Original Article

Influence of the 2020 pandemic on speedway – managerial implication

SYLWESTER BEJGER

The Faculty of Economic Sciences and Management, Nicolaus Copernicus University, Toruń, POLAND

Published online: April 30, 2021

(Accepted for publication April 15, 2021)

DOI:10.7752/jpes.2021.s2121

Abstract:

The season of the year 2020 was characterized by a number of difficulties and extraordinary organizational and management conditions for many sports disciplines. One of the disciplines most affected by the restrictions caused by the pandemic was motorcycle speedway. Suffice it to say that out of all European leagues, only the Polish league and the Swedish league completed the 2020 season. Moreover in both countries the games started with a long delay and were not preceded by typical preparatory activities for the season, such as sparring matches or individual tournaments. Focusing on Polish highest level league, the PGE Ekstraliga, the paper contains careful statistical and exploratory analysis of data sample encompassing seasons 2015 to 2020, aiming on discovery if and how season 2020 differed from previous ones. This analysis was carried out in the scope of three data sets with different levels of aggregation, concerning meetings, teams and heats of individual riders. The analysis covered the general performance of a whole sample of riders, performance and demographic structure of teams, and results of heat per rider, taking account rider's nationality and age. The results of the analysis made it possible to identify changes in selected metrics describing the effectiveness of teams and players, probably caused by structural difference of 2020 season. This allows for the formulation of conclusions supporting the management of the teams and clubs in decision making process concerning team's profile and riders and track preparation. Such conclusions can help coaches mitigate the impact of unfavorable external conditions on team quality and individual player preparation in similar circumstances in a future.

Key Words: - team management, motorcycle speedway, data mining, demographic and internationalization, impact of pandemic on sport

Introduction

Motorcycle speedway (or speedway in short, when misunderstanding is impossible) was created in Australia at the beginning of the 1920s (May, 1978). Speedway in Poland started in 1930 (Błaszowska, 2018). Currently, speedway in Poland plays an especially important role in the economic, social and sports sense. Contrary to other countries, this sport is very popular in Poland and ranks 9th on the list of sports by popularity (Report, 2021). The meeting of the Polish highest level league, PGE Ekstraliga alone gathered around 700,000 fans in the 2019 season (PGE statistics, 2021). PGE Ekstraliga is considered as one of the most important speedway leagues in a world. Speedway is an individual sport in which trained skill plays a central role, allowing us to understand speedway not as mere sport, but as art (Siitonen et al. 2020). However, the success of the team is the sum of not only the individual efforts of the players, but also depends on the proper strategy of team building and preparation, as well as preparation of a home track. As all of the motorsports, it is an expensive discipline either.

There is no doubt that the pandemic season of 2020 has seriously disrupted the functioning of speedway leagues and the training and competition cycle of the riders all over the world. Due to worldwide lockdown only PGE Ekstraliga and Swedish Eliteserien started and finished 2020 season in Europe (season started with significant delay in that countries, 12.06 and 01.08 respectively). In Poland, due to restriction imposed by authorities there were no usual spring, preliminary speedway events (individual tournaments, sparring meetings) and the number of trainings in a track for riders was significantly limited, either. Moreover, even after starting of the league the meetings were held partly without fans and partly with 50% participation of them.

Taking into consideration the obvious abnormalities of season 2020 and above mentioned, special importance of a speedway and speedway league for Polish sport and socio-economical life in general, the question of impact of the season's 2020 conditions on performance of the teams and riders seems to be very important. The main research goal of the paper is to find some distinctive changes of 2020 season's parameters by mining a statistical data. Such findings could lead to conclusions that could be useful for the management of the teams and decisions making process of building team's profile, taking into account both demographic aspect and internationalization. As in the other disciplines of sport (see for example Gulak – Lipka, 2020) role of globalization and internationalization in speedway is increasing. Improper management may lead to financial

difficulties, which in turn might result in bankruptcy of the club (Lis, Tomanek, 2020), especially in a macroeconomic crisis environment.

Material & methods

The following research methods have been used to achieve the assumed research goal: statistical data mining, descriptive, comparative and subject literature analysis. The main body of the research is statistical examination and inference leading to comparative analysis.

