

TourVis: Narrative Visualization of Multi-Stage Bicycle Races

J. Díaz¹ , M. Fort²  and P. Vázquez³ 

¹Digital Care Research Group, Universitat de Vic - Universitat Central de Catalunya, Vic, Spain

²Graphics and Imaging Lab, Universitat de Girona, Girona, Spain

³ViRVIG Group, Universitat Politècnica de Catalunya, Barcelona, Spain

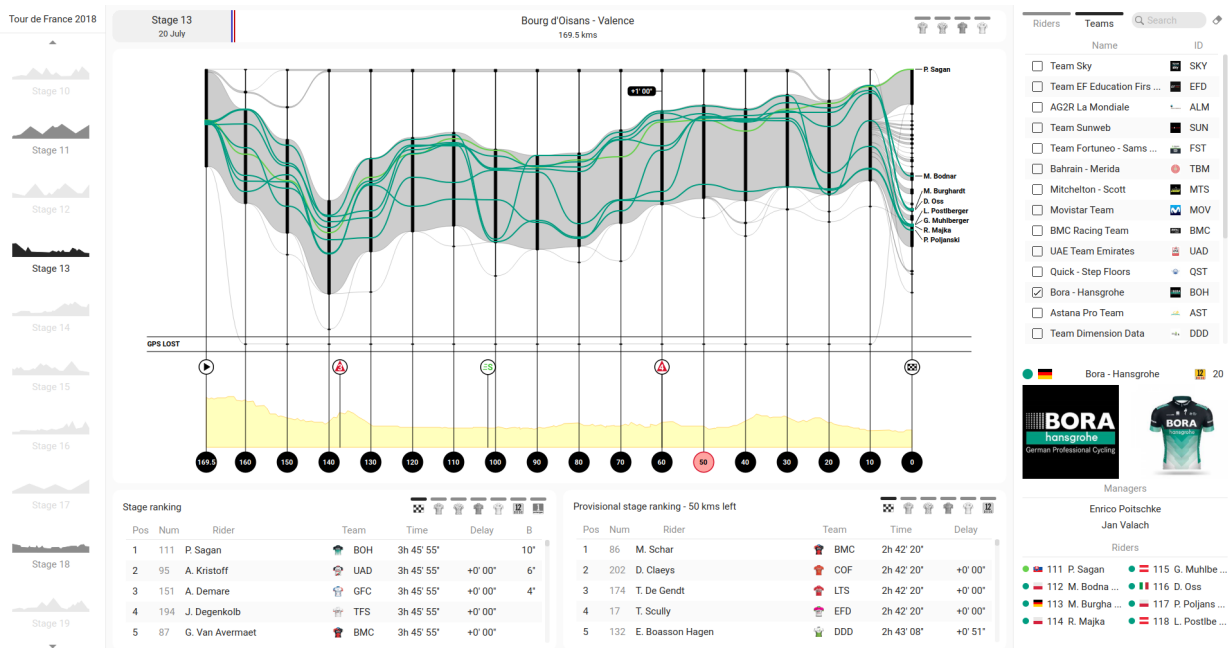


Figure 1: Visual analysis of the Tour de France 2018. Our multiple-view application allows for the easy exploration of the stages of a race. Within a stage, a large set of relevant events that happened through its development can be seen: leader changes, attacks and breakaways, intermediate sprints, categorized mountain passes, and rankings of different categories among others. Moreover, the performance analysis of each contender, a group of riders, or a team, can also be performed.

Abstract

There are many multiple-stage racing competitions in various sports such as swimming, running, or cycling. The wide availability of affordable tracking devices facilitates monitoring the position along with the race of all participants, even for non-professional contests. Getting real-time information of contenders is useful but also unleashes the possibility of creating more complex visualization systems that ease the understanding of the behavior of all participants during a simple stage or throughout the whole competition. In this paper we focus on bicycle races, which are highly popular, especially in Europe, being the Tour de France its greatest exponent. Current visualizations from TV broadcasting or real-time tracking websites are useful to understand the current stage status, up to a certain extent. Unfortunately, still no current system exists that visualizes a whole multi-stage contest in such a way that users can interactively explore the relevant events of a single stage (e.g. breakaways, groups, virtual leadership...), as well as the full competition. In this paper, we present an interactive system that is useful both for aficionados and professionals to visually analyze the development of multi-stage cycling competitions.

CCS Concepts

• **Human-centered computing** → Visualization systems and tools; Visual analytics;

1. Introduction

Multi-stage participatory races are competitions that last for several days where the contenders must cover a distance every day. They are quite popular in several sport disciplines such as motorsport rallies (e.g. the *Dakar*), running (e.g. *Jungle Marathon* in Brasil), or cycling (e.g. the *Tour de France*, the *Giro d'Italia* or *La Vuelta*). The increased use of GPS location trackers allows the organizers to follow the development but also provides a trove of information that can be useful to better understand the performance of different participants. In our case, we concentrate on cycling races, which are quite popular, especially in Europe, and gather great attention from the media. During event streaming, it is common to have access to several sorts of infographics that provide the situation of every single stage. However, there is no tool for the exploration of whole stage data, that lets users analyze the behavior of riders or teams along with it. The questions regarding the stage are multiple and are related both to the proper stage development, but also to the overall situation in the whole competition since the timings of the participants accumulate at the end of each stage. In this paper, we present a new tool that facilitates the exploratory analysis of single and multi-stage participatory cycling events where GPS positions have been sampled along with the race.

Line charts can be used to depict the evolution of participants in the different rankings across multiple-stage competitions. However, these charts just convey the information related to the rankings of the stages at the finish line. If we are interested in the events that occurred *during* a single stage, one must refer to a newspaper or dedicated website chronicles that provide the highlights of each stage. The goal of our application is to provide a new visualization tool that allows the user to see at a glance what happened in a single stage, not just who are the winners of the different rankings or the positions of the riders in every one of them, but also the full story of the stage: attacks and breakaways, change of leaders, the evolution of the different groups related to the stage profile, etc. Thus, we have made two contributions, that address two major challenges:

- A novel visual design for conveying the evolution of riders in a stage (Sect. 3.3).
- A multi-view system that uses the stage view and a set of interrogation techniques to facilitate the exploratory analysis of a full race (Sect. 3.4).

The resulting application can be used by the general public, who found it very intuitive and useful, to get an understanding of the most relevant facts that happened along a race. But also practitioners at the professional level. In particular, two experts (a professional cycling director assistant and race mechanic, and a team director) praised it as a very powerful tool. They thought it could greatly help directors to analyze stages. It could also help in real-time decisions, better than with their current tools if data could be acquired on the fly (NTT gathers them for the Tour organization, but data are private).

The rest of the paper is organized as follows: Section 2 covers the previous work in this area, while Section 3 describes our system. Data extraction and processing is explained in Section 4. The details of the different features and the implemented interactions are described in Section 5. The evaluation of our visualization tool

is presented in Section 6. Finally, we discuss advantages and shortcomings and conclude with some ideas for the future in Section 7.

2. Related Work

The gathering of data for analysis of athlete's performance is quite common nowadays in sports. Some visualization systems have been created for this purpose. Unfortunately, most of those systems analyze the behavior of participants in playgrounds [GH17, PVS*18], such as soccer, tennis, basketball, etc. When dealing with races, most visualization techniques devote most of the efforts to adequately present the final results [CF13, GQR16].

Perin et al. [PVS*18] classify the data used in sports visual analytics in three categories: box-score data, tracking data, and metadata. The first type includes goals and all sorts of scores, as well as ranking data. The second one refers to motion and position information obtained through sensors or video technologies, whilst the third category is related to information that includes physical characteristics of players or dimensions of grounds among others. Some of them have a tight relation with sport development, but some others are more loosely related. In cycling races, all of these ingredients are of high relevance: *i*) position of the rider, which is tracked continuously in real-time (except GPS malfunctioning or accuracy issues), *ii*) current ranking in the whole race, *iii*) information of the participants and the teams, and, finally, *iv*) metadata such as the geographic profile of the stage (slope is a crucial element here). Thus, we need to include all of them in our tool.

