

2.1.1. ANOVA táblák a K-means módszerre (2-6 rendre)

ANOVA						
		Sum of Squares	df	Mean Square	F	Sig.
Zscore(budget)	Between Groups	42,886	1	42,886	97,135	,000
	Within Groups	33,114	75	,442		
	Total	76,000	76			
Zscore(revenue)	Between Groups	46,451	1	46,451	117,897	,000
	Within Groups	29,549	75	,394		
	Total	76,000	76			
Zscore(popularity)	Between Groups	41,952	1	41,952	92,412	,000
	Within Groups	34,048	75	,454		
	Total	76,000	76			
Zscore(runtime)	Between Groups	8,175	1	8,175	9,040	,004
	Within Groups	67,825	75	,904		
	Total	76,000	76			
Zscore(vote_average)	Between Groups	18,440	1	18,440	24,028	,000
	Within Groups	57,560	75	,767		
	Total	76,000	76			
Zscore(vote_count)	Between Groups	55,533	1	55,533	203,503	,000
	Within Groups	20,467	75	,273		
	Total	76,000	76			
Zscore(age)	Between Groups	7,710	1	7,710	8,467	,005
	Within Groups	68,290	75	,911		
	Total	76,000	76			

ANOVA						
		Sum of Squares	df	Mean Square	F	Sig.
Zscore(budget)	Between Groups	42,042	2	21,021	45,807	,000
	Within Groups	33,958	74	,459		
	Total	76,000	76			
Zscore(revenue)	Between Groups	48,130	2	24,065	63,899	,000
	Within Groups	27,870	74	,377		
	Total	76,000	76			
Zscore(popularity)	Between Groups	39,956	2	19,978	41,017	,000
	Within Groups	36,044	74	,487		
	Total	76,000	76			
Zscore(runtime)	Between Groups	12,001	2	6,001	6,938	,002
	Within Groups	63,999	74	,865		
	Total	76,000	76			
Zscore(vote_average)	Between Groups	18,106	2	9,053	11,572	,000
	Within Groups	57,894	74	,782		
	Total	76,000	76			
Zscore(vote_count)	Between Groups	56,717	2	28,359	108,830	,000
	Within Groups	19,283	74	,261		
	Total	76,000	76			
Zscore(age)	Between Groups	48,197	2	24,098	64,139	,000
	Within Groups	27,803	74	,376		
	Total	76,000	76			

