Evaluating Transitions for Streaming Big Data

Filipa Castanheira[1], João Moreira[1,*], Daniel Mendes[2], Daniel Gonçalves[1]
[1] *INESC-ID, Instituto Superior Técnico, Universidade de Lisboa*
[2] *INESC TEC, Faculdade de Engenharia, Universidade do Porto*
[*] Email: joaopedro.moreira@inesc-id.pt

***Abstract*—Visualizations for Streaming Big Data convey high volumes of information in real-time, making it challenging for people to grasp significant data changes. One solution could be having visualizations that change themselves according to the incoming data. However, these changes would need to be effectively conveyed. In this work, we propose a set of transitions between different pairs of visual idioms, aiming to aid users in keeping track of the information in real-time and notice relevant changes. We target transitions between Line charts, Heat maps, and Stream graphs. We conceived seven transitions that modify different properties of the visual elements for each pair of visual idioms, following a novel taxonomy for their conceptualization. To assess the performance of the transitions, we conducted an online user study with 100 participants. Results suggest that animations are indeed better to change between different visualization idioms than abrupt transitions. We also suggest transition techniques for each visualization pair, between those proposed, according to participants' preferences. Lastly, we identify which concepts of our taxonomy were more present in our suggested transitions.**

***Index Terms*—information visualization, big data, streaming, vertical transitions, animation, user testing**

I. Introduction

Every second, an enormous amount of information is created. These volumes of data are called Big Data, which can be generated continuously, without interruptions, and streamed. Therefore, it is not possible to download it in advance and prepare the visualization beforehand. Currently, visualization techniques for streaming data are strongly related to the time context of the information [9]. Furthermore, the rate of new information received often exceeds the human perception limits if the data is presented in raw formats because it has too much detail. So, by introducing real-time visual analysis in such areas of application, many benefits arise [3].

With Big Data, analysts require visualizations to explore data more quickly and make essential decisions in time. One of the significant challenges in the visualization of data flows corresponds to the characteristics of the process of physical sedimentation. Data can arise at unpredictable moments, and, in those cases, it needs to be buffered. Also, it needs to be aggregated to show contextualized information over time [5]. Unfortunately, the creation of this type of visualization is not simple. Traditional techniques for visualizing information generally deal with previously known data sets of small dimensions. Thus, they are not adapted to transmit the temporal evolution of complex data sets. Thus, if people want to visualize ever-changing data with large dimensions and varieties, such as Streaming Big Data, they need to explore new strategies. For example, using different visual idioms with different purposes can be a solution. Therefore, having transitions between the various visual idioms becomes something essential. However, there must not be any information loss due to a poorly used transition.

It is also fundamental to consider possible adjustments to the entire visualization when divided into several temporal moments. Therefore, two critical topics must be addressed: horizontal transitions and vertical transitions. The former occurs between different temporal moments, and the latter occurs within the exact temporal moment. Vertical transitions are helpful when the visualization needs to adapt to newly-arrived data with different characteristics. Horizontal transitions are helpful when there is the need to aggregate the data into a different visual idiom. In both, there must be a common thread to avoid loss of information. The use of animations, for example, can aid in presenting the information and helping users in learning and decision-making, thus allowing them to follow the changes more closely. However, if misused, animations have a powerful distracting effect. In this work, we propose several vertical transitions to allow an effective perception of data changes. Each one is composed of one animation between two different visual idioms. To evaluate them, we carried out a user study to test several techniques and strategies for animated transitions. As such, the main contributions of our work are: (1) a conceptualization of a set of transition techniques; and (2) A perceptual study of the proposed transition techniques.

II. Background and Related work

Our work builds on top of previous research in three main areas: big data visualization, streaming big data visualization, and animated transitions between visualizations.

A. Big data visualization

In Big Data, it is essential to balance the available data with its presentation. However, due to the 5V [6] and the data-specific characteristics, there is no generic solution for presenting Big Data. Therefore, it is essential to find solutions that are gradually directed towards the total automation of the analysis process. From the 5V, two stand out in the data complexity: High Variety and Huge Volume. Regarding **High Variety**, ScrAnViz [18] for example, is a software capable of converting different data types. Regarding **Huge Volume**, the

978-1-6654-8343-8/21/$31.00 ©2021 IEEE

Heat Map Scope [4] stands out for its ability to represent a large amount of information using a Heat map and a Stream graph. Dimensional reduction techniques are essential to address both V's. They allow the control of the information, and they avoid context loss [2], [8], [14].

B. Streaming big data visualization

The information available is not always based on static data. Some systems receive it continuously, from multiple sources, in large quantities, and at high speeds. For Big Data Streaming, the most popular data types are the time series [10], [12]. Like traditional big data systems, it is a priority to assess whether the systems have dimensional reduction techniques. However, now the information is received at short intervals. Therefore, there must be an extra focus on visualization techniques. In particular, for time-series visualizations, some alternatives have been proposed. For example, Line charts [20], [21], Bar charts [15], [19] or Scatter plots [15]. In particular, interactivity is a technique often associated with improving and facilitating the understanding of one visualization. It maintains an active association between the user and the visual elements [15], [19]–[21]. However, although some systems handle data streaming, they are still limited to small data sets. VisMillion [15] is one of the most recent contributions in the direction of efficiency when merging Big Data with Data Streaming.

C. Animated transitions between visualizations.

Animations in visualizations are associated with a change in a visual representation over time, applied in various contexts as transitions. Their main objective is to facilitate the perception of such changes and guide a visualization in a particular direction. One of the most popular forms of motion analysis in visualization is the use of Object Tracking. For example, by animating object trajectories, the behavior of specific elements can become more evident and easy to understand. However, zoom [17] and other transitions techniques in visualizations can be an ineffective way for analyzing trends, even when they seem helpful [16]. In animation design, the law of Common Fate, for example, constitutes one of the design principles of visual animation. Chabli et al. [1] applied it for trend analysis in real dynamic visualization scenarios. Also, there are design guidelines to facilitate the identification of a set of data aggregation operations, such as the minimum, mean or median [7]. In these cases, it is essential to focus on illustrating all the world changes. Also, it is important how they can be analyzed without losing context or relevant data and avoiding dealing with sudden or distracting modifications for the user.

