POSSIBILITIES OF USING SELECTED VISUALIZATION METHODS FOR HISTORICAL ANALYSIS OF SPORTING EVENT – AN EXAMPLE OF STAGE CYCLING RACE TOUR DE FRANCE

Vladimír Bačík, Michal Klobučník

Department of Human Geography and Demography, Comenius University, Bratislava, Slovakia

ABSTRACT: Nowadays, we are facing an enormous amount of data which are being produced by different systems. These must be effectively stored and visualized in order to enable their proper interpretation. In our research, we attempted to use circular data visualization applied on the Tour de France cycling race. Since this competition has over 100 years of history, a variety of data is available. Using the chord diagrams we tried to illustrate the long-term development of this race, with an emphasis on its spatial and sport aspects. Spatial dimension is characterised by the large number of stages and mountain climbs in such locations which altogether enhance the meaning of this major sporting event, while sport aspects aim to capture the significant personages on the scorecard. Circular visualization has found its application in many disciplines (genetics, demography, medicine, etc.). In our contribution, we point out its importance also in the visualization of the historical milestones of the most important multi-stage cycling event in the world.

KEY WORDS: visualization, mapping, circular visualization, database, Tour de France

Corresponding author: Vladimír Bačík, vladimir.bacik@uniba.sk

Introduction

The Tour de France is one of the most significant global sporting events, and also a worldwide social phenomenon. Since its first edition in 1903, many changes, which are related to the technology, organizational and social aspects of the relevant time segment have been made. From the very beginning, the Tour de France has enjoyed a large audience among both the general public and professionals. One of the results of this is many publications dealing with different aspects of the Tour de France. A considerable emphasis on the spatial perspective of Tour de France is observed in the contributions of Fumey (2006), Campos (2003), and Reed (2007). The authors of these papers devote significant attention to linking the race and its impact on the space where the Tour de France is organized. In addition to the spatial aspect, also studies analyzing the event in specific broad areas, mainly the economic aspect Reed (2003), Desbordes (2007), the social aspect (Balduck et al. 2011), or last but not least, the physiological aspect (Lucia et al. 2003), get into the issue. The history of Tour de France can be recorded in different ways. Most of the publications map the overall history of the race (Mignot 2015, McGann and McGann 2006, 2008), whereas several were published on the occasion of the 100th anniversary of the event (Dauncey and Hare 2003, Wheatcroft 2005,

Whittle 2003). Other publications focused on specific aspects of the Tour de France, which are necessarily linked with this event. They are publications and contributions dealing with particular stages (Allchin and Bell 2003), important legends and cyclists (Fife 2008), or demanding mountain stages (Sidwells 2009) and significant mountain crossings and hills (Yates 2006), which make this event attractive even from the viewers' point of view. One of specific issues in the Tour de France history is also that of doping, which was detailed in the papers of Schneider (2006) and Mignon (2003). This topic is also discussed very frequently today and presents a specific research issue not only in the field of cycling but also in the field of sport. The geographic dimension of the event is accented in the work of P. Boury (1987), where he focused primarily on selected spatial aspects of Tour de France. Therefore the Tour de France history going back to the early 20th century provides us with preconditions to be recorded and presented in the most transparent possible form. With its grandiosity and interest, the Tour de France has markedly gone beyond its national sphere, thus becoming a global event, albeit one which always keeps its basic identity. Scientific contributions focus considerably on analyses of this sporting event from different research perspectives. One of them is the spatial perspective that we will attempt to present in the paper, using technologies applied for the distribution of spatial data. The actual relation between geography and sports is strong attention devoted in the works of Bale (1994, 2003), High and Hinch (2006). These studies accentuate the importance of geographic research in the study of sport events of global importance. In addition to our cartographic interpretation, we also visualized spatial and statistical data with diagrams developed based on the Circos project. It allows the data circular visualization in various research fields. These diagrams are often used by visualization in the field of biology and genetics, as exemplified by the studies of Krzywinski (2009), Russell (2010), Darzentas (2010), and Ledford (2010). However, recently this method of visualization has also been applied in social sciences where space plays an important role. An example could be the studies of Abel and Sander (2014), as well as Sander et al. (2014), in which migration data are shown in this interesting manner.

Research objectives

In the preceding part, we pointed out many different approaches revealing the Tour de France's position in the global context. One of the methods with which the history of Tour de France can be studied is the geographically based approach.

Geography as a disciplinary field uses many advanced methods for the collection and distribution of various data. One of the main means of expression for geographical information is the cartographic interpretation of phenomena in the form of a map output. At present, a high quantity of tools and techniques can be used for the development, preparation and final processing, only in standard printed form, but also in the online environment.

From its origination in 1903 to the present, the Tour de France has included many stages and mountain ascents in its itinerary. Significant stage centres and legendary mountains passes contribute considerably to the exceptional position of this event in the sporting world. Not many sporting events have such a high variability of possible places for direct sports competitions. Many sporting events require direct investments in the development of necessary sports grounds and relevant infrastructure (stadiums, race circuits, golf courses, etc.). On the contrary, Tour de France races are held in the places of everyday life, utilizing the existing infrastructure in the country with minimum direct investments in the region concerned. However, from the point of organization and logistics, it is very difficult to organize such an event as there are struggles of cyclists in a different environment on a daily basis. This is preceded by many months' preparation of the complete itinerary covering, among other things, a detailed organization of transport in the places concerned, the routing and an exact schedule for the movement of an accompanying convoy, the development of the "stage city" at the stage finishes or starts, providing the technical infrastructure required for the transmission, and many other activities. Thus the area where the Tour de France is performed presents an important part of this global event. In the context of sporting performance, mainly the relevant cyclists' performance is monitored, either in transit flat terrain stages with the mass finish of the entire peloton, or,

on the contrary, in competitions in sharp Alpine or Pyrenean ascents. From 1903 to the present, several thousands of cyclists stood at the start of these races, but only some of them indelibly went down in the history of this event.

When developing this paper, we asked the key question: How is it possible to present the most important historical aspects of the selected sporting event (the Tour de France) in an effective and interesting manner? These aspects can be perceived in the form of two basic dimensions – both spatial and sports. The spatial dimension is understood as the identification of the most important stages and mountain ascents which have a significant presence in the history of this race. The sports dimension can be perceived as the identification of the most successful rides and countries in the Tour de France history. By linking these two dimensions (the spatial and the sports) we will reveal a new, possible presentation of Tour de France history[1]. In addition to this output, the developed website (tdfrance. eu), which shows many historical data from the Tour de France in interactive form, is an important part of the paper. It is the online environment which presents one of possibilities for dynamic representation of the results of such an extensive data set. Such an approach can be used for the analysis of any other sporting event in which space constitutes a significant component of its history. Naturally, the methods used can find their place in any research in which effective processing and subsequent visualization of (not only) geographical data in printed or online form is inevitable.

A brief description of the use and an active presentation of methods used in geographical research and in analysing the rich history of a sporting event can be understood as the key objective of the paper. We will attempt to reveal certain specific features of this important event with these methods and to discover the time and spatial peculiarities of Tour de France from its origination to the present. The main output, in addition to the contribution, is also a cartographic interpretation of Tour de France history, as well as the presentation of the results obtained. This

we also consider useful in the dissemination of research results as the Internet has a major social impact and the visual interpretation of results in such an environment is available for any population group.

Methods

The presentation of the methodology forms the key part of the paper. We can also include the map, which expresses the above stated linking of the spatial and sports dimensions of Tour de France (see the Chapter Map design), as the main output. The methods used in geographical research are not limited only to the printed (or offline) expression of a phenomenon under review. Various methods of the visualization of phenomena on the Internet are part of advanced methods. Here, tools designed for the storage and effective handling of data (the database system), and various visualization methods allowing clients to see the characteristics dynamically, can be considered as highly important. In this chapter, we will deal briefly with the most important characteristics of particular methods which preceded the development of the final map output.

Data processing

The final map development was preceded by an extensive process covering the processing of historical data from particular Tour de France editions. For the purpose of efficient data handling, as well as due to the attempt to distribute the obtained results on the Internet in the form of a website dedicated to the Tour de France[2], a database containing 10 tables and 76 attributes (Fig. 1) has been developed. In order to visualize particular statistics on a map, the most important step was to allocate geographical coordinates to particular stages and mountain ascents. To this end, we primarily applied the Google geocoding service. Google Earth environment was used to localise mountain ascents or incorrectly allocated coordinates with specific places, in which case particular coordinates were corrected manually. In general,

---

[1]   Such linking is shown in a map output which can be downloaded at http://www.tdfrance.eu/download/tdf_map_bk.pdf.

[2]   The website dedicated to the Tour de France prepared by the authors of the paper, is available at: www.tdfrance.eu.
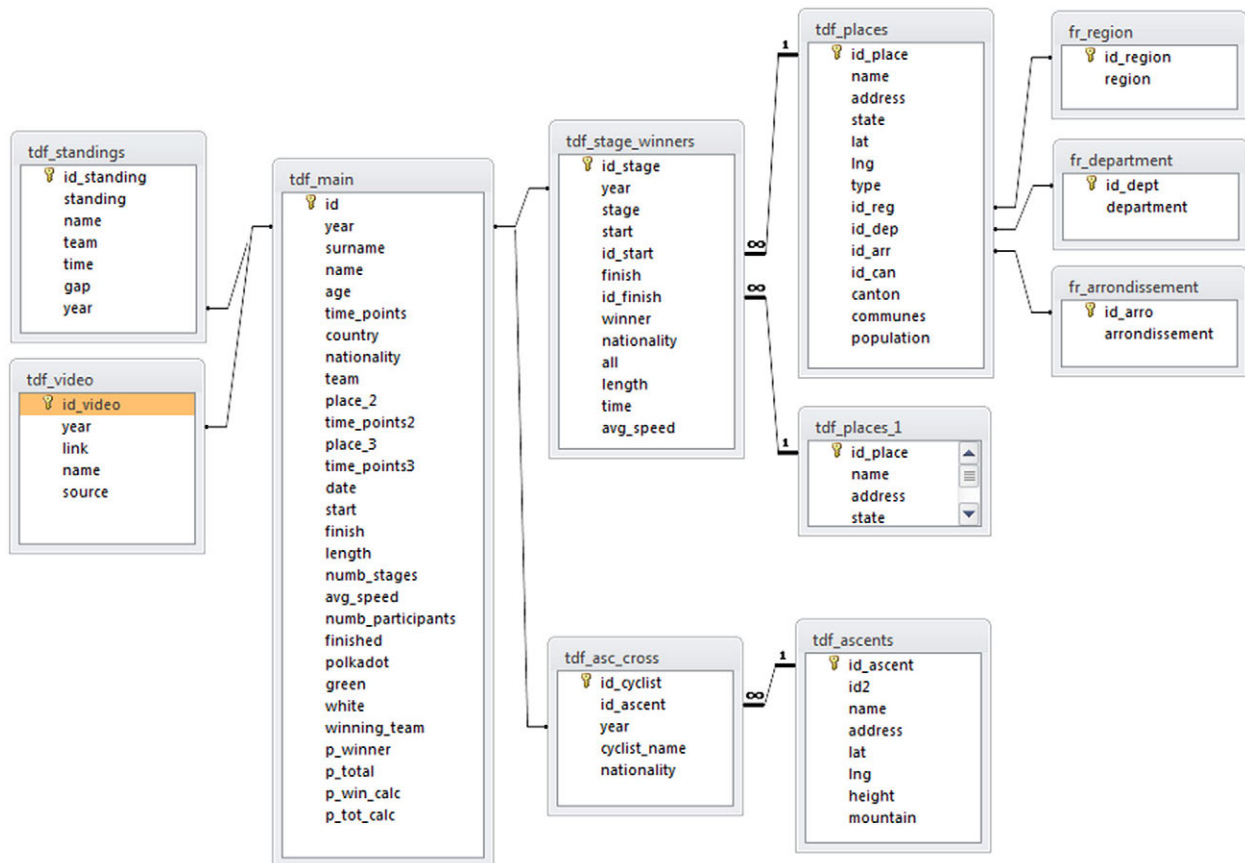
Fig. 1. A scheme of final database tables (the MS Access environment).

