Big data analysis as a tool for assessing professionalization of computer sports

UDC 004.42:796.011

PhD **A.V. Ermakov**[1]
PhD **E.N. Skarzhinskaya**[2]
[1]Russian University of Sport (SCOLIPE), Moscow
[2]Moscow University for Industriy and Finance "Synergy", Moscow

Corresponding author: bigbr@mail.ru

Abstract

Objective of the study was to substantiate the criteria and assess the level of professionalization of the sport "Computer Sports" using big data analysis tools.

Methods and structure of the study. The analysis of organizational and economic characteristics of eSports tournaments of various levels (Tier 1 and Tier 2) was carried out using a Python script and library script for this study. A statistical analysis of the numerical variables of the Tier 1 and Tier 2 teams by years (2011-2022) is also presented.

Results and conclusions. The signs of professionalization of the studied eSports discipline were revealed, which indicated the need to build scientific and methodological support for sports training.

Keywords: *artificial intelligence, big data, sports economics, professional sports, e-sports, computer sports, Dota2.*

Introduction. Modern sports science cannot avoid turning to high-tech solutions that allow obtaining data, managing data and analyzing them to gain new knowledge in the field of theory and methodology of physical culture and sports [1]. Cybersport is a relevant research ground both from the point of view of digital technologies and strategically - in the Strategy for the Development of Physical Culture and Sports until 2030, approved by the Decree of the Government of the Russian Federation of November 24, 2020 No. 3081-r, the promotion of computer sports is written in a separate line. Therefore, the study of managerial and economic aspects of esports in general [1], as well as various problems in the most popular esports disciplines [2, 3, etc.] is relevant.

Objective of the study was to substantiate the criteria and assess the level of professionalization of the sport "Computer Sports" using big data analysis tools.

Methods and structure of the study. The professionalization of sports, as a process of the maturity of social institutions and their economic indicators, allows us to assess the degree of development of a new sport, to which we include computer sports. The establishment of professionalization criteria determines the need for both the number of athletes and the quality of their training.

To analyze the economic indicators of the development of the sport "computer sports" (the identical term for cybersport), prize payments in the Dota2 game were chosen for highly qualified athletes, namely Tier 1 and Tier 2 teams of the world ranking. Payments were fixed for a ten-year period from 2011 to October 2022 for each level. For Tier 1 teams, 177 tournaments were analyzed (using artificial intelligence technology), and for Tier 2 teams, 322 tournaments, respectively. In each of the analyzed datasets, the following variables were recorded as variables: the name of the tournament, the time of its holding, the prize fund, the number of participating teams, the venue, the team that won first place, the team that took second place. The data were collected by scraping (webscraping) using a script specially written for this study.

Table 1. *Quantitative variables characterizing Dota2 tournaments*

Tournament end year	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Tier 1												
Number of tournaments	3	19	28	30	23	20	16	11	8	7	6	6
Average prize pool	34361.11	6954.12	10708.22	37114.83	89027.23	116389.09	136443.04	193154.93	287309.50	54166.67	397217.08	198236.88
Average number of teams in the tournament	14.67	13.84	12.32	8.90	9.22	10.70	10.50	15.64	16.75	14.29	14.00	13.83
Tier 2												
Number of tournaments	2	12	22	37	19	32	20	18	23	59	24	32
Average prize pool	1524.88	518.53	1459.63	4132.74	8355.11	13792.83	18973.42	26122.73	23927.29	12311.63	19536.57	22311.62
Average number of teams in the tournament	8.00	14.42	12.05	9.05	11.32	10.50	9.90	10.39	8.49	9.09	8.38	7.59

Results of the study and their discussion. The distribution of quantitative variables characterizing tournaments according to the years of holding is presented in Table 1.

Looking at the variable data, one can notice differences in absolute values between Tier 1 and Tier 2 tournaments for all the given variables. For a more accurate assessment, we calculate descriptive statistics for our dataset and place them in Table 2.

When considering the obtained descriptive statistics, the rather large range of all of them attracts attention, and for a better understanding of the current situation, it is necessary to make a visual analysis of the data under study. First, we visualize the dynamics of the development of Dota2 tournaments for Tier1 teams (Fig. 1)

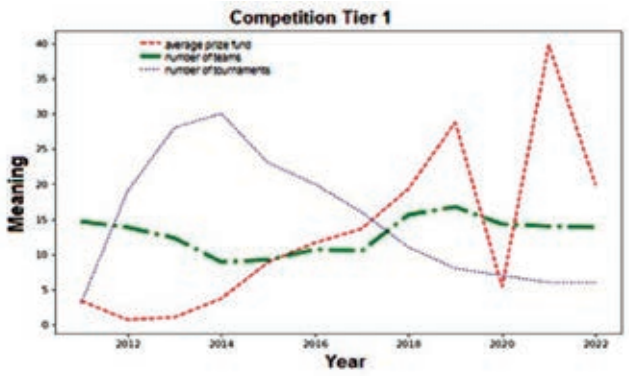

Figure 1. *Dynamics of the development of Dota2 tournaments for Tier 1 teams (the average prize pool is measured here at $10,000)*

Table 2. *Descriptive statistics of Dota2 tournaments for Tier 1 and Tier 2 teams*

Variable data	Tier 1			Tier 2		
	Total prize pool	Average prize pool	Number of teams in the tournament	Total prize pool	Average prize pool	Number of teams in the tournament
Mean	1487808.05	93683.92	11.79	110109.1	13678.85	9.67
Standard deviation	5243244.42	295263.4	5.04	97846.95	12568	4.83
Minimum	6000.00	512.25	4.00	1500	145.83	4
Percentile 25%	50000.00	5375.75	8.00	40000	3750	8
Percentile 50%	208239.00	23770.25	10.00	80000	9045.46	8
Percentile 75%	500000.00	50999.75	16.00	185064	25000	10.25
Maximum	40018195.00	2223233	32.00	700000	70000	36