Time period of the research covers seasons 2015 – 2020 of speedway PGE Ekstraliga in Poland. Author used data¹ collected by the portal <http://gurustats.pl>. Dataset is divided in three logical parts, namely data on meetings, data on teams and data on riders per meeting. For the purposes of the study, author chose an appropriate subset of the features. Detailed description of variable used in particular part of the research is provided in a Results section.

For the research the following exploratory methods have been used. Statistical visualization (histograms, scatter plots and a kernel density estimate (KDE) plot). A histogram aims to approximate the underlying probability density function that generated the data by binning and counting observations. Kernel density estimation (KDE) plot smooths the observations with a Gaussian kernel, producing a continuous density estimate (Hastie, et al. 2009). Selected statistical tests used: D'Agostino and Pearson's (1973) omnibus test of normality, Epps-Singleton (1986) ES test whether two samples are generated by the same underlying distribution, equality of variance Levene (1960) test. Levene's test is an alternative to Bartlett's test in the case where there are significant deviations from normality. In a part of the research connected with cluster analysis the Kruskal-Wallis H-test (2001) was adapted. This test is a non-parametric version of ANOVA and tests the null hypothesis that the population median of all of the groups sampled from are equal. For correlation assessment the Spearman (1903) rank-order correlation coefficient was used. This test is a nonparametric measure of the monotonicity of the relationship between two datasets. Unlike the Pearson correlation, the Spearman correlation does not assume that both datasets are normally distributed. To calculate Spearman r_s for a sample of size n , values of variables X_i and Y_i are converted to its ranks rg_X and rg_Y and use to compute:

$$r_s = \rho_{rg_X, rg_Y} = \frac{\text{cov}(rg_X, rg_Y)}{\sigma_{rg_X} \sigma_{rg_Y}}, \quad (1)$$

where: $\text{cov}(rg_X, rg_Y)$ is the covariance of the rank variables, σ_{rg_X} , σ_{rg_Y} are the standard deviations of the rank variables.

Important part of the research is the cluster analysis. For that task K-Means algorithm (1957, 1965) was used. The K-Means algorithm divides a set of N samples X into k disjoint clusters C , each described by the mean μ_j of the samples in the cluster. The means are called the cluster "centroids". The K-Means algorithm aims to choose centroids that minimize the criterion so called inertia, which is within-cluster sum of squares:

$$\sum_{i=0}^n \min_{\mu_j \in C} (\|x_i - \mu_j\|^2) \quad (2)$$

Inertia can be recognized as a measure of how internally coherent clusters are. The number of clusters k can be determined according to inertia or mean silhouette coefficient over all the instances. An instance's silhouette coefficient is equal to $(b-a)/\max(a,b)$ where a is the mean distance to the other instances in the same cluster (it is the mean intra-cluster distance), and b is the mean nearest-cluster distance, that is the mean distance to the instances of the next closest cluster (defined as the one that minimizes b , excluding the instance's own cluster). The silhouette coefficient can vary between -1 and +1: a coefficient close to +1 means that the instance is well inside its own cluster and far from other clusters, while a coefficient close to 0 means that it is close to a cluster boundary, and finally a coefficient close to -1 means that the instance may have been assigned to the wrong cluster.

Results

The first part of the empirical research starts with analysis of the meetings in a sample period. The main research question was whether the 2020 season was statistically significantly different from other seasons in terms of overall results and general performance of the riders. As a proxy measure of performance the Best Heat Time of a meeting was used. Three variables connected with results were also included in the study. Descriptive statistics of examined features are depicted in Table 1².

¹ Data available commercially on request.

² All of the Tables and Figures in a paper present author's own results.

