Conference contribution :
Tong, C., McNabb, L., Laramee, B., Lyons, J., Walters, A., Thayer, D. & Berridge, D. (2017). *Time-Oriented Cartographic Treemaps for the Visualization of Public Healthcare Data.* Proceedings of the Computer Graphics and Visual Computing (CGVC) Conference 2017, Manchester, UK: The Computer Graphics and Visual Computing (CGVC) Conference 2017.
http://dx.doi.org/https://doi.org/10.2312/cgvc.20171273

EG UK Computer Graphics & Visual Computing (2017)
Tao Ruan Wan and Franck Vidal (Editors)

Time-oriented Cartographic Treemaps for Visualization of Public Healthcare Data

Chao Tong[1], Liam McNabb[1] , Robert S. Laramee[1], Jane Lyons[2], Angharad Walters[2], Damon Berridge[2] and Daniel Thayer[2]

[1] Visual and Interactive Computing Group, Swansea University
[2] Medical School, Swansea University

Abstract

Cartographic treemaps offer a way to explore and present hierarchical multi-variate data that combines the space-efficient advantages of treemaps for the display of hierarchical data together with relative geo-spatial location from maps in the form of a modified cartogram. They offer users a space-efficient overview of the complex, multi-variate data coupled with the relative geo-spatial location to enable and facilitate exploration, analysis, and comparison. In this paper, we introduce time as an additional variate, in order to develop time-oriented cartographic treemaps. We design, implement and compare a range of visual layout options highlighting advantages and disadvantage of each. We apply the method to the study of UK-centric electronic health records data as a case study. We use the results to explore the trends of a range of health diagnoses in each UK healthcare region over multiple years exploiting both static and animated visual designs. We provide several examples and user options to evaluate the performance in exploration, analysis, and comparison. We also report the reaction of domain experts from health science.

1. Introduction

The Cartographic Treemap, combines geo-spatial information, a novel interactive neighborhood preservation metric, and space-filling geometry for the interactive visualization of geo-spatial, and high-dimensional data [TRL*17]. As a hybrid visualization, it combines the advantages of both cartograms and treemaps. Tong et al. implement and demonstrate this visual design with real-world high-dimensional healthcare data collected by the NHS to support clinical commissioning groups (CCGs) and healthcare service providers.

In this paper, we extend cartographic treemaps with time as an additional variate, in order to develop time-oriented cartographic treemaps. Based on a three year time span of healthcare data collecting by the NHS in the England, UK, we present and compare a range of visual design options highlighting advantages and disadvantages of each. We provide several user options to evaluate the performance in exploration, analysis, and comparison based on a given set of prerequisites and user tasks. The contributions of this paper include:

- A new time-oriented cartographic treemap that enables the user to explore hierarchical multi-variate data over a range of years.
- Both static and animated visual designs for cartographic treemaps: presenting the temporal trends of data.
- Interactive user-options that enable users to customize the visual layout.
- The application of our time based visualization to complex, real-world NHS data from England, UK.
- The reaction of domain experts from health science.

In order to achieve this, several challenges must be overcome. The first challenge is to develop several visual designs for incorporating time into cartographic treemaps. A second requirement is to compare the visual designs and present the relative advantages and disadvantages of each. Another is to provide user-options to facilitate both exploration, analysis, and comparison of time-dependent hierarchical, multi-variate UK-based healthcare data. This paper extends the work of Tong et al. [TRL*17] by adding time as a variate.

The rest of the paper is organized as follows. Section 2 presents previous related work on cartograms and treemaps based on time dependent data. Section 3 presents a description of the time dependent UK-based NHS data. Section 4 presents several tasks and requirements for the visual design. Section 5 describes different visual designs and user options in exploration, analysis and comparison of time-dependent hierarchical, multi-variate data in a standalone application. Section 7 reports the reaction from health science domain experts. And the final section presents conclusions and future work within the field.

2. Related Work

Some very helpful survey papers provide an overview of information visualization research in healthcare research [KM02,RWA*13,

Cartograms	Geo-spatial information	Neighborhood Preservation	Multi variate	Hierarchical	Space-filling	time-dependent
Raisz [Rai34]	■	■				
Dorling [Dor11]	■					
Auber et al. [AHL*11]				■		
Tobler [Tob04]	■	■				
Gastner et al. [GN04]	■	■				
Keim et al. [KNP04]	■	■				
Heilman et al. [HKPS04]	■	■			■	
Panse et al. [PSKN06]	■	■				
Van Kreveld et al. [vKS07]	■	■				
Slingsby et al. [SDW09]	■		■	■	■	
Slingsby et al. [SDW10]				■	■	
Alam et al. [AKV*15]	■	■				
Eppstein et al. [EvKSS15]	■				■	
Meulemans et al. [MDS*17]	■	■				
Treemaps						
Shneiderman and Johnson [Shn92]				■	■	
Bruls et al. [BHVW00]				■	■	
Shneiderman [SW01]				■	■	
Itoh et al. [IYIK04]			■	■	■	
Balzer et al. [BDL05]				■	■	
Irani et al. [ISS06]				■	■	
Tu and Shen [TS07]				■	■	
Mansmann et al. [MKN*07]				■	■	
Wood and Dykes [WD08]	■		■	■	■	
Jern et al. [JRA09]		■		■	■	
Slingsby et al. [SDWR10]	AP		■	■	■	
Buchin et al. [BEL*11]	AP	■		■	■	
Wood et al. [WBDS11]	■			■	■	
Wood et al. [WSD11]	■				■	
Duarte et al. [DSF*14]	■	■			■	
Ghoniem et al. [GCB*15]	■	■		■	■	

Table 1: *This table shows characteristics of related work. It includes six visualization properties: geo-spatial information, neighborhood preservation, multi-variate data, hierarchical data, space-filling and time-dependent. Geo-spatial information indicates whether a visualization conveys geographic information and AP in the column represents adjacency preservation only. Neighborhood preservation indicates an algorithm that features a distance metric to preserve neighborhood relationships. Multi-variate indicates the dimensionality of abstract data. Hierarchical indicates a type of hierarchical data. Space-filling indicates how well the output visualization fills the screen. And time-dependent indicates whether a visualization contains time as an attribute. "Older references are provided for thorgoughness." Our time-dependent cartographic treemaps consider all six properties.*

WBH15, ML17]. However, we would like to couple geo-spatial information with healthcare data.

Geo-spatial related work falls into the areas of cartograms and spatially-ordered treemaps. We focus on whether the literature includes time-dependent data and visualizations.

2.1. Cartographic visualization

Cruz et al. [CCM15] define a cartogram as "*a technique for displaying geographic information by resizing a map's regions according to a statistical parameter in a way that still preserves the map's recognizability*". They can display geo-spatial information and another data attribute (such as population or disease prevalence) in one visualization. Tobler [Tob04] and Nursat and Kobourov [NK16] survey general cartograms. They present the development of value-by-area cartogram algorithms and performance in computer science. In general, we don't find time-dependent cartograms in the previous surveys.

Auber [AHL*11] propose a method based on a geographic map metaphor, which facilitates the visualization and navigation through a hierarchy and preserves the order of a hierarchy's nodes.

Gastner et al. [GN04] present a diffusion cartogram for constructing value-by-area cartograms, which provides a valuable tool for the presentation and analysis of geographic data. Keim et al [KNP04] develop a faster algorithm for cartograms. It enables display of dynamic data with cartogram visualizations. These two algorithms are categorised as contiguous area cartograms. Their performance depends on the corresponding value in each area. If the value does not correspond to the area, the cartogram may be difficult to recognize. Also, there is no time variate in their visualization.

Raisz [Rai34] presents the rectangular cartogram, using rectangles instead of real area shapes. Dorling [Dor11] presents the Dorling cartogram which uses circles instead of geographic area shape, similar to the modified cartogram we present. They are categorized as non-continuous area cartograms. They can display statistical information well, regardless of original shape of area, and preserve relative position. However, they are not necessalily space-efficient. Van Kreveld and Speckmann [vKS07] present the first algorithm for rectangular cartograms. They formalize region adjacencies in order to generate layouts that represent the positions of the geographic regions. It converts a rectangular cartogram to a contiguous area cartogram. Our modified cartogram does not fall into the category of continuous cartograms, but resembles a cross between rectangular and Dorling cartograms [NK16]. Our algorithm can be considered as a modified space-filling rectangular cartogram with the addition of a hierarchical structure and multi-variate data.

Heilman et al. [HKPS04] propose a novel visualization technique for geo-spatial datasets that approximates a rectangular partition of the (rectangular) display area into a number of map regions preserving important geo-spatial constraints. They use elongated rectangles to fill the space whereas we use uniform rectangles to fill the space such that regions can easily be compared with one another. Their work focuses on univariate, non-hierarchical and static data.

Panse et al. [PSKN06] combine a cartogram-based layout (global shape) with PixelMaps (local placement), obtaining benefits of both for improved exploration of dense geo-spatial data sets. Their work also focuses on univariate, non-hierarchical without time-dependent data.

Slingsby et al. [SDW09] explore the effects of selecting alternative layouts in hierarchical displays that demonstrate multiple aspects of large multivariate datasets, including spatial and temporal characteristics. They demonstrate how layouts can be related through animated transitions to reduce the cognitive load associated with their reconfiguration, whilst supporting the exploratory process. No metric for neighborhood preservation is described in this work. Also, time is not a variate in this work.

Slingsby et al. [SDW10] present rectangular hierarchical cartograms for mapping socio-economic data. They present a detailed map of 1.52 million GB unit postcodes in their spatial hierarchy, sized by population and coloured by the OAC(Output Area Classifier) category that most closely characterises the population. However, no algorithm for preserving geo-spatial information is provided. No metric for neighborhood preservation is described. Also, evolution over time is not a variate in this work.