Table 1. Examples of selection of necessary data based on query above database tables.

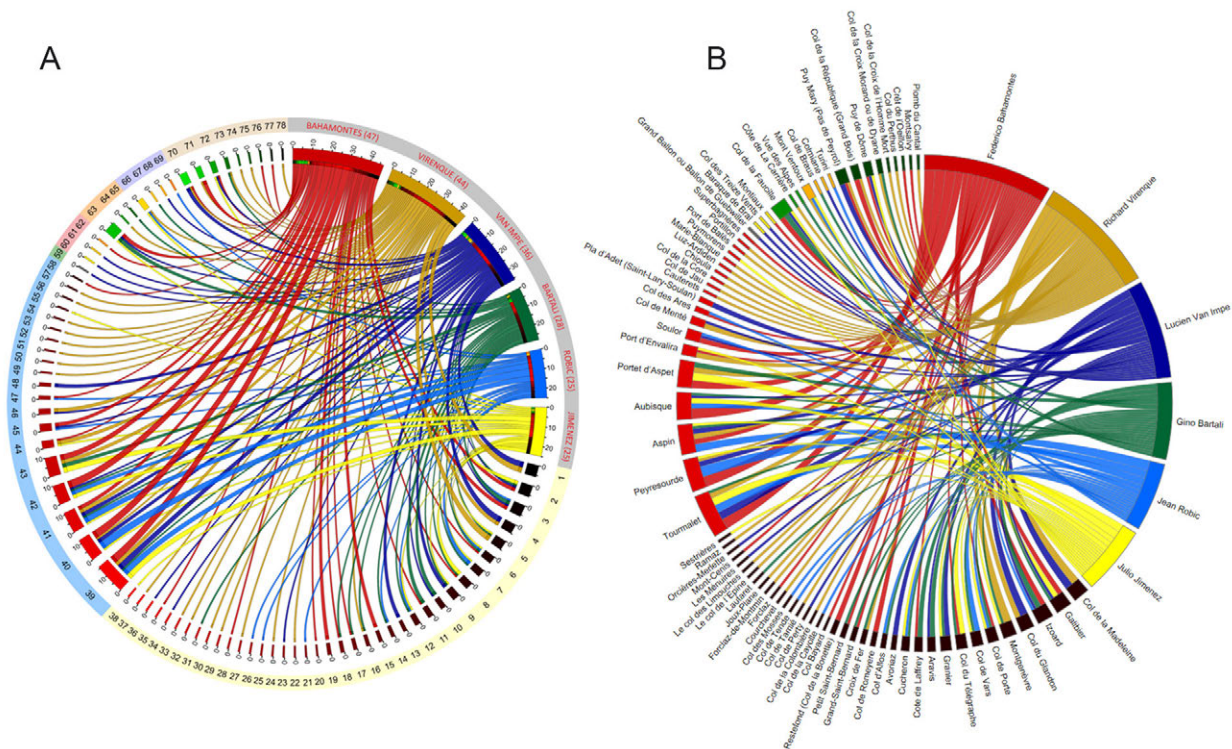| Question	SQL syntax
How many stage finish places were included in the itinerary of the Tour de France before World War II?	Select *,COUNT(finish) AS count_places from tdf_stage_winners,tdf_places WHERE tdf_stage_winners.id_finish=tdf_places.id_place AND year >= 1903 AND year <= 1939 GROUP BY id_finish ORDER BY `count_places` DESC");
How many cyclists won in individual stages in all editions of the Tour de France?	Select winner,nationality,COUNT(winner) AS countw from tdf_stage_winners WHERE winner != ,' AND year >= 1903 AND rok <= 2015 GROUP BY winner ORDER BY `countw` DESC");
What are the most successful nations as regards the number of triumphs in mountain crossings?	Select nationality,COUNT(cyclist_name) AS countc from tdf_asc_cross WHERE year >= 1903 AND year <= 2015 GROUP BY nationality ORDER BY `countc` DESC");
Which cyclists triumphed more than 1 time in Paris?	Select winner,COUNT(*) AS countw from `tdf_stage_winners` WHERE id_finish = 434 GROUP BY winner HAVING countw>1 ORDER BY `countw` DESC
Which are the 10 most frequently climbs included in the Tour de France editions after 1980?	Select *,COUNT(cyclist_name) AS countcn from tdf_ascents,tdf_asc_cross WHERE tdf_ascents.id_ascent=tdf_asc_cross.id_ascent AND rok >= $1980 AND rok <= $2015 GROUP BY id_ascent ORDER BY `countcn` DESC LIMIT 10
Which are the 5 most frequently starting places included in the history of the Tour de France?	Select *,COUNT(start) AS counts from tdf_stage_winnrs,tdf_places WHERE tdf_stage_winners.id_start=tdf_places.id_place AND year >= 1903 AND rok <= 2015 GROUP BY id_start ORDER BY `counts` DESC LIMIT 5

Fig. 2. An example of data circular visualization using a chord diagram (In this picture places of the triumphs of the most successful climbers in the Tour de France history are shown).
A – Offline presentation (available on the main map at: http://www.tdfrance.eu/download/tdf_map_bk.pdf).
B – Online presentation (available at http://www.tdfrance.eu/chord_diagrams_ascents_cyclists.php).

within 103 Tour de France editions organized up till now, we can identify a total of 682 stages and 287 categorised mountain ascents[3]. The design of database tables and the spatial reference of particular places was a basis for the cartographical preparation of the final map. For the purposes of selecting the required data from the database, standard SQL orders were applied, whereas particular selections of required data shown on the map were prepared using PHP and MySQL.

Database tables were designed to allow effective handling of these data on the created website. An important consideration in the design was the fact that these attributes saved in tables must allow answering any questions related to the history of the race under review. The standard SELECT command, based on which we obtain the final presentation of necessary data, was used for particular queries. A basic example of the use for practical questions can be found in Table 1.

The website (tdfrance.eu) is designed so that a majority of answers to these questions is

available to clients who are interested in such information. Certain functions and attributes are available to the administrators only; however, their implementation is only a question of the presentation of the function on the specific subpage.

Chord Diagrams

The chord diagrams were used to illustrate the quantity of achievements of relevant cyclists or nations in a specific stage or mountain ascent on a main map document. In this case, data from the database were stored (after the relevant selection taking into account the necessary criteria) as simple *.txt files. The diagrams were developed using the Circos project, the author of which is Martin Krzywinski (Genome Sciences Centre, Vancouver). The primary aim of the Circos project was circular genome visualization, however, such an interpretation also has a wide-ranging application in other scientific disciplines. This form of data processing can be interpreted easily and intuitively and, in combination with the map visualization, constitutes an interesting tool

---

[3]   Our analysis processes only ascents included in the second, first and "hors catégorie".
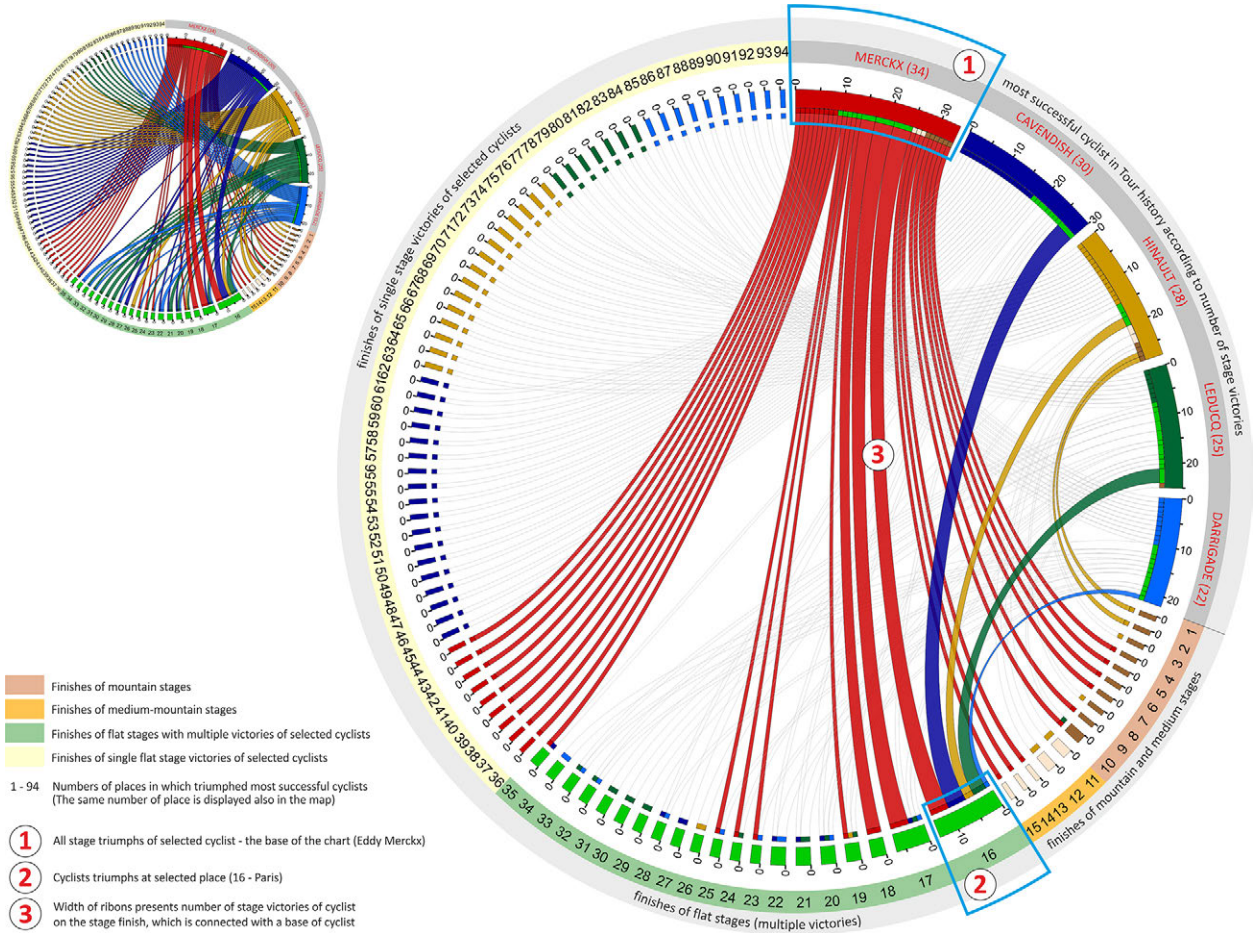
Fig. 3. Interpretation of the chord diagram (The original diagram is visually adjusted so that readers can see the main elements clearly).