ANOVA

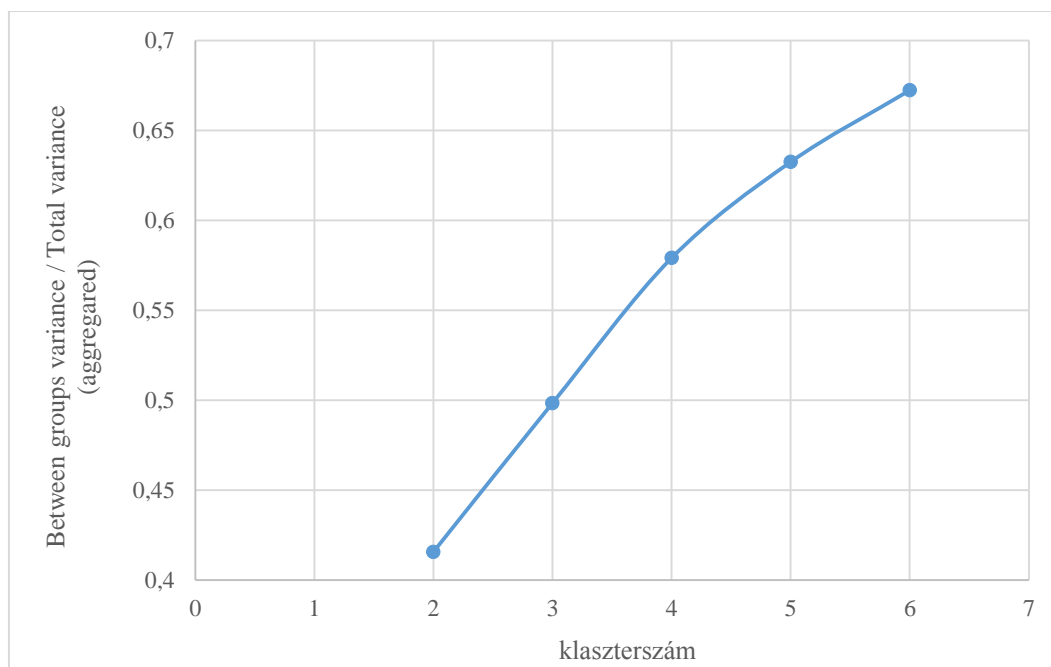
		Sum of Squares	df	Mean Square	F	Sig.
Zscore(budget)	Between Groups	43,523	3	14,508	32,609	,000
	Within Groups	32,477	73	,445		
	Total	76,000	76			
Zscore(revenue)	Between Groups	49,736	3	16,579	46,080	,000
	Within Groups	26,264	73	,360		
	Total	76,000	76			
Zscore(popularity)	Between Groups	43,692	3	14,564	32,907	,000
	Within Groups	32,308	73	,443		
	Total	76,000	76			
Zscore(runtime)	Between Groups	36,862	3	12,287	22,918	,000
	Within Groups	39,138	73	,536		
	Total	76,000	76			
Zscore(vote_average)	Between Groups	30,934	3	10,311	16,703	,000
	Within Groups	45,066	73	,617		
	Total	76,000	76			
Zscore(vote_count)	Between Groups	57,775	3	19,258	77,138	,000
	Within Groups	18,225	73	,250		
	Total	76,000	76			
Zscore(age)	Between Groups	45,526	3	15,175	36,353	,000
	Within Groups	30,474	73	,417		
	Total	76,000	76			

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
Zscore(budget)	Between Groups	51,495	4	12,874	37,825	,000
	Within Groups	24,505	72	,340		
	Total	76,000	76			
Zscore(revenue)	Between Groups	61,109	4	15,277	73,870	,000
	Within Groups	14,891	72	,207		
	Total	76,000	76			
Zscore(popularity)	Between Groups	47,093	4	11,773	29,325	,000
	Within Groups	28,907	72	,401		
	Total	76,000	76			
Zscore(runtime)	Between Groups	31,609	4	7,902	12,817	,000
	Within Groups	44,391	72	,617		
	Total	76,000	76			
Zscore(vote_average)	Between Groups	30,125	4	7,531	11,820	,000
	Within Groups	45,875	72	,637		
	Total	76,000	76			
Zscore(vote_count)	Between Groups	63,905	4	15,976	95,104	,000
	Within Groups	12,095	72	,168		
	Total	76,000	76			
Zscore(age)	Between Groups	51,124	4	12,781	36,994	,000
	Within Groups	24,876	72	,345		
	Total	76,000	76			

ANOVA						
		Sum of Squares	df	Mean Square	F	Sig.
Zscore(budget)	Between Groups	51,578	5	10,316	29,989	,000
	Within Groups	24,422	71	,344		
	Total	76,000	76			
Zscore(revenue)	Between Groups	58,817	5	11,763	48,605	,000
	Within Groups	17,183	71	,242		
	Total	76,000	76			
Zscore(popularity)	Between Groups	64,027	5	12,805	75,933	,000
	Within Groups	11,973	71	,169		
	Total	76,000	76			
Zscore(runtime)	Between Groups	34,483	5	6,897	11,794	,000
	Within Groups	41,517	71	,585		
	Total	76,000	76			
Zscore(vote_average)	Between Groups	34,285	5	6,857	11,671	,000
	Within Groups	41,715	71	,588		
	Total	76,000	76			
Zscore(vote_count)	Between Groups	64,614	5	12,923	80,580	,000
	Within Groups	11,386	71	,160		
	Total	76,000	76			
Zscore(age)	Between Groups	49,840	5	9,968	27,054	,000
	Within Groups	26,160	71	,368		
	Total	76,000	76			

