



Improving survey research using diagrams of data structure

Hsi-Lin (Wayne) Liu ^{a,1}, Yann-Jou Lin ^{b,2}, Yu-Wen Wang ^{c,3}, Wu-Chung Wu ^{d,*}

^a Taiwan Tourism Bureau, and Department of Horticulture and Landscape Architecture, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei 106, Taiwan

^b Department of Horticulture and Landscape Architecture, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei 106, Taiwan

^c Yu Da University, No. 168, Hsueh-fu Rd., Chaochiao, Miaoli 36143, Taiwan

^d School of Hospitality management, National Kaohsiung University of Hospitality and Tourism, 1, Sung-Ho Road, Hsiao-Kang, Kaohsiung City 812, Taiwan

ARTICLE INFO

Article history:

Received 1 September 2009

Received in revised form 1 January 2011

Accepted 1 February 2011

Available online 31 August 2011

Keywords:

Data structure diagram

Survey research

Data modeling

Databases

If-then conditions

ABSTRACT

Using diagrams of data structure, or conceptual models, is important in businesses. Survey research often has if-then data structure, but discussion of diagramming survey data structure is rare. This study uses the U.S. Fishing, Hunting and Wildlife-Associated Recreation (FHWAR) survey data and a Taiwan survey in the analysis of benefits of using data structure diagrams in survey research. Examples of data structure diagram use show how diagramming can support consistent and logical data collection, as well as improved data storage and analysis. Analysis also shows how storing if-then (conditional) data in entities/tables allows simple and intuitively meaningful unconditional variable names and can facilitate consideration of conditions that should/can affect analysis. A general conclusion is that the time has come for tourism and business survey researchers to benefit from using diagrams of data structure in planning data accumulation and to benefit from using modern systems in data collection, storage and analysis.

© 2011 Published by Elsevier Inc.

1. Introduction

Chen (1976) introduces entity-relation data diagrams in data structure modeling, or ER-modeling, as a way of thinking about data structure. Avedon (1991) claims that ER-modeling opens the door to “using inherent logical associations found in ‘real world’ information” (p. 40), eliminating the need to put all data from a respondent in a row of a table. A researcher can conceptualize data as real-world facts that they can appropriately structure. “Implementation complications are (as you might expect) in the cognitive tasks” (p. 42) of correctly defining information structure. Avedon sees the cognitive task as difficult and creative.

The value of data modeling in understanding and using data is influencing business data storage and processing (e.g., see Beynon-Davies, 2004; Shiflet, 2002). Avedon (1991) lists social science applications of entity-relation modeling, but survey researchers rarely use entity-relation and other models of data structure. Avedon’s research is linked to work in Parks Canada and includes research on survey data structure (Jaro, Stanley, & Beaman, 1992) and on understanding and using park use related data on operations (Grimm & Beaman, 1989). Beaman and

Vaske (2008) highlight the benefits of restructuring Fishing, Hunting and Wildlife Associated Recreation (FHWAR) into entity relationships.

In this paper, if-then or conditional response data refers to data having variables with responses that can only be understood based on the value of another response. Woodside and Wilson (2000) discuss if-then in relation to complexity of decision making information. Even asking people about trips they have taken results in conditional information in the sense that when reporting on two or more trips, a variable (e.g., duration), can have different values. To store data on multiple trips in one record (in a flat-file), notations like duration_1 and duration_2 are used. Woodside, MacDonald, and Burford (2004) show collecting data from leisure travels should involve obtaining conditional responses. They use a diagram in preparing to probe conscious and subconscious factors in long interviews, and they use diagrams in presenting interview results. Woodside and Wilson (2003, p. 500) review several points in a debate between large-sample researchers (e.g., $n > 100$) and case study researchers. Two points about much of the large sample research are (1) reporting is only based on the respondent and (2) researchers do not obtain the detail necessary “for gaining deep understanding of the mechanics and reasons embedded in the processes examined.” Martin and Woodside (2008) deal with the need for getting complex information from international travelers. The general implication is that by structured data collection researchers often do not collect the if-then information needed to understand decision making and behavior.

Saying that data contain if-then information is a statement about data structure. Avedon (1991) claims that thinking is constrained if researchers store complex data in a table or spreadsheet. Fig. 1 is a diagram that researchers can use to structure data without thinking in

* Corresponding author. Tel.: +886 7 8060505 1600.

E-mail addresses: wayne@tbrc.gov.tw (H.-L.(W.) Liu), yannjlin@ntu.edu.tw (Y.-J. Lin), hughwang@ydu.edu.tw (Y.-W. Wang), wuchung@mail.nkuht.edu.tw (W.-C. Wu).

¹ Tel.: +886 2 23491501 1603.

² Tel.: +886 02 33664860.

³ Tel.: +886 03 765 1188 8030.