III. Transitions between visual idioms

We developed a collection of transitions to be applied between specific combinations of two visual idioms. The goal was to emphasize when data significantly changed while keeping the visualization comprehensible.

A. Visual idioms

We chose three visual idioms, each with a unique color [11]. **Line charts** for the identification of **trends**, allowing the analysis of behavior and information changes in specific periods. **Heat maps** to observe the different **volumes** of data and to analyze its **flow**, and to observe and analyze **patterns** within each specific value range. **Stream graphs** to identify more **general patterns** of all the information such as the **dispersion/variability**, **minimum and maximum values**, and **quartiles and median**, making it possible to observe the dispersion of the data more reliable.

B. Concepts behind the transitions

We created a taxonomy (Fig. 1) as a theoretical starting point to design vertical transitions. It can be summarized into four keywords: fade, shape, cardinality, and position. Then, each concept is applied to the elements that make up a visualization—for example, the line chart's line or the heat map's squares. The fade is applied to the contour or fill of the elements. The shape of the elements is changed via contraction or expansion. The cardinality is changed by merging or dividing elements. Finally, the position of the elements is changed with translations or rotations.

C. Proposed Transitions

We proposed a set of seven transitions between pairs of visual idioms. Two of them were applied in all cases, **No Animation** (**NA**) and **Fade**. **NA** does not apply any animation. Instead, there is a sudden cut between idioms. When using the **Fade**, the first idiom only fades to the second.

1) Line chart to Heat map: The squares on the heat map are formed according to the data density in each time interval on the line chart (Fig. 2). The **Lines** transition extends the line format until the squares are formed. Because a line is an infinite set of points, the **Points** technique divides the line into tiny points until there are enough squares. To agglomerate into squares, the **Squares** transition creates small sections of the line chart that form overlapping squares that move to their place on the heat map. The **Rectangles** transition divides the line into rectangles that later change to squares. Finally, the **Columns** technique changes the rectangles tilt coming out of the line chart to fit the tilt of each square of the heat map.

2) Line chart to Stream graph: The stream graph areas are delimited by five lines representing different statistical values of the data coming from the line chart at each time interval (Fig. 3). One of the options is to expand the initial line's thickness until it reaches the entire width of the stream graph, as the **Expand** technique does. As the expansion takes place, the respective areas arise. Due to the color element's importance for better identifying a visual idiom, this visual element was highlighted to maintain the context. The **Color** technique carries out an Expand but maintains the color of the initial line throughout the transition. As the areas of the stream graph are bypassed via lines, the **Lines** transition causes the initial line chart line to replicate on the five lines needed for the stream graph. For the expansion to be gradual, regarding the

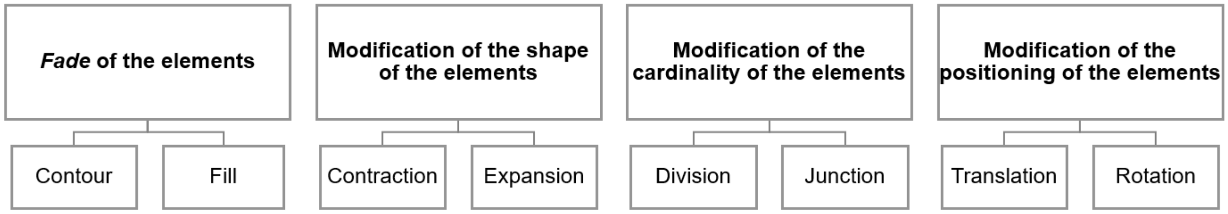
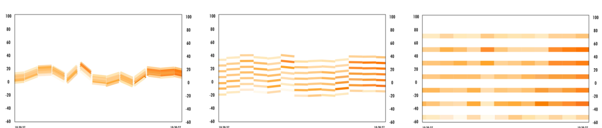


Fig. 1. Taxonomy used for the design of the transitions.

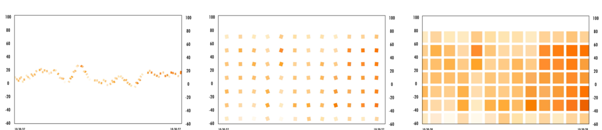
color of the areas, the **Fade Fill** carries out a Lines transition, but with the areas to be filled in Fade. Finally, the **Fade Total** technique applies an Expand transition, but with the entire visual idiom expanding into Fade, lines, and filling areas.

3) Heat map to Line chart: The line chart line is formed from the points that are concentrated in the squares of the heat map in each time interval (Fig. 4). The **Lines** technique makes a contraction of the squares until they reach the thickness of a line to give the user clues as to what the following visual idiom will be. Then, they overlap into a single line. Because a line is an infinite set of points, the **Points** technique reduces the heat map squares until they reach the point size, which will join together and form the line chart. Between the points and squares, the **Rectangles** transition appears. This event reduces the squares to the width of a rectangle before reducing their thickness to a line. The **Squares** technique keeps the squares on the heat map longer until they overlap to form a thick line and get reduced to a line chart line. Finally, the **Columns** applies a Squares transition, but in which the squares still overlap with the vertical alignment of the heat map, and only then rotate and reduce to line thickness.

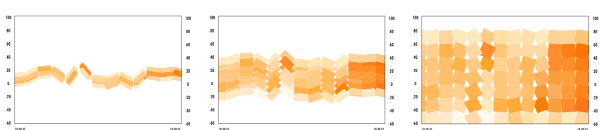
4) Heat map to Stream graph: The stream graph boundary lines represent different statistical values of the data. They are concentrated in the squares of the heat map in each time interval (Fig. 5). The squares may shrink in lines, which will be the five boundaries of the stream graph. With the **Lines** technique, the areas are filled with the respective colors. For the transformation not to be so fast, instead of lines, the areas'



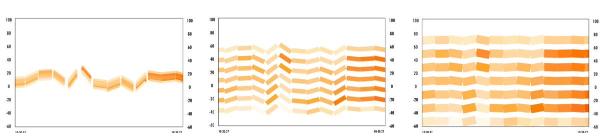
(a) Transition Lines.