for the presentation of a high quantity of not only genome data but also spatial data. Diagrams generated through the online service[4] could be later modified as necessary in appropriate software (for example CorelDraw).

The internal structure of a cyclist´s victories (and nationalities) in each finish places and mountain ascents, is captured by using a circular visualization, in the form of chord diagrams. This form of presentation is interesting and allows the display of large amounts of data. Data from the database (MySQL tables) were adjusted to the required form using standard SQL commands. The resulting datasets were transformed into the output format (*.txt file in the case of an offline chord diagram on the map, *.json file in the case of interactive online presentations). D3 JavaScript library was also used in the online presentation as

well as on the online maps (see section D3 and TopoJSON). The result is a circular visualization of selected historical data from Tour de France (Fig. 2).

Circular visualization is an useful tool for a displaying large amount of data. The interactive online environment allows the client to read a specific database using a variety of scripting languages (e.g. JavaScript). In offline mode, the client must properly interpret the displayed data. Basic interpretation of this circular visualization is shown in a chord diagram showing the most successful cyclists in the history of the Tour de France, according to the number of winning stages (Fig. 3).

The outer circumference of diagram is formed by individual matrix entities (cyclists and stage finishes, where the selected cyclists triumphed). Cyclists are ordered by the number of winning stages. Finish places are divided into four basic categories (see Fig. 3 legend). The base width of

the relevant cyclist (number 1 in Fig. 3) corresponds to the number of stage triumphs (e. g. Eddy Merckx – 34 stage triumphs). Under the base are small boxes of different colours, with the colour corresponding to the type of finish place in which the cyclist triumphed and the thickness expresses the number of stage wins at a given location. The cyclist and the finish place are connected with a ribbon of appropriate width (number 3 in Fig. 5). In the case of a finish place, the width of the base expresses the number of the observed cyclist's triumphs at this place. The individual segments reflect the triumphs of concrete cyclists (number 2 in Fig. 5). In this example we can see that the observed cyclists reached altogether 14 stage victories in Paris (Eddy Merckx and Mark Cavendish – 4 victories, André Leducq – 3, Bernard Hinault – 2 and André Darrigade – 1 triumph). The city of Paris is the only common place where all of the above cyclists triumphed (Merckx – 4 victories, Cavendish – 4, Leducq – 3, Hinault – 2, Darrigade – 1). Here, however, we can say that it is just the Paris triumph that is one of the most valuable victories for cyclists participating in the Tour de France. At this point, it is appropriate to draw attention to the basic principle of the final cyclist´s selection. There was only one criterion of this selection and it was the total number of stage triumphs. Based on data organized in a database, it would be easy to add similar characteristics to any group of cyclists. For example, five and three-time Tour de France champions (Eddy Merckx, Bernard Hinault, Jacques Anquetil, Miguel Indurain, Louison Bobet, Philippe Thys and Greg Lemond) triumphed altogether in 119 stages at 82 different places. In addition to basic quantitative characteristics of chord diagram we can also indirectly reveal the specific nature of these cyclists' racing techniques. Mark Cavendish (30 stage victories) and André Darrigade (22) are typical representatives of sprinters. On the contrary, the triumphs of Eddy Merckx (34), Bernard Hinault (28) and André Leducq (25) demonstrate their universality and abilities to win even demanding mountain stages. A specific feature is the triumphs of Bernard Hinault who gained many of his wins in time trials and in initial prologues (20).

Circular visualization of these historical data represents an attractive form of data representation and its interactive design on Internet allows to the client direct identification of values. Integral part of the resulting map are four chord diagrams, while in the legend is also a reference to the online version of each chord diagram.

D3 and TopoJSON

D3 (Data Driven Documents) is a JavaScript library for producing data visualizations of extensive amounts of data. D3 allows direct control of data while using the standard model of an object, the so-called DOM (Document Object Model). Developers can link input data to any elements of a document using D3, apply dynamic transformation, and modify the content (Bostock et al. 2011). The TopoJSON format is a text file that encodes the topology between particular map objects. At present, the universal nature of the D3 library in a combination with the presentation of spatial data in the TopoJSON format, can be considered as an effective and user-friendly method of cartographic visualizations of data. The previous section of the paper mentions the D3 library with respect to the online version of circular data visualization. This library was also used jointly with the TopoJSON format for an online presentation of maps (final points, mountain ascents). Such an online presentation has several benefits in comparison with a static map. We consider easy identification of selected places (final points, mountain ascents), and the linking of these objects to the database with relevant attributes, as the greatest benefits. Thus the client can easily display relevant data related to the selected place, and identify historical statistics. For the presentation of such a map, we must integrate basic libraries (topojson.js, d3.js) into the source file, while the libraries contain the complete functionality required for the display of the map (or of another object) on the part of the client. In addition to the above libraries, it is necessary to download the source data required for the map. The presentation of the map and the definition of variables corresponds to the standards of JavaScript. A model presentation of such a code is as follows[5]:

---

[5]   A detailed record of source code including defined variables can be shown directly on the website http://www.tdfrance.eu/json_finish.php.

```
<script>
d3.json(„bin/world.topojson", function(error, topology) {
d3.csv(„bin/finishes.csv", function(error, data) {
 g.selectAll(„circle")
.data(data
.sort(function(a, b) { return b.finishes- a.finishes; }))
.enter()
.append(„a")
.attr(„xlink:href", function(d) {
return „http://www.tdfrance.eu/town.php?id="+d.ID;}
                        )
.append(„circle")
.attr(„cx", function(d) {
return projection([d.lon, d.lat])[0];
})
.attr(„cy", function(d) {
return projection([d.lon, d.lat])[1];
})
.attr(„r", function(e) {
return radius(e.finishes);
})
.attr(„class", „bubble")
.attr(„fill", function(f) { return color(f.finishes); })
.append(„title")
.text(function(f) {
return f.place
+ „\nstate: „ + f.stat
+ „\nNumber of finishes: „ + f.finishes;
});
});
</script>
```

The supporting map of the countries in the world is in the format *.topojson (world.topojson). Final points with relevant attributes are saved in *.csv file (finishes.csv). This contains several important attributes. The latitude (lat)

and longitude (lon) defines the exact position of the relevant final point on the map. The attribute "ID" is the basic identifier of the given place which can show detailed statistics of the place as this value is identical to the attribute "id_place" in the database table "tdf_places" (see the section Data processing). A model final presentation of such a map is on Figure 4.

Map design

The basis for creating the final map was obtaining and efficiently redistributing all data. All necessary data were transformed into map layers (*.shp format). This map, consisting of four partial maps and diagrams, we have attempted to present the most important sport achievements of individuals, including the countries in terms of stage triumphs and victories in mountain crossings, and also to illustrate the places of such historical moments. Such cartographic processing can be viewed as innovative in terms of the perception of Tour de France history and could be inspiring also for other analyses covering the history of sports in all sectors where – as for the Tour de France – space plays an important role.

On the final map (Fig. 5) we intended to visualize the main spatial-historical characteristics of the Tour de France, while showing also the
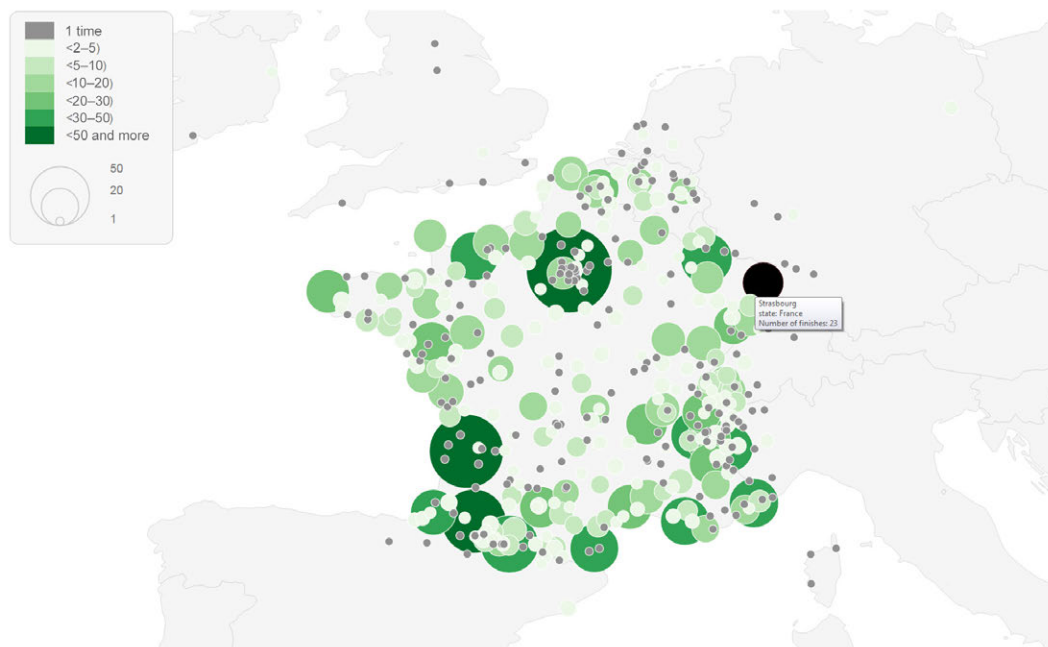

