

Motivation, information and success of personal Internet shopping in 1997

Jaume Benseny,

jaume.benseny@aalto.fi, student ID 463922

***Abstract.** This visualization presents a comprehensive comparison of motivations, searched information and success rate of shopping activities of Internet users in 1997. The visualization displays each of the mentioned topics in the context of a narrative that explains the shopping behaviour of Internet users in 1997. Each of the stages is made of two diagrams; one shows an overview of how distributed answers are among all users that participated in the survey in a form of treemap. The second one shows what concepts were more important to the users compared with the rest using horizontal bars. In chapter Results we analyse the main findings this visualization helps to discover from the data.*

I. INTRODUCTION

The visualization presented in this article shows some results of a survey carried out by the GVV's WWW Surveying Team from the Georgia Institute of Technology about Internet usage during 1997 in the US. More precisely, it focuses its analysis on a concrete dataset called Internet Shopping in the context of Electronic commerce [1].

A. Internet Shopping dataset

The Internet Shopping dataset gathers multivariate data corresponding to direct answers of the interviewed Internet users classified as professional and personal. Questions belonging to the dataset [2] have as answers a limited subset of possibilities displayed in the following list, no free text was allowed:

1. Yes or No
2. All, Most, Half, Few, None, Not Applicable, (Empty)
3. Never, <1month, 1-3months, 4-10months, Not Applicable, (Empty)
4. <5min, 5-15min, 15-30min, 30-60min, Not Applicable, (Empty)

B. Information displayed by the visualization, the 3 stages

From the Internet Shopping dataset, only personal answers were taking into account which means the information depicted is only about personal Internet shopping. Moreover, answers were also filtered limiting the scope of the visualization to those facts related to: motivation, type of information related to shopping activity and result which define the 3 stages of the visualization. Furthermore, answers used are only from type 1 and 2.

The visualization is composed of 3 different stages, stage 1 the motivation of users, stage 2 type of information related to the shopping activity and stage 3 results of the shopping activity. Each of the different stages contains a treemap and an horizontal bar chart.

II. RATIONALES

Multivariate data is encountered by researchers, engineers, financial managers in numerous situations faced, etc. A plethora of visualization techniques are available for multivariate data: geometric, icon-based, pixel-oriented, hierarchical, graphbased, and hybrid[3].

C. Narrative approach and Gestalt law of similarity

The three stages are placed in ascendant order to be seen from top to bottom building a narration that follows the logic of time and action and reaction. The three stages are perceived as different and independent because of the designed grid layout. **Code row is created by the use of low-level visual channel properties**; in this case, the same colour map is used in all elements in each stage. **This way they appear to be grouped together as described by the Gestalt law of similarity.** [5]

D. Display of proportions and use of tree maps

Treemap chart was used in the visualization to display answers of type 2: All, Most, Half, Few, None, Not Applicable, (Empty). Multi choice questions required of a design that could show accurate distribution of proportions overall answers.

Treemaps display hierarchical data, this visualization implements containment metaphor diagram as layout design. Each branch of the tree is given a rectangle, which is then tiled with smaller rectangles representing sub-branches. [6] **This design was selected because leaf node's rectangle has an area proportional to a specified dimension on the data.** This fact minimizes distortion of the data presented with insignificant values of lie factor.[7] Treemaps become a more effective tool compared to pie charts because perceived area of a circle grows more slowly than the actual measured area. [4]. Leaf nodes are coloured to show improve discrimination of answers.

E. Comparison of proportions and usage of horizontal bars

Horizontal bars chart was used in the visualization to compare answers of type 1: yes or no. Binary answers show the percentage of people that agrees with a statement. **Horizontal bars allow comparison among different levels of agreement regarding questions within the same topic.** Bars are ordered from top to bottom showing first the topics where users agreed more.

Horizontal bars are constructed to maximize data-ink ratio thus axis and grid are not displayed. Width of the bars was designed in order to balance need of easy comparison and user's attention claim in order to minimize redundant data-ink. Percentage is displayed within the bar in white reducing ink, improving perception by increasing contrast. [4]

III. METHODS

Visualization was mainly created with MATLAB. Each of the different stages that contain a treemap and a horizontal bar chart has been created independently with MATLAB from a specific subset of the original Internet Shopping dataset. Final composition, alignment of the three stages as well as titles was performed with GIMP.

A. Stages and colour map

Each stage has its own colour map to guarantee grouping of all elements (charts and titles) based on similarity. Selected MATLAB colour maps: stage 1- winter colour map, stage 2- summer colour map and stage 3- autumn colour map. Colour maps were chosen to guarantee distinctness among stages and contrast with white background.

Due to the fact that MATLAB's colour maps provide uniform colour spaces, RGB colours displayed in the Treemaps rectangle were selected maximizing their distance within the scale to achieve better distinctness within the treemap chart.

B. Tree map library

Treemap was built using MATLAB code of Joe Hicklin. It's available in MATLAB CENTRAL site under BSD License.[8]

IV. RESULTS

By looking at the visualization (see next page) we can easily conclude that regarding motivation, we can see that the majority of searches were intentional but not as much as we could expect. Moreover, we can add that the main reasons triggering Internet shopping were "Not pressure", "Save time" and "Convenience".

Regarding type of information searched, we can see that users were not visiting websites of stores. The more interesting type of information was "Details" about the product to acquire. Regarding success rate, we can see that very few were always successful but the majority of most of the time found what they were looking for. Main reasons for dissatisfying experiences were caused by slow browsing.