Alam et al. [AKV*15] present a set of seven quantitative measures (Average Cartographic Error, Maximum Cartographic Error, Adjacency Error, Angular Orientation Error, Hamming Distance, Average Aspect Ratio, Polygonal Complexity) to evaluate performance of cartograms based on the accuracy of data and its readability. They compare previous cartogram algorithms based on statistical distortion, geography distortion and algorithm complexity and evaluate their performance with respect to different properties.

Eppstein et al. [EvKSS15] introduce a new approach to solve the association challenge for grid maps by formulating it as a point set matching problem. They present algorithms to compute such matchings and perform an experimental comparison that also includes a previous method to compute a grid map. Their work focuses on geo-spatial information and filling space. Multi-variate, hierarchical, time-dependent data are not considered.

Meulemans et al. [MDS*17] design a comprehensive suite of metrics that capture properties of the layout used to arrange the small multiples for comparison (e.g. compactness and alignment) and the preservation of the original data (e.g. distance, topology and shape). Their work focuses on geo-spatial information and neighborhood preservation. Multi-variate, hierarchical data, time-dependent are not considered.

Nursat and Kobourov [NK16] survey cartogram research in visualization and present design guidelines as well as research challenges. They state that mapping multi-variate data is still a challenge in cartogram research. In general, previous cartographic visualizations focus on flat, univariate data. Whereas we process hierarchical, multi-variate and time dependent-data.

We note that visualizing multivariate data is one of the top future research challenges in the latest survey by Nursat and Kobourov [NK16]. Also cartograms, in general, are not space-filling and do not necessarily make the best use of screen space. In addition, time is not a variate in previous cartogram research. See table 1 for an overview. In previous work, Tong et al. [TRL*17] develop a layout algorithm for cartographic treemaps. We extend this to include time-variate data.

2.2. Geo-Spatial Treemaps

Mansmann et al. [MKN*07] present HistoMaps for visual analysis of computer network traffic visualization with a case study showing that a geographic treemap can be used to gain more insight into these large data sets. However, the visualization is essentially univariate (one scalar per level in the hierarchy). It is not adjacency preserving. Also, time is not a variate in the visual layout.

Wood and Dykes [WD08] provide a squarified layout algorithm that exploits the two-dimensional arrangement of treemap nodes more effectively. It is suitable for the arrangement of data with a geographic component and can be used to create tessellated cartograms for geo-visualization. They convert a geographic distribution of French provinces to a spatial treemap layout and preserve the corresponding geo-spatial relationships to some extent. However, they demonstrate that it is impossible to preserve local region adjacencies if nodes are constrained to a standard rectangle parent node. For example, a region map can only have one or two neighbors on a geographic map. Also, time is not a variate in this work. We preserve geo-spatial relationships with less error by allowing gaps in screen space at the different levels of the data hierarchy.

Jern et al. [JRA09] demonstrate and reflect upon the potential synergy between information and geo-visualization. They perform this through the use of a squarified treemap dynamically linked to a choropleth map to facilitate visualization of complex hierarchical social science data. It conveys the neighborhood relationships by using a second view. Also, time is not a focus in this work.

Slingsby et al. [SDWR10] develop an OAC (Output Area Classifier) explorer that can interactively explore and evaluate census variables. There is no inherent information preserving the geo-spatial location of regions because a grid is used to sub-divide space. It is not possible to derive any information about the geography of the UK regions.

Buchin et al. [BEL*11] describe algorithms for transforming a rectangular layout without hierarchical structure, together with a clustering of the rectangles, into a spatial treemap that respects the clustering and also respects to the extent possible the adjacencies of the input layout. The work of Buchin et al. is similar to ours with a few differences. First, they do not demonstrate their layout algorithm on a full geo-spatial map, e.g. the UK. Second, the space-filling requirement results in elongated rectangles that are difficult to compare. Third, the data is univariate.

Wood et al. [WBDS11] present Ballotmaps that use hierarchical spatially arranged graphics to represent two locations (geographical areas and spatial location of their names on the ballot paper) that affect candidates at very different scales. But their work does not contain any neighborhood preservation algorithm.

Wood et al. [WSD11] identify changes in travel behavior over space and time, aid station rebalancing and provide a framework for incorporating travel modeling and simulation by using flow maps. Their work focuses on univariate, non-hierarchical data.

Duarte et al. [DSF*14] propose a novel approach, called a Neighborhood Treemap (Nmap), that employs a slice-and-scale strategy where visual space is successively bisected in the horizontal or vertical directions. The bisections are scaled until one rectangle is defined per data element. Nmap achieves good space-filling visualization that couples related rectangles using a distance metric. However the distance metric is not geo-spatial, it is also neither a treemap of multivariate data nor a hierarchical visualization.

Ghoniem et al. [GCB*15] present a weighted maps algorithm, which is a novel spatially dependent treemap. They present a quantitative evaluation of results and analyze a number of metrics that are used to assess the quality of the resulting layouts. The work of Ghoniem et al. is similar to ours with some important differences. They place emphasis on evaluating adjacency relationships between nodes rather than geo-spatial positions. Requiring 100%

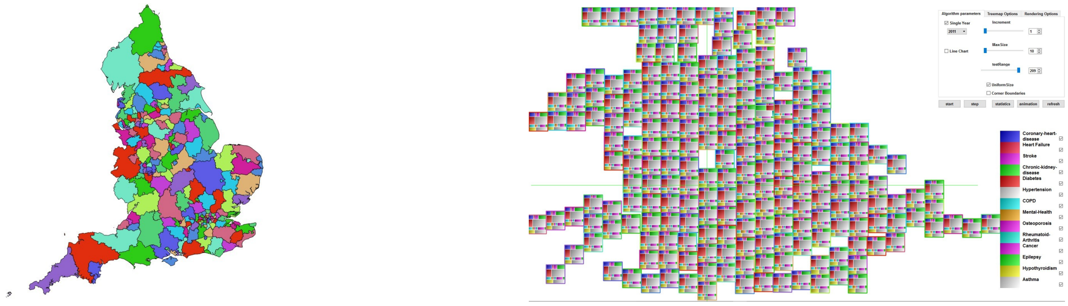

Figure 1: *The left map shows the original 209 CCG regions (Clinical Commissioning Groups) provided by Public Health England [NHS] (left). The original map only occupies 18% of screen space. The original visual design of cartographic treemap based on a single year (right) [TRL*17]. The cartographic treemap occupies 60% of screen space. This color map is from a published color-map from Setlur and Stone [SS16].*

space-filling results in higher geo-spatial error and elongated nodes. Also the data is not multivariate.

The work we present here differs from previous work in that it attempts to combine the space-filling, hierarchical characteristics of ordered space-filling treemaps together with the geo-spatial information conveyed by a cartogram. It add time as a variate into the cartographic treemap. Table 1 compares the current work with the work presented here. No previous algorithm combines all six properties. Especially, no other works contain a time variate. Time-dependent Cartographic Treemaps convey geo-spatial information. They feature an error-driven distance metric between nodes. They visualize multi-variate hierarchical data. They give the user interactive control over how much screen space is used. And they present time-dependent information in several visual designs.

3. Time-Oriented Public healthcare Data Description

We study open NHS healthcare data as a case study for time-oriented cartographic treemap visualization. The UK government collects yearly diagnoses of region specific healthcare data [NHS]. The public health profiles website [NHS] is used for publishing the latest national healthcare data in the England, UK. The data archive is designed to support GPs, clinical commissioning groups (CCGs), and local authorities to ensure that they provide and commission effective and appropriate healthcare services. See Figure 1. Typically this data is displayed using standard line charts, bar charts and pie charts. The standard visualizations do not feature any geo-spatial information. Also, time-related information is generally presented in isolation.

The dataset consists of 14 Excel files of around 10 Megabytes in total together with a CCG region map containing 209 regions (See Figure 1). There are more than 60,000 rows and an average of 100 columns in each file. We extract 14 healthcare disorders over three years prevalence indicators 2011-2013 from the dataset and present the information in our time-oriented cartographic treemap system. "The whole cartogram is resembles a treemap that represents a two-level hierarchy: geographical and various diagnoses in each box."

Our goal is to combine hierarchical, multi-variate healthcare data with complex geo-spatial information using the cartographic treemap algorithm of Tong et al. [TRL*17] and add time-oriented

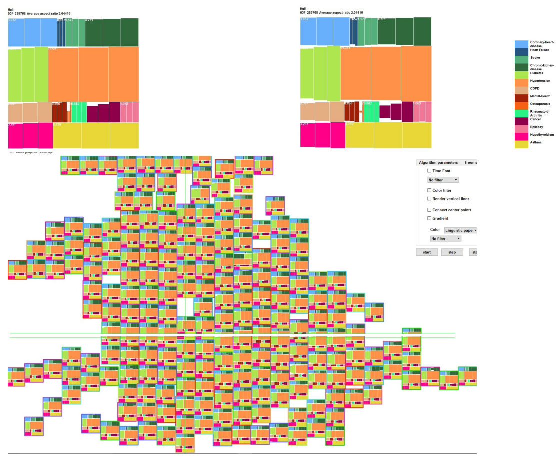

Figure 2: *This visualization shows the output of time-oriented cartographic treemaps with bar charts inside each healthcare variate, and with a details-on-demand window for one region node (top area of main map). It also shows the output of time-oriented cartographic treemaps with symmetric bar charts inside each healthcare variates (bottom half of UK cartogram), and with a details-on-demand window for one region node (top right). The three rectangles in each variates represent prevalence values over three years from 2011 to 2013. We observe that hypertension and diabetes are the most prevalent diagnoses over this time-period. The color map is derived from Colorgorical [GLS17]. Figure 9 in the appendix presents a high resolution version of this image.*

trends in a unified visual design. The challenge is not only to show the overview of hierarchical, multi-variate healthcare data based on regional information, but also depict the temporal evolution trends of data inside each region. We use the NHS healthcare data from 2011 to 2013, and the NHS healthcare regions map as input.

