

CS 416

Narrative Visualization

Jared Coberly
coberly2@illinois.edu

Introduction

The Missouri River 340 is the longest non-stop river race in the world. It is conducted each summer along the lower Missouri River from Kansas City to St. Charles, MO. This narrative visualization introduces the reader to the race and racers that participate in this four-day paddle ultramarathon through an exploration of race results. At the end of the visualization, the reader will have an appreciation for the evolution of the race and evolution of its racers, as well as opportunities to explore the race by investigating individual racers, race divisions, among others, across time. The visualization utilizes HTML, CSS, JavaScript, d3, d3-annotation, and topoJSON libraries. No other toolkits are used. Race result data was obtained from publicly available websites and combined/transformed into a single comma separated value flat file.

Narrative Structure

The visualization uses an integrated narrative using an **interactive slide-show** structure. Primary navigation of the slide show is achieved through **pagination controls** (Next, Previous). This forces the reader to follow the narrative in a specified sequence and imposes the narrative structure. Each slide allows readers to **drill-down** and explore specific elements of the race or racer. Depending on the data being explored, this interaction may be **pop-up** tool tips, **filtering and re-rendering** within a scene, or launching a **new slide** that displays a deeper, richer, and focused view of the race or racer. The narrative begins linear but bi-directional. Readers can move forward and backward through the slides and can choose to explore different aspects of the race as they form questions about the race. As the slide show progresses, non-linear exploration may occur.

Visual Structure

The visual structure combines traditional visualization elements. The first slide is an **establishing shot**, providing context for the location and scope of the race; specifically, the first slide visualizes the entire state of Missouri, draws the path of the Missouri River across the state. **Animated transitions** highlight race checkpoints in the order of the race, giving the reader a sense of participating in the race themselves. In subsequent slides, **charts** are the primary visual structure used to guide the viewer through the data story. By using purposely selected chart styles, each scene filters the race data to narrate a specific part of the race-story and guides the user through data exploration. **Highlighting**, using annotations and occasionally animation, is employed to draw attention to significant points in the data. The visual **structure remains constant** for each slide: a simple header, a brief section of prose orienting the reader to the scene, a chart, and if applicable, the bottom of the slide displays a data-rich **table**. Additionally, the **color scheme is standardized** across the entire narrative. The “Men’s Solo (paddle)” division, for example, is always displayed with the same color regardless of which scene a reader is viewing. Specifically, an **accessible color palette** was chosen to have four divergent colors with different saturations for grayscale. By using a **template** with consistent visual structure and color scheme, the user can anticipate where data will be displayed in the scene and can connect data between slides.

Scenes

The scenes are described below. The scenes are ordered to take the reader through an exploration of the race result data. The scenes are ordered so that each new slide filters-and-focuses the data to a narrower view. This allows the reader to first gain an appreciation for the race, each race by year, the race divisions, and finally the racers.

- 1) The narrative opens with an establishing shot. This scene orients the reader to the race, the location, and the scope. The goal is to establish an appreciation for the immense distance the race covers.
- 2) The next scene provides a “first chart” overview of the race history, giving a very high-level view of the evolution of the race: participants and winning times. The stacked bar chart also introduces the concepts of “did not finish” and “did not start.” Annotations highlight the two sentinel data points for this high-level view: the course record (fastest time) and peak popularity. Clicking on a bar section allows the reader to explore those categories for each race year.
- 3) The third scene recapitulates the second scene, but focuses the view to finishers only. Where the previous scene introduced the reader to the race by year, and highlighted significant milestones (course record, popularity), this scene brings the focus back to the average racer and highlights mean versus median finish time for each race and annotates the overall average time to complete the 340 mile challenge.

- 4) Having brought the readers mind toward the racer in scene three, scene four further brings the racer to forethought: “How many racers attempt this for the first time and how many racers are MR340 veterans?” The chart style changes to a stacked line, helping the reader better appreciate the balance of newbie-to-veteran racers over the evolution of the race. With the idea of veteran racers raised, the scene also adds a table of racers sorted by the number of races in descending order. Introducing individual (and veteran) racers in this context is intentional. It helps frame the readers’ appreciation for the intense dedication many racers have for this challenge and, hopefully, gives context to the following scenes.
- 5) The fifth scene presents the race in a traditional manner. This slide displays a bar chart for each race by divisions and finish status. A table displaying racers, overall place, division place, and finish time is also presented. This scene provides the most robust opportunity for the reader to interact with the data. It can be filtered by year, division, and finish status. Clicking on the racer name launches a racer-specific scene described below.
- 6) This sixth scene is a summary of an individual racer. This charts every race this person has participated in, their finish status and time. A table displays discrete data for each race. A miniature map of the racecourse is also displayed. Clicking on a specific bar or a table row cross-filters the chart and table and annotates the last checkpoint on the race map. For “did not finish” status, this better helps the reader understand how far along the racecourse the racer was when they exited the field.
- 7) Scene seven steps back from the racer and looks more at the course. It is an investigation into where racers exit the field early. These “did not finish” data are charted by checkpoint to help the reader appreciate how far into the race participants are exiting the field.
- 8) If the first scene was an establishing shot and the middle scenes guided the reader through the race from high level to the individual and back to a higher level, this last scene is the resolving action. The slide communicates to the reader the narrative is complete and provides a personal touch to the narration by connecting the author to the event. Navigation buttons to return to the previous slide or start the narrative over.

Annotations

Annotations were templated across the narrative. The primary forms were **callouts**, used to draw focus on specific data. **Consistent shapes and colors** were used in structuring these annotations. In the first scene these annotations draw focus on the checkpoints. Additional examples of annotations include the course record and highest number of participants in the overview scene. In the third scene, annotations are used to plot the average finish time. On the racer data slide, annotations are triggered based on the interaction of the reader:

highlighting the data selected (cross filter) and as highlights on checkpoints on the course map. In this scene the interaction drives the annotation, so the annotations change based on what the user is exploring.

Annotations support the messaging by emphasizing sentinel events in the data. This helps the reader better analyze the information being presented. Simple prose in each scene adds context to the visualization. Axis labels and data point labels, simple forms of annotation, are also employed.

Parameters

The primary parameters used in the visualization are **scene**, **year**, **division**, **finish time**, **finish status**, **checkpoint**, and **racer**. As described in the scene descriptions, the parameters are used to focus each scene to a specific view of the data. The states of the visualization are therefore defined by the parameters and the scene being viewed. A scene may focus on a specific division (“Men’s Solo (paddle)”, for example) in a specific year. Here the state of the narrative is the **scene** being viewed, the **year**, and the **division**. In later scenes the state is dependent on the specific **racer**.

Triggers

The primary trigger is that connects the user action is mouse clicks within each scene of the interactive slide show. **Navigation buttons** are the main trigger for each scene. These buttons follow **standard format** for hyperlinks, providing intuitive use by the reader. When the reader triggers the next scene, the narrative changes states to load that scene with the default parameters. Within a scene, the user can trigger state changes by **clicking on elements** within the scene (e.g., a bar in a stacked bar chart, a row in a table, etc.). This secondary trigger provides the most robust and engaging user interaction and is how the user “drills-down” and explores the data. These triggers may filter the visualization within the scene or may launch a new scene. **Visual cues** are used to communicate that these elements can be interacted with, namely by animating a color change when the mouse hovers over the element. These triggers are implemented via event listeners within the JavaScript code.