

Project Report: Users' satisfaction about Dalarna University's Homepage

ST3012 Data Collection

Authors: Péter Tempfli, Tobias Weiß
{v19pette, v18tobwe}@du.se



May 14, 2019

Contents

1	Abstract	3
2	Introduction - Tobias	3
2.1	Research Question	3
3	Methodology	3
3.1	Research design	4
3.1.1	Population - Tobias	4
3.1.2	Sampling size - Tobias	4
3.1.3	Designing the experiment - Peter	4
3.1.4	Data collection plan - Tobias	5
3.1.5	Statistical Indicators - Tobias	5
3.1.6	Data Quality Assurance - Tobias	6
3.1.7	Ethical considerations - Peter	6
3.1.8	Legal considerations - Peter	6
3.1.9	Analysis plan	6
4	Analysis	7
4.1	Pilot	7
4.2	Absolute and Relative Frequencies	7
4.3	(Co-)variances	7
4.4	Independence	7
5	Conclusion	7

1 Abstract

We write as very last thing **Keywords:** foo bar

2 Introduction - Tobias

Not long time ago, Dalarna University decided to update the website layout to cutting-edge web technologies. To set up a website is easy but who guarantees that the users are satisfied with it? With this project we help to increase the understanding of user needs and how well these needs are suited by the prevailing system.

2.1 Research Question

Users' satisfaction is a very broad topic and the scope has to be narrowed down carefully. A survey about general satisfaction would be doomed to fail. It either would be too big and comprehensive or too general and vague. Both would lead to missing and/or imprecise information. Therefore, this project focuses on functionality aspects of the "search personal" page. Following main questions were raised:

1. If "find personal" is accessed, then the path is intuitive and quick?
2. If "find personal" page is visited, then the required data is found quickly?
3. If "find personal" page is visited, then a up-to-date front-end is displayed?
4. If a filter is used, is it intuitive to add the needed one?
5. If more than one filter is needed, is it clear what is the relation between two filters?
6. Is it intuitive to remove a filter?

In order to answer these questions a survey among the website users is conducted.

3 Methodology

A qualitative survey [1] with quantified questions is suggested as functionality remains a subjective impression. For quantification a five-point Likert scale [2] is used. The scale is not directly visible to the subject as questions can be answered with sentences like "I fully agree", "I agree", etc. There is a discussion [3] whether five-point Likert scales are a sound method or not. This research sticks to it as more sophisticated methods [4] seem unhandy and are not easily adaptable in the system, which is used for data collection.

3.1 Research design

A Google docs online form is chosen to conduct the research. It seems to be used in other educational context [5] but is not suggested in research beyond a Master thesis level. Dalarna University offers a survey tool which should be used instead in a real context.

3.1.1 Population - Tobias

The population are all human beings who have visited the Dalarna University website in the past and who will visit it in the future. One can estimate their number by analyzing server logs but still has to narrow down the time (visitors during the last month) and make certain assumptions.

3.1.2 Sampling size - Tobias

As no previous data is available, one has to assume the most conservative P value of $P = 0.5$. Furthermore, the standard values of 10% for error level and 95% for the confidence interval are assumed. Therefore, the sampling size is calculated as shown in equation 1.

$$n = Z_{\alpha}^2 \frac{P(1-P)}{(\hat{p}-P)^2} = 1.96^2 \frac{0.5(1-0.5)}{0.1^2} = 96 \text{ subjects} \quad (1)$$

3.1.3 Designing the experiment - Peter

In order to have a correct, measurable outcome, the environment of the experiment should be designed using well known, scientifically reliable techniques. We have to ensure the following principles:

1. **Randomization.** We need to have objective effect, yet also need to give room for statistical inference
2. **Replicable experiment** Our experiment should be replicable; so we can study different outcomes from the *same experiment*.
3. **Controllable** A designed experiment makes sure that the effect are caused by factors, which are known by the researchers.

We are testing one environment, but we have more than one sub-sections in our population group. We can argue that we have *at least* one dimension of sub-group in this population (status: student, teacher, researcher...); and it is also easy to see, why we could create more than one dimensions of the sub-populations (age, sex, etc..)

The other dimension is the behavior which the site visitors conduct on the site – not every visitor behaves the same way, thus some visitors experience aspect of the site, which others might not. The list of possible behaviors is probably endless, so we need to focus on very broad, but still distinctive usage patterns, such as:

- Simple person search, user found
- Simple person search, filter used
- Simple person search, combination of filters used

So we can see that we have *at least* 2 sources of variability : user group and user behavior. It is impossible to test every combination of these blocks, but if we still want to know, which block is responsible for some effect in the final outcome. For our case, if we stick with the 2-dimensional population model, the Latin Square Design is a good candidate.

The LSD design (see formula 2) orders outcome data into tables, and from these matrices builds a model, which is used for a linear regression:

$$y_{i,j,c} = \mu + \tau j + \beta i + \gamma c + \epsilon_{i,j,c} \quad (2)$$

If we only calculate the mean of the whole survey, we can't tell anything specific about sub-groups. However, once we have the regression ran, we have all the coefficients for every block; so in our case, we know, how much a sub-group likes more, or likes less a given feature. The linear model also tells the significance levels of these coefficients.