4. Tasks and Requirements

The visual design of our application supports the following requirements and user tasks:

1. **T1**: To provide an overview, both temporal and spatial, of the prevalence rates for each diagnosis coupled with the geography.

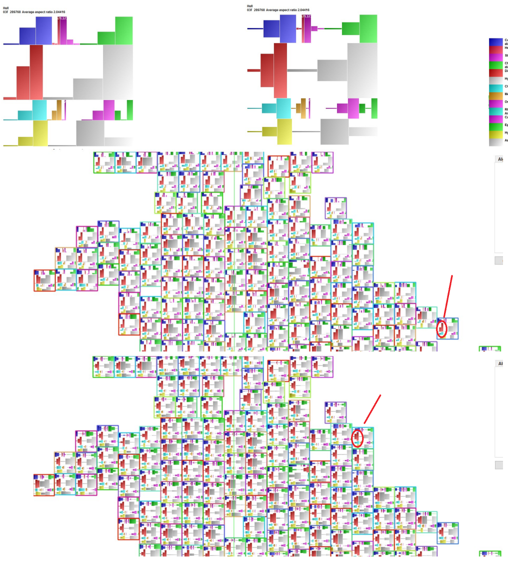

Figure 3: *This visualization shows the output of time-oriented cartographic treemaps with gradient-oriented bar charts (middle), and with a details-on-demand window for one region node (top left). It also shows the output of time-oriented cartographic treemap with the combinations of symmetric bar charts (bottom), and with a details-on-demand window for one region node (top right). Only the northern half of the UK is shown for presentation space purpose. The gradient-oriented bar charts really emphasize the increase in diabites over time. The visual design support task 1 and task 3. Figure 10 in the appendix presents a high resolution version of this image.*

2. **T2**: To provide selection and filtering options with a special focus on time-oriented trends, behaviors and patterns.
3. **T3**: To provide details on demand after exploration, filtering and selection have been performed.

These tasks mirror those outlined by Shneiderman [Shn96] in 1996 and are customized for this particular setting.

5. Time-Oriented Cartographic Treemap

This section describes the visual designs we used to support tasks 1-3 adding a time variate to previous cartographic treemaps. We use the previous cartographic treemap algorithm [TRL*17] for static data as our starting point and then implement several visual designs and user options for displaying time-oriented information in one visual system. The visual designs and user options are presented in the following subsections. First, we introduce time-oriented bar charts, symmetric bar charts, and gradient-oriented bar charts. We compare and discuss the relative advantages and disadvantages of each. Then we add the option of animation, showing increasing versus decreasing diagnoses over time, we describe line charts and other user-options for further exploration including observations based on the visual designs. Finally, we develop an attribute selection option which enables the user to turn individual healthcare variates on or off.

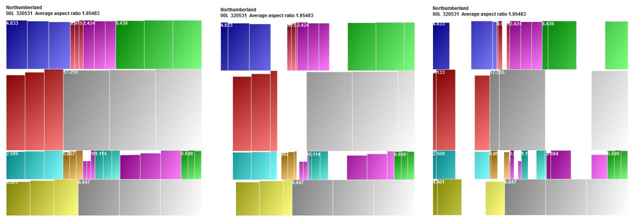

Figure 4: *This visualization shows three frames of the details-on-demond view using animation.*

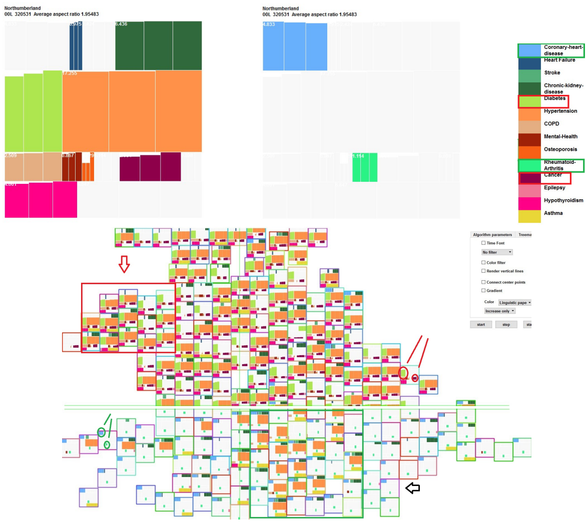

Figure 5: *This visualization shows the output of time-oriented cartographic treemaps with increasing only (top half) and decreasing only (bottom half) prevalance value filters to support task 2. Only the northern half of the UK is displayed for increasing and southern half of the UK is displayed for decreasing values is shown for presentation space purposes. We can observe a region in the noth-east with a group of increasing health dignoses including storks, diabetes, rheumatoid, COPD, ostsoporosis, cancer, and hypothyoidism. Also the London region reports a decrease in hypertension. The color map is derived from Colorgorical [GLS17]. Figure 11 in the appendix presents a high resolution version of this image.*

5.1. Time-Oriented Bar Charts

One of the most obvious choices for mapping time to a visual primitive is using a bar chart. Each column can represent one year and one bar chart can represent the prevalence of each diagnosis. The bar chart is a traditional tool to visualize categorical data. We start off by using bar charts to display yearly data (2011-2013). Each bar chart fits inside the rectangular output of region node and treemap node from Figure 2 well. We integrate each bar chart into a single treemap node for displaying the temporal trend of each healthcare variate inside a single CCG region (See Figure 2). The result addresses task 1 by providing the user with an overview of the data.

The evolution of data over time is difficult to observe using standard bar charts, neither the bars nor the data vary in height very much. To make the difference between each bar more clear, we introduce **symmetric bars** as a modification to the traditional bar chart (See Figure 2). A symmetric bar chart varies the height of each bar from the top while raising the bottom of each bar by the same amount simultaneously. This emphasizes the differences between bars. A details-on-demand window for one region node showing a magnified view of the different style of bar charts is also provided. This supports task 3. By using two styles of bar charts, the time-oriented, hierarchical, multi-variate healthcare data 2011-2013 is presented in single visual design and an overview of yearly healthcare information can be derived from the output. The users can see both an overview of all regions and the details-on-demand for a single region. As we can observe from the result, hypertension is the most prevalent health disorder over the time-span with the largest proportion throughout the UK while the second most prevalent health disorder during the time-period is diabetes. And both are generally increasing over time.

5.1.1. Gradient-Oriented Bar Charts

As the relative difference in height between bars over the three years is small, it is difficult to obtain a clear understanding of temporal trends inside each region from the previous visual design. We introduce a gradient-oriented version of the bar chart as a user option in order to highlight only the *changes* in prevalence rate during three years (See Figure 3). As opposed to the absolute values, in this version, the height of each bar represents the change between minimum and maximum data values. Both the standard and symmetric bar charts can be used to depict the gradient information. The trends of increasing and decreasing diagnosis over time are depicted clearly from the gradient-oriented bar chars. The gradient-oriented bar charts really emphasize the increase of diabetes overtime. This supports task 1. However, with this design too much information is packed into a small area. Distinguishing increasing trends from decreasing trends is difficult. We introduce animation to further clarify the trends.

Symmetric, Gradient-Oriented Bar Charts Symmetric bar charts are also enabled in gradient-oriented user options to further highlight the difference between bars to reflect trends over time (See Figure 3). In this version, the changes in value over the three years are presented with heightened emphasis. An overview of trends for all regions and all healthcare variates can be obtained from this visual design. Because the changes in prevalence rates over time are exaggerated, the user is cautioned when interpreting the graphs. From gradient-oriented bar charts and symmetric gradient-oriented bar charts, the trend is increasing for the majority of healthcare diagnoses. From this visual design, we can observe that for a given CCG, e.g. Hull, all prevalence rates increase over time with the exception of asthma and stroke. This supports both tasks 1 and 3.

5.2. Animation

With bar charts, symmetric bar charts and gradient-oriented user options, the overview of time-oriented healthcare information is presented in various visual designs to support the domain expert user requirements. However, we can add another user option that distinguishes increasing trends from decreasing in the visual design display easily as an approach related to task 2. Thus we introduce an animation option to present increasing trends and decreasing values in different directions. See Figure 4, we animate the bars depicting increasing trends through translation from left-to-right. Decreasing trends are animated by translating the bars from right-to-left. A white gap is inserted between last and first year to ensure the users can decipher where the first bar is. From the animation, the trends of time-oriented values are emphasized even further. This supports task 1. In order to view the animation we encourage the reader to visit the video demonstration at supplementary video URL `https://vimeo.com/223316576`.

