Analyzing Student Travel Patterns With Augmented Data Visualizations

Carl Skelton
OCAD University
Toronto ON M5T 1W1, CA
cskelton@faculty.ocadu.ca

Manpreet Kaur Juneja
OCAD University
Toronto ON M5T 1W1, CA
jmanpreet@faculty.ocadu.ca

Cody Dunne
Northeastern University
Boston MA 02115, USA
c.dunne@northeastern.edu

Jeremy Bowes
OCAD University
Toronto ON M5T 1W1, CA
jbowes@faculty.ocadu.ca

Steven Szigeti
OCAD University
Toronto ON M5T 1W1, CA
sszigeti@ocadu.ca

Minsheng Zheng
OCAD University
Toronto ON M5T 1W1, CA
mzheng@faculty.ocadu.ca

Marcus Gordon
OCAD University
Toronto ON M5T 1W1, CA
mgordon@faculty.ocadu.ca

Sara Diamond
OCAD University
Toronto ON M5T 1W1, CA
sdiamond@ocadu.ca

DIS'17 Companion, *June 10-14, 2017, Edinburgh, United Kingdom*

ACM ISBN 978-1-4503-4991-8/17/06.
http://dx.doi.org/10.1145/3064857.3079140

Abstract

Visualization and visual analytics tools can provide critical support for experts and stakeholders to understand transportation flows and related human activities. Correlating and representing quantitative data with data from human actors can provide explanations for patterns and anomalies. We conducted research to compare and contrast the capabilities of several tools available for visualization and decision support as a part of an integrated urban informatics and visualization research project that develops tools for transportation planning and decision making. For this research we used the data collected by the StudentMoveTO (Toronto) survey which was conducted in the fall of 2015 by Toronto's four universities with the goal of collecting detailed data to understand travel behaviour and its effect on the daily routines of the students. This paper discusses the usefulness of new software which can allow designers to build meaningful narratives integrating 3D representations to assist in Geo-spatial analysis of the data.

Author Keywords

Data visualization; urban systems; transportation.

ACM Classification Keywords

H.5.m. Information interfaces: Miscellaneous

StoryFacets (2) is a data analytics dashboard, designed to provide data analysis and communication through a combination of interactive infographics that can be arranged in narrative sequences. It is an exploratory data analysis and presentation tool that integrates general-purpose data loading, provenance analysis, qualitative markdown contents, an interactive infographics dashboard, and dynamic presentation options. Currently it provides three types of visualizations, bar chart, pie chart, and gather plot (a variation of scatterplot).

StudentMoveTO

The first iteration of StudentMoveTO was an online single-day travel survey of students, conducted by the four Toronto universities in the fall of 2015. (1)
Email invitations to participate in the survey were sent to student email accounts of the participating institutions (approximately 184,000 invitations). Responses were received from 10,203 female and 4,891 male students about home location; factors influencing home location choice; scheduling and travel related to work, study, and other daily activities. (1)

The infographics provided in the initial StudentMoveTO report (1) were informative, but limited to a presentation of standard bar and pie charts, graphs, and color coded pattern maps. Static representation of patterns, and statistical information was that it provided minimal insight into the changing nature of the patterns and relied heavily on text and activity lists.

Using Data Visualization

The research team focused on visualizing data from OCAD U student answers, in part in collaboration with an expert user team of administrators, and comparing it to other universities' data. The team conducted an analysis of the survey data using two different exploratory data visualization software (StoryFacets and Betaville's WebGL) applications. (2). The research team is working with StoryFacets and the new webGL version of Betaville as rapid prototyping platforms, to build and test visualization types and hybrids (3). We describe both tools below, including their analysis of the StudentMoveTO data specific to OCAD U. The data analysis undertaken with both platforms was framed by a key question: ***What data visualization types best represented the salient information?***

StoryFacets and Betaville

Using StoryFacets, the research team filtered and presented the survey data in order to address key questions about OCAD U's distinctive profile as a "commuter school." We were able to undertake these tasks within the research group and then add queries as observations and questions emerged during discussions with a user group of administrators. A strength of the tool is its rapid response to queries. For example, of students who commute by public transit, how many choose to "bunch" their commutes, and hence their courses into three or four days a week to deal with long travel times?(1)

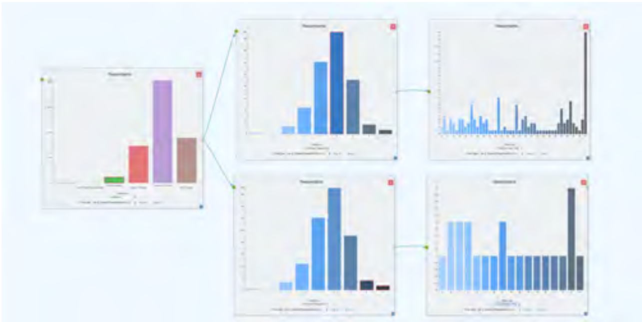

Fig.1. Of students who live 1-30 min. away, 25% come to campus 6-7 times/week compared to 11% who live 31-60 min. away and only 6% of those over an hour away. Likewise, of students who live an hour away, 45% come to campus only 4 times/week vs. 26% of students 1-30 min. away.