Fig. 4. An identification of stage finish places using an interactive map in the TopoJSON format
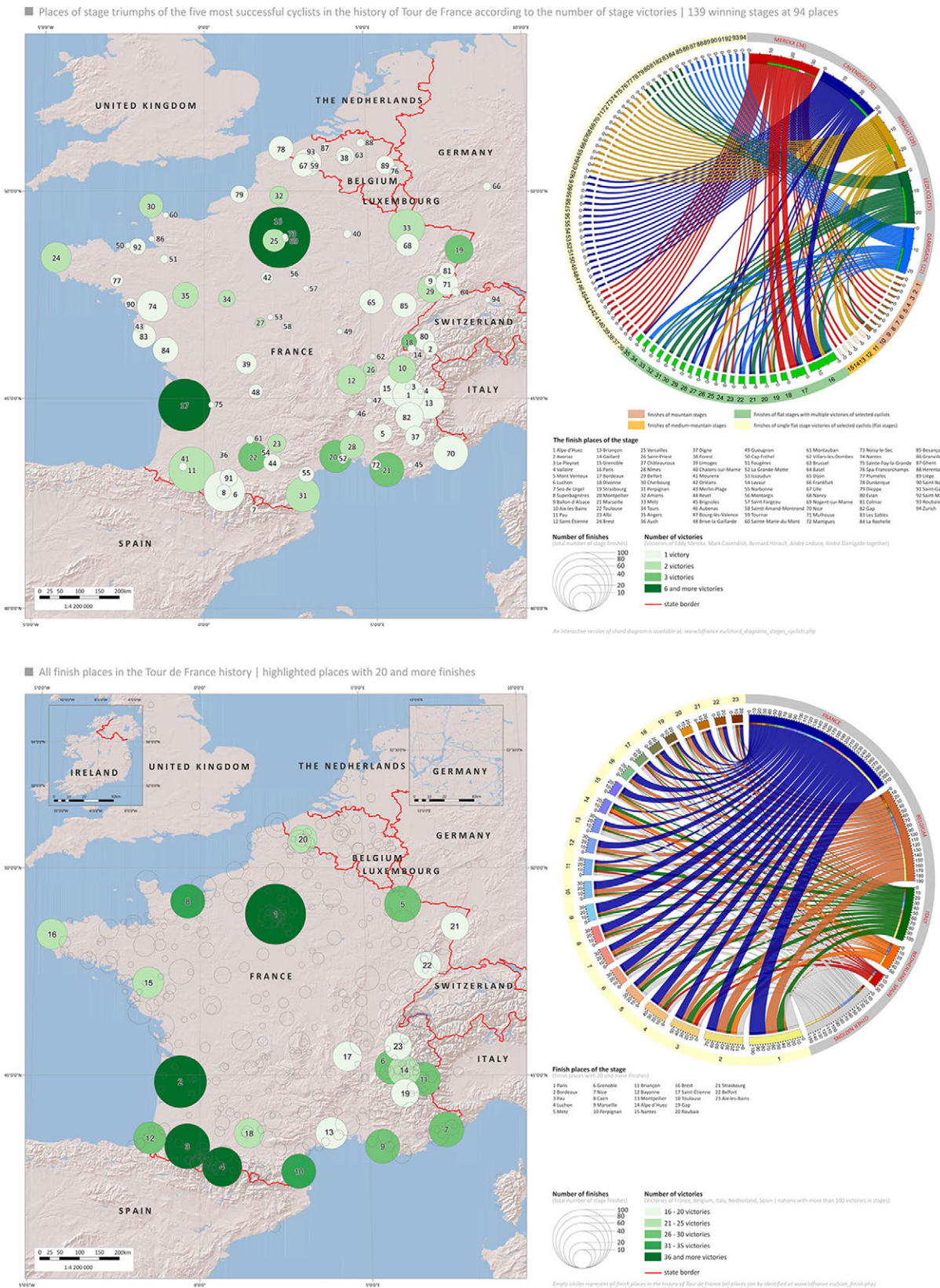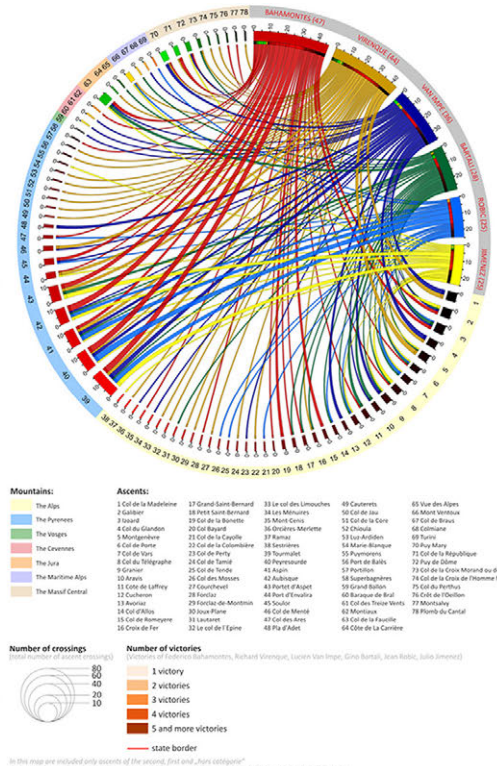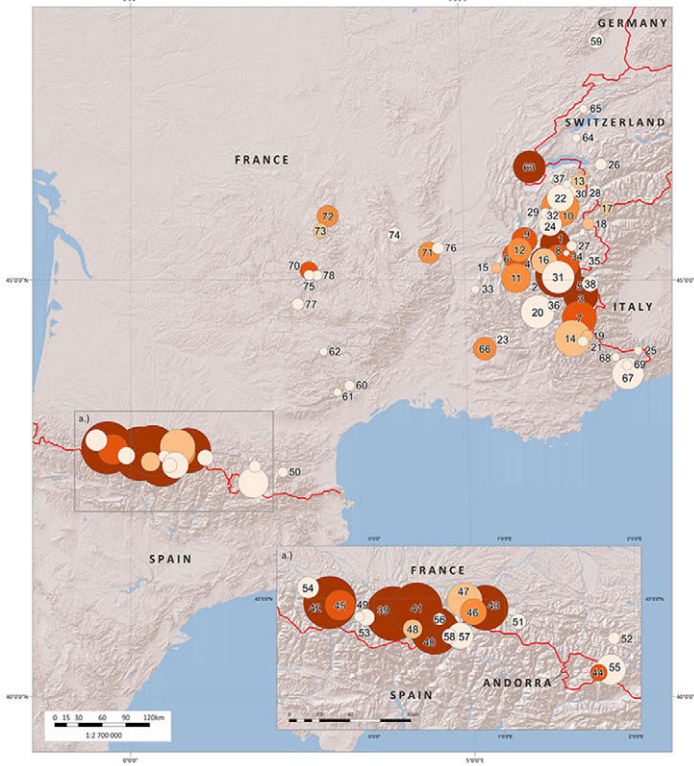(Interactive versions is available at http://www.tdfrance.eu/json_finish.php).

Fig. 5. Preview map of selected historical statistics of Tour de France (1903–2016).
Final map designed for A0 format (is available at: http://www.tdfrance.eu/download/tdf_map_bk.pdf) shows
historically the most successful cyclists and nationalities according to the number of stages victories and the winning
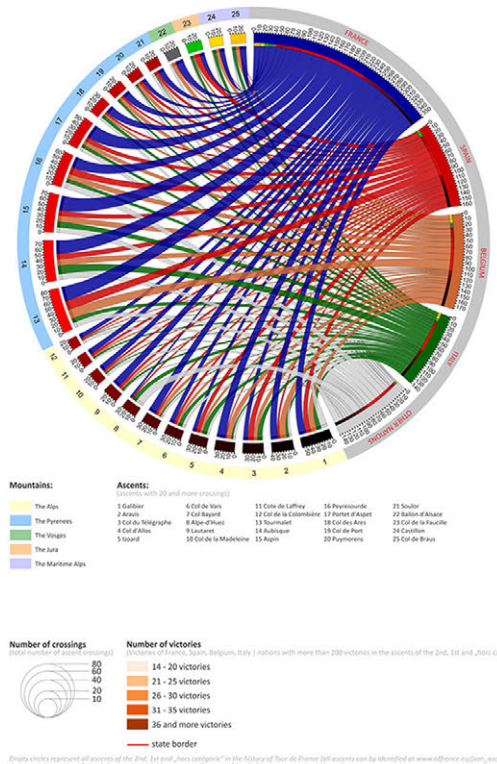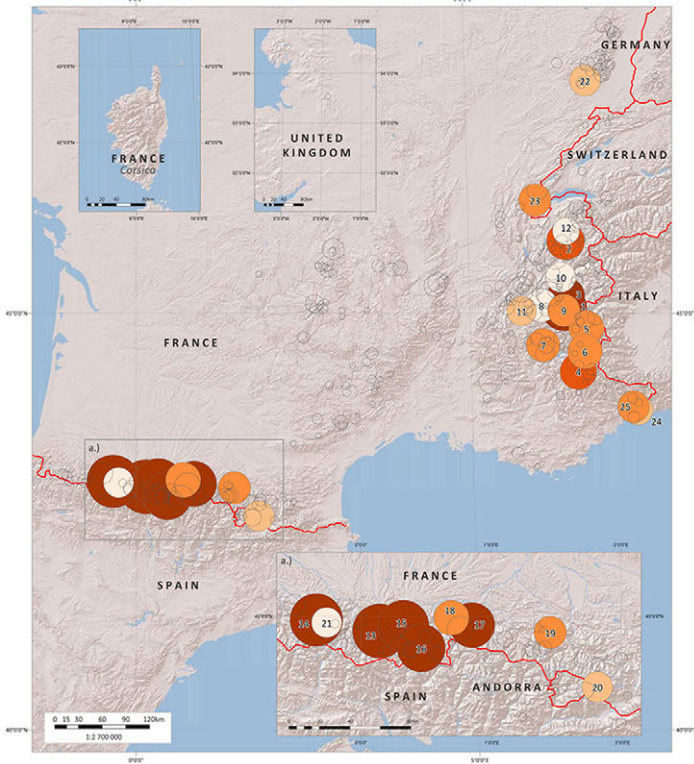mountain ascents crossing othe second, first and "horse" category.