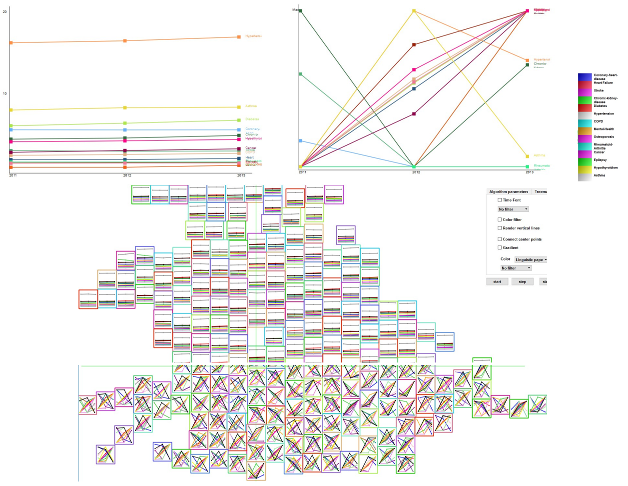

Figure 6: *This visualization shows the output of time-oriented cartographic treemaps with the line charts visual design user option (middle), and with a details-on-demand window for one region node (top left). It also shows the visual design with the gradient-oriented user option (bottom), and with a details-on-demand window for one region node (top right). Only the northern half of the UK and the southern half of the UK is shown for presentation space purposes. Figure 12 in the appendix presents a high resolution version of this image.*

5.3. Filtering and Focus+Context Rendering

Even though we can obtain a direct overview of healthcare diagnosis trends from animation, animation requires video output to be observable. As an alternative, we implement filtering options based on increasing and decreasing prevalence rates combined with focus+context rendering options. Using these options, we can emphasize increasing and decreasing trends in the output visualization and support task 2. See Figure 5 and 8. The user may choose to focus on increasing or decreasing diagnoses over time. Focus attributes are then rendered in color while context rectangles are rendered in grey-scale. And we may observe some useful patterns from the result. For example, most healthcare variates are increasing during 2011-2013, and Coronary-heart-disease is the most decreasing variate among 209 CCG regions except for the mid-east of England. Using animation and increasing and decreasing focus+context rendering user options, we can easily observe that coronary heart disease and rheumatoid-arthritis are the top two decreasing trends among CCG regions and approximately half of the hypertension di-

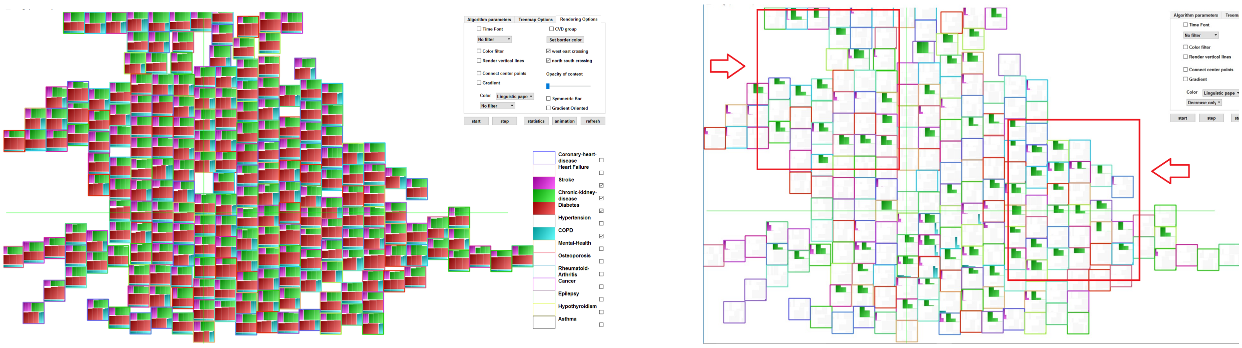

Figure 7: *This visualization shows the attributes selection user option to support task 2 with only four attibutes selected (left) and the decreasing only filter (right). We can observe that kidney disease is decreasing in the north west and the mid east of the UK.*

agnoses are decreasing too. The majority of diagnoses are increasing.

5.4. Line Charts

Bar charts are space-filling by nature and too many bars may crowd the display. Therefore we also experiment with line charts as an alternative visual design. We introduce line charts as a supplementary tool to simplify the time-oriented visualization. They also support task 1. By connecting a series of data points, line charts can present the trends of diagnoses occupying less visual design color and space. We implement line charts inside regions to replace the treemap layout (See Figure 6) as a user option. If we use standard line charts in a similar fashion as standard bar charts, it is difficult to observe trends. This is due to the very gradual change in diagnoses over time. Thus we incorporate a gradient-oriented version of the line chart as well. Gradient-oriented and details-on-demand user options are both provided for the line charts view. The user can filter and observe increasing and decreasing trends of all regions from overview and also focus on the details of a single region. As we can observe from line chart design, the increasing trends dominate diagnoses over time.

5.5. Interactive User-options

For further exploration and analysis, several user options are available, to explore and present the results focusing on different requirements such as choosing individual years and attributes collectively, which support task 2.

Choosing Years To simplify the standard output of the time-oriented cartographic treemaps, choosing individual years enables the user to focus on a single year of information rather than multiple years. The users can extract one year of information from the single year overview and switch between years and observe the differences over time.

The size of treemap nodes can mapped to the population of CCG regions. See Figure 15, choosing an individual year also enables the users to observe the changes to the population in 2011-2013.

Filtering Diagnoses For further simplifying the result and drawing the users attention to the information they require, we implement filtering attributes options. This enables the users to turn on and off specific attributes, and recompute the treemap layout with fewer attributes. In Figure 7, only four attributes are selected with an overview layout and details-on-demand output. The trends of only those four diagnosis in all CCG regions can be focussed on and observed more clearly. Figure 8 shows another important filtering option, depicting increasing or decreasing only prevalence values in a focus + context visual design style.

6. Observations and Discussions

Based on the time-oriented cartographic treemap visualization, several observations can be derived from the public healthcare data.

1. Diabetes and hypertension are the most prevalent diagnoses over 2011-2013, as can be observed in figure 2.
2. Diabetes and cancer are increasing over time in most UK regions. See Figures 3 and 5.
3. Coronary heart disease and Rheumatoid arthritis are decreasing over time in most UK regions. See Figure 5.
4. Kidney disease is decreasing in the north west and the mid east of the UK. See Figure 7.
5. A group of 11 connected CCGs in north west exhibit noticeable increase in Hypertension and diabetes. The CCGs regions are South Sefton, Liverpool, Blackpool, Southport and Formby, Knowsley, Fylde Wyre, St Helens, Halton, Bolton and Warrington. See Figure 5.
6. Hypertension is decreasing in the London area. The relevant CCG regions are Haringey, Islington, Wandsworth, Sutton, Herts Vallys, Richmond, Kingston, Surrey Downs, Brent, Hammersmith Fulham, Hounslow, North West Surrey, Guildford Waverley, Harrow and Ealing. See Figure 5.

7. Domain Experts Feedback from Health Science

The following is feedback directly from collaborators in health science. Time-oriented data, which are variously known as repeated, longitudinal or event history data, present analysts with a range of challenges. These issues become even more challenging when the data also vary spatially. The authors of this paper have developed an eye-catching interactive tool with which data analysts may use animation (please see our later comments) to explore spatial and temporal trends in the values of one or more attributes, as well as to identify salient features such as outliers or extreme values.

We feel that potential users of this tool would require some guidance on using the various facilities, for example, filters to query the data, and exporting the equivalent numerical summaries into table or output format. Advice would also be welcomed on interpreting the visualizations in an efficient and effective manner. For example, the developers of this tool have implemented an algorithm that maximises the use of space by distorting the original shape of the outline of the area under scrutiny. Users will need to be advised on how best to avoid becoming disorientated by this particular feature of the tool. This guidance may need to vary depending on the user group, for example, data analysts compared to clinicians.

We envisage a wide range of possible applications for this tool. The authors of the current paper have used animation to represent time. By using animation, the developers of this visualization tool have injected an element of dynamism into the analytical process, thereby enhancing the exploratory analysis of spatial longitudinal data.

8. Future Work

In this paper, we have focused on presenting graphical summaries of a single categorical attribute (levels of prevalence rate) through the use of arrays of tiles which contain simple bar charts and symmetrical bar charts. Here, we consider various possible extensions.

Adding a longer period of time is future work. Assessing the utility of animation is also future work. A data analyst may wish to examine how prevalence rate varies by age group; in other words, to assess the degree of association between two attributes. A second categorical attribute such as age group could be accommodated readily within this tool by using clustered or stacked bar charts, pyramidal bar charts or heat maps.

Other possible extensions would involve the graphical representation of other types of attribute (e.g. histogram for a continuous measure or score variable) and combinations of different types of attribute (e.g. box and whisker plot to compare the distribution of a continuous measure or score variable between two or more age groups). Other issues that may arise in the analysis of longitudinal data include state dependence and the mover-stayer problem [PB10]. These issues could be explored by displaying a heat map within each tile in order to represent the matrix of transition probabilities at each location on the cartogram.

9. Acknowledgments

Thanks to Richard Roberts, Dylan Rees for proofreading this paper.

References

[AHL*11] Auber D., Huet C., Lambert A., Sallaberry A., Saulnier A., Renoust B.: Geographical Treemaps. *Technical Report* (2011). 2

[AKV*15] Alam M., Kobourov S. G., Veeramoni S., et al.: Quantitative Measures for Cartogram Generation Techniques. In *Computer Graphics Forum* (2015), vol. 34, Wiley Online Library, pp. 351–360. 2, 3

[BDL05] Balzer M., Deussen O., Lewerentz C.: Voronoi Treemaps for the Visualization of Software Metrics. In *Proceedings of the 2005 ACM Symposium on Software Visualization* (2005), ACM, pp. 165–172. 2

[BEL*11] Buchin K., Eppstein D., Löffler M., Nöllenburg M., Silveira R. I.: Adjacency-Preserving Spatial Treemaps. In *Algorithms and Data Structures*. Springer, 2011, pp. 159–170. 2, 3

[BHVW00] Bruls M., Huizing K., van Wijk J. J.: Squarified Treemaps. In *Data Visualization 2000*. Springer, 2000, pp. 33–42. 2

[CCM15] Cruz P., Cruz A., Machado P.: Contiguous Animated Edge-Based Cartograms for Traffic Visualization. *IEEE Computer Graphics and Applications 35*, 5 (2015), 76–83. 2

[Dor11] Dorling D.: Area Cartograms: Their Use and Creation. *The Map Reader: Theories of Mapping Practice and Cartographic Representation* (2011), 252–260. 2