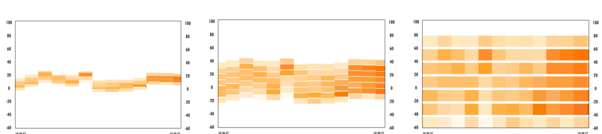
(b) Transition Points.



(c) Transition Squares.

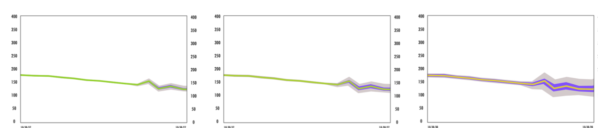


(d) Transition Rectangles.

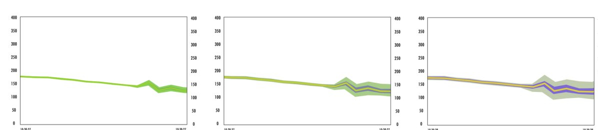


(e) Transition Columns.

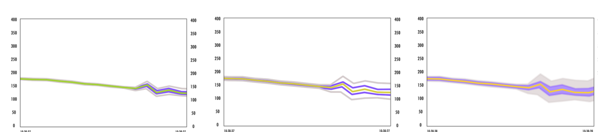
Fig. 2. Evolution of the transitions Line chart - Heat map.



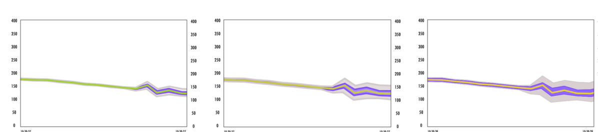
(a) Transition Expand.



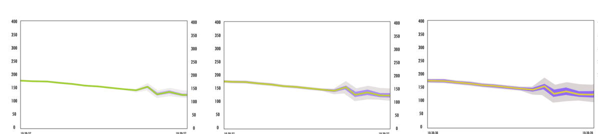
(b) Transition Color.



(c) Transition Lines.

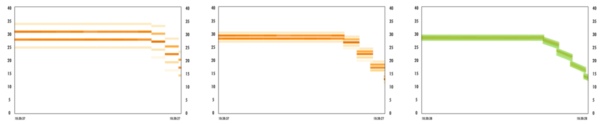


(d) Transition Fade Fill.

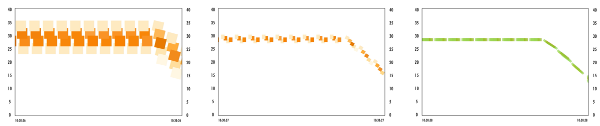


(e) Transition Fade Total.

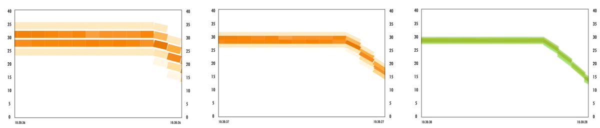
Fig. 3. Evolution of the transitions Line chart - Stream graph.



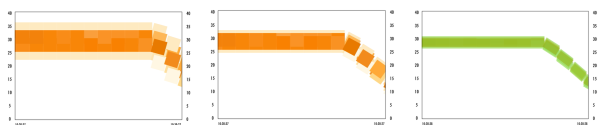
(a) Transition Lines.



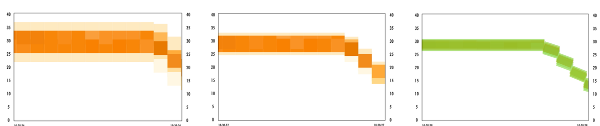
(b) Transition Points.



(c) Transition Rectangles.

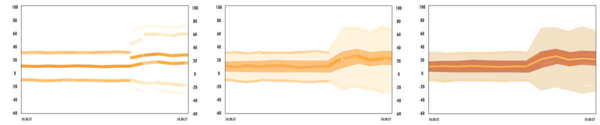


(d) Transition Squares.

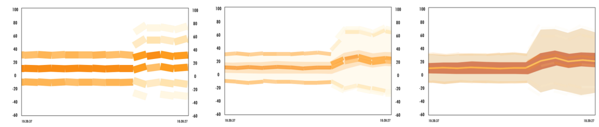


(e) Transition Columns.

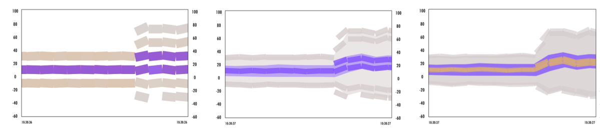
Fig. 4. Evolution of the transitions Heat map - Line chart.



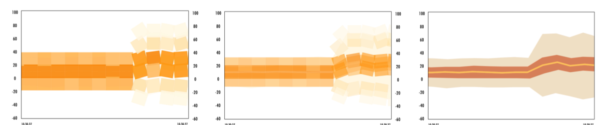
(a) Transition Lines.



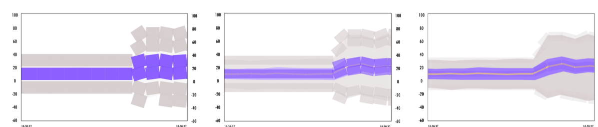
(b) Transition Rectangles.



(c) Transition Rectangles Color.



(d) Transition Squares



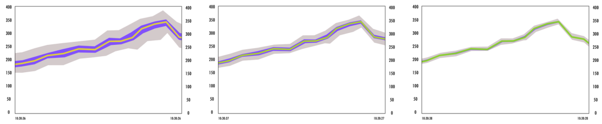
(e) Transition Squares Color.