Table 3. *Statistical analysis of numerical variables of Tier 1 and Tier 2 teams by years*

Tournament Characteristics	Mann-Whitney U test		H-test Kruskal-Wallis	
	The meaning of statistics	Reliability of differences	The meaning of statistics	Reliability of differences
Number of teams	33758.5	p≤0,001	26.3	p≤0,001
Number of tournaments	16504.0	p≤0,001	48.2	p≤0,001
Average prize pool	35192.0	p≤0,001	35.3	p≤0,001
Total prize pool	35914.0	p≤0,001	41.5	p≤0,001

The visualization allows us to conclude that the number of tournaments for Tier 1 teams had its moment of growth, the peak of which was passed in 2014 and is now in the stabilization stage. The number of teams turned out to be the most stable variable and also came into some equilibrium with the other variables. The prize fund turned out to be the most variable variable. Global social factors have a great influence on it, for example, we see a significant drop in 2020 associated with the COVID-19 pandemic, which led to restrictions, including mass events, although it can be seen that the decrease in the number of tournaments is insignificant compared to a decrease in the prize fund. A somewhat different situation is observed in Dota2 tournaments for Tier 2 teams (Fig. 2)

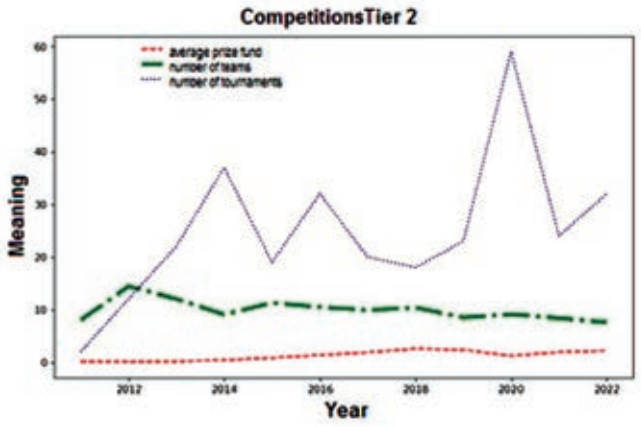

Figure 2. *Dynamics of the development of Dota2 tournaments for Tier 2 teams (the average prize pool is measured here at $10,000)*

For Tier 2 teams, the number of tournaments is the most variable, and you can see this number increase in 2020, where Tier 1 teams have seen a drop. The number of teams in the tournament and the average prize fund of the tournament can be considered stabilized. For reasonable conclusions, let's check the statistical relationship between these two levels. To do this, we apply an assessment according to two nonparametric criteria, the Mann-Whitney U-test and the Kruskal-Wallis H-test (Table 3).

It can be seen that all variables are significantly different and therefore we can conclude that there is no formed general model for holding Dota2 tournaments for Tier 1 and Tier 2 teams. When trying to create a regression model within each of the described levels, it also showed that such models will have low productivity. The highest coefficient of determination R2 was 0.067, which also allows us to state that there is no clearly visible productive model within the Tier 1 and Tier 2 levels themselves.

Conclusions. An analysis of the socio-economic characteristics of Dota2 tournaments for Tier 1 and Tier 2 teams showed that there are prerequisites for the formation of professional computer sports and its individual characteristics, such as: the number of tournaments; number of teams in the tournament for Tier 1; the number of teams in the tournament and the average prize pool for Tier 2 teams. At the same time, we see that there remains a strong dependence on the influence of global factors, such as the prize pool for Tier 1 teams and the number of tournaments for Tier 2. At the same time, the lack of productive models within esports and the absence The statistical relationship between the variables between the groups suggests that although professionalization is quite active in the Dota2 esports discipline, it is still far from complete. In the future, with the application of certain efforts, it is possible to have a significant impact on its architectonics or hinder the professional development of computer sports.

When predicting the sports results of domestic cybersportsmen at sports competitions of various levels, one should take into account the socio-economic conditions and the quality of scientific and methodological support for sports training.

References

1. Zalilov M.A. Istoriya kibersportivnogo menedzhmenta v Rossii [History of eSports management in Russia]. Kompyuternyy sport (kibersport): sostoyaniye i perspektivy razvitiya [Computer sports

(eSports): state and development prospects]. Proceedings national scientific-practical conference. M.A. Novoselov [ed.]. 2019. pp. 20-24.
2. Skarzhinsky N.S., Grebenkov V.S., Martynenko I.V. Modelnyye kharakteristiki kibersportsmenov v distsipline «Boyevaya arena» (na primere Dota2) [Model characteristics of cybersportsmen in the discipline “Combat Arena” (on the example of Dota2)]. Fizicheskaya kultura: vospitaniye, obrazovaniye, trenirovka. 2020. No. 5. p. 31.
3. Talan A.S., Babkin A.A., Talan M.S. Vzaimosvyaz vozrasta i roli v igre Dota2 [The relationship of age and role in the game Dota2]. Kafedralnaya nauka RGUFKSMiT [Department of Science RGUFKSMiT]. Proceedings scientific-practical conference of the faculty, 2019. pp. 56-59.
4. Esports job spotlight: Coach/Analyst Available at: https://britishesports.org/careers/coach-analyst/ (date of access: 09.26.2022).