Unfortunately, bicycle races (or other multi-stage events) have not received a large amount of attention from the visualization community, with the notable exception of Wood [Woo15], which is reviewed below. Some techniques have been developed to analyze some variables sets that can be acquired in real-time, such as pedaling forces [KYY*16] or the aerodynamic flows around a cyclist [MLT*20]. There is also a host of systems that visualize some other aspects of team sports [AAB*17, CEGH17, FMB*15, LOC*16, LTB16, PSBS12, SSS*14, SHJ*15]. However, most of the techniques can only be applied to closed and relatively small playgrounds since they rely on analyzing repeated actions of single players or the whole team structure, which is not the case of bicycle races. Some tools have been created to improve the analysis by semi- or fully automatic data extraction and annotation of video footage [BLC*14, DKVS14, SSN*16, SJB*16, SJL*17, SJS*17]. But these are of little utility for cycling since placing vision systems or sensors along a 100+ kilometer road would be unfeasible.

2.1. Narrative visualization

We use the word *narrative* to refer to visualization systems that emphasize the evolution along a match, stage, or season, not only final results. Chen et al. [CLX*16] summarize the basketball season using a time series component and a score summary. This is very useful to determine the wins and losses of a team, as well as the changes over time, but it does not provide visual cues to the concrete classification in each moment. Chung et al. [CPG*15] assist the video data organization procedure that ranks key instances of video footage for later analysis. Demaj also creates a set of vi-

visualizations for the analysis of the performance in tennis matches [Dem13], but the data is not presented in a time-dependent way.

One of the key components in our visualization tool is the one intended to visually communicate the positions of riders along a stage. Media coverage usually provides instant classification infographics, as well as other data, such as the position of leaders. However, at the end of a stage, only the ranking of the different classifications (stage, individual, climbers, points, etc.) are provided, with no information on the stage development. To get more insights into the evolution of a concrete race (e.g. a marathon), some tools provide abstract visualizations that reproduce the participants' movements through animation. However, these are typically not interactive (only speed can be changed), and in many cases the individual riders' names are unavailable. Moreover, users need to watch the whole animation to make sense of what happened. Basdere et al. [BCC*19] created an application, named SAFE, tailored to get situational awareness for mass-participation endurance events. In this case, the position of the runners is important, but they focus on assisting organizers in making decisions such as enhancing marathon logistics and supply, or quick responses to unexpected events, to preserve public health. Bačák and Klobučník [BK19] have proposed the use of chord diagrams to visualize the results of a certain stage. However, to understand the evolution of the stage, we need to provide intermediate positions, not just the classification in the finish line. Also, other data, such as the delay relative to the leader, performance of teams, or the virtual position of riders on the race, or the formation of groups are also crucial.

The closest research, according to the authors' knowledge, is the inspiring work by Wood [Woo15]. He proposes a series of visualization techniques to illustrate the behavior of a participant in multi-stage events. However, the analysis is centered on individual riders and the amount of interaction is quite limited. In our case, our goal is a fully interactive tool that allows a broader range of queries.

2.2. Ranking visualization

For ranking visualization, several techniques have been proposed. Rank clocks, for example, display the ranking of cities under different parameters along time [Bat06]. Perin et al. presented the À Table system [PVF14], which has a novel ranking visualization that facilitates temporal exploration of soccer teams with advanced interactions to answer questions related to the number of goals or points. A follow-up work [PBV16] focuses on emphasizing the separation between soccer teams with gaps. Gratzl et al. [GLG*13] facilitates the exploration of multiple attributes, some of them as headers of the tables, and can visualize several occurrences up to some extent. Han et al. [HPC*19] aggregate temporal data to gain insights on the changes of rankings for basketball teams. Thus, they do not provide a narrative on a certain season. Wall et al. [WDC*17] also provide static ranks for data elements and let the user explore the attributes with some interactive tools that facilitate highlighting some features such as changes. Shi et al. [SCL*12] adapt the ThemeRiver visualization technique to summarize a large number of Bing queries and provide several glyphs to illustrate rank changes. Miranda et al. [MLKS17] focus on rapidly extracting top-k elements in massive spatiotemporal data. Gousie et al. create

some glyphs to facilitate the identification of trends and clusters in ranked time-series data [GGB14].

In some cases, they have been designed for several elements to rank relatively small, such as [Bat06, PVF14, GGB14, PBV16], or focus on analyzing trends [SCL*12, GGB14]. In our case, we need to deal with a number of participants over one hundred, and we need to perform more complex queries involving several contenders at a time. Moreover, one fundamental feature of cycling stages is the formation of groups (several cyclists can work together to increase the speed with less fatigue), which must be properly identified, and visually isolated from other cyclists. Besides, we are also interested in queries related to multiple elements (e.g. position of all the cyclists of a team) and idiosyncratic metadata (e.g. relative position of a contender with respect to the leader of the individual general classification, or the arrival order to a certain geographic position of a stage, such as a categorized climb). Finally, as mentioned, several elements of metadata may be key in the stage development, such as the geographic profile (e.g. slope, length of uphill tracks), which have to be dealt with along the other aspects of the stage (i.e. current positions of the participants).

3. Overview

Our main goal is the development of an exploratory analysis tool that facilitates understanding the evolution of single or multi-stage sporting events, including the actual information of individual stages and not just their final rankings. We focus on road cycling due to its particularities, but most of the work presented here can be applied to other sports where participants are distributed in groups during the race, such as some athletics events (e.g. marathon, race-walking, or triathlon), open water swimming competitions, etc.

Bicycle races are very popular, especially in Europe. For instance, the *Tour de France* 2020 gathered more than 40M viewers [Bas20] only in France and, according to different media (e.g. BBC), it reaches around 3-4 billion television viewers all over the world. The most important multiple-stage events in men's professional road cycling are the *Tour de France*, the *Giro d'Italia* and *La Vuelta*, in Spain. While for women, the most important multiple-stage race is the *Giro Rosa* in Italy. We have real data from different stages of the Tour's 2018 edition to test our tool. But before we dive into the concrete requirements of our application, we need to better understand the data we need to visualize, and why.

3.1. Background

The professional road cycling season is formed by different events. Some of them are a single-day race, such as the World Championship or the so-called *Monuments: Milan-San Remo, Tour of Flanders, Paris-Roubaix, Liège-Bastogne-Liège* and *Giro di Lombardia*. Others are multiple-stage races that span from one week (e.g. *Paris-Nice* or *Tour de Romandie*) to three weeks duration: *Tour*, *Giro* and *Vuelta*. In multiple-stage events, the winner is the rider that accumulates a smaller amount of time throughout the stages (at the end of some stages, 10, 6, and 4 seconds are discounted to the accumulated time of the three first ranked riders of the stage as a bonus). There are two types of stages: the regular ones and time trials. In regular stages, all riders start at the same time. In individual time

trials instead, each cyclist starts after two minutes of the previous one, in reverse order with respect to the individual general ranking at the beginning of the stage, being the leader the last one to start. That gives the leader the advantage of knowing the (partial) timings of previous cyclists. Furthermore, some races include team time trials, where the members of each team compete together.

The profile of the stages, whether they are flat, hilly, or mountain, has a very strong impact on the performance of the cyclists. Mountain stages, particularly those that finish at summits, typically yield larger time differences between the riders at the end of the stage, due to the extreme effort required for climbing. Cyclists also present different athletic conditions: heavier riders are capable of developing more power in short distances, which makes them more suitable to win massive sprints in flat stages, while lighter riders are better-suited to climb faster long mountains. Thus, the profile of the stage is usually an indicator of the type of cyclist that might win. Weather conditions (rain, wind, cold, ...) also have influence (some cyclists are more skilled to run on wet pavement, for example).