Fig. 5. Evolution of the transitions Heat map - Stream graph.

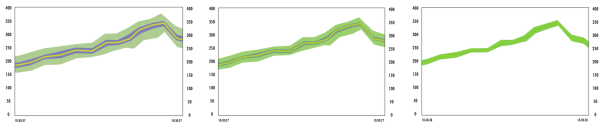
contours can first consist of rectangles, with the **Rectangles** technique. Due to the color's importance for better identifying a visual idiom, we highlighted elements, thus maintaining the context. The **Rectangles Color** technique carries out a Rectangles transition. However, it quickly changes all the elements to the final colors of the stream graph throughout the transition. To keep the heat map present for more time, instead of the squares moving to lines or rectangles, they can be kept as squares, taking advantage of their areas for the stream graph. This technique is applied by the **Squares**, also having the **Squares Color** variant.

5) Stream graph to Line chart: The line chart is formed from the lines that delimit the areas of the stream graph (Fig. 6). One option is to contract the thickness of the initial line until it reaches the line chart's width, as the **Contract** technique does. As the contraction takes place, the respective areas disappear. This visual element was highlighted because of the color element's importance for better identifying a visual idiom. The entire stream graph gradually changes to the line chart color. The **Color** technique carries out a Contract but with the color of the future line throughout the transition. As the stream graph areas are bypassed via lines, the **Lines** transition makes the filling disappear. The contraction is made only in terms of lines that bypass the areas, overlapping and forming the line chart. For a gradual contraction, in terms of the color of the areas, the **Fade Fill** carries out a Lines transition, but with the areas disappearing in Fade. The **Fade Total** applies a Contract transition, but with the entire visual idiom contracting into Fade, lines, and filling areas.

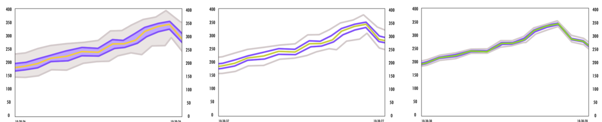
6) Stream graph to Heat map: The heat map squares are formed according to the data density in each time interval of the stream graph (Fig. 7). The stream graph can lose the filling of its areas and leave only its contours, which with the **Lines** technique will then thicken until they form squares. This technique can be applied, but with the emphasis given to color. **Lines Color** immediately changes the outlines for the heat map colors before the lines form squares. For the transformation not to be so fast, instead of lines, the contours of the areas can first consist of rectangles. Hence, it increases the size into rectangles before they reach squares with the **Rectangles** technique. The Stream graph can also expand itself to the area occupied by the heat map without first dividing it into different elements. In the **Expand** transition, the squares are created because of the stain created by the heat map expansion. This technique has also been developed focusing on color, where the spot expands while maintaining the color of the stream graph by applying the **Expand Color** transition.



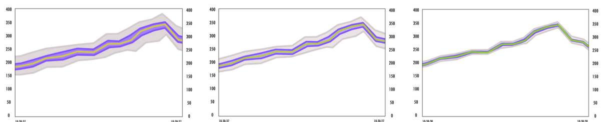
(a) Transition Contract.



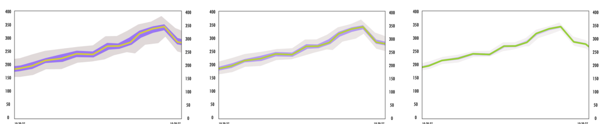
(b) Transition Color.



(c) Transition Lines.



(d) Transition Fade Fill.

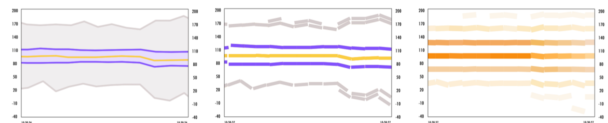


(e) Transition Fade Total.

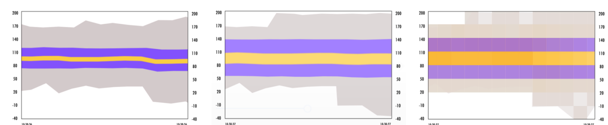
Fig. 6. Evolution of the transitions Stream graph - Line chart.

(a) Transition Lines.

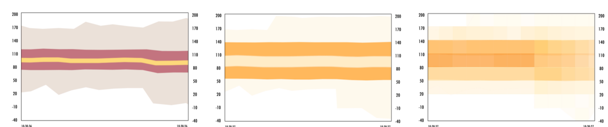
(b) Transition Lines Color.



(c) Transition Rectangles.



(d) Transition Expand.



(e) Transition Expand Color.

Fig. 7. Evolution of the transitions Stream graph - Heat map.

IV. USER EVALUATION

After developing the transition techniques, we carried out a study with real users to understand which ones worked best. The study process consisted of three moments: preparation, implementation, and analysis of results.

A. Method

Our goal was to find which transitions better emphasized data changes while allowing people to understand the data being shown as fast as possible. We developed a set of short videos (using Adobe After Effects), one per transition, and created questionnaires to validate people's comprehension. The questionnaires validated how accurately people could perform specific visual tasks after they watched each transition video.

To prevent an excessively long questionnaire, we prepared seven different versions. Each version contained one transition for each visualization pair that was analyzed in detail (Q1 to Q16 presented in Section IV-C). In each version of the questionnaire, we also asked participants to see the videos of all transitions for a single visualization pair and rank them according to their preference (Q17 in Section IV-C). Since we had seven transitions for each visualization pair and six visualization pairs, one version of the questionnaire did not have this last question. In each version, the techniques were presented following a Latin Square distribution.

B. Data sets

Each transition used a different data set, simulating changes that would justify the transition for a more appropriate visualization idiom. For this purpose, data sets were obtained through a time series generator. Therefore, we ensured that incoming data values increased, decreased, oscillated, or remained with the same initial trend, according to the objective for each transition. In addition, we sought to create drastic changes in the characteristics that each idiom best represented, trying to emphasize its variation.