3.1.4 Data collection plan - Tobias

As the survey is not carried out on the full sample size but just on a few show cases, we send an E-Mail or WhatsApp message with link to the Google forms document and explanation to some of our classmates.

In reality, an implementation of sequential sampling as described by Fan et al. [6] is suggested directly on the website. If users search persons a binomial approach can be applied. If the user is selected, a overlay with an invitation to the survey can be displayed. Still, it remains questionable if the structure of respondents is representative and delivers an unbiased view as humans tend only to report unwanted results [7]. Maybe incentives to answer the questionnaire have to be introduced.

3.1.5 Statistical Indicators - Tobias

The analysis will heavily relate on cross tables. Besides means and medians, χ^2 coefficients [8] (as shown in formula 3) and contingency coefficients [9] (as shown in formula 4) are calculated in order to determine the level of independence among the variables. These measures are important to see whether the questions are answered in a different way among the user groups (pupils, students, researchers, etc). Technically, the calculations will be performed in R using the `assocstats` function from the package `vcd`.

$$\chi^2 = \sum_{r=1}^k \sum_{s=1}^t \frac{(m_{r,s} - n_{r,s})^2}{m_{r,s}} \quad (3)$$

$$K = \sqrt{\frac{\chi^2}{n + \chi^2}} \quad (4)$$

@Misc, Owner = homaar, Timestamp = 2019.05.14

A pairwise correlation coefficient [10] (as shown in equation 5) is computed to show if some of the questions can be abandoned in the future.

$$r_{x,y} = \frac{s_{x,y}}{\sqrt{s_x^2 \cdot s_y^2}} \quad (5)$$

3.1.6 Data Quality Assurance - Tobias

The DESAP checklist is used to keep track of the Project's data quality. "DESAP is the generic checklist for a systematic quality assessment of surveys in the European Statistical System (ESS). It has been designed as a tool for survey managers and should support them in assessing the quality of their statistics and considering improvement measures. During its development, the checklist has been tested in a pilot study covering a large variety of survey areas. It is fully compliant with the ESS quality criteria and comprises the main aspects relevant to the quality of statistical data [11]".

3.1.7 Ethical considerations - Peter

How do we store the data? What is the problem with Google forms. (Maybe write that in a real research we would use our own form/tool!)

3.1.8 Legal considerations - Peter

GDPR checkbox in the survey

3.1.9 Analysis plan

All questions follow the same scheme. A dummy cross table is represented by table 1.

Table 1: Dummy cross table for all questions

	I fully agree	I agree	I mainly agree	I partly disagree	I disagree	Σ
pupil	0	0	0	0	0	0
student from other institution	0	0	0	0	0	0
researcher from other institution	0	0	0	0	0	0
student @DU	0	0	0	0	0	0
researcher/employee @DU	0	0	0	0	0	0
other	0	0	0	0	0	0
Σ	0	0	0	0	0	0

4 Analysis

4.1 Pilot

4.2 Absolute and Relative Frequencies

4.3 (Co-)variances

4.4 Independence

5 Conclusion

References

- [1] S. Sofaer, "Qualitative methods: what are they and why use them?," *Health services research*, vol. 34, no. 5 Pt 2, p. 1101, 1999.
- [2] R. Likert, "A technique for the measurement of attitudes.," *Archives of psychology*, 1932.
- [3] R. A. Cummins and E. Gullone, "Why we should not use 5-point likert scales: The case for subjective quality of life measurement," in *Proceedings, second international conference on quality of life in cities*, vol. 74, p. 93, 2000.
- [4] C. J. Chimi and D. L. Russell, "The likert scale: A proposal for improvement using quasi-continuous variables," in *Information Systems Education Conference, Washington, DC*, pp. 1–10, 2009.

- [5] E. F. Gehringer, “Daily course evaluation with google forms,” in *ASEE, American Society for Engineering Education Annual Conference & Exposition*, vol. 27, 2010.
- [6] C. Fan, M. E. Muller, and I. Rezucha, “Development of sampling plans by using sequential (item by item) selection techniques and digital computers,” *Journal of the American Statistical Association*, vol. 57, no. 298, pp. 387–402, 1962.
- [7] R. Bergstrand, A. Vedin, C. Wilhelmsson, and L. Wilhelmsen, “Bias due to non-participation and heterogenous sub-groups in population surveys,” *Journal of chronic diseases*, vol. 36, no. 10, pp. 725–728, 1983.
- [8] K. P. F.R.S., “X. on the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling,” *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 50, no. 302, pp. 157–175, 1900.
- [9] K. Pearson, J. A. Harris, A. E. Treloar, M. Wilder, and K. Pearson, “On the theory of contingency,” *Journal of the American Statistical Association*, vol. 25, no. 171, pp. 320–327, 1930.
- [10] J. Lee Rodgers and W. A. Nicewander, “Thirteen ways to look at the correlation coefficient,” *The American Statistician*, vol. 42, no. 1, pp. 59–66, 1988.
- [11] Eurostat, “The europeanself assessment checklistfor survey managers.” https://ec.europa.eu/eurostat/documents/64157/4373903/07-Checklist-for-Survey-Managers_DESAP-EN.pdf, 2008. [Online; accessed 14-May-2019].

Appendices

Survey Guideline

Information about the questions

Survey Questions

Only absolute relevant questions

DESAP checklist

Checklist can be found in the git folder