Another important feature that sets apart cycling races from others is groups. At the beginning of regular stages, all riders start together, but along the route, this initial big group may split into smaller ones or regroup. When the splits successfully open a gap between them and other riders, are called breakaways. And the riders trying to catch them are the so-called chasing group. The formation of groups is key to several aspects of the stage. The two more important are time and fatigue. First, whenever a certain number of contenders reach the finish together, all of them are assigned the time of the first one in the group, even if the group takes many seconds to cross the line. Second, a group of several people can collaborate and increase the speed with a lower fatigue toll because the rider who is in the lead, is the one receiving a larger amount of aerodynamic pressure. This technique, riding very close to make the maximum use of their slipstream is called drafting. Therefore, if cyclists alternate pulling (leading the group), they can go faster, or at least, maintain the speed with less effort. Of course, very small groups can just use this strategy in a limited fashion. The main group is the largest and it is known as the *bunch*, the *pack*, or the *peloton*, and it is commonly another reference to take into account since it is the one with greater possibilities of chasing any breakaway by groups or solo riders.

Multiple-stage cycling races usually include different ranking categories. They may vary depending on the event, but the most common are *individual* general ranking (or *general classification*), *mountain*, *points*, and *youth*. Finally, there is also a prize for the best team as a whole. In some races such as the *Tour de France*, the considered most combative rider of the race is also rewarded and is elected by a jury. The leaders of each category are easily recognizable since they wear distinctive colored jerseys: yellow, white with red polka dots, green and white for the individual, mountain, points and youth ranking leaders respectively in the *Tour*. The *Giro* and *La Vuelta*, for instance, change the colors and the individual general ranking leaders wear pink and red jerseys, respectively. Although most of the categories are based on the actual riding time of cyclists, points and mountain are point-based rankings where contenders receive a certain amount of points depending on their rank at intermediate sprints and the finish line in the former case, and

the summit of significant climbs in the latter. Thus, obtaining the ranking information at those landmarks is also important.

Though times are measured at the finish line for classification purposes, stages are long journeys that last for several hours. Therefore, delays can easily be in terms of minutes. As a consequence, throughout a stage, a contender may have an advantage that, if kept until the finish line, might make the rider leader of one of the rankings. Some media call this situation *provisional* or *virtual* leaderships. Because team managers and cyclists may take different decisions in route according to virtual leadership changes, this is another important aspect to analyze and visualize.

3.2. Requirements

Once we know the kind of data we are dealing with, we need to define our requirements. This was completed as a multi-stage process. First, two authors, who are bicycle enthusiasts, defined a set of initial questions that would be interesting to answer through such an application. Then, we proceeded to analyze how users typically obtain this information. To know the outcome and highlights of a stage, there are four typical sources: *i)* TV and radio live broadcasts, *ii)* real-time information websites, *iii)* sports newspapers chronicles, and *iv)* infographics developed by the communication media.

Thus, we analyzed the contents provided by TV/radio broadcasts with the chronicles of online versions of several newspapers such as *l'Équipe* (France) or *La Gazzetta dello Sport* (Italy), the information published in specialized websites, and the infographics that are available through different means. This led us to create a list of available information bits (Table 1 in Section 7 summarizes the information that can be found in the media, and compares it to the data provided by our visualization tool).

After the in-depth analysis of the information present in different media, we can conclude that the amount of data provided varies largely. The highest amount of information is supplied by TV broadcasts or real-time websites that stream live data. For example, the Tour's official website [Ama20], provides real-time information during the stages and lets users perform some queries. The website consists of a view that shows the profile of the stage together with information on the classification of the riders at different points, and other data. Unfortunately, it is only a snapshot of the stage at the current instant and does not allow the user to know how the stage has evolved to arrive in the present situation. Other insights, such as virtual leaders are missing. Besides, the application only works in real-time. After the race, the only information regarding the evolution of a stage that remains is the average speed of the front of the race and the final rankings.

Newspaper chronicles can be classified into two different types: simple summaries that just provide the highlights, and larger chronicles that narrate, in a more detailed way, the events of the stage. The latter conveys a higher amount of information, but they are still very far from the details supplied in TV broadcasts. Important events, such as rider withdrawals, the group members, or relative rankings at specific kilometers, do not appear. Lastly, many different media use infographics to show the summary of a stage but they are commonly simple rankings with a scarce amount of details on the riders (normally rank and delays from the leaders). Another

used technique is drawing the stage profiles with the positions of the riders at some point. In this case, groups are shown, but with little details (e.g. only the number of components, not the individual riders), and are typically valid only for a specific instant of time. None of the previous sources lets us check, for instance, what a specific participant did along a stage or the full race, or inspect how many cyclists were virtual leaders during the stage. Even the most detailed sources, TV/radio broadcasts, and real-time applications are quite limited in usage: if users want to get an understanding of what happened, they have to watch the whole broadcast or use the website during the stage, which lasts for several hours.

To see whether this available information corresponds to what enthusiasts would expect, we conducted a questionnaire with 20 practitioners, most of them belonging to two different cycling clubs. In it, we asked about the relevance that they would give to the different facts that we deem as important, and to help us find others. The results are summarized as column *U_{sr}* in Table 1. After analyzing all this information, we set up a set of global requirements, as well as a set of queries that we must address. Global features are:

- The application must let the user navigate through several stages, required to visualize not single but multiple-stage events.
- For each stage, the system must communicate the relevant facts (e.g., position and timing) at the finish, but also during the stage.
- For each stage, the visualization system must support the visual analysis and exploration of groups (what groups existed at a certain point, its members, etc.).
- The system must provide simple access to several queries to relate contenders using the current stage, but also general rankings information.

Since the visualization of one-day cycling races is the same as a single stage of multiple-day events, our application is focused on these latter types of competitions, where data of the different rankings have to be accumulated among the stages. In this way, we ensure that our application can deal with both types of events. Following the results of the questionnaires to users, we implemented a multi-view visualization system with several elements coordinated to facilitate the narration of a race, and different interrogation mechanisms to facilitate getting insights on each stage. Some examples of relevant facts (basic questions, **BQ**) highlighted by bicycle races enthusiasts, are:

- **BQ1:** Who won a specific stage?
- **BQ2:** What was the final ranking?
- **BQ3:** Which are the delays between riders at a specific instant of the stage?
- **BQ4:** Who is the leader of the race after a certain stage?
- **BQ5:** Who leads the other rankings after a certain stage?
- **BQ6:** Which cyclists belong to each team?

Most of these questions could be answered by adapting some of the ranking visualization systems mentioned in the previous work, but there are other, more sophisticated, that are quite specific to multi-stage cycling races. In the following list, we provide a small set of specific queries (**SQ**) that users would like to answer with the help of the application:

- **SQ1:** Who is the *virtual leader* of the race at a certain point of the route?

- **SQ2:** Which riders arrived first at a certain categorized climb or intermediate sprint?
- **SQ3:** Has there been attacks and breakaways in the stage?
- **SQ4:** What groups appeared in a certain stage? Do they represent a threat to the leader of the race?
- **SQ5:** What is the time difference between a contender and the leader at a certain point of a stage?
- **SQ6:** How much did a certain categorized climb affect the development of the stage?
- **SQ7:** How long did a certain cyclist lead a specific ranking?

The above questions are just some examples of data that users can only get to know with a lot of contextual information (from previous stages, from the individual general ranking, from how stages have developed...). Catching up the TV streaming of a certain stage will only provide the audience with a number of facts that is clearly limited, and only a system that conveys information of the full stage can help understanding how it evolved.