C. Questions

Regarding the **visualization elements**, in terms of data sets, visual idioms, and the data information, we asked participants two questions. First, if both charts were showing the same data set (**Q1**). Second, if the information transmitted by both charts was the same (**Q2**). The first question would tell us if participants understood that the data changed. The second would tell us if participants thought that each visual idiom conveyed the same information.

Then, for each transition presented, the user had the main task of evaluating the information behavior relative to the data, possible to observe in the displayed visualization. Participants were asked to observe how **information trends** varied such as **mean** (**Q3**), **median** (**Q4**), **dispersion** (**Q5**), **minimum** (**Q6**), **maximum** (**Q7**) and data **flow/volume** (**Q8**), according to a set of given options. Also, there was an "I do not know" option. This option helped us understand if the user could identify the cases in which certain information could not be observed in each visualization and avoid random answers.

To understand which transition was best at **facilitating and helping** in the analysis of the information presented in the various charts, we asked three different questions. (**Q10**): Did the transition help to understand the change from the first to the second chart?; (**Q11**): Did the transition interrupt the analysis that took place in the data in the first chart? Finally, (**Q12**): Is the transition of adequate duration?

Then, we presented three questions to assess the **workload of visualization and analysis** felt by the participant during the questionnaire. First, **Mental demand** (**Q13**): What was the degree of difficulty you felt?. Second, **Time demand** (**Q15**): How do you rate the duration of the transition? Finally, **Overall effort** (**Q14**): How many times have you played the video? Next, for the evaluation of their **individual characteristics**, participants rated the appealing level of each transition (**Q9**). In the end, we asked them to classify each technique globally (**Q16**) and to sort all the techniques for each visual idiom pair according to **preference** (**Q17**).

D. Participants

The questionnaires were answered by **100 participants**, distributed electronically in a balanced way across seven different versions. Among the participants, 39 were male, 61 were female, and their ages ranged from 18 to 26 years old (69%), 27 to 44 years old (7%), and 45 to 62 years old (24%). At least 81% of the participants had a degree. In terms of frequency of analysis of data charts, only 4% said they analyzed every day, while 22% say they did so at least once a week and 27% at least once a month. The rest never saw a chart. In addition, 54 people said they have never analyzed data in real time. The visual idiom most recognized by the participants was the Line chart, with 99%, and the least recognized was the Stream graph, with 47%. 67% of participants recognized the Heat map.

E. Results

We assessed differences between seven groups of an independent variable (transition applied). We tested three dependent variables separately, two with a between-subjects design and one with a within-subjects design. We used the chi-square test of homogeneity in the dichotomous variable (right or wrong, between-subjects, from Q1-Q9). When significant differences were found ($p < .05$), we applied multiple z-tests of two proportions with a Bonferroni correction. We used the Kruskal-Wallis H test in the quantitative variable (task load, between-subjects, from Q10-Q16). We used Dunn's procedure with a Bonferroni adjustment when significant differences were found ($p < .05$). Finally, we used the Friedman test in another quantitative variable (user preference, within-subjects, from Q17). We used Wilcoxon signed-rank tests using a Bonferroni correction when significant differences were found ($p < .05$). The statistically significant differences found are reported in Tables I and II.

1) Line chart to Heat map: For understanding mean variation (**Q3**), the best transition was the Lines. The NA transition was considered worse than Fade, Lines, Rectangles, and Columns. On the other hand, the Fade transition was better classified than Points, Rectangles, and Squares. The Lines transition was more preferred (**Q17**) than Points, which was also less preferred than Rectangles.

Summary: NA was the least preferred techniques. For mean variation's identification and preferences, **Lines** and **Fade** stood out, respectively.

2) Line chart to Stream graph: The Fade, Expand, Color, and Fade Fill transitions were more appealing (**Q9**) than NA. Fade, Color, and Fade Total proved to be more helpful in comprehension (**Q10**) than NA. The duration of transitions Fade, Expand, Color, Fade Fill, and Fade Total were more appropriate (**Q12**) than the duration of the transition NA. Fade, Color, Fade Fill, and Fade Total were globally ranked better (**Q16**) than NA. When all transitions were ordered, NA was less preferred than all other transitions, and Fade was the second-worst. Expand was considered more preferred (**Q17**) than the transition Color.

Summary: NA was significantly the worst in several questions. Considering user preferences, **Fade** also had low classifications, while **Expand** had significantly better results than another three techniques.

3) Heat map to Line chart: All techniques were similar for trend identification. They were also similar for their characteristics and their visualization/analysis workload. The only significant differences occurred in preferences (**Q17**), where **Columns** was ranked the best.

Summary: Columns was the most preferred transition.

4) Heat map to Stream graph: All techniques were similar for the identification of trends. NA's duration was significantly considered not so adequate (**Q15**) than Lines, Rectangles, Rectangles Color, and Squares transitions. When all transitions were ordered by preference (**Q17**), NA was considered less preferred than the Fade, Lines, Rectangles Color, Squares, and Squares Color transitions. The Rectangles transition was less preferred than Fade, Lines, Rectangles Color, Squares, and Squares Color. Also, the Fade transition was more preferred than Lines and Rectangles, and the Lines transition was more preferred than Rectangles.

Summary: NA and **Rectangles** had the worst results. **Fade** was preferred over three alternatives.

5) Stream graph to Line chart: All techniques were similar for the identification of trends (**Q9**). NA was less appealing than all the others transitions. Contract, Lines, Fade Fill, and Fade Total proved to help understanding changes (**Q10**) significantly more than NA. It was also pointed out as having a less adequate duration than all the other transitions (**Q12**). In terms of overall rating, NA was again significantly worse ranked than all the other transitions (**Q16**). All transitions were more preferred (**Q17**) than NA. Contract and Color were more preferred than the Fade and Lines transition.