Table 1. Descriptive statistics - meetings

stat\season	2015				2016				2017			
	BHT*	Pts_H	Pts_A	Pts_diff	BHT	Pts_H	Pts_A	Pts_diff	BHT	Pts_H	Pts_A	Pts_diff
count	64	64	64	64	64	64	64	64	64	64	64	64
mean	62.77	49.17	40.63	11.55	62.98	49.19	40.53	12.75	62.04	47.72	41.20	10.08
std	3.60	6.01	6.00	9.09	4.06	6.74	6.31	8.89	3.16	6.16	6.15	7.78
min	57.41	36	27	0	56.40	29	28	0	56.50	31	17	0.00
25%	60.24	45	37	5.5	59.56	45	35.75	6	59.52	44	38	3.75
median	61.09	49	40.5	10	61.46	49.5	40.5	11	61.06	48	41	9.00
75%	65.24	53	45	16	66.18	54	45	20	64.81	52	45.25	14.00
stat\season	2018				2019				2020			
count	64	64	64	64	64	64	64	64	63	63	63	63
mean	62.73	48.80	41.09	11.20	62.55	48.16	41.78	11.31	62.51	47.83	41.22	13.56
std	3.67	5.47	5.54	7.30	2.70	6.19	6.20	8.04	2.85	8.17	8.05	9.67
min	58.42	28	29	0	58.29	31	26	0	57.43	29	25	0
25%	59.44	46	37	6	60.57	44	38	6	60.07	42.5	35.5	4.5
median	61.82	49	40.5	11	61.49	49	41	10	62.16	48	40	14
75%	65.47	52.25	44	16	64.70	52	46	14	65.15	53	47	20

*BHT - Best Heat Time of a meeting (seconds); *Pts_H* - number of points collected by home team; *Pts_A* - number of points collected by away team; *Pts_diff* - absolute value of a difference ($Pts_H - Pts_A$)

Analyzing Table 1 one can observe similarities in a mean and median of *BHT* across all seasons and some differences in *Pts_diff* descriptive statistics in season 2020. In a next step visual explorations of the features was conducted.

Fig.1 Histograms and KDE - BHT

Fig.2 Histograms and KDE – Pts_diff

Fig.3 Scatter plots and regression lines – Pts_diff vs Meeting_no

Fig.4 Scatter plots and regression lines – Pts_H vs Meeting_no

Fig.5 Scatter plots and regression lines – Pts_A vs Meeting_no

The visual explorations shows that the empirical distributions of both the studied variables differed over the seasons (Fig. 1 and 2).

From Fig. 3 to 5 one can learn that variance of difference in meeting's score was exceptionally high in a season 2020, mostly due to higher variance of home points. That suggest a more balanced ratio of home and away wins. Content of Table 2 confirms that observation.

Table 2. Meeting results - counts

season	Win_H*	Win_A	Draw	ratio
				Win_A/Win_H
2015	47	13	4	0.255
2016	47	14	3	0.280
2017	38	19	7	0.422
2018	49	10	5	0.185
2019	44	17	3	0.362
2020	41	19	3	0.432

*Win_H – number of home wins, Win_A – number of away wins

To explore if season 2020 was structurally different in terms of the mean, variance and empirical distributions statistical test were performed next. Usual normality test was performed at first to check if standard normality assumption could be use in further examination. The results of this test and the values of third and fourth central moments reports Table 3.

Table 3. Normality test, skewness and kurtosis

season	BHT		Pts diff		BHT		Pts diff	
	test stat	p-value	test stat	p-value	skewness	kurtosis	skewness	kurtosis
2015	12.24	0.002	10.74	0.005	0.598	-1.019	1.003	0.368
2016	62.19	0.000	6.41	0.041	0.227	-1.475	0.498	-0.821
2017	9.41	0.009	10.92	0.004	0.579	-0.927	0.867	1.063
2018	12.34	0.002	7.39	0.025	0.544	-1.040	0.697	0.726
2019	9.94	0.007	10.03	0.007	0.423	-1.008	0.892	0.688
2020	23.41	0.000	3.79	0.150	0.209	-1.274	0.571	-0.332

The null hypothesis of that test is: feature comes from a normal distribution. Assuming standard significance level of 5% only in one case we **cannot reject** hypothesis of normality (yellow marked). Similarity of empirical distributions was checked next. The null hypothesis that two seasons have the same underlying probability distribution of examined features was tested by the ES test. The equality of variance and mean was tested in following stage of research. The strategy of testing was testing all season in pairs. Testing season 2020 against cumulative sample of seasons 2015 to 2019 was considered as less restrictive, so driving to less clear conclusions. The results of that test contain Table 4. It reports p-values of tests of season 2020 against all other season in a sample period.