3.3. Stage encoding

It is the central component, and it lets the user understand how the riders have fared along the route. We have designed this view as follows. Two components show the evolution of the stage in terms of individual contenders' positions along the route (see Figure 3-c), and conveys both, the groups (top) and the geographic profile of the stage (bottom), with the sampled kilometers and points of interest. The sampled route points are labeled as it is common in bicycle races, with the number of kilometers to finish. Though information regarding every meter of the course could theoretically be obtained, such an amount of detail is unnecessary for our purpose, since relevant events happen now and then and span for a longer time and distance. Therefore, sampled kilometers (e.g. every 10 kilometers) act as reference, and we can get the whole amount of details at these points (current ranking, current timing, height, the time between groups or riders, virtual positions in the general classification...). The riders' positions have two rendering modes: lines, and Sankey diagrams. Line rendering (similar to [Woo15]) can be used to identify individual riders. However, instead of creating straight lines between samples, we use Bézier curves, which are more aesthetically pleasing, and to avoid giving the impression of sharp speed changes along the course. This technique does not convey groups. Our second mode (enabled by default), renders the riders as a Sankey diagram, where the nodes are groups. At a glance, users can see whether the stage has been calm, with (almost) no breakaways or, on the contrary, it has had a lot of activity. The vertical axis encodes both time and groups' size. The bottom view also informs about categorized climbs (red, as they count for the mountain ranking), with their location in the route and the category of the climb. Intermediate sprints (green, since they serve for the points ranking) are also displayed. Individual riders can be highlighted by name, team, or through dedicated leadership buttons (race leader, mountain leader, and others). As we will see, by selecting kilometer samples, intermediate sprints, or categorized climbs, users get insights on the classification at those points.

One important aspect of the stage view is how groups and time are encoded. In bicycle races, the encoding of groups is as important as time differences. Therefore, we use the vertical dimension

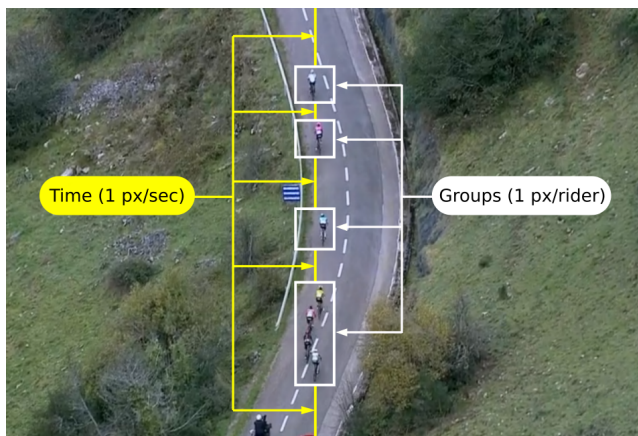


Figure 2: The vertical axis encodes both groups and time between them, in the same way riders are arranged on the road. The three riders at the top form individual groups, while the last make another group.

to encode both. We start with the following two observations: *i*) the number of riders is standard for the 3 big races (176 riders in 22 teams for the period 2018-2020), *ii*) we have around 500-600 pixels in vertical for the stage view, for a standard screen (typical vertical resolutions in 2020 range from 1080 to 1440, but grow up to 2880 for larger monitors). After evaluating different strategies to assign different space to groups and time, we ended up with the following method: each second is depicted with one pixel, and each rider in a group too. As a result, time gaps can be visually compared, as well as the group sizes (see Figure 2). This works for most of the stages since delays of several minutes can be encoded without requiring any vertical scrolling. Mountain stages, especially those with uphill finishes, may exhibit large time variations from the first to the last rider. In those cases, the view is initially centered on the riders that ended up first, and vertical scrolling (hiding the profile, if necessary) is allowed to grant more space to analyze the events.

3.4. Application layout

Our application consists of six linked views (see Figure 3). The *stage view* (*b*) is the central component, and contains the information regarding a single stage. This can be changed using the left widget (*race overview* (*a*)), which contains all the stages of the race. The *selection panel* (*c*) is used to select individual riders or teams, and it is intended to facilitate queries, though other filtering methods are also provided. At the bottom, we have three different views: *stage rankings* (*d*) (left), *details on demand* (*e*) (center), and *riders/teams information* (*f*) (right).

Race overview. The race view shows the profiles of all the stages that compose the race. As mentioned earlier, the profile is crucial for the development of a stage. The different grey levels indicate whether there is data available for the stage. Black indicates which of the stages is currently selected. By clicking on one stage, the main view shows the evolution, and the other views are also updated. Two scrolling buttons let the user inspect the other stages.

Stage view. It is the main component of our system. As described above, it is a multi-view widget with two main regions, the top one depicts the riders' positions throughout the course, and the profile is shown at the bottom. This elevation map helps giving users more context on the events that happen along the course. Note that the slope of the route crucially influences the riders' behavior (they may decide to attack if they feel strong, for example), in addition to the actual positions of other contenders (i.e., a participant in the main group will not attack if any of his/her teammates is in a break-away group). Therefore, the profile may be kept visible even if we need to scroll the riders' view. The top view has some buttons to highlight relevant facts such as the performance of the leaders of the different categories, illustrated with the corresponding icons. The bottom view also provides interactive elements: both sampled kilometer points, category climbs, and intermediate sprints can be clicked to get the current ranking at those points.

Riders/teams selection panel. It is placed at the right of the application. It shows the list of cyclists, ordered by their id, but it can be toggled to see the participating teams. All the elements can be clicked to see the details (e.g. performance, result, race time, time lost/won, etc.). This view also has a search box to facilitate looking for a concrete rider or team.

Stage rankings. This view (bottom left) is devoted to display the different rankings at the end of the stage. Buttons to show each specific classification are also provided.

Details-on-demand view. The bottom center view displays the details of the selected elements. It shows, for instance, the position of the riders at the sampled kilometers (including the start and the finish line), the points obtained at intermediate sprints or categorized climbs, the names of the riders in a specific group, whether it is a leading group (front of the race), the peloton or another chasing group. Moreover, for each contender, it also shows his/her current position in the stage, his/her current position in the general classification, as well as the virtual rank in the different classifications. In addition to this, the time he/she is gaining or losing with respect to the leader is also depicted.

Rider/team information. Finally, the last view (bottom right) presents the information of the selected riders or teams. For the participants, we show the picture, full name, number, team, some biographic data, and the details regarding the ranking in all categories (Figure 3 (*f*)). The team details show the full name, the country, the managers' names, the jersey (including the sponsors), and the eight members of the team, as shown in Figure 1.

4. Data extraction and processing

The information used to create the *stage view* and the one displayed in the other components of our application (see Figure 3) has been obtained as follows: the tracking data (rider name, number, team, group, position and delays from the front of the race), which is used to visualize the evolution of the stage, was downloaded from the application *racecenter.letour.fr*, at the starting and finishing lines, and every 10 kilometers. We tried to get at *La Rosière* climb (stage 11) the information at every kilometer, but the GPS signal was lost