[DSF*14] Duarte F. S., Sikansi F., Fatore F. M., Fadel S. G., Paulovich F. V.: Nmap: A Novel Neighborhood Preservation Space-filling Algorithm. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 2063–2071. 2, 3

[EvKSS15] Eppstein D., van Kreveld M., Speckmann B., Staals F.: Improved Grid Map Layout by Point Set Matching. *International Journal of Computational Geometry & Applications 25*, 02 (2015), 101–122. 2, 3

[GCB*15] Ghoniem M., Cornil M., Broeksema B., Stefas M., Otjacques B.: Weighted Maps: Treemap Visualization of Geolocated Quantitative Data. In *IS&T/SPIE Electronic Imaging* (2015), International Society for Optics and Photonics, pp. 93970G–93970G. 2, 3

[GLS17] Gramazio C. C., Laidlaw D. H., Schloss K. B.: Colorgorical: Creating Discriminable and Preferable Color Palettes for Information Visualization. *IEEE Transactions on Visualization and Computer Graphics 23*, 1 (2017), 521–530. 4, 5, 12, 14

[GN04] Gastner M. T., Newman M. E.: Diffusion-based Method for Producing Density-equalizing Maps. *Proceedings of the National Academy of Sciences of the United States of America 101*, 20 (2004), 7499–7504. 2

[HKPS04] Heilmann R., Keim D. A., Panse C., Sips M.: Recmap: Rectangular Map Approximations. In *IEEE Symposium on Information Visualization, 2004. INFOVIS 2004.* (2004), IEEE, pp. 33–40. 2

[ISS06] Irani P., Slonowsky D., Shajahan P.: Human Perception of Structure in Shaded Space-filling Visualizations. *Information Visualization 5*, 1 (2006), 47–61. 2

[IYIK04] Itoh T., Yamaguchi Y., Ikehata Y., Kajinaga Y.: Hierarchical data visualization using a fast rectangle-packing algorithm. *Visualization and Computer Graphics, IEEE Transactions on 10*, 3 (2004), 302–313. 2

[JRA09] Jern M., Rogstadius J., Astrom T.: Treemaps and Choropleth Maps Applied to Regional Hierarchical Statistical Data. In *IEEE Information Visualisation 2009 Conference* (2009), pp. 403–410. 2, 3

[KM02] Kosara R., Miksch S.: Visualization Methods for Data Analysis and Planning in Medical Applications. *International Journal of Medical Informatics 68*, 1 (2002), 141–153. 1

[KNP04] Keim D. A., North S. C., Panse C.: Cartodraw: A Fast Algorithm for Generating Contiguous Cartograms. *IEEE Transactions on Visualization and Computer Graphics 10*, 1 (2004), 95–110. 2

[MDS*17] Meulemans W., Dykes J., Slingsby A., Turkay C., Wood J.: Small Multiples with Gaps. *IEEE Transactions on Visualization and Computer Graphics 23*, 1 (2017), 381–390. 2, 3

[MKN*07] Mansmann F., Keim D. A., North S. C., Rexroad B., Sheleheda D.: Visual analysis of Network Traffic for Resource Planning, Interactive Monitoring, and Interpretation of Security Threats. *IEEE Transactions on Visualization and Computer Graphics 13*, 6 (2007), 1105–1112. 2, 3

[ML17] McNabb L., Laramee R. S.: Survey of Surveys (SoS) - Mapping The Landscape of Survey Papers in Information Visualization. *Computer Graphics Forum 36*, 3 (2017). 1

[NHS] NHS:. http://fingertips.phe.org.uk/profile/general-practice. 4

[NK16] Nusrat S., Kobourov S.: The State of the Art in Cartograms. *Computer Graphics Forum 35*, 3 (2016), 619–642. 2, 3

[PB10] Penn R., Berridge D.: *Social Statistics-Four-Volume Set*. Sage Publications Ltd, 2010. 8

[PSKN06] Panse C., Sips M., Keim D., North S.: Visualization of Geo-spatial Point Sets via Global Shape Transformation and Local Pixel Placement. *IEEE Transactions on Visualization and Computer Graphics 12*, 5 (2006), 749–756. 2

[QGI] QGIS:. http://www.qgis.org/en/site/. 17

[Rai34] Raisz E.: The Rectangular Statistical Cartogram. *Geographical Review* (1934), 292–296. 2

[RWA*13] Rind A., Wang T. D., Aigner W., Miksch S., Wongsuphasawat K., Plaisant C., Shneiderman B., et al.: Interactive Information Visualization to Explore and Query Electronic Health Records. *Foundations and Trends in Human–Computer Interaction 5*, 3 (2013), 207–298. 1

[SDW09] Slingsby A., Dykes J., Wood J.: Configuring Hierarchical Layouts to Address Research Questions. *IEEE Transactions on Visualization and Computer Graphics 15*, 6 (2009), 977–984. 2

[SDW10] Slingsby A., Dykes J., Wood J.: Rectangular Hierarchical Cartograms for Socio-economic Data. *Journal of Maps 6*, 1 (2010), 330–345. 2

[SDWR10] Slingsby A., Dykes J., Wood J., Radburn R.: OAC Explorer: Interactive Exploration and Comparison of Multivariate Socioeconomic Population Characteristics. *proceedings of GIS Research UK* (2010), 167–174. 2, 3

[Shn92] Shneiderman B.: Tree Visualization with Tree-maps: 2-D Space-Filling Approach. *ACM Transactions on graphics (TOG) 11*, 1 (1992), 92–99. 2

[Shn96] Shneiderman B.: The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *IEEE Symposium on Visual Languages, 1996.* (1996), IEEE, pp. 336–343. 5

[SS16] Setlur V., Stone M. C.: A Linguistic Approach to Categorical Color Assignment for Data Visualization. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 698–707. 4, 18

[SW01] Shneiderman B., Wattenberg M.: Ordered Treemap Layouts. In *Proceedings of the IEEE Symposium on Information Visualization 2001* (2001), vol. 73078. 2

[Tob04] Tobler W.: Thirty Five Years of Computer Cartograms. *ANNALS of the Association of American Geographers 94*, 1 (2004), 58–73. 2

[TRL*17] Tong C., Roberts R., Laramee R. S., Berridge D., Thayer D.: Cartographic Treemaps for the Visualization of Public Health Care Data. *Proceedings of the Conference on Computer Graphics and Visual Computing (CGVC), 2017* (2017). 1, 3, 4, 5

[TS07] Tu Y., Shen H.-W.: Visualizing Changes of Hierarchical Data Using Treemaps. *IEEE Transactions on Visualization and Computer Graphics 13*, 6 (2007), 1286–1293. 2

[vKS07] van Kreveld M., Speckmann B.: On Rectangular Cartograms. *Computational Geometry 37*, 3 (2007), 175–187. 2

[WBDS11] Wood J., Badawood D., Dykes J., Slingsby A.: BallotMaps: Detecting Name Bias in Alphabetically Ordered Ballot Papers. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (2011), 2384–2391. 2, 3

[WBH15] West V. L., Borland D., Hammond W. E.: Innovative Information Visualization of Electronic Health Record Data: A Systematic Review. *Journal of the American Medical Informatics Association 22*, 2 (2015), 330–339. 1

[WD08] Wood J., Dykes J.: Spatially Ordered Treemaps. *IEEE Transactions on Visualization and Computer Graphics 14*, 6 (2008), 1348–1355. 2, 3

[WSD11] Wood J., Slingsby A., Dykes J.: Visualizing the Dynamics of London's Bicycle-hire Scheme. *Cartographica: The International Journal for Geographic Information and Geovisualization 46*, 4 (2011), 239–251. 2, 3