Table 4. Comparative statistical tests - meetings

season	EPS test		test of variance		test of mean	
	BHT	Pts diff	BHT	Pts diff	BHT	Pts diff
	p-value	p-value	p-value	p-value	p-value	p-value
2015	0.001	0.413	0.317	0.320	0.654	0.230
2016	0.000	0.651	0.001	0.881	0.450	0.626
2017	0.148	0.131	0.868	0.037	0.374	0.027
2018	0.028	0.178	0.036	0.005	0.707	0.124
2019	0.662	0.426	0.515	0.045	0.940	0.157

Yellow marked p-values signal rejection of a null hypothesis on 5% significance level.

Second part of the research was the examination of data connected with teams. Considering the aim of the examination two important measures were selected for analysis: number of points collected by team in a meeting (*Pts*) and average age of a team in a meeting (*AAT*), calculated as average age of the riders participating in each meeting. *Pts* is considered as a proxy measure of average team performance in meeting and season, *AAT* is demographic information connected with team management and policy. Main goal of that part was to discover any patterns and changes in that patterns in a relation of age of a team and team performance. Descriptive statistics of examined features are depicted in Table 5.

Table 5. Descriptive statistics - teams

stat\season	2015		2016		2017	
	Pts	AAT	Pts	AAT	Pts	AAT
count	8	8	8	8	8	8
mean	44.70	26.74	44.56	27.21	44.15	27.14
std	2.86	1.92	3.34	1.92	2.84	2.43
min	38.36	23.93	39.00	24.56	40.71	24.08
25%	44.44	25.43	42.68	25.99	41.88	25.52
median	45.04	26.80	44.07	27.43	43.44	26.85
75%	46.17	28.21	47.38	28.19	47.01	28.27
stat\season	2018		2019		2020	
count	8	8	8	8	8	8

mean	44.68	26.22	44.69	26.81	44.31	27.01
std	2.90	1.58	3.59	1.83	4.92	2.00
min	41.50	22.93	39.50	23.99	34.93	24.61
25%	42.36	25.83	42.91	25.80	43.18	25.33
median	44.34	26.29	44.51	27.12	44.84	26.99
75%	46.29	27.13	45.49	27.70	47.15	28.86

In a next step visual explorations of the features was conducted.

Fig. 6 Histograms and KDE – Pts

Fig. 7 Histograms and KDE – AAT

Figures 6 and 7 show interesting oddity of both variables' empirical distributions in season 2020. In case of *Pts* empirical distribution has heavier tails and is more platykurtic than in the other seasons, in a case of *AAT* there is distinctive bimodal distribution.

Table 6. Normality test, skewness and kurtosis - teams

season	Pts		AAT		Pts	
	test stat	p-value	test stat	p-value	skewness	kurtosis
2015	0.10	0.951	83.47	0.000	0.000	0.011
2016	3.87	0.144	5.97	0.051	-0.037	-0.647
2017	8.93	0.012	8.79	0.012	-0.406	1.210
2018	2.28	0.320	7.46	0.024	-0.007	-0.551
2019	0.19	0.912	1.22	0.543	0.006	-0.259
2020	3.66	0.160	5444.97	0.000	-0.028	-0.641

As expected from visual inspection, there is no reason to reject a hypothesis of normality of distribution of *Pts* population (except 2017 season), and one could reject a hypothesis of normality of *AAT* distributions in almost every case. Furthermore, the value of kurtosis in 2020 confirms observation of platykurtic shape of distribution of *Pts* in 2020 season.

Table 7. Comparative statistical tests - teams

season	EPS test		test of variance		test of mean	
	Pts	AAT	Pts	AAT	Pts	AAT
	p-value	p-value	p-value	p-value	p-value	p-value
2015	0.162	0.082	0.021	0.081	0.712	0.151
2016	0.625	0.000	0.189	0.002	0.747	0.216
2017	0.007	0.000	0.002	0.737	0.949	0.725
2018	0.021	0.000	0.005	0.000	0.666	0.001
2019	0.067	0.000	0.009	0.023	0.653	0.664

The main object of that part of research is detection of demographic pattern in teams' performance. Moving to this step author started from visual inspection of relation between variables *Pts* and *AAT*. Fig 8 depicts bivariate KDE plot of *Pts* and *AAT* features by seasons.