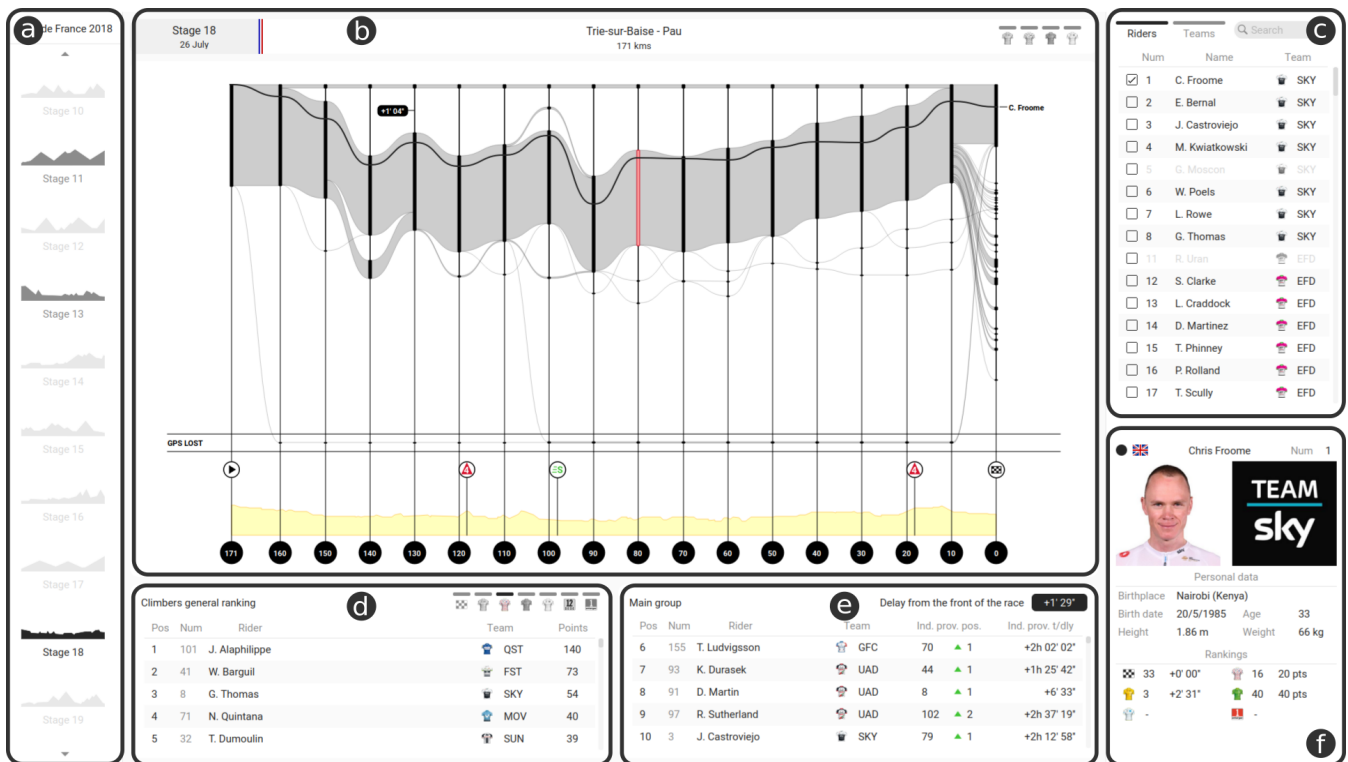


Figure 3: Application layout. Our application consists of six linked views: (a) Race overview, (b) stage view, (c) riders/teams selection panel, (d) stage rankings, (e) details-on-demand view and (f) rider/team information.

frequently. The stage data (name, day, kilometers, and rankings) was obtained from *Le Tour* website, together with the riders' icons, logos of the teams, stage profiles, and jerseys' icons that appear in other views. Flag icons were obtained from *www.gosquared.com* and the rest of them are our design. Extra information regarding the points and rankings at categorized climbs were extracted later from *www.cyclingnews.com*. Bio information from the riders shown in the *Rider/team information view*, comes from Wikipedia.

The data had some problems: some positions were wrong (e.g. two riders assigned to the same position), some times were also wrong (e.g. rider in position $n + 1$ assigned a time larger than rider $n + 2$ and n). Sometimes, riders in a group were assigned to the main group and others. All these issues were corrected manually from a small group of stages, which are the ones used for the application demonstrations. Once the data was cleansed, we derived provisional positions and rankings (calculated at every sampled point): ranking position, ranking gap, and general riding time.

5. Interactive data exploration

This section introduces the interaction techniques present in our application, and how the different views are used coordinately to answer questions about the evolution of a stage or the whole race.

Once we open the application, the race view shows all the stages, and we start selecting one. Once we select a race, the stage ranking view is updated with the final classification (by default), where we

can see the winner of the stage and the position of all the cyclists at the finish line (BQ1, and BQ2). Riders or teams will appear once we select them in the rightmost panel. Users can also click on the buttons indicating the different leadership categories to know who leads each one (BQ5). This is illustrated in Figure 3 (d), where the climbers ranking is shown. Relevant events that happened along the stage, are sometimes encoded explicitly, such as whether there was a breakaway (SQ3) that can be easily seen as groups that appear (and might disappear along the track) or can be found by exploring the views in different ways. In general, users will mainly pivot around the central view, but, as we will see, the exploration could start by searching for a specific participant or team. Next, we analyze how the designed linked views can be used to provide the answers to the main questions users can have about the race.

Main view. The *stage view* is central to all the analysis. It works as an overview because the whole stage is summarized there. As mentioned, if the stage has had really long differences between the contenders, more vertical space is needed. We highlight this through a small button at the bottom that can be used to go down. Right-click and dragging also let the user move the stage view. Besides, zooming is also allowed with the mouse wheel. To inspect an individual rider, the right menu can be used, through the interactive search or by scrolling down the list of participants. Multiple riders can be selected at once, as well as whole teams. Moreover, the stage view also has buttons to highlight the behavior of the leaders of the different categories. The stage profile is also interactive, let-

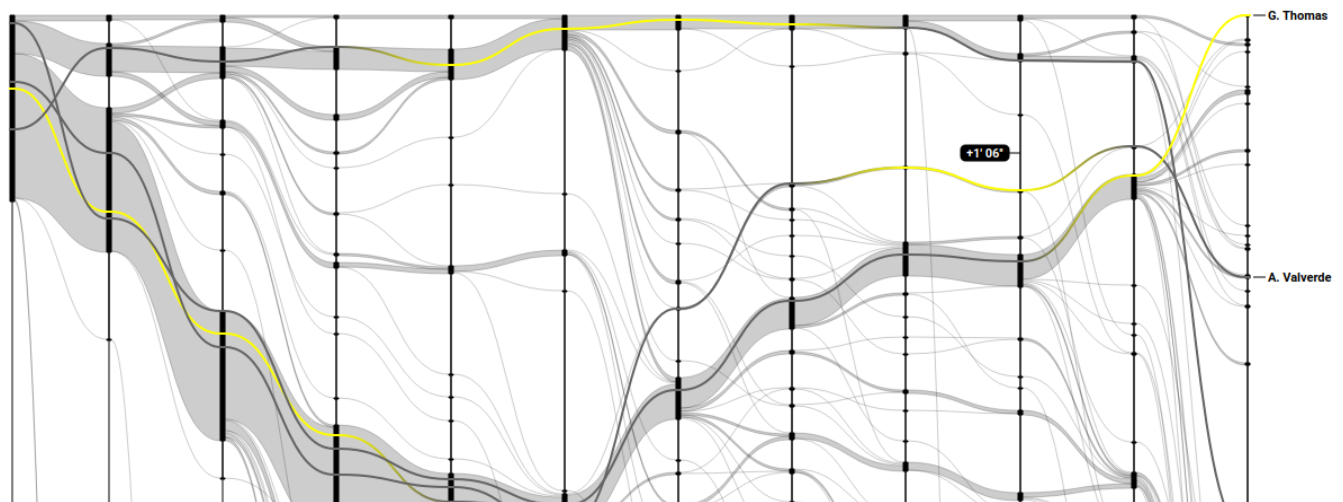


Figure 4: Analysis of virtual leaders along a single stage. Van Avermaet was the leader for the first 30 km of the stage. After that, he lost track, and Pauwels took over for 40 km. Then, Valverde shortened the delay and, since he was better in the general classification, caught up as the virtual leader for 20 km. Finally, Thomas surpassed Valverde, and ended up as the leader of the race.

ting the users select kilometer points as well as categorized climbs or intermediate sprints. This causes the details view to update with the ranking of the riders passing through those points first.