10. Appendix

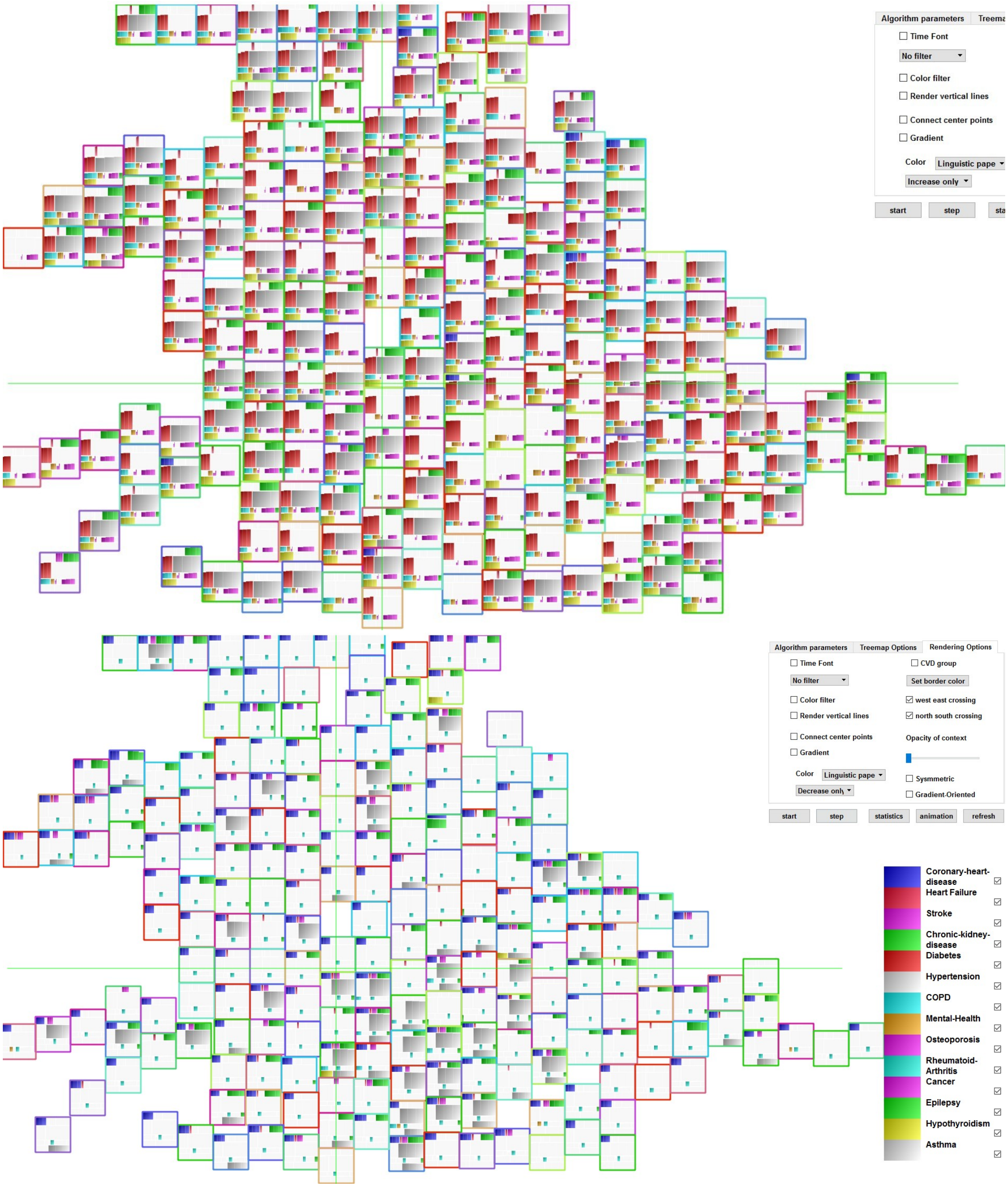

Figure 8: *This visualization shows the output of time-oriented cartographic treemaps with increasing only and decreasing only prevalance values filters. The selection user option is shown in focus, while other attributes are left as context information.*

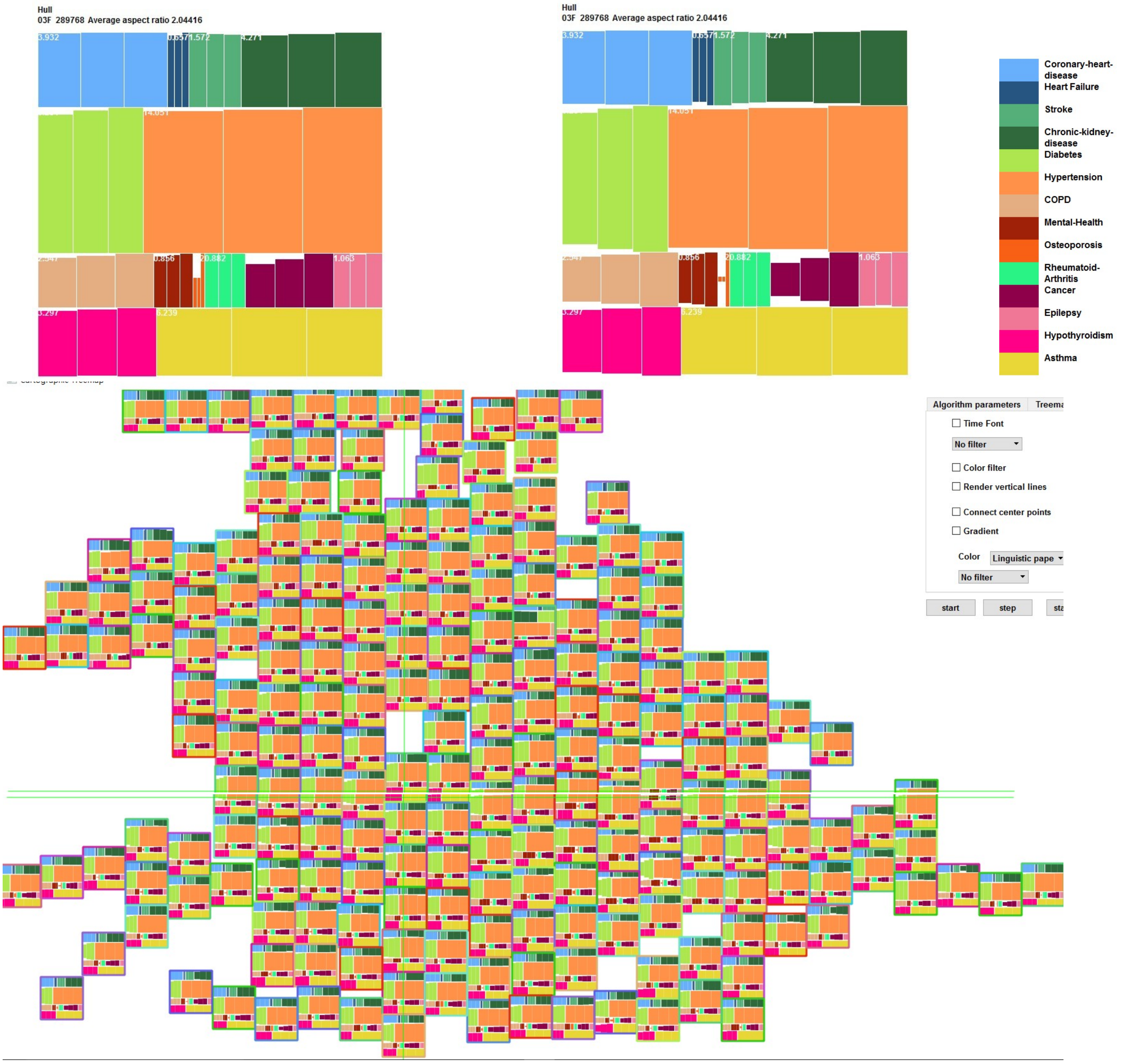

Figure 9: *This visualization shows the output of time-oriented cartographic treemaps with bar charts inside each healthcare variate, and with a details-on-demand window for one region node (top area of main map). It also shows the output of time-oriented cartographic treemaps with sysmmetric bar charts inside each healthcare variates(bottom half of UK cartogram), and with a details-on-demand window for one region node (top right). The three rectangles in each variates represent prevalence values over three years from 2011 to 2013. We observe that hypertension and diabetes are the most prevalent diagnoses over this time-period. The color map is derived from Colorgorical [GLS17]. See also Figure 2.*

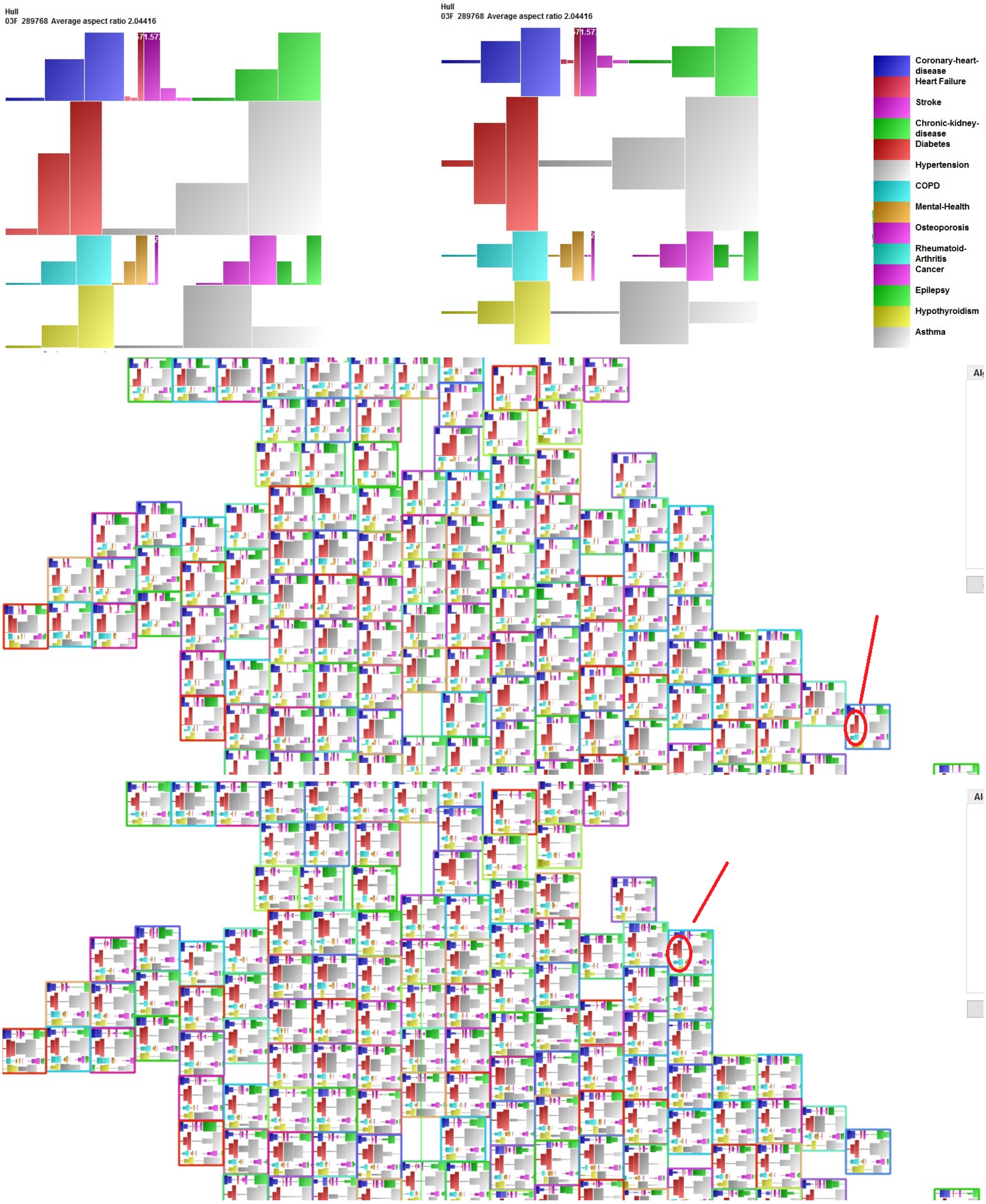

Figure 10: *This visualization shows the output of time-oriented cartographic treemaps with gradient-oriented bar charts (middle), and with a details-on-demand window for one region node (top left). It also shows the output of time-oriented cartographic treemap with the combinations of sysmmetric bar charts (bottom), and with a details-on-demand window for one region node (top right). Only the northern half of the UK is shown for presentation space purpose. The gradient-oriented bar charts really emphasize the increase in diabites over time. The visual design support task 1 and task 3. See also Figure 3.*