V. ENCOUNTERED DIFFICULTIES AND ESTIMATED TIME

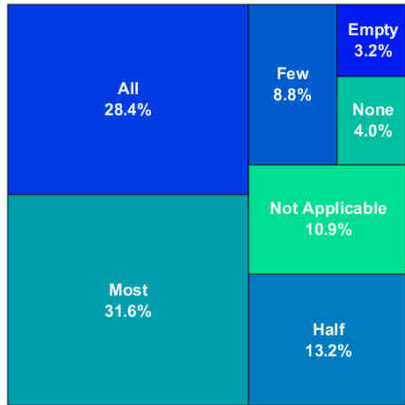
Main difficulties encountered during the execution of the assignment were:

- i) Dimensionality reduction was tried with MATLAB but as a result of working with binary variables made the calculation of distances unviable.
- ii) Due to the fact that Internet Shopping dataset is the result of a user survey, data is not continuous and therefore comparisons should be based on proportions. Analysis on possible diagrams to show proportions was performed.
- iii) Regarding the treemap, we experienced challenges when names needed to be displayed in small rectangles.
- iv) When applying colour map in treemap, we had to make sure that colours of rectangles were distinct enough.

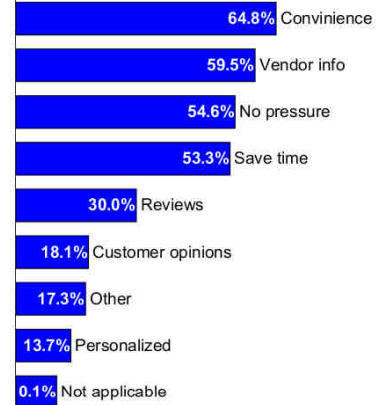
Estimated total time for the assignment was aprox. 25h.

Motivation, information and success of personal Internet shopping in 1997

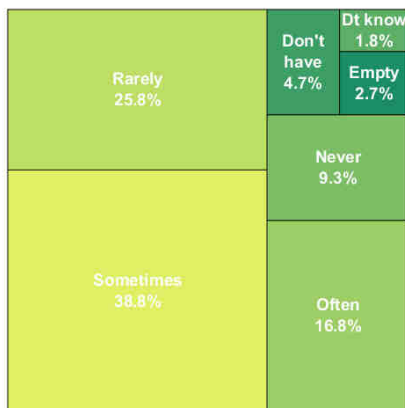
Intentional searches?



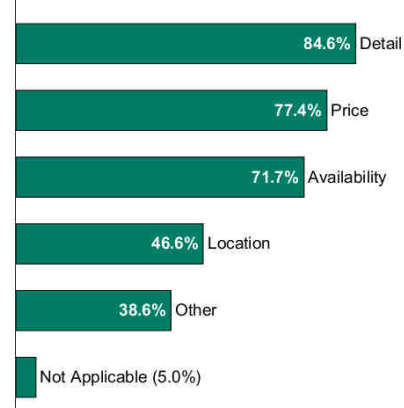
Reasons for using the web



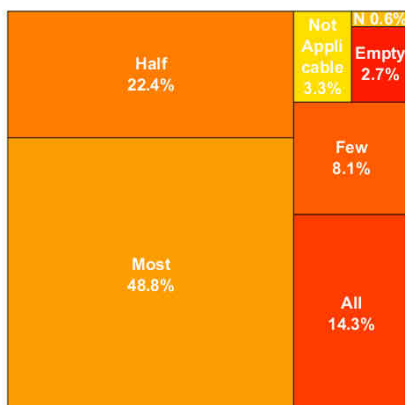
Visiting websites of stores?



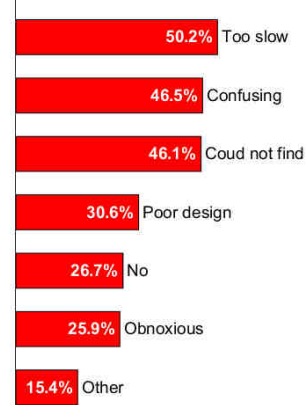
Type of information searched



Success Rate



Reasons for dissatisfying experiences



REFERENCES

- [1] Internet Shopping dataset in the context of Electronic commerce
http://www.cc.gatech.edu/gvu/user_surveys/survey-1997-10/datasets/
- [2] Internet Shopping bullet list http://www.cc.gatech.edu/gvu/user_surveys/survey-1997-10/bulleted/shopping_bullets.html
- [3] A survey on information visualization: recent advances and challenges. Shixia Liu, Weiwei Cui, Yingcai Wu, Mengchen Liu. Springer-Verlag Berlin Heidelberg 2014
- [4] The visual display of quantitative information. Edward R. Tufte. Cheshire (CN) : Graphics Press, 2001.
- [5] Information visualization: perception for design. Colin Ware. Boston : Morgan Kaufmann, 2012.
- [6] Visualizing Business Data with Generalized Treemaps. Roel Vliegen, Jarke J. van Wijk, Erik van der Linden IEEE transactions on visualization and computer graphics vol12.
- [7] USER CONTROLS FOR IMPROVED DATA VISUALISATION: A COMBINATION OF DISCIPLINES APPROACH. Peter O'Conner. BSc (Hons) Computer Science – 2013/14
- [8] Treemap code. MATLAB CENTRAL.
<http://www.mathworks.com/matlabcentral/fileexchange/17192-treemap>