Detailed insights. The (multiple) selections of individual riders using the rightmost view allow us to get detailed information on their performance and their time at sampled points. When cyclists are selected, their path is highlighted in the main view, and the details panel is filled with their id, name, team, and distance (in time) to the leader of the stage at this point (SQ5). It also shows the information on whether the contender is currently improving or worsening its position in the general ranking. By clicking on an uphill finish, and selecting the riders there, and inspecting their evolution for the rest of the stage, one can see how the finish affected the overall stage (SQ6). By clicking on any of the groups the rider has been part of, the group view is populated with the members of the groups (see Figure 3 (d)), their provisional time in the race, and other details. This can be used to identify whether a certain group poses a threat to the leadership (SQ4). When a group is selected (including the peloton), the paths of all the members along the stage can be made visible at once by clicking 'g'. If two riders are selected, we can obtain their stage time at any moment of the course and calculate their difference by hovering over the sampled kilometers (BQ3). The bottom left panel lets us see the general classification at the end of the stage (BQ4) or the rankings in other categories. Teams selection is also possible, yielding the performance of all the riders of the team, as well the above-mentioned details (BQ6).

Other interaction techniques. The stage view can be dragged and zoomed. When a stage has had very large time differences between groups, the vertical axis might cover a larger space than available. In this case, instead of compressing the view, we prefer to provide scrolling, so that the differences in time are proportional to the real ones. This way, when the user changes the stage, a better impression of the real differences at the finish is visualized. Zooming can

be used to better see the relative positions of selected riders. The main view also has button shortcuts to see the different leadership categories. By clicking on any of those buttons, the riders in the category are highlighted. These include virtual leaders too, so if in a certain stage the leadership has changed, it is explicitly communicated (SQ1). We can filter out cyclists by team or leadership by using the controls in the right toolbox. This way, we can answer questions regarding leadership in different categories (BQ4, BQ5), or the riders belonging to a team (BQ6). Another filtering operation happens with the combination of the stage profile. By clicking over a kilometer point, we can see which contenders have been the first to reach a certain summit (SQ2) or intermediate sprint. More details can be seen in the accompanying video.

To illustrate these features, we propose two use cases: the changes of virtual leaders along a stage (an example of SQ7), and the analysis of how teams controlled a stage to get a sprint finish.

Use case 1 - Virtual leader analysis. One of the complex scenarios is the analysis of virtual leaders during a stage. Some stages stay quite calm, but some days the general classification changes continuously and it is difficult to follow. Though we can get sometimes such information in TV broadcasting, it is only given now and then, and there is no tool that summarizes the evolution of very complex stages accurately. Aficionados may be interested in how many virtual leaders a stage has had, who were those, and for how long. Our stage analysis solves this by selecting the leader in the main view. By doing so, the system emphasizes the different riders that have been the virtual leader with the common yellow color. The portion of the stage where they were not leading, is also highlighted so that we can understand what happened before and after. This is helpful to reveal interesting scenarios where a cyclist has broken away, escaped alone or in a group for several kilometers, but then the effort does not pay off and is even unable to end with the main group due to the exhaustion at final kilometers before the finish. This is

clearly depicted in Figure 4, which belongs to stage number 11 of the Tour 2018. The stage was highly complex, to the point that, up to four participants were virtual leaders along the course. After we inspect the overview depicted in the main view, we can further analyze how these changes happened along the stage by clicking on specific kilometers and consulting the virtual general classification there in the *details-on-demand view*. In this case, G. Van Avermaet was virtually leading the race while S. Pauwels was not even in the top 5. Then, Pauwels appeared and led the virtual classification for 40 kilometers, where A. Valverde took over the virtual leadership. Near the end, we can see that G. Thomas surpasses Valverde and finishes the stage as the leader. The *details-on-demand view* also provides information on the changes in the general classification (how many positions a participant is currently escalating or losing) that are likely to happen.

Use case 2 - Shaping the stage for a sprint finish. Team members may have highly different physical qualities: some are climbers, others are good descendants (may be able to escape in downhill tracks), the rouleurs who perform well in hilly and flat terrains, and some are sprinters. Those are specialized in short but intense efforts and achieve high finishing speeds. If a team has good sprinters, they are interested in chasing the breakaways, so that the sprinters have possibilities of winning the stage. To this end, team members take the lead in the peloton and pull, so that gaps with escaped groups are bridged before the end of the stage. We can analyze if this strategy is followed in a certain stage by studying whether teams with good sprinters had some of their members pulling the main group or not. In TV broadcasts, the peloton lead is usually mentioned, but commonly just when teams are doing a relevant paper. It is difficult to get more information (e.g. the actual members of the team or teams that are pulling, or how long the effort has lasted). With our system, the user can check who is leading the peloton at every sampled point, and the teams can be compared to the ones that finished first quite simply. By selecting the stage ranking in the bottom left panel, we can see the cyclists who took part in the massive finish sprint. By iterating through different kilometer points and selecting the main group at those points, the *details on-demand view* is populated with the riders that lead the peloton, and their delay with respect to the front of the race. Thus, we can easily see that some riders are continuously pulling, and analyze if their efforts are yielding results. There is another alternative to solve this problem: team-wise. Teams with good sprinters can be selected using the *riders/teams panel*, and see how all their members fared along the stage. Therefore, we can easily see whether they were pulling in the main group or not. By combining this information with the one provided in the *stage rankings view* and finishing positions (*stage view*), we can determine whether the efforts were successful.

6. Evaluation

The limitations to movement and social contacts in Catalonia while preparing this paper prevented us to show the application in person to users (except for 2 people). Moreover, since the application was developed in Linux, we could not send them a binary, as they were Windows users. Thus, for each user (8 bicycle enthusiasts and practitioners), we performed the evaluation in the following way: using Zoom, which preserves quite well the video streaming quality, we

demonstrated the application for around 10 minutes. Then, we had an interactive session with them, that lasted from 20 to 90 minutes, where users asked us about the application, and performed different tasks by telling us what to do at every moment. Then, they completed a questionnaire offline where the grading was from 1 (low) to 5 (high). The average of all answers was:

- It serves to understand the evolution of a stage: **5**.
- The most relevant information of the race is shown: **4.5**.
- Would you use it to know of what happened in a stage? **4.875**.
- Do you think the system is useful? **Yes** (everybody).

In the final observations field, users made some suggestions, such as adding more color to the peloton. Moreover, they also found the system useful to visualize other kinds of races, such as marathons and triathlons, and other users mentioned F1 races, where the application should be suitable to predict overcuts and undercuts.

In addition to this, in the final track of our development, we also showed the application to two domain experts: a person with several years of experience as assistant sports director and race mechanic in different professional cycling teams, and the sports director of a women's professional road cycling team. The first expert was able to evaluate the application in person. The process went as follows: he first thought and defined some use cases he wanted to see, and then, he used the tool to solve them. The use cases he solved were: *i*) determining the position of riders at any moment, *ii*) riders' classification at the end and during a stage, *iii*) full ranking at the start of a stage, *iv*) querying for team components in groups during a stage, *v*) intermediate rankings at points of interest, and *vi*) highlighting riders of interest. Initially, he thought it was a finished commercial application. The second domain expert was not available in person (he is living in Switzerland), and we carried out the process as with the other users: first a demo, then interacting on his behalf, and finally, with a dialog analyzing the possibilities that lasted for more than one hour. He mentioned names of companies that gather the data, and suggested us to create the final product with further analysis on parameters that the team can gather from their riders. Both experts expressed their interest in having an application such as this one available. In its current stage, it could be used for post-stage analysis. In addition to this, both experts believed that, with real-time tracking data, it would be even more useful. Since it could help teams' directors to take better and quicker decisions than they currently can now. At the moment, their main source of information during a stage is the *Tour Radio*, which broadcasts in seven languages, and the information they obtain is scarce and delayed. Both of them found the application easy to use and learn. They also believe that for real-time use, a tablet version would be ideal.