Fig. 5. continued

most successful cyclists in history as well as the most successful countries. Therefore the key map layout (designed for A0 format) is divided into two sections (4 maps, including relevant chord diagrams). The left section presents stage triumphs; the right section presents triumphs in mountain ascents. The upper parts of particular sections show the places of stage triumphs of the most successful cyclists in terms of the number of stage triumphs, or winning crossings through mountain ascents of the second, first and "hors catégorie". The bottom parts present the victories of relevant countries, in which they show all the places of stage finishes or of mountain ascents in the history of the race (1903–2016). However, the most important stage finishes (20 and more stage finishes) and mountain ascents (20 and more crossings) are highlighted. The symbols of these places are colour-coded by the quantity of stage triumphs or of triumphs in the mountain ascents of the most successful countries in the history of Tour de France (for stages it was France, Belgium, Italy, the Netherlands and Spain – all countries with more than 100 winning stages; for mountain ascents it was France, Spain, Belgium and Italy – all countries with more than 200 winning crossings in mountain ascents). To present an internal structure of triumphs, we used a non-traditional, but in our opinion, an adequate interpretation using chord diagrams. The figures presenting places in the diagram and on the map are identical, and thus the number of triumphs of the countries under review (cyclists) in the relevant place (the map) as well as the share of wins of particular countries (cyclists) in the relevant place (the chord diagram), is easily identifiable. The size of diagrams is expressed on all maps based on the total number of stage finishes or of mountain ascents. All of the maps applied the Robinson projection; the scales of the main maps of stage triumphs are 1: 4 200 000; for mountain ascents, the scale was 1: 2 700 000. Since the numerical or verbal identification of particular places on the maps would result in a lack of clarity of such maps, at www.tdfrance.eu we developed an interactive map presentation showing the maps. These are linked to the database, so also an exact identification of the triumphs of particular cyclists in these places is possible. An online version of the corresponding chord diagrams which are part of the map (the hyperlink is in the map)

was also created. However, the maps have been designed to visualize key historical achievements (most successful nations and cyclists). The final map showed several historically relevant statistics. For the achievements of particular racers in stages or in mountain ascents, part of the legend is a list of all triumphs of such cyclists shown on the map. For the wins of particular countries, it is possible to identify directly on the map only stages with 20 and more stage finishes (or ascents with 20 and more crossings) in the whole history. However, the map shows all stages and mountain ascents included in the race itinerary at least once throughout history. For the purpose of identifying particular places and mountain ascents, an interactive map was inserted in the above stated website. These data are linked to the MySQL database, which keeps all collected attributes, and in addition to the identification of the place, it is also possible to get a detailed presentation of stages ending in the relevant stage place.

Important historical moments of the Tour de France (1903–2016)

The database and the visualization methods described in the preceding part of the paper were designed and used in such a manner that it is possible to identify the basic historical background of the history of Tour de France from its origination to the present when using them. Therefore, in this chapter we would like to stress certain interesting facts related to the development of the quantity of stage centres. They represent the above stated spatial dimension of the race whereas the inclusion of particular places in the itinerary of particular years significantly affects the overall character of the race. In the more than 100-year history of Tour de France we can identify a total of 682 places, which gained the status of stage places (either as a start or as a finish). 569 places are French, and the remaining places are stage places in the surrounding countries (most of them in Belgium – 39). In the first Tour de France editions, stages repeated. The route was almost identical, and new centres were included in the itinerary relatively sporadically. This fact also strongly affected the final classification of stage centres based on the frequency at which the relevant centre was included in the Tour de France itinerary as a stage place (Fig. 6)
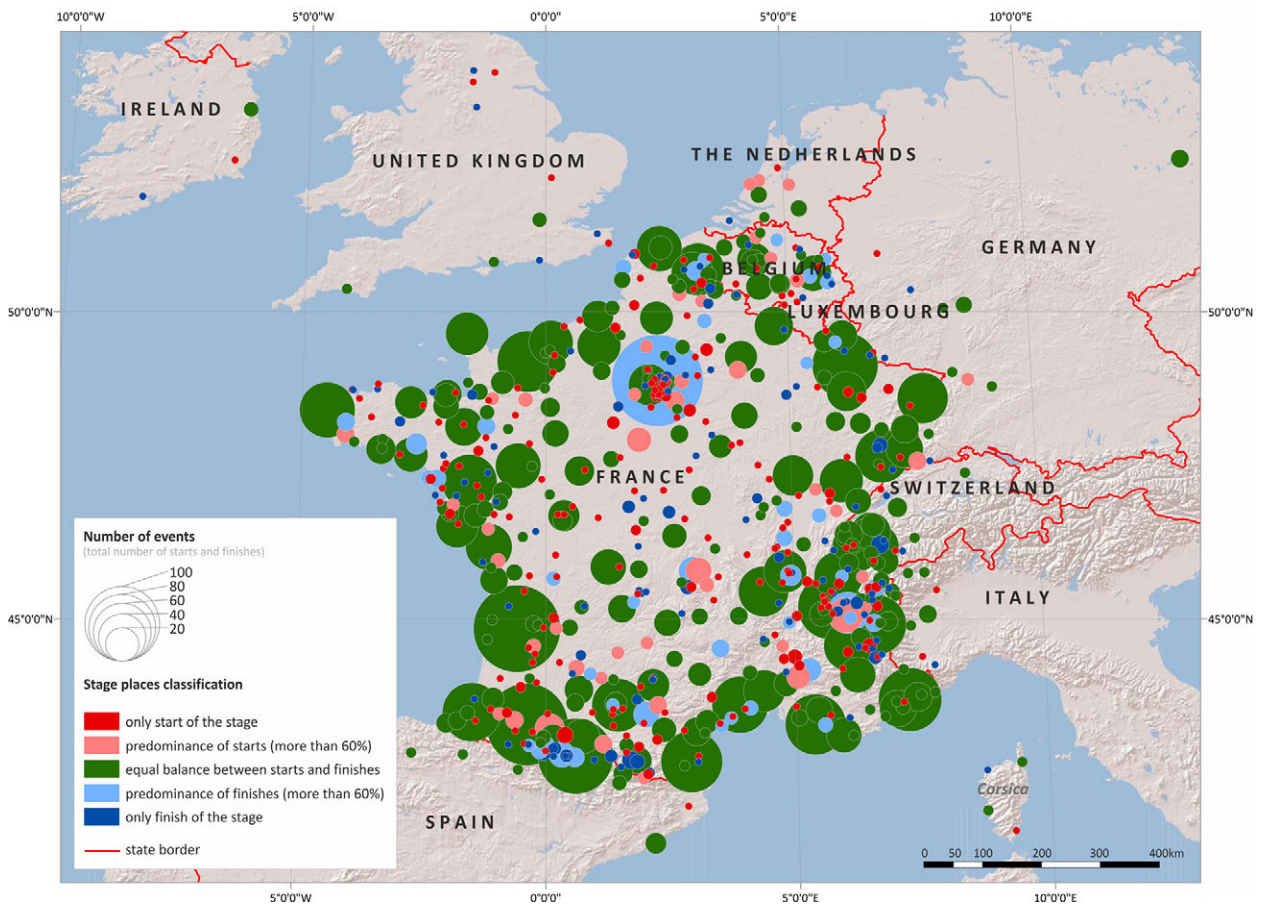
Fig. 6. Tour de France stage places' overall classification (1903–2016).
All places can by identified at www.tdfrance.eu/json_places.php.

With regard to the significance we can see that many of the places included in the first editions remain continuously at the top of the rankings in terms of the quantity of performed stages (e.g., Paris, Bordeaux, Metz, Grenoble, and Lyon). However, in the recent decades, there has been a notable trend towards increasing the share of new stage places to the detriment of stable centres. This is also connected, among other things, with the above stated use of the Tour de France for the promotion of particular regions in France, which thanks to television rights can be watched by millions of viewers throughout the world and which has, naturally, a great importance in promoting tourism in the surrounding region. The stability of stage centres included in the Tour de France itinerary in selected time periods is illustrated by three chord diagrams showing the most stable stage finishes (more than 20 times in the selected time interval) whereas each diagram shows the most successful countries in terms of the number of stage triumphs (Fig. 7).

The stability of stage finishes in the period before World War II (11 places with more than 20 finishes, a total of 80 stage finishes) can be seen on the diagrams. After World War II, we see an attempt of race organisers to expand the quantity of stages finishes, whereas only 3 towns (1947–1980: Bordeaux, Paris, Pau, 1981–2016: Paris, Alpe d´Huez, Pau) can be classified as stable finishes with more than 20 finishes in both following periods. However, the total quantity of stage finishes increased to 232 (1947–1980) or 338 (1981–2016), respectively. Also this graphical presentation is evidence that the organisers attempted to include the maximum possible number of stage finishes in the Tour de France itinerary. This fact is also visible in the localisation of particular finishes below the relevant diagrams on the same Figure 7. On the relevant diagrams we can also see the development of the sport dimension of the race. The clear dominance of French and Belgian cyclists in the first period is markedly balanced by the triumphs of cyclists from other countries.