Fig. 8 Bivariate KDE – PTS and AAT

Figure 9 contains scatter plots (plus regression lines) of *Pts* vs. *AAT* in all of the seasons in a sample.

Fig. 9 Scatter plots and regression lines - Pts vs. AAT

In Fig. 8 one can notice very interesting and distinctive pattern of KDE in 2020 season. There are two clear clusters of average age and team/points distribution. We cannot observe similar, in terms of size and shape, clusters in the previous seasons. Fig. 9 confirms such pattern via explicit grouping of points collected by teams around two values of average age. Moreover, one can preliminary asses direction of correlation between two variables of interest. In all seasons, except season 2016, one can observe weak, negative correlation between number of points and average age. In a season 2020 there is no correlation at all. Above mentioned observations need be confirmed by deeper investigation. To conduct the research, in a next steps the author utilized statistical test of correlation coefficient and unsupervised learning method of clustering. Table 8 contains values of Spearman rank-order correlation coefficient ρ and p-values for null hypothesis of value of zero of that coefficient.

Table 8. Spearman correlation coefficient - Pts vs. AAT

season	Spearman rho	p-value
2015	-0.23	0.009
2016	0.12	0.168
2017	-0.23	0.010
2018	-0.25	0.004
2019	-0.16	0.077
2020	0.02	0.848

Results reported in Table 8 confirms visual assessment. There is a weak, negative, statistically significant correlation between points collected by team and team's average age in all of the seasons, except seasons 2020 and 2016. In a season 2020 the lack of correlation is evident in terms of both ρ and p-value, when the result for 2016 is more vague.

In a next step the author performed cluster analysis to determine if the age clusters in fact exists, what is the number of clusters and if the median of points collected by the teams in clusters is significantly different. The well-known K-Means algorithm was used. Taking into account the results of visual analysis and analysis of correlation, the cluster study was carried out against two samples - sample of season 2020 and sample encompassing seasons 2015 to 2019. First task was to determine k (number of clusters) in samples. For that two methods were applied: inertia plot and silhouette coefficient plot. Inertia measure the distance between each instance and its centroid. Plotting the inertia as a function of k allow visually determine optimal number of clusters.

Mean silhouette coefficient plot shows value of coefficient for various k . One look for $\text{argmax}(\text{silhouette})$.

Fig. 10 Inertia and silhouette scores – season 2020 (left panel) and seasons 2015-2019 (right panel).

Clearly, in a season 2020 existed two clusters according to the average age of the teams (both inertia and silhouette point to $k = 2$). In the rest of a sample one can observe 5 to 8 clusters (silhouette), with 5 supported by inertia. The next phase of cluster analysis was determination of centroids and testing if the medians of number of points collected by teams in clusters were significantly different.

For medians equality test the Kruskal-Wallis H-test was performed. The null hypothesis of that test states that the population median of all of the clusters in a sample are equal.

Table 9. Analysis of clusters in the samples

Season 2020		Seasons 2015 - 2019			
Centroid cluster 1	28.83	Centroid cluster 1	25.64	Median number of points cluster 3	45.00
Mean number of points cluster 1	44.32	Mean number of points cluster 1	45.77	Centroid cluster 4	27.33
Median number of points cluster 1	44.00	Median number of points cluster 1	45.50	Mean number of points cluster 4	44.74
Centroid cluster 2	25.11	Centroid cluster 2	28.98	Median number of points cluster 4	44.00
Mean number of points cluster 2	44.71	Mean number of points cluster 2	43.05	Centroid cluster 5	31.16
Median number of points cluster 2	45.50	Median number of points cluster 2	43.50	Mean number of points cluster 5	42.29
		Centroid cluster 3	23.79	Median number of points cluster 5	41.00
		Mean number of points cluster 3	45.86		
The Kruskal-Wallis H - statistics	0.099			The Kruskal-Wallis H - statistics	15.87
p-value	0.753			p-value	0.003

Findings presented in Table 9 confirm preliminary observations made on the basis of visualization.

In the final part of a research author focus on individual meeting/rider data. As the results of that part are extensive, this paper contain only the most important ones from the point of view of the aim of examination. The analysis was done in two main directions – influence of a an age of a rider and nationality of a rider on performance in seasons. Each individual meeting/rider data record was tagged as belonging to rider of age (taken at the date of the heat) under or equal 25 years (“Younger” group) and above 25 years (“Older” group) and as “Polish” or “other” (polish nationality or other nationality of a rider). The split age was chosen on a basis of cluster analysis. At first the influence of age of a rider on performance by seasons was examined. Table 10 contain the most important findings.