7. Conclusions and Future Work

In this paper, we have presented a new application that facilitates the exploration of multi-stage cycling races. With the application, users can see the important events that happened along the stage, not only the final classification. Our multi-view, interactive system lets the user navigate through stages, and getting a lot of insights that would not be possible unless the aficionado watches the full

TV broadcasting. Table 1 shows the elements as analyzed by the different communication techniques, as well as what users consider useful, according to our previous study (see Section 3.2). Although TV broadcasting offers the largest amount of data compared to the other media, its information always refers to relevant riders. Little information regarding other riders or team components is provided. The other big source of information is newspapers' chronicles. We analyzed the contents from several newspapers (online versions): l'Équipe (France), la Gazzetta dello Sport (Italy), As and Marca (Spain), as well as two dedicated websites: Cycling News (UK), and Cycling Tips (Australia). As it can be seen, the amount of information we provide can only be compared to TV broadcasts. However, access to the full data can only be obtained if the user spends several hours watching the broadcasting, while our application facilitates getting insights in just seconds. Since our system has been developed around the evolution of a stage, and the groups, it can also be applied to other sports where belonging to a group may be a strategic move, such as car races, motorbikes, or running.

Our system is able to match the amount of information provided by other sources, such as TV broadcasting, at a lower time cost. However, there are still some aspects that can be improved. Our data gathering process is manual since no public databases are available. As a result, we sample the data at certain kilometer points. This is enough to provide a narrative on the stage. However, some events, such as short-lived breakaways or attacks, might be undetected. This is not a problem, because these events have almost no effect on the stage. If necessary, a finer sampling might be carried out to get this information, but the general representation should not change to avoid cluttering. Since the data gathering is manual, we also did not get information of all the stages, but only a number of them, to illustrate the features of our application. The inaccuracy of GPS positioning also prevents tracking all riders without problems. Sometimes, the signal of a cyclist is lost. We represent it in the main view since it is all the information we have. A better solution might be possible, such as keeping the position of the participant at the same pace as the group he/she belonged to before losing the signal, but it has its shortcomings, that may be important if the person losing the signal is some relevant contender. Another information that is difficult to obtain is when a rider falls. If they fall and catch the main group later, it may be reflected in the GPS positioning, but it is not sure.

In the future, we want to address some of those limitations. First, we would like to have a denser sampling. A one-kilometer sampling would probably be enough, but the main view would become too cluttered, so maybe dynamic adaptation (to relevant facts or zooming level) might be a solution. The inclusion of other meta-data could also be interesting (e.g., previous participation in the same race and achievements), could be shown. Finally, we would like to create an automated system for data gathering in real-time, coupled with the proper cleansing and data derivation tasks, so that the application can be updated instantly. In this case, we would need to rethink the stage layout view to tackle dynamic data.

Acknowledgments

The authors want to thank L. Díaz-Iriberry and A. González-Tablas for their insightful comments. As well, we thank all the people that

Data	Usr	TvB	LCh	Ours
Main group evolution	F	✓	F	✓
Rider pulling main group	F	F	F	✓
Team pulling main group	F	✓	F	✓
Individual leader (name)	✓	✓	F	✓
Individual leader pos.	✓	F	R	✓
Change individual leader (Y/N)	✓	✓	✓	✓
Virtual individual leader	F	✓	✓	✓
Breakaway (Y/N)	✓	✓	✓	✓
Breakaway (riders)	F	✓	✓	✓
Chasing group(s)	F	✓	✓	✓
Straggler group(s)	F	F	F	✓
Crashes and flat tyres	F	✓	✓	✓
Mountain pass category	✓	✓	F	✓
Mountain pass ranking	F	✓	R	✓
Intermediate sprints	F	✓	R	✓
Intermediate sprints ranking	R	✓	R	✓
Bonus/points sprints	F	✓	R	✓
Withdrawals	F	F	R	✓
Team members	R	R	✗	✓
Riders ind. ranking pos.	✓	R	✗	✓
Riders team pos.	R	R	✗	✓
Weather conditions	✓	✓	F	✓
Gap between groups	✓	F	R	✓
Stage winner	✓	✓	✓	✓
Stage ranking	✓	✓	✓	✓
Individual ranking	✓	✓	✓	✓
Other rankings	✓	✓	F	✓
Time won/lost 1st	✓	✓	✓	✓
Time won/lost 2nd	F	✓	✓	✓
Time won/lost 3rd	F	✓	✓	✓
Individual leader	✓	✓	✓	✓
Mountain leader	F	✓	R	✓
Points leader	F	✓	R	✓
Teams leader	F	✓	R	✓
Most combative rider	R	✓	R	✓
Youth leader	F	✓	R	✓
Bonus end stage	✓	✓	R	✓

Table 1: Types of relevant facts during a stage from different sources. The used acronyms are: Usr (cycling enthusiasts), TvB (TV Broadcasting), LCh (Long newspaper Chronicles), and Ours. Websites streaming live stages are not considered because the information provided, is only available a specific instant. Infographics appear as part of the information provided by TV broadcastings and chronicles. Pink items refer to information that changes throughout the stage, while blue encodes static data known at the finish line. Check and cross marks indicate whether the information is provided or not, F means frequently, and R indicates rarely.

participated in the user study. This work has been supported in part by project TIN2017-88515-C2-1-R(GEN3DLIVE), from the Spanish Ministerio de Economía y Competitividad, by 839 FEDER (EU) funds and by the Grant No. PID2019-106426RB-C31 / AEI / 10.13039/501100011033 from the Spanish Government.