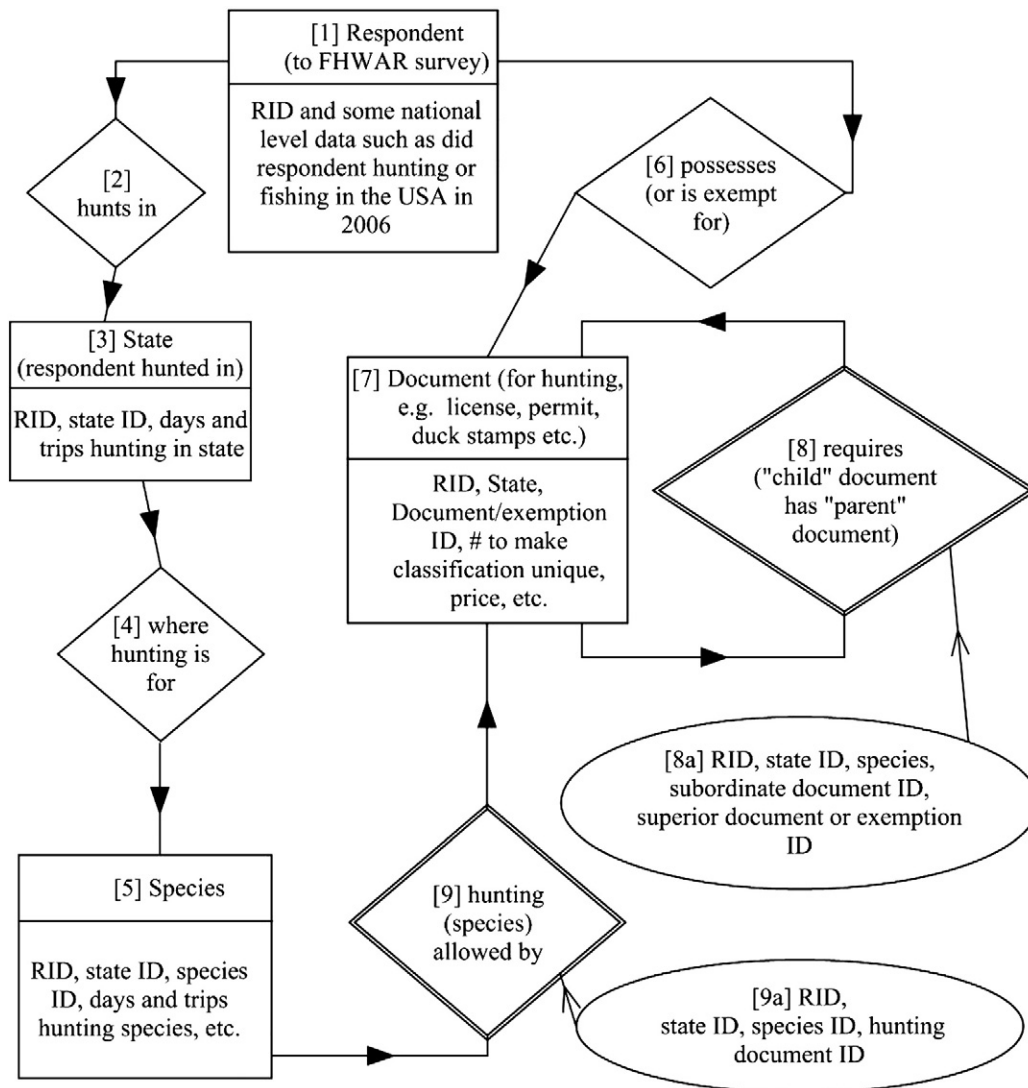


Fig. 1. Entity-relation model of relational structuring of part of the Fishing Hunting and Wildlife Associated Recreation (FHWAR) survey data and with relations that should exist.

terms of a table or spreadsheet. Beaman and Vaske (2008) used a diagram similar to Fig. 1 in restructuring parts of the 2001 and 2006 FHWAR data. In the figure, paths have shapes like □ called entities and ◇ called relations.

Entities are tables like spreadsheets with variables as columns and with values of variables in rows. For path 3–4–5, the text reads “[1] Respondent [2] hunts in [3] state” indicating a respondent hunts in a state. Attributes, or variables, of “[3] State” therefore have to do with that hunting. The figure lists attributes RID, days and trips. The respondent identification, RID, is a variable, referred to as a key, that allows linking the respondent to her/his hunting data (e.g., see Shiflet, 2002). If a respondent hunts in a state, a row of “[3] State” with his/her RID identifies the state and has total days and trips for the respondent’s hunting in the state (e.g., variables could be RID, STATE, DAYS and TRIPS). The text associated with path 3–4–5 shows that “[5] Species” is for recoding information about hunting particular species in a state. For data to be consistent, a respondent reporting hunting a species in a state must have a record in “[3] State” showing an adequate number of days and trips for the report on hunting species to be valid.

Response data could show that a respondent spent 10 days on 5 trips hunting ducks in Iowa. Days and trips hunting in Iowa would respectively only need to be greater than or equal to 10 and 5 because

one trip or day can apply to more than one species. Given that one does not hunt moose and duck on the same trip, logic implies that a stronger condition exists. The logical conditions mentioned cause the data to reflect reality and would be imposed on data. Researchers need to have checking processes identified in models to enforce such conditions.

This article is about the value of data modeling in survey research and the reasons to use data structure diagrams, not about how to create models. The last paragraph gives a certain amount of detail about Fig. 1 to allow readers to understand diagrams of data structure as models of reality. Researchers can use Fig. 1 to communicate with clients for information. Chen (1997) gives an interesting perspective on conceptual modeling. The information science (IS) literature makes clear that conceptual models such as Fig. 1 are important, but the literature is often built on the idea that support by an EDP specialist is necessary in database development (e.g., see Davies, Green, Rosemann, Indulska, & Gallo, 2006; and material in Kirchberg & Link, 2009). However, having such support need not apply to tourism or business survey researchers. Regardless, data structure diagrams in this paper are conceptual. They do not conform to conventions used by IS professionals (e.g., Beynon-Davies, 2004; Shiflet, 2002). The figures present entities and relations in a way that the authors of this paper find useful in

Download English Version:

<https://daneshyari.com/en/article/1018156>

Download Persian Version:

<https://daneshyari.com/article/1018156>

[Daneshyari.com](https://daneshyari.com)