Table 10. Impact of an age of a rider on exeptions and points collected

Exceptions					Points collected				Correlation coefficient			Binomial test
season	rider group	defects	exclusions	falls	count	sum	mean	median	season	Spearman rho	p-value	p-value
2015	Older	33	44	7	460	3534	7.68	8	2015	0.33	0.000	0.464
	Younger	28	50	7	429	2213	5.16	4	2016	0.45	0.000	0.989
2016	Older	34	32	7	459	3790	8.26	9	2017	0.37	0.000	0.126
	Younger	36	35	2	431	1952	4.53	3	2018	0.53	0.000	0.630
2017	Older	29	41	6	483	3731	7.72	8	2019	0.46	0.000	0.201
	Younger	23	36	8	412	1960	4.76	4	2020	0.48	0.000	0.000
2018	Older	21	39	6	503	4021	7.99	9				
	Younger	25	25	5	492	1732	3.52	2				
2019	Older	18	32	4	504	3857	7.65	8				
	Younger	18	22	6	489	1899	3.88	3				
2020	Older	18	27	4	516	3959	7.67	8				
	Younger	12	27	11	457	1651	3.61	3				

Analyzing exceptions the author assumed that only number of falls could be connected with potential abnormalities of season 2020 (as defects and exclusion are rather caused by endogenous or technical reasons). In fact in a 2020 season one can observed the highest number of falls of “Younger” riders in a whole sample. Column “Binomial test” of Table 10 contain p-values for null hypothesis of equality of probabilities of fall in a group of “Younger” and “Older”, when probability of fall is modelled by binomial distribution. As one can see, only in season 2020 probability of fall of a young rider was significantly greater than probability of fall of older rider.

Moving to analysis of performance of groups of younger and older riders, measured by the number of points collected by season, one can notice that in season 2020 older riders started 516 times (the highest number in a whole sample) and collected second highest sum of points in all of the seasons. Younger riders perform much worse, with the smallest sum of points collected and second smallest mean of points. It suggest that positive correlation between age and number of points collected by the rider should exists. The column “Correlation coefficient” of Table 10 contains both values of Spearman Correlation coefficient and p-values of test for no correlation. In all of the seasons there were statistically significant, positive correlation between variables in consideration. This result may seem to contradict the findings of the second part of the study, which showed a negative correlation or no correlation between the mean age and the number of points (see. Table 8). However, the contents of Table 8 refer to the average age of the teams, while those of Table 10 refer to the

relationship between the average age of the rider and the points scored. So in author's opinion these results supports each other, especially for 2020 season.

In Table 11 some preliminary results on relation of nationality and sport performance of a rider are shown.

Table 11. Impact of nationality on performance

season	nationality	count	sum	mean	std	coefficient of variation
2015	other	324	2694	8.31	4.02	48.3%
	Polish	565	3053	5.40	4.22	78.1%
2016	other	328	2698	8.23	3.89	47.3%
	Polish	562	3044	5.42	4.06	75.0%
2017	other	313	2455	7.84	3.83	48.8%
	Polish	582	3236	5.56	4.02	72.3%
2018	other	372	2637	7.09	4.63	65.3%
	Polish	623	3116	5.00	4.28	85.5%
2019	other	411	2974	7.24	4.31	59.6%
	Polish	582	2782	4.78	4.19	87.7%
2020	other	397	3044	7.67	4.22	55.0%
	Polish	576	2566	4.45	4.10	92.1%

If better average score of international riders are obvious to some extent (one of the reason is that junior riders in PGE Ekstraliga must be Polish), then two numbers of season 2020 (yellow marked) are something extraordinary. In that season international riders collected maximal sum of points in a whole sample and Polish riders demonstrated highest level of variation of points collected in the meetings (international riders showed level of that measure fairly similar to previous ones).