References

- [AAB*17] ANDRIENKO G., ANDRIENKO N., BUDZIAK G., DYKES J., FUCHS G., VON LANDESBERGER T., WEBER H.: Visual analysis of pressure in football. *Data Mining and Knowledge Discovery* 31, 6 (2017), 1793–1839. 2
- [Ama20] AMAURY SPORT ORGANIZATION: Official website of Tour de France 2021. <https://www.letour.fr/en/>, 2020. last accessed, November 2020. 4
- [Bas20] BASSAM T.: Report: Tour de france draws 40m domestic tv viewers. <https://www.sportspromedia.com/news/tour-de-france-2020-tv-ratings-ftv-eurosport>, 2020. last accessed, November 2020. 3
- [Bat06] BATTY M.: Rank clocks. *Nature* 444, 7119 (2006), 592–596. 3
- [BCC*19] BASDERE M., CANIGLIA G., COLLAR C., ROZOLIS C., CHIAMPAS G., NISHI M., SMILOWITZ K.: Safe: A comprehensive data visualization system. *INFORMS Journal on Applied Analytics* 49, 4 (2019), 249–261. 3
- [BK19] BAČÍK V., KLOBUČNÍK M.: Possibilities of using selected visualization methods for historical analysis of sporting event—an example of stage cycling race tour de france. *Quaestiones Geographicae* 37, 3 (2019), 5–22. 3
- [BLC*14] BIALKOWSKI A., LUCEY P., CARR P., YUE Y., SRIDHARAN S., MATTHEWS I.: Identifying team style in soccer using formations learned from spatiotemporal tracking data. In *2014 IEEE international conference on data mining workshop* (2014), IEEE, pp. 9–14. 2
- [CEGH17] CHAWLA S., ESTEPHAN J., GUDMUNDSSON J., HORTON M.: Classification of passes in football matches using spatiotemporal data. *ACM Transactions on Spatial Algorithms and Systems (TSAS)* 3, 2 (2017), 1–30. 2
- [CF13] CAVA R., FREITAS C. D. S.: Glyphs in matrix representation of graphs for displaying soccer games results. In *The 1st Workshop on Sports Data Visualization. IEEE* (2013), vol. 13, p. 15. 2
- [CLX*16] CHEN W., LAO T., XIA J., HUANG X., ZHU B., HU W., GUAN H.: Gameflow: narrative visualization of nba basketball games. *IEEE Transactions on Multimedia* 18, 11 (2016), 2247–2256. 2
- [CPG*15] CHUNG D. H., PARRY M. L., GRIFFITHS I. W., LARAMEE R. S., BOWN R., LEGG P. A., CHEN M.: Knowledge-assisted ranking: A visual analytic application for sports event data. *IEEE Computer Graphics and Applications* 36, 3 (2015), 72–82. 2
- [Dem13] DEMAJ D.: Geovisualizing spatio-temporal patterns in tennis: An alternative approach to post-match analysis. In *Proceedings of the 26th International Cartographic Conference* (2013), vol. 9, p. 16. 3
- [DKVS14] DIETRICH C., KOOP D., VO H. T., SILVA C. T.: Baseball4d: A tool for baseball game reconstruction & visualization. In *2014 IEEE Conference on Visual Analytics Science and Technology (VAST)* (2014), IEEE, pp. 23–32. 2
- [FMB*15] FRANKS A., MILLER A., BORNN L., GOLDSBERRY K., ET AL.: Characterizing the spatial structure of defensive skill in professional basketball. *The Annals of Applied Statistics* 9, 1 (2015), 94–121. 2
- [GGB14] GOUSIE M. B., GRADY J., BRANAGAN M.: Visualizing trends and clusters in ranked time-series data. In *Visualization and Data Analysis 2014* (2014), vol. 9017, International Society for Optics and Photonics, p. 90170F. 3
- [GH17] GUDMUNDSSON J., HORTON M.: Spatio-temporal analysis of team sports. *ACM Computing Surveys (CSUR)* 50, 2 (2017), 1–34. 2
- [GLG*13] GRATZL S., LEX A., GEHLENBORG N., PFISTER H., STREIT M.: Lineup: Visual analysis of multi-attribute rankings. *IEEE transactions on visualization and computer graphics* 19, 12 (2013), 2277–2286. 3
- [GQR16] GROHJAN E., QUEALY K., ROBERTS G.: Fastest men in the world since 1896 - on the same track, 2016. <https://www.nytimes.com/interactive/2016/08/15/sports/olympics/usain-bolt-and-120-years-of-sprinting-history.html>. 2
- [HPG*19] HAN D., PAN J., GUO F., LUO X., WU Y., ZHENG W., CHEN W.: Rankbrushers: interactive analysis of temporal ranking ensembles. *Journal of Visualization* 22, 6 (2019), 1241–1255. 3
- [KYY*16] KAPLAN O., YAMAMOTO G., YOSHITAKE Y., TAKETOMI T., SANDOR C., KATO H.: In-situ visualization of pedaling forces on cycling training videos. In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (2016), IEEE, pp. 000994–000999. doi:10.1007/s12650-016-0349-7. 2
- [LOC*16] LAGE M., ONO J. P., CERVONE D., CHIANG J., DIETRICH C., SILVA C. T.: Statcast dashboard: Exploration of spatiotemporal baseball data. *IEEE computer graphics and applications* 36, 5 (2016), 28–37. 2
- [LTB16] LOSADA A. G., THERÓN R., BENITO A.: Bkviz: A basketball visual analysis tool. *IEEE computer graphics and applications* 36, 6 (2016), 58–68. 2
- [MLKS17] MIRANDA F., LINS L., KLOSOWSKI J. T., SILVA C. T.: Topkub: A rank-aware data cube for real-time exploration of spatiotemporal data. *IEEE Transactions on visualization and computer graphics* 24, 3 (2017), 1394–1407. 3
- [MLT*20] MIAU J.-J., LI S.-R., TSAI Z.-X., VAN PHUNG M., LIN S.-Y.: On the aerodynamic flow around a cyclist model at the hoods position. *Journal of Visualization* 23, 1 (2020), 35–47. 2
- [PBV16] PERIN C., BOY J., VERNIER F.: Using gap charts to visualize the temporal evolution of ranks and scores. *IEEE Computer Graphics and Applications* 36, 5 (2016), 38–49. 3
- [PSBS12] PILEGGI H., STOLPER C. D., BOYLE J. M., STASKO J. T.: Snapshot: Visualization to propel ice hockey analytics. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (2012), 2819–2828. 2
- [PVF14] PERIN C., VUILLEMOT R., FEKETE J.-D.: À table! improving temporal navigation in soccer ranking tables. In *Proceedings of the SIGCHI conference on human factors in computing systems* (2014), pp. 887–896. 3
- [PVS*18] PERIN C., VUILLEMOT R., STOLPER C. D., STASKO J. T., WOOD J., CARPENDALE S.: State of the art of sports data visualization. *Computer Graphics Forum* 37, 3 (2018), 663–686. 2
- [SCL*12] SHI C., CUI W., LIU S., XU P., CHEN W., QU H.: Rankexplorer: Visualization of ranking changes in large time series data. *IEEE Transactions on Visualization and Computer Graphics* 18, 12 (2012), 2669–2678. 3
- [SHI*15] STEIN M., HÄUSSLER J., JÄCKLE D., JANETZKO H., SCHRECK T., KEIM D. A.: Visual soccer analytics: Understanding the characteristics of collective team movement based on feature-driven analysis and abstraction. *ISPRS International Journal of Geo-Information* 4, 4 (2015), 2159–2184. 2
- [SJB*16] STEIN M., JANETZKO H., BREITKREUTZ T., SEEBACHER D., SCHRECK T., GROSSNIKLAUS M., COUZIN I. D., KEIM D. A.: Director's cut: Analysis and annotation of soccer matches. *IEEE computer graphics and applications* 36, 5 (2016), 50–60. 2
- [SJL*17] STEIN M., JANETZKO H., LAMPRECHT A., BREITKREUTZ T., ZIMMERMANN P., GOLDLÜCKE B., SCHRECK T., ANDRIENKO G., GROSSNIKLAUS M., KEIM D. A.: Bring it to the pitch: Combining video and movement data to enhance team sport analysis. *IEEE transactions on visualization and computer graphics* 24, 1 (2017), 13–22. 2
- [SJS*17] STEIN M., JANETZKO H., SEEBACHER D., JÄGER A., NAGEL M., HÖLSCH J., KOSUB S., SCHRECK T., KEIM D. A., GROSSNIKLAUS M.: How to make sense of team sport data: From acquisition to data modeling and research aspects. *Data* 2, 1 (2017), 2. 2
- [SSN*16] SHAO L., SACHA D., NELDNER B., STEIN M., SCHRECK T.: Visual-interactive search for soccer trajectories to identify interesting game situations. *Electronic Imaging 2016*, 1 (2016), 1–10. 2
- [SSS*14] SACHA D., STEIN M., SCHRECK T., KEIM D. A., DEUSSEN O., ET AL.: Feature-driven visual analytics of soccer data. In *2014 IEEE*

- conference on visual analytics science and technology (VAST)* (2014), IEEE, pp. 13–22. [2](#)
- [WDC*17] WALL E., DAS S., CHAWLA R., KALIDINDI B., BROWN E. T., ENDERT A.: Podium: Ranking data using mixed-initiative visual analytics. *IEEE transactions on visualization and computer graphics* 24, 1 (2017), 288–297. [3](#)
- [Woo15] WOOD J.: Visualizing personal progress in participatory sports cycling events. *IEEE Computer Graphics and Applications* 35, 4 (2015), 73–81. [2](#), [3](#), [5](#)