1903–1939   1947–1980   1981–2016

Colours of states in chord diagrams:

■ France ■ Belgium ■ Italy ■ Luxembourgh ■ Switzerland ■ Nederlands ■ Spain ■ Other states
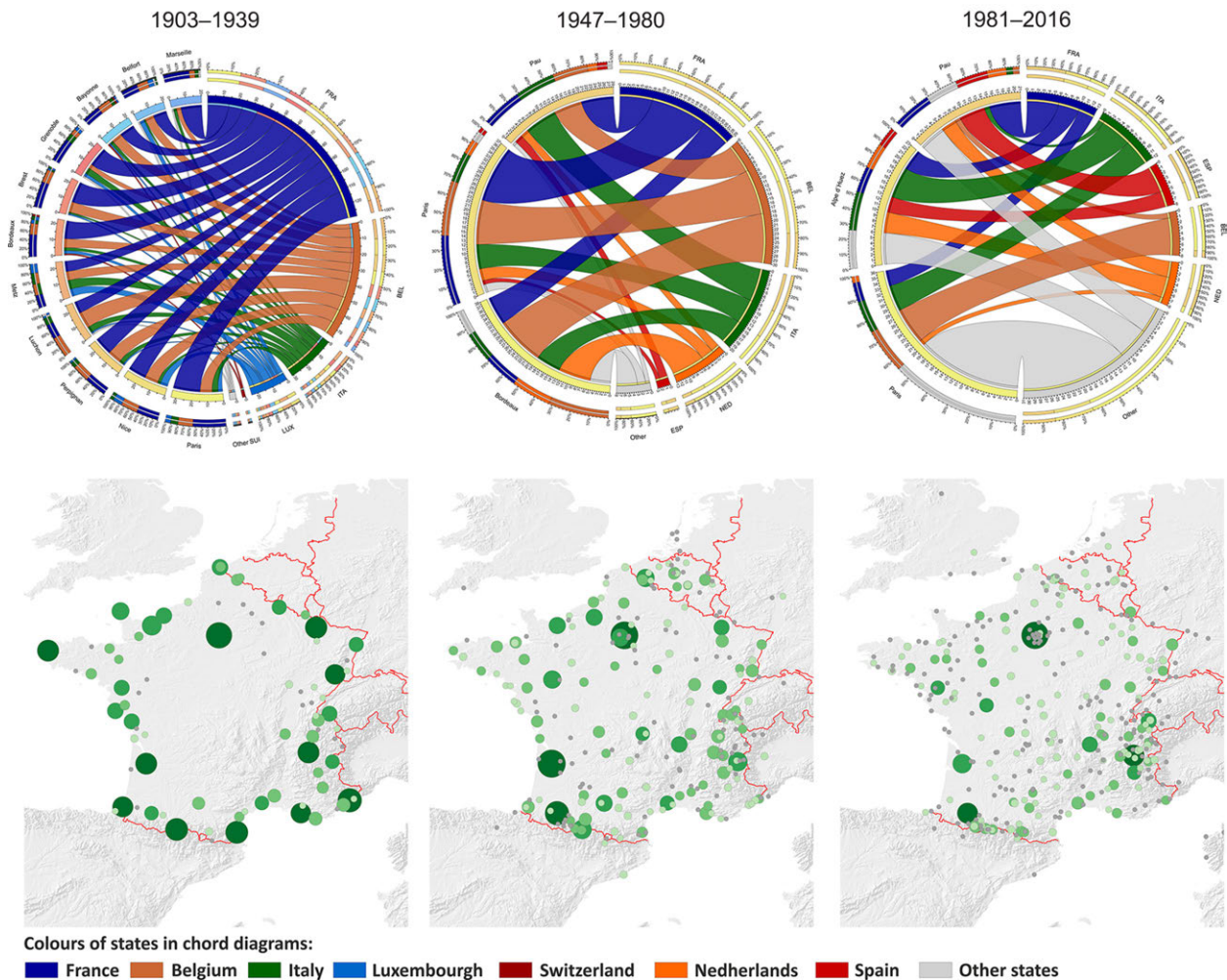
Fig. 7. Stable finish places of the Tour de France with more than 20 finishes within the given time span (below the diagrams, the maps show all stage finishes in the relevant period).
Stable finish places of the Tour de France with more than 20 finishes:*
1903–1939: Paris, Nice, Perpignan, Luchon, Metz, Bordeaux, Brest, Grenoble, Bayonne, Belfort, Marseille,
1947–1980: Bordeaux, Paris, Pau,
1981–2015: Paris, Alpe d´Huez, Pau.
Interactive map for each period could be found at following links:
1903–1939: http://www.tdfrance.eu/json_finish2.php,
1947–1980: http://www.tdfrance.eu/json_finish3.php,
1981–2016: http://www.tdfrance.eu/json_finish4.php.
*Below the diagrams, the maps show all final points in the relevant period (the ones that were included as the most stable are in dark green colour). Here we can see a change in stage centres from stable (1903–1939) up to the attempts to include more and more quantities of new stage finishes (1981–2016).

Based on the database with results, we can see that before 1939, cyclists from 9 countries were recorded on the scoreboard, while after 1981 it was even 32 countries[6].

The linking of spatial and sport dimensions is shown on the final map (Fig. 5). On this map we have only illustrated stage finishes because we covered the achievements of particular

countries or of individual racers in the Tour de France stages. In the historical statistics of the Tour de France, there is a record of a total of 831 cyclists as the winners of stages (including time trials, prologues and including winning teams in team time trials). On the main map we reflected the places of the triumphs of the most successful cyclists in history in terms of the number of stage triumphs reached. The five most successful cyclists in history won a total of 139 stages, the finishes of which were in 94 different places (Fig. 8).

---

6   These statistics can be obtained directly from the website http://www.tdfrance.eu/custom.php.
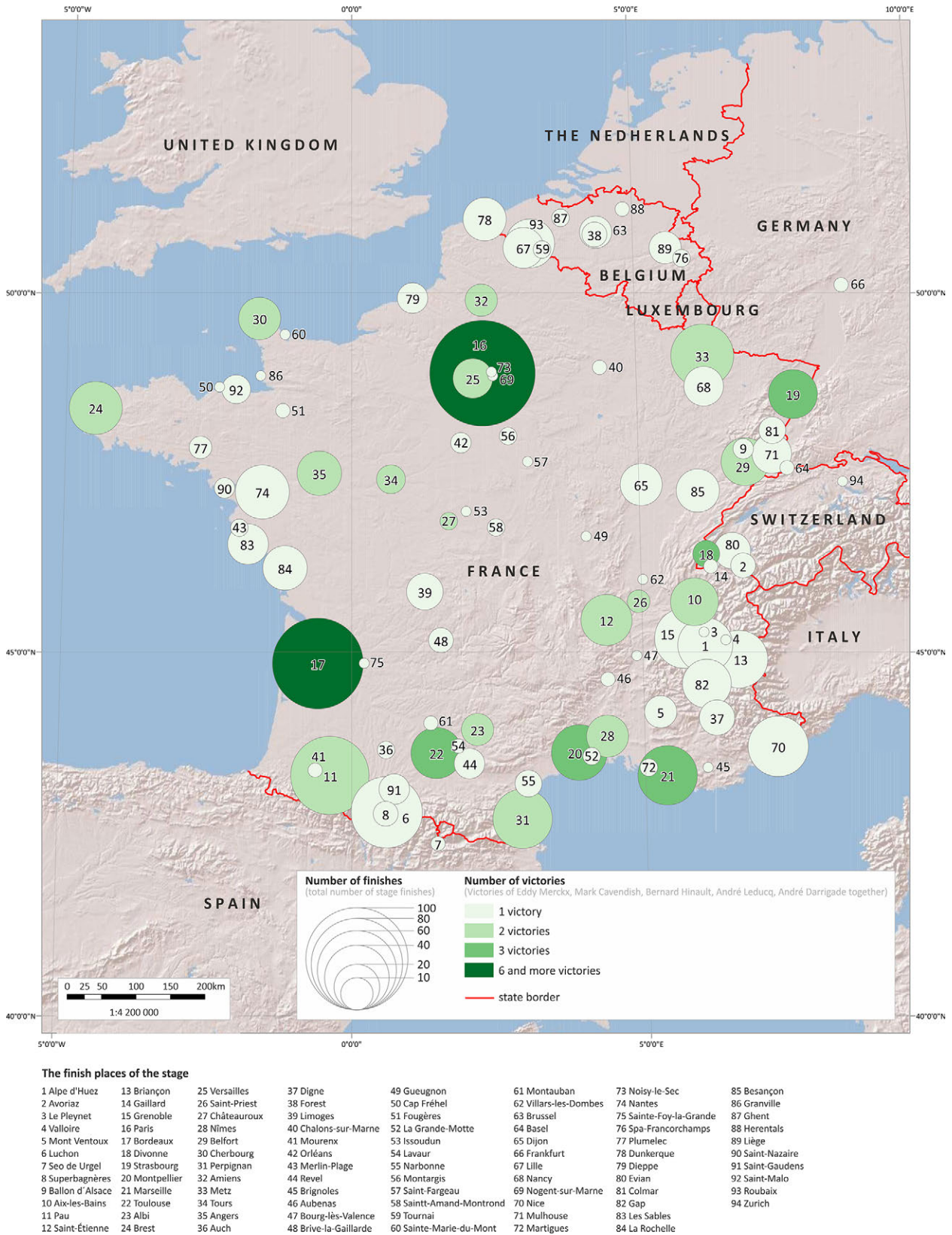
Fig. 8. Places of stage triumphs of the five most successful cyclists in the history of Tour de France according to the number of stage victories.
(All other partial maps are available at: http://www.tdfrance.eu/download/tdf_map_bk.pdf).

The map above (Fig. 8) expresses the stage triumphs of individual cyclists using standard cartographic methods. The size of the diagrams represents the total number of stage finishes in observed places (in 103 editions of the Tour de France cycling race altogether). Colour filling of the diagrams shows the number of triumphs of observed cyclists in these finishes. To express the internal structure of cyclist´s victories in the finish places, we used the above-described chord diagram, where the numbers of places on the map are identical to the numbers of places in the appropriate chord diagrams.

When evaluating the stage sections of particular countries, both the map and the diagram shows the victories of the five historically most successful countries (France, Belgium, Italy, the Netherlands, Spain), which reported more than 100 winning stages. The number of winning stages of these countries naturally reflects the long-term participation of cyclists from these countries in particular Tour de France editions. We must mention that in recent years, the scoreboard has been enriched by several countries in which cycling does not have a very rich history. However, globalisation also affects this sport, and therefore the movement of cyclists from various countries between professional teams with high-quality services is not exceptional. The map shows all stage finishes in the long Tour de France history. With a the coloured presentation of the number of triumphs of the countries under review, as well as with a diagram, we wanted to stress the historically most significant stage finishes, which are included in the itinerary of particular editions on a relatively regular basis. Thus we stressed 23 places, where more than 20 stage finishes were reported. In all of these stable centres, there is an apparently clear dominance of cyclists from the countries under review, whereas when looking at them in more detail, we can highlight the dominance of French and Belgian cyclists. Together, they triumphed in 1166 Tour de France stages, which is more than a 50% share of all stages performed (2215). Characteristics similar to stage finishes are also shown for the triumphs of cyclists or of particular countries in mountain ascents. It is important to stress that it is much more difficult to collect correct data for mountain ascents than for stage places. Therefore, in our analysis we used the official statistics of Tour de France,

which record cyclists' triumphs in mountain ascents of the second, first and "hors catégorie". In this case as well we wished to highlight, in addition to a spatial aspect, the success rate of particular cyclists or of the most successful countries. For cyclists, we considered the limit to be 25 winning crossings (6 cyclists), and for the triumphs of particular countries, we considered the limit to be 200 winning crossings. (Here between Italy (the fourth) and the Netherlands (the fifth and historically most successful country), is a significant drop in the number of wins). Our analysis covers a total of 287 mountain ascents. In the presentation of the number of wins using chord diagrams, the coverage of the ascent by geomorphological unit was preferred to the total frequency of the relevant ascent included in the stage itinerary. This fact also reveals the frequency at which particular Tour de France ascents are included in the itinerary. Here we can see a frequent inclusion of stable Pyrenean ascents in comparison with Alpine hills. With respect to these particular Pyrenean or Alpine ascents which appeared first in almost the same period (1910 or 1911), the difference is mainly in the shape of the relevant geomorphological unit. The Pyrenees are line-shaped, whereas the Alpine massif is much more extensive, and thus the general variability of possible ascents is higher. The six most successful cyclists in history who are considered in our contribution triumphed a total of 205 times in 78 different ascents. A joint triumph was recorded for all cyclists in two ascents (Tourmalet and Aspin), which in the rich history of Tour de France belong to the most frequently covered ascents (Tourmalet is on the 1st place, Aspin the 3rd place). Among the wins of particular countries the situation is similar to that of stage triumphs. Historical statistics are considerably affected by the long-term participation of cyclists from traditional road cycling countries (France, Spain, Belgium, and Italy). There is also an evident marked dominance of cyclists from the countries under review in all of the observed mountain crossings (25 ascents, whereas the boundary value was 20 crossings for the analysis). An exception is the Alpe d´Huez ascent, which is, thanks to a considerable dominance of Dutch cyclists in the 1970s and in the 1980s, nicknamed "the Dutch Mountain".