Discussion

The present study is, to the best of the author's knowledge, the first of its kind to be undertaken after the first pandemic season in speedway. Focusing on one of the most important speedway leagues in the world, the Polish PGE Ekstraliga, the study led to the formulation of several important observations regarding the impact of the turbulence of the 2020 season on the results of teams and players.

The analysis showed that season 2020 was similar to the rest of the sample in terms of average results of the teams. However, deeper investigation revealed some distinctive and important patterns.

Examination confirmed that average performance of the best group of riders in league, measured by BHT, was not significantly different in season 2020 and previous 5 seasons in terms of mean time of a heat. It means that winner's group of riders was performing similar to non-covid seasons. However empirical distribution of BHT in 2020 was significantly different in 3 cases of 5, which suggests that there was structural changes in performance distribution. In fact, analyzing skewness and kurtosis of empirical distributions one can say that BHT in was distributed more flat than in previous years (apart 2016 season), which can means that preparation of tracks in 2020 season was more diverse (supports both faster and slower wins) than in previous seasons. Moving to difference in scores (Pts_diff), this metric describes an average sport level of teams in a season. One can observe that internal distributions of that measure are not statistically different across all seasons, however there are significant difference in variability of it (in 3 from 5 seasons) and differences in a mean value. Although only one p-value in a mean test allows to reject the null in 5% level, other p-values are relatively small and can suggest important differences in mean. One can support that findings by results of Table 2 and visualizations of Fig 3, 4 and 5. All together leads to conclusion that 2020 season was more heterogeneous than the rest of the sample in terms of teams and track preparation. Some teams did win by great number of points when others did lose highly, as well. Home track was not so important handicap as usually, which could mean that there were lack of time for training in home objects and lack of careful preparation of home track as well. That findings are in accordance with 2020 season delay and national lockdown in Poland.

Moving to part two of the research, analysis of points and age of teams showed significant structural difference of season 2020. Shape of empirical distribution of points collected by teams confirmed increased variation in scores and a lack of concentration around the average number of points. There was a weak negative, statistically significant correlation between points collected by team and team's average age in all of the seasons, expect season 2020 and 2016. In a season 2020 the lack of correlation is evident in terms of both ρ and p-value, when result for 2016 is more vague. Cluster analysis confirmed that in a season 2020 there were two clusters according to the average age of the teams. One can describe cluster 1 as "older" teams' cluster, with average age of a team equal to 28.83 years. Cluster 2 can be named as "younger" teams' cluster with average age of 25.11 years. Median number of points collected by the teams in clusters were: 45.5 for "younger" and 44 for "older", however from the Kruskal – Wallis test one can learn that the difference in medians is not statistically significant. It implies that both groups of teams performed equally well. From managerial point of view it could mean that there were two equally good team building strategies in terms of age structure in season 2020. In a contrary to season 2020, in a second sample there were 5 clusters of teams (on average) with significantly different effectiveness measure as median of points collected.

The most important conclusions from the third part of the research were as follow. Only in season 2020 probability of fall of a young rider was significantly greater than probability of fall of an older rider. It could

mean that lack of track trainings and sparing meetings in a pandemic season of 2020 was particularly severe for the preparation for riding in a group of younger riders. In a season 2020 older riders started in 516 heats (the highest number in a whole sample) and collected second highest sum of points in the all seasons. Younger riders performs much worse, with the smallest sum of points collected and second smallest mean of points.

Taken together results of part two and three, it implies two different but equally effective team building strategies - younger teams, which consisted of only a few older riders who scored above average, and older teams consisting of several older riders, who scored on average. Taking into consideration internationalisation of teams, examination showed that international riders were similarly intensively exploited in season 2020 as in season 2019 but their performance were better and their shape was more stable. Polish riders showed highly unstable shape (high variance of points collected). These two facts lead to the conclusion that from the managerial point of view, in a lock-down conditions managers and coaches need to pay more attention to the preparation of the Polish riders and should exploit stable shape of international riders.

Conclusions

The significant impact of the extraordinary external conditions of the 2020 season on various sports seems obvious, but its exact determination requires scientific research. This article covers such preliminary research on speedway.