2.1.2. Klaszterkönyököt bemutató ábra



2.2.1. K-means végleges klaszterközépponttól vett távolság és klasztagságok száma

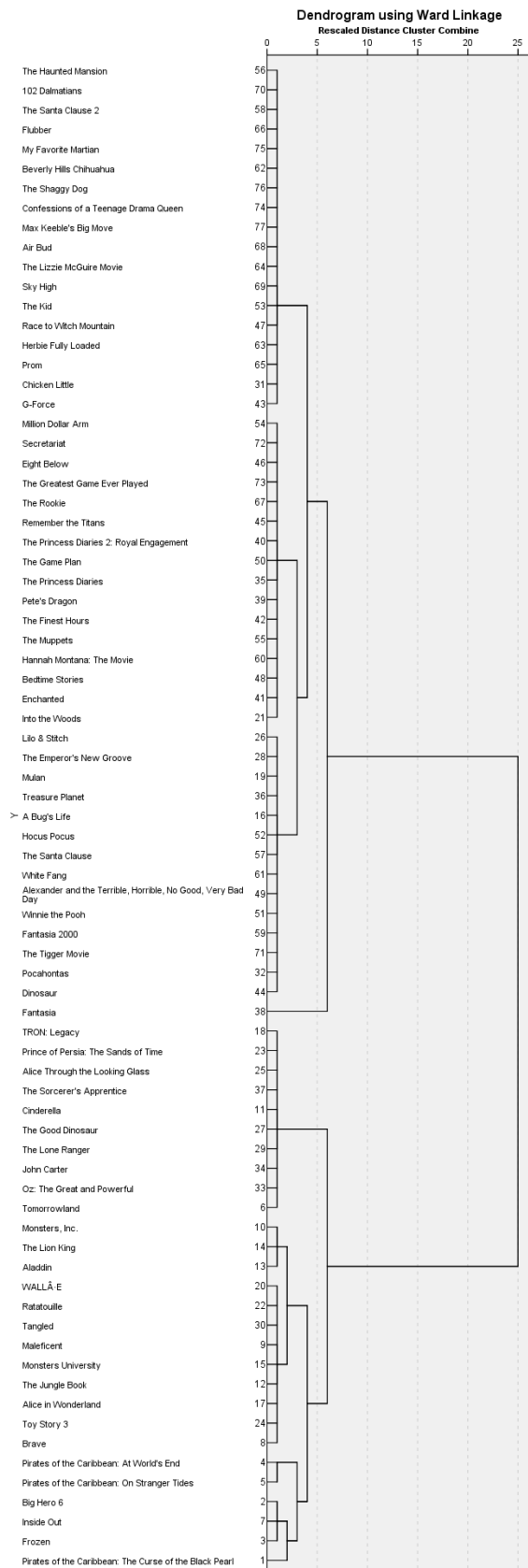
Final Cluster Centers

	Cluster			
	1	2	3	4
Zscore(budget)	-1,22567	-,47711	,63279	1,71140
Zscore(revenue)	-,74295	-,55782	1,29036	,67594
Zscore(popularity)	-,35831	-,53909	,97418	1,22220
Zscore(runtime)	1,12496	-,23456	-,29043	1,97886
Zscore(vote_average)	,98564	-,40924	1,09498	-,02915
Zscore(vote_count)	-,52908	-,61743	1,30111	,99760
Zscore(age)	6,50475	,03941	-,24598	-,50592