The bicycle stage race Tour de France was created in 1903. The 103 editions of Tour de France

organized up till now have brought many un-repeatable sport performances which have permanently gone down in the history of sports. In addition to the sport aspect, it is also unique thanks to its spatial coverage. In addition to the first editions when particular stages repeated, we can say that each edition is unique with its spatial arrangement. The organizers attempt to gradually present new places, which with their attractiveness and visual representation during broadcasting transmission, have the potential to be the most important places for the tourism in the given region. The inclusion of 666 stage places in the Tour de France itinerary is really a unique phenomenon in the field of sporting events. This trend of extending the list of stages can be expected even in the future as this aim is also declared by the representatives of this event at the A.S.O. (Amaury Sport Organisation). In the context of stage triumphs we can also expect that the list of countries that reported a winning stage (until now 33) will grow further. Cyclists from less traditional countries are increasingly found at the start of the race. A typical example can be the participation of the first African team in Tour de France 2015 (MTN-Qhubeka).

Discusion

The acquisition and subsequent processing of necessary data was an essential precondition for the whole research devoted to historical characteristics of the Tour de France. The objective was not a detailed description of particular years or an analysis of significant sports moments, but an overview of the history of the race using selected visualization techniques. Therefore, in this place it is adequate to briefly discuss the suitability of the methods used in similar analyses.

The database system, designed to ensure replies to any question related to space and the history of the relevant race, was the imagined core of the research. Today it is not necessary to discuss the significance of database systems in the online environment. Database systems are used in all tasks, in which data of any nature constitute an essential component in the implementation of particular activities. In our case, the said database system (see the chapter Methods – data processing) presented the input for all other performed

graphical and cartographical visualizations, either in printed or online form. The advantage of online presentation (the website www.tdfrance.eu) is the possibility to obtain detailed information on the required characteristics, based simply on a query created by the client, and subsequent processing through the database system on the part of the server. The printed form has, naturally, certain limitations but the tools of graphical and cartographical visualization allow relatively adequate presentation of the historical facts obtained.

In the cartographical presentation, standard methods of proportional symbols were used, with the relevant colour based on selected intervals. Interpretation options of this method are relatively intuitive, without any need to know the methodological details entering the development of such a created map. It would certainly be possible to use several relevant methods for the expression of the characteristics under review (e.g., a pie chart, stacked chart, etc.). The method preferred by authors of this contribution in the form of proportional symbols with the relevant range of colours was also affected by the fact that all visualizations have been performed by us in the online Internet environment. For the purpose of dynamic visualization, we applied the D3.js library and the TopoJSON format (see the chapter Methods – D3 and TopoJSON). The possibility of obtaining the visualization of almost identical printed and online forms through this library was a key reason for its application. There are many interesting tools that could be used in a similar task; we mention, for instance, map javascript libraries jQueryMapael, amMaps, jVector Map, and others. Again, their use is affected significantly by the personal preferences of the authors of the application. One the other hand, we must say that the concept of the D3 library is described precisely in several scientific works focused on the visualization of extensive data sets (Zhu, 2013, Bao and Chen, 2014, Jain, 2014). With respect to the extensiveness of the said library, the use of other cartographical methods is limited only by the authors' skills.

Cartographic interpretation of the results is accompanied by a graphic presentation in the form of chord diagrams. Our aim was to link the spatial dimension (the position of the place expressed as a position on the map) and the sports dimension
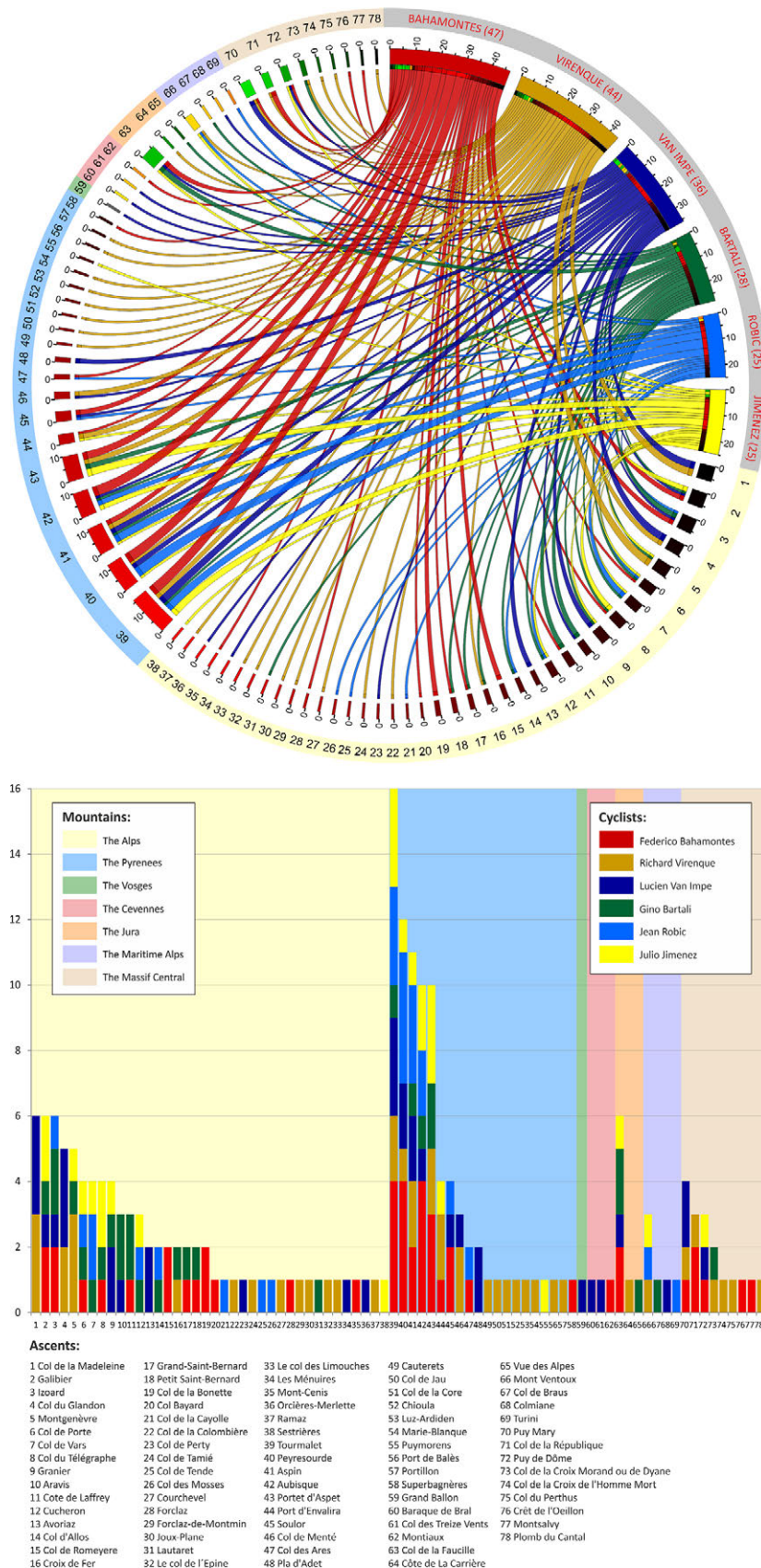
Fig. 9. Comparison of circular visualization and stacked chart on an example of the most successful cyclists according to the number of winnings ascents crossings.

(achievements of a specific cyclist or of a nationality in the stage or in a mountain ascent). Again, we must note the fact that, for the purposes of graphic presentation, some standard methods (a column chart, a stacked chart) could also have been used. An example of comparison between circular visualization and stacked chart is shown at Figure 9.

One of the key advantages of the concept of circular visualization is, according to Auber et al. (2005), its own neutrality. It means that all tops (entities – cyclists, towns) are at an identical distance one from another and from the centre of the presentation. None of these entities has a privileged status as the reader tends to perceive objects that are closer to the centre as the most important (Huang et al. 2007). Equally, the advantage of circular visualization is the possible presentation of a large amount of data. The use of standard graphical presentation (e.g., a column chart) would undermine the readability of the output. In the opinion of the author of the project CIRCOS, M. Krzywinsky, the circular form of a diagram is also better for perception by the human eye, for which circular movement is more natural than movement along breaks or sharp edges. The usability of circular data visualization is versatile and there are many studies that use this form. This method is primarily used in the area of genetic engineering (Saw et al. 2013, Martis et al. 2013). The historical context of the origin of Europeans is presented in an interesting way in the paper of Bollongino et al. (2013). In the field of social sciences, this concept is elaborated with the visualization of population migration flows in the world (Sander et al. 2014). Such a form of visualization can be used for the graphic analyses of any phenomena in which we wish to illustrate mutual links between particular entities. A good example of such use of chord diagrams can be used by monitoring of mutual balance of matches among the selected teams, or individuals in any sport sector[7]. Here we can see a depiction of tennis matches among all world leaders since 1973. From such kind of visualization, we see mutual balance and number of matches between players, the player who was the most successful from this point of view and at the same time. The methods

of graphic and cartographic visualization presented in the study for the purpose of analysis of the Tour de France history can find their place to a large extent in any research task which includes visual data interpretation. There are many methods with which it is possible to present the characteristics under review and data. Usually the method is markedly affected by the preferences of the authors. Nevertheless, it must be underlined that it is only one of possible methods of such interpretation. Our concept of the visual processing of the history of Tour de France is based on the knowledge base of the authors of the paper, as well as on the enthusiasm for this interesting topic.

