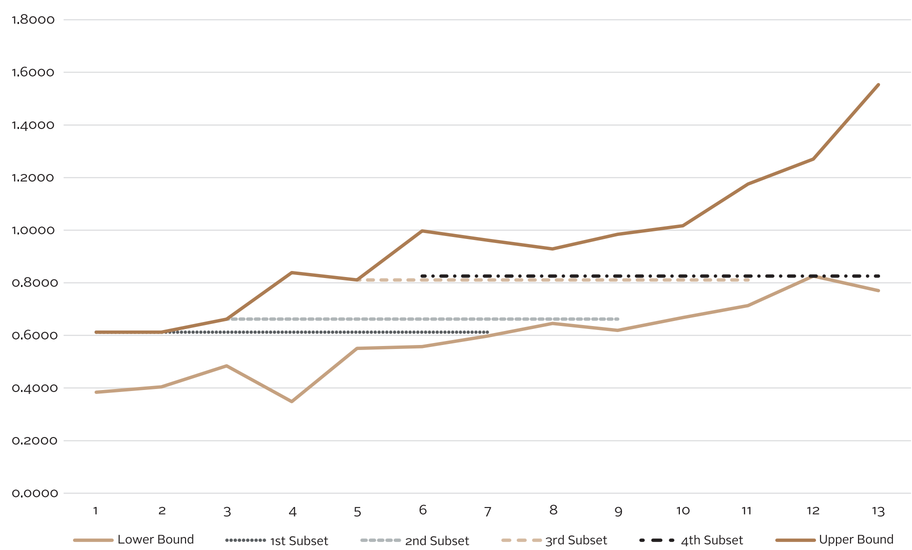
				Final Homogeneous Subsets				
Nr.	**Subject Category**	**Mean**	**Lower Bound**	**1.**	**2.**	**3.**	**4.**	**Upper Bound**
7	Marketing	0,842	0,668			0,811	0,826	1,017
8	Organizational Behavior and Human Resource Management	0,944	0,713			0,811	0,826	1,175
11	Economics and Econometrics	1,048	0,826				0,826	1,271
1	Accounting	1,162	0,770				0,826	1,553

Source: Own compilation based on SPSS28.

The visualization of the results is illustrated in Figure 1. The figure clearly shows that in the second homogeneous subset, the constant value falls within the range of 0.662 to 0.811, meaning it lies between a lower and an upper boundary. Naturally, similar lower and upper boundary values can be assigned to the other two homogeneous subsets as well. However, the aforementioned values are easier to quantify and effectively represent the homogeneous subsets identified based on the results.

The visual representation in the figure helps to further clarify the differences and overlaps between the subject categories, making them easier to interpret.

1. Figure: Confidence intervals of the averages obtained with the t-test and their comparison



Source: Own compilation.

The visualization can also be applied to the results obtained using the Kruskal-Wallis ANOVA method, which will be presented in the following section.

7. Post Hoc Analysis for Kruskal-Wallis ANOVA

In the case of Kruskal-Wallis ANOVA, the analysis transitions from a higher measurement scale to a lower, categorical/ordinal scale. Accordingly, we apply the Kruskal-Wallis H-test, which operates based on the mean ranks of the data.

The results of the analysis indicated that the mean ranks and their distribution differ significantly, necessitating a Post hoc analysis for further investigation.

Similar to the Welch ANOVA, we excluded pairwise comparisons in the Post hoc analysis, as the lack of transitivity prevents the establishment of a clear ranking order. Instead, we examined the homogeneous subsets generated by SPSS, with the results presented in Table 11.

The results of the analysis are very similar to those presented in Tables 9 and 10, suggesting that there is no significant difference between the three Post hoc analyses applied. This also implies that in this case, choosing only one Post hoc method for identifying differences between subject categories does not lead to major errors.

Since the Kruskal-Wallis ANOVA is based on medians, the values in Table 11 represent the medians of the subject categories rather than their means. As a result, the homogeneous subsets obtained through this method are based on the central tendencies of the data distribution, which may differ from the results of the other two ANOVA methods. However, the overall conclusions remain consistent across all three approaches.

11. Table: Three homogeneous subsets of Kruskal-Wallis ANOVA

Homogeneous Subsets			
	Homogeneous Subsets		
	1	2	3
Economics and Econometrics (miscellaneous)	960,093		
Industrial Relations	1080,281	1080,281	
Business, Management and Accounting (miscellaneous)		1093,539	
Business, Management and Accounting		1118,136	
Finance		1160,439	1160,439
Management Information Systems		1168,048	1168,048
Management of Technology and Innovation		1184,940	1184,940
Strategy and Management		1238,706	1238,706
Economics and Econometrics			1269,640
Marketing			1282,187

Homogeneous Subsets			
	Homogeneous Subsets		
	1	2	3
Accounting			1298,468
Tourism, Leisure and Hospitality Management			1306,048
Organizational Behavior and Human Resource Management			1355,197
Test Statistics	1,162	5,938	8,560
Kruskal-Wallis Two-Sided Significance	**0,281**	**0,430**	**0,381**

Source: Own compilation based on SPSS28. The homogeneous subsets are determined based on asymptotic significance. The significance level is 0.05.