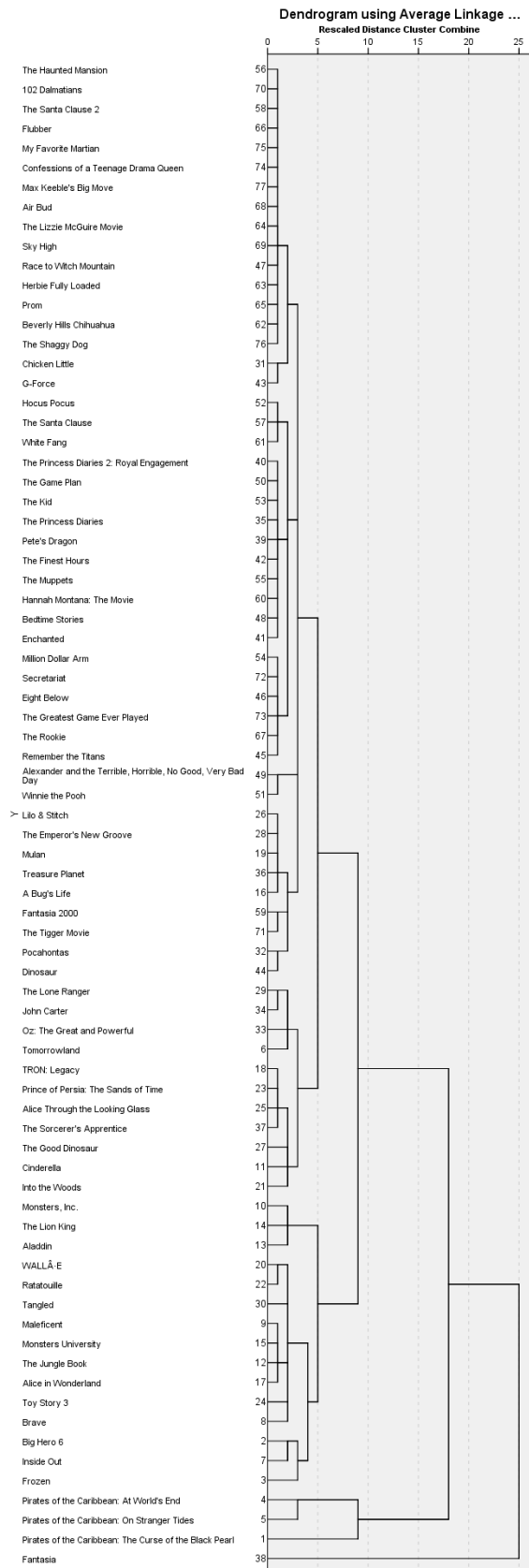
Number of Cases in each Cluster

Cluster	1	1,000
	2	50,000
	3	18,000
	4	8,000
Valid		77,000
Missing		,000

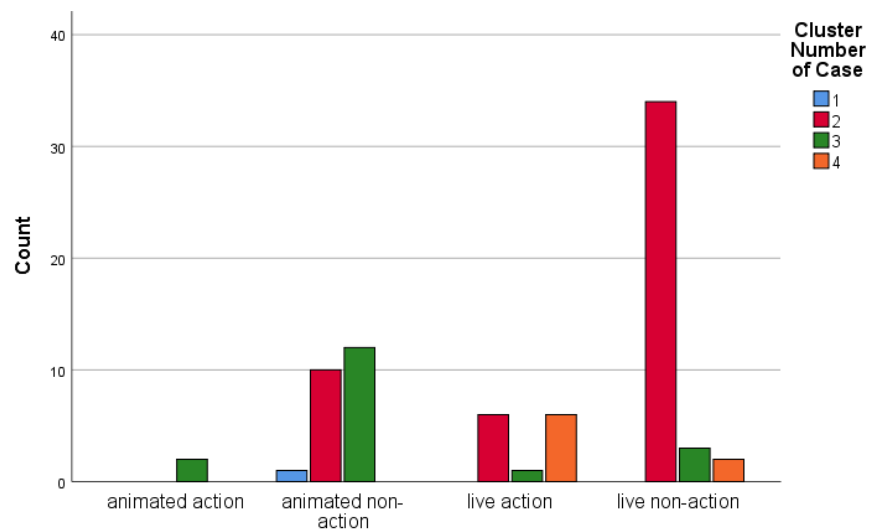
2.3.1.a Ward-linkage módszerrel kapott dendrogramm