Summary: **NA** had the overall worst classifications given by the participants. **Contract** and **Color** were preferred over most of the other transitions.

6) Stream graph to Heat map: Fade, Rectangles, Expand and Expand Color were more appealing (**Q9**) than NA. NA presented the less suited duration (**Q15**). When all transitions were ordered, NA was less preferred (**Q17**) than all the others, and Expand was more preferred than the Lines Color.

Summary: Once again, the **NA** was the worst transition. **Expand** was preferred over two techniques.

F. Discussion

We noticed that the NA transitions were the ones that were worst rated in most questions. This result met our expectations, demonstrating that using animation is a valuable tool for presenting and stimulating the understanding of the transition between different data representations in a visualization [13]. Indeed, animation allows users to follow each change more closely and spontaneously can improve their orientation between the data and analyze the different information.

Regarding how participants perceived specific metrics (mean, median, dispersion, among others), only identifying the mean in the pair Line chart-Heat map obtained significantly better results. We thus concluded that our proposed transitions proved to be similar in how they convey information. In any case, identifying the flow/volume and the dispersion of the data were the tasks that had the lowest accuracy. These results may have happened because of the participants' unfamiliarity with some statistical measures. On the other hand, it could also have been because of unfamiliarity with some visual idioms, how they work, and how to extract accurate information.

The use of **Fade** proved to be a good option in some of the visual idioms pairs, namely Line chart - Heat Map and Heat Map - Stream graph. Since both pairs include the Heat Map, the Fade transition may be significant in this visual idiom, probably because it consists of colored squares. Therefore, when squares need to appear or disappear from the screen, it may be easier for them to do so gradually.

In response to the transitions with the best results, we now suggest the most adequate for each pair of visual idioms. In particular, those that proved to be suitable for a more significant number of evaluation situations in each. If the transitions passed the data information the same way, we chose to consider those less demanding in terms of workload and those classified as more preferred. The transitions are:

- From Line chart to Heat map: **Lines** or **Fade**;
- From Line chart to Stream graph: **Expand**;
- From Heat map to Line chart: **Columns**;
- From Heat map to Stream graph: **Fade**;
- From Stream graph to Line chart: **Contract** or **Color**;
- From Stream graph to Heat map: **Expand**.

Lastly, we analyzed which concepts (Fig. 1) were most used by each of the suggested animated transitions. The one that was almost always present (five out of six) was the "Fade-Fill" concept. Although this does not indicate that the others should not be used, it shows that possibly it will be to the users' liking. On the other hand, the least common was the "Fade-Contour" concept. As with the "Fill," this does not mean it cannot be

TABLE I
SUMMARY OF THE STATISTICAL ANALYSIS (PT. 1). ONLY STATISTICALLY SIGNIFICANT RESULTS ARE REPORTED.

	Chi-Square / Kruskal-Wallis / Friedman		Multiple z-test of two proportions / Holm-Bonferroni / Wilcoxon		
	χ^2(6)	p	Pair	Z	p
Line chart to heat map					
Q3	15.239	.018	Lines (64%) - Rectangles (6%)	Count: 9 - 1	
Q17	30.134	<.0005	NA (1, 2.75) - Fade (7, 1)	-3.404	.001
			NA (1, 2.75) - Lines (5, 2.75)	-2.263	.024
			NA (1, 2.75) - Rectangles (4, 2)	-2.013	.044
			NA (1, 2.75) - Columns (5, 3.25)	-2.183	.029
			Fade (7, 1) - Points (2, 1.75)	-3.091	.002
			Fade (7, 1) - Rectangles (4, 2)	-1.990	.047
			Fade (7, 1) - Squares (4, 1.75)	-2.581	.010
			Lines (5, 2.75) - Points (2, 1.75)	-3.143	.002
			Points (2, 1.75) - Rectangles (4, 2)	-2.164	.030
Line chart to stream graph					
Q9	16.137	.013	NA (3, 1.5) - Fade (4.5, 1)	-2.681	.028
			NA (3, 1.5) - Expand (5, 1)	-2.819	.025
			NA (3, 1.5) - Color (5, 0.25)	-3.818	<.0005
			NA (3, 1.5) - Fade Fill (4, 1.25)	-2.058	.040
Q10	14.602	.024	NA (4, 1.5) - Fade (4.5, 1)	-2.583	.040
			NA (4, 1.5) - Color (5, 1)	-3.189	.006
			NA (4, 1.5) - Fade Total (4.5, 1)	-2.769	.030
Q12	12.645	.049	NA (4, 1) - Fade (5, 1)	-2.782	.025
			NA (4, 1) - Expand (5, 1)	-2.408	.032
			NA (4, 1) - Color (5, 1)	-2.998	.028
			NA (4, 1) - Fade Fill (5, 1)	-2.499	.048
			NA (4, 1) - Fade Total (5, 1)	-2.499	.048
Q16	17.181	.009	NA (3, 2) - Fade (4, 1)	-2.897	.020
			NA (3, 2) - Color (5, 1)	-3.786	<.0005
			NA (3, 2) - Fade Fill (4, 1)	-2.705	.021
			NA (3, 2) - Fade Total (4, 1)	-2.897	.020
Q17	38.357	<.0005	NA (1, 0) - Fade (2, 1.25)	-2.590	.010
			NA (1, 0) - Expand (7, 3.25)	-3.108	.002
			NA (1, 0) - Lines (3, 3)	-3.203	.001
			NA (1, 0) - Color (5, 2)	-2.647	.008
			NA (1, 0) - Fade Fill (4, 2)	-2.486	.013
			NA (1, 0) - Fade Total (5, 1.50)	-3.139	.002
			Fade (2, 1.25) - Expand (7, 3.25)	-2.975	.003
			Fade (2, 1.25) - Lines (3, 3)	-2.718	.007
			Fade (2, 1.25) - Fade Fill (4, 2)	-1.976	.048
			Fade (2, 1.25) - Fade Total (5, 1.50)	-2.392	.017
			Expand (7, 3.25) - Color (5, 2)	-1.965	.049
Heat map to line chart					
Q17	37.978	<.0005	NA (1, 0) - Fade (3, 3)	-3.224	.001
			NA (1, 0) - Lines (3, 2.50)	-2.997	.003
			NA (1, 0) - Points (4, 2)	-2.603	.009
			NA (1, 0) - Rectangles (4, 1.50)	-3.114	.002
			NA (1, 0) - Squares (6, 1.50)	-3.238	.001
			NA (1, 0) - Columns (7, 2)	-3.239	.001
			Fade (3, 3) - Columns (7, 2)	-2.937	.003
			Lines (3, 2.50) - Rectangles (4, 1.50)	-2.303	.021
			Lines (3, 2.50) - Columns (7, 2)	-2.574	.010
			Points (4, 2) - Columns (7, 2)	-2.288	.022
			Rectangles (4, 1.50) - Columns (7, 2)	-2.183	.029
			Squares (6, 1.50) - Columns (7, 2)	-2.500	.012
Heat map to stream graph					
Q15	16.509	.011	NA (2,1) - Lines (3, 1)	-2.297	.044
			NA (2,1) - Rectangles (3, 0.25)	-2.898	.020
			NA (2,1) - Rectangles Color (3, 1)	-3.694	<.0005
			NA (2,1) - Squares (3, 0)	-2.857	.016
Q17	35.286	<.0005	NA (1, 0) - Fade (7,4)	-3.502	<.0005
			NA (1, 0) - Lines (4, 2)	-2.007	.045
			NA (1, 0) - Rectangles Color (5,3)	-3.195	.001
			NA (1, 0) - Squares (4, 3)	-2.797	.005
			NA (1, 0) - Squares Color (5, 3)	-2.917	.004
			Fade (7, 4) - Lines (4, 2)	-1.976	.048
			Fade (7, 4) - Rectangles (3, 2)	-2.951	.003
			Lines (4, 2) - Rectangles (3, 2)	-2.177	.029
			Rectangles (3, 2) - Rectangles Color (5, 3)	-2.823	.005
			Rectangles (3, 2) - Squares (4, 3)	-2.352	.019
			Rectangles (3, 2) - Squares Color (5, 3)	-2.926	.003