Conclusion

In this study, we compared the mean SJR values of 13 subject categories within the two main economic fields of the SCImago journal list, aiming to determine whether significant differences exist between subject areas.

Our findings indicate that there are substantial differences among the economic subject categories in international journal rankings. Specifically, we identified eight subject categories where the mean SJR values were statistically similar, meaning there was no significant difference between them. In contrast, the remaining five categories exhibited statistically distinct means, with three having below-average SJR values and two having above-average values. The lowest mean SJR was observed in Economics and Econometrics (miscellaneous), while the highest mean SJR was found in Accounting.

The SJR metric primarily serves to indicate which quartile (Q) a journal belongs to within a given subject category. Journals are classified into four quartiles based on their SJR values:

- Q1 (0-25%) – Top-ranked journals
- Q2 (25-50%) – Upper-middle tier
- Q3 (50-75%) – Lower-middle tier
- Q4 (75-100%) – Bottom-ranked journals

In subject categories with lower average SJR values, researchers can achieve higher quartile rankings (Q1 or Q2) with relatively lower SJR scores. Conversely, in categories with higher average SJR values, attaining a Q1 or Q2 classification requires a significantly higher SJR score.

This distinction is particularly important because in some universities undergoing model transitions, research funding is directly linked to the quartile ranking (Q classification) of published articles. Consequently, understanding disciplinary differences in journal impact rankings is crucial for academic evaluation and funding allocation.

This study contributes to a better understanding of the ranking of academic journals, particularly in the field of economics. The results highlight significant differences in SJR values across subject categories, which can impact researchers' publication strategies and university rankings. Given that some universities' funding models are partially based on Q-category classifications, a more accurate classification can support strategic decision-making for both institutions and researchers. ■

References

1. Ágoston N. (2022): Mesterséges intelligencia és gépi tanulási módszerek a vállalati fizetésképtelenség becslésére= Artificial intelligence and machine learning methods to estimate firm insolvency. Statisztikai Szemle, 100(6), 584-609. https://doi.org/10.20311/stat2022.6.hu0584
2. Desai, T. A. (2022): One-Way ANOVA with Possible Heteroscedasticity of Variances and Non-Normality of Errors. In Important Applications of the Behrens-Fisher Statistic and the False Discovery Rate (pp. 3-13). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-99888-2_2
3. Dobos I. – Halmai P. – Sasvári P. (2023): A nemzetközi folyóiratlista összeállítása az MTA IX. Osztály Gazdaságtudományi Doktori Minősítő Bizottságában (GMDB). Pénzügyi Szemle, 23(2), 67-80. https://doi.org/10.35551/PFQ_2023_2_4
4. Erdei E. – Popp, J. – Neményi M. – Oláh J. (2021): Az Ipar 4.0 technológiák szerepe az élelmiszergyártó vállalatok innovációs tevékenységében. Statisztikai Szemle, 99(10), 978-996. https://doi.org/10.20311/stat2021.10.hu0978
5. Field, A. (2024): Discovering statistics using IBM SPSS statistics. Sage Publ. Ltd.
6. Islam, T. U. – Abbas, E. (2022): Validity of ANOVA under Non-normality & Heterogeneity. https://doi.org/10.21203/rs.3.rs-2071136/v1
7. Lantz, B. (2013): The impact of sample non-normality on ANOVA and alternative methods. British Journal of Mathematical and Statistical Psychology, 66(2), 224-244. https://doi.org/10.1111/j.2044-8317.2012.02047.x
8. SCImago, (n.d.). SJR — SCImago Journal & Country Rank [Portal]. http://www.scimagojr.com, letöltés: 2023. február 13.
9. Tarjáni A. J – Kalló N. – Dobos I. (2022): A nemzetközi digitális gazdaság és társadalom index 2020. évi adatainak statisztikai elemzése= Statistical analysis of 2020 data for the international digital economy and society index. Statisztikai Szemle, 100(3), 266-284. https://doi.org/10.20311/stat2022.3.hu0266
10. Tokodi P. (2023): A turizmus félelemföldrajza A belföldi turisták biztonságérzetének dimenziói= The fear geography of tourism–dimensions of domestic tourists' perception of safety. Statisztikai Szemle, 101(1), 53-74. https://doi.org/10.20311/stat2023.01.hu0053
11. Wikipedia, (n.d.): Post-hoc-Test, https://de.wikipedia.org/wiki/Post-hoc-Test, letöltés: 2023. február 13.