2.3.1.b Average-linkage módszerrel kapott dendrogramm



3.1. Kereszt táblás elemzés



4.1.1. Korrelációs mátrix

Correlation Matrix^a

		budget	revenue	age	popularity	runtime	vote_average	vote_count
Correlation	budget	1,000	,700	-,284	,427	,214	-,025	,551
	revenue	,700	1,000	-,162	,622	,211	,187	,775
	age	-,284	-,162	1,000	-,184	,134	,200	-,190
	popularity	,427	,622	-,184	1,000	,139	,270	,740
	runtime	,214	,211	,134	,139	1,000	,363	,226
	vote_average	-,025	,187	,200	,270	,363	1,000	,375
	vote_count	,551	,775	-,190	,740	,226	,375	1,000

a. Determinant = ,048

4.1.2. KMO és Bartlett teszt

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,734
Bartlett's Test of Sphericity	Approx. Chi-Square	6521,285
	df	21
	Sig.	,000

4.1.3. Anti-imázs mátrix

Anti-image Matrices								
		budget	revenue	age	popularity	runtime	vote_average	vote_count
Anti-image Covariance	budget	,442	-,175	,125	,013	-,126	,130	-,027
	revenue	-,175	,283	-,061	-,043	,000	,023	-,130
	age	,125	-,061	,828	,042	-,110	-,144	,062
	popularity	,013	-,043	,042	,443	,025	-,018	-,159
	runtime	-,126	,000	-,110	,025	,798	-,234	-,001
	vote_average	,130	,023	-,144	-,018	-,234	,654	-,136
	vote_count	-,027	-,130	,062	-,159	-,001	-,136	,253
Anti-image Correlation	budget	,735 ^a	-,494	,207	,030	-,212	,242	-,079
	revenue	-,494	,755 ^a	-,125	-,120	,000	,054	-,486
	age	,207	-,125	,630 ^a	,069	-,136	-,195	,135
	popularity	,030	-,120	,069	,834 ^a	,042	-,034	-,474
	runtime	-,212	,000	-,136	,042	,646 ^a	-,324	-,001
	vote_average	,242	,054	-,195	-,034	-,324	,570 ^a	-,334
	vote_count	-,079	-,486	,135	-,474	-,001	-,334	,738 ^a

a. Measures of Sampling Adequacy(MSA)

4.3.1. PCA és PAF kommunalítások (rendre)

Communalities		
	Initial	Extraction
budget	1,000	,651
revenue	1,000	,800
age	1,000	,542
popularity	1,000	,653
runtime	1,000	,474
vote_average	1,000	,678
vote_count	1,000	,832

Extraction Method: Principal Component Analysis.

Communalities		
	Initial	Extraction
budget	,558	,549
revenue	,717	,789
age	,172	,180
popularity	,557	,533
runtime	,202	,171
vote_average	,346	,771
vote_count	,747	,838

Extraction Method: Principal Axis Factoring.

4.3.2. A rotált komponens mátrix (PCA) és a rotált faktor mátrix (PAF) (rendre)

Rotated Component Matrix^a

	Component	
	1	2
budget	,797	-,129
revenue	,888	,105
age	-,408	,613
popularity	,792	,162
runtime	,211	,655
vote_average	,187	,802
vote_count	,879	,243

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.^a

a. Rotation converged in 3 iterations.

Rotated Factor Matrix^a

	Factor	
	1	2
budget	,738	-,073
revenue	,875	,150
age	-,313	,287
popularity	,690	,239
runtime	,161	,381
vote_average	,076	,875
vote_count	,847	,347

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.^a

a. Rotation converged in 3 iterations.

4.3.3. PCA és PFA módszerek teljes varianciáját magyarázó táblák (rendre)

Total Variance Explained

Component	Total	Initial Eigenvalues		Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
		% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3,155	45,074	45,074	3,155	45,074	45,074	3,069	43,845	43,845
2	1,474	21,056	66,130	1,474	21,056	66,130	1,560	22,285	66,130
3	,821	11,729	77,859						
4	,710	10,149	88,008						
5	,422	6,025	94,033						
6	,252	3,594	97,627						
7	,166	2,373	100,000						

Extraction Method: Principal Component Analysis.