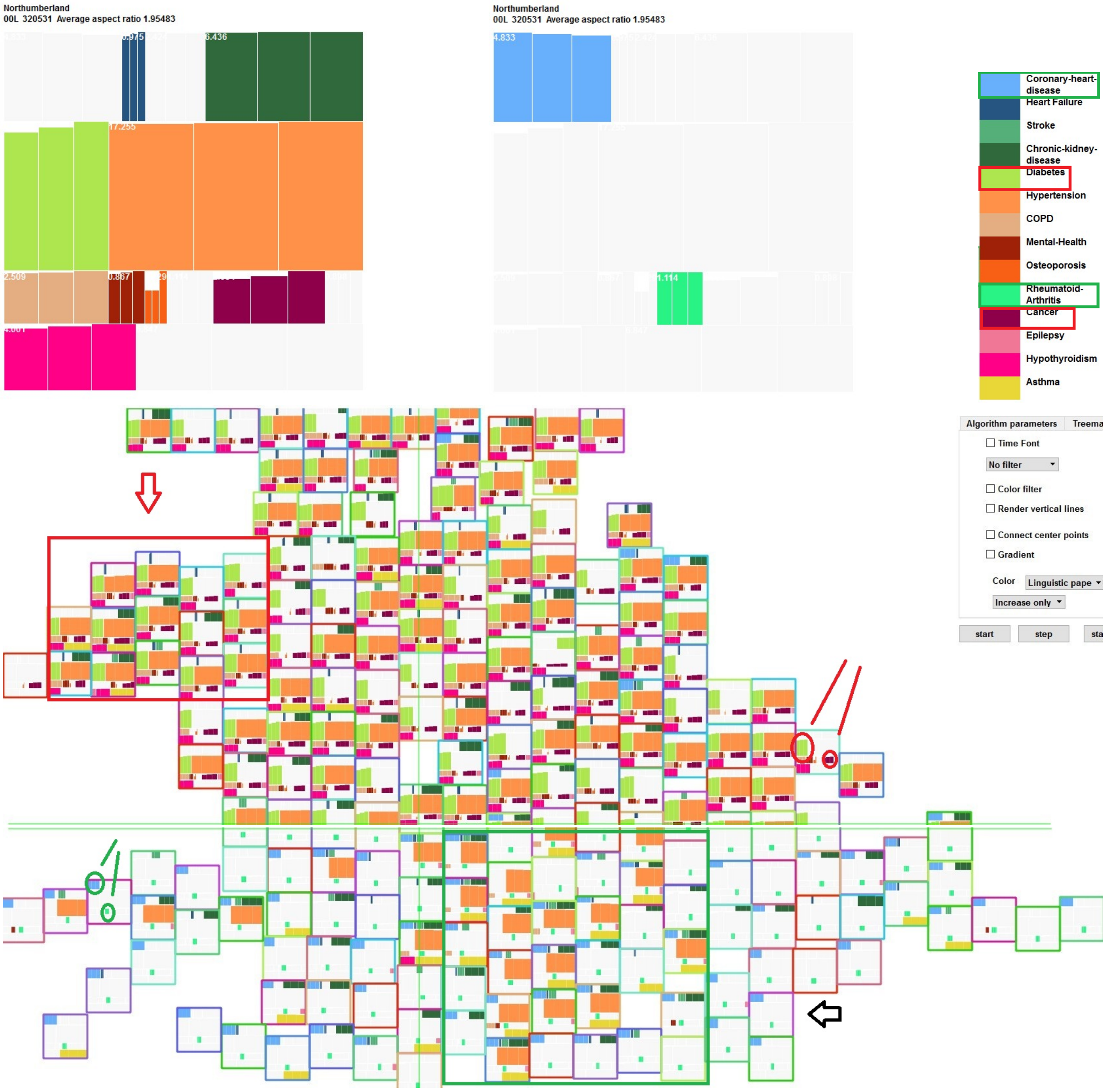

Figure 11: *This visualization shows the output of time-oriented cartographic treemaps with increasing only (top half) and decreasing only (bottom half) prevalance value filters to support task 2. Only the northern half of the UK is displayed for increasing and southern half of the UK is displayed for decreasing values is shown for presentation space purposes. We can observe a region in the noth-east with a group of increasing health dignoses including storks, diabetes, rheumatold, COPD, ostsoporosis, cancer, and hypothyoidism. Also the London region reports a decrease in hypertension. The color map is derived from Colorgorical [GLS17]. See also Figure 5.*

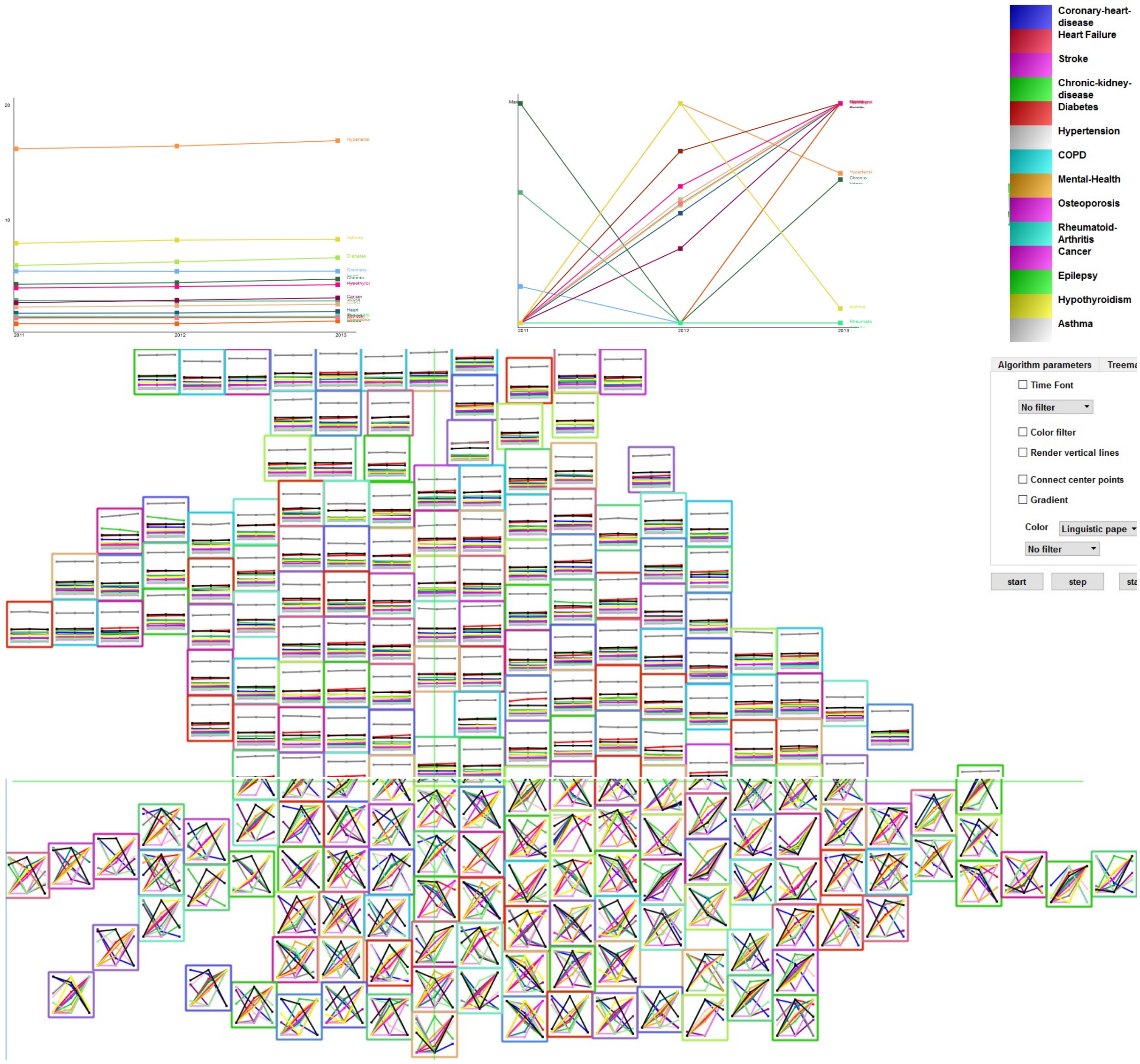

Figure 12: *This visualization shows the output of time-oriented cartographic treemaps with the line charts visual design user option (middle), and with a details-on-demand window for one region node (top left). It also shows the visual design with the gradient-oriented user option (bottom), and with a details-on-demand window for one region node (top right). Only the northern half of the UK and the southern half of the UK is shown for presentation space purposes. See also Figure 6.*