TABLE II
SUMMARY OF THE STATISTICAL ANALYSIS (PT. 2). ONLY STATISTICALLY SIGNIFICANT RESULTS ARE REPORTED.

	Chi-Square / Kruskal-Wallis / Friedman		Multiple z-test of two proportions / Holm-Bonferroni / Wilcoxon		
	$\chi^2(6)$	p	Pair	Z	p
Stream graph to line chart					
Q9	20.960	.002	NA (3, 2) - Fade (4, 1.25)	-3.001	.009
			NA (3, 2) - Contract (4.5, 1)	-3.673	<.0005
			NA (3, 2) - Lines (4, 1.25)	-2.729	.012
			NA (3, 2) - Color (4, 2)	-2.226	.026
			NA (3, 2) - Fade Fill (5, 1)	-3.945	<.0005
			NA (3, 2) - Fade Total (4, 1)	-3.108	<.0005
Q10	15.171	.019	NA (3, 1.75) - Contract (4, 1.25)	-1.994	.046
			NA (3, 1.75) - Lines (4, 1)	-3.239	.006
			NA (3, 1.75) - Fade Fill (4, 1)	-3.239	.006
			NA (3, 1.75) - Fade Total (4, 1)	-2.797	.020
Q12	23.315	.001	NA (3, 1.75) - Fade (4, 1)	-2.949	.006
			NA (3, 1.75) - Contract (4, 2)	-2.693	.007
			NA (3, 1.75) - Lines (4.5, 1)	-3.320	.004
			NA (3, 1.75) - Color (5, 1.5)	-3.102	.006
			NA (3, 1.75) - Fade Fill (5, 1)	-4.317	<.0005
			NA (3, 1.75) - Fade Total (5, 1)	-3.702	<.0005
Q16	22.322	.001	NA (3, 1.50) - Fade (4, 0.25)	-3.035	.008
			NA (3, 1.50) - Contract (4, 0.25)	-3.035	.008
			NA (3, 1.50) - Lines (4, 0.25)	-3.035	.008
			NA (3, 1.50) - Color (4, 2)	-2.961	.003
			NA (3, 1.50) - Fade Fill (4, 1)	-4.118	<.0005
			NA (3, 1.50) - Fade Total (4, 1)	-3.834	<.0005
Q17	36.612	<.0005	NA (3, 1.50) - Fade (4, 0.25)	-3.439	.001
			NA (3, 1.50) - Contract (4, 0.25)	-3.310	.001
			NA (3, 1.50) - Lines (4, 0.25)	-3.076	.002
			NA (3, 1.50) - Color (4, 2)	-3.007	.003
			NA (3, 1.50) - Fade Fill (4, 1)	-3.075	.002
			NA (3, 1.50) - Fade Total (4, 1)	-3.130	.002
			Fade (4, 0.25) - Contract (4, 0.25)	-2.950	.003
			Fade (4, 0.25) - Color (4, 2)	-2.206	.027
			Contract (4, 0.25) - Lines (4, 0.25)	-2.093	.036
			Lines (4, 0.25) - Color (4, 2)	-1.963	.050
Stream graph to heat map					
Q9	13.893	.031	NA (2, 2) - Fade (4, 1)	-2.863	.020
			NA (2, 2) - Rectangles (4, 2)	-3.379	.006
			NA (2, 2) - Expand (4, 1)	-2.636	.032
			NA (2, 2) - Expand Color (3, 1.75)	-2.019	.043
Q15	17.278	.008	NA (2, 1) - Fade (3, 1.25)	-3.222	.006
			NA (2, 1) - Lines (3, 2)	-3.302	.005
			NA (2, 1) - Lines Color (3, 1.5)	-3.182	.004
			NA (2, 1) - Rectangles (3, 2)	-2.159	.031
			NA (2, 1) - Expand (3, 0)	-3.213	.003
			NA (2, 1) - Expand Color (3, 1)	-2.977	.006
Q17	21.765	.001	NA (1, 1.25) - Fade (4, 4.25)	-2.828	.005
			NA (1, 1.25) - Lines (4, 2.50)	-2.886	.004
			NA (1, 1.25) - Lines Color (3, 2.25)	-2.064	.039
			NA (1, 1.25) - Rectangles (6, 4)	-2.493	.013
			NA (1, 1.25) - Expand (5, 3)	-3.317	.001
			NA (1, 1.25) - Expand Color (5, 1.25)	-2.462	.014
			Lines Color (3, 2.25) - Expand (5, 3)	-2.115	.034