Data analysis with StoryFacets revealed that as commute time increases, the number of OCAD U students who come to campus five to seven times per week decreases [Fig. 1]. StoryFacets provided an accessible and responsive web-based interface with an intuitive work flow to quickly bring together groups and sequences of graphics, to use storytelling as a medium to extrapolate trends and connect data, identify gaps in the data collection, or fill in the blanks by using specific queries. StoryFacets also retains the history of dataset exploration as linked charts, facilitating overview and recall for advanced analysts.

Betaville (2) [Figs. 2,3] is an interactive 3D “fly-through” geo-visualization tool that provides an efficient way to procedurally spatialize the initial tabular data outputs in a web browser. It relates the numbers intuitively to real-world commuting experience for those users with a background in urban geography of the Greater Toronto Area, indicating the students most affected by the commutes, and the institutional and city agency decision-makers most likely to be in a position to use the new knowledge to inform plans for change.

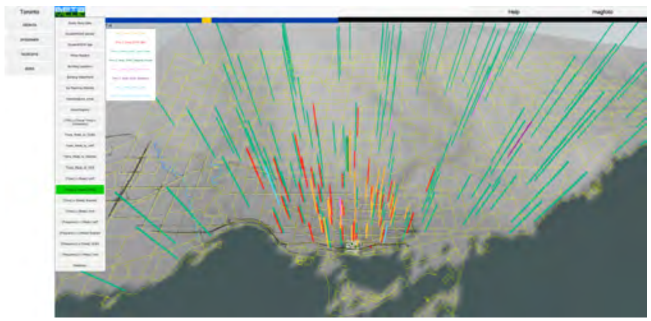

Fig. 2: 2D map with interactive 3D infographics representing StudentMoveTO data generated using Betaville

Betaville's interactive 3D spatial navigation of vertical bar graphs as an overlay on a map of the city provided for improved legibility and pattern recognition for expert and non-expert users alike: the very different patterns of travel distance and mode choice between the participating schools could make it immediately clear that the challenges are very different from institution to institution. The aggregate data reveals pattern trends at the municipal level, where the information is relevant to zoning and land use near the schools, and to transit planning across the metropolitan region. Blending 2D cartographic representation and interactive 3D infographics captured the information value of both types, without distorting either. [Fig. 2].

Fig. 3: Each 3D column visualization is coupled with a bar chart, representing the distribution of respondents by institutions with “OCAD University” selected as the subset. This representation illustrates Transit modes, Commute frequency and Commute time for OCAD University students. The bar charts generated in Story Facets give aggregate quantitative information, while the 3d visualization gives a corresponding spatial understanding, therefore enhancing the qualitative dimension of the data.

The greater value of the Betaville visualization of the StudentMoveTO data, though, is contextual and qualitative: the tight clustering of pedestrian and bicycle commuter home locations relative to the distribution of public transit users is straightforward, but the representation’s correlation of data point distribution with the real city layout adds a qualitative understanding, capturing real-world implications of the underlying information. The relatively tight 3 km and 6 km circles within which active transport options are practical imply very different daily student experience options according to the particular patterns of adjacent land use, and very different strategy options for institutions located in areas with distinct (and/or changing) patterns of housing, service, and employment availability (4, 5, 6).

While **StoryFacets** gave the analyst capability to build a narrative by connecting snippets of data representations together in a sequence, which then helped in seeing larger patterns and identifying relationships through permutations and combinations of data, plugging in **Betaville**, a 3D visualization software to StoryFacets, provided an added layer of clarity and narrative capability. Using relative and absolute spatial distribution representations provided stakeholders and decision-makers with a clear understanding of the different commute patterns. The representation revealed the patterns of clustering and the relative scales of the ranges of the different modes of commuting, but also related the numbers to their meaning: different lived experiences of commuting. [Fig.3]

Conclusions

The integrated use of StoryFacets and Betaville led to findings both at the narrow scale of a particular academic institution and then related these findings to the broader scale of the larger metropolitan area. At the larger metropolitan scale of the GTA and beyond, the visualization of the complexity of student data and travel patterns provided information and support for decisions around the tradeoffs that students make around housing location and transportation. The research suggests that the combinatory effects of a 2D system that allows fast queries and the recording of provenance with a 3D system which specializes results supports a depth of analysis than either conventional or "heat map" infographics. StoryFacets and Betaville can support collaboration between domain experts and analysts through sharing specific insights in the format of users' comments and qualitative content to further our understanding about the transportation and commuting behaviors. The portability and cross-platform compatibility of both StoryFacets and Betaville allows for distributed collaboration.