A statistical analysis of the meetings in Polish PGE Ekstraliga in the 2020 season and in the years 2015-2019 revealed the facts about the difference between the first season of the pandemic and the previous ones. An examination showed that average performance of the best group of riders in the league, measured by best time of heat in a meeting, was not significantly different in season 2020 and previous 5 seasons in terms of mean time of a heat. It means that the leading group of riders was performing similar to non-covid seasons. However empirical distribution of best times in 2020 was significantly different from previous years in 3 cases of 5, which suggests that there was structural changes in performance distribution. Values of both skewness and kurtosis of empirical distribution suggest more flat distribution than in previous years (apart 2016 season), which can mean that preparation of tracks in 2020 season was more unequal. Teams' average sport level measured by difference in scores occurred to be more diverse in pandemic season as well. It leads to conclusion that 2020 season was more heterogeneous than the rest of the sample in terms of teams and track preparation. That findings are in accordance with 2020 season delay and national lockdown in Poland which caused lack of time and conditions for trainings and careful track maintenance. Another significant consequence of the disruptions of the 2020 season was the greater number of falls of younger riders. The probability of fall of a young rider was significantly greater than probability of fall of an older rider in pandemic season only. It could mean that lack of active preparation was particularly severe for the younger riders. Taking into consideration nationality of riders in the teams, examination showed that foreign riders were similarly intensively exploited in season 2020 as in season 2019 but their performance were better and their shape was more stable. The shape of Polish riders was more diverse.

To summarize, the study revealed that the first pandemic season were in fact different than the other seasons in Polish PGE Ekstraliga taking into considerations various measures describing performance of the teams and the riders. Observed oddities drive to managerial implications formulated in a discussion section.

The results of this and similar studies seem to be important because they allow us to answer the question of how clubs, teams and players adapted to the changed conditions of functioning during pandemic. These findings can help club managers and coaches formulate more effective policies for such circumstances in a future.

References:

- Biggest Global Sports report. (2021). retrieved from: <http://www.biggestglobalsports.com/by-countryregion/4585580681>, on 09.02.2021
- Błaszowska, A. (2018). *Toruński sport żużlowy w latach 1930–1989*, [Speedway sport in Toruń in years 1930 – 1989], Rozprawy Naukowe Akademii Wychowania Fizycznego we Wrocławiu
- D'Agostino, R. and Pearson, E. S. (1973). Tests for departure from normality, *Biometrika* 60, pp.613-622
- Epps, T. W., Singleton, K. J. (1986). An omnibus test for the two-sample problem using the empirical characteristic function, *Journal of Statistical Computation and Simulation* 26, pp. 177–203
- Forgy, E., W. (1965). Cluster analysis of multivariate data: efficiency versus interpretability of classifications, *Biometrics* 21 (3), pp. 768–769
- Gulak-Lipka, P. (2020). Internationalization and managing diversity on the basis of professional basketball clubs, *Journal of Physical Education and Sport @ (JPES)*, Vol.20 (6), Art 484, pp. 3591 - 3598
- Hastie, T. Tibshirani, R., Friedman, J. (2009). *Elements of Statistical Learning Ed. 2*, Springer
- Levene, H. (1960). In *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, I. Olkin et al. eds., Stanford University Press, pp. 278-292
- Lis, A., Tomanek, M. (2020). Sport management: Thematic mapping of the research field. *Journal of Physical Education and Sport*. Vol 20 (Supplement issue 2), Art 167, pp. 1201 – 1208

-
- Lloyd, S. P. (1982). Least squares quantization in PCM (PDF). *IEEE Transactions on Information Theory*. 28 (2), pp.129–137
- May, C. (1978). *Ride it ! The Complete Book of Speedway*, Haynes Publishers, UK.
- PGE Ekstraliga Statistics. (2021). retrieved from: <https://speedwayekstraliga.pl/statystyki/frekwencja/?y=2019>, on 09.02.2021
- Siitonen, A., Mutanen, A., Halonen, I., Värri, V-M. (2020). *Regarding the philosophy of speedway*, in: Human and Nature, eds. Mutanen, A., Friman, M., Kantola, M., Konst, T., Research Reports from Turku University of Applied Sciences 50.
- Spearman C. (1904). The proof and measurement of association between two things, *American Journal of Psychology*, 15 (1), 72–101.
- Sprent, P. Smeeton, N.C. (2001). *Applied nonparametric statistical methods*. 3rd ed., Chapman and Hall/CRC. 2001