Total Variance Explained

Factor	Total	Initial Eigenvalues		Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
		% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3,155	45,074	45,074	2,835	40,502	40,502	2,633	37,620	37,620
2	1,474	21,056	66,130	,997	14,244	54,746	1,199	17,126	54,746
3	,821	11,729	77,859						
4	,710	10,149	88,008						
5	,422	6,025	94,033						
6	,252	3,594	97,627						
7	,166	2,373	100,000						

Extraction Method: Principal Axis Factoring.

5.1. Standardizáltuk a lineáris változókat

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
budget	2149	1	380000000	38949047,96	43253541,11
revenue	2149	5	2787965087	120873576,7	189144147,2
popularity	2149	,0370730000	875,5813050	28,61171914	36,59566046
runtime	2149	63	238	108,87	18,832
vote_average	2149	2,300000000	8,500000000	6,287947883	,8522884765
vote_count	2149	1	12002	950,17	1370,447
age	2149	1,052054795	101,1342466	15,99604021	13,84223797
Valid N (listwise)	2149				

5.2 A bent maradt változók szignifikancia alapján

Variables in the Equation							
		B	S.E.	Wald	df	Sig.	Exp(B)
Step 5 ^a	drama(1)	2,102	,492	18,228	1	,000	8,182
	Zscore(budget)	-1,455	,173	70,837	1	,000	,233
	Zscore(revenue)	-,408	,155	6,922	1	,009	,665
	Zscore(runtime)	4,452	,370	145,128	1	,000	85,800
	Zscore(vote_average)	-1,014	,166	37,439	1	,000	,363
	Constant	5,923	,408	210,691	1	,000	373,526

a. Variable(s) entered on step 5: Zscore(revenue).

5.3 Korrelációs mátrix

Correlation Matrix							
		Constant	drama(1)	Zscore (budget)	Zscore (revenue)	Zscore (runtime)	Zscore (vote_average)
Step 1	Constant	1,000	,034	-,521	-,141	,931	-,195
	drama(1)	,034	1,000	-,040	-,041	,155	-,243
	Zscore(budget)	-,521	-,040	1,000	-,549	-,513	,321
	Zscore(revenue)	-,141	-,041	-,549	1,000	-,105	-,282
	Zscore(runtime)	,931	,155	-,513	-,105	1,000	-,268
	Zscore(vote_average)	-,195	-,243	,321	-,282	-,268	1,000

5.4 Omnibus teszt-együttes szignifikancia

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 5	Step	7,244	1	,007
	Block	623,451	5	,000
	Model	623,451	5	,000

5.5 Wald tesztnél egyesével is szignifikánsak

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 5 ^a	drama(1)	2,102	,492	18,228	1	,000	8,182
	Zscore(budget)	-1,455	,173	70,837	1	,000	,233
	Zscore(revenue)	-,408	,155	6,922	1	,009	,665
	Zscore(runtime)	4,452	,370	145,128	1	,000	85,800
	Zscore(vote_average)	-1,014	,166	37,439	1	,000	,363
	Constant	5,923	,408	210,691	1	,000	373,526

a. Variable(s) entered on step 5: Zscore(revenue).

5.6 Cox & Snell R2 és Nagelkerke R2

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
5	395,884 ^a	,252	,667

a. Estimation terminated at iteration number 9 because parameter estimates changed by less than ,001.

5.7 Hosmer and Lemeshow teszt

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
5	55,718	8	,000

5.8 Hosmer and Lemeshow teszt

Contingency Table for Hosmer and Lemeshow Test

		animation = Animatio		animation = Live		Total
		Observed	Expected	Observed	Expected	
Step 5	1	117	114,545	98	100,455	215
	2	13	15,604	202	199,396	215
	3	3	4,432	212	210,568	215
	4	2	1,568	213	213,432	215
	5	0	,547	215	214,453	215
	6	1	,209	214	214,791	215
	7	0	,072	215	214,928	215
	8	1	,019	214	214,981	215
	9	0	,004	215	214,996	215
	10	0	,000	214	214,000	214