used, but it should be more carefully applied. The remainder concepts were applied equally.

V. CONCLUSIONS AND FUTURE WORK

We proposed transitions between three visual idioms to emphasize significant data changes. To evaluate them, we conducted a user study with 100 participants. As a result, we understood that animation improves people's perception, and we suggest a few transitions according to users' preferences. As future work, we intend to study additional techniques to highlight data changes. We plan to explore machine learning techniques so that a transition is triggered when needed.

ACKNOWLEDGEMENTS

This work was partially supported by FCT through projects PTDC/CCI-CIF/28939/2017, UIDB/50021/2020, SFRH/BD/143496/2019 and POCI-01-0145-FEDER-030740 – PTDC/CCICOM/30740/2017.

REFERENCES

[1] A. Chalbi, J. Ritchie, D. Park, J. Choi, N. Roussel, N. Elmqvist, and F. Chevalier. Common fate for animated transitions in visualization. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):386–396, 2020.

[2] N. Elmqvist, P. Dragicevic, and J. Fekete. Rolling the dice: Multidimensional visual exploration using scatterplot matrix navigation. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1539–1148, 2008.

[3] F. Fischer, F. Mansmann, and D. A. Keim. Real-time visual analytics for event data streams. In *Proceedings of the 27th Annual ACM Symposium on Applied Computing*, SAC '12, page 801–806, New York, NY, USA, 2012. Association for Computing Machinery.

[4] Y. Hashimoto and R. Matsushita. Heat map scope technique for stacked time-series data visualization. In *2012 16th International Conference on Information Visualisation*, pages 270–273, 2012.

[5] S. Huron, R. Vuillemot, and J.-D. Fekete. Visual Sedimentation. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2446–2455, Dec. 2013.

[6] X. Jin, B. Wah, X. Cheng, and Y. Wang. Significance and challenges of big data research. *Big Data Research*, 2, 02 2015.

[7] Y. Kim, M. Correll, and J. Heer. Designing animated transitions to convey aggregate operations. *Comput. Graph. Forum*, 38:541–551, 2019.

[8] H. Kobayashi, K. Misue, and J. Tanaka. Colored mosaic matrix: Visualization technique for high-dimensional data. In *2013 17th International Conference on Information Visualisation*, pages 378–383, 2013.

[9] M. Krstajić and D. Keim. Visualization of streaming data: Observing change and context in information visualization techniques. pages 41–47, 10 2013.

[10] Y. Luo, X. Qin, N. Tang, and G. Li. Deepeye: Towards automatic data visualization. In *2018 IEEE 34th International Conference on Data Engineering (ICDE)*, pages 101–112, 2018.

[11] F. Mansmann, M. Krstajic, F. Fischer, and E. Bertini. Streamsqueeze: a dynamic stream visualization for monitoring of event data. In *Visualization and Data Analysis*, 2012.

[12] P. McLachlan, T. Munzner, E. Koutsofios, and S. North. Liverac: Interactive visual exploration of system management time-series data. pages 1483–1492, 01 2008.

[13] T. Pereira, J. Moreira, D. Mendes, and D. Gonçalves. Evaluating animated transitions between contiguous visualizations for streaming big data. In *2020 IEEE Visualization Conference (VIS)*, pages 161–165, 2020.

[14] V. V. Pham and T. Dang. Mtdes: Multi-dimensional temporal data exploration system. In *2018 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pages 100–101, 2018.

[15] G. Pires, D. Mendes, and D. Gonçalves. Vismillion: A novel interactive visualization technique for real-time big data. In *2019 International Conference on Graphics and Interaction (ICGI)*, pages 86–93, 2019.

[16] G. Robertson, R. Fernandez, D. Fisher, B. Lee, and J. Stasko. Effectiveness of animation in trend visualization. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1325–1332, 2008.

[17] M. Shanmugasundaram and P. Irani. The effect of animated transitions in zooming interfaces. pages 396–399, 01 2008.

[18] K. Sriraghav, S. Jayanthi, N. Vidya, and V. S. Felix Enigo. Scranviz — a tool to scrap, analyze and visualize unstructured-data using attribute-based opinion mining algorithm. In *2017 Innovations in Power and Advanced Computing Technologies (i-PACT)*, pages 1–5, 2017.

[19] L. Stopar, P. Skraba, M. Grobelnik, and D. Mladenic. Streamstory: Exploring multivariate time series on multiple scales. *IEEE Transactions on Visualization and Computer Graphics*, 25(4):1788–1802, 2019.

[20] J. Traub, N. Steenbergen, P. Grulich, T. Rabl, and V. Markl. I2: Interactive real-time visualization for streaming data. In *Advances in Database Technology - EDBT 2017, 20th International Conference on Extending Database Technology - Proceedings, March 21-24, Venice, Italy*, pages 526–529. OpenProceedings.org, 2017.

[21] Y. Wu, Z. Chen, G. Sun, X. Xie, N. Cao, S. Liu, and W. Cui. Streamexplorer: A multi-stage system for visually exploring events in social streams. *IEEE Transactions on Visualization and Computer Graphics*, 24(10):2758–2772, 2018.