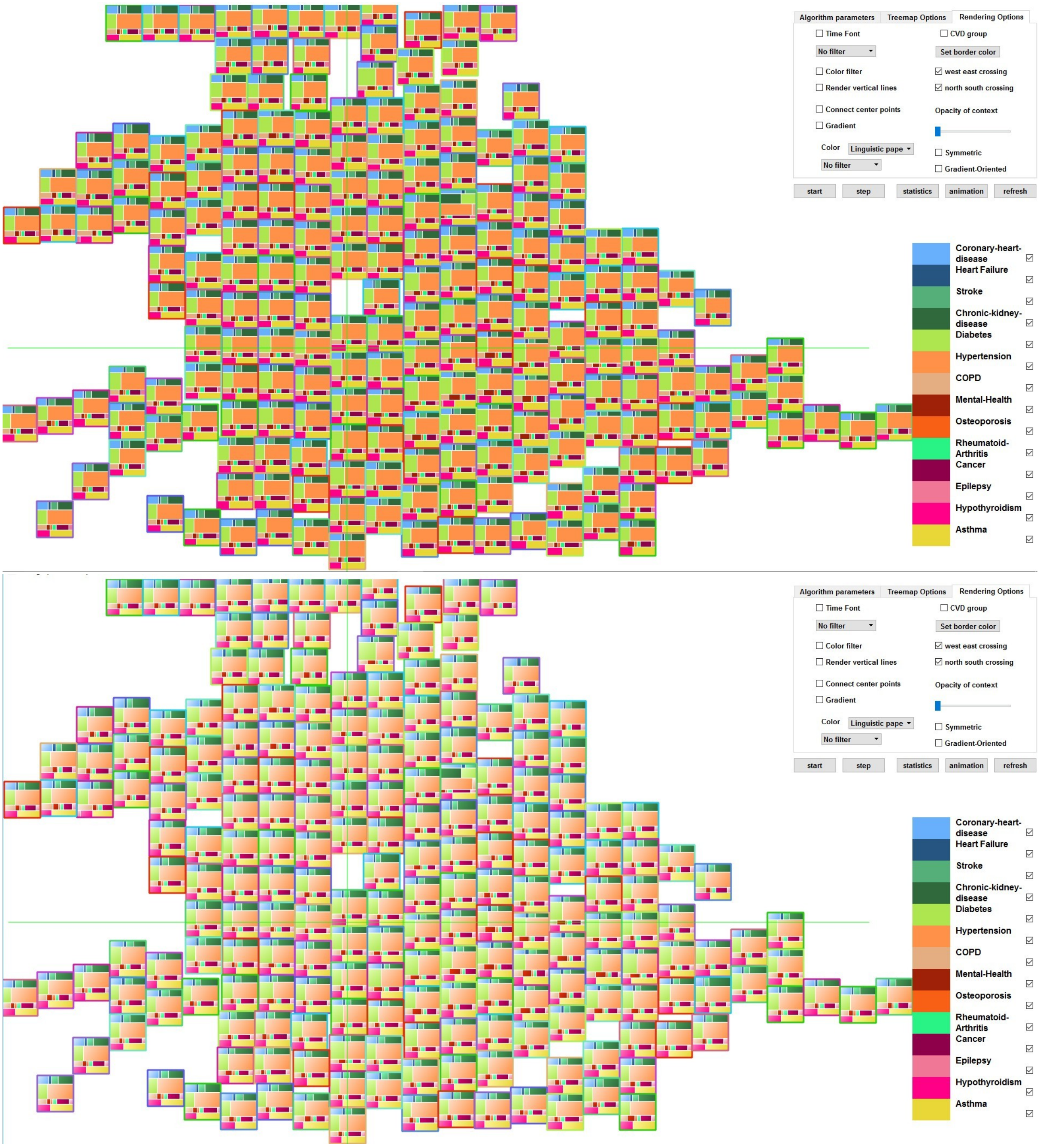

Figure 13: *Optional color map for visual design.*

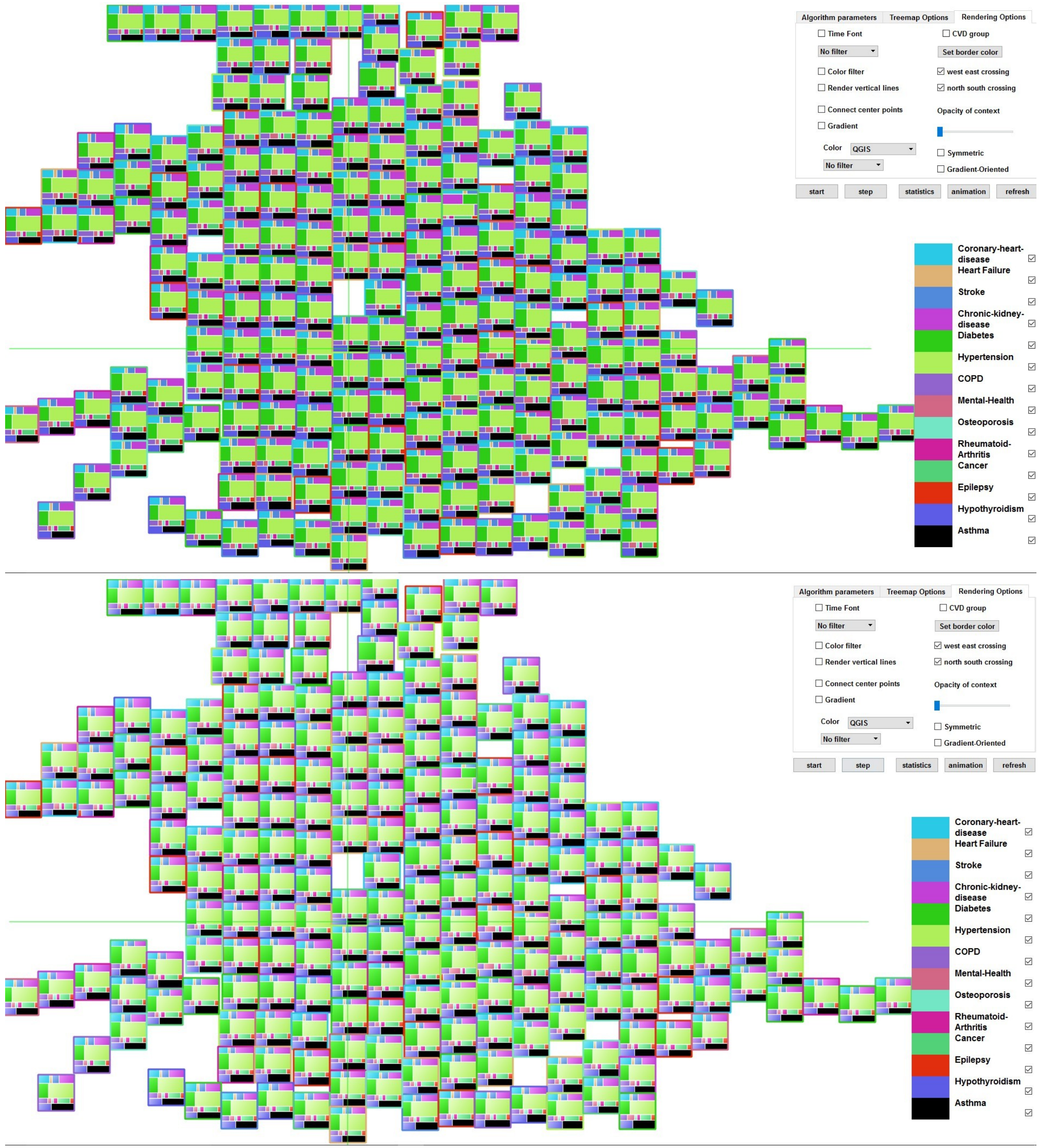

Figure 14: *Optional color map for visual design. This color map is derived from QGIS [QGI].*

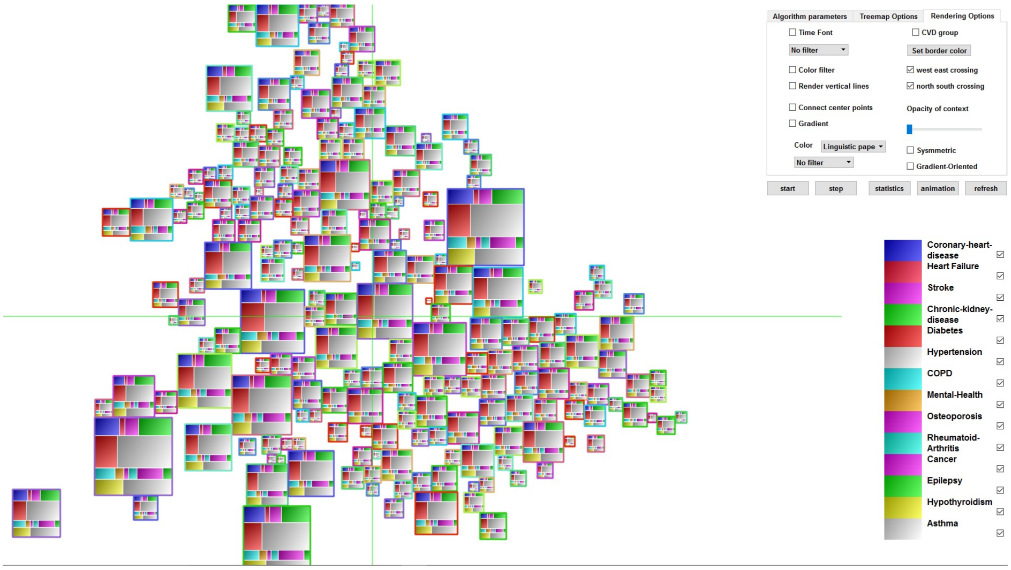

Figure 15: *This graph shows a single year with node size mapped to population. This color map is from a published color-map from Setlur and Stone [SS16].*