5.9. Roc görbe alatti terület

Area Under the Curve

Test Result Variable(s): Predicted probability

Area	Std. Error ^a	Asymptotic Sig. ^b	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
,968	,008	,000	,953	,984

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

6.1.1. Wilks' Lambda csoport átlagra

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
drama	,959	90,711	1	2147	,000
Zscore(budget)	,914	201,180	1	2147	,000
Zscore(revenue)	,929	164,054	1	2147	,000
Zscore(runtime)	,934	152,809	1	2147	,000
Zscore(vote_average)	,997	6,700	1	2147	,010

6.1.2. Normál eloszlás teszt

Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Zscore(budget)	,184	2149	,000	,771	2149	,000
Zscore(revenue)	,261	2149	,000	,604	2149	,000
Zscore(runtime)	,094	2149	,000	,924	2149	,000
Zscore(vote_average)	,044	2149	,000	,988	2149	,000
Zscore(drama)	,367	2149	,000	,633	2149	,000

6.1.3. Normál eloszlás teszt LN értékekkel

Tests of Normality ^{a,c,d,e,f}							
	animation	Kolmogorov-Smirnov ^b			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
In_Zbudget	1	,145	72	,001	,919	72	,000
In_revenue	1	,106	72	,042	,951	72	,007
In_vote_average	1	,210	72	,000	,787	72	,000
In_runtime	1	,212	72	,000	,746	72	,000
In_drama	1	.	72	.	.	72	.

6.1.4. Box's M teszt

Box's Test of Equality of Covariance Matrices

Log Determinants		
animation	Rank	Log Determinant
0	5	-3,945
1	5	-3,010
Pooled within-groups	5	-2,879

The ranks and natural logarithms of determinants printed are those of the group covariance matrices.

Test Results		
Box's M		409,318
F	Approx.	26,928
	df1	15
	df2	212216,742
	Sig.	,000

Tests null hypothesis of equal population covariance matrices.

6.2.1. Vizsgált elemszám

Analysis Case Processing Summary

Unweighted Cases		N	Percent
Valid		2149	100,0
Excluded	Missing or out-of-range group codes	0	,0
	At least one missing discriminating variable	0	,0
	Both missing or out-of-range group codes and at least one missing discriminating variable	0	,0
	Total	0	,0
Total		2149	100,0

6.2.2. Relatív szórás

Group Statistics					
				Valid N (listwise)	
animation		Mean	Std. Deviation	Unweighted	Weighted
0	<u>budget</u>	87455669,365	53206531,3956	137	137,000
	<u>revenue</u>	313952081,496	281376952,4147	137	137,000
	<u>runtime</u>	90,270	9,1076	137	137,000
	<u>vote_average</u>	6,470	0,8342	137	137,000
	<u>drama</u>	0,066	0,2487	137	137,000
1	<u>budget</u>	35646161,712	40443364,5331	2012	2012,000
	<u>revenue</u>	107726581,148	173625509,5541	2012	2012,000
	<u>runtime</u>	110,135	18,6552	2012	2012,000
	<u>vote_average</u>	6,276	0,8523	2012	2012,000
	<u>drama</u>	0,476	0,4995	2012	2012,000
<u>Total</u>	<u>budget</u>	38949047,961	43253541,1086	2149	2149,000
	<u>revenue</u>	120873576,749	189144147,2437	2149	2149,000
	<u>runtime</u>	108,868	18,8318	2149	2149,000
	<u>vote_average</u>	6,288	0,8523	2149	2149,000
	<u>drama</u>	0,450	0,4976	2149	2149,000
					szórás / átlag

6.2.3. Kanonikus korreláció

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	,316 ^a	100,0	100,0	,490

a. First 1 canonical discriminant functions were used in the analysis.

6.2.4. Wilks' Lambda

Wilks' Lambda				
Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	,760	588,658	5	,000

6.2.5. Koeffisiensek

Canonical Discriminant Function Coefficients

	Function 1
drama	-,242
Zscore(budget)	,744
Zscore(revenue)	,186
Zscore(runtime)	-,943
Zscore(vote_average)	,490
(Constant)	,109

Unstandardized coefficients