Further Research

Technology: Further research in employing machine-learning algorithms to enhance research productivity using StoryFacets to generate recommendations on data with similar structure, automating the information discovery process. Researchers will build and test for effective ways to represent dense data in 3D and 4D (animated 3D with time as the fourth dimension) cartography. An extensive usability test of the Betaville WebGL prototype will soon launch. We will test the approach described to extend and enhance professional 3D platforms such as Esri's CityEngine (7).

Analysis: Betaville's spatial distribution of the StudentMoveTO data is a beginning stage for a more detailed representation of the commuting and transportation narratives yet to be visualized. Further representations include the ability for analysts to use Betaville to visualize with "time scale" factors that allow for clustered narratives to be revealed within specific time frames, such as the relationship between optimal travel time and the location of affordable student housing, and immersive "fly-through" experience of particular itineraries. The integration of Betaville and StoryFacets will allow analysts to see other relationships and patterns that will bring to light narratives outside of the scope of commuting and transportation, to make the most of new levels of availability of current local data, rather than inference from generalized patterns or historical data from elsewhere (8, 9).

Users: We shall adapt the interface to serve the intersecting requirements of transit analysts and university administrators, while providing students with a dashboard and calculator to help them plan where to live, form transportation strategies, make curriculum choices, factoring in work (and childcare) access, to support self-management.

At maturity, the combination of rich user experience of the environment with accessibility for stakeholders, planners, and system operators, with feature-rich data analytics including text analysis of narrative survey responses over time, will offer new levels of decision support in and beyond the domain of urban transportation planning.

References

1. http://www.studentmoveto.ca/wp-content/uploads/2016/04/StudentMoveTO.Handout_4Uni.v2.pdf (downloaded 13/03/2017)
2. Dunne, C., C. Skelton, S. Diamond, I. Meirelles, M. Martino (2016). Quantitative, Qualitative, and Historical Urban Data Visualization Tools for Professionals and Stakeholders in: Distributed, Ambient and Pervasive Interactions: 4th International Conference Proceedings. Norbert Streitz, Panos Markopoulos Eds.
3. Chow, E. (2011) Geography 2.0– A Mash-UP Perspective in: Advances in Web-based GIS, Mapping Services and Applications p. 24 edited by Songnian Li, Suzana Dragicevic, Bert Veenendaal
4. Klöckner, C. A., and T. Friedrichsmeier (2011). A multi-level approach to travel mode choice – How person characteristics and situation specific aspects determine car use in a student sample. *Transportation Research Part F*, Vol. 14, pp. 261–277.
5. Habib, K. (2014). Household-level commuting mode choices, car allocation and car ownership level choices of two-worker households: The case of the city of Toronto. Transportation 41(3): 651–672.
6. Shannon, T., B. Giles-Corti, T. Pikora, M. Bulsara, T. Shilton, and F. Bull (2006). Active commuting in a university setting: Assessing commuting habits and potential for modal change. *Transport Policy,* Vol. 13, No. 3, pp. 240-253.
7. Esri CityEngine | 3D Modeling Software for Urban Environments. Retrieved 20 March 2017, from http://www.esri.com/software/cityengine
8. Grimsrud, M., and A. El-Geneidy (2014). Transit to eternal youth: lifecycle and generational trends in Greater Montreal public transport mode share. *Journal of Transportation*, Vol. 41, pp. 1–19.
9. Lavery, T. A., A. Páez, and P. S. Kanaroglou (2013). Driving out of choices: An investigation of transport modality in a university sample. *Transportation Research*, Vol. 57, pp. 37–46.

Acknowledgements

The authors gratefully acknowledge the support of the Ontario Ministry of Research & Innovation through the ORF-RE program for the iCity Urban Informatics for Sustainable Metropolitan Growth research consortium; IBM Canada, CreateDAV, and MITACS for support of graduate and post-doctoral research; Esri Canada and MITACS for support of graduate internships; Michael Carnevale at the OCADU Visual Analytics Laboratory for asset design and development work with Betaville; Artjem Disterhof and Martin Koplin at the Media2Culture (M2C) Institut für Angewandte Medienforschung at the University of Applied Sciences of Bremen for development work on the Betaville prototype.