Conclusions

The bicycle race Tour de France is a global sporting event. As one of the most watched sporting events on the planet, it has a history going back more than 110 years. From its origin in 1903 to the present, the Tour de France has experienced many major and minor changes (organisational, sports, logistic, etc.). The original plan to organise this event to increase the sales of the French newspaper "L'Auto" has turned into a global event which is watched today by millions of spectators directly along the route and by hundreds of millions of viewers on TV screens. It is no surprise, then, to see that it enjoys unprecedented popularity among the audience, sports fans and the scientific community. There are many scientific contributions dealing with various aspects of this event (sporting, cultural, social, physiological, historical, etc.) Cycling fans can also enjoy a multitude of available literature, which usually presents historical statistics, significant personalities or iconic achievements of cyclists in legendary stages. As regards geography as spatial science, however, the space in which the event is organised is the most interesting aspect. Every year the event consists of many stages, which have their start, finish and routes, such as significant mountain ascents, that the cyclists must overcome. It is this variability of significant stages that was the primary impulse for the development of the submitted study. With respect to the total number of years organised up till now (103), we can work with data in a wide

---

7   This example is available at: http://sodbtn.sk/chord/chord22/index.html.

time frame. The extensive quantity of data collected and the spatial aspect resulted in the basic research question. We were interested in how it is possible to capture the most important historical aspects of Tour de France correctly and transparently and distribute them to persons interested in this topic. For the purpose of visualization, we used methods allowing graphic and cartographic interpretation of the results in two basic forms - printed and online (the Internet). Both of these forms constitute a sort of specific output of this paper. The use of the database system (MySQL), designed so that it can effectively store the maximum possible quantity of data, constitutes the backbone of the entire research. The data must be organised in such a manner that it is possible to answer any questions related to the history of Tour de France. Then from such a database, we extracted data required for specific visualization of selected attributes, using SQL.

In the context of graphical visualization, we preferred a less traditional, but in our opinion interesting, method of circular visualization in the form of chord diagrams. The advantage of this concept is the possibility of displaying of a large quantity of data, as well as a certain neutrality resulting from the uniform arrangement of entities under review in a circular form. We used the visualization to express the internal structure of stage triumphs and of triumphs in mountain ascents of the most significant cyclists and nationalities in the history of Tour de France. The spatial presentation was illustrated using the standard method of a proportional symbol. The result is a comprehensive map composition designed for the A0 format, which depicts all significant historical moments of the Tour de France. In the context of graphic and cartographic interpretation, other methods used for similar analyses could also have been used, but our objective was the above stated visualization of results in the online environment in the form of a functional website developed by the authors of this paper (www.tdfrance.eu). The methods used by us are easily applied with the use of JavaScript library D3. At present, it is a standard for data visualization in the Internet environment, while for cartographic interpretation, this library is used in combination with TopoJSON format. The said methods have allowed us to capture the rich history of Tour de France both at sports and spatial levels. Such

linking can be inspiring for other analyses in the area of sport, and for other analyses in which the time and spatial aspects play an important role.

The bicycle event Tour de France is inherently connected to the performance of cyclists participating every year in this prestigious event. This part of history is accentuated in many professional and lay contributions and books dealing with this significant sporting event. Our objective was to explain these aspects in the context of the space where the event is organised. This space forms an integral part in the long-term development of the exceptional position of this race in the sporting world. Today many stage centres and mountain ascents are very sought-after by tourists. Enthusiasts from all over the world often follow the routes of the legends of this sport in ascents such as Tourmalet, Mont Ventoux, Alpe d´Huez, and many others at a later date and try to compete with the top cyclists' performances. Therefore, the Tour de France has long ceased to be a mere sporting event. Thanks to its nature and position, it significantly exceeds this primary role. Thanks to broadcasting rights, broadcasts from particular stages are transmitted to many countries in the world and many images feature cultural, historical and natural monuments around which cyclists pass by. Thus the race has become a significant tool for promoting tourism in the relevant regions. It will also be certainly interesting to monitor in the future how the Tour de France influences the visit rate of particular regions, the number of overnight stays in hotels, etc.

From its origin to the present, the Tour de France has experienced many changes. With respect to the present form of this race, we can expect other significant moments in the future, which will go down in the history of this event not only in the context of sporting performances, but also in the context of race organisation, the inclusion of other interesting places in the stage itinerary, etc. The Tour de France is a global phenomenon which is interesting from the geographical point of view thanks to the location where it takes place. The space will continue to expand, as evidenced through yearly debates on the possibility of the localisation of the "Grand Départ" for this event. More and more frequently, foreign towns host this event (3 out of the last 5 years; once on the island of Corsica), which contributes significantly to strengthening its position outside

France. In the context of constant development of this event in all aspects (the technological aspect of the development of cycling as such, broadcasting technology, logistical possibilities, etc.) we can say that "La Grande Boucle" will never grow old, but it deserves the attention of geography, which can bring a certain specific view of the subject and of its rich history.

Acknowledgements

References

Allchin R., Bell A., 2003. *Golden stages of the Tour de France: tales from the legendary stages of the world's greatest bike race*. Mousehold Press, London.

Balduck A. L., Maes M., Buelens M., 2011. The Social Impact of the Tour de France: Comparisons of Residents Pre- and Post-event Perceptions. *European Sport Management Quarterly* 11(2): 91–113.

Bale J., 1994. *Landscapes of Modern Sport*. Leicester University Press, Leicester.

Bale J., 2003. *Sports Geography*, 2nd edition. Taylor & Francis, London.

Bao F., Chen J., 2014. Visual framework for big data in d3.js. *Electronics, Computer and Applications*: 47–50.

Bollongino R., Nehlich O., Richards M.P., Orschiedt J., Thomas M.G., Sell C., Fajkošová Z., Powell A., Burger J., 2013. 2000 Years of Parallel Societies in Stone Age Central Europe. *Science* 342: 479–481.

Bostock M., Ogievetsky V., Heer J., 2011. D³ data-driven documents. *IEEE transactions on visualization and computer graphics* 17(12): 2301–2309.

Boury P., 1997. *La France du Tour. Le Tour de France, un espace sportif à géographie variable*. Harmattan, Paris.

Campos C., 2003. Beating the bounds: The Tour de France and national identity. *The International Journal of the History of Sport* 20(2): 149–174.

Darzentas N., 2010. Circoletto: visualizing sequence similarity with Circos. *Bioinformatics* 26(20): 2620–2621.

Dauncey H., Hare G., 2003. *The Tour De France 1903–2003: A Century of Sporting Structures, Meanings and Values*. Routledge, London.

Desbordes M., 2007. A review of the economic impact studies done on the Tour de France: methodological aspects and first results. *International Journal of Sport Management and Marketing* 2(5–6): 526–540.

Fife G., 2008. *Tour De France: The History, the Legend, the Riders*. Mainstream Publishing, Edinburgh.

Fumey G., 2006. Le Tour de France ou le vélo géographique. *Annales de Géographie* 115(4): 388–408.

Huang W., Hong S. H., Eades P., 2007. Effects of sociogram drawing conventions and edge crossings in social network visualization. *Journal of Graph Algorithms and Applications* 11(2): 397–429.

Jain A., 2014. Data visualization with the D3.JS Javascript library. *Journal of Computing Sciences in Colleges* 30(2): 139–141.

Krzywinski M., 2009. Circos: an information aesthetic for comparative genomics. G*enome research* 19(9): 1639–1645.

Ledford H., 2010. Big science: the cancer genome challenge. *Nature News* 464(7291): 972–974.

Lucia A., Earnest C., Arribas C., 2003. The Tour de France: a physiological review. *Scandinavian Journal of Medicine & Science in Sports* 13(5): 275–283.

Martis M. M., Zhou R., Haseneyer G., Schmutzer T., Vrána J., Kubaláková M., König S., Kugler K. G., Scholz U., Hackauf B., Korzun V., 2013. Reticulate evolution of the rye genome. *The Plant Cell* 25(10): 3685–3698.

McGann B., McGann C., 2006. *The Story of the Tour de France Volume 1: 1903–1964: How a Newspaper Promotion Became the Greatest Sporting Event in the World*. Dog Ear Publishing, Indianapolis.

McGann B., McGann C., 2008. *The Story of the Tour de France, Volume 2: 1965–2007: How a Newspaper Promotion Became the Greatest Sporting Event in the World*. Dog Ear Publishing, Indianapolis.

Mignon P., 2003. The Tour de France and the doping issue. *The International Journal of the History of Sport* 20(2): 227–245.

Mignot J. F., 2015. *Histoire du Tour de France*. La Découverte, Paris.

Reed E., 2003. The economics of the tour, 1930–2003. *The International Journal of the History of Sport* 20(2): 103–127.

Reed E., 2007. The Tour de France in the Provinces: Mass Culture and Provincial Communities' Relations with the Broader World. *French Historical Studies* 30(4): 651–684.

Russell P. J., 2010. *iGenetics: A Molecular Approach*. Benjamin Cummings, San Francisco.

Sander N., Abel G. J., Bauer R., Schmidt J., 2014. *Visualising migration flow data with circular plots*. Vienna Institute of Demography Working Papers.

Saw J. H., Schatz M., Brown M. V., Kunkel D. D., Foster J. S., Shick H., Christensen S., Hou S., Wan X., Donachie S.P., 2013. Cultivation and complete genome sequencing of Gloeobacter kilaueensis sp. nov., from a lava cave in Kīlauea Caldera, Hawai'i. *PloS one* 8(10): e76376.

Schneider A. J., 2006. Cultural nuances: doping, cycling and the tour de France. *Sport in Society* 9(2): 212–226.

Sidwells C., 2009. *Tour Climbs: The complete guide to every mountain stage on the Tour de France*. Collins, New York.

Wheatcroft G., 2005. *Le Tour: A History of the Tour De France*. Simon & Schuster UK, London.

Whittle J., 2003. *Le Tour: A Century of the Tour de France*. MBI, London.

Yates R., 2006. *Ascent: The Mountains of the Tour De France*. Van Der Plas Publications, San Francisco.

Zhu N.Q., 2013. *Data Visualization with D3. js Cookbook*. Packt Publishing Ltd, Birmingham.